International Journal of Security and Its Applications
Vol. 12, No. 2 (2018), pp.29-44
http://dx.doi.org/10.14257/ijsia.2018.12.2.04

Query-Dependency-Aware Location Privacy Protection for Road Networks

Hui Chen[1] and Xiaolin Qin[2]

[1,2]College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[1]School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, China
[1,2]Jiangsu Key Laboratory of Internet of Things and Control Technology, Nanjing, China
[1]chenhuinuaa@nuaa.edu.cn, [2]qinxcs@nuaa.edu.cn

Abstract

Effective location privacy protection policies could protect mobile users from suffering location privacy leakage threat when they enjoy location-based services (LBS). The anonymization effectiveness will suffer great deterioration if attacker obtains more background knowledge with respect to user's historical queries in conventional location privacy protection approaches mainly based on spatial cloaking. Most mobile users move over road networks and the dependency of their queries significantly affect the anonymous space privacy degree. This paper presents a query-dependency-based location privacy protection method for road networks. In the proposed method, the personalized privacy requirement of user is adequately considered. The feasibility and validity of the proposed method are verified through experiments for many scenarios.

Keywords*: Query dependency, privacy protection, road networks*

1. Introduction

With the rapid developments of wireless communication technology, positioning technology, and mobile object database management technology, especially, as well as the rise of Internet of Things, location-based services (LBS) have being extended from the early application of quickly positioning patient in emergency to wider fields [1,2]. More and more users use smart phones or other mobile terminals to obtain personalized LBS by providing their current location information. For instances, individual user can issue a request of finding the nearest restaurant; traffic administrative department can master the road traffic report any time; merchants can send electronic ads or coupons to customers nearby. LBS can also provide location-based weather forecast services and emergency medical services, *etc*.

LBS brings great convenience to people's production and living, which continually expresses its potentials practical value and development prospect in wide fields. However, LBS meanwhile poses serious threat to user's privacy [3]. In LBS, the server processes user's service request in terms of the received location information sent to it by user. Hence, the quality of LBS is determined by the accuracy of user's location information. Although the service providers claim that they are safe and reliable, the user is bound to reveal his location information as enjoying services in this kind of service providing mode. According to the revealed location information, the attacker can further deduce other private information such as identity, behavior, habits, and *etc.* of the user while he does not want to reveal. If the server suffers malicious attack, or even the server itself is

Received (September 5, 2016), Review Result (November 17, 2016), Accepted (January 17, 2017)

ISSN: 1738-9976 IJSIA

an attacker, it will cause a greater loss once the attacker had got user's privacy information from the server. Effectively protecting user's privacy is particularly important to the prosperity and development of LBS market, and only in this way the user can safely enjoy LBS. Clearly, it is extremely necessary to and great important to investigate privacy protection techniques for LBS.

The topic of privacy protection in LBS has garnered wide attention by researchers, governments, enterprises, and *etc*. One aspect of LBS-related privacy threat is location privacy which mainly refers to user's private information directly related to his location, such as user's accurate location and his interests or habits that may be deduced from his location contained in LBS request [3].

So far, location obfuscation method with *k*-anonymous idea is mainly used for location privacy protection in LBS [4-12]. In this method, the accurate location of user who sends LBS request is generalized to a region and then referred as an anonymous space to be sent to the LBS server. By doing so, location of the user sent request is indistinguishable from that of at least other *k*-1 users, which reduces the possibility of user's accurate location being deduced by the attacker. However, with the rapid growth of Internet applications, attackers are able to obtain more information relevant to the users and thus construct new background knowledge to reduce the privacy degree of anonymous space, which makes it is easier to deduce user's location for them. This situation brings a new challenge to existing location privacy protection methods and the problem briefly stated here may be illustrated through an example given as follow.

Example 1 Tom sends a LBS request to find the nearest bar. In order to protect location privacy, his accurate location is generalized to an anonymous region which contains locations of two mobile users who are here named as Tom and Jan. Suppose the attacker has learnt from the obtained background knowledge that Tom and Jan just respectively issued a request to find the nearest restaurant and clinic. Because it is significantly less possible to issue a request of finding a bar after finding a clinic than a restaurant, the attacker could easily exclude Jan from the anonymous region and then deduce Tom's accurate location.

In example 1, the constructed anonymous space is of low privacy degree since the attacker has obtained more background knowledge, *i.e.*, request history, relevant to the user. According to the dependency of two successive queries in the user's request history, the attacker could reduce the anonymous space and thus improve reasoning accuracy. Hence, one can see that the background information obtained by attacker directly affects the privacy degree of anonymous space. Here, the query dependency is obviously an important form of that background information. In the location privacy protection method proposed in this work, he query dependency is introduced and regarded as an important factor for constructing anonymous space.

On the other hand, in real life, activities of mobile users are more restricted on road networks, while most existing privacy protection methods are designed in Euclidean space and they are not suitable for road networks environment. In addition, most of these methods cannot meet user's personalized privacy requirement well. Although the parameter *k* is allowed to be set by user in these methods, in practice, there are many other factors such as user's identity, background, habits, and *etc*. affect the privacy degree of anonymous space. However, particularly, these factors are often personalized.

To problems illustrated above, contributions of this work can be summarized as following several aspects:

(1). Defining a new function to define the privacy degree by fully considering the influence of query dependency on measuring privacy degree of anonymous space.

(2). Meeting personalized privacy requirement of user.

(3). Proposing a feasible location privacy protection method for road networks environment.

The remainder of this paper is organized as follows. Section 2 highlights related techniques for location privacy protection. Section 3 gives brief introductions to the background and definitions on related concepts and problems. Section 4 describes the technique of query-dependency-aware location privacy protection for road networks. Section 5 conducts relevant experiments for many scenarios and provides analyses on the results. Section 6 finally concludes the paper.

