Advances in Computer Science Research, volume 87

3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019)

Decentralized Location Privacy Protection Method of Offset Grid

Jie Ling [a], Junyi Xu [b, *]

School of Guangdong University of Technology, Guangzhou, 510000, China.

[a]jling@gdut.edu.cn, [b, *] xujunyitk@163.com

Abstract. Location services bring convenience to people's lives, but also easily lead to the disclosure of personal privacy information. Many location privacy protection methods are based on trusted anonymous servers, but in real life, the anonymous servers are not trusted. This paper proposes a decentralized location privacy protection method of offset grid for untrusted anonymous servers. The method divides the location area into grids, and the user calculates the offset grid area according to the periodically updated grid information and sends the offset grid area to the anonymous server. The anonymous server selects K-1 distributed grid coordinates according to the history of other user's query, the anonymous grid points set satisfies the multi-diversity. The method in this paper has lower time overhead while protecting user privacy information, and the location distribution of anonymous location sets is more decentralized. Experiments were carried out on the simulated data set. The experimental results verify the effectiveness of the method.

Keywords: Location Privacy Protection; Offset Grid; K-Anonymity; Decentralized Position; Untrusted Anonymous Server.

1. Introduction

With the rapid development of mobile communication technology, users can accurately locate their location through mobile devices. Location-Based Service (LBS) integrates location and other information of mobile communication devices to provide value-added services for users [1], such as access to the nearest gas station information or access to coffee shop information within 1 km, etc.

When requesting service, user sends the location to the Location Service Provider (LSP) to obtain high quality service. In the process of obtaining location service [2], the user leaks the location information to LSP, or attacker intercepts the user's request to obtain the corresponding information in the process. Attackers can obtain other key information through location information [3]. For example, attacker collects the location information of each user for analysis, and knows that the user's location trajectory is always active in a certain area so as to know the area of user's activity; Or attacker can know the user's interest by analyzing the content of the request (such as the golf course or hospital in which the user inquires) to gain interesting or sensitive physical problems. When using LBS, personal privacy issues should not be ignored, especially personal sensitive information.

Gruteser first introduced k-anonymity technology in relational database into LBS location privacy protection technology: when a mobile user's location is extended to be difficult to distinguish from other k-1 users, the user's identity is also difficult to distinguish from these users, so the location satisfies K-Anonymity [4]. At present, there are two kinds of scheme of privacy protection methods for LBS: distributed Point-To-Point scheme and third-party Anonymous Server scheme.

Distributed Point-To-Point scheme is a collaboration among users, which implements location concealment on mobile devices and sends hidden request to LSP for service. The third-party Anonymous Server scheme is to add an Anonymous Server (AS) between users and LSP. Users only need to send request to AS. AS is responsible for anonymous processing of users' requests. After obtaining the result from LSP, the content that meets the user's needs is filtered and sent to the user.

In the distributed Point-To-Point scheme, user can search neighboring users by multi-hop communication [5], collecting the location information of neighboring users to construct anonymous regions. After searching other users by multi-hop communication, use also can continues the process of anonymity with the continuous transmission of request [6], so as to finally achieve the purpose of constructing anonymity, which can achieve anonymity as well as allocate consumed time to other adjacent users;

Without anonymous area, users and other users collaborate to construct anchors position [7] in order to protect location privacy; Users can encrypt request to protect location privacy [8]; Document [9] proposes a method to realize K-Anonymity by dividing user requests into blocks and exchanging blocks with neighboring users. User collaboration does not need a third-party anonymous server, but conceals its own privacy information through the collaboration of adjacent users, which can avoid the malicious behavior of untrusted third parties.

In the structure of Anonymous Server, document [10] constructs multiple anonymous regions in AS, each of which satisfies K-Anonymity, L-Diversity, and each anonymous region can be far apart; User use Hilbert curve [11] to hide location information. AS aggregates users with the same Hilbert value to form K-Anonymity, which makes it difficult for AS to know the real location of users; The search region can be divided into several grids [12] and AS selected some grids to form K-Anonymity. By using asymmetric and symmetric encryption, only users can know the exact location information and request information. We can divide the region into grids, achieves K-Anonymity at AS, and then uses symmetric encryption (such as AES) [13] and IBE to encrypt sensitive query information and results; To uses encryption method, we can use Order-Preserving Encryption method [14] to hide the location of users, and both AS and LSP are difficult to know the exact location information and query information of users. The central server can become the performance bottleneck and the attack point of the system [15], if the AS is broken by attacker or not trusted, it will lead to the leakage of user information. Many scholars have studied the non-central server and the untrusted central server more and more deeply [16].

