# Trajectory Privacy Protection Scheme for Different Travel Modes

Yanzi Li[1,2], Jing Zhang[1,2(✉)], Peng Gao[3], and Sitong Shi[1,2]

[1] School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, Fujian, China
[2] Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou 350118, Fujian, China
[3] Southest University, Nanjing 210096, Jiangsu, China

**Abstract.** The problem of trajectory privacy is getting more attention in recent years. It is unreasonable to adopt the same privacy protection scheme in diversified travel modes, since the trajectories have different constraints under different travel modes. Aiming at this problem, a mobile trajectory privacy protection scheme for different travel modes is proposed in this paper. At first, the overall privacy protection scheme is designed according to the classification of privacy level of road sections and different travel modes of users. And then, by considering both service and privacy, two false trajectory generation algorithms are designed, which can protect the real trajectory. The simulation experiments show that, the proposed scheme can process trajectory anonymity effectively. Compared with the algorithm in non-road network environment, the road network based algorithm can improve the average service accuracy by 36%, and decrease the privacy leak by 78%, which can provide users with more efficient and secure location services.

**Keywords:** Privacy protection · Trajectory · Road network · Anonymity

## 1 Introduction

Due to the rapid development of intelligent terminals with location function and mobile Internet, more and more LBS (Location Based Service) have been integrated into people's daily life [1]. For example, querying the hotel, hospital, gas station within one kilometer, the least time-consuming path to a place. However, in the case of obtaining the convenience of LBS intelligent service, privacy will certainly confront the danger of being maliciously used. The location information is provided by people to obtain accurate services. And this information will also be obtained by malicious users. The malicious users can even obtain users' trajectory information through continuous query, so as to analyze their sensitive information, such as personal preferences, behavior patterns, financial status, etc. How can people enjoy het convenience of intelligent services while avoiding information leakage? This has become an urgent problem to be solved in today's society, and also the focus of scholars in recent years [2].

At present, the main method of location privacy protection is to confuse the attackers by mixing real and false location [3]. Location $K$-anonymity is a location confusion


L. Lin et al. (Eds.): SocialSec 2021, CCIS 1495, pp. 133–144, 2021.
https://doi.org/10.1007/978-981-16-7913-1_10

privacy protection model first proposed by Marco Gruteser [4]. Later, scholars have studied a variety of protection schemes around this idea, mainly including schemes based on false location [5, 6], spatio-temporal correlation combined anonymity scheme [7], $K$-anonymity based on Hilbert curve value [8], etc. They focus on single point query in European space. For example, when shopping mall, it is possible to initiate the query service in any location of the mall. However, if people initiate the query on the vehicle, the situation will be different. At this time, the privacy scheme in the traditional European space cannot meet the privacy requirements of the road network [9]. Literatures [9–15] are all aimed at the privacy protection of vehicle information and consider the problem of trajectory leakage from the perspective of prediction, etc., but seldom consider the environmental information of road network. Throughout the existing studies, there is still a lack of trajectory privacy protection scheme for different travel modes. It is not feasible to use the same scheme for free European space and road network environment. It is also not feasible to use the same method for the protection of trajectory position as for the protection of single point position.

Based on the above information, trajectory privacy protection schemes under different environments are designed:

1) Formulate non-road network or road network protection methods according to travel forms and demands;
2) According to the privacy level of road network map, the trajectories on the road with high risk are protected;
3) Avoid trajectory matching and associating with user to prevent potential information leakage.

## 2 Model Description

### 2.1 System Model Description

$K$-anonymous trajectory set: an anonymous trajectory set composed of the user's real trajectory and $K - 1$ false trajectories generated by the false trajectory generation algorithm, represented by ATS $= \{L_1, L_2, \ldots, L_K\}$. In this paper, $L_1$ represents the real trajectory, and others represent false trajectories.

Location privacy protection system consists of three parts: user, privacy protection server and LBS. The privacy protection server includes a module for generating false trajectory set and a module for query feedback. Among them, the generation module of false trajectory set is responsible for generating multiple virtual trajectories, which are published after confusing with the real trajectory; the query feedback module is responsible for filtering the feedback results of LBS. Request query includes travel mode, location, request content and so on. according to the travel modes and privacy requirements, the privacy protection server selects the false trajectory generation algorithm to generate and process the anonymous set of trajectory, and sends the processed anonymous information to LBS. The matching results which calculated by LBS are fed back to the privacy protection server. The filtered real service is fed back to the requesting user by the privacy protection server.

