

Authenticating Location Based Skyline Queries in Mobile Environment

A.Sethupathi
Computer Science and Engineering
V.S.B Engineering College
Karur, India

A.P.V Raghavendra
Computer Science and Engineering
V.S.B Engineering College
Karur, India

Abstract: With the booming of Smartphone's and mobile devices, location-based services have experienced massive escalation in nowadays. The outsourcing data processing services to cloud service provider becomes very trending in recent years, which provides solution to the clients instead of data owner. However, we cannot expect real solutions from the data processing services; it may give dishonest results to the clients. Therefore, to provide the correct results some authentication techniques are requiring. In this paper, we learn the authentication techniques for location-based arbitrary-subspace skyline queries (LASQs), which signify an essential class of LBS applications. We suggest a basic Merkle Skyline R-tree method and a novel Partial S4-tree method to authenticate LASQs. For authentication process using this LASQ, the client can contact server frequently during movement and also verify the results by client itself.

Keywords: location based services, outsourcing data management, cloud service provider, Merkle Skyline R-tree, novel Partial S4-tree

1 INTRODUCTION

With the rapid development of mobile handset devices (such as smartphones and tablet computers), wireless communication, and positioning technologies in the past decade, Location-based services (LBSs) have prospered. Users carrying location-aware mobile devices are able to query LBSs for surrounding points of interest (POIs) anywhere and at any time. Among the many types of location-based queries, one important class is *location-based skyline queries*. These queries take into account both the spatial and non-spatial attributes of the POIs. A representative example is finding nearby restaurants with good food, where the distance to the querying user is a spatial attribute and the goodness of the food is a nonspatial attribute. The query returns a set of restaurants that are closer to the querying user and/or have better food than those not returned. In general, while spatial objects can have a long list of non-spatial attributes—such as foodquality, service, hygiene, environment, and price—only a *small subset* of these attributes (termed a *subspace*) is of interest to a particular user in a single query. Moreover, different users may have different preferences—e.g., Mary prefers taste, whereas Tom is concerned about hygiene, environment, and price. In this paper, we call these skyline queries location-based arbitrary-subspace skyline queries (LASQs). To scale up LBSs along with their ever-growing popularity, a rising trend is to outsource data management and service provisioning to Cloud service providers (CSPs) such as Amazon EC2 and Google App Engine. More specifically, a data owner delegates its data to a CSP, which in turn provides query services to clients on behalf of the data owner. While such an outsourcing model is advantageous in terms of cost, performance, and flexibility in resource management, it brings a great challenge to query integrity assurance. If the CSP is untrustworthy or compromised, it may return incorrect or incomplete query results to clients (intentionally or not) for various reasons:

The CSP may return incorrect results unintentionally because of bugs in the implementation of query processing algorithms.

- The CSP (or the adversary who compromised it) may intentionally tamper with the query results. For example, in the restaurant-finding scenario mentioned above, a restaurant may be ranked higher than other restaurants just because the CSP is sponsored by that restaurant.
- To cut costs or avoid performance bottlenecks in peak hours, the CSP may return incomplete results by carrying out the query evaluation process partially.

Therefore, in the data-outsourcing model, there is a need for clients to authenticate the *soundness* and *completeness* of query results, where *soundness* means that the original data is not tampered with by any third party (including the CSP), and *completeness* means that no valid result is missing. This leads to a problem known as *authenticated query processing* which has been studied for various spatial queries, including range queries top-*k* queries *k*NN queries and shortest-path queries.



Fig. 1. Authenticated query processing.

Fig. 1 shows a general framework of authenticated query processing. The data owner obtains, through a certificate authority (e.g., VeriSign), a pair of private and public keys of digital signatures. Before delegating a spatial dataset to the CSP, the data owner builds an authenticated data structure (ADS) of the dataset. To support efficient query processing, the ADS is often a tree-like index structure, where the root is signed by the

data owner using his/her private key. The CSP keeps the spatial dataset, as well as the ADS and its root signature. Upon receiving a query from the client, the CSP returns the query results, the root signature, and a verification object (VO), which is constructed based on the ADS. The client can authenticate the correctness of the query results using the returned VO, the root signature, and the data owner's public key.

