Query-Privacy-Aware Location Cloaking for Mobile P2P System

Min Li[1, 2], Zhiguang Qin[1] and Cong Wang[1]

[1]*School of Computer Science & Engineering, University of Electronic and Technology of China, Chengdu, China*
[2] *College Of Computer Science, Sichuan Normal University, Chengdu, China*
Email: lm_turnip@126.com, qinzg@uestc.edu.cn, wongcong@gmail.com

Abstract

Location-based Services (LBS) have brought the potential threats of mobile users' sensitive personal information. Privacy-aware in LBS is including ***location privacy*** *and* ***query privacy****. Unfortunately, existing privacy protection algorithms rarely pay attention to both of them. This paper proposes a novel privacy protection method which combines* ***K****-anonymity and* ***L****-diversity to protect both location privacy and query privacy. Two effective query-privacy-aware methods are introduced into the cloaking algorithm. One is the* ***history sharing scheme*** *which confuses history queries within tolerance time. Another is the* ***batch query scheme*** *which confuses real queries presented by the peers. Our technique is suitable for P2P mobile networks, which can effectively eliminate the bottleneck of system brought by the anonymizer. In addition, we develop an* ***imprecise location scheme*** *to prevent the* ***inference attack*** *of few malicious peers. The experiments show that the proposed algorithms are effective to protect users' location privacy and query privacy in the mobile P2P system.*

Keywords: *Location-based Services, location privacy-preserving, query privacy-preserving,* ***K****-anonymity,* ***L****-diversity, homogeneity attack, inference attack, P2P networks*

1. Introduction

With the development of mobile equipment and location technique, more and more mobile users chose to access LBS for obtaining real-time query services, which can not be reached by off-line map. To get higher quality of service (QoS), mobile users have to provide more accurate location, which might bring privacy leak threat. This is a trade-off between privacy-preserving and QoS in LBS.

Privacy protection in LBS mainly focuses on the following two issues: *location privacy* and *query privacy*. If the adversaries collect a user's privacy information, it is very easy for them to deduce the user's some sensitive property. For example, if a user issues queries in a school, the adversaries can associate the querier with a student or a teacher. On the other hand, if a user issues queries for a hospital may lead the adversaries to infer her medical conditions. Recently, the *K*-anonymity Spatial Region (*K*-ASR) [1-7] has been widely used to protect location privacy through blurring the querier's exact location into a spatial region, while only considering *K*-ASR can not effectively work.

Existing privacy protection architecture in LBS can be classified into two main categories: three-tier centralized architecture and two-tier decentralized architecture. In three-tier centralized architecture, the anonymizer (*i.e.*, a third party) becomes the bottleneck and the attack target of the whole system. Therefore, we can take it into consider to apply the decentralized architecture in mobile peer-to-peer (P2P) system.

Several problems can be proposed about decentralized P2P architecture including: *homogeneity attack* and *inference attack*. In this paper, we propose a novel privacy protection

method which combines anonymity and diversity, and in which two effective schemes are introduced to enhance the query diversity and prevent *homogeneity attack*. One is the *history sharing scheme* which allows query initiator to confuse history queries cached by it and its neighbour peers within tolerance time. But it can take extra overhead for the Location Services Provider (LSP) to process the dummy history queries. To improve the rate of effective inquires, *batch query scheme* is proposed to confuse with real queries presented by the peers. In addition, because that neither neighbor peers nor the LSP are trusted, an *imprecise location scheme* can be proposed to blur exact location of peers with a large blurring region. So, *inference attack* of the malicious peers can also be effectively coped with.

2. Related work

Spatial cloaking based on *K*-anonymity [8, 9] is the main location privacy-preserving technique, termed *K*-ASR [1, 2], which makes the adversaries to identify query users with probability *1/K*. In three-tier centralized framework, some typical cloaking methods have been studied, such as Casper [3] and Interval [2] are based on quadtree, but both of them have not reciprocity. To deal with the shortage, Hilbert-based cloaking [7] is proposed. But the trusted third-party can know users' exact locations, thus, users' location privacy may still be compromised.

In contrast, in two-tier decentralized framework, no one else can know any user's exact locations. In this framework, there are generally two kinds of approaches: PIR-based location preserving [18-19] and P2P-based location preserving [11-16]. Because of the PIR performance has not been get bigger promotion, thus PIR-based privacy framework has not able to work effectively.P2P Cloak [13] is first discussed in P2P-based framewok, in which cloaked area is formed through the cooperation of the peers, but it is vulnerable to "the center of cloaked area" privacy attack. Ghinital, G. proposed PRIVE [12] and MobiHide [11], both of them pay close attention to the index structure. PRIVE forms clustering and hierarchy index structure using B^+ tree. Besides, it is effective to improve the retrieval efficiency, but the upper nodes are easy to be the bottleneck of the system. To overcome the problem, MobiHide organizes a circular storage structure using the Chord P2P system. Meanwhile it can bring a new problem that the cluster heads are needed to keep complex index structure, so it is not applicable for the P2P system which has constrained storage resources and limited computing power. The works [14, 15] put forward information sharing mechanism to reduce the communication cost of P2P system. In above works, all neighbour peers cooperate through providing their exact location, without taking into account the malicious peers. Hashem, T. considers the problem and proposes "local confusion area method" to realize confusion and anonymous. The ref [16] proposes that sending queries with an anonymous chain, but it is just to hide query users' identification, and not to hide query users' exact location.

