JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 38, 1-16 (2022)
DOI: 10.6688/JISE.202201_38(1).0001

K-Implicit Tracking Data Publishing Scheme Against Geo-Matching Attacks

KUN NIU[1,2], CHANGGEN PENG[1,2,+], YOULIANG TIAN[1,2] AND WEIJIE TAN[1]
[1]*State Key Laboratory of Public Big Data*
College of Computer Science and Technology
[2]*Institute of Cryptography and Data Security*
Guizhou University
Guiyang, 550025 P.R. China
E-mail: cgpeng@gzu.edu.cn

The dynamic queries by wireless mobile network users, will generate the social data with location tags and tracking data sequences, which enable the adversary can infer privacy information combined background knowledge, especially strong geographic correlation information. Therefore, we propose the Geo-matching privacy inference attack methods based on road network and sensitive semantic location. To address these issues, a *k*-implicit data publishing scheme with adaptive privacy budget is presented, which is based on the road network topological graph and the sensitivity quantification of grid unit, and also realizes an optimized dynamic anonymous region construction. Finally, the proposed Geo-matching attack algorithm is simulated to verify the effectiveness of the *k*-implicit data publishing scheme. The experiment results show that the proposed scheme can better resist the Geo-matching attack under different privacy budget thresholds.

Keywords: area sensitivity, geo-matching attacks, implicit privacy protection, tracking data, spatial anonymit

1. INTRODUCTION

With the development of the mobile location technology and wireless communication network technology, mobile users can obtain a large number of accurate and reliable information at any time, any place and any environment. However, at these interactive communication, it will produce the users' mobile tracking information (*i.e.*, a sequence of points with timestamps sorted by sampling time) and the social data with location tags, that will be used by the attacker, lead to serious privacy issues. For example, in intelligent transportation, the server continuously collects global positioning system (GPS) sampling information of vehicle equipment, which will leak the driver's trajectory. Some information can be analyzed by adversary such as home addresses, workplaces, and health conditions, *etc*. In location based service (LBS) services with continuous queries (such as finding points of interest), the user needs to continuously upload the location to LBS server, the privacy of mobile users may be exposed through real-time tracking. Therefore, the goal of dynamic location privacy protection of wireless communication network is:

Received July 10, 2020; revised September 3, 2020; accepted September 25, 2020.
Communicated by Xiaohong Jiang.
+ Corresponding author.

the premise of satisfying user query accuracy and low network energy consumption, the communication mode of users can be hidden through path disguise, anonymous publication and other technologies, in order to prevent the leakage of network sensitive trajectory sequence information [1], at present, trajectory privacy protection based on geo-sensing has been widely studied and concerned by the academic and business communities. Research combined with road network graph data is often used in the geo-sensing privacy protection technology [2]. It is essential to design a new privacy protection model that not only protects the precise location information of user (geographic coordinates), but also protects any side information that could lead to make a rough estimate of the location.

The concept of location K anonymity was first proposed in reference [3], it is mainly divided into spatial anonymity and virtual anonymity. The idea is to ensure that, the publishing area contains at least another k–1 user besides the user who initiated the query, and the server cannot accurately identify the user who actually initiated the query from the k users. The spatial anonymous method usually needs the help of fully-trusted third party. Therefore, this method has certain limitations, which can not meet the needs of adaptive user privacy protection, and the data quality is reduced, unable to meet the query accuracy. In the research of spatio-temporal data publishing for road network, K anonymity is also important [4]. Memon *et al*. [5] proposes a mixed regional scheme for multiple pseudonym transformation strategies, which provides trajectory privacy protection for real-time operation in the traffic network. However, multiple query request by a single user are discrete and discontinuous. Therefore, Yiu *et al*. [6] proposes spatial segmentation method according to the road network density, then generates k anonymity zone according to the user's own privacy preference, and uses the anchor point optimization algorithm to generate the anchor points to achieve the anonymous query effect.

For continuous trajectory data publishing scenarios, cached data is often used to improve a spatial K anonymous algorithm for multi-level caching based on the predicted location, unit cache contribution rate, so as to enhance user privacy [7]. There are also theory and methods research, such as information entropy theory, spatial encryption method, and the optimal trajectory publishing under the utility constraint of Markov model method [8–10]. In reality, multi-user aggregate query is often published in real applications. Therefore, relevant researches use user collaboration to publish false query tracks, and put forward the aggregation nearest neighbor query algorithm based on strategy optimization to confuse the adversary [11–13].

