KNN Query Processing 0f Secured Multi Data Owner Using Voronoi Diagram

Jemi Florence[1], Kanya Devi J[2]

[1,2]*Department of Computer Science and Engineering*
[1,2]*Sri Shakthi Institute of Engineering and Technology*
L&T Bypass Road, Coimbatore, India

Abstract— **an economical and flexible way to data owners for delivering spatial data to users of location-based services would be Outsourcing spatial databases to the cloud. In this data outsourcing model the service provider can be untrustworthy. It may return incorrect or incomplete query results to clients intentionally or not. Ensuring spatial query integrity is critical. Here an efficient road network is proposed, with k-nearest-neighbor query verification technique which utilizes the network Voronoi diagram and neighbors to prove the integrity of query results. To deal with data integrity and authentication problem for outsourced databases Digital Signature is used. This Technique verifies both the distances and the shortest paths from the query point to its kNN results on the road network. Data owner builds an authenticated data structure before providing the data to the service provider. K-nearest-neighbor query verification technique utilizes the network Voronoi diagram and neighbors to prove the integrity of query results in ROAD Network.Skyline queries take into account both the spatial and non-spatial attributes of the points. For example instead of considering all the hotels, a nearest neighbor query would instead ask for the hotel that is closest to among those who provide services such as pool, internet at the same time.**

Keywords— **data outsourcing, location based services, service provider, road network, voronoi diagram, skyline queries.**

I. INTRODUCTION

Location-based services (LBS) are a general class of computer program-level services that use location data to control features. As such LBS is an information service and has a number of uses in social networking today as an entertainment service, which is accessible with mobile devices through the mobile network and which uses information on the geographical position of the mobile device. This has become more and more important with the expansion of the smart phone and tablet markets as well.LBS are used in a variety of contexts, such as health, indoor object search, entertainment work, personal life, etc.LBS include services to identify a location of a person or object, such as discovering the nearest banking cash machine. In spatial database real entities are modeled in geometric manner, for example location of hotels, hospital, restaurants are represented as points on maps, while larger area such as landscapes, lakes, parks are represented as a combination of rectangles. Spatial database system can used in geographic information system, in this range search can be utilized to find all cash machines in a certain area, while nearest neighbor retrieval can find the cash machines closer to a given address. 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 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. Queries in spatial database have become increasingly important in recent years with the increasing popularity of some services such as Google Earth and Yahoo Maps, as well as other geographic applications. Today, widely used of search engines has made it realistic to write spatial queries in a new way. Traditionally, queries focus on objects only geometric properties, for example whether a point is in rectangle or how two points are close from each other. Some new application allows users to browse objects based on both of their geometric coordinates and their

Associated texts. Such type of queries called as spatial keyword query.

II. SYSTEM ANALYSIS

Existing system allows processing of NN queries in an un trusted out sourced environment, and protecting both the POI and querying users positions.

This technique rely on mutable order preserving encoding (mOPE) which guarantees indistinguishability under ordered chosen-plaintext attack (IND-OCPA). It provides performance optimizations to decrease the computational cost inherent to processing on encrypted data, and we consider the case of incrementally updating datasets. This approach verifies *k*-nearest-neighbor results in the Euclidean space. We propose the VD-*k*NN method for secure NN queries which works by processing encrypted Voronoi diagrams. The method returns exact results, but it is expensive for *k*>1, and may impose a heavy load on the data owner. To address the limitations of VD-*k*NN, Top *k*NN method is introduced, that works by processing encrypted Delaunay triangulation that supports any value of *k* and decreases the load at the data owner. Main drawbacks of existing system are missing of authentication query processing and single data owner is used to find data path. To overcome these drawbacks authentication and multi data owner technique are introduced in the proposed system. Fig.2 illustrates the framework of data outsourcing model
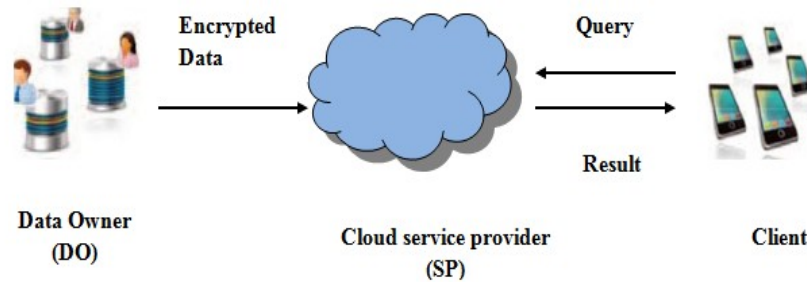
Fig.2. Data outsourcing model framework

