A Generalized Location Privacy Protection Scheme in Location Based Services

Jing-Jing Wang[1(✉)], Yi-Liang Han[1], and Jia-Yong Chen[2]

[1] Key Laboratory of Network and Information Security of APF, Engineering College of APF,
Xi'an 710086, China
344505421@qq.com, yilianghan@hotmail.com

[2] Laboratory of Information and Network of APF, Engineering College of APF,
Xi'an 710086, China
260840527@qq.com

Abstract. When the user getting location based services by the traditional technology, his location information of region is always be exposed. However, in modern mobile networks, even the current geographical region is a part of privacy information. To solve this problem, a new generalized k-anonymity location privacy protection scheme in location based services (LPPS-GKA) with the third trust servicer is proposed. And it can guarantee the users get good location-based services (LBS) without leaking the information of the geo-location region, which has protected the perfect privacy. Analysis shows that LPPS-GKA is more secure in protecting location privacy, including region information, and is more efficient than other similar schemes in computational and communicational aspects. It is suitable for dynamic environment for different user's various privacy protection requests.

Keywords: Location privacy protection · Generalized k-anonymity · Location based service

1 Introduction

Nowadays, social communication method has evolved dramatically along with the big data era and the development of data networks. So the new secure schemes or systems should be 'application-oriented' with strong reality backgrounds. And more and more people join in the social networks to share sources and information, and generate a lot of data, namely big data, at the same time. What's more, the users are always not willing to let other people know these data. And the location information is one of the most important contents of the user's privacy data, which should be protected properly. However, most of the new-type applications of the mobile social networks, especially the location based services, are always publishing the user's location information at any time [1], and the traditional opinion of anonymous privacy protection is not valid any more in modern society.

As a result, some newer methods for location data preservation are proposed successively for protecting location privacy comprehensively, such as the technology of

W. Chen et al. (Eds.): BDTA 2015, CCIS 590, pp. 211–217, 2015.
DOI: 10.1007/978-981-10-0457-5_20

randomization [2], space-vagueness [3] and time-vagueness [4]. In these methods mentioned above, space-vagueness technology has its dominant position in real applications in big data and social networks relatively because of its moderate computation cost and easy realization.

2 Related Works

In the modern social networks, the researches on location-preserving have gotten lots of outcomes. Just as described in paper [5], the location privacy protection technology based on k-anonymity has made some remarkable developments in these outcomes.

k-anonymity is one of the specific ways to realize space-vagueness. When publishing data, the real data should be disposed first, then published together with other $(k - 1)$ data simultaneously [6]. k-anonymity technology is first used in location-privacy preserving by Gruteser in 2003 [7], which means to fuzzy the user into k adjacent access points in one region. But it cannot protect the located region from leaking when each space-vagueness. From then on, it has been used widely in the location privacy protection and gotten great developments.

However, the research about how to protect location data in location based services has a rather late start. Since 2011, Huang Yi, etc. proposed a method for location privacy preserving in location services, but the time for user to get location services increases greatly when the privacy request k increasing [8]. And it is also supported by querying adjacent points in one region, so the geo-location region is exposed too.

In 2013, Damiani and Cuijpers first pointed out the issue that the specific information of user's region is also an important part of location privacy [9]. And they proposed one protection scheme based on privacy policy, but its efficiency is not high enough for mobile Internet. In 2014, Peng etc. proposed a method to judge the location privacy attacks according with the region located [10]. Then in 2015, Wang etc. proposed a location privacy protection approach named KAP, which aimed at the privacy issue of location service under the mobile Internet combing the concept of k-anonymity [5]. The approach had stronger security, but it needs a number of other access points' data around the user during every location service, which may leak the user's real located region again.

To achieve the perfect location privacy, we must protect the specific location. On the other hand, we must pay attention to the content about the located region including the user too. And the method realizing this idea should be feasible and efficient.