However, as mentioned above, storage and publishing of any sensitive data can be extracted effective correlation knowledge and mined user privacy by data mining technology [14,15]. Common attacks include background attack, probability attack and semantic attack. It can take advantage of the uniqueness and regularity of mobility and recover the personal trajectory from the aggregated mobile data publishing with an accuracy of 73%-91% without any priori knowledge [16]. Therefore, the improvement of the privacy protection methods against aggregation association and background knowledge attacks are still hot research topics [17]. Most of the existing location privacy protection mechanisms are mainly aimed at homogenous attacks with privacy disclosure of sensitive attribute by users, query related attacks, and privacy attacks based on the correlation between user trajectories [18–20]. In order to increase the space-time uncertainty of location data, the information of user identity is hidden by data distortion, suppression publishing, encryption tools and other methods [21–23]. But it increases the overhead of communication

and hardware, and reduces the query service quality.

Most of the studies focus on spatio-temporal sequence data publishing scheme [24, 25]. In these scheme, noise disturbance is only added for resistance user accurate identification of the path, but the problem is that unbounded random noise is unable to protect user privacy. Differential privacy theory is commonly used to control the noise generation mechanism, the noise generation algorithm is adopted by finite laplace function [26, 27]. However, noise data, real data and tracking spatio-temporal sequence data combines with the road network will generate location region links, which will reduce its anti-attack capability to different degrees [28].

Privacy protection based on user tracking data is mainly aimed at two aspects: Firstly, protect the user's exact location information based on the implicit publishing of the road network index. Secondly, quantify grid area sensitivity, and optimize the effectiveness and availability of publishing locations. The main contributions of this paper are summarize as follows:

- *This paper defines two kinds of Geo-matching privacy attack methods, which can effectively attack the tracking data published anonymously, and infer the user's identity attributes such as route and address. In addition, the adversary attack algorithm based on road network matching is constructed.*

- *According to the frequency of cache users and the adaptive regional sensitivity of users' preferences, a k-implicit privacy protection scheme based on the road network index is proposed, which realizes the anonymous region optimization publishing algorithm, and can resist the correlation attack of the road network and the sensitive semantic location of users.*

- *Simulation experiment is carried out, the implicit tracking data publishing scheme can effectively protect the correlation between users and the real geographical characteristics under the different thresholds.*

The remainder of this paper is organized as follows. The preliminary knowledge of road network graph and position tracking data model are discussed in Section 2. In Section 3, we define a new Geo-matching privacy attack based on the road network and the sensitive semantic. The system framework and algorithm are proposed for the online interactive location tracking privacy protection method in Section 4. Simulation and experimental results are provided in Section 5, followed by conclusions in Section 6.

2. PRELIMINARY

The path link attack caused by the geographical element matching in the continuous publishing scenario can also lead to serious privacy leakage [29]. Therefore, how to take the correlation road network environment as the strong correlation background knowledge of the tracking data, effectively utilize the geographic matching information and reasonably conceal the geographical location under the dynamic publishing scenario of user location. In this section, we introduce relevant important concepts and basic knowledge of road network graph, location tracking, and path recognition accuracy.

Definition 1 *Road network graph.*

The road network topology structure is described as a weighted undirected incomplete sparse graph RoadWork **RW**.

$$\begin{aligned} &\mathbf{RW} = (\mathbb{V}, \boldsymbol{E}, \boldsymbol{W}) \\ &\mathbb{V} = \{x | x \in \text{Nodeset}\} \\ &\boldsymbol{E} = \{\langle \text{x}, \text{y} \rangle | \boldsymbol{L}(x,y) \cap (x,y) \in \mathbb{N}\}, \end{aligned} \tag{1}$$

here, $\mathbb{V}$ is the nodes set of the road network graph, the unordered pair $< x, y >$ represents the edge between nodes x and y, and $\boldsymbol{E}$ represents the links set between different nodes in the RoadWork, the predicate $\boldsymbol{L}(x,y)$ represents the path from node x to y, $\boldsymbol{W}$ is the weights set of the edges in the graph.

Definition 2 *Location tracking.*

In the three-dimensional space (X,Y,T), the trajectory **Tr** of a moving object is determined by the sequence points $(x_1,y_1,t_1),(x_2,y_2,t_2),\cdots,(x_n,y_n,t_n)$ $(t_1 < t_2 < \cdots < t_n)$. The trajectory defines the position of the moving object as an implicit time function. The position between the moving object on the trajectory $\mathbf{Tr}(x_i,y_i)$ and (x_{i+1},y_{i+1}) points can be obtained by linear interpolation.

Definition 3 *The point of interest datas (**Poi**$_s$).*

The point of interest data set (**Poi**$_s$) is expressed as

$$\mathbf{Poi}_s = \{\mathbf{Poi}_1, \cdots, \mathbf{Poi}_i, \cdots, \mathbf{Poi}_n\}, \tag{2}$$

here, **Poi**$_i$ in the collection is generally captured by keywords on the i real location data. Its geographic structure can be defined as:

$$\mathbf{Poi}_i = (< \text{lon}, \text{lat} >, \text{Address}, \text{semType}, \text{Con}, \text{Time}), \tag{3}$$

where $< lon, lat >$ is the longitude and latitude of the **Poi**, *address* is the semantic description, *semType* is the semantic type, and *Con* is the confidence degree.

