The New Casper: A Privacy-Aware Location-Based Database Server[*]

Mohamed F. Mokbel[1] Chi-Yin Chow[1] Walid G. Aref[2]

[1]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN

[2]Department of Computer Science, Purdue University, West Lafayette, IN

Abstract

This demo presents Casper; a framework in which users entertain anonymous location-based services. Casper consists of two main components; the location anonymizer that blurs the users' exact location into cloaked spatial regions and the privacy-aware query processor that is responsible on providing location-based services based on the cloaked spatial regions. While the location anonymizer is implemented as a stand alone application, the privacy-aware query processor is embedded into PLACE; a research prototype for location-based database servers.

1 Introduction

The wide spread of location-detection devices (e.g., GPS-like devices) enables new applications in which users continuously send their location information to a location-based database server. Examples of these applications include store finders, traffic reports, and location-based advertisements. Although location-based applications promise safety and convenience, they threaten the privacy and security of their customers. In order to get a location-based service, a user has to report her private location information to the server. With untrustworthy servers, such model provides several privacy threats. For example, an employer may check on her employee behavior by knowing the places she visits or the personal medical records can be inferred by knowing which clinic a person visits.

This demo presents *Casper* [1, 3]; a novel framework that turns traditional location-based servers to provide anonymous services to their customers. *Casper* consists of two components, namely, the *location anonymizer* and the *privacy-aware query processor*. The *location anonymizer* is a trusted third party that receives continuous location updates from users, blurs the location updates to cloaked spatial areas, and sends the cloaked areas to the location-based database server. While *cloaking* the location information, the *location anonymizer* also removes any user identity to ensure the pseudonymity of the location information. Similar to the exact point locations, the *location anonymizer* also blurs the query location information before sending a cloaked query area to the location-based database server. The *privacy-aware query processor* is embedded inside the location-based database server to tune its functionality to deal with anonymous queries and cloaked spatial areas rather than the exact location information.

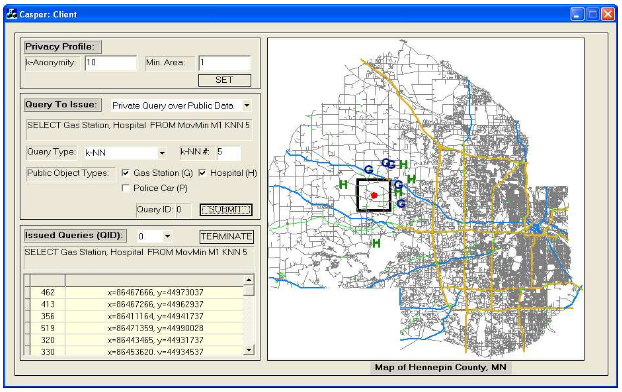
Mobile users register with *Casper* with a *privacy profile* that has the form (k, A_{min}), where k indicates that the user wants to be k-anonymous, i.e., not distinguishable other k users, while A_{min} indicates that the user wants to hide her location information within an area of at least A_{min}. Mobile users have the ability to frequently change their privacy profiles to adjust a personal trade-off between the amount of information they reveal about their locations and the quality of service that they obtain from *Casper*.

2 The Location Anonymizer

Casper employs two types of location anonymizers, namely, the *basic* location anonymizer and the *adaptive* location anonymizer. The *basic* location anonymizer employs a grid-based complete pyramid data structure that hierarchically decomposes the spatial space into H levels where a level of height h has 4^h grid cells. Each pyramid cell is represented as (cid, N) where cid is the cell identifier while N is the number of mobile users within the cell boundaries. The cloaking algorithm basically takes the exact location information as an input and finds its corresponding grid cell at the lowest pyramid level. If such cell satisfies the user privacy requirements, then it will be sent to the location-based database server as the cloaked spatial region. However, if such cell does not satisfy the privacy requirements, the algorithm is recursively evaluated on higher pyramid layers till a suitable grid cell is found.

On the other hand, the *adaptive* location anonymizer employs an incomplete pyramid structure that maintains only those grid cells that can be potentially used as cloaking regions for the mobile users. For example, if all mobile users have strict privacy requirements where the lowest pyramid level would not satisfy any user privacy profile, the *adaptive*

[*]M. Mokbel is partially supported by Grants-in-Aid, U. of Minnesota. W. Aref is supported in part by NSF Awards IIS-0093116 & IIS-0209120.



