2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science

# *Semantic-Based Customizable Location Privacy Protection Scheme*

*Xin Lv*[*1], *Haitao Shi*[1]
[1]College of Computer and information, Hohai University
Nanjing, China
[*]Corresponding author: lvxin.gs@163.com

*Aili Wang*[2], *Tao Zeng*[1], *Zhongzhong Wu*[1]
[2]Information Center, Ministry of Water Resources
Beijing, China

***Abstract*—In recent years, with the popularization of mobile intelligent terminals, location-based services (LBS) have been widely used. When users enjoy the convenience of LBS, they also face with the risk of leakage of location privacy. Therefore, it is very important to provide effective privacy protection during service application. The previous methods of location privacy protection lack consideration of background information such as road network information and location semantics, which leads to the weakness of anonymity. For this reason, a semantic-based customizable location privacy protection scheme is presented in this paper. According to the road network environment, this scheme introduces the regional popularity and combines with the user-defined sensitivity to calculate the privacy of adjacent segments, then so as to obtain the anonymous candidate road segments efficiently until the anonymous requirements are satisfied. Experimental results show that anonymous set construction is efficient and the proposed scheme can effectively reduce the influence of relevant background knowledge on anonymity.**

***Keywords—Location Semantics; Privacy-preservation; Road Network; Popularity***

## I. INTRODUCTION

With the rapid development of mobile devices, applications based on Location Based Services (LBS) have become more and more widespread [1, 2]. In order to get services from LBS provider, the user have to share its personal location information, however, the services provider is not completely trusted in the most of scenario. Also, the location information can be stolen by the attackers, then more personal information, such as identity, behavior, and habits [3], maybe leaked by the mining methods. Therefore, the personal information should be protected during the process of interacting with the service provider.

The primary methods, such as $k$-anonymous algorithm, are usually designed in Euclidean space [4-7]. The ideal of $k$-anonymous algorithm [8] is to construct an anonymous set containing $k$ users, instead of the real location of the service requester, making the attacker hard to distinguish the objection from the set. However, with background knowledge possessed by the attacker, the security of the $k$-anonymous algorithm is compromised. Road network information is a kind of common knowledge to the publics, thus, it is necessary to design the privacy-preserving algorithm under road network environment. The knowledge of road network includes road information, also semantic information of hot locations on the road, such as shopping mall, hospital, and so on, which is related to users' privacy. Besides, the sensitivity is hierarchy, for instance, the semantic like hospital is more sensitive than shopping mall to the user.

Most of the current location privacy protection strategies are based on the Euclidean space, and the research based on the road network environment is not sufficient [9]. Chow [10] proposed a location-based privacy protection approach to the road network, which fuzzified the user's location into several adjacent road segments and considered the query cost and the query quality in constructing anonymous set. In 2006, Machanvajjhala A et al [11] proposed the concept of "$L$-diversity." In 2009, Ting Wang et al [12] applied "$L$-diversity" to the network environment and proposed the concept of "road segment diversity". In order to raise the difficulty of inferring attacks in the road network environment, they considered that the anonymous set must satisfy the $k$-anonymity and the $L$ road segments at the same time. Literature [13] aimed at the nonuniform distribution of users under the road network and the possible inference attacks, a location privacy protection algorithm for road network restriction is designed. By sorting the edge right of the road network and combining the geographical location distribution of the road, the construction of the hidden edge set is carried out to reduce the risk of inferring the attack by the nonuniform edge weight. However, none of these methods consider the influence of the semantic information of the location on the generated anonymous area in the road network environment.

Based on the research results above-mentioned, a privacy protection method based on location semantics in road network environment is proposed in this paper. On the premise of $k$-anonymity and $L$-diversity, the semantic information of the location is considered as an important evaluation factor to measure the privacy of anonymous space, and the user can customize the privacy requirements, which improves the efficiency and success rate of the construction of anonymous space.

## II. PRELIMINARIES

The location privacy protection method proposed in this paper aims to discuss the influence of the semantic information on location privacy protection and protect the semantic based location privacy protection under the road network environment. The definition of location semantic information involved in this paper is given as follow.

### *A. The central server architecture*

This paper is based on the central server architecture [14] (Figure 1). A trusted third party (anonymous server) exists in the client and the location server. Users send their own location, inquiry content, and privacy requirements to anonymous servers.

