PROCEEDINGS OF SCIENCE

AP2P Spatial-cloaking Algorithm without Exposing the Collaborators for LBS

Zhengyuan Zhai[1]
Communication and Network Research Center, Beijing University of Posts and Telecommunications
Beijing, China
E-mail: zhaiyuan@bupt.edu.cn

Chunxiao Fan[2]
Communication and Network Research Center, Beijing University of Posts and Telecommunications
Beijing, China
E-mail: cxfan@bupt.edu.cn

User's privacy issue is an important concern in Location-based Services (LBS) with widespread use. A well-known privacy preserving architecture for LBS is the Peer to Peer (P2P) model which allows users to collect other collaborators' locations through P2P network. Thus, it would incur the serious privacy disclosing if the users in P2P network maliciously collect those collaborators' locations. In this paper, we propose a spatial-cloaking algorithm which considers the being of the malicious collectors in P2P network. The key difference between our program and the existing schemes is that the collaborator map his location into an index which is not one-one mapping, so it can avoid the aforementioned issue in P2P network efficiently. In addition, our algorithm can also provide a personalized privacy-preserving for LBS in the road network. The experiment results show that our proposed algorithm is more efficient when it's compared to some existing work in terms of the communication cost, the anonymous time, the size of generated cloaking region and the success rate of anonymizing a location.

ISCC 2015
18-19, December, 2015
Guangzhou, China

[1]Speaker

[2]This study is supported by the National Great Science Specific Project (Grants No. 2014ZX03002002-004) and National Natural Science Foundation of China (Grants No. NSFC-61471067 and No. NSFC-61170176)

PoS(ISCC2015)003

1. Introduction

Recently, due to the ubiquity of smart phones and mobile devices, as well as the rapid development of the wireless communication network, location-based services (LBS) have been widely used. In LBS, the mobile users utilize their precise location to gain the desired services. However, to get a service of high quality, mobile users must pay for the loss of their privacy. Because when a LBS query and an exact location are submitted, user's individual information (such as home address, work place, and interests) can be linked with the location query using some techniques, i.e. infer the identity by tracing locations users submitted [1]. Additionally, the LBS users are always moving on the road and require different levels of privacy as well as quality of services (QoS). Therefore, it is imperative to design a feasible privacy-preserving mechanism to address the privacy concerns of LBS.

To solve the privacy issue of LBS, many approaches based on spatial-cloaking have been well studied [2-4, 8-12], which mostly employ a privacy metric called k -anonymity. To achieve the k -anonymity, queries related with k locations are submitted to the LBS server mainly through the centralized or Peer to Peer (P2P) architecture. The centralized architecture builds a cloaking region (CR) enclosing other $k-1$ users' locations with the help of the trusted location anonymizer, while the P2P architecture collects other $k-1$ collaborators' locations to form a CR via multi-hop communication. However, the location anonymizer is vulnerable to large-scale computing by anonymizing location and extracting accurate results for all users. Moreover, it's likely to be attacked by an adversary for having too much knowledge of the users' locations and queries [5-7]. The P2P model is first suggested by Chow et al. [8], and greatly extended in some papers [9-12], for example, Che et al. proposes the dual-active model enabling users to collect neighbor peers' gathered location [11], which avoids the deficiencies caused by the centralized anonymizer. Nervertheless, most existing approaches of the P2P model holding all peers in the P2P network are trustworthy. In this case, an adversary would easily collect others' locations if pretending to be a normal user. To this end, some cloaking algorithms without exposing the collaborators' locations to other peers in the P2P network have been proposed [13, 14]. They all have limiations for relying on Wi-Fi received signal strength of adjacent peers or time difference of arrival of beacon signal and the collaborators' blurred region to form a CR. Besides, these methods are performed in Euclidean space, which is not efficient to protect the users' locations when they are moving on the road. Motivated by this, we present a novel spatial-cloaking approach to preserve users' locations privacy in the mobile P2P network, which considers road constraints and personalized privacy requirements. To our knowledge, this is the first paper proposed to solve the location privacy issue in the P2P architecture of LBS under road-network. The main contributions of this paper are summarized as below: (1) supporting road-network environment. We divide the whole road-network into many cells based on the density of the road, so our work is appropriate in road-network environment; (2) resistance to location leaking in P2P model of privay-preserving LBS. In our program, the collaborators compute and share their location indexes based on their position inside a cell. In this way, when a LBS query peer receives a collaborator's location index, he can determine where he is inside the cell, but doesn't know the exact location; (3) providing personality privacy requirements. Based on the same user's loaction index, we can retrieve some dummy locations. By covering these different dummy locations, we can generate different CRs to meet a personalized privacy requirement.