Definition 4 *Path recognition accuracy* ρ.

The attack efficiency of the proposed attack algorithm based on Geo-matching is represented by path recognition accuracy ρ, the sample space is represented by path set Tr, Tr' is the sample number of the tracking data points identified as the wrong path after the enhanced privacy release scheme, then,

$$\rho = \frac{\mid \text{Tr}\mid - \mid \text{Tr}' \mid}{\mid \text{Tr} \mid} \times 100\%. \tag{4}$$

3. GEO-MATCHING PRIVACY ATTACKS

Geo-matching refers to the process of matching the publishing location sequence with the digital map. In continuous query, the user publishes the anonymous area (or

anonymous set) continuously, the attacker has strong background knowledge, including the road network vector data, historical anonymous area (or anonymous set) R_{u,t_i}, geographic unreachable location *Gul* and other information.

Definition 5 *Geo-matching privacy attack based on road network.*

Hypothesis the anonymous areas of the user U at the time $(t_1, t_2, \cdots, t_n)$ are ($\mathbf{R}_{u,t_1}$, $\mathbf{R}_{u,t_2}, \cdots, \mathbf{R}_{u,t_n}$), respectively. Each anonymous region $\mathbf{R}_{u,t_i}$, it consists of K perturbation points. If the anonymous area is $\mathbf{R}_{u,t_i}$ and the centroid of t_i is the point Q_i, there is,

a: $R_{u,ti} \cap Gul \neq 0$, the privacy of the user u at the time t_i is threatened, and the disturbance point $P_{u,i}$ at that moment can be recognized by the attacker.

b: Users continuously publish query requests and generate anonymous region of location data through anonymous technology before submission query server. The dynamic anonymous region centers are mapped one by one on the road network vector data, which can generate the continuous nearest road network location points $(H_1, H_2, \cdots, H_n)$. As a result, the user's tracking privacy is threatened, and then the user's path way, identity attributes, behaviors and other information are inferred, which is shown in Algorithm 1.

Algorithm 1 : Geo-matching privacy attack based on road network.

Input: $\mathbf{R} = \{R_{u,t1}, \cdots, R_{u,t_i}, \cdots, R_{u,t_n}\}$, $routeNet$, R_{u,t_i} is the perturbation cluster published at time t_i

Output: getPath

```
1: U_i ⟵ the centroid of R_{u,t_i};
2: Step 1: Go through the information of each point.
3: for i = 1; i < n; i++ do
4:     optimal Route;
5: end for;
6: Step 2: Construct the tree topology structure of road network.
7: treeInfo = creatTree(routeNet);
8: Step 3: Extract the path closest to the U_i.
9: for all Edge such that Edge ∈ netWork do
10:    if (distance < min) then
11:        min = distance;
12:        path = edge;
13:    end if
14: end for;
15: Step 4: Complete the path lookup according to the index.
16: getPath=(startPos,treeInfo,endPos)
```

Definition 6 *Geo-matching privacy attacks based on sensitive semantic.*

Mobile communication network location technology uses a symbolic language to describe a geographic location, which is essentially a semantic abstraction of location [30, 31]. Privacy attack based on semantic matching refers to the attacker uses the anchor (location reference point) location or the sensitive semantic location associated with the

anonymous region, as well as other side information in sensitive region to estimate accurately the target user privacy preferences. Consider the following semantic sensitivity quantification and attack methods:

a: Location aggregation based on the nearest neighbor analysis (NNA).

The location sensitivity of the grid unit is defined by the spatial aggregation or dispersion degree of the historical location points, and the spatial aggregation degree is quantified based on NNA [32].

$$\mathrm{NNA} = \frac{\overline{D_o}}{\overline{D_E}}, \tag{5}$$

where $\overline{D_o} = \sum_{i=1}^{n} d_i/n$ is the average of the distance between the feature point and its nearest neighbor. d_i is the distance between the feature point and the nearest neighbor. n is the number of position points in the grid unit. D_E is the average distance of randomly distributed features.

b: Privacy attacks based on sensitivity.

It is assumed that the probability density function of the buffer position distribution in space is denoted as $P(r)$, and when $P(r) = 0$, the region r is unreachable. The user sets the semantic position ft to be sensitive and non-sensitive according to privacy preferences, and defines an acceptable minimum regional sensitivity threshold q. The sensitivity of the position ft in the region r is defined as $P(ft, r)$.

$$P(ft, r) = \frac{\mathrm{Area}(ft \cap r)}{\mathrm{Area}(r)}, \tag{6}$$

$\mathrm{Area}(a)$ represents the area of region a, if the position sensitivity satisfy equation $P(ft, r) \geq q$, then the area generates a privacy attack with sensitive semantic matching.

