Future Generation Computer Systems 49 (2015) 125–132

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

A novel attack to spatial cloaking schemes in location-based services

Ben Niu [a,b], Xiaoyan Zhu [a], Qinghua Li [c], Jie Chen [a], Hui Li [a,*]

[a] National Key Laboratory of Integrated Networks Services, Xidian University, Xi'an 710071, China
[b] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[c] Department of CSCE, University of Arkansas, AR, USA

HIGHLIGHTS

- We design a variance-based attack to exploit weaknesses of some existing algorithms.
- We verify the effectiveness of our VBA on several existing algorithms.
- We propose a random walk-based cloaking algorithm to mitigate the proposed attack.

ARTICLE INFO

Article history:
Received 1 March 2014
Received in revised form
12 October 2014
Accepted 25 October 2014
Available online 18 November 2014

Keywords:
Internet of things
Mobile computing
Wireless communications
Location-based services
Location privacy

ABSTRACT

Location-based services (LBSs) have been one of the novel uses and most popular activities in internet of things (IoT). In such location-based applications, mobile users enjoy plenty of conveniences at the cost of privacy. To protect user's location privacy, many research solutions have been proposed. In this paper, we focus on an important class of solutions, short-range communication-based spatial cloaking algorithms, which achieve *k*-anonymity within some collaborative groups. We first analyze the inherent drawbacks of existing P2P-based and encounter-based spatial cloaking approaches and propose a Variance-Based Attack (VBA) against them. Then we study the proposed attack on several existing spatial cloaking solutions. Finally, we propose a countermeasure *R*-cloak, which can mitigate VBA for current P2P cloaking algorithms. Our empirical evaluations further verify the effectiveness and efficiency of *R*-cloak.

1. Introduction

Internet of things (IoT) aims to extend computing and connectivity to anything, anyone, anywhere and anytime. Based on modern mobile devices with GPS module, all the "things" are connected by some wireless technologies, such as RFID, WiFi, Bluetooth, ad hoc and 3G/4G networks, etc. As a result, location-based services (LBSs) become increasingly popular and bring us more opportunities to learn about the environment. By submitting location-related queries to particular LBS servers, users can obtain and enjoy the service data easily, such as navigation, finding nearby point of interests (POIs), receiving location-based advertisements or traffic alerts, etc.

However, since the submitted queries contain some sensitive information (e.g., user interests and current location), the untrusted LBS server can obtain these information about users such as where the users are, what they are doing, and so on. As a result, the LBS server may track users directly or just release the sensitive information to third parties (e.g., advertising agencies) for monetary purpose. This situation conflicts with the increasing privacy concern of mobile users. We thus need to protect users' privacy.

To address the privacy problem, many approaches [1–5] have been proposed. Besides policy-based solutions and cryptography primitive-based schemes [6], most of them are based on the widely used location perturbation and obfuscation [7,8] techniques, which aim to protect location privacy by adding some noises or shifting the original location. They can be roughly divided into two main categories: trusted anonymization server-based schemes [9–11] and mobile device-based schemes [12–15]. With the help of a trusted server (e.g., *location anonymizer* [16]), *k-anonymity* [17] can be achieved by enlarging the exact location of real user into a bigger cloaking region covering several other users (e.g., $k-1$) geographically. As a result, the untrusted LBS server's uncertainty of distinguishing the real location from all the submitted lactations increases. However, the drawbacks of the third-party server are obvious since it knows all the submitted information of mobile users and becomes the single point

* Corresponding author.
E-mail addresses: xd.niuben@gmail.com (B. Niu), xyzhu@mail.xidian.edu.cn (X. Zhu), qinghual@uark.edu (Q. Li), jccucn@gmail.com (J. Chen), lihui@mail.xidian.edu.cn (H. Li).

http://dx.doi.org/10.1016/j.future.2014.10.026
0167-739X/

of failure. It is also the bottleneck of the system performance. Different from server-based schemes, mobile device-based solutions for privacy protection are mainly achieved through P2P-based schemes [18,19] or the encounter-based schemes [13,20,15] which do not rely on any trusted anonymization server. Based on the information shared and exchanged between users, they can construct a suitable size of cloaking region for mobile users under protection of *k-anonymity*. Unfortunately, due to constrains of short-range communication (e.g., peer discovery strategy and communication range) and users' mobility pattern, existing P2P-based and encounter-based schemes suffer from two problems. **First**, the chosen locations may be around the location of the real user with high probability. **Second**, the cloaking region may not be big enough to satisfy the minimum requirement of each mobile user. As a result, the adversary can achieve a higher probability than allowed to determine the user's real location.

In this paper, we analyze the typical properties of existing short-range communication-based spatial cloaking algorithms and propose a Variance-Based Attack (*VBA*) against them, which can detect the real location of user with higher probability than allowed. Then, with several case studies, we verify the effectiveness of *VBA*. Finally, we propose a random walk-based cloaking algorithm to protect location privacy from *VBA*. The major contributions of this paper are as follows.