The above works mainly consider space cloaking approaches, without considering the query diversity. Aniket Pingley [17] proposes the dummy technical to confuse the history queries whose frequency is highest, but this can take "position correlation attack". Moreover, due to without anonymizer, it is a difficult problem for P2P system to get the frequency of queries. Obviously, an effective method is needed to realize query privacy-preserving.

In this paper, we propose a query-preserving-aware cloaking algorithm, which adopts the ref [15] as the generation of *K*-ASR through the peers searching method, meanwhile realizing effective query *L*-diversity through "confusing query scheme".

3. System Model

Figure 1 describes the P2P LBS system that includes two entities: 1) **mobile user.** Who can obtain their own coordinates (x, y) through the positioning device, *e.g.*, GPS, and be willing to cooperate with others. 2) **LBS server** termed as **LSP**, who can process the query services. When the mobile users issue queries, they first present cooperation request to their neighbor peers to form anonymous spatial region and confusing query list, which satisfies their privacy profile. Then, a requester agent can be elected to send the queries to LSP along with the (K, L)-ASR. Finally, the LSP processes the queries and broadcasts the encrypted result back to the initiators.

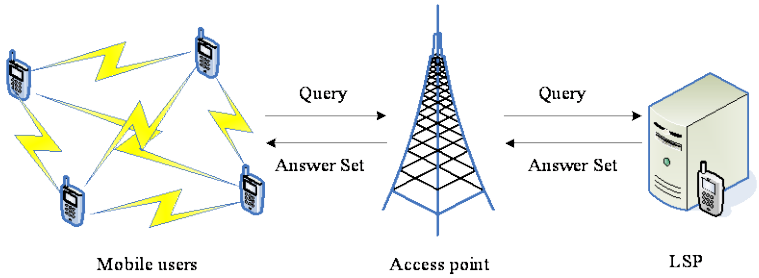


Figure 1. System Architecture

Attack Model. Firstly, we assume most of the mobile users are security as well as the base station. However the location services provider (LSP) and the communication channel are not our trusts. Moreover, there are very few malicious mobile users who can obtain neighbour peers' position through initiating request. However, these malicious peers' power is limited, so they can not obtain the coordinate of all nodes.

Privacy Model. It is important for location privacy-preserving to realize spatial anonymity. Meanwhile it is important for query privacy-preserving to realize the diversity of the service attribute. A mobile user can express her privacy profile by specifying the value of (K, L), where K indicates the spatial anonymity level and L indicates the query diversity level.

4. Definitions

Firstly, we definition service attributes of queries as $C = \{c_1, c_2 \ldots c_m\}$, such as hospital, school, restaurant *etc*. To realize the diversity of the queries, a query set should include a certain number of service attributes. But in a query set, some service attributes may have a dominant, so the adversaries associate the query with the dominant service attributes. Therefore, the number of service attributes is not sufficient to quantify its diversity. Therefore, the entropy of Shannon's information theory is introduced to measure about the diversity of a query set, and shown by definition 1.

Definition 1 Query *L*-diversity. Given a query set $Q = \{q_1, q_2 \ldots q_w\}$, let $C = \{c_1, c_2 \ldots c_m\}$ be a set of service attribute associated with Q, where w indicates the number of total query request, m indicates the number of total service attribute. Let the attribute set of a query be $Att(Q) \in C$, n_i $(1 \leqslant i \leqslant m)$ be the number of the ith service attribute in $Att(Q)$, and

$N = \sum_{i=1}^{m} n_i$ .We define the entropy of Q as $E(Q) = -\sum_{i=1}^{m} \frac{n_i}{N} \log \frac{n_i}{N}$, and define the diversity of C as $D(Q) = 2^{E(Q)}$.

Property 1 (*K, L*) privacy. An anonymous spatial region *A* along with the query set *Q* are said satisfy (*K*, *L*) privacy, if the probability of distinguishing everyone in *A* does not exceed 1/*K*, and the query diversity of *Q* meets the inequality $D(Q) \geqslant L$.

In the paper, we focus on the (*K*, *L*) cloaking algorithm for a snapshot LBS. In addition, due to the few of malicious peers, we also focus on how to protect the peer node position and inquire when providing cooperation.

5. Imprecise Obfuscation Scheme

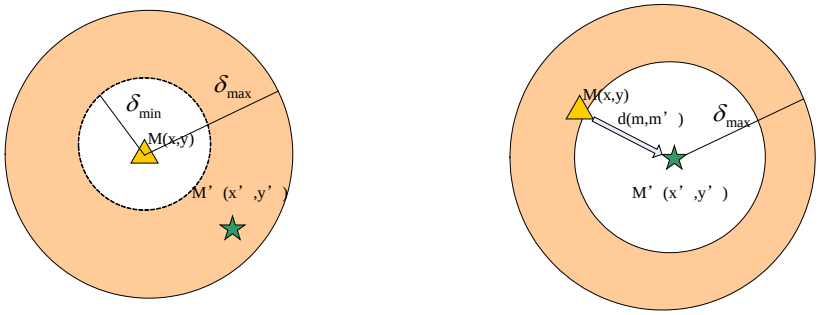
In P2P networks, to form the anonymity and confusing, peers need to reply with their location and queries to the initiator in cloaking stage. In order not to expose their exact location, this paper introduces a simple and efficient imprecise obfuscation scheme to blur their precise location.

