A Full Privacy-Preserving Scheme for Location-Based Services

Fei Shao, Rong Cheng and Fangguo Zhang*

School of Information Science and Technology,
Sun Yat-sen University, Guangzhou 510006, China
tianxin0120@163.com, chengrongada@163.com, isszhfg@mail.sysu.edu.cn

Abstract. Location based services(LBS) pose risks to user's privacy, as they have access to user's identity, location and usage profile. Many approaches have been made up to deal with the privacy problems. But few of them meet the requirement of full privacy. In this paper, we propose a protocol that does not require a trusted third party and provides full privacy. We use group anonymous authentication to fulfill identity privacy, while using program obfuscation to satisfy the privacy requirement of usage profile. And we assume that there exist some geography or geometry methods to form a cloaking region to meet location privacy.

Keywords: privacy, location based service, anonymous credential system, program obfuscation.

1 Introduction

Location based services involve the collection, using, and sharing of location data[6]. Which may pose a great risk to user's privacy[4]. The malicious location services providers(LP) may do something against the users' willing. Users may have the feeling of being followed up and queries may disclose sensitive information about individuals. So privacy issue is something that must be concerned[1].

All the exiting privacy preserving techniques can be divided into three categories, that is two-tier spatial transformations, three-tier spatial transformations and cryptographic transformations[6]. Methods in Category 1 do not require any trusted third party, and the query anonymization is performed by the mobile user itself. [12, 5] are all of that category. Those methods offer an amount of privacy, however, none of them can prevent re-identification of the query source if an attacker has knowledge about specific users' locations. Category 2 assumes the presence of a trusted third-party anonymizer server, and offers better protection against background knowledge attacks. [10, 9, 11] are all of that category. They offers a better privacy guarantee, but has several drawbacks: (i) The anonymizer is a single point of attack. (ii) A large number of cooperating, trustworthy users are needed. (iii) Privacy is guaranteed only for a single snapshot, users are not

* Corresponding author.

protected against correlation attack. Category 3 offers a stronger privacy guarantees, and protects user's privacy even against powerful adversaries. A novel LBS privacy approach based on Private Information Retrieval (PIR) was introduced in [8].It can resists the correlation attack, but can hardly protect user's identity privacy. It occupies a large computational and communication cost even for a small databases.

In this paper, we propose a full privacy LBS scheme without the anonymizer. We use the anonymous credentials system to protect user's identity privacy, while we use obfuscation of a program to protect user's usage profiles. We assume that there exists some geography or geometry methods to form a CR, in which all users are indistinguishable.

The rest of this paper is organized as follow: section 2 gives some basic preliminaries, section 3 provides the system architecture and presents our scheme, the security and privacy properties are considered in section 4, and the paper concludes in section 5 finally.

2 Preliminaries

In this section, we present the building blocks we are using to construct the proposed scheme.

The Anonymous Credentials System. Anonymous credentials system consists of users and organizations. Organizations know the users by pseudonyms. It allows users to authenticate themselves in a privacy-preserving manner. There are many protocol of that kind [3]. In this paper, we propose a anonymous credentials system based on Jan Camenisch and Anna Lysyanskaya (CL) signature scheme [3].

Program Obfuscation. An obfuscation $\mathcal{O}$ can obfuscate a code or program to create an obfuscated code, which is difficult for humans to understand. Recently in [7, 2], they use multilinear map and fully homomorphic encryption to realize security confusion of arbitrary polynomial circuit. That is, any function that can be realized using polynomial circuit can be obfuscated into an obfuscation $\mathcal{O}$. The obfuscation $\mathcal{O}$ will not reveal anything about what the function is.

3 Full Privacy-Preserving LBS Scheme

In this section, we describe the details of the system architecture and the proposed scheme of our full privacy-preserving LBS scheme.

3.1 The System Architecture

Fig.1 depicts the architecture of our proposed scheme. It consists of two parts, the mobile user(Bob) and LP. LP has a database(DB), the data stores in DB in the form of $(attribute, coordinate, content)$ short for (att, coo, con). The attribute of the data includes hospital, school and so on. The coordinate of data is a location

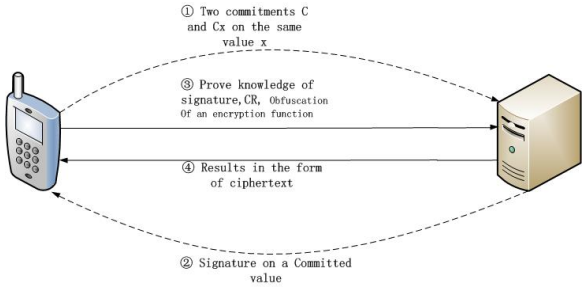

Fig. 1. The System Architecture

in the form of (longitude, latitude), while the content is the detail of the data. Typically a user asks NN queries in the form of $query(attribute, coordinate, \theta)$ short for $q(att, coo, \theta)$. The message flow in that architecture can be divided into three phases: System Setup, User Join, Getting services. The Fig.1 shows only the last two phases.