Data owner provides the encrypted data to the cloud service provider. Cloud service provider or unauthorized third party could not change the data provided by the data owner. Data owner obtain private and public key from key generation center. When client queries the cloud service provider gives the encrypted result set with the help of private key client decrypts the result. An efficient road network k-nearest-neighbor query verification technique is proposed, which utilizes the network Voronoi diagram and neighbors to prove the integrity of query results. Both the distances and the shortest paths from the query point to its kNN results on the road network are needed to be verified. This approach is evaluated on real-world road networks together with both real and synthetic points of interest datasets. The k-nearest-neighbor search is a very important query type for supporting Geospatial applications. Since most mobile users move on roads in reality, nearest neighbor search algorithms have been extended to support queries on spatial networks.

An approach based on the network Voronoi diagram is used for evaluating NN queries on spatial networks. The key idea is to partition a large network into a number of small Voronoi regions and then pre-compute distances both within and across the regions. In order to speed up NN search in a spatial network a solution is introduced which explore the entire spatial network by pre-computing the shortest paths between all the vertices in the network and using a shortest path quad tree to capture spatial coherence. By using this approach the shortest paths between various vertices can be computed only once to answer different NN queries in a given spatial network.

III. RELATED WORK

In this section, we briefly introduce the related work on Nearest neighbour search technique, authenticated data Structure, voronoi diagram, Road networks.

A. Nearest neighbour search technique

This method is used to retrieve a group of spatial web objects such that the query's keywords are cover by group's keywords and objects are near to the query location and have the lowest inter object distances. This method addresses the two instantiation of the group keyword query. First is to find the group of objects that cover the keywords such that the sum of their distances to the query is minimized. Second is to find a group of objects that cover the keywords such that sum of the maximum distance among an object in group of objects and query and maximum distance among two objects in group of objects is minimized. Both of these sub problems are NP-complete. Greedy algorithm is used to provide an approximation solution to the problem that utilizes the spatial keyword index IR-tree to reduce the search space. But in some Application query does not contain a large number of keywords, for this exact algorithm is used that uses the dynamic programming. [1]

B. Authenticated data structure

To support the query verification, we need a well-defined *Authentication data structure* (**ADS**) built on the outsourced Data, which should be cryptographically signed by DO to ensure data integrity. Consider the outsourced database with a set of POIs P over an underlying road network G. An elaborate authentication data structure is used in order to support k-nearest-neighbor query verification on road networks where the distances from the query point to objects are measured by the shortest path on the graph with regard to the edge weight W. Given a set of POIs and the graph G, the network Voronoi diagram can be computed by applying parallel Dijkstra's algorithm where POIs are treated as multiple sources. In this algorithm, we employ a Fibonacci heap to expand the shortest path tree from all POIs in the graph until shortest path trees meet. For each POI, a set of border points is discovered and all the road segments in between the border points form the network Voronoi cell for that POI (generator). Note that the network Voronoi cell for any generator on the graph is unique because it contains a unique set of road segments.

C. Voronoi Diagram

Voronoi diagram is a partitioning of a plane into regions based on distance to points in a specific subsets of the plane. Partitioning of the plane into regions of equal nearest neighbors is the main feature of voronoi diagram. Basic properties of voronoi diagram (1) a planar graph where every vertex is of degree 3. (2) dual graph forms a special kind of triangulation called the Delaunay Triangulation. Algoriths for Voronoi diagrams (i) compute the intersection of n-1 half planes for each site and merge the cells into the diagram (ii) Divide and conquer (iii) Plane Sweep (iv) Randomized incremental construction.

IV. MODULES

In this section, the proposed system consists of five modules: User interface design, Road network and skyline query, Data provider and data owner process, authenticated query processing, shortest path on road network using voronoi diagram.

A. User interface design

User interface design or user interface engineering is the design of computers, appliances, machines, mobile communication devices, software applications and websites with the focus on the user's experience and interaction. The goal of user interface design is to make the user's interaction as simple and efficient as possible, in terms of accomplishing user goals what is often called user-centered design. To run remote control system GUI application in J2EE is developed. User can easily execute the project with the help of GUI. A social networking service is a platform to build social networks or social relations among people who, for example, share interests, activities, backgrounds or real-life connections. A social network service consists of a representation of each user (often a profile), his social links, and a variety of additional services.