The constants $\delta_{\min}$ and $\delta_{\max}$ determine the minimum and maximum distance of the peer's position from the center of the obfuscation region. If the coordinate of a peer node is *m*(*x*, *y*), given the offset tolerant ($\delta_{\min}, \delta_{\max}$), then the center of the obfuscation area is m' (x', y'), which meets the formula 1. Then the radius of the obfuscation region meets the formula 2. Figure 2(a) describes the restricting of the obfuscation center's coordinate via $\delta_{\min}$ and $\delta_{\max}$, and Figure 2(b) describes the restricting of the obfuscation radius.

$$\delta_{\min} \leq d(m', m) = \sqrt{(x'-x)^2 + (y'-y)^2} \leq \delta_{\max} \quad (1)$$

$$R = d(m', m) + < 0,1 > *(\delta_{\max} - \delta_{\min}) \quad (2)$$

Where < 0, 1> means a random number between 0 and 1.



(a) the restricted of the obfuscation center (b) the restricted of the R

Figure 2. Imprecise Obfuscation Scheme

6. Query-privacy-aware Location Cloaking Algorithm

In order to realize (*K*, *L*) privacy profile, we use the peer searching to form *K* spatial anonymous, along with confusing the query *L* diversity. It is the key for query privacy protection to generate query perturbation consistent with the query context, that is, the

location where the query is issued. For example, when users locate on the in the plain area, they may not issue a query about mountain. If such queries are confused, the adversary may easily to eliminate the perturbation. In this paper, we propose two ways to generate query perturbation as follows.

6.1. History Sharing Scheme

The algorithm mainly includes two phases, to search peers phase and to confuse history queries phase. In the first phase, a query initiator communicates with the peers via multi-hop to find at least *K*-1 peers. Then she can choose the *K*-1 nearest peers and herself to determine her cloaking region. The cloaking region meets the *K*-anonymous because the initiator is not distinguishing with the rest of *K*-1 peers. In the second phase, the initiator can confuse the *L*-diversity history query cached by their self and neighbors within the time tolerance. Due to the limitation of time interval, the distance of the query regional will not too far, so it can ensure the high reliability of confusion queries. Algorithm 1 describes the pseudo code of the history-sharing-based cloaking algorithm.

Step 1: Searching Peers Phase

When the user *U* want to get help from her neighbor peers, *U* broadcasts a request message, which is in the format, M_s={pseudonym, *U*.key, hop}, where the pseudonym is *U*'s identifying, *e.g.*, encrypted IP address, the *U*.key is *U*'s public key, and the hop is the transmission distance of the request message which is initialized one (line 4), then the user can increase the hop distance by one to broadcast until the number of replied peers is equal or greater than *K* or the number of hops reaches the maximum hop distance (h_{max}) (line 5-8). The friendly peers will response with the encryption replied message, which is in the format, $M_R=E_{U.key}$\{pseudonym, imprecise-loc, $<HQ, T_{send}>$\}, where the imprecise-loc calculated by imprecise obfuscation scheme, which can prevent against leaking of the exact location, and $<HQ, T_{send}>$ is the sharing history query list along with the send time T_{send}, and $HQ=\{<E_{LSP.Key}\ (q_1),\ c_1>, \cdots <E_{LSP.Key}\ (q_m),\ c_m>\}$, where $E_{LSP.Key}\ (q_i)$ is query request encrypted with the LSP public key, which can prevent against leaking to query initiator, and c_i ($1 \leqslant i \leqslant m$) is the corresponding service attribute associated with q_i. Then the initiator *U* can form her cloaking candidate peers list of all replied peers along with their sharing history query list (line 6).

Step 2: Confusing History Query Phase

In our algorithm, each peer may be the query initiator or the query requestor, so each peer is required to catch a recent launch or send query request list, along with a send time stamp T_{send}. Our system will set a tolerant time interval (*tol*) to accept history query list, *i.e.*, determine the acceptable terms as follow:

$$T_{send} \geq T_{current} - tol \qquad (3)$$

Where $T_{current}$ is the representative of the current time by inquiring service.

The initiator *U* candidates the historical queries from candidate set (Candiset) to join dummy set (Dumset) that meet the inequality 3(line 13-14), then the queries of dummy set can be chosen into the query list based on two conditions: one is that T_{send} is the recent time; the other is whose service attribute is not chosen(line 17-18). If meeting *L*-diversity the algorithm will be returned success (line 20-21), otherwise, it is for the customer to choose to give up, to send, or to repeat line 1 to line 5 by increasing *tol* (line 24).

Due to each of confusion service attribute is not identical and uniqueness, to meet the query *L*-diversity, the number of confusing history queries is *L*, i.e. $n_i = 1$, $N = \sum_{i=1}^{m} n_i = m$, then $E(Q) = -\sum_{i=1}^{m} \frac{n_i}{N} \log \frac{n_i}{N} = -\sum_{i=1}^{m} \frac{1}{m} \log \frac{1}{m} = -\log \frac{1}{m}$, $D(Q) = 2^{E(Q)} = m = L$.

