POIsafe: a Privacy-Conscious System for Retrieval of Points of Interest

Daniele Riboni, Linda Pareschi, and Claudio Bettini

Università degli Studi di Milano, DICo, Milan, Italy
{riboni,pareschi,bettini}@dico.unimi.it

1 Introduction

Services for retrieval of points of interest (POIs) are becoming increasingly popular due to the widespread diffusion of GPS-enabled mobile devices having access to fast wireless networks. We have developed a context-aware service to share, manage, and retrieve geo-referenced resource descriptions enriched with multimedia content [1]. The access to such services is prone to potentially serious privacy issues, since requests include sensitive information or can lead to the disclosure of sensitive information, and they are often handled by untrusted parties, or sent through insecure channels. Context data, including user location, is in some cases sensitive information that users prefer not to be associated with their identity. In other cases, the interest for specific resources is considered sensitive and the issuer of such a request uses a pseudonym not to be identified; however, context data present in the same request or in a sequence of requests may be used by an adversary to re-identify the issuer. We are not aware of any context-aware service for retrieval of POIs with an effective and comprehensive privacy protection mechanism, and we believe this is a challenging research goal. In this paper, we focus on one particular kind of context data, location, but we plan to extend our techniques to tackle the general problem illustrated above.

Different techniques have been proposed for protecting against the disclosure of location information in location-based services (LBS). Cryptographic approaches inspired by Private Information Retrieval (e.g., [2]) provide very strong guarantees in terms of privacy; however, they determine a relevant overhead in network and power consumption and service response time, especially when applied to services that consider a wide set of context data. Obfuscation-based techniques are based on a perturbation of the user's location. In techniques based on generalization the exact location is enlarged to a region; in other cases, a fake user's location is communicated instead of the real one. In particular, the latter approach is adopted by *SpaceTwist* [5] to enforce location privacy while guaranteeing that k-nearest neighbor (kNN) queries are correctly answered even if the user provides a fake location, at the cost of computation and communication overhead. Indeed, according to SpaceTwist, the client issues a sequence of requests from the same fake location asking for more close-by POIs until it is sure that those provided by the service include the kNN set corresponding to its real location. Based on the request-response sequence, an adversary can

only identify an area (called *twisted space*) from which the requests may have been sent. More recently, a technique (derived from SpaceTwist) to couple location privacy with identity anonymity, named *AnonTwist* [4], has been proposed. Given a density map to estimate the number of people in a given area, AnonTwist provides a probabilistic guarantee that, even if an adversary has access to presence information, the twisted space contains at least N individuals; hence, the request issuer is indistinguishable among at least N individuals. However, both SpaceTwist and AnonTwist rely on the assumption that the function that generates the fake location is unknown to the adversary.

In this paper we take into account the realistic case in which the function that generates the fake location is known to the adversary. In Section 2 we illustrate a specific technique to protect privacy under this assumption. In Section 3 we present *POIsafe*, an extension of our system for POIs retrieval to enforce location privacy and identity anonymity as a first step towards a comprehensive solution considering other context data. Section 4 concludes the paper.

2 Enforcing privacy in *POIsafe*

Even if we are investigating alternative solutions, our current approach is inspired by AnonTwist [4]. However, differently from AnonTwist and SpaceTwist, our proposed algorithm assumes that the function for generating fake locations could be known to the adversary. Under this assumption, by observing the fake location, the adversary may reconstruct the possible area A from which it originated. Moreover, based on the request-response sequence, the adversary is able to understand that the area from which the request originated corresponds to the intersection I between the twisted space and area A. Hence, the goal of our technique is to ensure that area I is greater than a specified threshold, and that it includes at least N potential issuers.

In particular, the maximum radius r of perturbation (e.g., 1 Km) and the minimum number of potential issuers N are chosen according to the user's preferences about privacy and quality of service (QoS). Then, before issuing a request, a random *distance* ranging from 0 to r is chosen, and a random point having that distance from the real user's location is chosen as the fake location. Note that with this technique, A is the area contained within the circle having center in the fake location, and radius r. Then, the client incrementally asks for nearest POIs until *i)* the exact kNN set is retrieved, *ii)* area I is greater than the chosen threshold, and *iii)* I contains at least N users according to the density map.

In general, the farther is the fake location from the real one, the higher is the user's privacy. However, large perturbations of the real location determine poor QoS. Indeed, even if the service guarantees to correctly answer the user's query, a very high number of POIs may be communicated to the client before obtaining the correct kNN set, determining an increase in communication and computational costs, and response time. With respect to usability, users can be provided with an intuitive interface to set their preferences; i.e., a single slider with *response time* on the left-hand side and *privacy* on the right-hand side.

The value of the slider influences the value of radius r and threshold N. Before submitting a request, density information is used to control if the value of r is adequate to provide a sufficient level of anonymity with high probability (e.g., according to the density map the corresponding area includes a number of users that doubles threshold N). In this case, a green signal appears; in the other case, a red one is used. Note that the red signal does not mean that the technique will necessarily fail to preserve privacy, but only that anonymity might be at risk.