Casper: Client

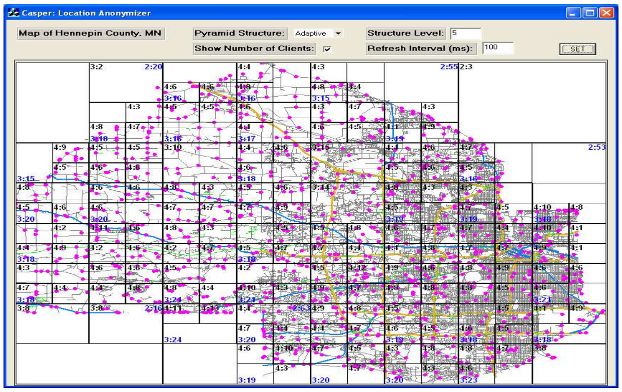
Privacy Profile:
k-Anonymity: 10 Min. Area: 1
SET

Query To Issue: Private Query over Public Data
SELECT Gas Station, Hospital FROM MovMin M1 KNN 5
Query Type: k-NN k-NN #: 5
Public Object Types: ☑ Gas Station (G) ☑ Hospital (H) □ Police Car (P)
Query ID: 0 SUBMIT

Issued Queries (QID): 0 TERMINATE
SELECT Gas Station, Hospital FROM MovMin M1 KNN 5

462	x=86467666, y=44973037
413	x=86467266, y=44962937
356	x=86411164, y=44941737
519	x=86471359, y=44990028
320	x=86443465, y=44931737
330	x=86453620, y=44934537

Figure 1: The Client GUI



Casper: Location Anonymizer

Map of Hennepin County, MN
Pyramid Structure: Adaptive
Show Number of Clients: ☑
Structure Level: 5
Refresh Interval (ms): 100
SET

Figure 2: The Adaptive location anonymizer

location anonymizer will not maintain such level, hence, the cost of maintaining the pyramid structure is significantly reduced. The cloaking algorithm of the *adaptive* location anonymizer is similar to that of the *basic* location anonymizer. However, in the *adaptive* location anonymizer, we always search in less number of levels as we maintain less grid cells than that of the *basic* location anonymizer.

3 Privacy-aware Query Processing

Two data types are stored in the *privacy-aware* location-based database server, *public* data and *private* data. *Public* data includes stationary objects such as hospitals, restaurants, and gas stations or moving objects such as police cars and on-site workers. *Private* data mainly contains personal information of mobile or stationary users with a *privacy profile* of non-zero k or non-zero A_{min}. Such data are represented as cloaked spatial regions rather than exact point location information. Based on the stored data, three novel query types are supported in *Casper* through its *privacy-aware query processor*: (1) *Private queries over public data* where the query issuer location is private while the data objects are public, (2) *Public queries over private data* where the query issuer location is public while the data objects are private, and (3) *Private queries over private data* where both the query and data object locations are private.

Considering private nearest-neighbor queries over public data as a case study, the *privacy-aware* query processor employs a filter-refine approach in which a set of data objects at the server side are selected as filters. Using these filters, the whole set of data is pruned to only a *candidate* list. The *candidate* list is sent back to the client where the query can be evaluated locally using the candidate list. The size of the candidate list is greatly affected by the user privacy profile. The more strict the privacy profile, the larger the candidate list size. The *privacy-aware* query processor of *Casper* is proved to provide an *inclusive* candidate list, i.e., the candidate list includes the exact query answer, and a *minimal* candidate list, i.e., given a set of filters, the candidate list is of minimal size.

4 Demo Scenario

Figure 1 gives a GUI screen shot of the client module. The GUI is divided into four parts: (1) A road network map that displays the position of the client roaming on the road network, (2) The privacy profile interface where the user can update the values of k and A_{min} at any time and submit the updated privacy profile to the location anonymizer, (3) The query submission interface in which the client has the ability to choose either range or nearest neighbor queries from four different query categories, public query over public data (i.e., traditional queries), private query over public data, public query over private data, and private query over private data, and (4) The query result interface in which the query answer is tabulated to the user in addition to being shown on the road network map.

Figure 2 gives the GUI screen shot for the *adaptive* location anonymizer. The GUI mainly shows the current exact locations of all participating mobile users over the road network. One level of the pyramid data structure is displayed as a grid over the road network. The number of mobile users at each grid cell is displayed along with the grid level. As a default, the GUI displays only the lowest-level maintained cells. Finally, there will be a GUI screen shot for the privacy-aware location-based database server which is implemented inside the PLACE server [2]. The GUI will show the road network map with the exact locations of public data and the cloaked regions for private data.

References

[1] M. F. Mokbel. Towards Privacy-Aware Location-Based Database Servers. In *International Workshop on Privacy Data Management, PDM, co-located with ICDE*, Apr. 2006.

[2] M. F. Mokbel and W. G. Aref. PLACE: A Scalable Location-aware Database Server for Spatio-temporal Data Streams. *IEEE Data Engineering Bulletin*, 28(3):3–10, 2005.

[3] M. F. Mokbel, C.-Y. Chow, and W. G. Aref. The New Casper: Query Procesing for Location Services without Compromising Privacy. In *VLDB*, 2006.