Journal of Network Intelligence ©2021 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **6**, Number **3**, August **2021**

Trajectory Privacy Protection Based on Spatial-time Constraints in Mobile Social Networks

Zekun Zhang, Siyang Chen, Xiaoting Sun, Yongquan Liang*

College of Computer Science and Engineering
Shandong University of Science and Technology
Qingdao, 266590, China
zekunzhang2018@163.com, 386621631@qq.com, iamsxt103@163.com
*Correponding Author: lyq@sdust.edu.cn

Received April 2021 , Revised June 2021

Abstract. *Massive trajectory data generated with the boom of mobile social networks may contain some private information of users. Attackers can mine active scenes, location and other attribute information of mobile users based on trajectory data. The privacy protection of trajectory data in mobile social networks is of great significance. In this paper, we propose a trajectory privacy protection method based on spatial-time constraints in trajectory publishing in mobile social networks to settle the leakage of privacy in trajectory data publishing. When generating a dummy trajectory, the selection of pivot and the overall movement direction of trajectory are first restricted to prevent sensitive locations from being selected as pivot and to ensure that the overall movement direction of trajectory is similar. Secondly, initial dummy locations generation algorithm based on spatial-time constraints algorithm (IDG) and trajectory generation algorithm based on spatial-time constraints algorithm (TGC) are proposed to generate the next adjacent locations of pivots and other locations, respectively, to ensure that the generated dummy trajectory has the same motion pattern as real trajectory. Experiments are performed to verify the performance of the proposed algorithms.*

Keywords: Location Privacy Protection, Trajectory Publishing, Spatial-time Constraints, Dummy Trajectory Generation

1. **Introduction.** With the continuous evolution of mobile devices, wireless communications [1], and positioning technology, a new application model of location-based service (LBS) has emerged and advanced [2–4]. LBS is a location-based value-added service, clients can request LBS to obtain information services closely related to specified geographic location [5, 6], such as nearby hotels, route planning, and map navigation. At present, LBS has penetrated into all aspects of public lives, transformed public lifestyles and brought great convenience to people [7, 8].

However, while enjoying convenience brought by LBS, people are also confronted with the risk of privacy leakage [9, 10]. Because users' trajectory data contains rich spatial and temporal information, research and mining of trajectory data can obtain abundant information related to mobile users. LBS providers usually publish collected trajectory data to the third parties, such as government departments employ trajectory data for road planning. However, malicious attackers can analyze these trajectory data and combine their own background knowledge to illegally obtain personal privacy such as users' hobbies, religious beliefs, physical conditions, home and work addresses, and even bring economic losses or threaten users [11, 12]. Therefore, in view of this situation, it is indispensable to propose an efficient trajectory privacy protection method that can not only prevent users' trajectory privacy from divulging, but also ensure that published trajectory data has high availability [13, 14].

The current location privacy protection schemes in trajectory publishing are roughly divided into three categories: dummy trajectory generation methods [15–17], trajectory generalization methods [18] and trajectory suppression methods [19–21]. Compared with the latter two methods, dummy trajectory generation methods do not require a trusted third party and can retain complete trajectories information. Therefore, it is often applied to protect user trajectories privacy during trajectory publishing.

The frequently-used dummy trajectory generation methods contain random generation method and rotation generation method [22]. Since random generation scheme method does not consider actual situation, generated dummy trajectories have a large randomness. For example, the generated locations may be inaccessible places such as rivers or swamps, so rotation method is generally utilized to generate dummy trajectories. However, existing dummy trajectory rotation generation method does not combine actual landscape, road conditions, and spatial-time constraints between trajectories in the process of generating dummy trajectories, and the overall movement direction of dummy trajectory is significantly different from true trajectory. Therefore, if users apply the existing dummy trajectory rotation generation method to protect they real trajectory, attackers can identify certain dummy trajectories by combining acquired geographic knowledge, spatial-time constraints between adjacent locations in a single trajectory, and spatial-time correlation between trajectories. Even attain user's true trajectory directly.

An example diagram of dummy trajectory rotation generation method is demonstrated in Figure 1, where solid line represents real trajectory and two dashed lines represent dummy trajectories generated by rotation method. Obviously, this method has shortcomings. First, we can see from this figure that the overall movement direction of dummy trajectory and real trajectory are significantly disparate. Secondly, single location after real trajectory rotated does not consider surrounding environment and adjacent locations may not be reachable within a prescriptive time. In addition, the selected pivot may be a sensitive location such as a hospital, which cannot achieve the objective of user privacy protection. Attackers can eliminate some dummy trajectories based on their background knowledge, increase the probability of identifying true trajectory, and even in some cases can directly guess users' true trajectory.

In order to generate a dummy trajectory similar to real trajectory, this paper elevates traditional dummy trajectory rotation generation method from the perspective of trajectory overall direction, the selection of pivots, and the time accessibility constraints of adjacent locations in trajectory. Our proposed trajectory privacy protection algorithm reduces the probability of attackers identifying real trajectory, thereby achieving the privacy effect of protecting user's real trajectory. Our contributions primarily consist of four aspects:

1. The diversity between the start and end locations of generated dummy trajectories and real trajectory is weeny, ensuring the similarity of overall trajectory movement direction.
2. Through rotation method to generate dummy trajectories, avoid selecting sensitive locations (such as hospitals) as pivots, otherwise the intention of privacy protection cannot be achieved.
3. Adjacent locations in generated dummy trajectory are required to satisfy time accessibility. Ensure that the distance between the front and back adjacent locations in trajectory can be reached within a reasonable time, preventing attackers from