3 Our Scheme

In the existing outcomes in researches on the issue, most of schemes about location privacy protection are based on the center server and distributed P2P structure [11]. So in this paper, we will discuss about this type protection schemes to realize perfect privacy.

3.1 Basic Definition

Definition 1: CRT (Chinese Remainder Theorem): Suppose $m_1, m_2 \cdots m_k$ are positive integers that are pairwise co-prime. Then, for any given sequence of integers $a_1, a_2 \cdots a_k$, there exists an integer x solving the following system of simultaneous congruences.

$$\begin{cases} x \equiv a_1 \bmod m_1 \\ x \equiv a_2 \bmod m_2 \\ \qquad \cdots \\ x \equiv a_k \bmod m_k \end{cases} \tag{1}$$

If $M = m_1 m_2 \cdots m_k$, $M_i = \frac{M}{m_i}$ and $M_i M_i^{-1} = 1 \bmod m_i$, we can computer the solution set of the equations above:

$$X = M_1 M_1^{-1} a_1 + M_2 M_2^{-1} a_2 + \cdots + M_k M_k^{-1} a_k + K * M, \ \ K \in Z \tag{2}$$

Obviously, the original number x is included as one element in the solution set X.

K is an integer in Eq. (2). Choose K as different values, we can get different solutions to satisfy all the equations in Eq. (1).

Definition 2: Generalized k-Anonymity (GKA): In location based services, if one user's accurate location data is extend to k access points, and: (a) It is not necessary that these k access points are adjacent neighbours in one region and interact with the user during location. (b) These k access points must belong to a equivalence class, which means for the attacker, he can not tell which one is the user's real location data.

It can be defined as generalized *k-anonymity*.

From the definition above, we can see that GKA is more scientific and practical than the traditional definition in the background of big data and complex social networks.

3.2 LPPS-GKA

At present, most of the researches based on *k-anonymity* used the trusted third party, namely center servicer [8]. The main function of the center servicer is anonymity and agency query. As shown in Fig. 1, we begin with step 1:

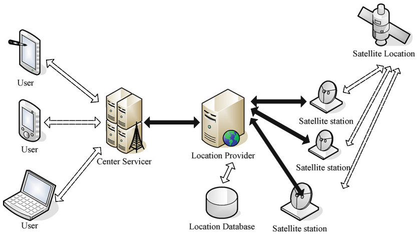

Fig. 1. Illustration for scheme with center servicer

Here,

x: The longitude-coordinate of the user's location;
y: The latitude-coordinate of the use's location;
t: The time when the user asks for service;
v: The velocity when the user asks for service;
con.: The query content of the user;
k: The privacy protection request parameter wanted by the user;
K, K':$K, K' \in Z$ It is the integer chosen in the process of CRT computation.

Step 1:
Query: The user sends his location information, queries contents and privacy request to the center anonymous servicer;

$$Q = (x, y, t, v, con., k)$$

Step 2:
GKA: The center servicer first extends the accurate location coordinate of user into an equivalence set including *k* different elements through CRT. Then it sends the query content and these *k* elements as a query set to the location provider.

- Choose $m_1, m_2, \cdots m_k$ randomly, satisfying the refine condition: $m_1, m_2, \cdots m_k \in Z^+, gcd\left(m_i, m_j\right) = 1, i, j \in Z^+$, and if the user doesn't want to leak his region information, make sure $m_1 * m_2 * \cdots m_k \gg 10^4$ (because we take the precision of coordinates up to four decimal places), then the results returned are far apart from each other; Or else, if the user doesn't care about his current area information, $m_1, m_2, \cdots m_k$ can be chosen randomly even its multiply is very small. So as to realize generalized *k-anonymity* defined in our paper, which is more scientific than the traditional definition. We give a toy example for the results from GKA, which is illustrated by Fig. 2.
- Send the query message *KAC* to the location provider