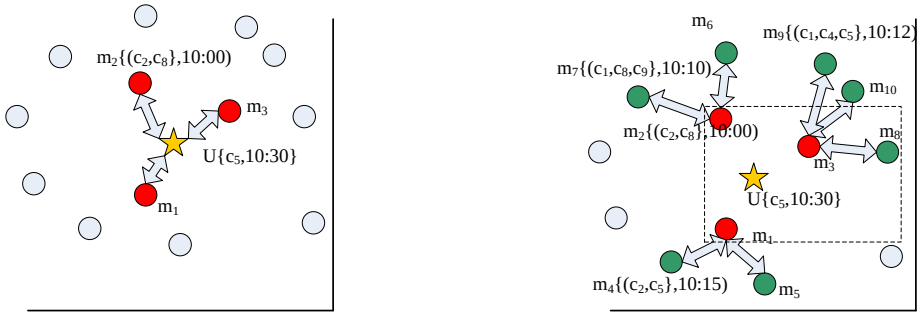
Algorithm 1 History-Sharing-based Cloaking Algorithm

Input: *U*
output: CandiSet
1) **Step 1**: **peer search phase**
2) CandiSet.PeerList ={*U*};
3) CandiSet.QueryList={*U*.query};
4) hop=1;
5) **while** *Num(candiset)*<*U.k* **do**
6) CandiSet.PeerList+={the response peers}
7) hop=hop+1;
8) **if** *hop*> h_{max} **then**
9) to return, or to continue;
10) **end if**
11) **end while**
12) **Step 2: confusing history query phase**
13) **if** $R.T_{send} >= T_{current} - tol \bigcap R \in CandiSet$ **then**
14) DumSet+= *R*;
15) **end if**
16) **while** $R \in DumSet$ **do**
17) **if** $R.T_{send}$ is the recent and $R.HQ.c \notin CandiSet.QueryList.C$ **then**
18) CandiSet.QueryList+=*R.HQ*;
19) **end if**
20) **if** $D(CandiSet) \geq L$ **then**
21) return CandiSet;
22) **end if**
23) **end while**
24) choosing to give up, or to return, or to repeat by increasing *tol*

Figure 3 depicts a running example to illustrate the algorithm. Figure 3(a) illustrates the peer searching phase with one hop. Within the transmission of *U*, *U* finds three neighbor peers, donated by m_1, m_2, and m_3 respectively, who described by red circles. Because that the number of response peers is lower than five (*U.K*=5). Then *U* has to continue broadcasts her request with *hop*=2 for searching more peers. Figure 3(b) describes that seven peers who described by green circles replies the request. Thus the anonymous candidate set, donated by $A = \{U, m_1, m_2 \ldots m_{10}\}$, meets the 5-anonymity, so the searching step can be stopped, and the confusing history query phase will be started.

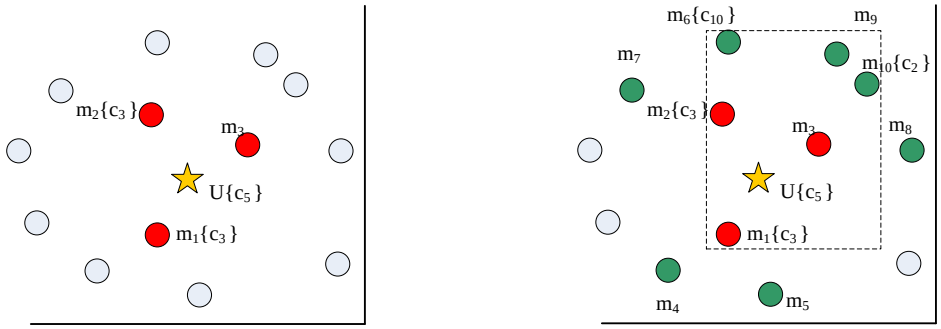
In figure 3(b), *U* can firstly candidate m_4's history query list ($\{c_2, c_5\}$), who has the recent time ($m_4.T_{send}$ =10:15). Due to consider the diversity, c_5 should be discarded. Then *U*'s dummy query list is $\{c_2, c_5\}$, which is not meet the query diversity requirement (*U.L*=3). Thus *U* continue chose m_9 whose time ($m_9.T_{send}$=10:12) is recent. So, the c_4 or c_1 can be chosen, i.e. $\{c_1, c_2, c_5\}$ or $\{c_4, c_2, c_5\}$. So, the confusion history query step can be stopped. Next, the

algorithm forms a smallest anonymous rectangle region including the nearest *K-1* peers from *U*, which are depicted by the dotted line. Since the encrypted response message can only be decrypted by *U*, therefore, flooding attacks can be effectively prevented.



(a) One-hope peer search (b) Two-hope peer search

Figure 3. A Running Example of History-sharing-based Cloaking Algorithm (*U* is the Query User and *U.K*=5, *U.L*=3, *tol*=20min)



(a) One-hope peer search (b) Two-hope peer search

Figure 4. A Running Example of Batch-query-based Cloaking Algorithm with Uniform Privacy Profile (*K*=5 and *L*=3)

6.2. Batch Query Scheme

Since algorithm 1 meets *L*-diversity through confusing history queries, which can spend the LSP extra process time. To address the problem, we propose the confusing batch queries cloaking scheme. Compared to the previous scheme, confusing queries are not historical queries, but real-time query requests coming from the cooperation peers' response.

Considering that the privacy profile can be described by uniform *(K, L)* and personalized *(K, L)*, so the two aspects are discussed respectively as follows.

