P³RN: Personalized Privacy Protection Using Query Semantics over Road Networks

Xiao Pan, Lei Wu, Chunhui Piao, and Xiaoshuo Xu
Shijiazhuang Tiedao University, Shijiazhuang, 050043, China
{smallpx, wulei, pch, xuxs}@stdu.edu.cn

Abstract. Privacy protection has received considerable attention for location-based services. A lot of location cloaking approaches focus on the identity and location protection, but few algorithms pay attention to prevent the sensitive information disclosure using query semantics. In terms of personalized privacy requirements, all queries in a cloaking set, from some user’s point of view, are sensitive. These users regard the privacy is breached. We call this attack as personalized homogeneity attacks. We show that none of the existing location cloaking approaches can effectively resolve this problem over road networks. We propose a \((K, L, P)\)-anonymization model and a personalized privacy protection cloaking algorithm over road networks P³RN. The efficiency and effectiveness of P³RN are validated by a series experiments.

Keywords: Privacy protection, road networks, location based services, sensitive information.

1 Introduction

With advances in wireless communication and mobile positioning technologies, location-based services (LBSs) have seen wide-spread adoption. These applications provide users with a great convenience, and improve the quality of work and personal life significantly. However, the increasing collections of individual’s information (e.g., location) open the door for potential privacy disclosure.

In general, existing work on privacy preserving in LBSs protects users three kinds of information: identity, location, and sensitive information. In order to protect these information, different models and methods are proposed. To hide the user’s identity, location \(k\)-anonymity model is the most acceptable model. For example, in Fig. 1(a), \(u_1\), \(u_2\) and \(u_3\) constitute a cloaking set. The locations of \(u_1\), \(u_2\) and \(u_3\) are represented by a segments set \(S\). \(u_1\) is indistinguishable from \(u_2\) and \(u_3\) in the cloaking set, thus the users identities are successfully protected.

To hide the exact locations, based on the location \(k\)-anonymity model, location cloaking is the popular methodology. Its main idea is to reduce the spatial and temporal resolution of the user’s location. In Euclidean space, exact locations are usually extended to a rectangle or a circle. While in a constrained space (e.g. road networks), exact locations are usually published as a segments
In the previous example, the adversary cannot sure where is the exactly location of \( u_1 \) on the segments in \( S \).

Most existing work focuses on the identity and location protection, but few literatures pay attention to prevent the sensitive information disclosure. The sensitive information is disclosed using two kinds of published information: location semantics \([3,7,16]\) and query semantics \([14]\). Fig. 1(b) shows an example of the first case. Considering the public geographical context, the segments in \( S \) are totally covered by a hospital. Thus, each user’s visited place is disclosed. Fig. 1(a) illustrates the other case. When the adversary considers the query semantics, three users issue hospital-related queries. The attackers get the inquirer’s health conditions. In this paper, we consider to protect the sensitive information using query semantics over road networks.

In order to protect the sensitive information, besides location \( k \)-anonymity model, a cloaking set also follows \( l \)-diversity model \([9]\). In our scenario, according to the \( l \)-diversity model, the query contents issued from a cloaking set are at least \( l \) different. However, location \( k \)-anonymity model and \( l \)-diversity model are not sufficient to protect the sensitive information.

As we known, privacy requirements are personalized. Whether a query is sensitive or not depends on the personalized privacy requirements. Fig. 2(a) shows an example of the personalized privacy profiles. Fig. 2(b) is a cloaking set following location 3-anonymity model and 3-diversity model. In \( u_4 \)’s side, three queries issued from the cloaking set are all sensitive. \( u_4 \) doesn’t want any