2. Related Works

Location privacy protection strategy can effectively protect privacy information directly related to user's location from attackers. The existing location privacy protection strategies can be roughly classified into three catagories as follows.

(1) Spacial Cloaking [4, 5]. The basic idea of this technique is using a spatial region to replace user's true location, which makes the attacker unable to distinguish user's ture location with other users' locaitons in the cloaked region.

(2) Obfuscation Schemes [13-15]. In this technique, user's true location is obfuscated by adding dummy locations or some fixed locations (forks in the road networks for example) around the true location, or by transforming the true location with certain strategies (such as expanding or reducing the scope, transferring the gravity center, and *etc*.).

(3) Cryptography-based Schemes [16]. This kind of method completely transform user's location by setting encryption strategies at client. By doing so, the server could handle user's request but not learn his true locaiton.

Spacial Cloaking was proposed earlier and also is the most commonly used method. One so-called *k*-anonymous model firstly proposed by Sweeney [17] is the most commonly used in Spacial Cloaking method. The model of *k*-anonymous was firstly used in privacy preserving for relational databases publication when it was proposed. The main idea of this model is to make the personal information represented by a record which cannot be distinguished with at least other *k*-1 records. Gruteser and Grunwald [4] firstly introduced the idea of *k*-anonymous into location privacy protection and then proposed location *k*-anonymous model. In location *k*-anonymous model, user's accurate location is generalized to be an anonymous space where user's location cannot be distinguished with at least other *k*-1 users' locations.

So far, the location *k*-anonymous model has exposed its shortcomings. The privacy degree of the anonymous space constructed by this model will be significantly reduced when the attacker has mastered more background knowledge. Wernke et al [18] noted that an attacker is able to obtain user's personal background as background knowledge thereby threaten the privacy of user. Xue *et al.*, [19] investigated the impact of location semantics obtained by attacker on the privacy degree of anonymous space and put forward the idea of location diversity to ensure that there are at least *l* different location semantics in the anonymous space. Damiani *et al.*, [20] surveyed the effect of location sensitivity on the privacy degree of anonymous space and accordingly proposed a solution to sensitive location leak in location sharing application of social networks. Chen and Pang [1] classified background information obtained by attacker into static and dynamic classifications and emphatically analyzed the impact of user's static information including profiles and query dependency on user's privacy, and proposed a corresponding privacy protection method for query privacy protection.

The moving objects mainly move on road networks. Most of the proposed location privacy protection techniques mentioned before are designed for Euclidean space while only a few methods for road networks [21-23]. Palanisamy and Liu [21] proposed a location privacy protection method for road networks with dynamic

pseudonym mechanism and mix-zones. Chow et al [22] used adjacent road segments to construct anonymous space and proposed a location privacy protection algorithm for road networks. In this algorithm, request executing cost and query quality are surveyed in the process of constructing anonymous set. However, the methods given above do not consider the effects of background information on constructing anonymous set. Yigitoglu et al [23] proposed a location privacy protection method for road networks under considering the effects of sensitive positions. This method is mainly used for protecting sensitive information in location sharing application of social networks.

In the location privacy protection method proposed in this paper, the influence of background information on anonymous space is taken into account. In detail, the role of query dependency for privacy degree of anonymous space is analyzed and basing on which a feasible location privacy protection method is designed. In addition, the proposed method allows users to set personalized privacy requirements.

3. Preliminaries

3.1. System Architecture

For location privacy protection, there are three commonly used system architectures including non-cooperative structure, centralized structure, and peer-to-peer cooperative structure [24]. In this work, the centralized structure depicted in Figure 1 is adopted. This structure consists of three components, namely client, anonymizer, and LBS server. The anonymizer is introduced as a trusted third party between the client and LBS server. A user sends his precise location, query, and privacy requirement to the anonymizer when he issues a request. The anonymizer cloaks user's precise location according to his privacy requirement, and then sends the anonymous set with user's query to the LBS server. The LBS server calculates the candidate results and sends them to the anonymizer. The anonymizer then analytically processes the candidate results and finally returns the refined query result to the user. In such a system model, to construct effective anonymous set, the anonymizer should save current information of the map and road networks, and real-timely update the information of mobile users moving on road networks.

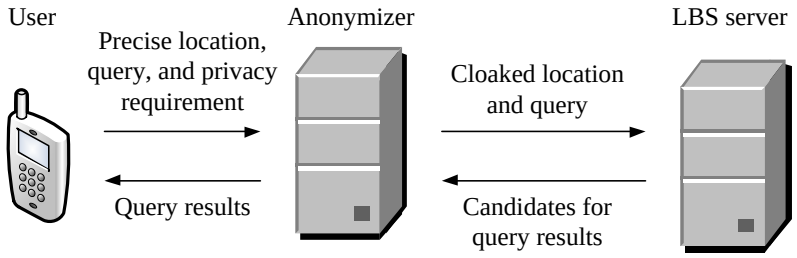

Figure 1. The Centralized Structure for Location Privacy Protection

3.2. Mobile Users

As a practical matter, the proposed location privacy protection method limits the activities of mobile users to road networks. The road networks are represented as a connected graph $G = (V, E)$, where V is the set of vertices representing the endpoints of the road segments and intersections of the adjacent road segments, while E is the set of edges representing the road segments. Each segment $seg(sid, s, e) \in E$ represents one edge in road networks, where *sid* is the identification (ID) of the segment, s and e respectively denote the start and end points of the segment.

Definition 1 (Mobile User). Let u be the mobile user who issues a request, then U denotes the set of such users. Using $l(x, y, sid)$ to represents user's location, where, x and y are coordinates of the location and sid denotes the ID of the segment which the location belongs to. Now, we can use L denotes the set of such locations. Using function $POS(u, t) \in L$ to give the location of user u at a given time t and function $SEG(l) \in E$ to obtain the segment which the location l belongs to.