6.2.1. Uniform *(K, L)*-privacy profile: Compared with the algorithm 1, the response message will be joined with the cooperation peers' real-time query information ($\{F, (Q_{LSP}, c)\}$), which is in the format, $M_R=E_{U.key}$\{pseudonym, imprecise-loc, F, (Q_{LSP}, c), $<HQ, T_{send}>$\}, where (Q_{LSP}, c) is the encrypted query and its non-encrypted service attribute, F is a flag (if F equals 0, (Q_{LSP}, c) is null), and Q_{LSP} =E_{LSP} (query) can prevent the leaking of the requirement to the initiator. The searching process can see algorithm 1.

Due to the uniform *(K, L)* -privacy profile, not especially considering the difference of privacy requirement, the (Q_{LSP}, c) of all nodes will be joined in *U*'s candidate set

according to the distance to *U* from near to far until meeting the requirements of the *L*-diversity, otherwise calling step 2 of the algorithm 1.

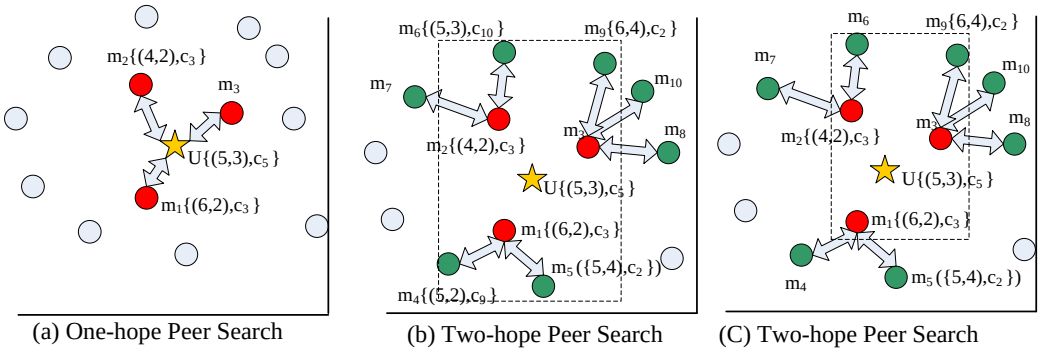
Figure 4 describes the running example, the searching stage of *hop*=1 is the same as Figure 3. In this figure, not the peers' history list, but the peers' real query is marked. In searching stage of *hop*=2, because that the candidate set gather 11 peers (*K*=5) and 5 query requests, meanwhile $E(Q) = -(\frac{2}{5}\log\frac{2}{5} + \frac{3}{5}\log\frac{1}{5})$ (the query type and corresponding quantity are respectively: c_2 (1), c_3 (2), c_5 (1), c_{10} (1)), $D(Q) = 2^{E(Q)} > 3$ (*L*=3), so it meets the privacy profile and the algorithm stops. Next, the algorithm forms a smallest anonymous rectangle region including the query peers, which are depicted by the dotted line.

6.2.2. Personalized *(K, L)*-privacy Profile: In practical applications, each user has a different privacy requirement, so this paper puts forward an algorithm, which confuses batch query cloaking with personalized *(K, L)* to meet the flexibility and personalization.

Compared with the algorithm mentioned above, the response message will be joined with the different privacy profile of each user, i.e. *MA*= {pseudonym, imprecise-loc, *F*, (Q_{LSP}, c), *(K, L)*, (*HQ*, T_{send})}. Algorithm 2 describes the pseudo code. If the candidate set meets the max of *(K, L)* among the response peers, the searching step should be stopped (line 9-10), otherwise *U* will continue to search by increasing hop count until the maximum hop count(line 12). If the *(K, L)*-privacy profile can not be satisfied, *U* will call step 2 of the algorithm 1(line 15).

Algorithm 2 Batch-Query-based Cloaking Algorithm with Personalized *(K, L)*

Input: *U*
output: CandiSet
1) CandiSet.PeerList ={*U*};
2) CandiSet.QueryList={*U*.query};
3) hop=1;
4) **while** $h<h_{max}$ **do**
5) CandiSet.PeerList+={the response peer *M*}
6) **if** *M.F=1* **then**
7) CandiSet.QueryList+=*M*.query;
8) **end if**
9) **if** CandiSet.QueryList meet the max of *(K, L)* **then**
10) return CandiSet;
11) **end if**
12) hop=hop+1;
13) **end while**
14) **if** CandiSet.QueryList do not meet the max of *(K, L)* **then**
15) call step 2 of the algorithm 1;
16) **end if**



(a) One-hope Peer Search (b) Two-hope Peer Search (C) Two-hope Peer Search

Figure 5. A Running Example of Batch-query-based Cloaking Algorithm with Personalized *(K, L)*

Shown as Figure 4, not the peers' history list, but the peers' real query is marked in figure 5. Figure 5(a) depicts the algorithm with *hop*=1, the anonymous candidate set A= {U, m_1, m_2, m_3}, and m_1 and m_2 have query requests. Due to the max of K and L are 6 and 3 respectively, $A.K = 4 < 6$, $E(A) = -(\frac{2}{3}\log\frac{2}{3} + \frac{1}{3}\log\frac{1}{3})$ **(**the query type and corresponding quantity are respectively: c_3 (2), c_5 (1)**)**, $A.L = D(A) = 2^{E(A)} < 3$, so continue to search with *hop*=2, which is illustrated by figure 5(b). In the second step, the max of K and L are 6 and 4 respectively, and $A.K$=11>6, $E(A) = -(\frac{3}{7}\log\frac{1}{7} + \frac{4}{7}\log\frac{2}{7})$ **(**the query type and corresponding quantity are respectively: c_2 (2), c_3 (2), c_5 (1), c_9 (1), c_{10} (1)**)**, $A.L = D(A) = 2^{E(A)} > 4$, so the algorithm stops. Figure 5(c) depicts that only m_9 has request with *hop*=2, if h_{max} =2, and $E(A) = -(\frac{2}{4}\log\frac{2}{4} + \frac{2}{4}\log\frac{1}{4})$ **(**the query type and corresponding quantity are respectively: c_2 (1), c_3 (2), c_5 (1)**)**, then $A.L = D(A) = 2^{E(A)} < 4$, so calling step 2 of the algorithm 1. Next, the algorithm forms a smallest anonymous rectangle region including the query peers, which are depicted by the dotted line.

