JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 38, 121-138 (2022)
DOI: 10.6688/JISE.202201_38(1).0007

K-Anonymous Based Anti-Positioning Security Strategy in Mobile Networks

LIANG ZHU, LI-PING YU, ZENG-YU CAI, XIAO-WEI LIU AND JIAN-WEI ZHANG
College of Computer and Communication Engineering
Zhengzhou University of Light Industry
Zhengzhou, 450001 P.R. China
E-mail: {lzhu; ing}@zzuli.edu.cn; yulipingl520@163.com; 8949934@qq.com; hellolxww@163.com

Location privacy-preserving for location-based service (LBS) in mobile communication networks has been received great attention. Traditional location privacy-preserving methods mostly focus on the researches while users publish the location information to LBS providers. However, it lacks the studies on location privacy-preserving when users acquire location information from location provider (LP). In this paper, we present a K-Anonymous based Anti-positioning Security Strategy (KA-ASS), in order to protect location privacy during the positioning service phase. Firstly, the anonymous knowledge base (AKB) is built to select the $k-1$ dummy location fingerprint for the original positioning request. Then, the Continuous Optimal Fingerprint Selection (COFS) algorithm is proposed to generate the dummy trajectory sequence, which has higher directional similarity and distance similarity with original trajectory sequence. Finally, experimental evaluation on the real-world datasets illustrate that our KA-ASS scheme can not only reduce the anonymous time cost, but also achieve better anonymous quality for location positioning service.

Keywords: location-based service (LBS), 5G networks, location privacy, positioning service, *k*-anonymous

1. INTRODUCTION

As a promising way of improving people's quality of life, location-based service (LBS) has emerged and been used frequently. With the rapid development of mobile communication systems and mobile terminals, more accurate positioning services in mobile networks are provided. Particularly in 5G networks, many new envisioned technologies, *e.g.* Ultra Dense Network (UDN), massive Multiple Input Multiple Output (MIMO) and millimeter Wave (mmWave) communication have great potential in achieving sub-meter accuracy positioning [1].

Fig. 1 shows the architecture of network positioning service, which includes location provider, mobile devices, LBS providers, and communication networks. The location provider provides the basis for determining the location of a mobile device, which can track its specific location through the built-in GPS chip or third-party network location provider and transmit location information to the application. Mobile devices refer to electronic devices that can connect to the network and transfer data. Mobile devices serve as the basis for collecting location data and sending LBS requests, usually including smart-phones, computers, smart-watches and internet connected vehicles devices. LBS provider refers to a third party that can provide LBS services for mobile devices. Usually, LBS providers own or can create location-based information content. Communication network is used to

Received December 8, 2020; revised March 6, 2021; accepted April 6, 2021.
Communicated by Xiaohong Jiang.

connect mobile devices with LBS service providers or network positioning providers to realize information transmission between them, including wireless communication network, satellite network and so on.

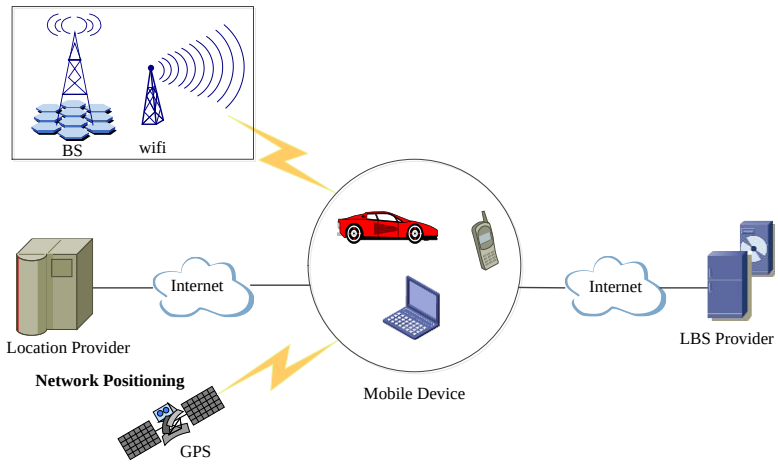


Fig. 1. The architecture of network positioning service.

However, mobile users enjoy the convenience of LBS at the risk of location privacy disclosure. The location information disclosure will directly or indirectly cause the attackers to know the current location of victims. What's more, the attackers can deduce the victims' workplace, life habits and so on. For example, an attacker can infer the victim's workplace according to the location information which frequently submitted during the working day. Sometimes, users may want to provide locations with varying precision. Thus, the accuracy of positioning is not always accurate for different demands of users. For example, a user wants to tell friends the exact location, while providing a rough location for the weather service. Moreover, location information is not just specific longitude and latitude information. For example, users do not want to publish their information in the hospital. In that case, the spatial information to be protected represents a location.

Location privacy-preserving in mobile networks has been received great attention. Traditional methods mainly focus on the researches while users publish the location information to LBS providers. It lacks the studies on location privacy-preserving when users acquire location information from location provider (LP). In this paper, we present the K-Anonymous based Anti-positioning Security Strategy (KA-ASS), in order to not only protect location privacy during the positioning service phase, but also ensure higher data utility. There are two key issues that need to be addressed. One is that how to select the dummy location fingerprint, which is indistinguishable from the real location fingerprint during the positioning service phase. Another one is that how to generate the dummy trajectory sequence considering on the location continuity of the original trajectory sequence.

The contributions of our work can be divided into three aspects as following.

(1) We construct the anonymous knowledge base (AKB) on the anonymous server, in order to select the $k - 1$ dummy location corresponding to the original location. Furthermore, the graph model of APs is utilized to ensure that the relationship between

APs is discovered with less space consumption.

(2) Considering that the attacker may launch an attack by utilizing the trajectory characteristics in the continuous location request, the Continuous Optimal Fingerprint Selection (COFS) algorithm is proposed to generate the dummy trajectory sequence according to directional similarity and distance similarity.

(3) Through the security analysis, we prove our KA-ASS scheme satisfies trajectory privacy-preserving demands. We conduct an extensive experimental study over the real dataset. The experimental results show that our KA-ASS scheme can privately provide positioning services with high efficiency.