3.3. Query Dependency

The query dependency and its effects on location privacy are to be discussed in this paper. Definitions concerned are given as follows.

Definition 2 (Query). We use q to denote one kind of query supported by LBS, then Q denotes the set of such queries. A LBS request issued by user u is represented by $req(u, l, t, q)$. Using function $QUE(req) \in Q$ to obtain the query of a request req.

According to the above analyses, if the attacker has fully collected user's request history and from which to get the dependency of queries as background knowledge, his destructive effect will be dramatically improved. Thus, the query dependency is related to user's request history especially two successive requests.

Definition 3 (Request History). For user u, his request history is denoted by a sequence of successive requests, *i.e.*, $H_u = < req_{u1}, \ldots, req_{ui}, \ldots, req_{un} >$. Using $H_u(i) = req_{ui}$, $req_{ui} = req\ (u, l_i, t_i, q_i)$ to represent the ith request in the request history of user u, where $t_i \le t_{i+1}$ $(1 \le i \le n\text{-}1)$. The total number of requests in H_u is denoted by $len(H_u)$.

Definition 4 (Successive Requests). Using $Suc_u(i, j) = \{<H_u(k), H_u(k+1)> \mid QUE(H_u(k)) = q_i \wedge QUE(H_u(k+1)) = q_j,\ 1 \le k \le len(H_u)\text{-}1\}$ to represent the set of successive query pairs q_i and q_j contained in the request history of user u.

Definition 5 (Query Dependency). It is assumed that current query issued by a user is only affected by his last query and the continuous queries are regarded as a Markov process [1,25], then the query dependency of the current query q_j issued by user u with respect to his last query q_i is denoted by $dep_u(q_i, q_j)$ and can be obtained by the conditional probability $p_u(q_j|q_i)$ as

$$dep_u(q_i, q_j) = p_u(q_j \mid q_i) = \frac{|Suc_u(i, j)| + \lambda}{\sum_{q_k \in Q} |Suc_u(i, k)| + |Q| \cdot \lambda} \quad (1)$$

where, λ is the smoothing parameter whose value is usually set to be 1 [1,25].

3.4. Privacy Requirement

The anonymous method proposed in this paper generalizes user's accurate location to an anonymous set RS which is composed of several adjacent road segments and satisfies user's personalized privacy requirement in the following two aspects.

(1) The number of mobile users in anonymous set (here denoted by $RS.UN$): It is required to make the user cannot be distinguished with at least other k-1 users in the anonymous set. This part of requirement stems from the classical k-anonymous method. It is also the most commonly used and basic requirement in location privacy protection methods.

(2) The number of road segments in anonymous set (here denoted by $RS.SN$): It is required that the number of road segments in anonymous set should not be less than the threshold value defined by the user. If the value of this number is too small, only one segment for example, then even there are many mobile users in the anonymous set, the difficulty of being attacked by the attacker will still be significantly decreased.

According to the analyses illustrated above, we can accordingly define user's privacy requirement as follow.

Definition 6 (Privacy Requirement). For user *u* who requests anonymity, his privacy requirement is represented by $PR_u(UN, SN)$, where, *UN* and *SN* are respectively user-defined lower limits of mobile users number and road segments number.

3.5. Problem Definition

In this system, some reasonable assumptions are made to the attacker as follows: (1) he masters the map and road segments information; (2) he masters mobile users' location information but not their IDs; (3) he can obtain the number of mobile users over any road segment; (4) he can get users' request history. In addition, we assume that the anonymizer is believable to users.

Therefore, the problem to be solved in this work falls into: according to the personalized privacy requirement as well as considering the influence of query-dependency, we construct an anonymous set *RS* for the user. The constructed anonymous set consists of several adjacent road segments and hides user's actual location in the set.

4. Query-Dependency-Based Location Privacy Protection

4.1. Privacy Measurement

Now, one can get that, the number of mobile users in similar behavior contained in anonymous set directly affects the privacy degree of the set. Hence, in the procedure of measuring the privacy degree of an anonymous set, besides numbers of both mobile users and road segments contained in the set, we should also consider the number of mobile users in similar behavior contained in the set. And the value of such number should be directly used to measure the region privacy degree.

Definition 7 (Region Privacy Degree). At time index *t*, the privacy degree of region $reg \subset E$ (denoted by PRM(*reg, t*)) is determined by the number of mobile users who are contained in the region and in similar behavior at the same time index (denoted by NumUSB(*reg, t*)). So, the value of region privacy degree can be expressed as

$$\mathrm{PRM}(reg,t) = \mathrm{NumUSB}(reg,t) \tag{2}$$

where, NumUSB(*reg, t*) is calculated in detail by

$$\begin{aligned} &\mathrm{NumUSB}(reg,t) = Card\{u' \mid u' \in U \wedge u' \neq u, \mathrm{SEG}(\mathrm{POS}(u',t)) \in reg, \\ &dep_{u'}(q_{u'i}, q_{u'j}) \geq dep_u(q_{ui}, q_{uj}), \\ &q_{u'i} = \mathrm{QUE}(H_{u'}(n)), q_{uj} = q_{u'j} = \mathrm{QUE}(H_u(n)), q_{ui} = \mathrm{QUE}(H_u(n-1))\} \end{aligned} \quad . \tag{3}$$

4.2. Query-Dependency-Based Anonymous Set Constructing

The main idea of constructing anonymous set is to elect proper adjacent road segments in terms of the privacy requirement defined by user. The key problem to be solved here is how to determine appropriate road segments and a way of judging while choosing is to be applied. The key point of this way is choosing a segment from candidates each time to keep current anonymous set be of optimum region privacy degree, and once choosing a segment then taking judgement on current anonymous set whether it meets the privacy requirement or not.