7. Experimental Evaluation

In this section, we evaluate the performance of the proposed query-privacy-aware location cloaking algorithm (donated as **QA**) with three scheme, *history sharing scheme* (donated as **QA-HS**), *batch query scheme with uniform (K, L)* (donated as **QA-BQ-U**), *batch query scheme with personalized (K, L)* (donated as **QA-BQ-P**). For comparison purpose, we have implemented one other approach, termed as P2P Cloaking, which form *K*-anonymity area only using the first phase in algorithm 1, not considering the confusing query.

In our experiments, we use a networked generator to randomly generate 100 to 500 mobile users (default value is 200) on a spatial space of 6280,000 square meters. Mobile users are randomly selected to issue queries and have a transmission distance of 200 meters; the ratio of the number of query users to the number of mobile users is 0.3. The privacy profile K and L are uniformly selected from a range [5, 30] and [1, 20] respectively, and their default value are 20 and 10 respectively. The algorithm will form history query list of some users after it running for a period of time, and the query categories can be defined by the system. The mobile users are moving at speeds between 50 to 80 miles per hour. The time tolerance for

dummy queries and h_{max} are 20 minutes and 5-hops respectively. The parameter settings are shown on the Table 1.

Table 1. Parameter Settings

Parameters	Default values	Range
Number of mobile users	200	[100,500]
Number of querying users	60	[20,80]
Transmission range	200m	200m
K- anonymity	20	[5,30]
L- diversity	10	[1,20]
Tol	20mimutes	[5,30]minutes
h_{max}	5	5

1) **Protection Level.** This measures query privacy attack probability. Figure 6 shows the resilience of our algorithms to the *homogeneity attack* with respect to varying the *L*-diversity level from 1 to 20, and the query users from 20 to 80. The results show that the P2P Cloaking algorithm is the most vulnerable to the *homogeneity attack*, while **QA-HS**, **QA-BQ-U**, **QA-BQ-P** are very close to 1/L. Thus, the three schemes taken in this paper can effectively prevent the *homogeneity attack*.

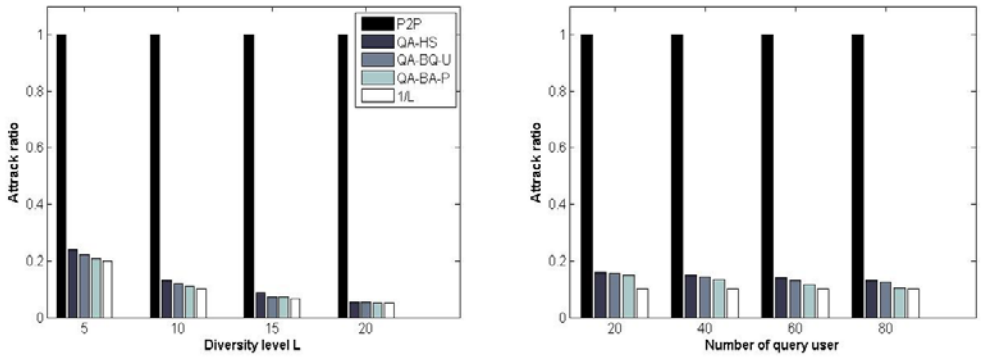


Figure 6. Query Homogeneity Attack

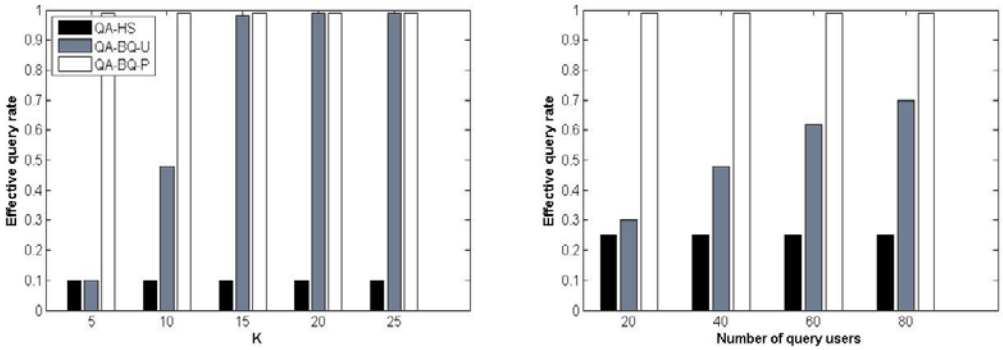


Figure 7. Effective Query Rate

2) **Effective Query Rate.** This is a ratio of the actual queries to the number of query list sent to the LSP. Figure 7 evaluates the LSP's effective overhead of our algorithm with respect

to *K*-anonymity level and number of query users. The higher the ratio is, the smaller the number of dummy query is. The results shows that the effective query rate of confusing batch query scheme is very high, especially the **QA-BQ-P** algorithm is very close to 1. Thus, the confusing batch query scheme does not increase the query overhead of LSP.