2. RELATED WORKS

The research of location privacy protection focuses on two stages, which are publishing location information to servers and providing location service to users. There are mainly four kinds of location privacy protect strategies, *i.e. k*-anonymous, position dummies, mix zone and encryption.

The method of *k*-anonymous is firstly proposed by Samarati and Sweeney [2] in 1998, which is the most frequently used method in location privacy protection. It requires a certain number (at least *k*) of indiscriminable records for the published datasets, so that the attacker cannot distinguish the specific privacy information of target user. Gruteser *et al.* [3] firstly taken the thought of *k*-anonymous into location privacy protection. For *k*-anonymous, the generated anonymous space containing the real location and other $k-1$ location was sent to LBS server, in order to protect the real location of target user. Mokbel *et al.* [4] proposed a Casper architecture to satisfy the need of self-defined privacy protection of users. Bamba *et al.* [5] proposed the *l*-diversity method based on *k*-anonymous, in order to ensure that the specific place and the target user can not be associated by attackers.

The main idea of position dummies is that the real location of target user should be converted to one or more fake locations. Then the fake locations should be published to the LBS provider by the trust third parties. The method of position dummies to protect privacy is firstly proposed by H. Kido *et al.* [6]. L. Y. Man *et al.* [7] proposed a SpaceTwist framework to examine the trade-off between location privacy, query performance, and query precision. Because the false position method is easy to cause a large overhead on mobile devices with limited resources, X. Liu *et al.* [8] built the Bayesian game model according to the generation process of dummy element position, and proposed a strategy selection algorithm to help users achieve the optimal overhead. However, due to the large amount of noise in the data sent to the LBS provider, a large number of redundant results were still returned to the mobile client, resulting in high communication costs.

The method of mix zone is to construct an anonymous space area according to the true location of target user. Then the anonymous space area is published to the LBS provider. Relying on trusted third parties, the method of mix zone can better reduce the load on mobile clients. B. Gedik *et al.* [9] firstly proposed a unified personalized privacy protection framework, which constructed the *k*-anonymity model according to the user's context-aware personalized privacy requirements. P. Kalnis *et al.* [10] studied a quantified privacy protection intensity of spatial anonymity, constructed an appropriate anonymous area, and discussed the trade-off between anonymous areas. Nevertheless, attackers can still pose a threat to users' privacy through location-dependent attacks. X. Pan *et al.* [11]

proposed an anonymous algorithm based on increased threshold by considering two privacy indicators, such as *k*-anonymity and anonymous granularity to defend against positional dependency attacks.

Encryption technique can effectively protect privacy of user location without utilizing the trust third parties. G. Ghinita *et al.* [12] firstly proposed a framework based on private information retrieval to support private location-dependent query. As the framework achieved privacy protection through encryption technology, the participation of trusted third parties was not required. At the same time, S. I. Ahamed *et al.* [13] proposed a P2P-LBS framework, which utilized a simple but robust generalization technology to protect users' location privacy, thereby removing the participation of trusted third parties.

Based on the above analysis, this paper adopts the idea of *k*-anonymous technology to protect the location privacy of target user. Privacy preservation is significant for LBSs in mobile communication networks [14]. Recent years, *k*-anonymous based location protection strategy has evoked enormous research interest. A new protection technology based on homomorphic encryption against privacy threats in WiFi fingerprint location service was proposed in [15], to address the location privacy challenge. In [16], the dummy location information was added to the location request, in order to make it impossible for the location server to distinguish the real request information and location information of the user.

3. OVERVIEW OF KA-ASS SCHEME

In this section, we give the problem definition, scheme design and attack model of KA-ASS scheme.

3.1 Problem Definition

Definition 1 (Location fingerprint): It is used to represent the AP collected at a location and the collection of RSSI corresponding to the AP, which represented as *fp*. For example, the location fingerprint of location *loc* can be denoted by $fp_{loc} = \{(ap_1, rrsi_1), (ap_2, rrsi_2), \ldots, (ap_n, rrsi_n)\}$. *n* denotes the number of AP that can be collected.

Definition 2 (Real location fingerprint): The fingerprint information submitted by users collected at a certain location contained in the location request, which represented as fp_{real}.

Definition 3 (Dummy location fingerprint): The fingerprint information is generated by the anonymous knowledge base of anonymous server through privacy protection algorithm, which represented as fp_{dummy}.

Definition 4 (Initial AP): The first AP selected by privacy protection algorithm at the time of building fp_{dummy}, which represented as $ap_{initial}$.

Definition 5 (Positioning request): It is used to represent the content of the query request sent by the user when using the location service. For example, the positioning query request in location *loc* can be denoted by $query = (user\_key, fp_{loc})$. *user_key* denotes the unique identifier for the user's identity.

Definition 6 (Anonymous positioning request): It is used to represent the positioning

query request that has been processed anonymously. The *k*-anonymous based privacy protection is realized by adding $k-1$ dummy location fingerprint that attackers cannot distinguish in the positioning query request, so as to the actual location fingerprint is failed to be discovered by attackers. The anonymous positioning query request can be denoted by $query^* = (user\_key, fp_1, fp_2, \ldots, fp_k)$. $[fp_1, fp_2, \ldots, fp_k]$ includes one fp_{real} and $k-1$ fp_{dummy}.

3.2 Scheme Design

Fig. 2 shows the system architecture of KA-ASS, which concludes mobile clients, anonymous server and location provider. Mobile clients are the devices which can collect the nearby AP information, *e.g.* smartphones, smartwatches or PAD, *etc.* Anonymous server is responsible for building, storing, and updating the anonymous knowledge base (AKB). Location provider is responsible for calculating the location information based on the location fingerprint information. First, mobile clients send user identification, collected location fingerprint and privacy parameters to anonymous server. Then, anonymous server generates $k-1$ fp_{dummy} according to AKB and sends the anonymous positioning request to location provider. Location provider computes each location information according to the location fingerprint and sends the results to anonymous server. When anonymous server receives the returned set of locations, the fp_{real} can be filtered out for target user. Finally, the anonymous server sends the location data (*lon*, *lat*) processed by privacy-preserving to mobile clients.