Some methods for untrusted anonymous servers mostly use cryptographic tools to ensure the security of user requests, which makes it difficult for LSP and AS to obtain accurate user information. However, there is a large amount of computation in encryption operation, which can easily lead to the decline of user service quality, such as the extension of response time. Secondly, the K-Anonymity of user's request in AS is to search k-1 locations near the request location. K locations satisfy L-Diversity. In some cases, if there is a sparse population, it needs a large range of searches to determine the anonymous area or directly fail anonymity. Therefore, this paper proposes a Decentralized Location Privacy Protection Method of Offset Grid of untrusted anonymous servers (DPOG). The method makes it difficult for LSP and AS to obtain the exact information of users and has low time overhead. AS, according to historical query data, select k-1 grid coordinates with different query contents to form an anonymous set, and then send it to LSP for query.

2. Preliminaries

2.1 System Architecture

Fig.1 show the system architecture of DPOG presented in this paper. It contains four entities: User, untrusted Anonymous Server (AS), Function Server (FS) and Location Service Provider (LSP).

User is a mobile device with location function. It can get location where user is located. User needs to construct a rectangular area based on real location and send the request to AS after calculating the offset grid area. User obtains periodically updated basic grid G from FS. G can be denoted as $\{(x_0, y_0), r, TIME\}$, (x_0, y_0) represents the starting real coordinates of the basic grid, r represents the width of a single grid, and TIME is the identifier of G.

The function of FS is to update the information of basic grid G periodically. For the request of obtaining G from User and LSP, FS get identity information and respond basic grid information G.

AS saves user's request and searches qualified grid coordinates in historical query data to construct K-Anonymous set. It is also responsible for filtering the data returned by LSP. In this method, AS is not credible.

LSP is a large location server provider, which records Point of Interest (POI) information, responds to the request, and sends the search result information to AS.

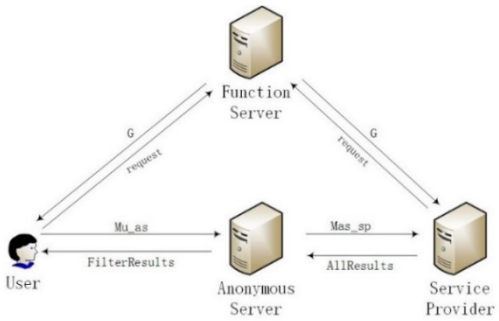

Figure 1. System Architecture Based on Untrusted Anonymous Server

2.2 Threat Model

This paper assumes that AS and LSP are not credible, and will try to mine the user's personal privacy information from the requested information. DPOG aims to make it difficult for AS to know the exact location of the user, and LSP to know the location of the user and the user's identity information. This paper also assumes that AS and LSP are independent service providers, and they do not collude. If AS and LSP collude, then users' privacy information can't be guaranteed.

Since the basic grid *G* of FS can be acquired by users who need to request location services, if the requesting user is malicious, malicious users will attempt to attack AS and LSP to obtain the user's request information, or intercept the user's request information in the communication link. The security of AS and LSP is beyond the scope of this paper. The external security of AS and LSP is guaranteed by the service provider. The request message sent by the user to AS cannot contain the exact information of the user. The user needs to generate a rectangular area containing the location in order to prevent the attacker from intercepting the information in the communication link and causing the location leakage.

2.3 DPOG Method

DPOG method includes five processes: user grid processing, AS anonymous processing, LSP search results, AS filter results, user filter results. Each process is described in detail below.

2.4 User Grid Processing

User sets parameters for requests, such as {*MinArea,Type,ID,K,Radius*}. *K* indicates that when AS performs anonymous operation, a total of *K* grid coordinates with different query contents are needed. *Type* indicates the type of query. *MinArea* specifies that the area of rectangular area that user needs to generate should be no less than this value. *Radius* indicates the size of the search area centered on the query locations.

As shown in Fig.2, it is the result of user gridding. User sends a message to FS to get the basic grid information *G*. User obtains its accurate location (x_u,y_u) through mobile devices and randomly generates a number *b* with a value range of 0~1. (x_u,y_u) is offset to random direction to get location *loc'*. The length is $\sqrt{MinArea} * b$. According to *loc'*, create rectangular region $H=\{(x_1,y_1),(x_2,y_2)\}$, (x_1,y_1) represents the coordinates of the lower left corner of the rectangular region, (x_2,y_2) represents the coordinates of the upper right corner of the rectangular region.