3) **Cloaking Area Size.** This is the average size of cloaking region. Figure 8 depicts that when the value of *K* increases, the average cloaking area under all schemes increases. The results indicate that the cloaking area of **QA-HS** and **QA-BQ-U** are very close to P2P Cloaking, the cloaking area of **QA-BQ-P** is larger than others. This is due to the fact that it is likely to search more query users than others to satisfy the level of *L* privacy requirement, not only considering the *K*-anonymous requirement. It also shows that the cloaking area of P2P Cloaking and **QA-HS** are not relevant to the level of *L*. This is because that they blur users only considering the level of *K*, the *L*-diversity of **QA-HS** is realized by using the confusing history queries.

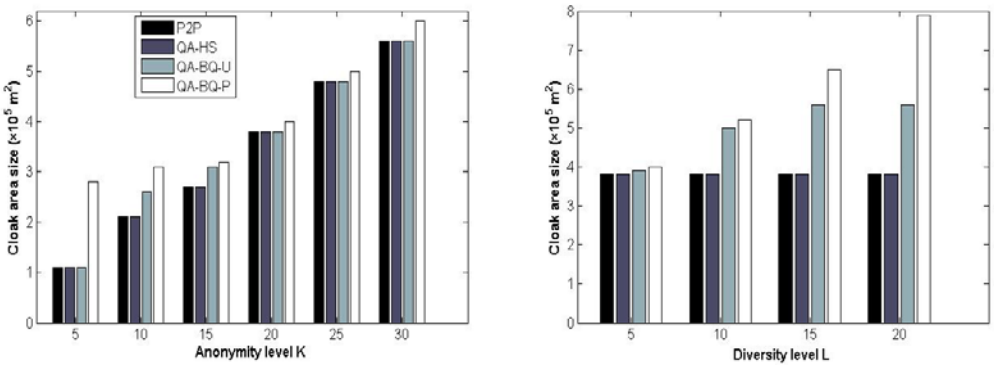


Figure 8. Cloaking Area Size

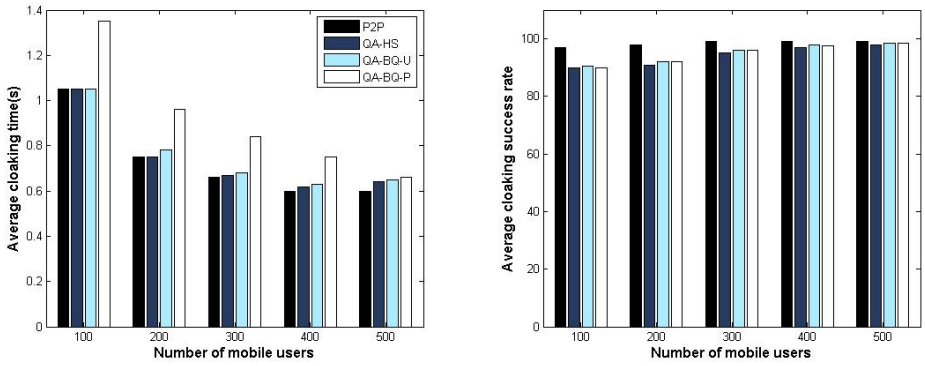


Figure 9. Scalability

4) **Scalability.** This measures the average cloaking time and the average success rate with respect to the number of mobile users from 100 to 500 respectively. Figure 9 shows that the cloaking time of all algorithms slightly decreases with increasing the number of users. This is because the user can find enough number of peers via fewer hops. Moreover, **QA-HS** and **QA-BQ-U** do not take extra time to search the peers, so they have almost the same cloaking time with **P2P** Cloaking. While **QA-BQ-P** has rises for obtaining highly effective query, but

QA-BQ-P is close to P2P when the number of users' value is 500. Figure 9 also shows that the performance of all algorithms gets better when there are more users. Moreover, there are more query users when the number of users increases. Thus, it is more likely that the user can find enough history queries and real queries to realize *L*-diversity. So the cloaking success rate of three algorithms taken in this paper improves more than P2P Cloaking with increasing the number of users.

8. Conclusion

In this paper, we present a novel query-preserving-aware cloaking method for mobile P2P system, which realizes both query-privacy-preserving and location-privacy-preserving. The main idea of the method is to realize the *L*-diversity along with cloaking through confusing history queries or actual queries, termed as *history sharing scheme* and *batch query scheme* respectively. It is important to note that forming the *(K, L)*-cloaking area does not rely on the anonymity server, which may bring the bottleneck of system. Moreover, *L*-diversity measurements rely on query entropy, rather than considering the differences in service attributes.

The experiments indicate that our algorithms combine high level of privacy with a high level of QoS. However, due to the lack of anonymous server for P2P system, if the query initiator directly access to LSP, it will be vulnerable to the *trackback attack*. So, in the future work, we can focus on fair and efficient requestor agent selection scheme.

Acknowledgements

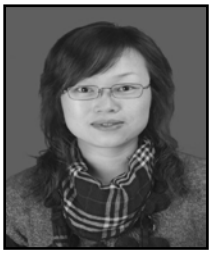
This work was suppoprted by Key Project on the Integration of Industry, Education and Research of Guangdong Province (Grant No. 2012B091000054), and Sichuan Department of Education (Grant NO.13ZB0152).