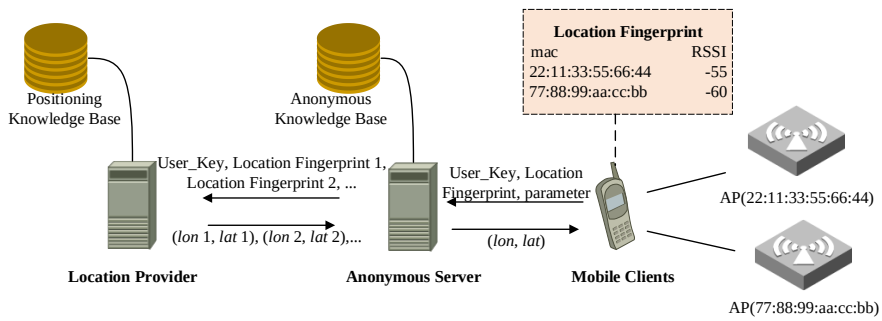


Fig. 2. The system architecture of KA-ASS.

The privacy parameter of users mainly includes anonymous degree *k* and location fingerprint capacity *n*. Anonymous degree represents the number of fp_{dummy} in the anonymous positioning request. The more the number of fp_{dummy}, the stronger the privacy protection. Location fingerprint capacity represents the maximum number of AP that can be included in the location fingerprint of users. The more the number of AP, the higher the positioning accuracy. However, when the number of AP in the location fingerprint is too much, the positioning accuracy will not increase with the increase of the number of AP.

3.3 Attack Hypothesis

(1) Trusted boundary

In this paper, it supposes that mobile clients are trusted and the positioning service requests can only be triggered by the user. The location fingerprint information collected

by mobile devices in the positioning requests is real. The location information returned by LP will not be stolen by other attackers through malicious software or other means. Network communication is reliable, which is guaranteed by the technology in the field of communication security. The third-party anonymous server is trusted. It will strictly enforce the location privacy-preserving policy and not collude with LP to launch attacks. Also, the third-party anonymous server will not record and analyse the location information of users.

(2) Ability boundary of attacker

In this paper, it supposes that LP is untrustworthy, which is the attacker that causes location privacy leakage. When the victim sends the positioning request to LP, it will honestly provide the positioning service for the victim. However, the geographical location information and continuous request sequence of the victim is recorded in LP. Meanwhile, the attack analyses the positioning request and positioning result of the victim, and infers the real location of the victim from the location fingerprint layer and location layer. First, LP owns the positioning service request data of all local users, so as to calculate the query probability of different regions. Second, LP can track the continuous positioning requests of victims and analyse their movement pattern. Finally, LP understands the realization process of location privacy-preserving based on anonymous server, that is, it knows the location privacy-preserving methods used by anonymous server.

3.4 Trajectory-Based Attacks

The continuous location requests will be sent when some applications are running in mobile clients, such as maps or POI recommendation. Suppose that users can acquire their current location and destination location by utilizing the Apps (*e.g.* baidu map, google map, *etc.*), and go to the destination according to the directions of the map. During this process, users can acquire the change of their location in real time and determine whether they are heading towards the destination correctly. When the users make the continuous location request in a short time, the real locations of the users satisfy the certain trajectory characteristics in space. However, the traditional anonymous algorithm does not consider the correlation between two consecutive anonymous location requests, which causes the generated false location is random.

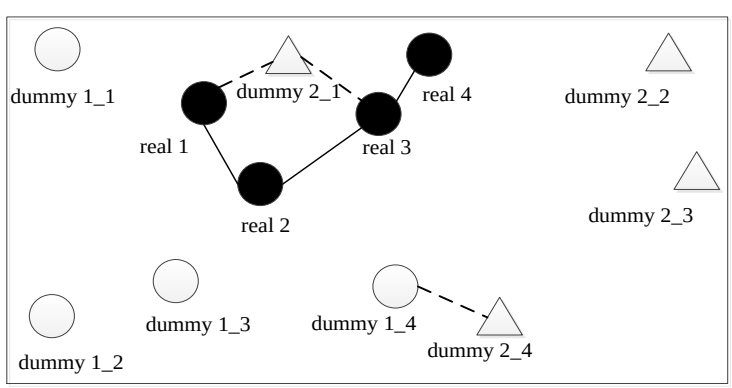


Fig. 3. The example of trajectory-based attacks.

As shown in Fig. 3, the user makes use of the anonymous positioning request generated by 3-anonymous for four consecutive times. The data of [real 1, real 2, real 3, real 4] is the real location obtained for the positioning request submitted by the user in order, while the dummy location generated by the anonymous algorithm may be [dummy 1_1, dummy 1_2, dummy 1_3, dummy 1_4] and [dummy 2_1, dummy 2_2, dummy 2_3, dummy 2_4]. Even though three location fingerprints are contained in each positioning request, the number of trajectory sequence may be 3^4 (*i.e.* 81). However, only [real 1, real 2, real 3, real 4] and [real 1 dummy 2 1, real 3, real 4] can meet the trajectory characteristics. In this case, the attackers can easily infer the three locations visited by victim according to the trajectory characteristics. What's more, the attackers can infer the possible area where victim sends the next positioning request according to the location of real 4.

4. MODELS AND ALGORITHMS

The definition of *k*-anonymous is that it requires the published data contain *k* indistinguishable identifiers. Also, the probability of a given individual being found is 1/*k*. For location privacy-preserving, a set of queries with *k* users is generated by *k*-anonymous technology. Users utilize the common anonymous area formed by the set of queries. In general, three parameters are required in *k*-anonymous technology, *i.e.* the degree of anonymity *k*, minimum anonymous region $A_{\min}$ and maximum delay time $T_{\max}$.

In order to make the anonymous trajectory sequence has high similarity with the real trajectory sequence, there are two characteristics need to be satisfied for anonymous trajectory sequence.

(1) Directional similarity

The real trajectory sequence of each user is directional because of the destination of target user. If the difference on moving direction between real trajectory and anonymous trajectory is high, the attacker can easily deduce the real trajectory. Thus, the directional similarity needs to be considered for *k*-anonymous.

(2) Distance similarity

The distance between two continuous dummy locations in anonymous trajectory should be similar with the distance between two corresponding real locations. Otherwise, the attacker can exclude the false trajectory according to the background knowledge of target, such as average moving speed. Thus, the distance similarity also needs to be considered for *k*-anonymous.