### 2.2 Trajectory Privacy Protection Scheme

This scheme is composed of three modules: area map matching module; section privacy level determination module; virtual trajectory generation module.

Area map matching: First, the query service is initiated by the user, single point or continuous query. Such as query point A to point B of the best driving path or several continuous queries surrounding food. The trajectory area is preliminarily determined through query, and the road network structure of this region is established.

Determine the privacy level of the road: each road section in the area is divided, and the privacy level of each road section is determined. For example, the urban main road has dense traffic flow and almost 90% of vehicles choose this section, so the privacy level can be set as a low level, and the probability of densely populated areas being located by attackers is low. The privacy level of the road with few people should be set higher. In extreme cases, this section has only one user at the moment, it must be the query user, the risk is very high. In this model, the privacy protection server assumes that information such as road privacy level can be obtained.

The false trajectory generation algorithms are determined according to the privacy level of road section and user demand: A. False Trajectory Algorithm without Road Net (FTA); B. False Trajectory Algorithm Based on Road Network (RNFTA). The algorithms, which are the focus of this article, will be discussed in detail in part 3.

## 3 False Trajectory Generation Algorithm

### 3.1 False Trajectory Algorithm Without Road Net (FTA)

FTA algorithm is suitable for walking users, without road network environment. Finding the appropriate distance between true and false trajectory points is the key of this algorithm: if the distance is too close, the generated false trajectory does not meet the anonymity principle; if the distance is too far, the generated false trajectory and the real trajectory feature are too different, and the requested service cannot be obtained. Therefore, both privacy and service requirements need to be taken into account.

$K$ is defined as the degree of anonymity, that is, the number of false trajectories; $\theta$ is the threshold value of the included angle between two adjacent sections in the false trajectory. In order to prevent the motion trajectory from being too small, the included angle between two adjacent sections should be greater than this threshold value; $d$ is the maximum distance threshold between the false trajectory point and the corresponding user's real trajectory point. If the threshold is too large, the false position is too far away from the real position, so it is difficult to get the required service. If it is too close, it is difficult to ensure privacy. Therefore, it can be set according to the customer's requirements. $(x_i, y_i)$ and $\left(x_i^{'}, y_i^{'}\right)$ respectively represent the coordinates of the $i$-th sampling position point in the true and false trajectories. The specific algorithm is described as Algorithm 1:

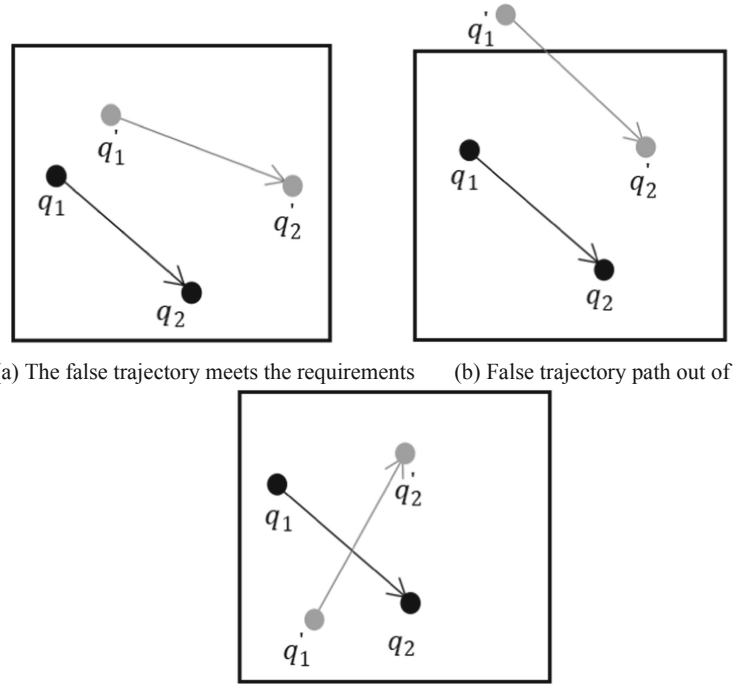
**Algorithm 1: False Trajectory Generation Algorithm without Road Net (FTA-r1)**