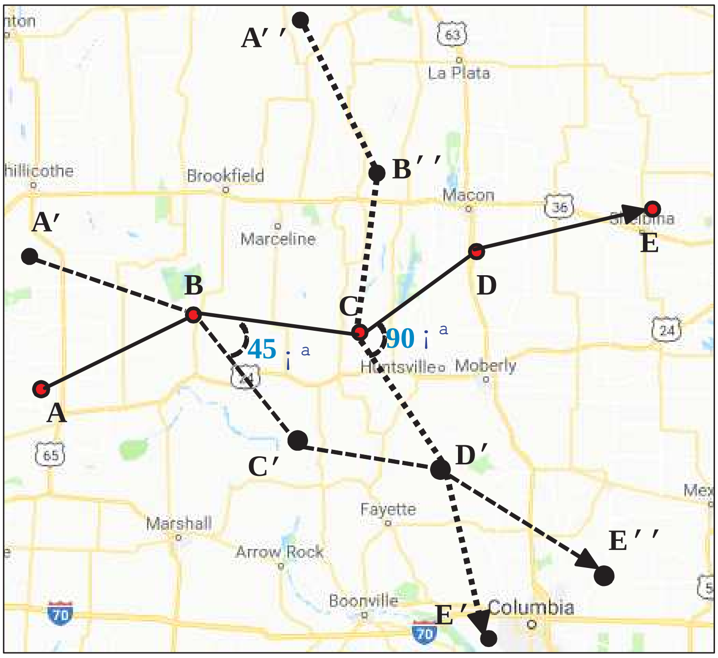

Figure 1. Dummy trajectory rotation method.

identifying dummy trajectories by comparing published time with real time based on map knowledge.

4. Compared with traditional dummy trajectory generation method, we only employ rotation to generate the next location of selected pivots, and propose an original generation method for other locations.

2. **Related Work.** In this section, we investigate trajectory privacy protection approaches in data publishing. When publishing trajectory data, it is essential to meet the high availability of trajectory data and prevent revealing users' sensitive information.

Currently, the most ordinary trajectory privacy protection method is to delete quasi-identifier of each trajectory in trajectory data publishing [20, 23]. There are other attributes in trajectory database, such as age, gender, ect. These attributes are quasi-identifiers. The deliberate attackers can guess user identity utilizing them. In order to resist this malicious attack, many privacy methods, such as k-anonymity [24, 25], l-diversity [26], and confidence bounding [27], have been proposed in the context of relational data.

The k-anonymity [24, 25] resists identity linkage attacks by requiring every quasi-identifier group contain at least k records. l-diversity [26] and confidence bounding aim at preventing attribute linkage attacks.

The dummy trajectory generation method employs a perturbation technique or a pseudonym mechanism to generate dummy trajectory instead of real trajectory or publish a trajectory set. The pivotal problem with this method is that trajectory privacy protection effect cannot be achieved when applying pseudonym technique, and it must be combined with perturbation technology, but the degree of perturbation is not easy to grasp. Once the scope of perturbation is too small, it is hard to achieve the purpose of user trajectory privacy protection. If the scope of perturbation is too large, data availability will be reduced.

In [15], Nergiz *et al.* suggested that utilizing a simple random reconstruction of the original database from the anonymity, to overcome possible drawbacks of generalization approaches. However, the effect of anonymity needs to further improve. In [16], Ryo *et al.* proposed a method of protection the actual user's location based on his/her movements with pauses. The proposed approach generated dummy locations that moved naturally while stopping at several locations like the user; the dummies also took into consideration geographical restrictions. Shuhei *et al.* [17] transformed user's real trajectory into a dummy trajectory utilizing a rotation method, without considering constraints such as movement speed and road network. An attacker can easily exclude a large number of dummy trajectories, resulting in the failure of trajectory privacy protection.

The generalization method leads in a trusted third party (central anonymous server) and employs it to divide locations in trajectory instead of users' real trajectory, thereby achieving the effect of protecting users' trajectory privacy. Among them, The most epidemic method is k-anonymity [24, 25]. The prime problem to be solved by this method is how to exploit an effective way to prevent attackers from inferring users' location in anonymous region based on background knowledge. Wang *et al.* [28] utilized a trusted third party, proposed a scheme to hide the starting point and destination of users' queries into a series of other users' queries to achieve the goal of protecting user trajectory privacy.

The trajectory suppression method [19, 29] refers to directly releasing non-sensitive data, and restricting or conversion certain sensitive locations or certain sensitive trajectory segments in trajectory during trajectory data publishing. In [20], Chen *et al.* introduced local suppression to trajectory data anonymity to improve the resulting data utility which is independent of the underlying data utility metrics and suitable for different trajectory data mining workloads. In order to reduce side effects and improve the performance of the disinfection process, Wu *et al.* [30] proposed an ant colony system-based algorithm called ACS2DT. The proposed algorithm introduces a useful heuristic function to monitor side effects and calculate hidden information. In order to adjust the selection strategy of deleted transactions.

In order to maintain a balance between data utilization and data privacy, Vincent *et al.* [31] proposed a trajectory privacy protection method that carried out generalization of sensitive attributes and local trajectory suppression. Different privacy protection schemes are provided according to the demands of users in trajectory publishing. The disadvantage of this method is that it does not consider the relationship between multiple trajectories and the effect of time on trajectories.

In order to use the deletion behavior of things based on evolutionary computing technology to hide sensitive information, Wu *et al.* [32] designed a sensitive pattern of different lengths within the framework based on genetic algorithm, used stricter thresholds to protect sensitive information, and used record deletion technology. Data cleaning can effectively hide sensitive information in medical situations. Wu *et al.* [33] improved the authentication protocol in the distributed cloud computing environment to ensure perfect forward secrecy and avoid attacks by privileged insiders. This solution improves efficiency and achieves higher security standards.

Wu *et al.* [34] proposed a multi-objective algorithm based on a grid method to find the best solution as a candidate for disinfection, which can significantly reduce the side effects of disinfection and speed up the calculation cost algorithm. Wang *et al.* [35] proposed a forward privacy protection scheme for IoT medical systems, using a searchable scheme to achieve mental protection and searchable functions, and solve the problem of using only partial search results to verify search results in top-k search scenarios. The question of correctness can protect the privacy of the healthcare system that supports the Internet of Things.