According to the basic grid *G* and position(x_i,y_i), the offset grid coordinates(c_i,r_i) can be calculate

$$(c_i, r_i) = (\frac{x_i - x_0}{\mathrm{r}}, \frac{y_i - y_0}{\mathrm{r}}) \quad (1)$$

r is the width of a single grid in *G*. The offset grid coordinates is $H'=\{(c_1,r_1),(c_2,r_2)\}$. Construct message M_{u_as}, M_{u_as}= *{ID, H', Type, K, TIME, Radius}*, and then sent M_{u_as} to AS for anonymous processing.

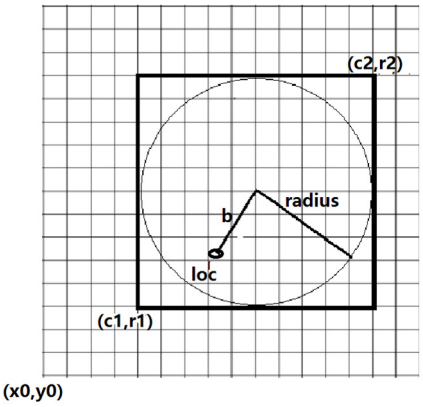

Figure 2. User Grid Processing

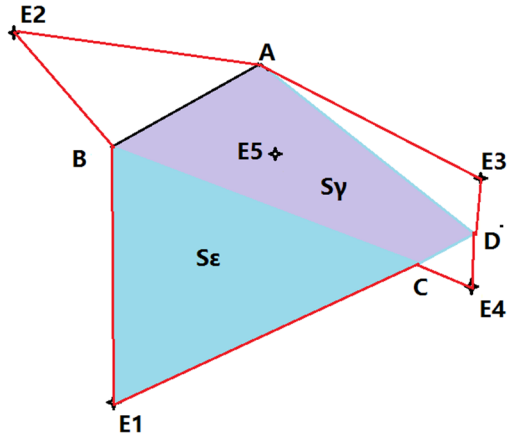

Figure 3. Select Most Decentralized Point

2.5 AS Anonymous Processing

AS extracts the offset grid coordinates H', privacy protection K and Type from the received M_{u_as} message, calculates the central grid coordinate (c_u,r_u) of the request according to H'. Adds$\{(c_u,r_u),Type\}$to γ, which denotes the anonymous set of grid points.

Search for K-1 request $type_i$,0<i<k, which is different from Type of user in historical request. Each $type_i$, include many grid coordinates (c, r) in historical request, and we select m grid coordinates points randomly. $type_i = \{(c_1,r_1),(c_2,r_2)\ldots(c_m,r_m)\}$. Calculate the most decentralized grid coordinates point in each $type_i$ distance from γ. The way to find the most decentralized grid point by the following methods.

Any N-sided polygon can be divided into n-2 triangles. The area of the polygon can be obtained by calculating the sum of the area of each triangle. The area of the triangle is obtained by

$$s = \sqrt{p(p-a)(p-b)(p-c)} \tag{2}$$

p is the half circumference of the triangle, and a, b and c are the three sides of the triangle. For each grid point in $type_i$ (0 <i < k), the grid point c with the largest polygon area is selected and added into γ.

$$c = arg \max_{\varepsilon \in type_i} (S_\gamma + S_\varepsilon) \tag{3}$$

For example, as shown in Fig 3, γ already contains four location points ABCD, select five grid points {E1, E2…E5}in the $type_i$ that are different from the four grid points types. Because the polygonal region$(S_\gamma+S_\varepsilon)$ formed by the grid point of E1 and γ is larger, E1 is selected to be added to γ.

After calculation, the most decentralized grid coordinates point(c_i,r_i), 0<i<k,with the $type_i$ of each grid coordinate point added to γcan be obtained. The set contains K different request contents and grid coordinates point, which is $\gamma=\{((c_u,r_u),type_u),\ldots,((c_i,r_i),type_i)\}$,0<i<k. Construct message M_{as_LSP}, $M_{as\_LSP}=\{\gamma,TIME\}$. Send M_{as_LSP} to LSP for query.