2.Preliminaries

In this section, we review some knowledge background related with our proposed program, then present the P2P system architecture and users' privacy requirements.

2.1 Road-network

Definition 1: Road-network graph. As depicted in Fig. 1, a road-network graph G is an undirected graph comprised of the set V and E , where $V=\{v_1, v_2, \ldots, v_n\}$ is the set of network nodes, while $E=\{e_{ij}\}$ is the set of edges. The edge e_{ij} connects two adjacent nodes v_i and v_j with a weight of ω_{ij} denoting the Euclid distance from v_i to v_j .

We define the number of edges connecting with the node v_i as its degree $d(v_i)$. When $d(v_i)$ equals to 1, 2, 3 or more, it's an end node, an intermediate node or a junction respectively, i.e. the node v_1 is an end node, the node v_5 is a junction as shown in Fig. 1.

Definition 2: Network distance from points to road junctions. The network distance $d(P,V)$ from a point P to the junction V , differing from the Euclid distance $d_E(P,V)$, is defined as the sum of weights from P to V ,as shown in Fig. 1, the distance from the node v_1 to the junction v_5 is the sum of weights between ω_{14} and ω_{45} .

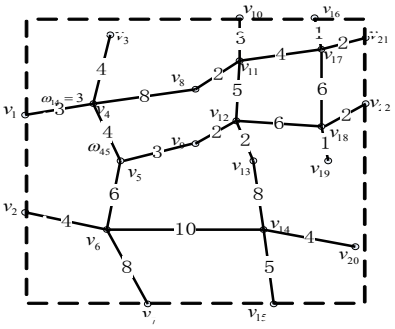


Figure 1: A Road-network Model

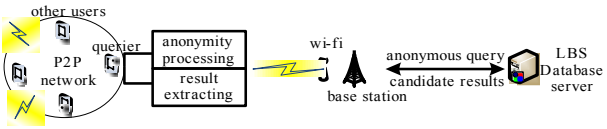


Figure 2: A Review of P2P LBS System Architecture

2.2 Hilbert-filling Curves

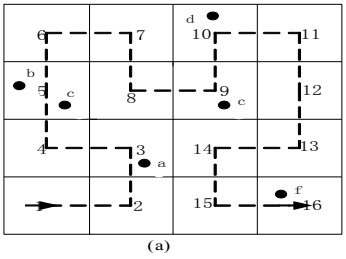
The Hilbert-filling curves, as showed in Fig. 3 (a), first partition a 2-d space into some cells of the same size, and then traverse through all cells in a certain order without crossing. Fig. 3 (b) is a variation of the Hilbert-filling curve partitioning the space into cells of various sizes based on the road density 4, which means each cell including at least 4 road segements. Some work has utilized the Hilbert-filling curve to achieve privacy-preserving, which has an obvious drawback for mapping users in the same cell to a H-value,as shown in Fig. 3 (a), the users b and c locate in the cell 5, so the H-value of their location is 5. This way can't differentiate people inside the cell well, resulting in a larger CR when the population density is low.

2.3 P2P System Architecture and Privacy Requirements

As illustrated in Fig. 2, there are two important entities in the P2P architecture of LBS: mobile users and LBS database servers. Mobile users are equipped with cellphones. The LBS user first communicates with other peers to organize a P2P network via ad hoc network rooting, then hide his location in a community by anonymity processing, and finally, he submits the query and CR to the LBS servers forwarding by the base stations. After receiving the query and CR, the LBS servers find a set of candidate results in their databases and return these results to the user. The user extracts demanding results from the candidate results with his precise location.

In the P2P system architecture of LBS, there are 3 types of adversaries: untrusted LBS servers, semi-trusted peers and eavesdroppers. Our paper mainly focuses on the untrusted LBS servers and the semi-trusted peers who can perform the inference attack or collude with other peers in P2P network. We provide that the semi-trusted peers can't deliberately send false information once they join a P2P network. In our framework, each mobile user has its own privacy profile which includes three parameters (k, l, A_{min}) in terms of k -a anonymity, l -diversity and a minimum area A_{min} of CR.