Chen *et al.* [36] proposed a dynamic solution to prevent information leaks from outsourced data or search keywords when updating data. It retains forward privacy and backward privacy as the two security features of the solution. Experiments prove that this is a The number one searchable encryption scheme, and will not cause data leakage. Wang *et al.* [37] proposed an incentive evolutionary game model that encourages cooperation between nodes. Through the self-evolution of the evolutionary game, both normal and abnormal nodes are encouraged to participate in network collaboration. This model can effectively improve network performance and can be extended to various networks.

As a whole, trajectory data privacy protection should comprehensively consider the time factor in trajectory, the overall movement direction of trajectory, and the speed of movement. Fortunately, our propose trajectory privacy protection algorithm taking the above factors into consideration. Through security analysis and simulation experiments, it is verified that our algorithm can not only successfully protect users' trajectory privacy, but also greatly improve data availability.

3. Related definitions and system model.

In this section, we give the related definitions and system model.

3.1. Related definitions.

In the process of generating dummy trajectories, we propose a trajectory privacy protection algorithm based on spatial-time constraints from a global and local perspective. To facilitate the description of the algorithm, several related definitions are given in this. Among them, individual location recognition rate in trajectory and trajectory recognition rate are widely applied as trajectory privacy protection measures in trajectory data publishing [38].

Definition 1 (Trajectory set (TC)) [38]: The trajectory set TC refers to a set that comprises m dummy trajectories $T^{'}$ and real trajectories T_{real} formed by mobile clients during trajectory publishing, and is formalized as:

$$TC = \{T_1^{'}, T_2^{'}, ..., T_m^{'}, T_{real}\}, \tag{1}$$

$|TC|$ denotes the number of trajectories in the set.

When trajectories are indistinguishable in TC, the divulged probability of true trajectory is $\frac{1}{|TC|}$.

From a local perspective, we define individual location recognition rate in the trajectory to measure the degree of privacy protection in local trajectory.

Definition 2 (Individual location recognition rate in trajectory (SR)) [38]: The individual location recognition rate SR in trajectories is the average of all location recognition rates on trajectory. It is assumed that trajectory set TC during trajectory data publishing, and each trajectory has n locations. Simultaneously, attackers have relevant background knowledge that can be employed to identify real trajectory. The set K_i is formed by location point on the real trajectory and location points on the dummy trajectory in the trajectory set TC at time t_i, and the probability of recognizing real location points at t_i is $\frac{1}{|K_i|}$. The calculation formula of SR is:

$$SR = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{|K_i|}. \tag{2}$$

From a global perspective, we define trajectory recognition rate to measure the degree that entire trajectory is protected.

Definition 3 (Trajectory recognition rate (TR)) [38]: Trajectory recognition rate is utilized to measure the probability that an attacker will recognize real user's trajectory in trajectory set. Let trajectory set TC contain m trajectories during trajectory data publishing, among them, there are k trajectories with intersections, and then there are

$(m-k)$ trajectories without intersections. Assume that the number of trajectories composed of k trajectories with intersections is T_k. The probability of attacker identifying true trajectory TR is:

$$TR = \frac{1}{T_k + (m-k)}. \tag{3}$$

This paper exploits trajectory distortion to measure the degree of data availability, and its definition is as follow.

Definition 4 (Trajectory distortion (TD)) [38]: Trajectory distortion is a measure of the availability of information after privacy protection processing of trajectory data. The greater trajectory distortion, the less useful information remains in trajectory. Assume that there are K dummy trajectories in trajectory set TC, and each trajectory has m locations. In this paper, the trajectory distortion TD is defined as the average distance between location in real trajectory and numerous locations in dummy trajectory, which is formalized as:

$$TD = \frac{1}{k}\sum_{j=1}^{m}(\frac{1}{K}\sum_{i=1} Kdist(u_j^i, u_j^{real})), \tag{4}$$

where u_j^i is the j-th location of the i-th trajectory, and u_j^{real} is the j-th location of the real trajectory.

3.2. **System Model.** In trajectory data publishing, our system structure adopts "collect, process, and publish" criteria. Specifically, data collection server (LBS server) collects the trajectory data of mobile users firstly, and then the original trajectory is processed by the trajectory privacy protection server. Finally the processed trajectory data is distributes to a third party (such as government, scientific research units, and universities). Therefore, our system model contains three phases: 1) trajectory data collection phase; 2) trajectory data privacy processing phase; 3) trajectory data mining and analysis phase, as shown in Figure 2:

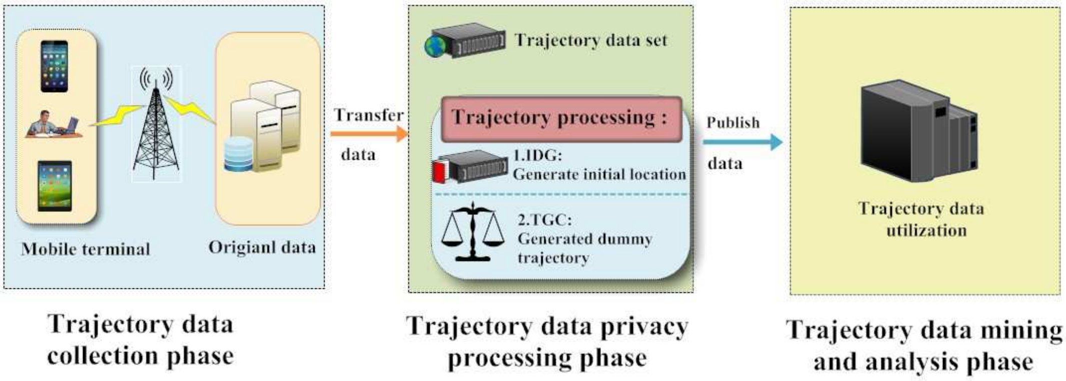

Figure 2. System model diagram.