In this section, COFS algorithm is proposed to generate the anonymous trajectory sequence. It requests that the generated fp_{dummy} must need satisfy the directional similarity and distance similarity with real trajectory sequence during the continuous two positioning queries.

4.1 Construction of Anonymous Knowledge Base

The target of KA-ASS scheme is to make the dummy location fingerprint fp_{dummy} generated by the anonymous server and the real location fingerprint fp_{real} submitted by the user indistinguishable from the LP. In this paper, the anonymous knowledge base (AKB) is

constructed based on the weighted undirected graph model, in order to store the information of AP and discover the spatial distribution of AP.

Definition 7: Graph $G = \{V, E, W\}$ is the weighted undirected graph, thereinto,

(1) V is the set of points in G, $v_i \in V$ denotes one AP.
(2) E is the set of edges in G, $e_{ij} \in E$ iff the signal region of v_i and v_j cover each other.
(3) W is the set of weights in the edges of G, $w_{ij} \in W$ denotes that v_i and v_j are the distance estimation of AP, respectively.

LP does not share the location of APs or graph G with anonymous server because the positioning base is the key resource of LP. In order to discover the dummy location for real location, the AKB is built for anonymous server. In AKB, the AP sets of fp_{real} submitted by all users and the corresponding RSSI are utilized to evaluate the relationship of AP in G. The detailed AKB construction method is shown as following.

(1) For any AP belonging to fp_{real}, add AP to the set of points V.
(2) For any AP_i and AP_j belonging to the same fp_{real}, the signal strength of AP_i and AP_j must cover each other, add e_{ij} to the set of edges E.
(3) For any AP_i and AP_j belonging to fp_{real}, the estimation of the distance of the acquisition position relative to AP_i and AP_j is approximately calculated according to the WiFi signal intensity attenuation model [17]. Eq. (1) shows the computational process.

$$P_L(d) = A - 10\lg(d), \tag{1}$$

where $P_L(d)$ denotes the RSSI measured at the target point, d represents the distance of target point, n is the signal attenuation factor, A is the value of wireless signal strength received by the signal receiver when it is 1 meter away from the signal sender.

The distance d_i between AP_i and target point can be computed as Eq. (2).

$$d_i = 10^{[A-P_L(d_i)]/10n} \tag{2}$$

Based on the collected fingerprints of all users, the weight w_{ij} of edge e_{ij} can be computed as Eq. (3),

$$w_{ij} = \min_{\forall fp \to (ap_i \in fp) \wedge (ap_j \in fp)} (10^{\lceil A-P_L(d_i) \rceil / 10n} + 10^{\lceil A-P_L(d_j) \rceil / 10n}). \tag{3}$$

4.2 Grid Based AP Division

The whole space can be divided according to the grid with AP graph model. The bi-directional index is utilized to record the set of AP in each grid and the network of each AP.

The grid based AP division is shown in Fig. 4. In the left part of Fig. 4, each record in Cell table includes a tuple (Cell ID, AP Set), where Cell ID represents the unique network and AP Set denotes the existed set of AP in this network. In the right part of Fig. 4, each record in AP table includes a tuple (AP ID, Cell ID), where AP ID represents the

access point with a unique MAC address.

In order to divide AP into the corresponding network, the location of each AP needs to be considered. However, LP cannot share the location of AP to anonymous server because of the protection of positioning base. Thus, the reverse triangle positioning model [18] is used to predict the location information of AP. The basic idea is that the location fingerprints submitted by multiple users may contain the signal emitted by each AP. The distance of AP relative to the registration point can be calculated according to the positioning results and the transmission loss model. The coordinates of the AP can be computed as follows.

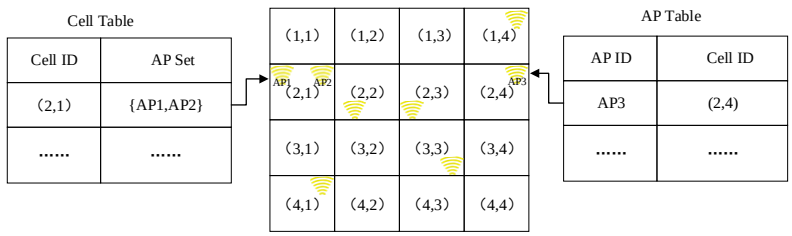


Fig. 4. The process of grid based AP division.

$$\begin{Bmatrix} (x-x_1)^2+(y-y_1)^2=d_1^2 \\ (x-x_2)^2+(y-y_2)^2=d_2^2 \\ (x-x_3)^2+(y-y_3)^2=d_3^2 \end{Bmatrix}, \tag{4}$$

$$(x,y)=\left(\frac{x_{12}+x_{23}+x_{13}}{3},\frac{y_{12}+y_{23}+y_{13}}{3}\right). \tag{5}$$

4.3 COFS Algorithm

In order to achieve *k*-anonymous, Continuous Optimal Fingerprint Selection (COFS) algorithm needs generate $k-1$ fp_{dummy} according to the anonymous knowledge base. At the same time, the directional similarity and distance similarity should be satisfied between real trajectory sequence and false trajectory sequence. Through COFS algorithm, the real trajectory sequence of target users cannot be discovered by the attackers.

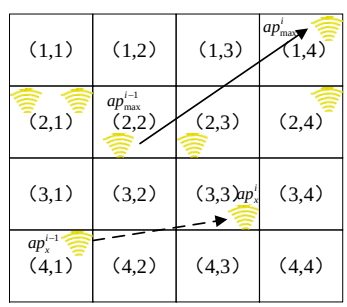


Fig. 5. The selection of initial AP.