$$KAC = \left((x_1, y_1), \left(x_2, y_2\right), (x, y), con., t\right)$$

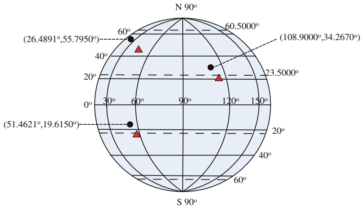

Fig. 2. Illustration for GKA

Step 3:
LBS: The location provider offers the query results set QC to the center servicer:

$$QC = \left(con_1., con_2., con., t^{'}\right)$$

The center servicer finds the accurate result $con.$ in the set, and computes r according the time interval and velocity. Through $r = v(t^{'} - t)$ and the original location of the user to judge the current area with radius r, then sends $con.$ to the user by locating this area. In this way, LPSS-GKA is completed with a center servicer.

4 Discussion and Analysis

The privacy protection technology of location big data not only needs to protect the user's location data, but also to balance the feasibility of services and overheads [12]. In this chapter, we will discuss the features and performance of our schemes from 3 aspects respectively: (1) Geo-indistinguishability; (2) Survivability of services; (3) Overheads in computation and communication.

4.1 Geo-Indistinguishability

Theorem 1. LPSS-GKA - has realized geo-indistinguishability.

Proof. As the user asks for service every time, we get an equivalence class including his real location data, and in this equivalence class, there are k different elements with equivalent relationship with each other. And in theory, these k elements have the uniform probability to be chosen for the attacker. Even in the continual services during a period of time, the attacker can't identify with greater advantage. Because in fact, the different regions located successively form an equivalence class too. So we can say it realized the objective of geo-indistinguishability.

4.2 Feasibility of Services

Theorem 2. LPPS-GKA is feasible. It supports high quality of services, reliability of query results.
We prove its feasibility mainly through two standards used universally: (a) Quality of services; (b) Reliability of query results.

(a) ***Quality of services***

Definition 3. The quality of service can be judged by the proportion of the number of the successful privacy requests $n^{'}$ as k changed in the whole number of all the privacy requests n, namely the success rate of anonymity. Its math expression is:

$$R_{SA} = \frac{n'}{n} \times 100\% \left(n' \le n\right)$$

Proof. Because we used the classic math tool CRT to realize *k-anonymity*, when k is changed by user, the only additional thing need to do is choose more or less K values in the solution set according to different k.

As we all known, the number of integers is countless, which means $R_{SA} \approx 1$ if the servicer and the terminal devices run normally.

(b) ***Reliability of query results***

Definition 4. The reliability of query results can be judged by the relationship of the distance between the user's original location p and the current location p_1 when receiving query result after a time interval Δt, and the radius of user's location region.

- For LPSS-GKA $r = v(t' - t) = v * \Delta t$. If $\frac{|pp_1|}{v*\Delta t} \le 1$, we say the query result is reliable and accurate. Or else, the query result can't be returned to the user's hand, because the user's location has gone beyond the communication range accepted by the center servicer. In this case, the location based service is invalid, but this case can be neglected if we choose proper parameters.

4.3 Performance

High efficiency is the main advantage for our scheme to real application in social networks or mobile networks.

(a) In LPSS-GKA, the space-vagueness degree, namely the privacy request k can be changed easily and smoothly without adding more computation overheads. Shown as Fig. 3, even k becomes larger, the overhead of computation and communication of both center servicer and user's terminal devices is almost fixed because the high efficiency of CRT.