1. Trajectory data collection phase: This phase is mainly employed to collect trajectory data, which formed by mobile clients continuously requesting LBS services for a period of time.
2. Trajectory data privacy processing phase: As the core part of system structure, in the privacy processing of trajectory data, it is indispensable to prevent attackers from deriving the sensitive information of target user via combining trajectory and

background knowledge they have mastered. The published trajectory data is required to have high availability as well. Only by considering this two goals can various types of applications based on trajectory data analysis and mining play an outstanding role. Here, the trajectory privacy protection algorithm based on spatial-time constraints is proposed for protection trajectory privacy. First, IDG algorithm is adopted to generate initial location that satisfy spatial constraints, and then TGC algorithm is applied to generate dummy trajectories that meet spatial-time constraints and mix them with real trajectories to form a trajectory set for trajectory data publishing.

3. Trajectory data mining and analysis phase: The trajectory data contains vast space-time-related information. The analysis and mining of trajectory data can support a large number of mobile-related applications. For example, the government can implement road planning based on the collected trajectory data; the travel company (Didi etc.) can optimize the driver's order-taking route based on the collected trajectory data; the merchant can analyze the user's behavior pattern and make business decisions based on user's trajectory information.

4. **Algorithm design.** It is essential to give consideration to the two goals of trajectory privacy protection and trajectory data availability when releasing trajectory data. In order to accomplish above goals, this paper proposes a trajectory privacy protection algorithm based on spatial-time constraints. The algorithm is mainly divided into two parts: 1)Initial dummy locations generation algorithm based on spatial-time constraints (IDG) is adopted to generate dummy locations corresponding to the next adjacent locations of non-pivot. 2) Trajectory generation algorithm based on spatial-time constraints (TGC) is employed to generate dummy trajectories. Below we will expound the specific process of algorithms in detail.

4.1. **Privacy protection algorithm for trajectory based on spatial-time constraints.** Given the real trajectory of mobile user, in order to ensure that the generated dummy trajectories and real trajectory have similar movement directions, the start and end locations of dummy trajectories need to be within a reasonable range from the real trajectory's, that is they are in the same circular region.

When generating the next adjacent location of non-pivot, this paper propounds an initial dummy locations generation algorithm based on spatial-time constraint as part of STC algorithm. The specific approach is: randomly generated a location point pos_0 near location l_i in real trajectory at time t_i is as the center of circle, which meets $r_{min} \leq$ dist(l_i, pos_0)$\leq r_{max}$, where r_{min} and r_{max} are the minimum and maximum radius from location l_i. These two parameters are customized according to user requirements.

Then, using the selected initial dummy location pos_0 and real location l_i where the straight line is rotated counterclockwise, select a dummy location $pos^{'}$ through angle $\theta = 2\pi/k$. In addition, it is that makes distance between the generated dummy position $pos^{'}$ and initial location pos_0 meeting $dist(pos^{'}, pos_0) = \rho \cdot r$, to generate total k dummy locations. Among them, ρ is a random number between 0 and 1 ($0 < \rho < 1$), and r is the average of r_{min} and r_{max}. In order to ensure the rationality of generating dummy locations, we also consider the actual road distribution. Algorithm 1 reveals the initial dummy locations generation algorithm based on spatial-time constraints (IDG) pseudo code.

4.2. **Trajectory generation algorithm based on spatial-time constraints.** In Algorithm 1, dummy locations set corresponding to real trajectory at time t_i is generated, that is, the initialization of the next adjacent location of non-rotational pivot is completed. After that, selecting pivot and the next neighbor location generated by the rotation of

Algorithm 1 : Initial dummy locations generation algorithm based on spatial-time constraints (IDG)

```
Input: Real trajectory $T$, $r_{min}$, $r_{max}$, location $l_i$ at time $t_i$, the number of lotions $m$ in initialized trajectory.
Output: Dummy locations set $L$ (The next adjacent location set of a non-rotating pivot) at time $t_i$.
Initialize the location collection $L$
for $i = 1$ to $m$ do
  $L \leftarrow L \cup \{l_i\}$
  Generate a random location $pos_0$, as the center of circle, satisfying $r_{min} \leq dist(l_i, pos_0) \leq r_{max}$
  for $j = 1$ to $k - 1$ do
    Rotate $pos_0$ and $l_i$ by $2\pi/k$ angle
    According to road distribution and formula $dist(pos', pos_0) = \rho \cdot r$, generate a dummy location $pos'$
    $L \leftarrow L \cup \{pos'\}$
  end for
end for
return $L$
```

pivot is implemented by trajectory generation algorithm based on spatial-time constraints (TGC). At the same time, this algorithm is another part of STC algorithm.

In the process of generating the next neighbor positions of pivot, first randomly pick a few locations as pivot. Since the location has sensitive semantics, determine whether the selected pivot is a sensitive location (hospital, church, etc.) based on the geographic semantic label of location. If it is a sensitive location, it needs to be reselected until the selected pivot is a non-sensitive location.

On the one hand, we randomly choose an angle within the scope of $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$ to rotate at pivot, so that the overall movement direction of dummy trajectories is similar to the overall direction of real trajectory.

On the another hand, according to the distance of real trajectory, the time interval of publication and the background knowledge of map, TGC determines the average speed $\overline{v}$ of the target user on the road, and calculates the distance user moves in a time period according to formula $s = \overline{v} \cdot t$. The reasonable location near s is the next location of dummy trajectory to satisfy time accessibility. Algorithm 2 gives trajectory generation algorithm based on spatial-time constraints (TGC) pseudo-code.

5. Simulation experiment and performance analysis.

Here, we perform simulation experiments on the proposed trajectory privacy protection algorithm based on spatial-time constraints and elaborate performance analysis.