978-1-5386-7445-1/18/$31.00 ©2018 IEEE
DOI 10.1109/DCABES.2018.00047

The anonymous server sends the user's location after anonymity to the LBS server. The location server queries the candidate result and returns it to the anonymous server, the anonymous server analyzes the query candidate set, and returns the screening valid results to the requester. The anonymous server mentioned in this article needs to store the current map information as well as the road information (including the semantic information on the road network location), also, the moving users' information on the road need to be updated timely.

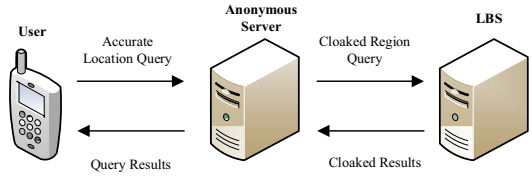


Fig. 1. The System Architecture

### B. Popularity and sensitivity

The paper discusses the influence of location semantics on location privacy protection. Based on the definition of location semantic information in literature [15], the influence of location semantics that are in the anonymous set is measured by popularity and sensitivity in constructing anonymous set, that makes the location semantics of the final anonymous set have minimum influence on location privacy. The definition of semantic information involved is given below.

In this paper, the road network is denoted as an undirected connectivity diagram $G=(V,E)$, $E=\{e_1,e_2,\cdots,e_m\}$ denotes the road segments in the road network, each road segment $Seg_i(sid,s,t)\in E$ is an edge in the road network, with $sid$ is the road segment number, $s$ and $t$ denote the starting and the end point of the road segment. $V=\{p_1,p_2,\cdots,p_n\}$ denotes the intersection of the road segment. Anonymous set $RS$ is comprised of multiple adjacent road segments $Segs=\{seg_1,seg_2,\cdots,seg_n\}$ and multiple mobile users $Users=\{user_1,user_2,\cdots,user_m\}$ on the road segments, in which the number of road segments $Segs$ and mobile users $Users$ should satisfy the personalized privacy demand of users.

**Definition 1** Location. $pos(lid,sid,x,y,tp_i)$ denotes the location in the road network, with $lid$ is the number of the location, $sid$ is the number of the road where the location is located, $(x,y)$ is the coordinate of the location, and $tp_i$ is the type of the location. The type of location is divided into $n$ types in total, and $TP=\{tp_1,tp_2,\cdots,tp_n\}$ is the set of $n$ location types.

**Definition 2** Location popularity. It is used to describe the popularity of a location type in the current road network. The popularity $pop_{tp_i}$ corresponding to the location of the type is

$$pop_{tp_i}=\frac{Numtp_i}{NumTP} \tag{1}$$

With $NumTP$ is the total number that the anonymous server system calculates of all types of locations $TP$ in the current network, $Numtp_i$ denotes the total number of each type of location $tp_i\in TP$. The set $POP=\{pop_{tp_1},pop_{tp_2},\cdots,pop_{tp_n}\}$ indicates the popularity of all location.

**Definition 3** Location sensitivity. It is used to describe the popularity of a location type in the current road network. For each user, corresponding to the same type of location sensitivity is different, the user according to their own circumstances, for each location type $tp_i\in TP$ set a sensitivity $sen_{tp_i}$, used to indicate the type of location for the user. The set $SEN_u=\{sen_{tp_1},sen_{tp_2},\cdots,sen_{tp_n}\}$ is the set of sensitivity of all location types relative to user $u$.

Based on definition 2 and definition 3, the popularity and the sensitivity of the region can be defined as follow:

**Definition 4** Regional popularity. The popularity $POP_{reg}$ of an area $reg$,

$$POP_{reg}=\sum_{i=1}^{|TP|}\frac{|pos.tp=tp_i|}{NumPos(reg)}pop_{tp_i} \tag{2}$$

**Definition 5** Regional sensitivity. The sensitivity $SEN_{reg}$ of an area $reg$,

$$SEN_{reg}=\sum_{i=1}^{|TP|}\frac{|pos.tp=tp_i|}{NumPos(reg)}sen_{tp_i} \tag{3}$$

$|TP|$ in Formulas (2) and (3) is the total number of types of positions contained in the region $reg$; and $NumPos(reg)$ is the number of positions contained in the region $reg$.