4. IMPLICIT LOCATION TRACKING PRIVACY PROTECTION

4.1 System Model

As mentioned in the previous section, the location sequence generated by tracking in physical LBS service can lead to the inference attack on user privacy due to geographical matching problems such as spatial road network and regional sensitivity. Therefore, we propose a system architecture that satisfies the privacy protection of tracking location, which is shown in Fig. 1.

- *Users: Each mobile client carries a mobile device with information processing, data storage, GPS and other functions to quickly access the internet of things. Users participating in the LBS query trust each other, and there is no collusion attack between the user and the LSP.*

- *Anonymous Server: The server is trusted. Its main functions include anonymous request, stealth area construction, buffer filtering query result, and the purpose of dynamically hiding the user's moving path in the tracking location publishing scene.*

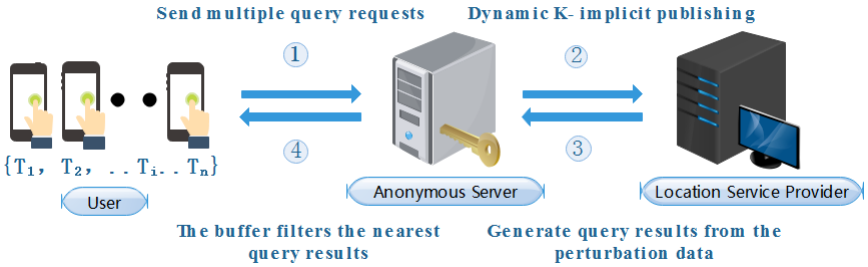

Fig. 1. Implicit location tracking privacy protection system architecture.

- *Location Service Provider (LSP): LSP stores map resources and POIs related to geographical locations, such as restaurants, hospitals, tourist attractions, to provide users with timely location-based query services. But, during the user query process, the tracking data stored by LSP has the risk of privacy leakage.*

In this system model, the anonymous server is used as the trusted third party between the client and the location server. In order to construct an effective anonymous area, the anonymous server needs to save the current map information and the road topology information, and update the information of the mobile users in the road in real time. In addition, the third-party anonymous server mainly divides and computes the cache user density threshold, as well as the anonymous partition construction operation, in order to reduce the query and computing cost of the user mobile terminal.

4.2 Scheme and Algorithm

In order to prevent Geo-matching tracking location privacy attacks, the location confusion is published through a trusted third-party server, dynamically adjusted according to the road network structure in real time, traversing the user network's road network segmentation space outsourcing unit at time t_i. According to the supply of space center $\rho(X_{ti},Y_{ti})$, quantize grid privacy sensitivity, and construct an area η that satisfies k anonymity. The proposed scheme is as follows:

Step 1: User u issues a query request at time t_i:

$$req(u,t_i) = (ID_u, < x_{ti}, y_{ti} >, P.type),$$

where ID_u is the user identifier, $< x_{ti}, y_{ti} >$ is the location information of the user current time, P.type indicates the type of the query $\mathbf{POI}_s$.

Step 2: After receiving the request $req(u,t_i)$, the third-party server creates a quadtree index of the node, constructs a network space K anonymous area η for the current user, and determines the privacy preference sensitivity of the anonymous area. $P(ft,r) < q$ (privacy threshold), the third step is performed, if it is greater than q, the publishing is suppressed.

Step 3: Replace the original point with the K anonymous disturbance point to issue a query request to the LBS server, or perform data distribution. LSP performs the

policy optimized nearest neighbor query for all requests, and returns the result to the third-party server.

Step 4: Based on the real user location, the third-party server traverses the $\mathbf{POI}_s$ linear index, finds and obtains the final point of interest, and returns to the user.

Algorithm 2 is described as follows:

Algorithm 2 : Implicit tracking data publishing.

Input: $P_{t_i} = \{P_{t_1}, P_{t_2}, \cdots, P_{t_i}, \cdots, P_{t_n}\}$, K, $h = 1$
Output: the disturbance tracking position S_i;

```
Step 1: create a network tree topology.
V ⟵ {∅}; S ⟵ {∅}; h ⟵ 1
for i = 1; i < n; i++ do
    Constructing the minimum road network cell V_i for user U_i;
end for
Step 2: the calculation of the anchor
for i = 1; i < n; i++ do
    Calculate the centroid S_i for user's network cell;
    S ⟵ {S_i};
end for
Step 3: construct anonymous area η and sensitivity quantification.
P(ft, r) ⟵ Area(ft ⋂ r)/Area(r)
if P < q then
    add V_i into the set, V ⟵ {V_i}, h = h + 1;
else
    Inhibition of publishing and going to line 11;
end if
Step 4: disturbance tracking position publication.
```