The anonymous set is composed of several adjacent road segments including the one where user issued cloaking request sits. Naturally, the segment where user sits should be firstly chosen as the first segment to constructing anonymous set *RS*. Taking judgement on current anonymous set *RS* whether it meets the privacy

requirement defined by user or not once electing a segment. If the privacy requirement is met, the current anonymous set is referred as the final to be returned. If not, finding adjacent segments of all segments contained in current anonymous set and putting the most appropriate one into current anonymous set. As mentioned before, here, the most appropriate adjacent segment is exactly the one which makes the new constructed anonymous set containing it be the optimum with respect to region privacy degree. The strategy applied here is taking query dependency as an important investigating factor for electing the segment with more mobile users who are in similar behavior to be an anonymous region. Following the definition of region privacy degree, the anonymous region is of higher privacy degree when it contains more mobile users in similar behavior. Obviously, the elected segment here should maximize the region privacy degree, PRM_{RS}, of current anonymous set. It is equivalent to choose the segment with maximum PRM_{seg} to be added into current anonymous set as adding a segment every time. Following the principle given above, adding a segment into current anonymous set every time until the set meets the privacy requirement.

Algorithm 1 gives the pseudocodes of electing the optimum road segment. The input parameters of this algorithm are adjacent road segments set *ES* and current time *t*. Firstly, algorithm 1 initializes related parameters (line 1); then computes the region privacy degree of each adjacent road segment and records the segment with greatest region privacy degree as current optimum segment (lines 2 to 6); finally returns the optimum road segment (line 7).

Algorithm 1. Optimum road segment electing (OPTRSE)

Inputs: adjacent road segments set *ES* and current time *t*.

Output: the optimum road segment *BESTE*.

1: *BESTE* = *Φ*; *MPD* = 0;
2: for each *seg* in *ES*
3: assigning *seg* to *e*;
4: *PD* = PRM(*e*, *t*);
5: using *MPD* to record current maximum value of *PD* and assigning corresponding *e* to *BESTE*
6: end for
7: return *BESTE*

Algorithm 2 gives the pseudocodes of query-dependency-based anonymous set constructing algorithm. The inputs of this algorithm are a LBS query $req(u, l, t, q)$ and privacy requirement PR_u respectively issued and defined by user *u*. In privacy requirement PR_u, according to his own actual situation, the user defines lower limits of mobile users number $PR_u.UN$ and segments number $PR_u.SN$ to be contained in the anonymous set. Algorithm 2 firstly puts the segment where user issued request into the anonymous set *RS* (line 2). If *RS* does not meet $PR_u.UN$ and $PR_u.SN$ in user privacy requirement this moment, the loop executes steps as follows until *RS* meets $PR_u.UN$ and $PR_u.SN$ (line 3): (1) putting adjacent road segments of all segments contained in current *RS* into set *R* (line 4); (2) calculating the region privacy degree of each segment in *R* and picking up the segment with greatest region privacy degree to be jointed in anonymous set *RS* (lines 5 and 6). The set finally returned is exactly the constructed anonymous set (line 9).

Algorithm 2. Query-dependency-based anonymous set constructing (QDBASC).

Inputs: A LBS query, $req(u, l, t, q)$, of user *u* and his privacy requirement PR_u.

Output: the anonymous set *RS*.

1: *RS* = *Φ*;
2: *RS* = SEG(POS(*req.u*));
3: while NumUser(*RS*) < $PR_u.UN$ or NumSeg(*RS*) < $PR_u.SN$
4: *R* = FindEdges(*RS*);

5: *BestEdge* = OPTRSE(*R*, *req.t*);
6: *RS* = *RS* ∪ *BestEdge*;
7: *R* = *Φ*;
8: end while
9: return *RS*

The anonymization method proposed above mainly takes the influence of user behavior on anonymous space privacy degree into account. User historical query dependency is referred to be a key factor in constructing anonymous space in the proposed method. Consequently, the privacy degree of anonymous set is principally determined by region privacy degree which is only related to the number of users with similar behavior contained in the anonymous set, and does not directly affect service quality. It should be noted here that, the anonymous set constructing with the proposed method is a process of continuously increasing segments number which directly determines the size of candidate set and finally affects service quality. Therefore, in detailed implementing procedure, an upper bound should be set for the segments number.

5. Experiments and Analyses

Experimental hardware environment for the algorithm proposed in this paper is 3.2 GHz quad-core CPU with memory of 4GB. The operation system platform is Microsoft Windows 7 Professional.

5.1. Experimental Data Set and Parameter Settings

Using German city of Oldenburg road networks data [26,27] to be the experimental data which totally includes 6105 roads and 7035 vertexes as shown in Figure 2. Further using a network-based mobile object generator designed by Brinkhoff [28] to generate 10000 mobile users in uniform distribution. In order to assess anonymous performance of the proposed algorithm, the parameter of users number being ranged from 5000 to 25000 is investigated. Meanwhile, generating 10000 points of interest in 6 types (such as hospital, bar, shopping mall, school, club, and restaurant) distribute over road networks in uniform. To investigate user's behavior, 10 times historical requests with Markov process for each user are generated through modifying the data generator. In addition, in experiments, 1000 mobile users are set to issue anonymous requests.

Figure 2. The Road Map of Oldenburg City, Germany

The privacy requirement of an anonymous request message send by a user contains only two parameters, *i.e.*, *PRu.UN* and *PRu.SN*. Considering the road segments number requested in privacy requirement cannot be increased without limitation in real-life privacy protection, the maximum value of the number is needed and expressed as $PRu.SN_{\max}$ in experiments. All experimental parameter settings are given in Table 1.

