Computer Communications 85 (2016) 1–13

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

ABAKA: A novel attribute-based k-anonymous collaborative solution for LBSs

Tooska Dargahi[a], Moreno Ambrosin[b], Mauro Conti[b,*], N. Asokan[c]

[a] Department of Computer Engineering, West Tehran Branch, Islamic Azad University, Tehran, Iran
[b] Department of Mathematics, University of Padua, Padua, Italy
[c] Department of Computer Science, Aalto University and University of Helsinki, Helsinki, Finland

ARTICLE INFO

Article history:
Received 19 August 2015
Revised 17 February 2016
Accepted 7 March 2016
Available online 4 April 2016

Keywords:
Location-based services
Privacy
k-anonymity
p-sensitivity
Ciphertext-policy attribute-based encryption

ABSTRACT

The increasing use of mobile devices, along with advances in telecommunication systems, increased the popularity of Location-Based Services (LBSs). In LBSs, users share their exact location with a potentially untrusted Location-Based Service Provider (LBSP). In such a scenario, user privacy becomes a major concern: the knowledge about user location may lead to her identification as well as a continuous tracing of her position. Researchers proposed several approaches to preserve users' location privacy. They also showed that hiding the location of an LBS user is not enough to guarantee her privacy, i.e., user's profile attributes or background knowledge of an attacker may reveal the user's identity. In this paper we propose ABAKA, a novel collaborative approach that provides identity privacy for LBS users considering users' profile attributes. In particular, our solution guarantees *p*-sensitive *k*-anonymity for the user that sends an LBS request to the LBSP. ABAKA computes a cloaked area by collaborative multi-hop forwarding of the LBS query, and using Ciphertext-Policy Attribute-Based Encryption (CP-ABE). We ran a thorough set of experiments to evaluate our solution: the results confirm the feasibility and efficiency of our proposal.

1. Introduction

With the rapid development of mobile devices and advances of telecommunications, mobile users tend to have ubiquitous access to information such as traffic prediction or location map data. Location-Based Services (LBSs) are the best examples of this new trend, allowing mobile users to receive information based on their geographical position [1]. Based on their location, mobile users can access several types of information and services, e.g., getting the position of the nearest gas station, restaurant or hospital.

An LBS consists of two major entities: a user (from now on referred also as *issuer* of a query) who is interested in acquiring location-based service, and a Location-Based Service Provider (LBSP) which provides the desired location-based service to the issuer. To obtain such a service, the issuer sends her geographical location, along with her identity and the query to the LBSP. Unfortunately, some queries (such as searching for the nearest hospital specialized in a particular disease) may reveal privacy-sensitive information about the issuer.

The growing interest of smartphone users in using LBSs leads to two major privacy concerns: *location privacy* and *identity privacy* (also known as *query privacy*). The former refers to preventing the disclosure of the exact location of an issuer, while the latter is the ability of concealing the link between her identity and her query. These two concepts are complementary, and therefore, guaranteeing both location and identity privacy for an issuer becomes a challenging task. Researchers proposed several solutions providing location and identity privacy in the context of LBSs (examples can be found in [2]). The location privacy problem has also been studied extensively in other contexts such as sensor networks [3], and cloud computing [4].

A popular tool used in the literature to guarantee user's identity privacy, in the context of LBSs, is the concept of *k*-anonymity [5]. This concept refers to a set of k users in which a target user is indistinguishable (with respect to her location) from the other $k-1$ individuals in the set. However, according to [6], in the presence of an attacker with background knowledge about a user's profile attributes, we can only guarantee *k*-anonymity by considering anonymity sets in which all the users have the same profile attributes. Furthermore, the authors in [7] proved that *k*-anonymity is not sufficient to protect the privacy of an individual's attributes in a dataset, and might not prevent the disclosure of sensitive attributes for the user. With respect to *sensitive attributes*, we refer to a precise definition in [8]: "*an attribute whose values may be confidential for an individual (subject to her/his preferences)*".

[*] Corresponding author. Tel.: +390498271488.
E-mail address: conti@math.unipd.it (M. Conti).

http://dx.doi.org/10.1016/j.comcom.2016.03.002
0140-3664/© 2016 Elsevier B.V. All rights reserved.

Indeed, in the context of LBSs, the semantics of an issued query might allow the LBSP to infer sensitive attributes of an issuer's profile, or even her identity [9].

In order to address this problem, researchers proposed a solution called p-sensitive k-anonymity [7,9,10], in which at least p different values for each group of sensitive attributes are used. In the context of LBSs, this translates in ensuring that the anonymity set for an issuer contains individuals with diverse values for a specific set of privacy-sensitive attributes. In this paper, inspired by the concept of "personalized privacy preservation" by Xiao and Tao in [8], we give the opportunity to the issuer of a query to decide her preferences in sensitive attributes, based on her query content and physical location. We provided this feature for the issuer, due to the fact that an attribute could be sensitive for a query in special location, and insensitive for another query in another location- (we will further clarify this matter in the following). Before introducing the key contribution of the paper, we present a running example.