For the first anonymous positioning request, the trajectory characteristic is not considered because of location anonymous. For the *i*th ($i > 1$) anonymous positioning request, it should consider the change of location in the continuous two positioning requests. The current location and movement trend cannot be acquired because of the security of AP. COFS algorithm measures the user's movement trend by using the AP with the highest signal strength in the grid in the fingerprint of two consecutive location requests. Let $ap_{\max}^{i-1}$ and $ap_{\max}^{i}$ denote the AP which the value of RSSI is maximal for the *i*–1th and *i*th fp_{real}, respectively. Let ap_x^{i-1} and ap_x^i denote the selected $ap_{initial}$ in the process of building fp_{dummy} for the *i*–1th and *i*th anonymous, respectively. Eq. (6) should be satisfied for COFS algorithm.

$$cell(ap_{\max}^{i}) - cell\left(ap_{\max}^{i-1}\right) = cell\left(ap_x^{i}\right) - cell\left(ap_x^{i-1}\right), \tag{6}$$

where $cell(ap)$ denote the grid coordinates of ap. COFS algorithm ensures that the location vector of $ap_{initial}$ is same with the location vector of actual AP in the continuous two anonymous requests. For example in Fig. 5, the location vector with maximal RSSI in fp_{real} is (1, 4) – (2, 2) = (–1, 2). Thus, the $ap_{initial}$ can be selected in the grid (4, 1) + (–1, 2) = (3, 3) while generating the fp_{dummy}.

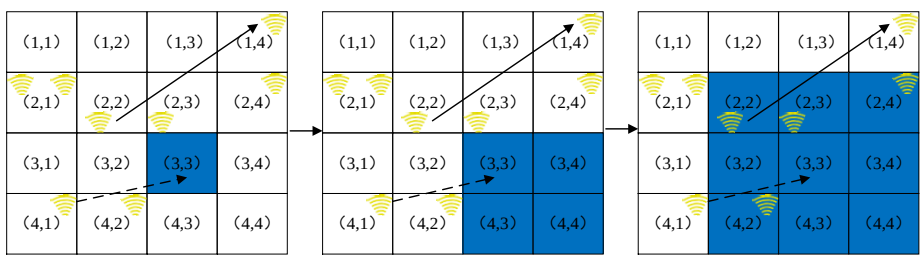


Fig. 6. The example of candidate grid extension.

However, it should be expanded when the candidate grid does not contain any APs. As shown in Fig. 6, The COFS algorithm increases the side length of the candidate grid by 1 each time. If the current grid side length is odd, it expands to the lower right direction. If the current grid side length is even, it expands to the upper left direction, until there is an AP to choose from in the candidate grid.

Algorithm 1: COFS

Input: User fingerprint $fp_{real}((ap_1, rssi_1), (ap_2, rssi_2), \ldots, (ap_n, rssi_n))$, $ap_{\max}^{i}$ is the AP with maximum RSSI in fp_{real}. Fingerprint capacity n, Degree of anonymity k. The list of AP with the largest RSSI value of each location fingerprint in the $i-1$ anonymous set $list_{i-1}: \left[ap_{\max}^{i-1}, ap_1^{i-1}, ap_2^{i-1}, \ldots, ap_{k-1}^{i-1}\right]$.

Output: Anonymous set *Ano*

1: Initialize $Ano \leftarrow \varnothing$

2: $(x, y) \leftarrow cell\left(ap_{\max}^{i}\right) - cell\left(ap_{\max}^{i-1}\right)$;

3: **for** AP in $\left[ap_1^{i-1}, ap_2^{i-1}, ..., ap_{k-1}^{i-1} \right]$ **do**
4: $(x', y') \leftarrow cell(ap) + (x, y)$;
5: Assign the AP set in the grid (x', y') to *ap_set*;
6: **while** *ap_set* == null **do**
7: Implement the grid expansion policy and assign the AP set of the expanded grid to *ap_set*;
8: Randomly select an AP from the *ap_set* as $ap_{initial}$;
9: Generate fp_{dummy} for each AP and add it into *Ano*;
10: **end for**
11: Add fp_{real} into *Ano*;
12: Store the anonymous $ap_{initial}$ and $ap_{\max}^{i}$ as the next input $list_i$;
13: Return *Ano*

Algorithm 1 shows the pseudo-code of COFS algorithm. The detailed steps are as following.

(1) It initializes the anonymous in order to store the anonymous result set, including the fp_{real} and $k-1$ fp_{dummy}. The fp_{real} can be cut out by retaining the first n APs with maximal RSSI.
(2) It determines whether there is a list $list_{i-1}$ consisting of $ap_{initial}$ and $ap_{\max}^{i-1}$, which were selected anonymously last time. If the $list_{i-1}$ doesn't exist, this is the first anonymous processing for the trajectory privacy-preserving. Thus, the anonymous location set can be directly built by COFS algorithm, then perform Step 4. If the $list_{i-1}$ exists, it performs Step 3.
(3) Firstly, the location vector (x, y) of $ap_{\max}^{i-1}$ and $ap_{\max}^{i}$ in the grid while user submits the location fingerprint with two continuous positioning requests are calculated. Then, for any AP in $\left[ap_1^{i-1}, ap_2^{i-1}, ..., ap_{k-1}^{i-1} \right]$, the location coordinates (x', y') of candidate set can be computed according to reverse triangular positioning. If the AP set in grid (x', y') is empty, it performs the grid expansion strategy. Finally, the fp_{real} is added in the anonymous result set.
(4) The $ap_{initial}$ and $ap_{\max}^{i}$ that are used anonymously are stored in $list_i$ as the input of the next anonymous. Then, it returns the generated anonymous result set.

4.4 Security Analysis

KA-ASS scheme generates dummy location fingerprints to implement *k*-anonymity, iff LP cannot distinguish the real location fingerprint from the *k* location fingerprints. That is to say, the probability of LP getting the real location of target user will be $1/k$. As discussed above, the dummy location fingerprints can be inferred according to the existence, similarity and selectivity. For the existence of APs, real location fingerprints are utilized to construct the AKB, in order to discover the relationship between locations. It means that the APs in dummy location fingerprints must be the actual location in the real world. For the similarity of APs, the APs with edges on the initial AP are chose in COFS algorithm. It leads the APs are close to each other in the actual environment. For the selectivity, the APs with higher weight on edges to the initial AP are used, the higher expected RSSI

means the similar between the two APs. In the KA-ASS scheme, the APs closest to the initial AP are select, and the reasonable RSSIs are assigned. Because the COFS algorithm satisfies the above three security points, the LP can not infer the real location information of target user by utilizing proposed COFS algorithm theoretically.