5.1. Experiment environment.

We compare STC algorithm with random trajectory generation method (Random) [22], trajectory rotation generation method (Rotation) [22], and SPP [31] algorithm proposed by Bindschaedler *et al.*

The hardware configuration of simulation experimental environment is shown in Table 1. The experimental data comes from trajectory data set collected by Microsoft Research Asia (MSRA) [4], which contains 17,621 trajectory data gathered from 182 users from March 2007 to August 2012, with a total distance of more than 1.2 million kilometers. Moreover, these trajectories are marked by a time stamped GPS point, and each location

Algorithm 2 : Trajectory generation algorithm based on spatial-time constraints (TGC)

```
Input: The number of pivot $m$, real trajectory $T$.
Output: Dummy trajectories set $ST_d$.
$C = \phi$ // $C$ is empty set, which will store the selected pivot
for $i = 1$ to $m$ do
  Randomly select a location $l_i$ among the $n$ locations of real trajectory $T$
  if $l_i$ is a sensitive location then
    Reselect location $l_i$
  else
    $C \leftarrow C \cup l_i$
  end if
end if
The pivot rotates $\theta$ angle ($\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$)
Calculate time interval $t$ between pivot and the next adjacent location of pivot according to real trajectory $T$
Calculate moving distance $s$ according to $s = \overline{v} \cdot t$
Randomly select a location near $s$ along rotation direction as the next adjacent locations of pivot to meet time accessibility
The generated next adjacent locations are randomly combined with locations $L$ generated by Algorithm 1 to form $k$ dummy trajectories
return $ST_d$
```

contains information such as latitude, longitude and altitude. The advantage of this trajectory data is that it covers a extensive range, not only the trajectory data of user's home address and work place, but also the trajectory of user's outdoor activities, such as travel and shopping. Therefore, this trajectory data is universally utilized in multiple research fields such as location-based privacy protection and location-based social network services.

The detailed parameters of experimental data we applied are given in Table 2. Among them, S represents the number of trajectories, n indicates the number of locations contained in each trajectory. Here we intercept each trajectory at 20 locations, $Period$ manifests the time span of trajectory data, and U_{num} represents the number of users included in the trajectory data.

5.2. **Experiment Result and analysis.**

5.2.1. *Individual location recognition rate in trajectory vs. number of dummy trajectories.* Figure 3 demonstrates the relationship between individual location recognition rate in trajectory and number of dummy trajectories. The individual location recognition rate in trajectory can also indicate the privacy protection degree of local trajectories to a certain extent. It can be seen from Figure 3 that the larger the number of dummy trajectories,

Table 1. Experimental environment hardware configuration

Operating system	Windows7 64 bits
RAM	32G
Processor	Intel(R) Core i7-8700k
Programming language	Python
Programming environment	Pycharm

the lower individual location recognition rate in trajectories of four algorithms. Assume that attackers possess background knowledge such as map information. Since the Random algorithm randomly selects locations to construct dummy trajectory, and the selection of pivot is also random, without considering background knowledge and spatial-time constraints. Therefore, the individual location recognition rate is the highest, and the degree of privacy protection of user is terrible. SPP algorithm can apply the geographic and semantic features of real trajectories to generate dummy trajectories, so individual location recognition rate in trajectory is low, but the algorithm does not take into account the temporal accessibility of adjacent locations. The STC algorithm proposed in this paper avoids the selection of sensitive locations as pivot, and generating the next adjacent location of pivot and other locations takes into account spatial-time constraints, so it has a low individual location recognition rate in trajectory. In other words, STC algorithm has higher local privacy protection capabilities.

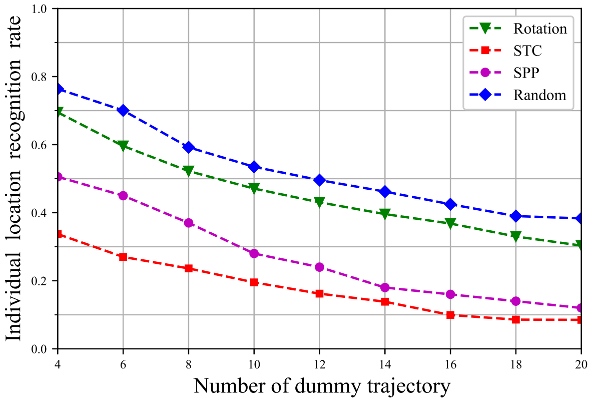

Figure 3. Individual location recognition rate in trajectory vs. Number of dummy trajectory.

5.2.2. *Trajectory recognition rate vs. number of dummy trajectories.* Figure 4 portrays the relationship between trajectory recognition rate and number of dummy trajectories. As you can see from this figure, with the increase of number of dummy trajectories, trajectory recognition rate of this four algorithms are decreasing. Because the Random and SPP algorithms have difficulty in generating intersections between trajectories, this two algorithms have higher trajectory recognition rate under the same number of dummy trajectories. However, SPP algorithm is superior, because SPP excludes some dummy trajectories generated by Random algorithm according to background knowledge.

In order to generate dummy trajectories, Rotation algorithm and STC algorithm proposed in this paper both select pivot to rotate, so it can generate more dummy trajectories in dummy trajectory set, reducing the probability of real trajectories being identified. Compared with STC algorithm, Rotation algorithm rotates the entire trajectory after

Table 2. Experimental data parameters

Parameters	Value
S	10000
n	20
$Period$	2007.3 $\sim$ 2012.8
U_{num}	182

selecting pivots, and some locations in trajectory after rotation may not meet the constraints of road network and are easily excluded by the attackers. The pivots selected by the STC algorithm are non-sensitive location, and the next location of rotation output expects to meet spatial-time constraints, which has a more outstanding strength of location privacy protection.

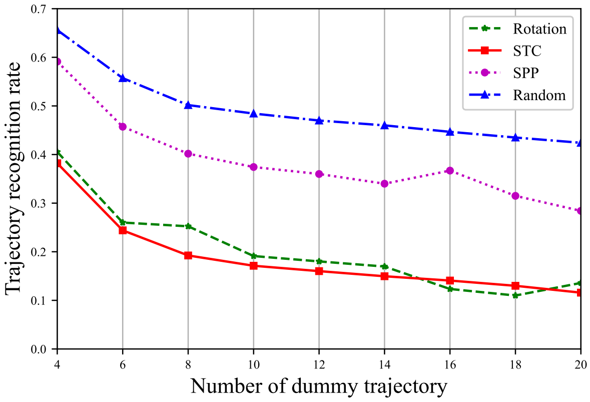

Figure 4. Trajectory recognition rate vs. Number of dummy trajectory.