4.3 Anonymous Area η Construction Algorithm

In order to achieve that a mobile user's publishing location is indistinguishable from other $k-1$ users' locations, in the proposed scheme, we improved the anonymous area η construction method to prevent strong background knowledge associated attacks of road network and Geo-matching. The grid unit is constructed according to the area where the road network data is located, and the segmentation is based on the closed road network area where the user is located. The specific construction scheme is showed in Fig. 2 and Algorithm 3 are described as follows:

1) User u sends a request to LBS server at the time t_i, and finds the road network segmentation area of the current location point according to the road network topology index.

2) Traverse and search the historical cache data points on the anonymous server. In the segmentation region, filter and identify the grid cells according to the probability density threshold that meet the privacy requirements. White grids mean that

the historical cache is smaller than the threshold, and gray grids mean that the historical cache meets the threshold. That goal is to improve the geographic semantic indistinguishability of anonymous publishing points. The higher the probability is, the weaker the geographic semantics will be, which can protect users' privacy information.

3) Find the centroid S_i of the road network segmentation region. Starting from the location of the anchor point, find the area contained the k identified grid cells closest to the centroid through the extended matrix, which region is the anonymous region η. It can protect path indistinguishability of the tracking data publishing.

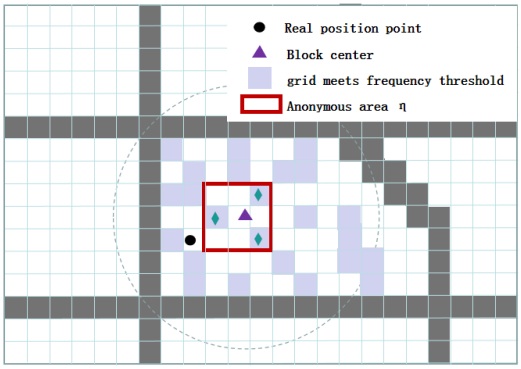

Fig. 2. Anonymous area η construction.

Algorithm 3 : Anonymous area η construction algorithm.

Input: $P_{t_i} = \{P_{t_1}, P_{t_2}, \cdots, P_{t_i}, \cdots, P_{t_n}\}$,$K$,$M$,$network$
Output: Anonymous area η

Step 1: Gets the user track point.
for $i = 1; i < n; i++$ **do**
    $P_{t_i} \longleftarrow$ List $< Vertex >$ user;
end for
Step 2: Find the block area center.
$S_i \longleftarrow \{cellX, cellY\}$;
Compute the grid buffered data frequency $\{m_i\}$;
Constructing the Road network block V_i for P_{t_i};
Step 3: Find grid cells that meet the sensitivity threshold M and extended matrix boundary that satisfies K anonymity.
for $i = 1; i < k; i++$ **do**
    if $m_i > M$ **then**
        Add m_i into the set, spread rectangular η;
    end if
end for

4.4 Complexity Analysis

Time complexity is an important indicator to measure the performance of the algorithm. The time cost of this algorithm is mainly reflected in two parts: The first part is to traverse the cached data set s to find the grid unit that meets the threshold. The time complexity of this operation part is $\mathcal{O}(s)$. The second part is to query the road network partition operation, find the regional centroid, and satisfy the indistinguishable to adjacent paths. The time complexity of this operation part is $\mathcal{O}(ablog(ab))$, where a is the road network density and b is the sampling point. The third part is the diffusion matrix to find the nearest k identification grid cells, the time complexity is $\mathcal{O}(k)$, so the total time complexity is $\mathcal{O}(s+ablog(ab)+k)$.

5. EVALUATIONS AND RESULTS

In this section, the tracking data publishing method based on road network topology is tested on the road network data set. The experiment uses the Thomas Brinkhof network data generator and generates moving object data [17]. It assumes that the road length is proportional to the number of users, and network communication is secure and trustworthy, regardless of the energy consumption caused by network communication.

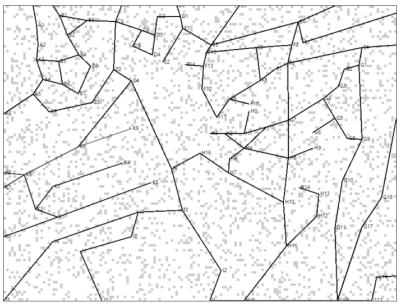

Fig. 3. Cache data frequency statistics grid graph.

Fig. 4. Privacy attack effect based Geo-matching.