3 System architecture

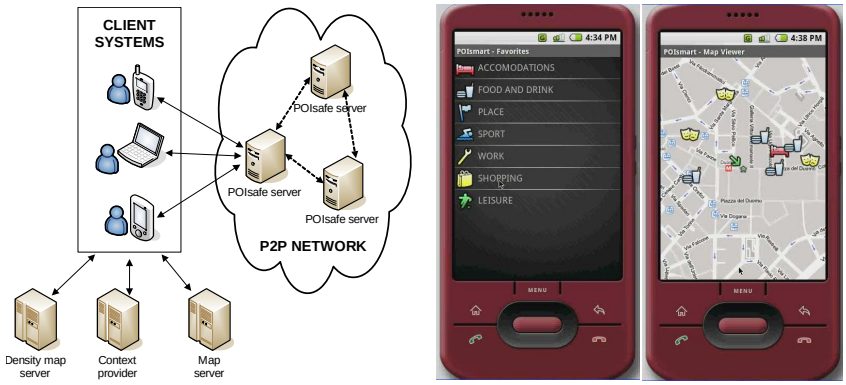
POIsafe is based on a peer-to-peer network of POISAFE SERVERS, which are in charge of managing POIs, and of searching POIs in the peer-to-peer network on the basis of the user's context and explicit search keywords. The mechanism of search in the peer-to-peer network has been presented in detail in [1]. An overview of the *POIsafe* network is shown in Figure 1(a).

Users can access the *POIsafe* network from a wide range of CLIENT SYSTEMS, which provide an interface for the user to browse her own POIs hierarchy, reorganize the hierarchy, add new POIs, search shared POIs in the peer-to-peer network, and set their preferences, including those regarding privacy. Before issuing a request, each CLIENT SYSTEM retrieves context information useful for service adaptation. This information can be retrieved either locally or from an external CONTEXT PROVIDER. Perturbation of location information and requests for POIs are executed as illustrated in Section 2. The density map is retrieved from a trusted density map server. In the previous version of *POIsafe*, ranking of POIs was performed only at the server side. However, since in the new version a fake user's location is communicated to the server, returned POIs are re-ranked at the client side considering the exact user's location. Moreover, an external MAP SERVER is queried by the CLIENT SYSTEM for obtaining maps showing the position of returned POIs and information for navigation support.

The POISAFE SERVER has been developed in Java, and implements the algorithms for POIs scoring and distributed search presented in [1], the modified AnonTwist algorithm, as well as various facilities for managing and searching POIs. The architecture adopts Web services for client/server and server/server communication. At the time of writing we have developed CLIENT SYSTEMS for laptops and smartphones. In particular, we have developed a novel Android client (see Figure 1(b)), which takes advantage of the integration with Google Maps.

4 Conclusions and future work

In this paper we presented the extension of an existing system for context-aware management and retrieval of points of interest. The extension is a first step towards a comprehensive privacy solution for this kind of services. In particular, future work includes a thorough investigation of the formal properties of the proposed algorithm. Several research issues remain open. In particular, we point out that the proposed technique may be ineffective if an adversary can observe



(a) *POIsafe* network (b) Client implementation

Fig. 1. *POIsafe* system and current implementation

histories of requests issued by users in different time granules. Indeed, as shown in [3], the frequency of a service parameter included in requests, matched with the presence of candidate issuers, can be exploited to associate a given user with that service parameter. The integration of techniques to protect against these kinds of attacks will be the subject of future investigation.

Acknowledgments

The authors would like to thank Song Wang and X. Sean Wang for providing a working implementation of the AnonTwist algorithm.

References

1. C. Bettini and D. Riboni. Context-aware Web Services for Distributed Retrieval of Points of Interest. In *Proc. of the 2nd International Conference on Internet and Web Applications and Services*, page 36–40. IEEE Computer Society, 2007.
2. G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan. Private Queries in Location Based Services: Anonymizers are Not Necessary. In *Proc. of SIGMOD 2008*, pages 121–132. ACM, 2008.
3. D. Riboni, L. Pareschi, C. Bettini, and S. Jajodia. Preserving Anonymity of Recurrent Location-based Queries. In *Proc. of the 16th International Symposium on Temporal Representation and Reasoning*. IEEE Computer Society, 2009.
4. S. Wang and X.S. Wang. AnonTwist: Nearest Neighbor Querying with Both Location Privacy and k-Anonymity for Mobile Users. In *Proc. of First International Workshop on Mobile Urban Sensing*, pages 443–448. IEEE Computer Society, 2009.
5. M. L. Yiu, C. S. Jensen, X. Huang, and H. Lu. SpaceTwist: Managing the Trade-offs among Location Privacy, Query Performance, and Query Accuracy in Mobile Services. In *Proc. of ICDE 2008*, pages 366–375. IEEE Computer Society, 2008.