**Input**: $K$, $\theta$, $d$
**Output**: $K$ false trajectories
begin
1. obtain the starting coordinate $(x_i, y_i)$ of the true trajectory, $i = 1$;
2. $k = 1$ , in the $d$ range of starting coordinate, the initial section $[(x_1', y_1')(x_2', y_2')]$ of false trajectory is randomly generated;
\\find the initial position of the $k$-th $(1 \le k \le K)$ false trajectory
3. obtain the true trajectory node coordinate $(x_{i+1}, y_{i+1})$ at the next moment;
4. randomly generate false trajectory point $(x_{i+1}', y_{i+1}')$;
5. obtain angle $\theta_k$ , which is the included angle between line segments $[(x_i', y_i')(x_{i+1}', y_{i+1}')]$ and $[(x_{i-1}', y_{i-1}')(x_i', y_i')]$;
6. if the included angle does not satisfy the condition $(\theta \le \theta_k \le 180°)$, return to 4;
7. $(x_{i+1}', y_{i+1}')$ is added to the false trajectory;
8. if the trajectory point is not the endpoint, return to 3;
9. this false trajectory L is added to the false trajectory set ATS, if $k < K$, return to 2;
end

FTA-r1 only takes into account the angle problem of false trajectory when generating position points. The random generation of false coordinates makes the feature difference between the false trajectory and the real trajectory relatively large. Therefore, the FTA-r2 algorithm specifies the following requirements to generate false trajectory points: the included angle between the direction angle of false trajectory segment and that of true trajectory segment must be less than a certain angle; the real trajectory section and the false trajectory section should meet the requirement of distance deviation.

The generation of false trajectory points is shown in Fig. 1, where the black point is the real trajectory point and the gray point is the false trajectory point. It can be seen from Fig. 1 that the false track segment generated in Fig. 1(a) meets the two requirements of the FTA-r2 algorithm and is qualified. The false track segment generated in Fig. 1(b) exceeds the set range and does not meet the requirements of distance offset. The included angle between the direction angle of the false trajectory segment generated in Fig. 1(c) and that of the true trajectory segment exceeds the set angle threshold, which does not meet the angle requirements.

Two parameters $\varphi_i$ and $\varphi_i'$ are added in algorithm FTA-r2, which respectively represent the direction angle of the $i$-th line segment of the real trajectory and the false trajectory. The line segment in the false trajectory has to satisfy $\left|\varphi_i - \varphi_i'\right| \le \theta$. FTA-r2 false trajectory generation algorithm is to replace the 5th and 6th lines of code in Algorithm 1 (FTA-r1) with:




(a) The false trajectory meets the requirements (b) False trajectory path out of range

(c) False trajectory direction angle deviation is too large

**Fig. 1.** Generation of false trajectory points

**Algorithm 2: False Trajectory Generation Algorithm without Road Net (FTA-r2)**

**Input**: $K$, $\theta$, $d$

**Output**: $K$ false trajectories

begin

1. obtain the starting coordinate $(x_i, y_i)$ of the true trajectory, $i = 1$;
2. $k = 1$, in the $d$ range of starting coordinate, the initial section $[(x_1', y_1')(x_2', y_2')]$of false trajectory is randomly generated;

\\find the initial position of the $k$-th $(1 \le k \le K)$ false trajectory

3. obtain the true trajectory node coordinate $(x_{i+1}, y_{i+1})$ at the next moment;
4. randomly generate false trajectory point $(x_{i+1}', y_{i+1}')$;
5. obtain the direction angles of true and false trajectories, $\varphi_i$ is the direction angle of line segment $[(x_i, y_i)(x_{i+1}, y_{i+1})]$, and $\varphi_i'$ is the direction angle of line segment $[(x_i', y_i')(x_{i+1}', y_{i+1}')]$;
6. if the included angle does not satisfy the condition $(|\varphi_i - \varphi_i'| \le \theta)$, return to 4;
7. $(x_{i+1}', y_{i+1}')$ is added to the false trajectory;
8. if the trajectory point is not the endpoint, return to 3;
9. this false trajectory L is added to the false trajectory set ATS, if $k < K$, return to 2;

end

### 3.2 False Trajectory Algorithm Based on Road Network (RNFTA)

For vehicle users, the road network environment is an important background for maintaining false trajectories. Trajectories are normal only when they are driven on the road, otherwise they can be easily eliminated, thus the probability of being targeted is increased. Digital maps contain a common type of structured road data, and the map matching function of planning the route can be processed to locate the location points on the real road. RNFTA algorithm introduces the idea of road network matching and route planning, and realizes map matching through the trajectory correction service and route planning service of Baidu server. The two endpoints of the false trajectory are randomly generated according to the two endpoints of the true trajectory. The $x$ value is introduced to represent the number of trajectory feature points of the true trajectory, and the trajectory feature points are used to represent the intermediate points of the trajectory. The starting point, the feature points and the end point are used for route planning and route matching, so as to obtain the false trajectory. The specific RNFTA algorithm is shown in Algorithm 3:

**Algorithm 3: False Trajectory Generation Algorithm Based on Road Network (RNFTA)**

**Input**: $K$, $x$, $d$
**Output**: $K$ false trajectories
begin
1. obtain the starting coordinate $(x_1, y_1)$ and the ending coordinate $(x_{-1}, y_{-1})$ of the true trajectory;
2. divide the true trajectory into $x$ trajectory segments;
3. calculate the feature points of each trajectory segment;
\\find the initial position of the $k$-th false trajectory
4. obtain the true trajectory node coordinate $(x_i, y_i)$ at the next moment, and determine whether to generate false positions according to the privacy level of the section where each feature point is located;
5. if it is at a lower level, no change is required and return to 4; if it is at a higher level, run step 6;
6. generate false trajectory points $(x_i', y_i')(x_{i+1}', y_{i+1}')$ randomly;
7. plan the route for $(x_i', y_i')(x_{i+1}', y_{i+1}')$;
8. the obtained planned route fragment is added to the false trajectory;
9. if the trajectory point is not the endpoint, return to 4;
10. trajectory correction is performed by the true trajectory timestamp set and false trajectory.
11. get the false trajectory after repair;
12. if (the number of false trajectory points after repair is closer to the number of true trajectory points)
{false trajectory = false trajectory after repair;}
13. correct the number of false trajectory points;
14. the false trajectory is added to false trajectory set;
15. this false trajectory L is added to the false trajectory set ATS, if $k < K$, return to 1;
end

## 4 Simulation Analysis

In this section, the solution is simulated and analyzed. The simulation platform is implemented by using PYTHON language. The computer platform used in the experiment is a Windows 10 64-bit computer with Intel core i5-6300HQ and 8 GB memory. The true trajectory data used in the experimental analysis are ten trajectories randomly labeled in Nanjing. The sum of the total length of the trajectory is about 53 km. The number of sampling points for each trajectory is between 10 and 35, and the sampling distance is about 200 m.

The platform first needs to import the real trajectory. Users can import the real trajectory into the simulation platform in the form of files or load the real trajectory from the database. After importing the trajectory, the simulation system will display the effect of the current trajectory, as shown in Fig. 2.

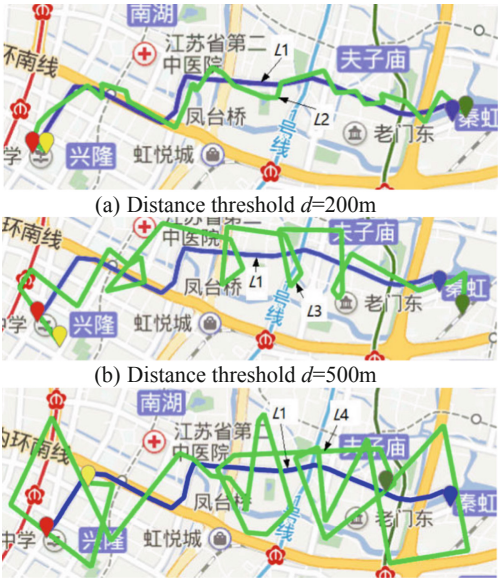

**Fig. 2.** Simulation system

### 4.1 Feasibility Analysis of False Trajectory Generation Algorithm Without Road Net

Firstly, the FTA algorithm and the influence of different generalization scope is analyzed. As shown in Fig. 3, three false trajectories with $K = 3$ are generated. $L_1$ is the real trajectory, $L_2$, $L_3$ and $L_4$ are false trajectories, and the distance threshold of the trajectory is set to 200 m, 500 m and 1000 m respectively.

As shown in Fig. 3, the larger the distance threshold, the greater the false trajectory deviation, which can well protect the real trajectory. However, the false trajectory has some problems that the route is too false and the service accuracy is insufficient. For example, walking to a few meters away in a second, driving too far off the road, and insufficient feedback when querying the surrounding gas stations.

The comparison of the two FTA algorithms is shown in Fig. 4. It can be seen that the false trajectory generated by FTA-r2 is smoother and closer to the reality, which is suitable for walking users. However, it is not applicable to the vehicle trajectory. The false trajectory deviates seriously from the road, and the attacker can easily locate the real trajectory.