Medical help example. Consider a set of smartphone users in a geographical area. We assume that each user is assigned a profile that consists of five attributes: {*Gender, Age, Nationality, Job, Zip-code*}. Suppose a user u_1 is a 19-year-old Finnish girl living in Italy. She is looking for a pregnancy help center near her house, where the doctors are able to speak English. She sends an LBS query $Q =$ *"where is the nearest pregnancy help center with English speaking doctors?"* and wants to cloak her location while being 9-anonymous. In this example, based on the content of the query, the attributes *Gender* and *Zip-code* should be identical between all the users in the anonymity set (i.e., providing profile k-anonymity). Moreover, based on the semantics of the issued query, *Age* and *Nationality* are sensitive attributes of u_1. It should be noted that age and nationality are not sensitive attributes per se, but due to the fact that the issuer is in Italy, her nationality could reveal her identity. Moreover, her query semantics (i.e., being pregnant) strongly relates to her age. Therefore, we consider these two attributes to be her sensitive attributes. Assume that she computes a cloaked area using one of the existing k-anonymity preserving methods, and sends her query to the LBSP. Given the fact that she is looking for an English speaking doctor, a malicious LBSP can infer that the issuer is foreigner. Moreover, suppose that there are only two foreign users in her cloaked area: one 19 years old (u_1) and the other 50 years old. In such case, if the attacker has this background knowledge, he can infer that the issuer is likely to be u_1. This example emphasizes the fact that, based on the query semantics and considering the attacker's background knowledge, some attributes could be sensitive in specific scenarios and reveal the identity of the issuer. A proper privacy preserving solution should take into account sensitive attributes of u_1, according to the semantics of the query. For example, a solution could provide an anonymity set in which all the k users are non-Italian (i.e., providing profile k-anonymity) and there are enough diversity in age attribute (i.e., providing p-sensitivity considering the more probable values for being pregnant).

Contribution. In this paper, we propose ABAKA (Attribute-Based k-Anonymous collaborative solution for LBSs), a novel solution to provide both identity, and location privacy for LBS users taking into account the profile attributes of the users. Our motivation is the existing limitations of the prior research in the area of LBS users' privacy: on the one hand, those researches which attempt to ensure k-anonymity considering the profile of the users (such as in [6]) are centralized; and on the other hand, the existing distributed approaches do not consider profile attributes of the LBS users (such as in [11]).

In this paper, we make the following contributions:

- We propose ABAKA, the *first* privacy-preserving LBS system that guarantees p-sensitive k-anonymity running a TTP-free protocol between participating users (Section 4). In particular, ABAKA has the following features:
  - It cloaks the exact location of a user into a cloaked area of arbitrary size, by ensuring that (at least) $k-1$ collaborating users will forward a query in a random multi-hop path within the cloaked area.
  - ABAKA guarantees p-sensitivity by ensuring that the collaborating users in the anonymity set, which will forward the query, have specific attributes selected by the issuer. Each issuer can select a desired set of attributes based on the semantics of the query she wants to send. In particular, with ABAKA she can decide: (i) which attributes need to be identical within an anonymity set; and (ii) which attributes are sensitive, and thus need to have p different values within the anonymity set.
  - ABAKA adopts Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [12], in order to apply fine-grained access control over encrypted data, by defining high-level access policies as a combination of attributes. CP-ABE allows the issuer to specify attribute-based policies on the query; in this way, she ensures that other $k-1$ collaborative users have the desired attributes.
  - ABAKA ensures the confidentiality of the query, by using public key encryption.
- We run a systematic performance evaluation of ABAKA using two different datasets (Section 5.1) and a thorough evaluation of the computational overhead imposed by cryptographic processing required by ABAKA (Section 5.2). Our evaluation demonstrates that ABAKA is feasible on both smartphone and PC platforms.

2. Background on attribute-based encryption

In what follows, we introduce the fundamental concepts about Attribute-Based Encryption (ABE), and Ciphertext-Policy Attribute-Based Encryption (CP-ABE) in particular. In 2005, Sahai and Waters introduced a Fuzzy Identity-Based Encryption scheme [13], called ABE. This scheme is a public key encryption protocol that allows an encryptor to specify fine-grained access control policies over data. In this scheme, each user is assigned a set of *attributes* (e.g., *Gender, Age*, or *Job*). The data owner encrypts a plaintext in such a way that all the users that have a specific set of attributes will be able to decrypt the ciphertext (i.e., if user's attributes *satisfy* the policy over the data). CP-ABE [12] is a type of ABE in which the access policy is included into the ciphertext, and expressed as a combination of attributes. An example of such a policy is: $(Age = 19 \wedge Gender = female) \vee (Nationality = Italian)$ (see Fig. 1).

Each user has a private decryption key, which represents the set of attributes she owns. She will be able to decrypt a ciphertext

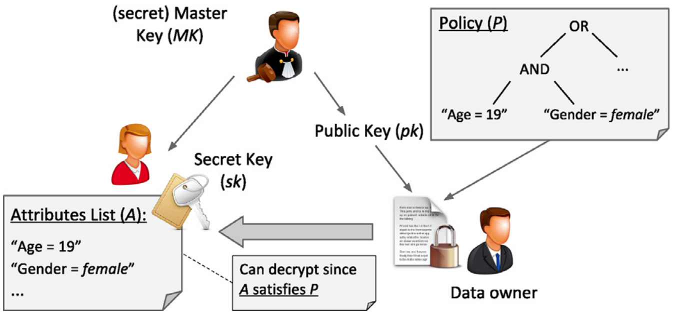

Fig. 1. Example of CP-ABE encryption and decryption.