2.6 LSP Search Results

After LSP receives the message M_{as_LSP}, it sends a request to FS according to TIME to obtain the basic grid information G, and restores the grid coordinates(c_i,r_i) of each $type_i$ in γ to the real coordinates

$$(x_i, y_i) = (c_i * r + x_0, r_i * r + y_0) \tag{4}$$

According to each coordinate(x_i,y_i) and $type_i$ as well as query radius, s POIs results are retrieved, $POI_i=\{(x_j,y_j), info_j\}$,0<=j<s. Each POI includes real position coordinates and relevant information. Since the real coordinates of POI cannot be exposed to AS, the true coordinates (x, y) of each POI are calculated according to the basic grid G and (1). Therefore, the offset grid coordinates of POI_i are $\beta_i=\{(c_j,r_j), info_j\}$,0<=j<s. The offset grid coordinates of each type of POI and type are sent to AS for filtering. $M_{LSP\_as}= \{\beta_i, type_i\}$,0<=i<k.

2.7 AS Filter Results

After receiving the message M_{LSP_as}, AS finds the $type_u$ requested by the requesting user in the history request list and sends the type of query result grid setβ_u to user, $M_{as\_u}=\{\beta_u\}$.

2.8 User Filter Results

User gets POI information from M_{as_u}. According to the basic grid G, the offset grid coordinates$(c_i,r_i)$are converted to the real coordinates (x_i, y_i) according to (4). Since the position submitted to AS is the offset grid coordinates, some POI does not meet the user's requirements. Therefore, it is necessary to calculate the distance from the user according to the user's location (x_u, y_u), discard POI which is out of radius and then conforming POI is presented to users in combination with relevant info.

2.9 Performance Evaluations

2.9.1 Safety Analysis

This paper assumes that LSP and AS will correctly respond to user requests and have good external security, but will try to mine user privacy information from data.

As far as AS is concerned, it can obtain user ID and grid offset information $H'=\{(c_1,r_1),(c_2,r_2)\}$. According to the grid offset information, AS cannot directly know user's location information, only the user's id and the relative grid coordinates of the basic grid G. If the AS maliciously acquires the basic grid G, because the user offsets the location and sends it to the AS in a rectangular area, the user can be in any grid point, so the probability that the AS can identify the user's location is

$$Pro = \frac{r^2}{MinArea} \tag{5}$$

When the basic grid is divided more finely, that is, the smaller the r, the larger the anonymous area generated by the user is than MinArea, the attacker needs more cost to identify the exact location of users.

For LSP, as AS transforms user id, it is difficult to know user's identity information. Because the request information has been anonymously processed by AS, LSP can determine the probability of user's request content is 1/K. After user's anonymous region generation and AS's anonymous processing, the probability of LSP identifying user's location is

$$Pro = \frac{r^2}{k * MinArea} \tag{6}$$

When K increases, the probability of identifying user's location decreases gradually, protecting user's privacy information.

2.9.2 Experimental Environment

The experiment of this paper uses Thomas Brinkhoff's Network-based Generator of Moving Object [17,18] to generate random moving objects from Oldenburg. The DPOG method in this paper,

are compared in terms of time consumed, network communication cost and anonymous location decentralized with OPEG [14] and ELPP [11].

DPOG method is implemented in Python language and runs on Windows 10 operating system. The CPU is Intel Core i7-6700. Memory is 4G and default parameter values are shown in Table 1.

City Oldenburg contains 6105 nodes, 7035 edges, 23572 long and 26915 wide. In this region, 5057 POI locations were generated randomly. Based on this region, 1000 random locations are generated, and each location simulates requesting location services.

Table 1. Default Parameters of Experiment

Parameter	Defaults
Minimum anonymous region	2
K	3
Number of Point of Interest	5057
Width of grid in basic grid	0.2
Search radius	1

3. Experimental Results and Analysis

(1) Time Consumed

The time consumed for DPOG, OPEG and ELPP methods were compared. The time here refers to the time consumed by the user to send a query request to the user to obtain the query results. As shown in Fig.4, with the increase of K value, the time consumed by DPOG increases gradually, while ELPP method does not increase with the increase of K value. The reason is that when ELPP method operates anonymously on AS, it only needs to select other locations with the same Hilbert value in historical query data. DPOG method needs to calculate the dispersion between locations determined by polygon area. The increase of K value needs to calculate more triangle area. The OPEG method uses Order-Preserving Encryption instead of K anonymity, so its time consumption is not related to k, it needs more time to encrypt and decrypt information.