Table 1. Parameter Settings

Parameter	Default value	Evaluation range
Users number	10000	5000 -- 25000
$PR_u.UN$	25	15 -- 35
$PR_u.SN$	6	3 -- 15
$PR_u.SN_{\max}$	20	
Historical requests times	10	
Points of interest	10000	
Number of users issued anonymous requests	1000	

5.2. Performance Evaluating Criteria for the Algorithm

In the following experiments, the proposed algorithm is evaluated from 4 aspects including anonymous success rate, average anonymizing time, relative anonymous degree, and relative spatial granularity.

(1) Anonymous success rate. It is the ratio between the number of messages cloaked successfully by the algorithm and the number of anonymous request messages issued by all mobile users [24]. This criterion indicates the responsiveness capability of the location cloaking algorithm with respect to the anonymous request. Obviously, the higher anonymous success rate, the better the cloaking algorithm.

(2) Average anonymizing time. It is the time taken to cloak user's actual location and executed by location anonymizer, in statistical mean sense. This criterion can be adopted to evaluate the efficiency of the anonymizing algorithm. Shorter average anonymizing time means higher efficiency of the algorithm when it is used for performing anonymity.

(3) Relative anonymous degree. It refers to the ratio between the number of users contained in the anonymous set constructed after cloaking and users number requested in location privacy requirement, *i.e.*, $PR_u.UN$, which can be expressed as [24]

$$relative\ anonymous\ degree = \frac{RS.UN}{PR_u.UN}. \quad (4)$$

Clearly, under condition of successfully cloaking, the value of the relative anonymous degree is greater than or equal to 1. Usually, with increasing of the value of relative anonymous degree, the efficiency of cloaking is better.

(4) Relative spatial granularity. It refers to the ratio between the tolerable maximum number of road segments and that of segments contained in anonymous set constructed after cloaking. It is calculated by [24]

$$relative\ spatial\ granularity = \frac{PR_u.SN_{\max}}{RS.SN} \quad (5)$$

where, the value of parameter $PRu.SN_{\max}$ is limited by the anonymizer in terms of user anonymous requirement. Greater relative spatial granularity means smaller

anonymous space under condition of meeting privacy requirement and it is closer to the optimum solution. Obviously, the value of relative spatial granularity is greater means better.

5.3. Results and Analyses

(1) Anonymous success rate. Figure 3 shows the change trend of anonymous success rate with respect to the number of mobile users over road networks, mobile users number requested in privacy requirement, and road segments number requested in privacy requirement. From Figure 3(a), one gets that the more mobile users over road networks, the more beneficial to perform anonymizing and the higher anonymous success rate. It is very easy to understand that since the number of users distributed on each road segment increases as the number of users over road networks becomes greater, which leads to the privacy requirement is more easily to be satisfied and the anonymous success rate be improved. Conversely, when the number of users over road networks decrease to a certain degree, users distributed on each road segment becomes little. Under this condition, during the process of performing anonymous request, it is difficult to ensure that the number of users contained in anonymous set simultaneously meets the privacy requirement as the road segments number reaches the tolerable maximum value, which causes failure anonymization.

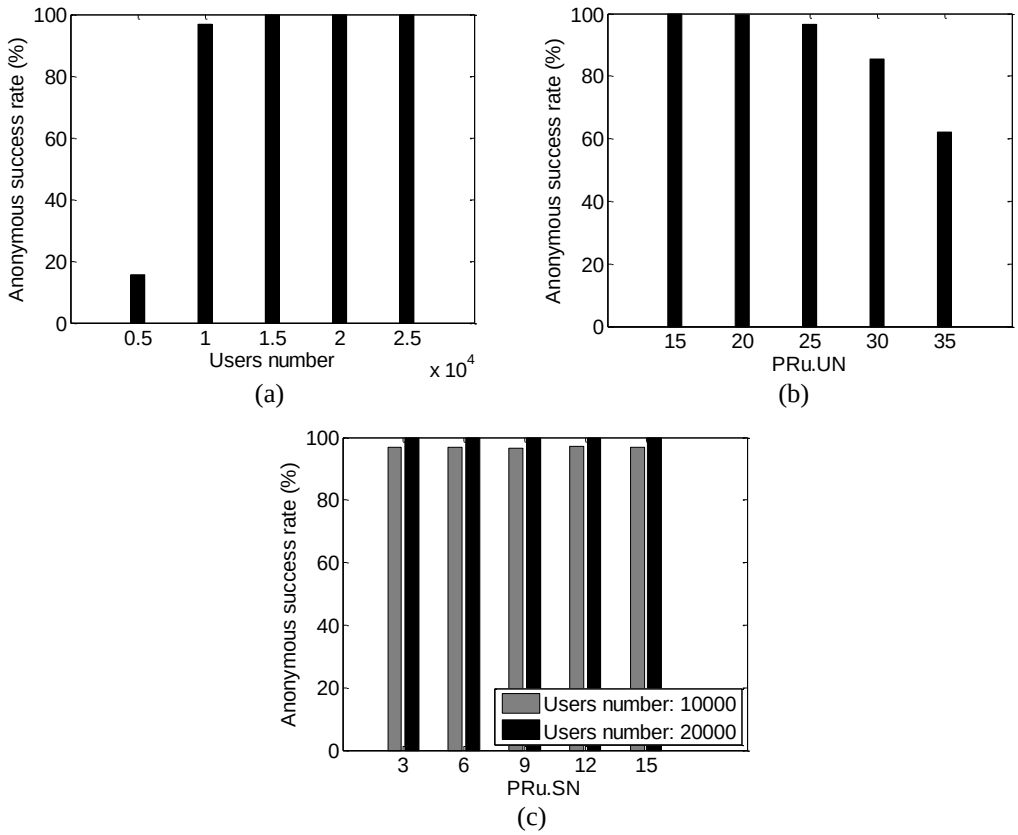

Figure 3. Anonymous Success Rate with respect To the Number of Mobile users Over Road Networks (a), *PRu.UN* (b), and *PRu.SN* (c)

Figure 3(b) tells us that the anonymous success rate declines as increasing of users number requested in privacy requirement, $PR_u.UN$. In the case of the number of users over road networks is 10000, parameters $PR_u.UN$, $PR_u.SN$, and $PR_u.SN_{\max}$

are respectively set to be 35, 6, and 20, the anonymous success rate only reaches to around 60%. It shows that users number requested in privacy requirement is inappropriately set to be too large. On the other hand, from Figure 3(c), one gets that the number of users over road networks is either relatively large or relatively small, the changing of road segments number, $PR_u.SN$, requested in privacy requirement does not bring significant changes of anonymous success rate. Therefore, from reducing computational cost in performing anonymity point of view, according to anonymous success rate, the parameter $PR_u.SN$ should not be set to be too large.