- We propose a variance-based attack which exploits the weaknesses of recent short-range communication-based spatial cloaking algorithms. This attack is constructed based on the following intuitions. Since users can only communicate with other users nearby through WiFi/Bluetooth, the real user is more likely located at the center part of the cloaking region, which indicates that, with a higher probability, the variance of distances from the real user to the other users within the cloaking region is the smallest. Thus, the variance of distances can be used to infer the real user's location.
- We verify the effectiveness of our *VBA* on several existing cloaking algorithms including the P2P-based spatial cloaking schemes in [18,19], and the encounter-based privacy-preserving schemes in [20,15]. The case studies show that those schemes are vulnerable to variance-based attack.
- To address the problem of current short-range communication-based spatial cloaking algorithms, we propose a random walk-based cloaking algorithm (*R-cloak*) which can mitigate the proposed attack. Instead of directly selecting all the $k-1$ users from other users nearby, we select part of them (e.g., m users, where $m < k-1$) first, and randomly choose one of the selected users to perform the aforementioned selection phase again until the number of chosen users reaches $k-1$. Evaluation results further indicate the effectiveness of *R-cloak*.

The rest of this paper is organized as follows. We present related work in Section 2. In Section 3, we introduce some preliminaries. In Sections 4 and 5, we describe the details of the proposed attack and our solution. Finally, we show evaluation results and draw conclusions in Sections 6 and 7, respectively.

2. Related work

2.1. Location privacy metrics

User location in LBSs can always be defined as a single coordinate, therefore, we can measure the location privacy by determining the probability of distinguishing the real coordinate from the observed information. Based on this principle, two typical types of location privacy metrics have been proposed: uncertainty-based and distortion-based metrics. The former metrics measure the probability (i.e., location exposing probability) of determining the real location from a set of candidates. The distortion-based metrics [21,22] are based on the estimation error between the real location to the observed location of the real user. The adversary quantifies this distortion and aims to minimize it to infer user's real location. *k-anonymity* [17,15,5] is one of the most popular metrics to quantify location privacy, it tries to hide the real information of a user into other $k-1$ users. User's location privacy is closely related to the adversary's ability to distinguish the real information from others. For example, the ability to link two pseudonyms of a particular user or distinguish the paths along which a user may travel has been investigated in [2,23], respectively. As a better measurement for the location privacy, entropy-based metrics, which borrow the concept of entropy from information theory to quantify the uncertainty of determining the real location from others, are proposed and have been widely used in current literature approaches [23,20].

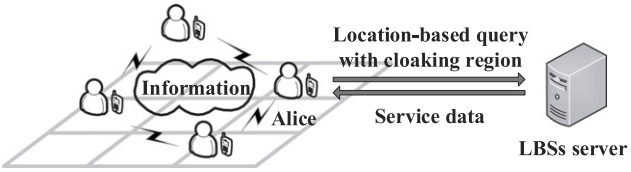

Fig. 1. The basic system architecture.

2.2. Location privacy preserving mechanisms

Location privacy preserving mechanisms (LPPMs) have become an important research area in location privacy over recent years, besides some policy-based approaches and cryptography primitive-based approaches [6], they always protect user's privacy through one of the following techniques. (a) *temporal and spatial cloaking*. Basically, the mobile users need to expand their exact locations into a bigger cloaking region [9,18,19,24], which covers more users (i.e., k users), geographically. (b) *Location obfuscation*. The main idea behind it is to blur the real location into another space in which the user's exact location and spatial information [25,3] can be maintained while enjoying the obtained service data from the LBS server. (c) *False locations*. Mobile users protect their privacy by submitting a set of locations, including both their real locations and some randomly [26] or carefully [15,5] generated dummies. In our work, we choose the *temporal and spatial cloaking*-based scheme due to that it is one of the most popular techniques, which supports kinds of environment settings such as centralized [9,27,16,24], distributed [28] and P2P [18,19], and many problem settings, including snapshot scenarios [9,18,28,16,29,24,15,30], which are used to protect user privacy in some independent cases (i.e., interests finder applications), continuous scenarios [27] and trajectories [31], which aim to preserve user privacy in some non-independent cases (i.e., location navigation). In our work, we pay much attention to provide privacy-preserving scheme for users in the snapshot scenario.

3. Preliminaries

3.1. System architecture

Fig. 1 illustrates the system architecture of short-range communication-based cloaking schemes. We employ two roles, mobile users and LBS server.

Mobile users: they can communicate with either the LBS server through the cellular networks (i.e., 3G/4G), or other mobile users by some short-range communication techniques, such as WiFi or Bluetooth. Mobile users can form a collaborative group, share