Definition 3 (k, l, A_{min}) **-indistinguishability**. An anonymous processing CR is (k, l, A_{min}) -indistinguishability, if the user's real location in CR is indistinguishable from other $k-1$ submitted locations and all locations of CR locate in at least l different road segments within the desired size A_{min} of CR.



Points	H-value
a	3
b	5
c	5
d	10
e	9
f	16

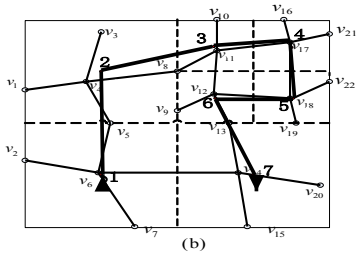


Figure 3 (a): A Hilbert Curve with Order 4(4×4 cells) **(b):** A Variation of Hilbert-filling Curve

3. Our Cloaking Algorithms

3.1 Initialization

In the initialization phase, trust authority (TA) recursively partitions the whole road network into N cells, and sets a central server CS which can be a base station in the center of each cell. In the process of our partition, each cell includes at least 4 road segements and 1 road junction. Algorithm1 shows a detailed process of partitioning and filling the road network. In the end, TA broadcasts the system parameter $param=(H_n, loc_{CS_n}, R_n, N_n, \{v_j^n\}, \{r_j^n\})$ to the mobile users. For simplicity, we arrange $\{v_j^n\}$ in an ascend order by $d_E(loc_{CS_n}, v_j^n)$.

3.2 Cooperative Sharing Phase

After initialization phase, each CS_n first organizes mobile users locating in the $cell_n$ to form a cooperative group, and then periodically updates a list of registered collaborators $Regist=(UID, h_{user}, t_k)$, where UID indicates the unique identity of the user, and h_{user} is the location index of user at time t_k. When a mobile user is willing to join the cooperative group, he computes his location as in Algorithm 2, and then shares a 3-tuple record (UID, h_{user}, t_k) with other mobile users by submitting to the CS_n.

3.3 Anonymous Query Phase

When a mobile user in the $cell_n$ launches a query, he first refers to the CS_n and the TA respectively for obtaining the list of registered collaborators $Regist$ in the $cell_n$ and the positions of nodes, the horizontal angles or vertical angles of each edge in G_n. After receiving these parameters, he can choose some collaborators whose time stamp t_k in $Regist$ is close to his current time and retrieve these collaborators' dummy locations as in Algorithm 3. Then, the query user can use these dummy locations and his location to generate a CR satisfying his privacy requirement.

Based on user's location index, our algorithm can retrieve dummy locations in one cell to meet different privacy requirements. Fig. 4 elaborates we can generate three different CR meeting different privacy levels based on a location index table, where CR_1 includes other four users and the LBS user at one road segments, CR_2 is a minimum area including 4 users and the real LBS user, while CR_3 meets (k, l, A_{min}) -indistinguishability.

Algorithm1: Partitioning and Hilbert-filling the Road Network

Partitioning phase

Input: road network $G=(\{v_i\},\{e_{ij}\},\{\omega_{ij}\})$, square G_0, defined num

Output: a sub-road graphs and relevant cells Table T $\{(G_n, cell_n)\}$

1: initialize G_0 including G and the Table T

2: if the number of junctions inside $G_0 > num$

3: recursively partitioning G_0 into four squares, $cell_k, k=1,2,3,4$

4: add end points at the interface of G_0 edges and $cell_k$

5: get sub-road network $G_k=(\{v_i^k\},\{e_{ij}^k\},\{\omega_{ij}^k\})$ for $cell_k$

6: if there is G_k, the number of junctions inside it $< num$

7: add G_0 and relevant cell into Table T

8: else for each sub-network $G_k, k=1,2,3,4$, return to progress 1

9: return Table T

Hilbert-filling phase

Input: Table T: $G_n=(\{v_i^n\},\{e_{ij}^n\},\{\omega_{ij}^n\})$

Output: parameters $(H_n, loc_{CS_n}, R_n, N_n, \{v_j^n\}, \{r_j^n\})$

1: use the Hilbert-filling curve connect cells in Table T

2: determine the H-value H_n of $cell_n$ in connection order

3: compute center position loc_{CS_n} and size R_n for each $cell_n$

4: find the number of junctions N_n, junctions set $\{v_j^n\}$ and compute the least network distance r_j^n to junction v_j^n for each G_n