5. PERFORMANCE EVALUATION

In this section, we conduct several experiments based on a real dataset and a simulated dataset to evaluate the performance of our proposed scheme in terms of anonymous time cost, anonymous quality and trajectory similarity.

5.1 Datasets and Experimental Setup

Datasets. We use the dataset to verify the validity and efficiency of our proposed algorithms. GeoLife datasets [19] has recorded the GPS trajectory of 182 users with 18670 trajectories in five years (from 2008/10 to 2012/8), which not only includes the daily activity, but also includes the recreational activity. Fig. 7 shows the distribution of some locations in GeoLife dataset by making use of Baidu maps. Also the data format of the collection is shown in Table 1, including the information of latitude and longitude collected by GPS, the corresponding AP and RSSI at the current position.

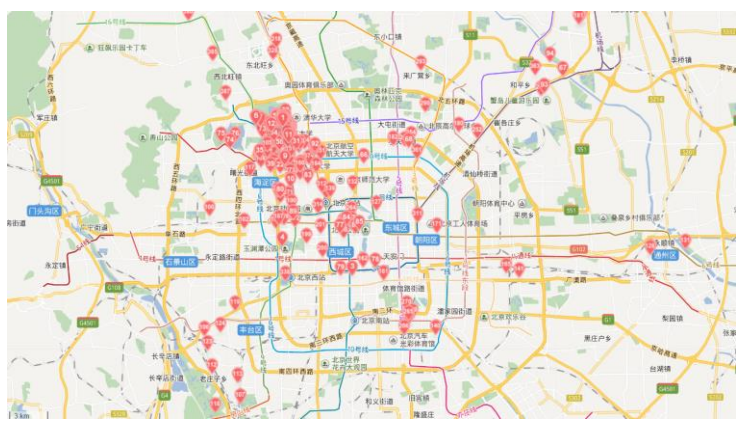


Fig. 7. The collected user location information in GeoLife dataset.

Table 1. The collected user fingerprint information.

Location Information (latitude, longitude)	User Fingerprint	
	AP(MAC)	RSSI(dB)
45.767182, 126.6143162	70:ba:ef:dc:5f:40	–55
	a8:15:4d:c3:92:4c	–57
	1a:dc:56:f0:d0:f3	–60
45.770261, 126.6133323	1a:dc:56:f0:d0:f3	–63
	70:ba:ef:d9:bf:50	–67
	70:ba:ef:dc:5f:40	–68
…	…	…

Experimental Setup. Firstly, the AKB is established by using all the collected location fingerprint information, including AP graph *G* and AP spatial grid distribution. Secondly, in order to simulate the access frequency of different location areas, we use Miller projection to convert the longitude and latitude coordinates of the positioning result information into two-dimensional plane coordinates. The 20m*20m range in the coordinate system was taken as the access region, and the location fingerprints of different positions in the region were used to simulate multiple location requests of users. The AP score list of the AKB is also constructed during the simulation of the location request, and the visiting number of each region is recorded. The visiting frequency p_{Region_i} of $Region_i$ can be computed according to Eq. (7).

$$p_{Region_i} = n_i \Big/ \sum_{j=1}^{N} n_j \tag{7}$$

5.2 Experimental Results and Analysis

In this part, we compare the performance of proposed KS-ASS scheme with KAP scheme [20]. There are three indexes to evaluate the two algorithms as follows.

(1) Anonymous time cost

The anonymous time represents the total time it takes from the system to receive the location request from the user to complete the anonymity of the location request. The shorter the anonymous time overhead, the less time the user needs to spend on location privacy protection, and the more efficient the algorithm is.

(2) Anonymous quality

The anonymity quality is to measure the similarity between fp_{dummy} and fp_{real} generated by privacy protection algorithm. The higher the quality of anonymity, the more difficult to distinguish for LP. In this paper, the anonymous quality is measured according to the accuracy radius and the fingerprint identification rate.

(3) Trajectory similarity

The higher the similarity between the user's real moving trajectory and the location of dummy, the harder it is for the attacker to distinguish the user's real trajectory, and the harder it is to find the user's real position. We used the following two indicators to assess the similarity of the false trajectory.

Position Distance Deviation (PD): The average value of the difference between the distance of a location result from fp_{real} and that of a location result from fp_{dummy} in any two consecutive location requests in the trajectory. The value of PD can be computed as Eq. (8).

$$PD = \frac{\sum_{i=0}^{k-1}\sum_{j=0}^{n-1}\left|dist\left(loc_{real}^{j}, loc_{real}^{j+1}\right) - dist\left(loc_{i}^{j}, loc_{i}^{j+1}\right)\right|}{(k-1)*(n-1)}, \tag{8}$$

where loc_{real}^{j} denotes the location of *j*th positioning request in the real trajectory, and loc_i^j denotes the location of *j*th positioning request with the *i*th trajectory in the anonymous set. $dist(loc_i, loc_j)$ represents the distance between loc_i and loc_j, which can be computed as Eq. (9).

$$dist\left(loc_i, loc_j\right) = 2R*\arcsin\left(\sqrt{\sin^2\left(\frac{lat_i - lat_i}{2}\right) + \cos\left(lat_i\right)*\cos\left(lat_j\right)*\sin^2\left(\frac{lon_i - lon_j}{2}\right)}\right) \tag{9}$$

PD is designed to reflect the similarity between the real trajectory and the false trajectory during the distance change of each location request. The smaller the PD value is, the higher the similarity is. The higher the value of PD, the lower the similarity.

Position Angle Deviation (PA): The average value of the angle difference at any three consecutive positions corresponding to the real trajectory and the false trajectory.