**Definition 6** Relative anonymous. The relative anonymity $RA_{RS}$ of anonymous set $RS$ is used to indicate the anonymity of the algorithm through the ratio of the number of mobile users $RS.UN$ after the anonymous algorithm is executed to the number of users $PR_u.UN$ in the privacy requirements of the user's location.

$$RA_{RS}=\frac{RS.UN}{PR_u.UN} \tag{4}$$

In the case of anonymous success, relative anonymity value greater than or equal to 1. As the relative anonymity is higher and higher, anonymous effect is better.

### C. Regional privacy definition

In the above, regional popularity and regional sensitivity are defined, and in the process of constructing the anonymous set, the regional privacy is defined to consider both of these two factors in the construction of the anonymous set. The privacy $PRM_{reg}$ of an area $reg$ is jointly determined by the regional popularity and the regional sensitivity, and its value is

$$PRM_{reg} = \frac{POP_{reg}}{SEN_{reg}} \quad (5)$$

That is to say that when the popularity of a region $reg$ is higher and the sensitivity is lower, the privacy of the region is higher, which means that the semantic information of the location in the region has the least influence on the privacy of the user. When constructing anonymous set, road segments and mobile users are selected to add to the anonymous set and should ensure that the added set makes the regional privacy of the current anonymous set is largest.

### D. Personalized privacy requirement

The method of constructing anonymous set in this paper is to allow user to customize personalized privacy requirements. When submitting a query request, user can customize the privacy requirements by submitting the lowest number of the users and the road segments. When constructing anonymous set, the number of the users and the road segments in anonymous set must satisfy user's privacy requirements at the same time, then, the anonymity can be considered as successful.

(1) The anonymous set contains the lowest number of mobile users $(RS.UN)$. It is required that at least $RS.UN-1$ mobile users that can't be distinguished from the current query user in this anonymous set $RS$. The method is mainly derived from the classical algorithm $k$-anonymity algorithm in location privacy protection.

(2) Anonymous set contains the lowest number of road segments $(RS.SN)$. It requires that the number of road segments can't be less than the user-set value $RS.SN$ in the anonymous set $RS$. When constructing the anonymous set, if the anonymous set contains too few sections, for example, if the constructed anonymous set contains only one road segment, then even if the number of users in the anonymous set satisfies $k$-anonymity, the attacker can easily determine the current user's road segment, greatly reducing attacking difficulty.

**Definition 7** User personalized privacy requirement. For a user $u$ that makes a query, his privacy requirements are expressed in $PR_u(RS.UN_u, RS.SN_u, SEN_u)$. In which, $RS.UN_u$ denotes the user-defined lowest number of anonymous mobile users; $RS.SN_u$ denotes the user-defined lowest number of anonymous road segments; and $SEN_u$ is the user-defined sensitivity of a group of different location types.

## III. Semantic-Based Customizable Location Privacy Protection Scheme

The method proposed in this paper defines and constructs an anonymous set based on the premise that an attacker has the following background knowledge: (1) the attacker owns the information of the current road network, road segments and the semantic information of the road segments; (2) the attacker has the location information of the mobile users in the current road network, but does not know each mobile user's identity; (3) the attacker can get the number of mobile users on the current road segment; (4) the attacker does not have the background knowledge of mobile users in the current road network. In addition, this article assumes that third-party servers, anonymous servers, is trustworthy to users and will not be compromised by attackers.

### A. Construct anonymous set

The anonymous set construction algorithm proposed in this paper consists of two algorithms. It solves the problems of determining which road segments are optional and how to select an optimal set of road segments to be added to the current anonymous set. The algorithm 1 solves the problem of how to select an optimal road segment set to add to the current anonymous set. The main ideals are: a set of road segments is selected from the set of candidate road segments each time (a set of road segments is a set of all the road segments that the vertex points at the same point) and is added to the current anonymous set. According to the location sensitivity $SEN_u$ of the user's privacy requirements, calculate the regional privacy of the anonymous set after each set of road segments is added, and select the set of road segments that makes the regional privacy of the current anonymous set is maximum as the optimal road segments set to add to the current anonymous set.