B. Road Network and Skyline Query

A road network can be considered as a graph with positive weights. The nodes represent road junctions and each edge of the graph is associated with a road segment between two junctions. The weight of an edge may correspond to the length of the associated road segment, the time needed to traverse the segment or the cost of traversing the segment. Using directed edges it is also possible to model one-way streets. Such graphs are special in the sense that some edges are more important than others for long distance travel. All of these algorithms work in two phases. In the first phase, the graph is preprocessed without knowing the source or target node. The second phase is the query phase. In this phase, source and target node are known. The location-based skyline queries are to find the nearby spatial object with good services where the distance to the querying user is a spatial attribute and the goodness of the objects. The query returns a set of result that is closer to the client current position.

C. Data provider and data owner process

To scale up services there has been a rising trend of outsourcing data management to service providers, which provide query services to clients on behalf of data owners. Outsourcing spatial databases to the cloud provides an economical and flexible way for data owners to deliver spatial data to users of location-based services. **Data as a Service**, or DaaS, is a cousin of software as a service. Like all members of the as a Service (aaS) family, DaaS is based on the concept that the product, data in this case, can be provided on demand to the user regardless of geographic or organizational separation of provider and consumer. Additionally, the emergence of service-oriented architecture (SOA) has rendered the actual platform on which the data resides also irrelevant. This development has enabled the recent emergence of the relatively new concept of DaaS. 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.

D. Authenticated query processing

Authenticated data structures are a model of computation where untrusted responders answer queries on a data structure on behalf of a trusted source and provide a proof of the validity of the answer to the user.

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 Cloud service providers CSP, the data owner builds an authenticated data structure (ADS) of the dataset. To support efficient query processing, the ADS are 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.

E. Shortest path on road network using voronoi diagram

The shortest path between two given locations in a road network is an important problem that finds applications in various map services and commercial navigation products. The normal road network define the distance between two points is measured by the road network distance instead of their Euclidean distance, assuming objects can only move along street systems.

A road network system can be modeled as a weighted graph $G(V,E,W)$ consisting of a set of vertices $V = \{p1, p2, . , pn\}$ and a set of edges $E = \{e1, e2, ., em\}$ connecting vertices to form a graph. W represents the cost of each edge in E

Fig 4.2 illustrate the original road network is represented as a graph where $p1$, $p2$, and $p3$ are points of interest and $p4$-$p16$ are intersections on the road network.
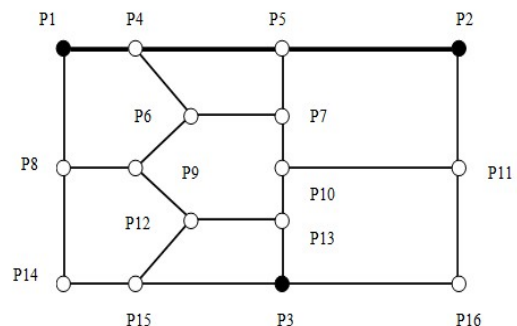

Fig.4.1 Road network

*p*1, *p*2, and *p*3 are points of interest and *p*4-*p*16 are intersections on the road network.The network Voronoi cell *V(pi)* contains all points on edges that are closer to *pi* than to any other POIs. It is actually a shortest path tree generated from *pi*, and hence *pi* is also called the *generator* of *V(pi)*. Note that, different from the Voronoi Diagram in the Euclidean space where each Voronoi cell is a continuous area, the network Voronoi cell of each generator contains a set of road segments.
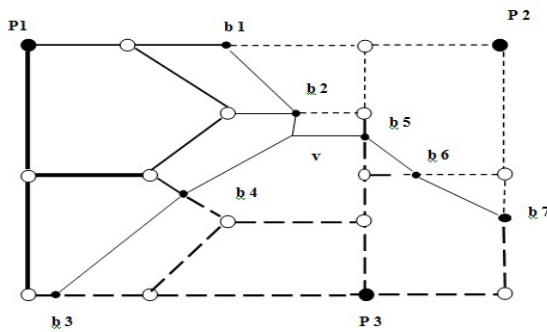
Fig.4.2 Network voronoi diagram

In Fig 4.2, V (p1), V (p2), and *V (p3)* are represented by line segments with different styles separated by point's *b*1-*b*7. Since objects are restricted on road segments, the network Voronoi cell of a specific generator is unique.