5.2.3. *Cost time vs. number of trajectories.* Figure 5 manifests the relationship between cost time and number of trajectories. It can be seen from this figure that as the number of trajectories increases, cost time by four algorithms also increases. Since Random algorithm randomly selects locations to construct dummy trajectories, it does not consider other constraints, so it takes the least time to protect privacy information. The SPP algorithm prudently considers the semantic characteristics of each location and synthesizes a dummy trajectory in real trajectory, of course, it takes the most time to execute anonymous process. The STC algorithm devotes more time because it is necessary to judge the sensitivity of pivot, but it is better than Rotation algorithm. The Rotation algorithm randomly selects pivots and rotates to generate dummy trajectories, and determines dummy trajectories that satisfy constraint conditions, so it takes relatively more time.

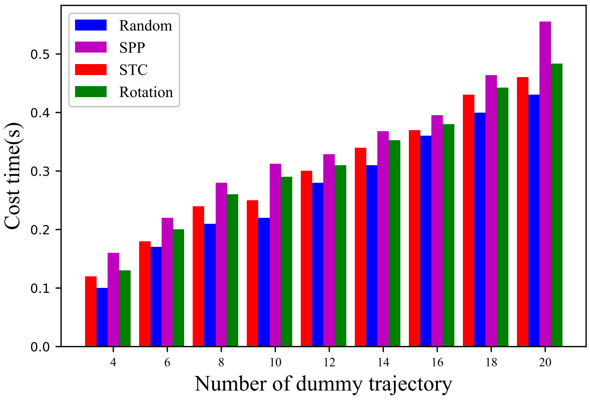

Figure 5. Cost time vs. Number of trajectory.

5.2.4. *Trajectory distortion vs. number of dummy trajectories.* Figure 6 demonstrates the relationship between trajectory distortion and number of dummy trajectories. Intuitively, trajectory distortion is supposed to controlled within a reasonable range, not too large or too small. Otherwise, excessive trajectory distortion affect the availability of trajectory data, and excessively small fake trajectory distortion will gather together trajectories, failing to achieve the purpose of trajectory privacy protection. It can be seen from the experimental results that trajectory distortion of other three algorithms increases with the increase of the number of dummy trajectories. This is because Random and Rotation algorithms have randomness largely when constructing dummy trajectories, which leads to excessive trajectory distortion and low availability of data. The SPP algorithm does not consider spatial-time constraints, and will generate locations with excessive offset distances. Therefore, as the number of dummy trajectories increases, the degree of information distortion will be too large. However, STC algorithm proposed in this paper imposes spatial-time constraints on the rotated next adjacent locations and other locations, so trajectory distortion is within a reasonable range.

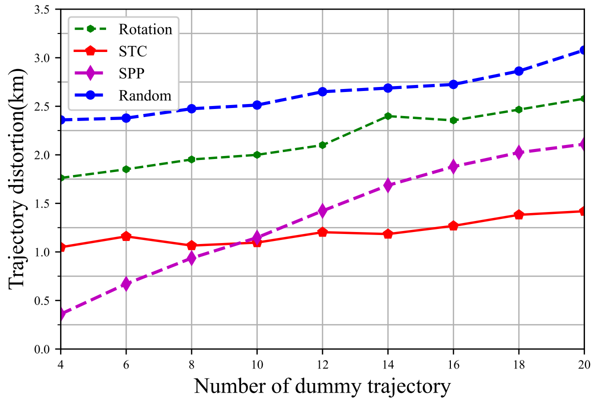

Figure 6. Trajectory distortion vs. Number of dummy trajectories.

5.2.5. *Rotation angle vs. the area of anonymous region.* Figs. 7 and 8 delineate the influence of k value and rotation angle θ on the anonymous region area of trajectory in our STC algorithm. In order to maintain the overall consistency of trajectory moving direction, trajectory rotation angle is specified at $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. Figure 7 describes the effect on the anonymous area of trajectory when rotation angle is $-\frac{\pi}{2} \leq \theta \leq 0$. It can be seen that as rotation degree decreases, the area of the anonymous region decreases.

Figure 8 depicts the effect on the anonymous area of trajectory when rotation angle is $0 \leq \theta \leq \frac{\pi}{2}$. Obviously, the area of anonymous region enlarges with the increase of rotation angle. In addition, at the same rotation angle, the larger k value, the larger anonymous region. The area of the trajectory anonymous region ought to be within a rational range. If the area of trajectory anonymous region is too small, it is ordinary to expose the true trajectory of target user. If the area of trajectory anonymous region is too large, system overhead is increase to a great extent. As can be seen in figure, in order to achieve a balance between trajectory privacy and system overhead, the absolute value of rotation angle is most logical between 40 and 50.

6. **Conclusions.** With the development of wireless communication equipment and sensor technology, users are producing a large amount of trajectory data. The sharing of these data can help to conduct user behavior analysis, route planning and road condition

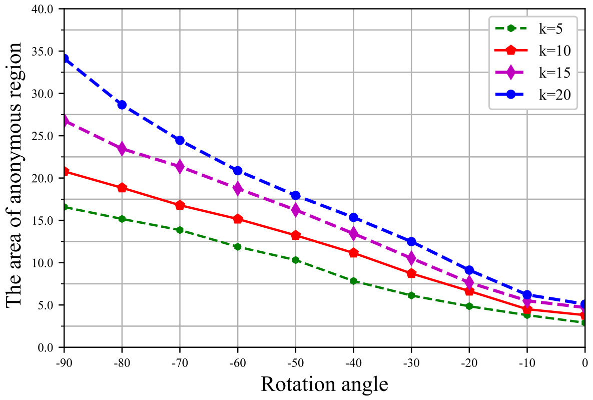

Figure 7. Rotation angle ($-\frac{\pi}{2} \leq \theta \leq 0$) vs. The area of anonymous region.