The algorithm 2 solves the problems of determining which road segments sets are optional and how to meet user's privacy requirements. The main ideals are: take the road segment of the current user as an initial anonymous set, the road segments set adjacent to the current anonymous set is set as candidate road segments. A set of optimal road segments is selected from the set of candidate road segments to add to the current anonymous set, and determine whether the number of mobile users and road segments in the current anonymous set meet the user's privacy requirements, if satisfied, the current anonymous set is taken as the final anonymous set, if not satisfied, the set of adjacent road segments of the current anonymous set is calculated again as candidate road segments set, and the algorithm 1 is halted until meet the user's privacy requirements. The steps of the algorithm 1 and the algorithm 2 are as follows.

Algorithm 1 is the optimal road segment set selection algorithm. The input parameters of the algorithm are the current anonymous set $CRS$, the set of candidate segments that is the set of adjacent segments $SEGS$ and the set of sensitivity $SENS$ that defined by user, the returned result is the set of optimal road segments $BESTSEGS$. The concrete steps are as follows:

(1) Initialize the input parameters;

(2) Add each of the candidate road segment sets to the current anonymous set.

(3) Calculate the regional popularity, regional sensitivity and regional privacy of the anonymous set after adding the set of road segments, and take the road segment with the largest degree of regional privacy as the current set of optimal road segments, that is, the location semantic information of the added road segments has the least influence on the anonymity effect;

(4) Return to the optimal set of road segments.

The pseudo-code of the algorithm is as follows:

**Algorithm 1** Optimal road segment set selection algorithm (OPTSEGS)

Input: the current anonymous set $CRS$, the candidate set of road segments $SEGS$, the sensitivity set $SENS$.

Output: the optimal road segment set $BESTSEGS$.

(1) $BESTSEGS = \varnothing$; $MAX = 0$;

(2) for each $Segs$ in $SEGS$
(3) assign $Segs$ to $e$;
(4) $PR = POP_{(CRS\cup e)}$;
(5) $SR = SEN_{(CRS\cup e)}$;
(6) $MR = \frac{PR}{SR}$;
(7) $MAX$ records the current maximum value of $MR$, and the corresponding $e$ assigned to $BESTSEGS$;
(8) end for
(9) return $BESTSEGS$

Algorithm 2 is an anonymous set construction algorithm based on location semantics. The algorithm's input parameters include user $U$, user's query request $Q$, and privacy requirements $PR_u$ for customization. The custom privacy requirements include the lowest number of mobile users $PR_u.UN$ and road segments $PR_u.SN$ contained in the anonymous set, and the given sensitivity $PR_u.SEN_u$ of each location type relative to himself. The concrete steps are as follows:

(1) Initialize the input parameters;
(2) Add the current user's road segment to the current anonymous set;
(3) Determine whether the current anonymous set meets the user privacy requirements, if the requirements are satisfied, the loop is ended; if the requirements are not satisfied, the loop continues;
(4) Calculate the set of adjacent road segments of the current anonymous set;
(5) Execute algorithm OPTSEGS, add the result to the current anonymous set and execute algorithm step 3;
(6) Return to the anonymous set.

The pseudo-code of the algorithm is as follows:

**Algorithm 2** Anonymous Set Construction Algorithm Based on Location Semantics (Enhance-LSBASC)

Input: User $U$, the user's query request $Q$, user-defined privacy requirements $PR_u$

Output: Anonymous set $RS$

(1) $RS = \varnothing$;
(2) The user where the road assignment to $BestEdge$;
(3) Put $BestEdge$ into anonymous set $RS$;
(4) while $NumUser(RS) < PR_u.UN$ or $NumSeg(RS) < PR_u.SN$
(5) $R = FindEdges(RS)$;
(6) $BestEdge = OPTSEGS(RS, R, PR_u.SEN_u)$;
(7) Put $BestEdge$ into anonymous set $RS$;
(8) $R = \varnothing$;
(9) end while
(10) return $RS$