Algorithm 2: Computing User's Location Index

Input: user Alice's current location loc_{Alice}, system parameter $param$

Output: Alice's current location index h_{Alice}

1: for each loc_{CS_n} and R_n, compute the Euclid distance $d_E(loc_{Alice}, loc_{CS_n})$

2: if $d_E(loc_{Alice}, loc_{CS_n}) < R_n$, compute $d(loc_{Alice}, v_j^n), j = 1,2,\ldots, N_n$

3: if $r_{j'-1}^n < d(loc_{Alice}, v_{j'}^n) \le r_{j'}^n$, compute $h_{Alice} = H_n + \dfrac{j'.\ d(loc_{Alice}, v_{j'}^n)}{N_n.\ r_{j'}^n}$ (3.1)

4: else find the least $d(loc_{Alice}, v_{j'}^n)$, compute $h_{Alice} = H_n + \dfrac{j'}{N_n}$ (3.2)

Algorithm 3: Retrieving Users' Dummy Locations

Input: $param$, collaborators' list CL_n, junction positions $loc_{v_j^n}$ and edges directions θ_{ij}^n of $cell_n$

Output: user's possible locations $\{(x_{user}, y_{user})\}$

1: if $H_n + \dfrac{j-1}{N_n} < h_{user} \le H_n + \dfrac{j}{N_n}$, $j = 1,2,\ldots, N_n$

2: compute $d(loc_{user}, v_j^n) = \dfrac{(h_{user} - H_n).\ N_n.\ r_j^n}{j}$ (3.3)

3: find all the possible paths from the user to v_j^n with $d(loc_{user}, v_j^n)$

4: determine the nearest node $v_{i'}^n$ to user on each path

5: compute $x_{user} = x_{v_{i'}^n} + (d(loc_{user}, v_j^n) - \omega_{i'j}^n).\cos(\theta_{i'}^n)$ (3.4)

6: and $y_{user} = y_{v_{i'}^n} + (d(loc_{user}, v_j^n) - \omega_{i'j}^n).\sin(\theta_{i'}^n)$ (3.5)

7: return all the possible locations $\{(x_{user}, y_{user})\}$

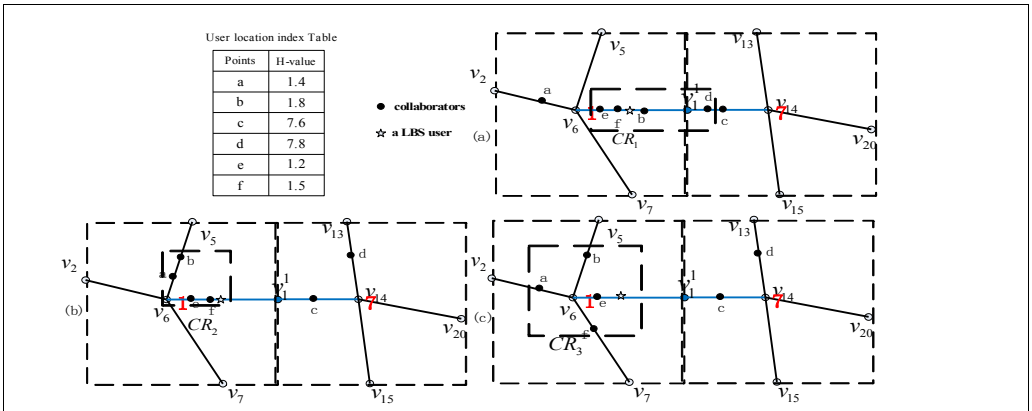


Figure 4: Examples of Generating Different CRs Based on a Same Location Index Table

4.Experiments

This section evaluates the performance of our cloaking algorithms compared with the Dual-active [11], the EDA [12], and theX-region algorithm [14]. Our simulated data are generated by the Thomas Brinkhoff generator with a road network of Oldenburg which includes 6105 nodes and 7035 edges. We build 1758 cells of various sizes with $l=4$ over this road network. Table 1 presents the parameter settings and defaults in our experiment. Our algorithm is performed in Java with an environment of Windows 7, Intel i3 2.50GHz CPU and 4GB RAM.

parameter	users' number N	users'number (a cell) n	anonymity k	degree	road density l	area of CR A
ranges	[5000,25000]	[5,40]	[5,20]		[4,8]	[400,20000]
defaults	10000	20	5		4	400

Table 1: Parameters and Defaults of our Experiment

We compare our algorithm with existing algorithms in respect of the communication cost, the anonymous time, the size of the CR and the success rate of anonymous. The communication cost is the average communitation hops and average size of transmitted messages in the phase of cooperative sharing, while the success rate of anonymous indicates the ratio of users in a cell who can be successfuly anonymized with a certain privacy level.