Given a set of points of interest, one can construct the network Voronoi diagram by expanding shortest path trees from each POI simultaneously until the shortest path trees meet.

The meeting points, termed as *border points*, are also on the edges of the road network with the property that the costs (e.g., road network distances) from the meeting point to the two neighboring POIs are equal to each other.

V. EXPERIMENTS

In this section, we present various experiments to evaluate the effectiveness of the proposed system.

A. Secured Query Processing

Data owner builds an authenticated data structure before providing the data to the service provider. To deal with data integrity and authentication problem for outsourced databases Digital Signature is used.Data integrity and Authentication is achieved via digital signature.

B. Shortest path based on voronoi diagram

The query verification problem for k-nearest-neighbor queries on road networks is addressed. With the help of a network Voronoi diagram-based verification both distance and shortest path to KNN query is achieved. ROAD network technique verifies both the distances and the shortest paths from the query point to its kNN results on the road network.

VI. CONCLUSIONS

The project identifies the problem of authenticating location-based skyline queries in arbitrary subspaces (LASQs). A basic MSR-tree authentication method is proposed by extending the previous work on skyline query authentication. To enable authentication for large scale datasets and subspaces, a Partial-S4-tree method is proposed, in which most of the redundant objects can be easily identified and filtered out from the VO. Data integrity and Authentication is achieved via digital signature.The query verification problem for k-nearest-neighbor queries on road networks is addressed.With the help of a network Voronoi diagram-based verification both distance and shortest path to KNN query is achieved.

For authenticating continuous LASQs, a perfecting-based solution is proposed to avoid frequent query issuances and VO transmissions. Extensive experimental results demonstrate the efficiency of the proposed methods and algorithms under various system settings. In future work (i) to reduce time, optimization methods can be used. (ii) to handle highly frequent network updates, such as traffic changes during rush hours.

REFERENCES

[1]   A. Chen, H. Hu, and J. Xu, "Authenticating Top-k queries in location-based services with confidentiality," in *PVLDB*, Hangzhou, China, 2014.

[2]   C. Chomicki, P. Godfrey, J. Gryz, and D. Liang, "Skyline with presorting," in *Proc. ICDE*, 2003.

[3]   D. Kossmann, F. Ramsak, and S. Rost, "Shooting stars in the sky: An online algorithm for skyline queries," in *VLDB*, Hong Kong, China, 2002.

[4]   E. Mykletun, M. Narasimha, and G. Tsudik, "Signature bouquets: Immutability for aggregated/condensed signatures," in *ESORICS*, Sophia Antipolis, France, 2004.

[5]   Feifei Li,Marios Hadjileftheriou ,George Kollios, and LeonidReyzin "Dynamic Authenticated Index Structures For Outsourced Databases" Computer Science Department, Boston University ,2 AT&T Labs Inc.

[6]   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.

[7]   H. Hu, J. Xu, Q. Chen, and Z. Yang, "Authenticating locationbased services without compromising location privacy,"in *SIGMOD*, 2012.

[8]   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).

[9]   L. Hu, W.-S. Ku, S. Bakiras, and C. Shahabi, "Spatial query integrity with voronoi neighbors," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 863–876, Apr. 2013.

[10]   M. Berg, O. Cheong, and M. Kreveld, "*Computational Geometry: Algorithms and Applications*," 3rd ed., Berlin, Germany: Springer,2008, ch. 7.

[11]   N. Huang, H. Lu, B. C. Ooi, and K. H. Tong, "Continuous skyline queries for moving objects," *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 12, pp. 1645–1658, Dec. 2006.

[12]   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.

[13]   R. C. Merkle, "A certified digital signature," in *CRYPTO*, Santa Barbara, CA, USA, 1989.

[14]   S. Borzonyi, D. Kossmann, and K. Stocker, "The Skyline operator," in *Proc. ICDE*, Heidelberg, Germany, 2001, pp. 421–430.

[15]   Y. Gao and B. Zheng, "Continuous obstructed nearest neighbor queries in spatial databases," in *Proc. ACM SIGMOD*, Providence, RI, USA, 2009, pp. 557–590.