Next, take User as an example to illustrate the above process. For convenience of analysis, only 6 road segments are shown here, as shown in Figure 2. A pair of values $(Seg_i, num)$ are used to indicate the number of road segments and the number of mobile users on the road segments. The location of the User in the figure is identified by the arrow. Set up $TP = \{tp_1, tp_2, tp_3, tp_4\}$, which, $tp_1$ is hospital, $tp_2$ is bar, $tp_3$ is shopping mall, $tp_4$ is school. Set up $pop_{tp_1} = 0.3$, $pop_{tp_2} = 0$, $pop_{tp_3} = 0.4$, $pop_{tp_4} = 0.3$. The user-defined privacy requirements are $PR_u(RS.UN_u, RS.SN_u, SEN_u)$, which, $RS.UN_u = 50$, $RS.SN_u = 3$, $SEN_u = \{0.5, 0.3, 0.2, 0\}$. According to the algorithm steps, we first add the $Seg_1$ of User in the anonymous set $RS$, because $NumUser(RS) = 10$ and $NumSeg(RS) = 1$ do not meet the privacy requirements, continue to perform the algorithm steps. The set of $\{Seg_2, Seg_3\}$ and $\{Seg_4\}$ that adjacent to $Seg_1$ are added to the set $R$, $PRM_{(RS\cup\{Seg_2, Seg_3\})} = 1.737$ and $PRM_{(RS\cup\{Seg_4\})} = 1.133$ are calculated respectively according to the regional privacy metrics. $\{Seg_2, Seg_3\}$ is added to the anonymous set $RS$ according to the calculation results. At this point, $NumUser(RS) = 60$, $NumSeg(RS) = 3$, to satisfy the privacy requirements, the algorithm ends and returns $RS = \{Seg_1, Seg_2, Seg_3\}$ as an anonymous set of User.

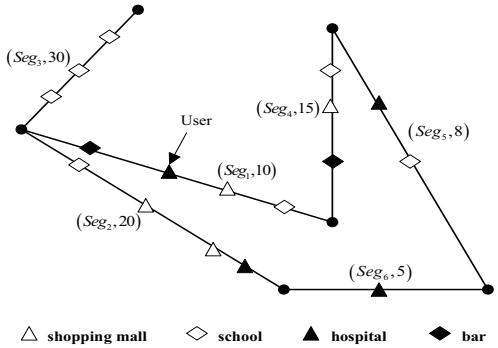


Fig. 2. The example

## IV. EXPERIMENT AND ANALYSIS

The environment of the experiment is Intel (R) Core i5-6300HQ CPU @ 2.30GHz; 8GB RAM; operating system is Microsoft Windows 10 Professional; and the algorithm is written in Java based on Eclipse environment.

### *A. Experimental data sets and parameter settings*

The experimental data contents two parts. The first part is the highway network data in Oldenburg, Germany [16], including 6105 roads and 7035 road vertices. The second part is 10000 uniform distribution mobile users obtained from Brinkhoff based network mobile object generator [17], by introducing the highway network of Germany Oldenburg city into Brinkhoff generator, the mobile users are distributed on the road segments, and a specific semantic information definition is labeled on the attribute of location type in location data generated by Brinkhoff generator, including 4 kinds of location semantics (hospital, bar, shopping mall and school).

The parameters in the experiment include the number of mobile users, the user-defined privacy requirements $(PR_u.UN, PR_u.SN, PR_u.SEN_u)$, the maximum number of experimental cycles $PR_u.Cycle_{\max}$, the number of locations, and the number of users that send out service request. The experiment randomly selects 1000 mobile users who request

service to simulate experiments. In real life, the number of privacy-required users and road segments cannot be increased without limit, so set a default value and a range of evaluation in the experiment. Considering the time complexity and quality of service, the maximum cycle number of the algorithm is set in the experiment. All experimental parameters in the experiment set as shown in Table 1.

TABLE I. PARAMETER SETTINGS

| ***Parameter*** | ***Defaults*** | ***Evaluation range*** |
|---|---|---|
| the number of mobile users | 10000 | |
| $PR_u.UN$ | 25 | 15~35 |
| $PR_u.SN$ | 6 | 3~15 |
| $PR_u.Cycle_{\max}$ | 20 | |
| $PR_u.SEN_u$ | user custom | |
| the number of locations | 10000 | |
| the number of users that request service | 1000 | |

## *B. Analysis of experimental results*

The experiment compares and evaluates the proposed Anonymous Construction Algorithm (Enhance-LSBASC) with the Anonymous Construction Algorithm (LSBASC) proposed in literature [15] from the aspects of anonymity success rate, average anonymous execution time and relative anonymous.