- System Setup: LP and Bob generate the system public keys and secret keys, LP manages DB in the form of (att, coo, con).
- User Join: Bob joins the system and obtains an anonymous credential from LP.
- Getting services: Bob first proves knowledge of the anonymous credential to LP, then sends his query and cloking area CR to LP. After verifying Bob's credential, LP feeds back Bob's query according to its database.

3.2 The Proposed Scheme

The target of this paper is to propose a full privacy LBS scheme. Our approach is twofold. One is to use the anonymous credentials system to hide the user's identity and the other is to use the obfuscation of a program to protect user's usage profile. The scheme is shown below in detail.

System Setup Phase: In this phase, LP does the things: Generate two safe primes p, q, and calculate a special RSA modulus $n = pq$. The length of p, q is $l_n = 2k$, k is a system parameter; Randomly choose $a, b, c \in \mathbf{QR}_n$; Set $PK = (n, a, b, c)$ and $SK = p$; Manage the database DB in the form of (att, coo, con). Bob does the following things: Generate a number $h_c \in \mathbf{QR}_n$ as a generator of group $< h_c >$; Randomly choose a number g_c from group $< h_c >$. Sets (n_c, g_c, h_c) as commitment public key; Generate $ElGamal$ public key $pk = (\mathbb{G}, q, g, h)$ and secret key $sk = x$.

User Join Phase: In this phase, Bob randomly chooses a value, and makes commitments on that value. LP signs on that committed value, and knows nothing about it. The detail is shown in Figure.2.

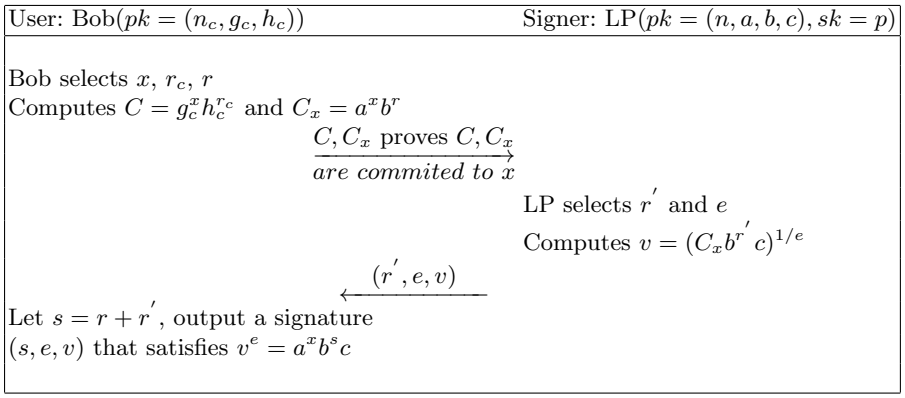

Fig. 2. *User Join* Protocol

Getting Services Phase : The getting services phase starts, when Bob wants to make a service query to LP. Bob does the following things: Contacts LP through proof knowledge of the signature to make LP believe he is a legal user. It goes like the phase of proof knowledge of the signature in [3], you can go to [3] for detail, thus it is omitted here; Gets himself located and generates a CR; Generates the following program F in figure 3; Uses obfuscation method to obfuscate F into an obfuscated code $\mathcal{O}$; Lastly sends $\mathcal{O}$ and CR to LP.

Input: $T = SizeofDB'$, $data_i (1 \leq i \leq T)$.
$data_i$ means the ith data in DB'.
Output:$Enc(data_i)$.

1. for $1 \leq i \leq T$;
2. if $data_i(att) = q(att) \wedge d(data_i(coo), q(coo)) \leq \theta$;
3. output $Enc(data_i)$;
4. else output $Enc(ans)$.

$d(a, b) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ is a simple function to compute the distance between bivector $a(x_1, y_1)$ and $b(x_2, y_2)$; ans is an initial random number known to Bob, unknown to LP; $Enc(data)$ is an El Gamal encryption scheme

Fig.3. Program of *Getting Service*

LP first verifies the zero knowledge proof. If passed, he does the things: Searches DB according to CR and θ to find all the available data DB'; Sets all the data as inputs of the obfuscated code $\mathcal{O}$; Sending all the outputs of $\mathcal{O}$ to Bob.
Bob uses the $ELGamal$ decryption function decrypt the results to obtain the

results in the form of plaintext. If none of the data satisfies Bob's query, $\mathcal{O}$ may give a result of $Enc(ans)$. Bob decrypts the $Enc(ans)$ will get ans, Bob will know that there is no place satisfying his requirement.