(a) Distance threshold $d$=200m

(b) Distance threshold $d$=500m

(c) Distance threshold $d$=1000m

**Fig. 3.** One simulation example of FTA algorithm


(a) FTA-r1 Algorithm (b) FTA-r2 Algorithm

**Fig. 4.** Comparison of two FTA algorithms

### 4.2 Feasibility Analysis of False Trajectory Generation Algorithm Based on Road Network

One example of RNFTA algorithm is shown in Fig. 5, where $L_1$ is the real trajectory, $L_2$ and $L_3$ are false trajectories generated by setting different privacy levels. In Fig. 5(a), R1 and R2 are set to a low level of privacy; in Fig. 5(b), R3 and R4 are set to a low level of privacy. These sections are main roads with dense vehicles and do not require trajectory anonymity. In both cases, the two false trajectories are both on the road, so the trajectories produced do not deviate from the road, and both are feasible. Another


(a) Case 1 (b) Case 2

**Fig. 5.** Example 1 of RNFTA algorithm


(c) Case 3 (d) Case 4

**Fig. 6.** Example 2 of RNFTA algorithm

example of the RNFTA algorithm is shown in Fig. 6. By comparing Fig. 5 and Fig. 6, it can be seen that RNFTA algorithm can adapt to different simulation environments.

### 4.3 Performance Analysis

Next, the algorithm performance will be analyzed from the perspective of trajectory coverage and crowding degree. These two indicators can well measure the ability of the constructed false trajectory to obtain accurate service and resist aggression. Trajectory coverage means that for a point in the real trajectory, the circular region of the distance threshold always contains false trajectory points, indicating that the real trajectory can be covered by the false trajectory. Trajectory coverage refers to the proportion of the covered true trajectory points $T^{'}$ in the total number of true trajectory points $T$, that is,

$$\text{Trajectory coverage} = \frac{T^{'}}{T} \times 100\% \tag{1}$$

A higher coverage indicates that accurate services can still be obtained under the condition of anonymity.

The distance threshold value $d = 50$, 100, 150 and 200 are taken respectively to compare their impacts, and the trajectory coverage of the algorithm is calculated as shown in Fig. 7. According to the analysis, the false trajectories generated by the RNFTA algorithm for road network matching can better cover the real trajectories and obtain accurate services. Average service accuracy has increased by 36%.

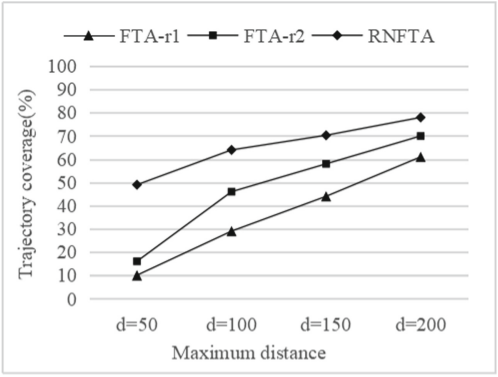



**Fig. 7.** Trajectory coverage analysis chart

The true trajectory points crowding degree $F$ refers to the number of false trajectory points contained in the circular region of the distance threshold value for a point in the real trajectory. The higher the true trajectory points crowding degree $F$ is, the lower the probability of the attacker locating the real location is, indicating that the anti-attack ability of the algorithm is better. The crowding degree is defined as the ratio of the sum of true trajectory points crowding degree $F$ to the total number of true trajectory points $T$, which is

$$\text{crowding degree} = \frac{\sum_{i=1}^{T} F_i}{T} \tag{2}$$

The distance threshold value $d = 50$, 100, 150 and 200 are taken respectively to compare their influences, and the trajectory crowding degree is calculated, as shown in Fig. 8. According to the analysis, the false trajectory and true trajectory set generated by the RNFTA algorithm for road network matching have higher crowding degree of trajectory points, which reduces the degree of privacy leakage by 78%. In conclusion, the scheme in this paper can be applied to users of different travel modes. Meanwhile, RNFTA based on road network has better ability to obtain accurate services and resist attacks.