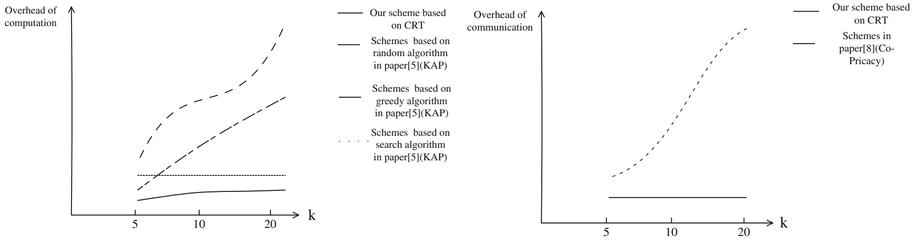

Fig. 3. Illustration for the overheads comparison of our scheme and classic schemes

When asking for location based services, the users in our scheme needn't to interact with other adjacent users for several times any more like most of the existing schemes to get other fuzzy access points. Therefore, it can spare more communication time and network source.

5 Conclusion and Future Work

A new LPPS-GKA for space-vagueness with center servicer in location services based on CRT is proposed in this paper. The scheme can be proved to get good LBS for users without leaking the information of the user's location region. Besides, it can meet different location privacy preserving requests of users with high efficient. When the privacy request is higher, namely k is larger, the time increased can be neglected.

Since the third trust servicer is not so trustable, so the future work is to design a better protection scheme without third center servicer.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China (Nos. 61272492 and 61572521), Shaanxi Province Natural Science Foundation of China (No. 2015JM6353) and the Basic Foundation of Engineering University of CAPF (No. WJY201521).

References

1. Jabeur, N., Zeadally, S., Sayed, B.: Mobile social networking applications. Commun. ACM **56**(3), 71–79 (2013). doi:10.1145/2428556.2428573
2. Suzuki, A., Iwata, M., Arase, Y., Hara, T., Xie, X., Nishio, S.: A user location anonymization method for location based services in a real environment. In: Proceedings of the 18th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, San Jose, pp. 398–401 (2010). doi:10.1145/1869790.1869846
3. Gredik, B., Liu, L.: Protecing location privacy with personalized k-anonymity: architecture and algorithms. IEEE Trans. Mob. Comput. **7**(1), 1–18 (2008). doi:10.1109/TMC.2007.1062
4. Pan, X., Xu, J., Meng, X.: Protecing location privacy against location-dependent attacks in mobile services. IEEE Trans. Knowl. Data Eng. **24**(8), 1506–1519 (2012). doi:10.1109/TKDE.2011.105
5. Wang, Y., Zhang, H., Yu, X.: KAP: location privacy-preserving approach in location services. J. Commun. **35**(11), 182–190 (2014). (in Chinese)
6. Wernke, M., Skvortsov, P., Durr, F., et al.: A classification of location privacy attacks and approaches. Pers. Ubiquit. Comput. **18**(1), 163–175 (2012)
7. Gruteser, M., Grunwald, D.: Anonymous usage of location-based services through spatial and temporal cloaking. In: Proceedings of the 1st International Conference on Mobile Systems, Applications and Services (MOBISYS 2003), San Francisco, California, pp. 31–42 (2003)
8. Huang, Y., Huo, Z., Meng, X.F.: Coprivacy: a collaborative location privacy preserving method without cloaking region. Chin. J. Comput. **34**(10), 1977–1985 (2011). (in Chinese)
9. Damiani, M.L., Cuijpers, C.: Privacy challenges in third-party location services. In: IEEE 14th International Conference on Mobile Data Management (MDM 2013), Milan, Italy, pp. 213–225 (2013)
10. Peng, Z.T., Kaji, K., Kawaguchi, N.: Privacy protection in Wi-Fi based location estimation. In: The 7th International Conference on Mobile Computing and Ubiquitous Networking (ICMU 2014), Singapore, pp. 62–67 (2014)
11. Yang, S.-T., Ma, C.-G., Zhou, C.-L.: LBS-oriented location privacy protection model and scheme. Chin. J. Commun. **35**(8), 116–127 (2014)
12. Wang, L., Meng, X.F.: Location privacy preservation in big data era: a survey. J. Softw. **25**(4), 693–712 (2014). doi:10.13328/j.cnki.jos.004551. (in Chinese)