(2) Anonymous Location Dispersion

DPOG takes into account the degree of dispersion of anonymous location points when constructing K-Anonymity area. The higher the degree of dispersion, the larger the area of location set. The selected location can be scattered in different places, not clustered in a range, which can improve the degree of privacy protection of users. Because OPEG does not have K-Anonymity process, so just compared with ELPP in anonymous area. The ratio $\beta = \frac{S_{DPOG}}{S_{ELPP}}$, S_{DPOG} and S_{ELPP} are anonymous area formed by DPOG and ELPP methods respectively. The larger the β, the more dispersed the location of DPOG is than that of ELPP.

As shown in Fig.5, as the value of K increases gradually, the value of βincreases gradually, which indicates that the anonymous set generated by DPOG has a larger area than ELPP, and the distribution of anonymous location of DPOG is more dispersed than ELPP, rather than clustered in an area. When the value of K is 10, the anonymous area of DPOG is 11% higher than ELPP, which indicates that DPOG can produce more anonymous set than ELPP.

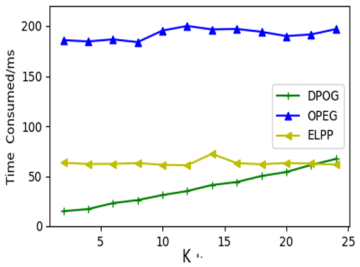

Figure 4. Comparisons of Time Consumed

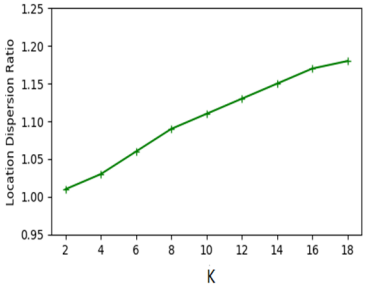

Figure 5. Location Dispersion Ratio

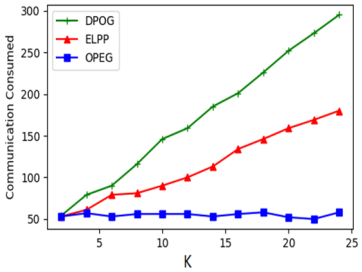

Figure 6. Comparison of Network Communication Cost

(3) Network Communication Cost

Network communication Cost refers to the amount of data transmitted over the network. Because we just simulate the location request process and the query process, and Communication Cost mainly refers to the number of POI returned by LSP. So, we just compare the number of POI returned to AS from LSP for different methods.

As shown in Fig.6, because the K anonymous area generated by DPOG is dispersed, the distance between them is farther, and POI searches are needed within the search radius for each location, so the results of searches will be more. While the ELPP has more concentrated anonymous location points, so there are more duplicate query results. The OPEG does not use K-Anonymity, but directly encrypts the query result to protect private information, and its network communication cost is small.

(4) Influence of Parameter MinArea on DPOG

As shown in Fig.7, as the parameter MinArea increases, the time consumed of DPOG has no effect. This is because MinArea affects the size of the offset grid area of the user and the offset of the user's location. For AS and LSP, no additional data processing is required. When the value of MinArea increases, the position of the user is more offset. As shown in Fig.8, the distance between the user's query result and the position of the user increases as the value of MinArea increases. The result of the query obtained by the user will gradually deviate from the location of the user. With the increase of MinArea, users' privacy protection will be better, and the probability of location being identified is smaller.

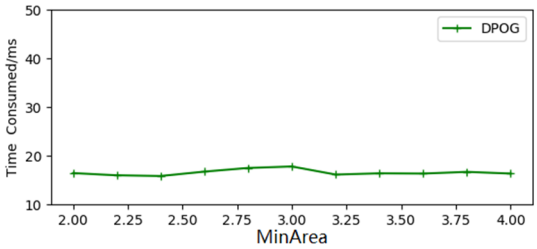

Figure 7. Effect of Minaaea on Time Consumed

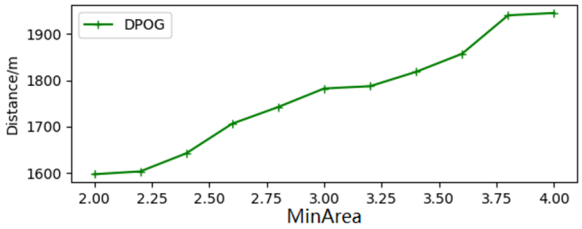

Figure 8. Effect of Minarea on Query Results

4. Conclusion

For the application environment of untrusted anonymous server, this paper proposes a Decentralized Location Privacy Protection Method of Offset Grid. In [11] and [14], in order to prevent the leakage of the user's private information to the anonymous server and the location server provider, a large time consumed is required, and the generated anonymous location set is centralized and easy to anonymous request single. The method proposed in this paper calculates the offset grid area based on the periodically updated grid information. The anonymous server selects the scattered grid coordinates point of the requested content according to the user query request. We make it difficult for AS to know the exact location of the user, and LSP to know the location of the user and the user's identity information. At the time, it is difficult to know the user's private information, and there is a lower time overhead and a more decentralized anonymous locations set.