Fig. 5 (a) shows the comparison between our algorithm with the EDA and the Dual-active algorithm in the average communication hops, which indicates our algorithms has little change, but the EDA and the Dual-active increase before the number of users reaching 15000. In the cooperative sharing phase, if there are k users in the cell, the broadcasting numbers follow the poisson distribution $\frac{(e^{-\lambda}).\lambda^k}{k!}$ for the EDA and the Dual-active ($\lambda=6$ and 8 respectively), while the user only communicates with the TA and some central servers in our model. Fig. 5 (b) reveals our algorithm is superior to the EDA and the X-region on the average size of transmitted messages. This is because the EDA transmit user's 2D coordinates in the message, whereas we transmit 1D value. The X-region algorithm has the worst result for transmitting an obscured region. Fig. 6 and Fig. 7 respectively show the response time and the size of region with different privacy levels when anonymizing a location. From these two Figures, we can draw that the response time and the size of cloaking region increase with more stricted privacy levels. We can also obersve that our scheme increases less than the other two methods for retrieving many dummy locations as shown in the Algorithm 3.

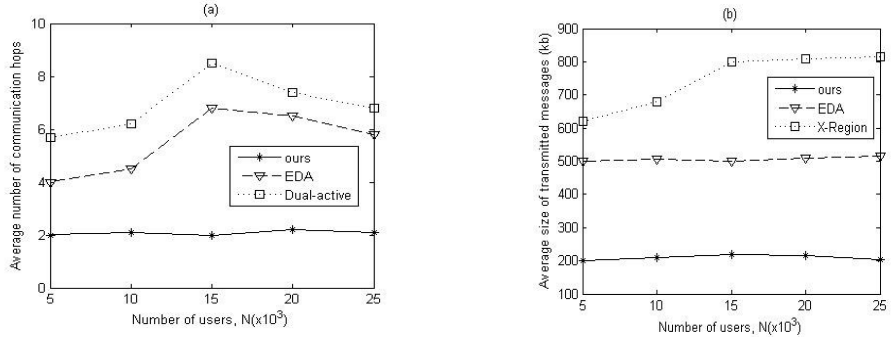


Figure 5: The Communication Cost, (a) the Average hops (b) the Average Size of Messages

Finally, we experimentally evaluate the success rate of anonymity for our scheme and compare it with the other two methods' result, as showed in Table 2. From Table 2, we can draw

a conclusion that our method reaches a higher success rate of anonymity with different privacy levels when 40 users scatter randomly in the cell. The cause accounting for such superior is that when there are k users in the cooperative sharing group, the user receives k locations to form a CR for the Dual-active and the EDA, while the user in our scheme can retrieve at most $4k$ dummy locations as illustrated in the Formulas (3.3.4) and (3.3.5), which greatly increases the probability of a user locating in an anonymous group.

	Ours	Dual-active	EDA
$k=5, l=4$	99.1%	97.6%	98.1%
$k=10, l=4$	95.4%	91.8%	92.6%
$k=15, l=4$	92.3%	90.1%	90.8%
$k=20, l=4$	90.4%	85.9%	87.1%