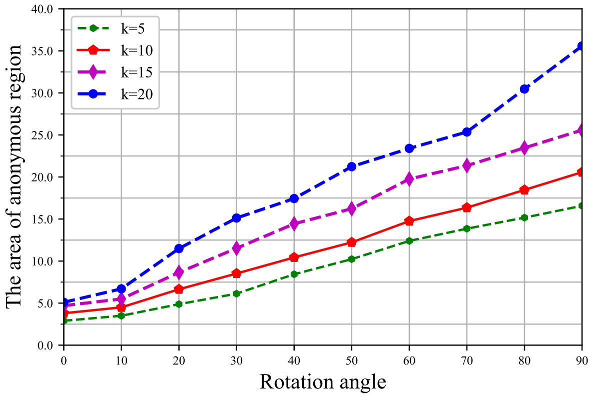

Figure 8. Rotation angle ($0 \leq \theta \leq \frac{\pi}{2}$) vs. The area of anonymous region.

analysis, and better provide users with rich services. However, a deliberate attacker can also mine the user's sensitive information from the published data and make a link attack, causing the user's property. Therefore, protecting published trajectory data can prevent the disclosure of privacy.

This paper focuses on privacy protection in trajectory data publishing, and proposes a trajectory privacy protection algorithm based on spatial-time constraints. The algorithm improves traditional rotation generation method in dummy trajectories generation, and makes spatial-time constraints on the overall movement direction of dummy trajectories, the selection of pivot, generating the next adjacent locations of pivot, and the generation of other locations. Finally, simulation experiments verify that our algorithm can achieve a balance between data availability and location privacy protection.

In future researches, we will further analyze user trajectories exploring for more other factors, such as unvarying locations and quasi-identifiers factor. These factors are employed to protect users' trajectory privacy and provide users with more convenient, fast and secure services. In addition, how to evaluate the usage of trajectory data by LBS providers can also be studied accordingly.

7. **Acknowledgments.** This work is supported by Humanity and Social Science Fund of the Ministry of Education under Grant 20YJAZH078 and 20YJAZH127, Open Project of Tongji University Embedded System and Service Computing of Ministry of Education

of China under Grant ESSCKF 2019-06 and ESSCKF 2019-08, National Innovation and Entrepreneurship Training Program for College Students under Grant 201910424014.

REFERENCES

[1] Y. Ping, X. Yue, X. Ming and S. Q. Li, 6G Wireless Communications: Vision and Potential Techniques, *IEEE Network*, vol. 33, no. 4, pp. 70-75, 2019.

[2] L. Chen, S. Thombre, K. Jarvinen, L. E. Simona, A. S. Anette, L. Helena, B. M. Zahidul H, B. Shakila, F. G. Nunzia, H. Salomon, Robustness, security and privacy in location-based services for future IoT: A survey, *IEEE Access*, vol. 5, pp. 8956-8977, 2017.

[3] H. R. Hwang, Y. L. Hsueh, H. W. Chung, A Novel Time-Obfuscated Algorithm for Trajectory Privacy Protection, *IEEE Transactions on Services Computing*, vol. 7, no. 2, pp. 126-139, 2013.

[4] C. Bettini, Privacy protection in location-based services: A survey, *Handbook of Mobile Data Privacy*, pp. 73-96, 2018.

[5] J. P. Lin, J. W. Niu, H. Li, M. Atiquzzaman, A Secure and Efficient Location-based Service scheme for smart transportation, *Future Generation Computer Systems*, vol. 92, pp. 694-704, 2019.

[6] J. Q. Zhang, X. Wang, Y. F. Yuan, L. N. Ni, RcDT: Privacy Preservation Based on R-constrained Dummy Trajectory in Mobile Social Networks, *IEEE Access*, vol. 7, pp. 90476-90486, 2019.

[7] J. Chen, H. Kun, Q. Yuan, M. Chen, Y. Xiang, Blind Filtering at Third Parties: An Efficient Privacy-Preserving Framework for Location-Based Services, *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2524-2535, 2018.

[8] S. Wang , X. Meng, J. Yu, R. Bie, Y. Sun, X. Cheng, N-in-One: A Novel Location-Based Service, *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5274-5286, 2018.

[9] L. N. Ni, L. Chao, X. Wang , H. L. Jiang, J. G. Yu, DP-MCDBSCAN: Differential privacy preserving multi-core DBSCAN clustering for network user data, *IEEE Access*, vol. 6, pp. 21053-21063, 2018.

[10] K. Emre, P. T. Brochmann, S. Erkay, S. Yucel, Privacy Risks in Trajectory Data Publishing: Reconstructing Private Trajectories from Continuous Properties, *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pp 642-649, 2008.

[11] L. N. Ni, F. L. Tian, Q. H. Ni, Y. Yan, J. Q. Zhang, An anonymous entropy-based location privacy protection scheme in mobile social networks, *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 93-111, 2019.

[12] J. Q. Zhang, Y. F. Yuan, X. Wang, L. N. Ni, J. G. Yu, M. M. Zhang, RPAR: Location Privacy Preserving via Repartitioning Anonymous Region in Mobile Social Network, *Security and Communication Networks*, vol. 2018, 6829326, 2018.

[13] M. Li, L. H. Zhu , Z. J. Zhang, R. X. Xu, Achieving differential privacy of trajectory data publishing in participatory sensing, *Information Sciences*, vol. 400, pp. 1-13, 2017.

[14] X. Zhe, J. J. Yang , M. Huang, P. Loganathan, X. J. Fu, G. Mong, QLDS: A Novel Design Scheme for Trajectory Privacy Protection with Utility Guarantee in Participatory Sensing, *IEEE Transactions on Mobile Computing*, vol. 17, no. 6, pp. 1397-1410, 2018.