Fig. 3 shows the road network data and cell grids that satisfies the area sensitivity query by mobile users' threshold settings. In the experiment, we extract the real-time track point and randomly generate noise point (false noise is removed from the geographic sensitivity evaluation). The Geo-matching attack algorithm takes the minimum outsourcing rectangle, add the edge nearest to the center of the rectangle into the attack edge set, and then cyclic query until all the tracks are traversed.

Fig. 4 shows the road network matching attack algorithm effect under the $K = 10$ anonymous scheme. It can be seen that the algorithm can still obtain the dynamic tracking sequence of the user when the random noise disturbance is published.

Therefore, based on the grid sensitivity and rectangular spreading algorithm, we improved implicit anonymous scheme combining with the road network index to find the optimal anchor point, searches the disturbance data points around the anchor point,

constructs the anonymous area, and realizes *k*-implicit publishing, which has good anti-attack ability.

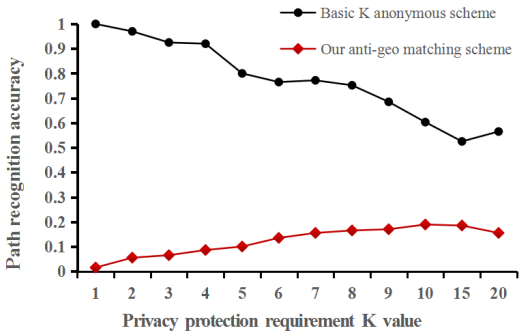

Fig. 5. Path recognition accuracy ρ under different privacy protection schemes.

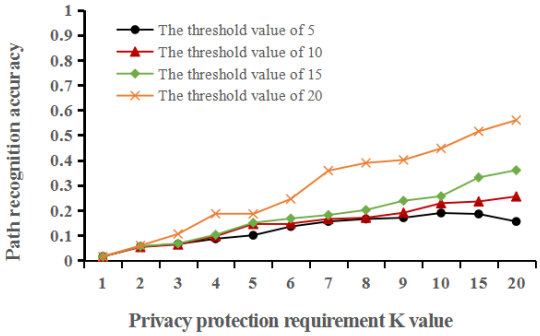

Fig. 6. Path recognition accuracy ρ of different cache threshold schemes.

Fig. 5 shows the effect of the improved implicit publishing algorithm for different values of K. The scheme compares the basis k-anonymous perturbation publishing schemes, and the recognition rate of the correct road segment after the disturbance is the quantification of its anti-attack capability. The scheme has a good statistical performance when the K value is small, so that the anti-attack capability reaches a very high peak.

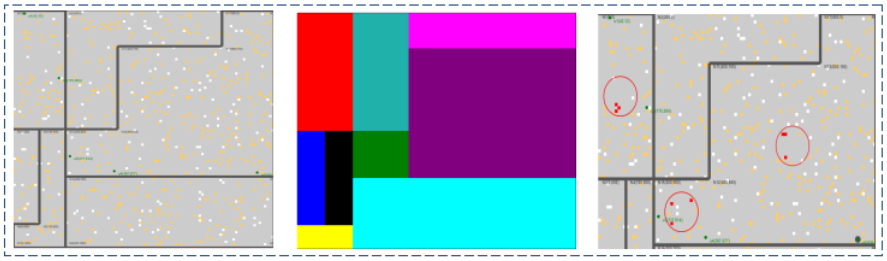

Fig. 7. K-implicit privacy protection tracking data publishing process.

In addition, the scheme verifies the path recognition accuracy under different privacy protection threshold settings. When the K value is low, there is no significant difference. With the increase of K value, the publishing scheme with higher threshold setting has higher dispersion degree and higher correct recognition efficiency, as shown in Fig. 6.

Fig. 7 shows the K-implicit privacy protection process of tracking data publishing. Fig. 7 (a) shows caching data preprocessing and privacy threshold settings. Fig. 7 (b) shows the partitioning operation for finding the anchor points. In Fig. 7 (c), the anonymous areas are constructed through rectangular spreading algorithm firstly by step-by-step searching, and then we can attain the optimal K implicit publishing data.

The experiment shows that the scheme can make the adversary not to distinguish the real moving path of publishing data, also ensure that the adversary can not identify the real position through the semantic sensitivity of user's continuous publishing locations.

6. CONCLUSION

This paper proposes a Geo-matching privacy attack algorithm, which considering the strong correlation problems between anonymous publishing and geographical environment, such as road network, geographical unaccessibility, and geo-coding semantics. Moreover, we propose an optimized anonymous area construction algorithm and a k-implicit publishing scheme against Geo-matching attack. The proposed scheme combines with the network characteristics and quantitative sensitivity of user preferences region, which can protect the user tracking sequence from association mining, realizes the user implicit publishing, and guarantees the user query efficiency. Simulation experiment tests the anti-attack capability of the proposed scheme, it shows that the proposed scheme has the performance of anti-matching attack and anti-association analysis.