Acknowledgments

This work is supported by the science and technology project of Guangdong Province (No.2017B090906003) and the project of Guangzhou Science and Technology (No.201802010043, 201807010058).

References

[1]. PAN Xiao, XIAO Zhen, MENG Xiao feng. Survey of location privacy-preserving[J]. Journal of Computer Science and Frontiers,2007,1(3):268- 281.

[2]. Wang L, Meng XF. Location privacy preservation in big data era: A survey. Ruan Jian Xue Bao/Journal of Software,2014,25(4):693-712 (in Chinese). http:// www.jos.org.cn/1000-9825/ 4551.htm.

[3]. FENG Deng-Guo,Zhang Min, Li Hao. Big Data Security and Privacy Protection[J].Journal of Computers,2014,37(01):246-258(in Chinese).

[4]. Gruteser M, Grunwald D. Anonymous usage of location-based services through LSPatial and temporal cloaking. In: Proc. of the 1st Int'l Conf. on Mobile Systems, Applications and Services. New York: ACM Press, 2003. 31–42.

[5]. Yang N, Cao Y, Liu Q, et al. A novel personalized TTP-free location privacy preserving method[J]. International Journal of Security and Its Applications, 2014, 8(2): 387-398.

[6]. Niu B, Zhu X, Li Q, et al. A novel attack to LSPatial cloaking schemes in location-based services[J]. Future Generation Computer Systems, 2015, 49: 125-132.

[7]. HUANG Yi,Huo Zheng, MENG Xiao Feng. Co Privacy: A Collaborative Location Privacy-Preserving Method without Cloaking Region[J]. Chinese Journal of Computers, 2011,34(10): 1976-1985(in Chinese).

[8]. Ghaffari M, Ghadiri N, Manshaei M H, et al. P4QS: A Peer to Peer Privacy Preserving Query Service for Location-Based Mobile Applications[J]. IEEE Transactions on Vehicular Technology, 2017, 66(10): 9458-9469.

[9]. Ma C, Zhang L, Yang S, et al. Achieve personalized anonymity through query blocks exchanging[J]. China Communications, 2016, 13(11): 106-118.

[10]. Li T C, Zhu W T. Protecting user anonymity in location-based services with fragmented cloaking region[C]//Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on. IEEE, 2012, 3: 227-231.

[11]. Peng T, Liu Q, Wang G. Enhanced location privacy preserving scheme in location-based services[J]. IEEE Systems Journal, 2017, 11(1): 219-230.

[12]. YANG Songtao, WANG Huiqiang, MA Chunguang. Location privacy protection method based on random mesh. Systems Engineering and Electronics, 2018, 40(2): 422-426.

[13]. Schlegel R, Chow C Y, Huang Q, et al. User-defined privacy grid system for continuous location-based services[J]. IEEE Transactions on Mobile Computing, 2015 (1): 1-1.

[14]. SHEN Nan, JIA Chun-fu, LIANG Shuang, LI Rui-qi, LIU Zhe-li. Approach of location privacy protection based on order preserving encryption of the grid[J]. Journal on Communications, 2017,38(07):78-88.

[15]. HUO Zheng,MENG Xiao-Feng,HUANG Yi.PrivateCheckIn:Trajectory Privacy-Preserving for Check-In Services in MSNS[J].Chinese Journal of Computers,2013,36(04):716-726.

[16]. Zhang XJ,Gui XL,Wu ZD.Privacy preservationfor location-based services: A survey. Ruan Jian Xue Bao/Journal of Software,2015,26(9):2373-2395 (in Chinese). http:// www. jos. org. cn/ 1000-9825/4857.htm.

[17]. Brinkhoff T. A framework for generating network-based moving objects[J]. GeoInformatica, 2002, 6(2): 153-180.

[18]. Brinkhoff T. Generating network-based moving objects[C]//Scientific and Statistical Database Management, 2000. Proceedings. 12th International Conference on. IEEE, 2000: 253-255.