4 Security

The proposed scheme has the property of full privacy. The full privacy includes location privacy, identity privacy and usage profiles privacy. Following we analyze the full privacy from three aspects.

Theorem 1 . The location privacy preserved if the user is indistinguishable from $K-1$ other users in the CR.

Proof: The location privacy of our scheme is based on the property of CR. If the user is indistinguishable from $K-1$ other users in that CR, So LP won't know the exact location of the user and the user's location privacy has been protected.

Theorem 2 . The identity privacy preserved if CL signature-based authentication is anonymous.

Proof: The identity privacy property means that LP are conceived to forestall the re-identification of anonymous users. In CL signature, it allows users to authenticate themselves in a privacy-preserving manner. He can prove to LP that he has a right credential without revealing anything else about his identity. So LP won't know who the user is, and the anonymous property holds in a snapshot. CL signature makes the user capable to prove knowledge of signature as mangy time as possible even to the same verifier. So users can make continuous services queries while LP won't know that the queries came from the same user. Therefore our scheme can resist correlation attack and keep identity privacy in single or continuous services queries.

Theorem 3 . The usage profiles privacy preserved if program obfuscation $\mathcal{O}$ satisfies virtual black-box property.

Proof: From the virtual black-box property of obfuscation $\mathcal{O}$, LP won't know anything about the function. The outputs of $\mathcal{O}$ are ciphertexts of $ElGamal$ encryption, LP won't know what results are they. When no data in DB matches, $\mathcal{O}$ outputs an encryption of an initial random number ans, so LP won't know that no data matches and get nothing from the execution of $\mathcal{O}$. The $ElGamal$ encryption is a scheme of CPA security, the same data encrypts twice will get different results. So LP cannot use the $\mathcal{O}$ as random oracle to test which result the user has got. The usage profiles have been preserved.

5 Conclusion and Comparison

In this paper we proposed a full privacy LBS scheme. We assume that there exist geographical or geometrical methods to form a CR. The efficiency of our scheme may be low, but we proposed the first scheme to fulfil full privacy. We use anonymous credentials system to hide the user's identity and use obfuscation of a program to protect user's usage profiles. The security of our scheme is based on the security of CL signature, CR, $obfuscation$ and so on. To find other efficient

and useful function families which can provide anonymous property is our future work.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61379154 and U1135001) and the Specialized Research Fund for the Doctoral Program of Higher Education.

References

1. A.Kofod-Petersen, J.Cassens, *Proxies for Privacy in Ambient Systems*, JoWUA, volume: 1, number: 4, pp.62-74, 2012
2. Z. Brakerski, G. N. Rothblum. *Virtual Black-Box Obfuscation for All Circuits via Generic Graded Encoding*, IACR Cryptology ePrint Archive, 2013
3. J. Camenish, A. Lysyanskaya, *A Signature Scheme with Efficient Protocol*, SCN2002, LNCS2576, pp.268-289, 2003
4. M. L. Damiani, *Privacy Enhancing Techniques for the Protection of Mobility Patterns in LBS:Research Issues and Trends*, European Data Protection: Coming of Age, 2013
5. M. Damiani, E. Bertino, and C. Silvestri. *PROBE: an Obfuscation System for the Protection of Sensitive Location Information in LBS*, Technique Report 2001-145, CERIES, 2008
6. G. Ghinita, *Understanding the Privacy-Efficiency Trade-off in Location Based Queries*, ACM APRINGL 2008, pp.1-5, 2008
7. S. Garg, C. Gentry, S. Halevi. *Candidate Indistinguishability Obfuscation and Functional Encryption for All Circuits*, EUROCRYPT 2013, LNCS 7881,Springer, pp.1-17, 2013
8. G. Ghinita, P. Kalnis, A. Khoshgozaran, C.Shahabi, K.Tan. *Private Queries in Location Based Services: Anonymizers are not Necessary*, SIGMOD 2008, pp.121-132, 2008
9. P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias. *Preserving Location-based Identity Inference in Anonymous Spatial Queries*, IEEE TKDE, 19(12), pp.1719-1733, 2007
10. M. F. Mokbel, C. Y. Chow, and W. G. Aref. *The New Casper: Query Processing for Location Services without Compromising Privacy*, In Proceedings of VLDB, pp.763-774, 2006
11. M. Dahl, S.Delaune, and G.Steel. *Formal Analysis of Privacy for Anonymous Location Based Services*, In Proceedings of the Workshop on Theory of Security and Applications (TOSCA'11), LNCS 6993, pp.98-112, 2012
12. M. L. Yiu, C. Jensen, X. Huang, and H. Lu. *SpaceTwist: Managing the Trade-Offs Among Location Privacy, Query Performance, and Query Accuracy in Mobile Services*, In International Conference on Data Engineering (ICDE), pp.366-375, 2008