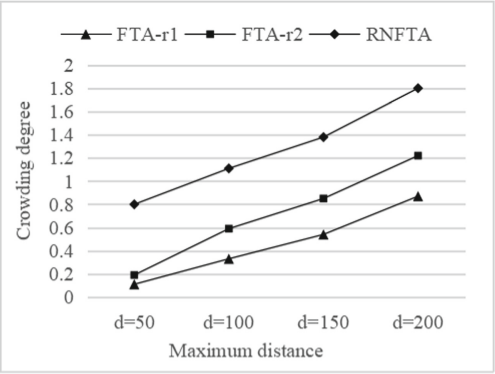



**Fig. 8.** Crowding degree analysis chart

## 5 Conclusion

In view of the problem that privacy protection cannot be mixed when people travel in different ways, this paper proposes a mobile trajectory privacy protection scheme facing different travel modes. Firstly, the overall privacy protection scheme is designed according to the classification of route section privacy level and different modes of user travel; then, a false trajectory generation algorithm for non-road environment (FTA) is designed, taking into account the requirements of service and privacy; finally, the false trajectory generation algorithm under road network environment (RNFTA) is designed, thus the real trajectory can be hidden and protected. The simulation experiment shows that the scheme can handle the anonymous trajectory quickly for the users of different travel modes, and can provide the users with more efficient and safe location services.

**Funding.** This work is funded by the National Natural Science Foundation of China (No.61902069 and U1905211), the Science Foundation of Fujian University of Technology (GY-Z21048, GY-Z18181, GY-Z21024), the Natural Science Foundation of Fujian Province of China (2021J011068).

## References

1. Feng, D., Zhang, M., Ye, Y.: Research on location trajectory publishing technology based on differential privacy model. J. Electron. Inf. Technol. **42**(1), 74–88 (2020)
2. Wu, Z., Wang, R., Li, Q.: A location privacy-preserving system based on query range cover-up or location-based services. IEEE Trans. Veh. Technol. **69**(5), 5244–5254 (2020)
3. Xu, C., Luo, L., Ding, Y.: Personalized location privacy protection for location-based services in vehicular networks. IEEE Wirel. Commun. **9**(10), 1633–1637 (2020)
4. Zeng, H., Zuo, K., Wang, Y.: Semantic diversity location privacy protection in road network environment. Comput. Eng. Appl. **56**(7), 102–108 (2020)
5. Gruteser, M., Grunwald, D.: Anonymous usage of location-based services through spatial and temporal cloaking. In: Conference: Proceedings of the First International Conference on Mobile Systems, Applications and Services, San Francisco, CA, USA, pp. 31–42 (2003)

6. Kato, R., Iwata, M., Hara, T., Arase, Y., Xie, X., Nishio, S.: User location anonymization method for wide distribution of dummies. In: Decker, H., Lhotská, L., Link, S., Basl, J., Tjoa, A.M. (eds.) DEXA 2013. LNCS, vol. 8056, pp. 259–273. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40173-2_22
7. Zhang, S., Liu, Q., Wang, G.: Trajectory privacy protection method based on location obfuscation. J. Commun. **39**(7), 81–91 (2018)
8. Li, Z., Li, W., Wen, Q.: A efficient blind filter: location privacy protection and the access control in FinTech. Futur. Gener. Comput. Syst. **100**, 797–810 (2019)
9. To, Q.C., Dang, T.K., Kueng, J.: A Hilbert-based framework for preserving privacy in location-based services. Int. J. Intell. Inf. Database Syst. **7**(2), 113 (2013)
10. Ye, A., Meng, L., Zhao, Z., Diao, Y., Zhang, J.: Trajectory differential privacy protection mechanism based on prediction and sliding window. J. Commun. **41**(4), 123–133 (2020)
11. Xu, Z., Zhang, J., Tsai, P., Lin, L., Zhuo, C.: Spatiotemporal mobility based trajectory privacy-preserving algorithm in location-based services. Sensors **21**(6), 113–134 (2021)
12. Squicciarini, A.C., Qiu, C.: Location privacy protection in vehicle-based spatial crowdsourcing via geo-indistinguishability. In: 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). IEEE (2019)
13. Arain, Q.A., Memon, I., Deng, Z., Memon, M.H., Mangi, F.A., Zubedi, A.: Location monitoring approach: multiple mix-zones with location privacy protection based on traffic flow over road networks. Multimedia Tools Appl. **77**(5), 5563–5607 (2017). https://doi.org/10.1007/s11042-017-4469-4
14. Ji, Y., Gui, X., Dai, H., Peng, Z.: A two-stage user interest zone construction method to support trajectory privacy protection. Chin. J. Comput. **40**(12), 2734–2747 (2017)
15. Kuang, L., Wang, Y., Zheng, X., Huang, L., Sheng, Y.: Using location semantics to realize personalized road network location privacy protection. EURASIP J. Wirel. Commun. Netw. **2020**(1), 1–16 (2019). https://doi.org/10.1186/s13638-019-1618-7