Privacy protection method and privacy quantification of tracking data with spatio-temporal attributes are still under constant research, and there are still many problems that need to be further studied and solved. In the next stage, we will continue to study the privacy protect and risk quantification of the tracking process.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of China (No. U1836205, 61662009, 61772008 and 11761020), The Science and Technology Program of Guizhou Province (No. Guizhou Science Contract Major Program [2018]3001, [2018]3007 and [2017]3002, Guizhou Science Contract Support [2019]2004, [2018]2162 and [2018] 2159, Guizhou Science Contract Foundation [2019]1049 and [2017]1045, Guizhou Science Contract Platform Talent [2020]5017), and the 13th Five-Year National Cryptography Development Foundation (No. MMJJ20170129).

REFERENCES

1. D. Wu, X. Wang, L. Sun, Y. Ling, and D. Zhang, "Identity privacy based reliable routing method in VANETs," *Peer-to-Peer Networking and Applications*, Vol. 7, 2014,

pp. 285-294.

2. C. Liu, Y. Tian, J. Xiong, Y. Lu, Q. Li, and C. Peng, "Towards attack and defense views to *k*-anonymous using information theory approach," *IEEE Access*, Vol. 7, 2019, pp. 156025-156032.
3. M. Gruster and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, 2003, pp. 31-42.
4. F. Zheng, X. Peng, and P. Li, "One more accuracy *k*-anonymity framework for lbs," *Mobile Information Systems*, Vol. 2019, 2019, pp. 1-11.
5. I. Memon, L. Chen, Q. A. Arain, and H. Memon, "Pseudonym changing strategy with multiple mix zones for trajectory privacy protection in road networks," *International Journal of Communication Systems*, Vol. 31, 2018, pp. 1-44.
6. M. L. Yiu, C. S. Jensen, X. Huang, and H. Lu, "SpaceTwist: managing the trade-offs among location privacy, query performance, and query accuracy in mobile services," in *Proceedings of IEEE 24th International Conference on Data Engineering*, 2008, pp. 366-375.
7. S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatial *k*-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Generation Computer Systems*, Vol. 94, 2019, pp. 40-50.
8. W. Zhang , M. Li, R.Tandon, and H. Li, "Online location trace privacy: an information theoretic approach," *IEEE Transactions on Information Forensics and Security*, Vol. 14, 2019, pp. 235-250.
9. B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving *k*-anonymity in privacy-aware location-based services," in *Proceedings of IEEE Conference on Computer Communications*, 2014, pp. 754-762.
10. D. Liao, H. Li, G. Sun, M. Zhang, and V. Chang, "Location and trajectory privacy preservation in 5G-enabled vehicle social network services," *Journal of Network and Computer Applications*, Vol. 110, 2018, pp. 108-118.
11. L. Zhang, C. Jin, C. Huang, and J. Fu, "A trajectory privacy preserving scheme in the CANNQ service for IoT," *Sensors*, Vol. 19, 2019, pp. 1-23.
12. T. Peng, Q. Liu, D. Meng, and G. Wang, "Collaborative trajectory privacy preserving scheme in location-based services," *Information Sciences*, Vol. 387, 2017, pp. 165-179.
13. X. Wang, Y. Luo, S. Liu, T. Wang, and H. Han, "Subspace *k*-anonymity algorithm for location-privacy preservation based on locality-sensitive hashing," *Intelligent Data Analysis*, Vol. 23, 2019, pp. 1167-1185.
14. J. Xiong, H. Liu, B. Jin, Q. Li, and Z. Yao, "A lightweight privacy protection scheme based on user preference in mobile crowdsensing," *Transactions on Emerging Telecommunications Technologies*, Vol. 6, 2020, pp. 1-16.
15. R. Mendes and J. P. Vilela, "Privacy-preserving data mining: methods, metrics and applications," *IEEE Access*, Vol. 5, 2017, pp. 10562-10582.
16. Z. Tu, F. Xu, Y. Li, P. Zhang, and D. Jin, "A new privacy breach: user trajectory recovery from aggregated mobility data," *IEEE/ACM Transactions on Networking*, Vol. 26, 2018, pp. 1446-1459.