[15] N. M. Ercan, A. Maurizio, S. Yucel, Towards Trajectory Anonymization: A Generalization-Based Approach, *Transactions on Data Privacy*, vol. 2, no. 1, pp. 52-61, 2009.

[16] K. Ryo, I. Mayu, H. Takahiro, S. Akiyoshi, N. Shojiro, A dummy-based anonymization method based on user trajectory with pauses, *Proc. of the 20th International Conference on Advances in Geographic Information Systems*, pp. 249-258, 2012.

[17] H. Shuhei, A. Daichi, H. Takahiro, X. Xie, Dummy generation based on user-movement estimation for location privacy protection, *IEEE Access*, vol. 6, pp. 22958-22969, 2018.

[18] H. Nattapon, N. Juggapong, R. Surapon, Privacy preservation for trajectory data publishing by look-up table generalization, *Proc. Australasian Database Conference (ADC)*, pp. 15-27, 2018.

[19] T. Manolis, P. Giorgos, M. Nikos, S. Spiros, Local suppression and splitting techniques for privacy preserving publication of trajectories, *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 7, pp. 1466-1479, 2017.

[20] R. Chen, C. M. Fung Benjamin, M. Noman, C. Desai Bipin, K. Wang, Privacy-preserving trajectory data publishing by local suppression, *Information Sciences*, vol. 231, pp. 83-97, 2013.

[21] Z. M. Zhao, J. B. Wang, C. L. Sun, Y. W. Yuan, B. Li, Method of trajectory privacy protection based on restraining trajectory in LBS, *International Journal of Information and Communication Technology*, vol. 13, no. 3, pp. 329-339, 2018.

[22] P. R. Lei, W. C. Peng, I. J. Su, C. P. Chang, Dummy-based schemes for protecting movement trajectories, *Journal of Information Science and Engineering* , vol. 28, no. 2, pp. 335-350, 2012.

[23] P. P. Sui, T. Y. Wo, Z. L. Wen, X. X. Li, Privacy-preserving trajectory publication against parking point attacks, *Proc. IEEE 2013 10th Inter. Conf. on Ubiquitous Intelligence and Computing and Autonomic and Trusted Computing* , pp. 569-574, 2013.

[24] T. Zhen, K. Zhao, F. L. Xu, Y. Li, S. Li, D. P. Jin, Protecting Trajectory From Semantic Attack Considering k-Anonymity, l-Diversity, and t-Closeness, *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 264-278, 2018.

[25] T. Zhen, K. Zhao, F. L. Xu, Y. Li, S. Li, D. P. Jin, Beyond k-anonymity: Protect your trajectory from semantic attack, *Proc. 2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1-9, 2017.

[26] M. Ashwin, G. Johannes, K. Daniel, M. Venkitasubramaniam, l-Diversity: Privacy Beyond k-Anonymity, *Proc. of the 22nd International Conference on Data Engineering*, 2006.

[27] K. Wang, B. J. Fung, P. Yu, Handicapping attacker's confidence: An alternative to k-anonymization, *Knowledge and Information Systems*, vol. 11, no. 3, pp. 345-368, 2007.

[28] S. Wang, S. Richard, N. Surya, Privacy-protected statistics publication over social media user trajectory streams, *Future Generation Computer Systems*, vol. 87, pp. 792-802, 2018.

[29] K. E. Ghasemi, A. Mahdi, D. Fatemeh, PPTD: Preserving personalized privacy in trajectory data publishing by sensitive attribute generalization and trajectory local suppression, *Knowledge-Based Systems*, vol. 94, pp. 43-59, 2016.

[30] M. T. Wu, J. Zhan, J. Lin, Ant Colony System Sanitization Approach to Hiding Sensitive Itemsets, *IEEE Access*, vol. 5, pp. 10024-10039, 2017.

[31] B. Vincent, S. Reza, Synthesizing plausible privacy-preserving location traces, *Proc. 2016 IEEE Symposium on Security and Privacy (SP)*, pp. 546-563, 2016.

[32] M. T. Wu, G. Srivastava, A. Jolfaei, P. Fournier-Viger and C. W. Lin, Hiding sensitive information in eHealth datasets, *Future Generation Computer Systems*, vol. 117, pp. 169-180, 2021.

[33] T. Y. Wu, Z. Lee, M. S. Obaidat, S. Kumari, C. M. Chen, An Authenticated Key Exchange Protocol for Multi-server Architecture in 5G Networks, *IEEE Access*, vol. 8, pp. 28096-28018, 2020.

[34] T. Y. Wu, J. Lin, Y. Zhang and C. H. Chen, A Grid-Based Swarm Intelligence Algorithm for Privacy-Preserving Data Mining, *Applied Sciences*, vol. 9, no. 4, 774, 2019.

[35] K. Wang, C. M. Chen, Z. Tie, M. Shojafar and S. Kumari, Forward Privacy Preservation in IoT-Enabled Healthcare Systems, *IEEE Transactions on Industrial Informatics*, https://doi.org/10.1109/TII.2021.3064691 , 2021.

[36] C. M. Chen, Z. Tie, E. K. Wang, M. K. Khan, S. Kumar and S. Kumari, Verifiable dynamic ranked search with forward privacy over encrypted cloud data, *Peer-to-Peer Networking and Applications*, https://doi.org/10.1007/s12083-021-01132-3, 2021.

[37] K. Wang, C. M. Chen, Z. Tie, M. Shojafar and S. Kumari, Incentive evolutionary game model for opportunistic social networks, *Future Generation Computer Systems*, vol. 102, pp. 14-29, 2020.

[38] Y. Zheng, X. Xie, W. Y. Ma, GeoLife: A collaborative social networking service among user, location and trajectory, *IEEE Data Engineering Bulletin*, vol. 33, no. 2, pp. 32-39, 2010.