In a preliminary study we have investigated the authentication problem for location-based skyline queries in a fixed space of attributes. In this paper, we extend this study to the general problem of authenticating location based skyline queries in arbitrary subspaces of attributes (*i.e.*, LASQs). Because a basic solution that returns all results in the full space is inefficient, we propose a new authentication method based on the notion of *signed subspace skyline scope (S4)*. We devise a data structure, called *Partial-S4-tree*, which pre-computes, signs, and stores the skyline scopes of some subspaces, so that many redundant objects can be easily identified and safely removed from the VO, thereby minimizing its size and saving the server processing time. To improve the filtering effects, we further propose a storage-budget allocation policy to construct the Partial-S4-tree for each spatial object. For continuous LASQs, the concept of *clear area* is introduced to enable a moving client to re-evaluate new results locally. Moreover, we propose an approach to prolong the client's residence time inside a clear area.

In summary, our contributions in this paper are four-fold:

- We identify the problem of authenticating LASQs in outsourced databases. To the best of our knowledge, this study is the first attempt to investigate this problem.
- For a one-shot LASQ authentication, we propose a basic Merkle Skyline R-tree method and a Partial-S4-tree method, aiming to reduce the server processing time and minimize the VO size.
- We develop a prefetching-based approach for authenticating continuous LASQs. This approach enables the clients to re-evaluate new LASQ results locally during movement, thus reducing both communication and computation costs.
- We conduct extensive experiments to evaluate the performance of the proposed methods and algorithms.

2.RELATED WORK

In this section, we review the related work on query authentication and skyline query processing.

2.1 Query Authentication

Authenticated query processing has been studied extensively. Most studies on query authentication are based on an AD called Merkle Hash Tree (MH-tree) In MH-tree, the digests of index nodes are recursively computed from the leaf nodes to the root. After that, the root digest is signed by the data owner's private key and stored on the outsourced database server. For each user query, this signature is returned to the client along with the query results and a VO for result verification. In contrast, an alternative method is to employ signature aggregation, which signs every object in the dataset and generates a VO by aggregating the signatures of the result objects along with some non-result objects (*e.g.*, the objects immediately beyond a query range). However, as the aggregate signature is generated on-the-

fly, this method incurs high overhead in query processing and client-side verification. Therefore, in this paper, we focus on authentication methods based on MH-tree.

Following the concept of MH-tree, the query authentication problem has been studied for relational databases data streams and text search engines Yang *et al.* first introduced this problem to the spatial database domain and studied the authentication of spatial range queries. They proposed an authenticated index structure called MR-tree, which combines the ideas of MB-tree and R*-tree Yiu *et al.* investigated how to efficiently authenticate kNN queries and shortest-path queries. In Hu *et al.* proposed a novel approach that authenticates spatial queries based on neighborhood information. More recently, in we developed new schemes for range and top-*k* query authentication that preserves the location privacy of queried objects.

In our preliminary studies, we investigated the authentication of location-based skyline queries in *fixed* subspaces. A new authenticated index structure called MRSky-tree (or MSR-tree) was proposed in. The main difference between MR-tree and MSR-tree is that the former indexes the spatial objects while the latter indexes the solution space of spatial objects (in form of a notion called skyline scope).

2.2 Skyline Query Processing

Skyline query processing was first introduced into the database community by Borzanyi *et al.*[4]. A number of algorithms have been developed since then. These algorithms can be divided into two categories. The first category is non-index-based algorithms. The representatives are Black-Nested-Loop (BNL) and Divide-and-Conquer (D&C). BNL scans the dataset sequentially and compares each new object to the skyline candidates obtained so far. D&C partitions the dataset into several parts, processes them part by part, and finally merges all partial skylines. SFS improves BNL by pre-sorting the dataset. In the Bitmap approach, each data point is encoded in a bit string and the skyline is computed on the bit matrix of all data points.

3. LASQ AUTHENTICATION METHOD

In this section, we propose a basic LASQ authentication method. We start with the authentication problem in a fixed subspace, and then extend it to arbitrary subspaces

3.1 LASQ Authentication

3.1.1 Design of Authenticated Index Structure. To expedite query processing, we index all the objects' subspace skyline scopes by an R*-tree, where the subspace skyline scopes are stored in the leaf nodes as data entries. Additionally, to support query authentication, we follow similar ideas of MB-tree and MR-tree to maintain a series of digests for all index nodes in the tree structure.