References

[1] B. Gedik and L. Liu, "Location Privacy in Mobile Systems", A Personalized Anonymization Model. Proceedings of 25th IEEE International Conference on Distributed Computing System, Columbus, OH, **(2005)**.
[2] M. Gruteser and D. Grunwald, "Anonymous Usage of Location-Based Services through Spatial and Temporal Cloaking", Proceedings of 1st International Conference on Mobile System, New York, **(2003)**, pp. 31-42.
[3] F. M. Mohamed, C. Y. Chow and G. A. Walid, "The new casper: Query processing for location services without compromising privacy", Proceedings of 32nd International conference on Very Large Data Bases, ACM Press, pp. 763-774, **(2006)**.
[4] T. Xu and Y. Cai, "Location anonymity in continuous location-based services", Proceedings of 15th ACM Symposium on Advances in Geographic Information Systems, New York, **(2007)**.
[5] T. Xu and Y. Cai, "Exploring historical location data for anonymity preservation in location- based services", Proceedings of IEEE INFOCOM 27th International Conference of the Computer and Communications Societies, Phoenix, AZ, **(2008)**, pp. 547-555.
[6] B. Bamba, L. Liu and P. Pesti, "Supporting anonymous location queries in mobile environments with PrivacyGrid", Proceedings of the 17th International Conference on World Wide Web, New York, **(2008)**, pp. 237-246.
[7] P. Kalnis, G. Ghinita and K. Mouratidis, "Preventing location-based identity inference in anonymous spatial queries", Proceedings of IEEE Transactions on Knowledge and Data Engineering, **(2007)**, pp. 1719-1733.
[8] P. Samarati, "Protecting Respondents Identities in Microdata Release", Proceedings of IEEE Transnaction on Knowledge and Data Engineering, IEEE Press, **(2001)**, pp. 1010-1027.
[9] L. Sweeney, "K-anonymity: A model for protecting privacy", International Journal of Uncertainty Fuzziness and Knowledge Based Systems, vol. 10, no. 557, **(2002)**.

[10] T. Hashem and L. Kulik, "Safeguarding location privacy in wireless ad-hoc networks", Proceedings of International Conference on Ubiquitous Computing (UBICOMP), Innsbruck, Austria, **(2007)** September 16-19.
[11] G. Ghinita, P. Kalnis and S. Skiadopoulos, "Mobihide: A mobile peer-to-peer system for anonymous location-based queries", Advances in Spatial and Temporal Databases Lecture Notes in Computer Science, vol. 4605, no. 221, **(2007)**.
[12] G. Ghinita, P. Kalnis and S. Skiadopoulos, "PRIVE: Anonymous location-based queries in distributed mobile systems", Proceedings of International World Wide Web Conference, New York, USA, **(2007)**, pp. 371-380.
[13] C. Y. Chow, M. F. Mokbel and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location based service", Proceedings of 14th annual ACM international symposium on Advances in geographic information systems, New York, USA, **(2006)**, pp. 171-178.
[14] W. S. Ku, R. Zimmermann and H. Wang, "Location-based spatial query processing with data sharing in wireless broadcast environments", Proceedings of International Conference on IEEE Transactions on Mobile Computing, IEEE Press, **(2008)**, pp. 778-791.
[15] C. Y. Chow, M. F. Mokbel and X. Liu, "Spatial Cloaking for Anonymous Location-based Services in Mobile Peer-to-Peer Environments", J. GooInformatica, vol. 15, no. 351, **(2011)**.
[16] J. Xu, X. X. Huang and M. Guo, "Location privacy through anonymous chain in dynamic P2P network (in Chinese)", J. Zhejiang University(Engineering Science), vol. 46, **(2012)**, no. 712.
[17] A. Pingley and Z. Nan, "Protection of Query Privacy for Continuous Location Based Services", Proceedings of IEEE INFOCOM, IEEE Press, Shanghai, China, **(2011)**, pp. 1710-1718.
[18] G. Ghinita, P. Kalnis and A. Khoshgozaran, "Private Queries in Location Based Services: Anonymizers are not necessary", Proceedings of SIGMOD, New York, USA, **(2008)**, pp. 121-132.
[19] A. Khoshgozaran and C. Shahabi, "Location privacy: going beyond K-anonymity, cloaking and anonymizers", Computer Science Knowledge and Information Systems, vol. 26, **(2011)**, no. 435.

Authors



Min Li, was born in Sichuan, China, in 1978. She received the M.S. degree in the University of Electronic Science and Technology of China, in 2005. She is currently working toward the Ph.D. degree in computer science at UESTC. Her current research interests include wireless sensor networks, privacy protection, specifically the location privacy in LBS.

Zhiguang Qin, received Ph.D degree from the University of Electronic Science & Technology of China. Now he is a professor, dean of Computer Science & Engineering Department, director of Computer Application Key Lab in Sichuan Province, member of IEEE. Currently his main research interests concern is the security of the networks.

Cong Wang, received the B.S. and M.S. degrees from Southwest University of China, Chong-qing, China. Currently he is pursuing the Ph.D. degree in computer science at the University of Electronic Science & Technology of China.His main research interests are the applications of machine learning techniques to computer networking problems, specifically the prediction of latency in large-scale networks.