(1) Anonymous success rate. Figure 3 and Figure 4 show the comparison of the anonymous success rate between the algorithm Enhance-LSBASC and algorithm LSBASC in different privacy requirements, the number of mobile users and road segments. From the experimental results in Figure 3, it can be seen that when the number of privacy-required users is within the range of the default value (25), the anonymous success rates of the algorithm Enhance-LSBASC and the algorithm LSBASC are basically the same, because the number of privacy-required users is within the default range, the maximum tolerated road segments set in LSBASC and the maximum number of cycles set in Enhance-LSBASC of the algorithm will not reach the upper limit, so the anonymous success rates of the algorithm Enhance-LSBASC and the algorithm LSBASC are basically the same; when the number of privacy-required users is larger than the default value, because algorithm LSBASC chooses the best one to join the anonymous set when selecting the optimal road segment, when the number of added road segments reaches the set upper limit of the road segment tolerance, that is, the maximum number of cycles of the optimal road segment selection algorithm , the number of anonymous mobile users can't meet the privacy needs, resulting in anonymous failure. The algorithm Enhance-LSBASC selects the optimal road segment set to join the current anonymous set each time, and each time the number of road segments added to the anonymous set and the number of mobile users is more than that of the algorithm LSBASC, so the algorithm Enhance-LSBASC has higher anonymity success rate than the algorithm LSBASC when the number of privacy-required users is increasing. It can be seen from Figure 4, since the algorithm Enhance-LSBASC selects the optimal road segment set to join the current anonymous set each time and the algorithm LSBASC selects a single optimal road segment to join the current anonymous set each time, the LSBASC algorithm has a high success rate. When the number of privacy-required users remains unchanged, the success rate of anonymity will not change as long as the number of road segments is within the number of cycles tolerated.

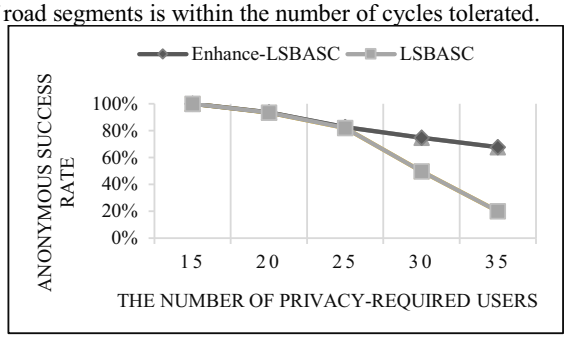


Fig. 3. Anonymous success rate under the different number of privacy-required users

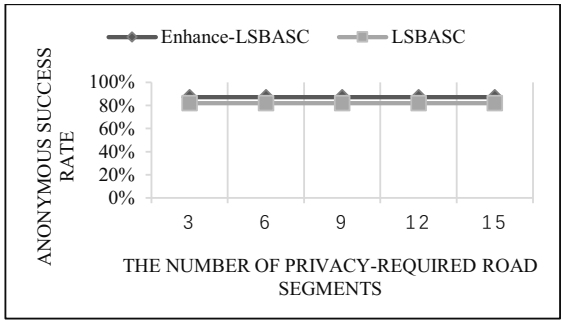


Fig. 4. Anonymous success rate under the different number of privacy-required road segments

(2) Average anonymous execution time. Figure 5 and Figure 6 show the comparison between the algorithm Enhance-LSBASC and the algorithm LSBASC for the average anonymous execution time of the number of the mobile users and road segments with different privacy requirements. From the experimental results of Figure 5, it can be seen that the average anonymous time of the algorithm Enhance-LSBASC is less than the algorithm LSBASC, and with the increase of the number of privacy-required users, the average anonymous execution time is getting bigger and bigger. This is because as the number of privacy-required users increases, the number of adjacent road segments in the set of candidate road segments increases more and more. When the algorithm selects the optimal road segment, the number of cyclic comparison increases. The algorithm LSBASC selects the optimal road segment each time to join the current anonymous set, and the algorithm Enhance-LSBASC selects the optimal road segment each time. When the final road segment is selected, the number of cyclic comparison is less than the algorithm LSBASC. Therefore, the average execution time of algorithm Enhance-LSBASC is smaller than that of algorithm LSBASC. From the experimental results of Figure 6, when privacy requires a certain number of users, the average anonymous time of the algorithm Enhance-LSBASC is better than the algorithm LSBASC with the increasing number of privacy-required road segments, and the average anonymity time is invariable. The main reason is that