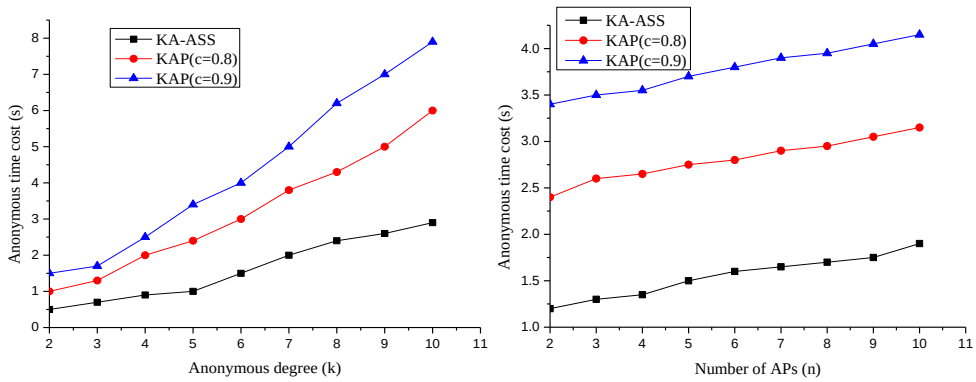


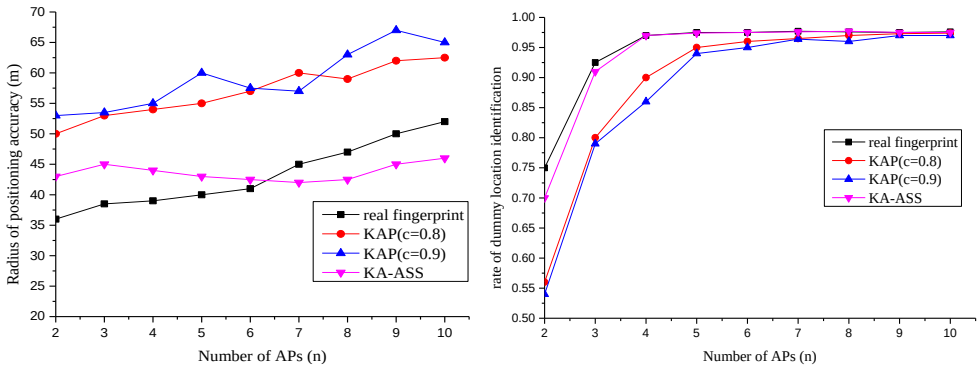
Fig. 8. Effect of anonymous time cost.

Fig. 8 shows the effect of anonymous time cost for anonymous degree and number of APs, respectively. From the experimental results, we can see that the anonymous time cost of KA-ASS and KAP increases with the increase of the anonymity degree. The reason is that the increase of k means the increase of privacy level, which causes more fp_{dummy} needs to be generated. Thus, it leads the cost of calculation time increasing. Under the condition of a certain anonymity degree, the anonymous time cost of KA-ASS and KAP also increases continuously with the increase of the number of APs. However, the number of APs in location fingerprint has little influence on the anonymous time cost.

The anonymous time cost of KAP is significantly higher than the proposed KA-ASS scheme. The anonymity time of KA-ASS is about twice than the KAP when the aggregation coefficient is 0.9. For KAP scheme, the time cost with the aggregation coefficient threshold of 0.8 is less than that with the aggregation coefficient threshold of 0.9. The reason is that the proposed KA-ASS can directly select the n–1 AP closest to the spatial distribution of $ap_{initial}$ from the AP graph G to form fp_{dummy}. However, the KAP scheme needs to check the aggregation coefficient corresponding to $ap_{initial}$ after selecting $ap_{initial}$. The $ap_{initial}$ needs to be reselected when the aggregation coefficient is less than the specified threshold, which takes more time. What's more, the larger threshold means the selection of $ap_{initial}$ is more stringent, which causes more time cost.

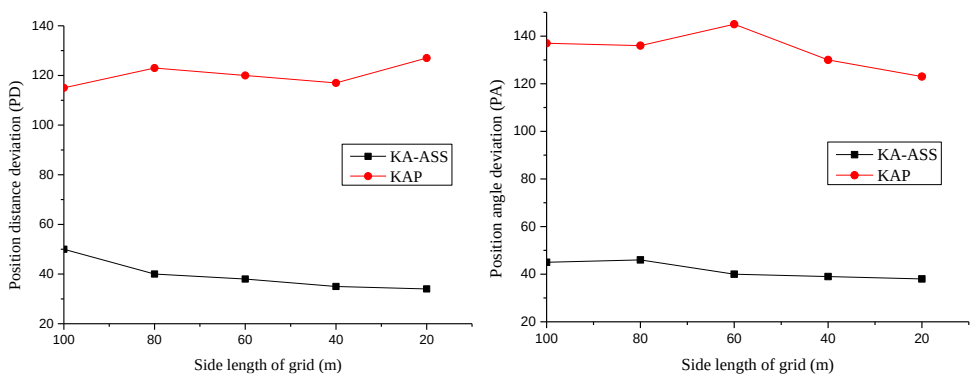
Fig. 9 shows the effect of anonymous quality for the number of APs. From the experimental results, we can see that the radius of positioning accuracy and rate of dummy location of KA-ASS and KAP increase with the increase of the number of APs. The fp_{dummy} constructed by KA-ASS is superior to KAP in positioning accuracy and fingerprint identification rate. The reason is that KAP scheme only considers the existence and proximity of APs during the process of constructing fp_{dummy}. On this basis, KA-ASS scheme considers

the construction mode of fp_{real}. Also, the fp_{dummy} is generated by choosing the closest APs with $ap_{initial}$, in order to ensure the same performance of fp_{real} in the positioning process. Thus, the proposed KA-ASS scheme has higher anonymous quality than KAP scheme.



(a) (b)

Fig. 9. Effect of anonymous quality.



(a) (b)

Fig. 10. Effect of trajectory similarity.

