MOBIIHIDE: A Mobile Peer-to-Peer System for Anonymous Location-Based Queries*

Gabriel Ghinita¹, Panos Kalnis¹, and Spiros Skiadopoulos²

¹ Dept. of Computer Science
National University of Singapore
{ghinitag,kalnis}@comp.nus.edu.sg
² Dept. of Comp. Science & Technology
University of Peloponnesse, Greece
spiros@uop.gr

Abstract. Modern mobile phones and PDAs are equipped with positioning capabilities (e.g., GPS). Users can access public location-based services (e.g., Google Maps) and ask spatial queries. Although communication is encrypted, privacy and confidentiality remain major concerns, since the queries may disclose the location and identity of the user. Commonly, spatial $K$-anonymity is employed to hide the query initiator among a group of $K$ users. However, existing work either fails to guarantee privacy, or exhibits unacceptably long response time.

In this paper we propose MOBIHIDE, a Peer-to-Peer system for anonymous location-based queries, which addresses these problems. MOBIHIDE employs the Hilbert space-filling curve to map the 2-D locations of mobile users to 1-D space. The transformed locations are indexed by a Chord-based distributed hash table, which is formed by the mobile devices. The resulting Peer-to-Peer system is used to anonymize a query by mapping it to a random group of $K$ users that are consecutive in the 1-D space. Compared to existing state-of-the-art, MOBIHIDE does not provide theoretical anonymity guarantees for skewed query distributions. Nevertheless, it achieves strong anonymity in practice, and it eliminates system hotspots. Our experimental evaluation shows that MOBIHIDE has good load balancing and fault tolerance properties, and is applicable to real-life scenarios with numerous mobile users.

1 Introduction

Consider the following scenario: Bob uses his GPS enabled mobile phone (e.g., iPAQ hw6515, Mio A701) to ask the query “Find the nearest AIDS clinic to my present location”. This query can be answered by a Location-Based Service (LBS), e.g., Google Maps, which is not trusted. To preserve his privacy, Bob does not contact the LBS directly. Instead he submits his query via a trusted

* This work has been partially supported by project PENED 03 funded by the European Social Fund (75%) and the General Secretariat of Research and Technology (25%).
pseudonym service which hides his identity (services for anonymous web surfing are commonly available). Nevertheless, the query still contains the exact coordinates of Bob. One may reveal sensitive data by combining the location with other publicly available information. If, for instance, Bob uses his mobile phone within his residence, the untrustworthy owner of the LBS may infer Bob’s identity (e.g., through a white-pages service) and speculate that he suffers from AIDS. Bob may even hesitate to ask innocuous queries such as “Find the nearest restaurant”, in order to avoid unsolicited advertisement.

Recent research on LBS privacy focused on the $K$-anonymity [17,20] technique, which is used in relational databases for publishing census, medical and other sensitive data. A dataset is $K$-anonymous, if each record is indistinguishable from at least $K-1$ other records with respect to certain identifying attributes. In the LBS domain, a similar idea appears in Ref. [7,12,15], which employ spatial cloaking to conceal the location of the querying user $u$: Instead of reporting the coordinates of $u$, they construct an Anonymizing Spatial Region (ASR or $K$-ASR) which encloses $u$ and $K-1$ additional users. Typically, a central trusted server (called location anonymizer, or simply anonymizer in the sequel) exists between the users and the LBS. All users subscribe to the anonymizer and continuously update their position while they move. Each user sends his query to the anonymizer, which constructs the appropriate $K$-ASR and contacts the LBS. The LBS computes the answer based on the $K$-ASR, instead of the exact user location; thus, the response may contain false hits. The anonymizer filters the result and returns the exact answer to the user.

The centralized approach has several drawbacks; for example, the anonymizer may become bottleneck since it must handle frequent location updates as users move [8]. Most importantly, the centralized anonymizer poses a serious security threat. If it is compromised by an attacker, or forced to cooperate with a government agency, the history of all user movements and their queries may be revealed. For these reasons, two fully distributed systems emerged: (i) cloakP2P [5] is a Peer-to-Peer system which constructs $K$-ASRs by considering users in the neighborhood of the querying user. (ii) PRIVÉ [8], on the other hand, clusters users in a hierarchical overlay network, resembling a distributed $B^+$-tree. Both systems minimize the security risk by distributing the sensitive information in numerous peers. However, we will show that cloakP2P fails to provide privacy for many user distributions, whereas PRIVÉ may suffer from slow response time, since root-level nodes constitute potential bottlenecks.

In this paper we propose MOBIHIDE, a Peer-to-Peer (P2P) system for anonymous location-based queries which addresses the problems of existing approaches. In MOBIHIDE the participating mobile devices form a hierarchical distributed hash table, based on the Chord P2P architecture [19], which indexes the locations of all users. In order to map the 2-D user locations to the 1-D Chord space, we employ the Hilbert space-filling curve [16]. $K$-ASRs are collaboratively assembled by peers in a distributed fashion, by choosing random groups of $K$ users (including the querying user) that are consecutive in the 1-D space. We prove that for uniform query distribution MOBIHIDE guarantees privacy,