10000 mobile users are evenly distributed on the road when the number of privacy-required users is fixed, so in order to satisfy the number of privacy-required users, the number of road segments in the anonymous set must reach a certain number. Therefore, the average anonymous time tends to constant.

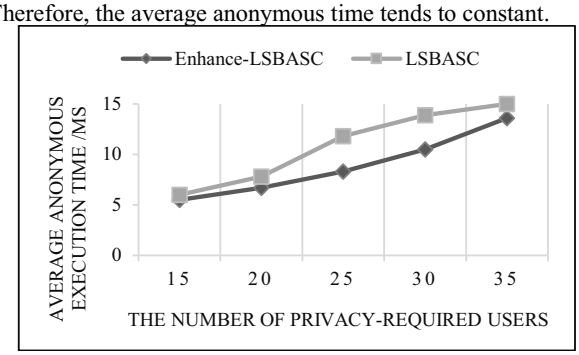


Fig. 5. Average anonymous execution time under the different number of privacy-required users

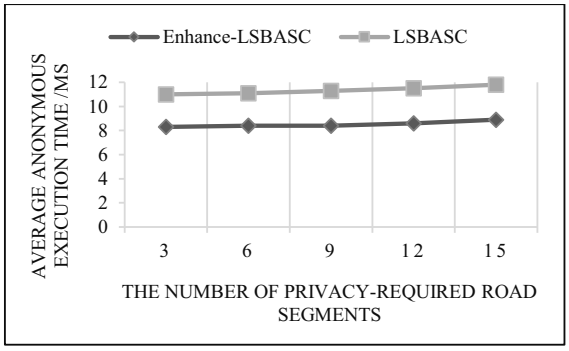


Fig. 6. Average anonymous execution time under the different number of privacy-required road segments

(3) Relative anonymous. Figure 7 shows the relative anonymity of the algorithm Enhance-LSBASC and the algorithm LSBASC for different numbers of privacy-required users. The results show that the relative anonymity of algorithm Enhance-LSBASC is higher than that of algorithm LSBASC, but with the increasing number of privacy-required users, the relative anonymity of algorithm Enhance-LSBASC and algorithm LSBASC will eventually tends to be gentle. The relative anonymity of the algorithm Enhance-LSBASC is higher than that of the algorithm LSBASC mainly because the algorithm Enhance-LSBASC selects the optimal road segment set to join the anonymous set each time, and the algorithm LSBASC selects the optimal road segment to join the anonymous set each time. So, the algorithm Enhance-LSBASC has more users than the algorithm LSBASC in the anonymous set. The relative anonymity of the algorithm Enhance-LSBASC shows a decreasing trend because the optimal set of road segments selected by the Enhance-LSBASC algorithm is added to the anonymous set. This leads to the occurrence that constructs the same anonymous set by two adjacent privacy-required users, and with the increasing number of privacy-required users, the occurrence of anonymous failures will be more and more, and in this case, the relative anonymity is less than 1, so the relative anonymity will decrease with the increasing number of privacy-required users, the final relative anonymity of the algorithm Enhance-LSBASC and algorithm LSBASC will tend to be gentle, because when the number of privacy-required users is too high, the optimal choice of algorithm will reach a loop limit, the anonymous success rate will tend to be gentle, making the relative anonymity tends to a stable value.