17. J. Chen, K. He, Q. Yuan, M. Chen, R. Du, and Y. Xiang, "Blind filtering at third parties: an efficient privacy-preserving framework for location-based services," *IEEE Transactions on Mobile Computing*, Vol. 17, 2018, pp. 2524-2535.
18. H. Hu, J. Xu, S. T. On, and J. Du, "Ng: privacy-aware location data publishing," *ACM Transactions on Database Systems*, Vol. 35, 2010, pp. 53-56.
19. J. Xu, X. Tang, H. Hu, and J. Du, "Privacy-conscious location based queries in mobile environments," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 21, 2010, pp. 313-326.
20. K. Zhao, Z. Tu, F. Xu, and Y. Li, "Walking without friends: publishing anonymized trajectory dataset without leaking social relationships," *IEEE Transactions on Network and Service Management*, Vol. 16, 2019, pp. 1212-1225.
21. H. Cai, Y. Zhu, and Z. Feng, "A truthful incentive mechanism for mobile crowd sensing with location-sensitive weighted tasks," *Computer Networks*, Vol. 13, 2018, pp. 1-14.
22. R. Shokri, G. Theodorakopoulos, P. Papadimitratos, E. Kazemi, and J. Hubaux, "Hiding in the mobile crowd: location privacy through collaboration," *IEEE Transactions on Dependable and Secure Computing*, Vol. 11, 2014, pp. 266-279.
23. I. Lien, Y. Lin, J. Shieh, and J. Wu, "A novel privacy preserving location-based service protocol with secret circular shift for KNN search," *IEEE Transactions on Information Forensics and Security*, Vol. 8, 2013, pp. 863-873.
24. S. Ji, T. Wang, J. Chen, W. Li, P. Mittal, and R. Beyah, "De-SAG: on the de-anonymization of structure-attribute graph data," *IEEE Transactions on Dependable and Secure Computing*, Vol. 16, 2019, pp. 594-607.
25. X. Wang, A. Pande, J. Zhu, and P. Mohapatra, "STAMP: Enabling privacy-preserving location proofs for mobile users," *IEEE/ACM Transactions on Networking*, Vol. 24, 2016, pp. 3276-3289.
26. M. Li, L. Zhu, Z. Zhang, and R. Xu, "Achieving differential privacy of trajectory data publishing in participatory sensing," *Information Sciences*, Vol. 400, 2017, pp. 1-13.
27. S. Zhang, X. Mao, K. K. R. Choo, T. Peng, and G. Wang, "A trajectory privacy-preserving scheme based on a dual-k mechanism for continuous location-based services," *Information Sciences*, Vol. 527, 2020, pp. 406-419.
28. S. Wang and R. O. Sinnott, "Protecting personal trajectories of social media users through differential privacy," *Computer Society*, Vol. 67, 2017, pp. 142-163.
29. X. Pan, J. Xu, and X. Meng, "Protecting location privacy against location-dependent attacks in mobile services," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24, 2012, pp. 1506-1519.
30. K. Wang, W. Zhao, J. Cui, Y. Cui, and J. Hu, "A k-anonymous clustering algorithm based on the analytic hierarchy process," *Journal of Visual Communication and Image Representation*, Vol. 59, 2019, pp. 76-83.
31. F. Deldar and M. Abadi, "PLDP-TD: personalized-location differentially private data analysis on trajectory databases," *Pervasive and Mobile Computing*, Vol. 49, 2018, pp. 1-22.
32. L. Ni, F. Tian, Q. Ni, Y. Yan, and J. Zhang, "An anonymous entropy-based location privacy protection scheme in mobile social networks," *Journal on Wireless Communications and Networking*, Vol. 93, 2019, pp. 1-19.

Kun Niu received her M.S. degree from China University of Mining and Technology, China, in 2011. Currently, she is a Ph.D. student in School of Computer Science and Technology, Guizhou University. Her main research interests include space information technology and privacy protection.

Changgen Peng received his Ph.D. degree from the Department of Computer Science and Technology, Guizhou University, China, in 2007. He is presently a Professor, Ph.D. Supervisor at the College of Computer Science and Technology, Guizhou University. He is an academic leader of data security and cryptography at State Key Laboratory of Public Big Data. His research interests include data privacy, cryptography and big data technology and security.

Youliang Tian received the B.S. degree in Mathematics and Applied Mathematics and the M.S. degree in Applied Mathematics from Guizhou University, in 2004 and 2009, respectively, and the Ph.D. degree in Cryptography from the Xidian University in 2012. From 2012 to 2015, he was a Postdoctoral Associate with the State Key Laboratory, Chinese Academy of Sciences. He is currently a Professor and the Ph.D. Supervisor with the College of Computer Science and Technology, Guizhou University. His research interests include algorithm game theory, cryptography, and security protocol.

Weijie Tan received the M.S. degree in Communication and Information System from the Communication University of China, Beijing, China, in 2011, and the Ph.D. degree in Information and Communications Engineering from Northwestern Polytechnical University, Xi'an, China, in 2019. From 2016 to 2017, he was a visiting researcher with the Audio Analysis Laboratory, AD:MT, Aalborg University, Denmark. He is currently with the faculty of the State Key Laboratory of Public Big Data, Guizhou University. His research interests include communication network security, communication signal processing, array signal processing, and sparse signal processing.