(2) Average anonymizing time. According to Figure 4, it is apparent that under condition of mobile users over road networks are fixed at a certain number, with the increase of users number $PR_u.UN$ or segments number $PR_u.SN$ requested in privacy requirement, the average anonymizing time increases. On the other hand, Figure 4 shows that in the case of users number $PR_u.UN$ requested in privacy requirement is fixed at a certain value, the larger number of mobile users over road networks the shorter average anonymizing time as shown in Figure 4(a) ($PR_u.UN$ = 25, $PR_u.SN$ = 6), under condition of road segments number requested in privacy requirement is relatively small (*e.g.* $PR_u.SN \leq 9$). The reason is that, when the number of mobile users over road networks becomes greater, users number requested in privacy requirement is more easily to meet the privacy requirement and hence saving computational time for judging as choosing the same segments. And under condition of road segments number requested in privacy requirement is relatively larger (*e.g.* $PR_u.SN \geq 12$), with increasing of the number of mobile users over road networks, the average anonymizing time becomes longer as shown in Figure 4(b). This is because in given conditions that users number requested in privacy requirement is fixed while the number of mobile users over road networks is relatively large, the algorithm should take more time to compute region privacy degree for electing the optimum segments, which is exactly caused by a relative large road segments number requested in privacy requirement.

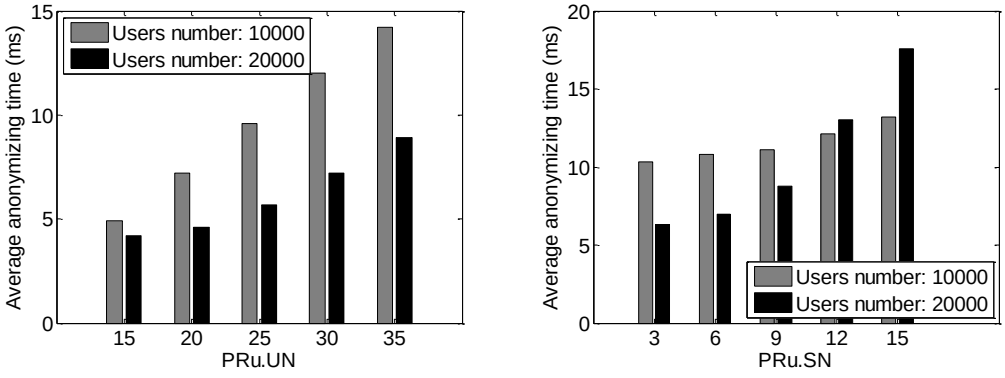

Figure 4. Average Anonymizing Time with respect to *PRu.UN* (a) and *PRu.SN* (b)

(3) Relative anonymous degree. Figure 5 gives the relationship between relative anonymous degree and users number requested in privacy requirement, $PR_u.UN$. Results shown in Figure 5 reveal that with increasing of $PR_u.UN$, the relative anonymous degree decreases lightly but still is greater than or equal to 1 in the case of successfully cloaking. On the contrary, it is less than 1 if failing to cloak that it is under condition of the number of mobile users over road networks is 10000 and $PR_u.UN$ is set to be 35 shown in Figure 5. The reason is, under this condition, cases

of failing to cloak are too many and counted. At the same time, we can see that, with increasing of the number of mobile users over road networks, anonymous effectiveness of the algorithm improves.

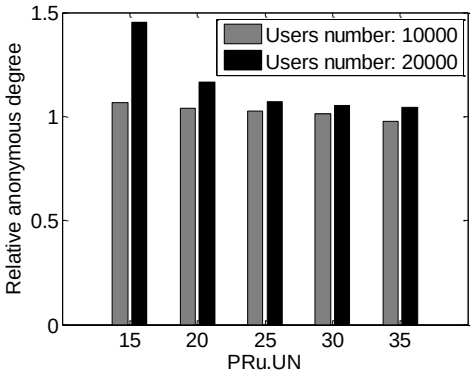

Figure 5. Relative Anonymous Degree with Respect to *PRu.UN*

(4) Relative spatial granularity. Figure 6 gives the relationship between relative spatial granularity and users number requested in privacy requirement, $PR_u.UN$. It is shown that, with increasing of $PR_u.UN$, the relative spatial granularity also decreases relative severely. It is because the increasing of $PR_u.UN$ directly results in the increasing of road segments contained in anonymous set and hence reduces the relative spatial granularity. Figure 6 still shows that, under condition of setting less users number requested in privacy requirement, the algorithm gets relatively high relative spatial granularity, which means properly low users number requested in privacy requirement can bring to relatively optimum anonymous space. On the other hand, the increasing of the number of mobile users over road networks can also give rise to larger relative spatial granularity.