Table 2: The Success Rate of Anonymous with Different Privacy Levels

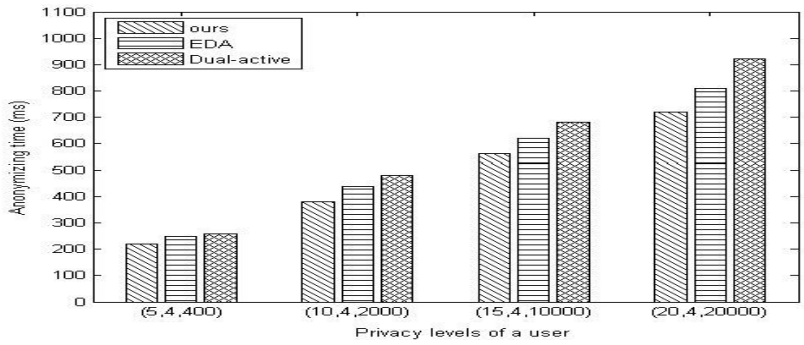


Figure 6: The Anonymizing Time against Different Privacy Levels

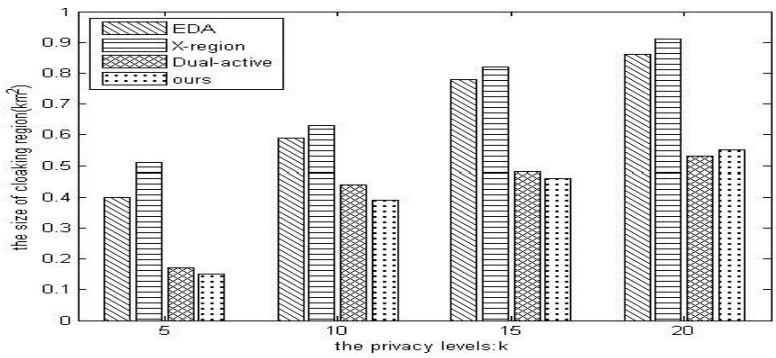


Figure 7: The Size of Generated Cloaking Region with Different k

5. Conclusion

In this paper, we propose a novel method of privacy-preserving LBS in mobile P2P network. Our program can solve the issue of exposing cooperators' location in traditional P2P architecture, also provides a personalized cloaking region for mobile users on the road by two mechanisms: mapping collaborators' location into an unrecovered index, and retrieving dummy locations based on an index. The experiment result turns out that our framework can reduce the

communication cost and anonymize time efficiently, but we don't consider the moving speed of users and the query efficiency at the server side, which will be our research focus in the future.

References

[1] L. Barkhuus, A. K. Dey. *Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns*[C]. Proc ACM INTERACT 2003, Zurich, Switzerland, July, 2003: 702-712

[2] B. Gedik, L. Liu. *Protecting location privacy with personalized k-anonymity: Architecture and algorithms*[J]. IEEE Transactions on Mobile Computing, 2008, 7(1): 1-18

[3] *T.* Wang, L. Liu. *Privacy-aware mobile services over road networks*[C]. Proceedings of the VLDB Endowment 2009, Lyon, France, August, 2009: 1042-1053

[4] K. Mouratidis, M. L. Yiu. *Anonymous query processing in road networks*[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(1): 2-15

[5] G. Ghinita, P. Kalnis, S. Skiadopoulos. *PRIVE: anonymous location-based queries in distributed mobile systems*[C]. Proceedings ACM World Wide Web'07, Alberta, Canada, May, 2007: 371-380

[6] R. Shokri, G. Theodorakopoulos, P. Papadimitratos, et al. *Hiding in the Mobile Crowd: Location Privacy through Collaboration*[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(3): 266-279

[7] A. Pingley, W. Yu, and N. Zhang, et al. *A context-aware scheme for privacy-preserving location-based services*[J]. Computer Networks Journal, 2012, 56(11): 2551-2568

[8] C. Y. Chow, M. F. Mokbel, X. Liu. *A peer-to-peer spatial cloaking algorithm for anonymous location-based service* [C]. Proc ACM International Symposium on Advances in geographic information systems (GIS'06), Arlington Virginia, USA, November, 2006: 171-178

[9] C. Y. Chow, M. F. Mokbel, X. Liu. *Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments*[J]. Advances of Computer Science for Geographic Information Systems, 2011, 15(2): 351-380

[10] G. Ghinita, P. Kalnis, S. Skiadopoulos. *MOBIHIDE: a mobile peer-to-peer system for anonymous location-based queries*[C]. Pro Advances in Spatial and Temporal Databases, Boston, USA, July, 2007: 221-238

[11] Y. Che, Q. Yang, X. Y. Hong. *A dual-active spatial cloaking algorithm for location privacy preserving in mobile peer-to-peer networks*[C]. Proc IEEE Wireless Communications and Networking Conference (WCNC 2012), Paris, France, April, 2012: 2098-2102.

[12] Y. Z. Che, K. Chiew, and X. Y. Hong, et al. *EDA: an enhanced dual-active algorithm for location privacy preservation in mobile P2P networks*[J]. Journal of Zhejiang University Science, 2013, 14(5): 356-373

[13] H. B. Hu, J. L. Xu. *Non-exposure location anonymity*[C]. Proc IEEE ICDE 2009, Shanghai, China, March, 2009: 1120-1131

[14] Y. Z. Che, Q. M. He, and X. Y. Hong, et al. *X-Region: A framework for location privacy preservation in mobile peer-to-peer networks*[J]. Communication Systems Journal, 2015, 28(1): 167-186