Fig. 10 shows the effect of trajectory similarity for the side length of grid. From the experimental results, we can see that KA-ASS scheme has smaller PD and PA compared with KAP scheme. The reason is that the COFS algorithm does not consider the relationship between the false trajectory and the real trajectory in the anonymous processing of two consecutive positioning requests, so the continuous false location cannot be guaranteed to meet the trajectory characteristics. KA-ASS scheme indirectly represents the change of user location for two consecutive times according to the location vector of AP in fp_{real}. Also, the fp_{dummy} is generated in this anonymous according to the fp_{dummy} generated by the previous anonymous and location vector. Thus, the value of PD and PA for KA-ASS is smaller than KAP. It represents that the similarity between real trajectory and false trajectory is higher by using KA-ASS scheme, so as to well protect the trajectory information of target users.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we study the problem of anti-positioning security strategy under the k-anonymous model in mobile networks. First, we introduce LBS which takes positioning service into account to satisfy the personalized demand of users. Then, we explore the possibility of designing a K-Anonymous based Anti-positioning Security Strategy (KA-ASS), which can ensure great performance of anonymous time cost and anonymous quality. By experiments, it shows that our KA-ASS scheme can privately provide positioning services with high efficiency. For the future work, we will further complete attack model by considering social connection graph or text of content, *etc.* Deep learning technology will be utilized to construct user model, in order to intelligently perceive users' practical location and select dummy location.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61902361, in part by the Henan Provincial Science and Technology Department under Grant Nos. 212102210095 and 202102210176.

REFERENCES

1. J. A. del Peral-Rosado, R. Raulefs, J. A. Lopez-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1g to 5g," *IEEE Communications Surveys & Tutorials*, Vol. 20, 2017, pp. 1124-1148.
2. P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: *k*-anonymity and its enforcement through generalization and suppression," in *Proceedings of IEEE Symposium on Research in Security and Privacy*, 1998, pp. 1-19.
3. M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st ACM International Conference on Mobile Systems, Applications and Services*, 2003, pp. 31-42.
4. M. F. Mokbel, C. Y. Chow, and W. G. Aref, "The new casper: Query processing for location services without compromising privacy," in *Proceedings of the 32nd International Conference on Very Large Databases*, 2006, pp. 763-774.
5. B. Bamba, L. Liu, P. Pesti, *et al.*, "Supporting anonymous location queries in mobile environments with privacy grid," in *Proceedings of the 17th ACM International Conference on World Wide Web*, 2008, pp. 237-246.
6. H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Proceedings of IEEE International Conference on Pervasive Services*, 2005, pp. 88-97.
7. L. Y. Man, C. S. Jensen, X. Huang, *et al.*, "SpaceTwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services," in *Proceedings of IEEE International Conference on Data Engineering*, 2008, pp. 366-375.
8. X. Liu, K. Liu, L. Guo, *et al.*, "A game-theoretic approach for achieving *k*-anonymity in location based services," in *Proceedings of IEEE International Conference on*

Computer Communications, 2013, pp. 2985-2993.
9. B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in *Proceedings of IEEE International Conference on Distributed Computing Systems*, 2005, pp. 620-629.
10. P. Kalnis, G. Ghinita, K. Mouratidis, *et al.*, "Preventing location-based identity inference in anonymous spatial queries," *IEEE Transactions on Knowledge and Data Engineering*, 2007, Vol. 19, pp. 1719-1733.
11. X. Pan, J. Xu, and X. Meng, "Protecting location privacy against location-dependent attacks in mobile services," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24, 2012, pp. 1506-1519.
12. G. Ghinita, P. Kalnis, A. Khoshgozaran, *et al.*, "Private queries in location based services: anonymizers are not necessary," in *Proceedings of ACM SIGMOD International Conference on Management of Data*, 2008, pp. 121-132.
13. S. I. Ahamed and C. S. Hasan, "A novel location privacy framework without trusted third party based on location anonymity prediction," *ACM Sigapp Applied Computing Review*, 2012, Vol. 12, pp. 24-34.
14. X. He, R. Jin, and H. Dai, "Leveraging spatial diversity for privacy aware location-based services in mobile networks," *IEEE Transactions on Information Forensics and Security*, Vol. 13, 2018, pp. 1524-1534.
15. H. Li, L. Sun, H. Zhu, X. Lu, and X. Cheng, "Achieving privacy preservation in WiFi fingerprint-based localization," in *Proceedings of IEEE Conference on Computer Communications*, 2014, pp. 2337-2345.
16. Y. Baseri, A. Hafid, and S. Cherkaoui, "K-anonymous location-based fine-grained access control for mobile cloud," in *Proceedings of the 13th IEEE Annual Consumer Communications and Networking Conference*, 2016, pp. 720-725.
17. J. Zhou, P. Zhang, and A. F. Xin, "Study on adaptive algorithm for wireless sensor network indoor positioning," *Journal of Geodesy and Geodynamics*, Vol. 32, 2012, pp. 74-77.
18. G. Caso, L. De Nardis, F. Lemic, V. Handziski, A. Wolisz, and M. D. Benedetto, "ViFi: virtual fingerprinting WiFi-based indoor positioning via multi-wall multi-floor propagation model," *IEEE Transactions on Mobile Computing*, Vol. 19, 2020, pp. 1478-1491.
19. Y. Zheng, X. Xie, and W. Y. Ma, "Geolife: A collaborative social networking service among user, location and trajectory," *Bulletin of the Technical Committee on Data Engineering*, Vol. 33, 2010, pp. 32-39.
20. Y. Wang, H. Zhang, and X. Yu, "KAP: location privacy-preserving approach in location services," *Journal of Communications*, Vol. 35, 2014, pp. 182-190.

Liang Zhu (朱亮) received Ph.D. degree in Computer Science and Technology from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in October 2017. He is currently a Lecturer with the Institute of Computer and Communication Engineering at Zhengzhou University of Light Industry, Henan, China. His current research interests include mobile social networks, personalized service recommendation, and privacy preserving.

Li-Ping Yu (余丽萍) received her master degree in Computer Application Technology from Henan Polytechnic University, Henan, China, in 2013. She is a Research Associate at Zhengzhou University of Light Industry. His research interests include distributed computing, user behavior analyzing and modeling.

Zeng-Yu Cai (蔡增玉) received his master degree in Computer Application Technology from Northeast Normal University, Changchun, China, in 2006. He is an Associate Professor at Zhengzhou University of Light Industry. His research interests include trusted computing, plan recognition and information security.

Xiao-Wei Liu (刘啸威) is currently a graduate student at Zhengzhou University of Light Industry. His research interests include location-based social networks, personalized service recommendation, and user privacy protection.

Jian-Wei Zhang (张建伟) received his Ph.D. degree in Computer Application Technology from PLA Information Engineering University in 2010. He is a Professor at Zhengzhou University of Light Industry. His research interests include broadband information network and network security.