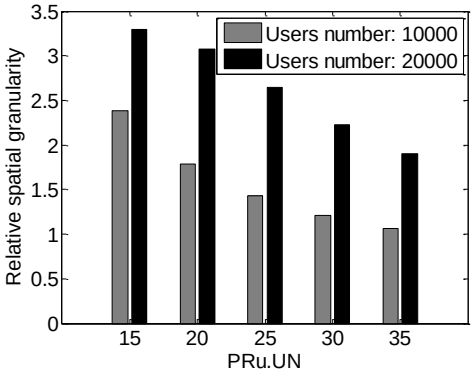

Figure 6. Relative Spatial Granularity with respect to *PRu.UN*

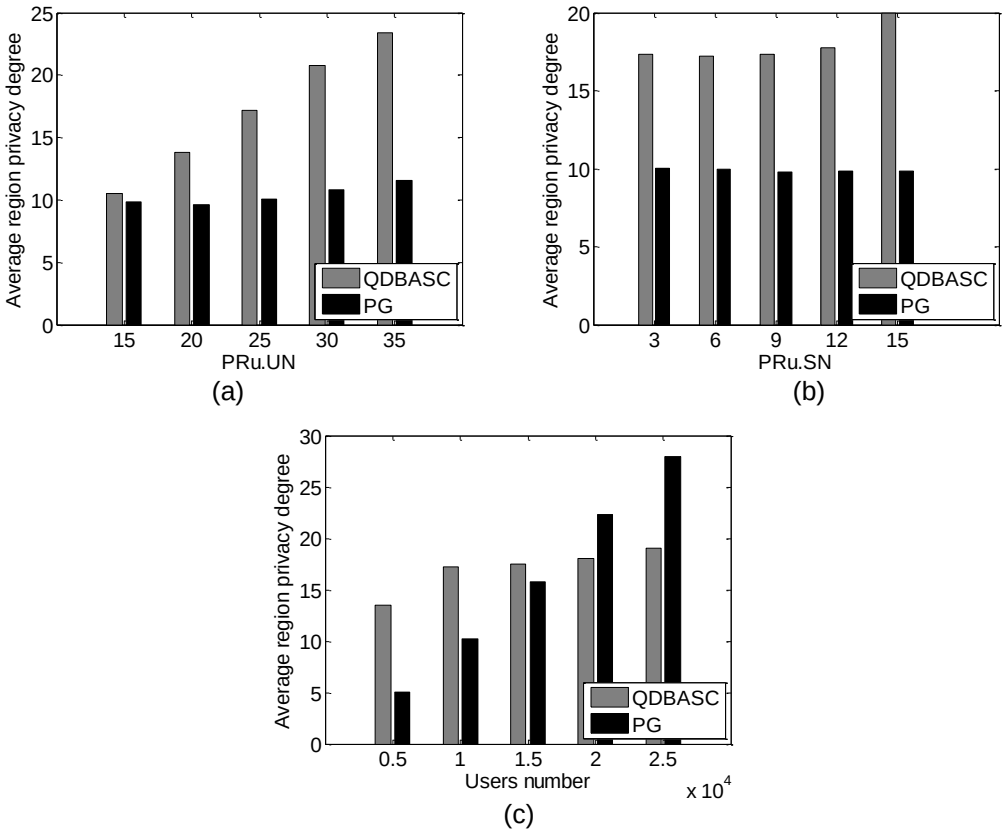

Figure 7. Average Region Privacy Degree with Respect to *PRu.UN* (a), *PRu.SN* (b), and the Number of Mobile Users over Road Networks (c)

Moreover, another anonymous set constructing algorithm, *i.e.* PG algorithm, for road networks LBS privacy protection and proposed by Chow [22] is chosen here to be conducted comparative analyses on performance. The relationships between average region privacy degree and users number requested in privacy requirement, $PR_u.UN$, road segments number requested in privacy requirement, $PR_u.SN$, and the number of mobile users over road networks are given in Figure 7. From Figure 7(a) and (b), in the case of users number requested in privacy requirement $PR_u.UN$ and road segments number requested in privacy requirement $PR_u.SN$ change, the privacy degree of anonymous region constructed by QDBASC algorithm proposed in this paper is higher than that by PG algorithm. The influence of user's behavior on constructing anonymous set is adequately considered in QDBASC algorithm but not in PG algorithm, which makes the anonymous region constructed by QDBASC algorithm gives higher privacy degree. From Figure 7(c), for PG algorithm, with increasing of the number of mobile users over road networks, the average region privacy degree accordingly improves. However, when this number changes, the average region privacy degree with QDBASC algorithm changes slightly, which indicates this algorithm is relatively stable with respect to the number of mobile users over road networks.

6. Conclusions

In this paper, we have presented a location privacy protection approach with query dependency for road networks. Unlike traditional location privacy protection strategies, the proposed algorithm investigated the influence of query dependency on anonymization effect. Current query issued by a user is assumed to be only affected

by his last query and the continuous queries are regarded as a Markov process which exactly conforms to the random characteristic in real world. At the same time, a definition of query-dependency-based anonymous region privacy degree is given and referred as the criterion for constructing anonymous set which keeps personalized privacy requirement satisfied. A great deal of experiments with real and simulated data verify the feasibility and validity of the proposed method. However, we have to point out that one future challenge of constructing anonymous set for road networks with the proposed method is furtherly investigating the influence of time on user behavior pattern which is exactly not investigated in this work.

Acknowledgment

This work is supported in part by National Natural Science Foundation of China (61373015, 61300052, 41301047), A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and The Important National Science and Technology Specific Project (BA2013049).