3.1.2 Server Query Processing and VO Construction. With the help of MSR-tree, an LASQ is reduced to a point-location

query on the indexed subspace skyline scopes. Specifically, starting from the root and going all the way down to the leaf nodes, the server checks whether any child of a node covers the query point. If it does, the node is *unfolded* and inserted into the VO for further checking; otherwise, the node is *pruned* and only its MBR and digest are inserted into the final VO. When visiting a leaf entry associated with an object o , if the corresponding S_o does not cover the query point, both S_o and H_o should be inserted into the VO; otherwise, o is an LASQ result and only S_o is inserted into the VO (H_o can be computed locally by the client based on the received result). It is noteworthy that as the nodes in the VO also form a tree structure, we call it a *VO-tree*. In general, a VO-tree contains four types of data:

- 1) the subspace skyline scopes of all objects in the visited leaf nodes;
- 2) the digests of non-result objects in the visited leaf nodes;
- 3) the MBRs of all visited non-leaf entries;
- 4) the digests of the pruned index nodes.

3.1.3 Client Result Verification. The VO-tree and the root signature ($Sig(H_{root})$), along with the skyline results, are sent to the client after query processing. To verify the correctness of the results, the client checks the following three conditions:

- 1) the subspace skyline scopes of all result objects should cover the query point q ;
- 2) no MBRs of the pruned index nodes and no subspace skyline scopes of the non-result objects cover q ;
- 3) the root signature matches the root digest computed from the VO-tree.

4 CONCLUSION

In this paper, we have studied the problem of authenticating location-based skyline queries in arbitrary subspaces (LASQs). We have proposed a basic MSR-tree authentication method by extending our previous work on skyline query authentication. To enable authentication for large scale datasets and subspaces, we have further proposed a Partial-S4-tree method, in which most of the redundant objects can be easily identified and filtered out from the VO. For authenticating continuous LASQs, we have proposed a prefetching-based solution to avoid frequent query issuances and VO transmissions. Extensive experimental results demonstrate the efficiency of our proposed methods and algorithms under various system settings. In particular, our proposed Partial-S4-tree method outperforms the basic authentication method by up to 69% in terms of the overall query latency and up to 74% in terms of the VO size.

5.FUTURE WORK

As for the future work, we will extend this work to road network environments. Since the query distance is defined by network distance in a road network, the skyline scope defined in this paper no longer works, which calls for new authentication methods. Moreover, we are also interested in studying the

authentication problem for dynamic objects, where how to guarantee the freshness of query results is a very challenging issue .

6. REFERENCES

- [1] (2011) AT&T to Launch Cloud-Based LBS Mobility Data Offering [Online]. Available: <http://www.mobilecommercedaily.com/2011/01/06/att-to-launch-cloud-based-lbs-mobility-data-offering>.
- [2] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: An efficient and robust access method for points and rectangles," in *SIGMOD*, Atlantic City, NJ, USA, 1990, pp. 322–331.
- [3] M. Berg, O. Cheong, and M. Kreveld, "Computational Geometry: Algorithms and Applications," 3rd ed., Berlin, Germany: Springer, 2008, ch. 7.
- [4] S. Borzsonyi, D. Kossmann, and K. Stocker, "The Skyline operator," in *Proc. ICDE*, Heidelberg, Germany, 2001, pp. 421–430.
- [5] Q. Chen, H. Hu, and J. Xu, "Authenticating Top-k queries in location-based services with confidentiality," in *PVLDB*, Hangzhou, China, 2014.
- [6] J. Chomicki, P. Godfrey, J. Gryz, and D. Liang, "Skyline with presorting," in *Proc. ICDE*, 2003.
- [7] H. Hu, J. Xu, and D. L. Lee, "A generic framework for monitoring continuous spatial queries over moving objects," in *SIGMOD*, Baltimore, MD, USA, 2005.
- [8] H. Hu, J. Xu, Q. Chen, and Z. Yang, "Authenticating locationbased services without compromising location privacy," in *SIGMOD*, 2012.
- [9] H. Hu, Q. Chen, and J. Xu. "VERDICT: Privacy-preserving authentication of range queries in location-based services," in *ICDE*, Brisbane, QLD, Australia, 2013 (Demo).