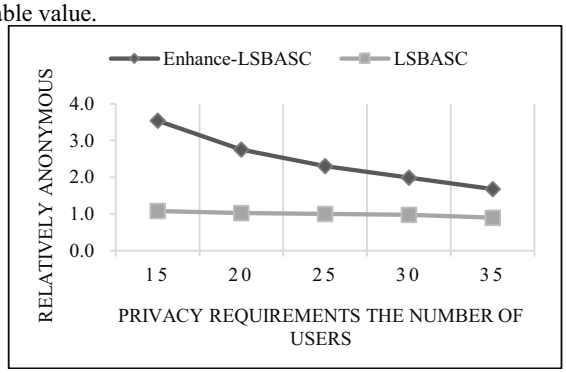


Fig. 7. The relative anonymity under the different number of privacy-required users

## V. CONCLUSION

A semantic-based customizable location privacy protection scheme is proposed in this paper, which allows users to customize the level of privacy protection. Considering the semantic-related background knowledge will cause privacy leaks, the concepts of regional popularity, regional sensitivity and regional privacy are introduced in the process of constructing anonymous set, which can effectively resist the semantic based attack. The simulation experiments show that the proposed method performs better than the existing method on efficiency.

## ACKNOWLEDGMENT

This work is partially supported by "The National Key Research and Development Program of China" (Grant No. 2016YFC0400910); "General Program of National Natural Science Foundation of China" (Grant No. 61272543); "NSF-China and Guangdong Province Joint Project" (Grant No. U1301252); "The Fundamental Research Funds for the Central Universities" (Grant No. 2016B11714).

## REFERENCES

[1] Wang B, Yang X, Wang G, et al. Energy efficient approximate self-adaptive data collection in wireless sensor networks[J]. Frontiers of Computer Science, 2016, 10(5): 936-950.

[2] Ying Z A, Shanghai. Location-Based Services: Architecture and Progress[J]. Chinese Journal of Computers, 2011, 34(7): 1155-1171.

[3] Mokbel M F. Privacy in Location-Based Services: State-of-the-Art and Research Directions[C]// International Conference on Mobile Data Management. IEEE, 2008: 228.

[4] Niu B, Li Q, Zhu X, et al. Achieving k-anonymity in privacy-aware location-based services[J]. Journal of Graph Algorithms & Applications, 2016, 20(2): 363-410.

[5] Gruteser M, Grunwald D. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking[C]// International Conference on Mobile Systems, Applications, and Services. DBLP, 2003: 31-42.

[6] Kalnis P, Ghinita G, Mouratidis K, et al. Preventing Location-Based Identity Inference in Anonymous Spatial Queries[J]. IEEE Transactions on Knowledge & Data Engineering, 2007, 19(12): 1719-1733.

[7] Lee H J, Hong S T, Min Y, et al. A new cloaking algorithm using Hilbert curves for privacy protection[C]// ACM Sigspatial International Workshop on Security and Privacy in Gis and Lbs. ACM, 2010: 42-46.

[8] L. Sweeney. k-anonymity: A model for protecting privacy[J]. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 2002, 10(5): 557-570.

[9] Min L I, Qin Z G. Survey of location privacy protection over road networks[J]. Application Research of Computers, 2014, 16(16): 1-10.

[10] Chow C Y, Mokbel M F, Bao J, et al. Query-aware location anonymization for road networks[J]. Geoinformatica, 2011, 15(3): 571-607.

[11] Machanavajjhala A, Kifer D, Gehrke J. L -diversity: Privacy beyond k -anonymity[J]. Acm Transactions on Knowledge Discovery from Data, 2007, 1(1): 3.

[12] Wang T, Liu L. Privacy-Aware Mobile Services over Road Networks.[J]. Proceedings of the Vldb Endowment, 2010, 2(1): 1042-1053.

[13] Sun Lan, Luo Jin, Wu Ying-Jie. An algorithm for protecting location privacy in road network[J]. Journal of Shandong University(Engineering Science), 2012, 42(5): 1042-1053.

[14] Cheng R, Zhang Y, Bertino E, et al. Preserving User Location Privacy in Mobile Data Management Infrastructures[J]. Proc of Workshop on Privacy Enhancing Technologies, 2006, 4258: 393-412.

[15] Chen Hui, Qin Xiao-Qi. Location-semantic-based location privacy protection for road network[J]. Journal on Communications, 2016, 37(8): 67-76.

[16] Chen S, Jensen C S, Lin D. A benchmark for evaluating moving object indexes.[J]. Proceedings of the Vldb Endowment, 2008, 1(1): 1574-1585.

[17] Brinkhoff T. A Framework for Generating Network-Based Moving Objects[J]. Geoinformatica, 2002, 6(2): 153-180.