References

[1] X. Chen and J. Pang, "Protecting Query Privacy in Location-Based Services", GeoInformatica, vol. 18, no. 1, **(2014)**, pp. 95-133.
[2] A. Y. Zhou, B. Yang, C. Q. Jin and Q. Ma, "Location-based services: architecture and progress", Chinese Journal of Computers, vol. 34, no. 7, **(2011)**, pp. 1155-1171.
[3] G. S. Kang, X. Ju, Z. Chen and X. Hu, "Privacy protection for users of location-based services", *IEEE* Wireless Commun., vol. 19, no. 1, **(2012)**, pp. 30-39.
[4] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking", Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, San Francisco, CA, USA, **(2003)** May 31–42.
[5] M. F. Mokbel, C. Y. Chow and W. G. Aref, "The new casper: query processing for location services without compromising privacy", Proceedings of the 32nd International Conference on Very Large Data Bases, Seoul, Korea, **(2006)** August 763–774.
[6] P. Kalnis, G. Ghinita, K. Mouratidis and D. Papadias, "Preventing location-based identity inference in anonymous spatial queries", IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 12, **(2007)**, pp. 1719–1733.
[7] S. Mascetti C. Bettini, X. Sean Wang, D. Freni and S. Jajodia, "ProvidentHider: an algorithm to preserve historical k-anonymity in LBS", Proceedings of the 10th IEEE International Conference on Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan, **(2009)** May 172–181.
[8] C. Zhang and Y. Huang, "Cloaking locations for anonymous location based services: a hybrid approach", GeoInformatica, vol. 13, no. 2, **(2009)**, pp. 159–182.
[9] M. T. Tran, I. Echizen and A. D. Duong, "Binomial-mix-based location anonymizer system with global dummy generation to preserve user location privacy in location-based services", Proceedings of the 5th International Conference on Availability, Reliability and Security, Krakow, Poland, **(2010)** February 580-585.
[10] 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, no. 8, **(2012)**, pp. 1506-1519.
[11] X. Pan and X. Meng, "Preserving location privacy without exact locations in mobile services", Frontiers of Computer Science, vol. 7, no. 3, **(2013)**, pp. 317-340.
[12] V. A. Kachore, J. Lakshmi and S. K. Nandy, "Location obfuscation for location data privacy", Proceedings of 2015 IEEE World Congress on Services, New York City, NY, USA, **(2015)** June 213-220.
[13] H. Kido, Y. Yanagisaw and T. Satoh, "An anonymous communication technique using dummies for location-based services", Proceedings of IEEE International Conference on Pervasive Services, Santorini, Breece, **(2005)** July 88–97.
[14] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy", Proceedings of the 3rd International Conference on Pervasive Computing, Munich, Germany, **(2005)** May 152-170.
[15] C. A. Ardagna, M. Cremonini, S. De Capitani Di Vimercati and P. Samarat, "An obfuscation-based approach for protecting location privacy", IEEE Transactions on Dependable and Secure Computing, vol. 8, no. 1, **(2011)**, pp. 13-27.

[16] G. Ghinita, P. Kalnis, A. Khoshqozaran, C. Shahabi and K. L. Tan, "Private queries in location based services: Anonymizers are not necessary", Proceedings of ACM SIGMOD Int. Conf. Management of Data, Vancouver, BC, Canada, **(2008)** May 121-132.
[17] L. Sweeeney, "Achieving k-anonymity privacy protection using generalization and suppression", International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, vol. 10, no. 5, **(2002)**, pp. 571-588.
[18] M. Wernke, P. Skvortsov, F. Dürr and K. Rothermel, "A classification of location privacy attacks and approaches", Personal and Ubiquitous Computing, vol. 18, no. 1, **(2014)**, pp. 163-175.
[19] M. Xue, P. Kalnis and H. K. Pung, "Location diversity: enhanced privacy protection in location based services", Proceedings of the 4th International Symposium on Location and Context Awareness, Tokyo, Japan, **(2009)** May 70-87.
[20] M. L. Damiani, C. Silvestri and E. Bertino, "Fine-grained cloaking of sensitive positions in location-sharing applications", IEEE Pervasive Computing, vol. 10, no. 4, **(2011)**, pp. 64-72.
[21] B. Palanisamy and L. Liu, "MobiMix: protecting location privacy with mix-zones over road networks", Proceedings of the 27th IEEE International Conference on Data Engineering, Hannover, Germany, **(2011)** April 494-505.
[22] C. Y. Chow, M. F. Mokbel, J. Bao and X. Liu, "Query-aware location anonymization for road networks", GeoInformatica, vol. 15, no. 3, **(2011)**, pp. 571-607.
[23] E. Yigitoglu, M. L. Damiani, O. Abul and C. Silvestri, "Privacy-preserving sharing of sensitive semantic locations under road-network constraints", Proceedings of the 13th IEEE International Conference on Mobile Data Management, Bengaluru, India, **(2012)** July 186-195.
[24] X. Pan, Z. Xiao and X. Meng, "Survey of location privacy-preserving", Journal Frontiers of Computer Science and Technology, vol. 1, no. 3, **(2007)**, pp. 268-281.
[25] C. D. Manning and H. Schuütze, "Foundations of Statistical Natural Language Processing", MIT Press, Cambridge, Massachusetts, **(1999)**.
[26] S. Chen, C. S. Jensen and D. Lin, "A benchmark for evaluating moving object indexes", Proceedings of the VLDB Endowment, vol. 1, no. 2, **(2008)**, pp. 1574-1585.
[27] http://www.comp.nus.edu.sg/~spade/benchmark.
[28] T. Brinkhoff, "A framework for generating network based moving objects", GeoInfomatica, vol. 6, no. 2, **(2002)**, pp. 153-180.

Authors

Hui Chen, she received her bachelor's degree of Computer Science and Technology in Anhui Polytechnic University, Wuhu (2003), her master's degree of Computer Software and Theory in Hefei University of Technology, Hefei (2006), China. She is a lecture in the School of Electronic & Information Engineering at Nanjing University of Information Science & Technology, Nanjing, China. She is studying her doctor's degree of Computer Application Technology in Nanjing University of Aeronautics and Astronautics, Nanjing, China. Her current research interests include Moving Objects Databases and Privacy Protection.

Xiaolin Qin, he is a professor in the College of Computer Science and Technology at Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests focus on Data Management, Internet of Things, Data Security and Privacy Protection, Big Data Management and Analysis, *etc*.

