A. Deena Mercy *et al*, International Journal of Computer Science and Mobile Applications,
Vol.3 Issue. 9, September- 2015, pg. 9-13 **ISSN: 2321-8363**

COLLABORATIVE LOCATION-BASED NOVEL SCHEMES FOR TRUSTED TOP-K QUERY RETRIEVAL

G. Angeline Prasanna M.Sc., M.Phil., M.C.A.,(Ph.D)[1], **A. Deena Mercy**[2]

1. HOD of CA&IT Department, Kaamadhenu Arts & Science College
2. M.Phil. Scholar, Kaamadhenu Arts & Science College

Abstract:

Due to the explosive growth of Internet-capable and location-aware mobile devices, information generation and sharing become increasingly popular. Collaborative location-based approach is used for selecting top-k query based on Point Of Interest (POI). This system consists of a data collector, data contributors, location-based service providers (LBSPs), and system users. The data collector collects reviews about points-of-interest (POIs) from data contributors, while LBSPs buy POI data sets from the data collector and allow users to do spatial top-k queries which invite the POIs in a certain region and with the highest k ratings for an interested POI attribute. But LBSPs are untrusted and may hark back fake query results for various bad motives, e.g., in favour of POIs willing to pay. Here novel schemes are used to detect fake query result and moving it to top-k query results set. The efficacy and efficiency of our schemes are thoroughly analysed and evaluated.

Keywords: Spatial top-k query, location-based service, security**.**

1. INTRODUCTION

The flourish of smart phones contributes prosperity to location based services (LBSs) in nearly all social and business sectors, such as geo-social networks, merchandizing, marketing, and logistics. As these LBSs cause new business opportunities, there is a developing need of the mobile users to affirm the legitimacy of service results, such as a heeling of recommended local restaurants grouped by location and user rating. This moment is even more critical in an outsourced model where businesses (or data owners) publish their data to a 3rdparty service provider (SP), who handles LBS queries based on these data. As the SP is supposed to spoof query consequences in grace of their "sponsors", to sustain growth among fierce competition, it will soon be compelled to provide exploiters not only the effects, but also the proof of rightness. The data owner publishes not only data (e.g., spatial objects) to the SP, but also the endorsements of the data establishing exposed. These endorsements are signed by the data possessor fiddling with by the SP. Given a query, the SP brings back both the query results and a proof, named verification object (VO).

The VO is used by the verification phase, to rebuild the endorsements and thus assert the rightness of the results. However, one key restriction of all these works is that throughout the verification phase, the client is presumed to be completely hoped and ennobled to receive any data values, even though they are not part of the consequences. Unfortunately, this presumption is flawed in LBSs whose data is often sensitive locations and ought to remain secret against the client. For example, in online real-estate sites, the address of an attribute is often suppressed as business confidentiality. There is a call for privacy-preserving query authentication proficiencies in LBSs that assure the confidentiality of location data against the client. There is also a proposed privacy-preserving authentication for location-based wander queries. A location-based advertisement and recommendation are often recognized as one of the most productive LBS businesses and thus provoke the greatest controversy with their ranking results, here it is examined the privacy-preserving authentication for location-based top-k queries, where the rank assess of an object is a linear compounding of distance penalty and non-spatial score (e.g., user average rating). This query resolution is like to an abstraction of several location-based topk queries defined in [11, 29] and even the k-nearest neighbour (k NN) queries. The first gainsay of

privacy-preserving location-based top-k queries is its security framework. The effects of a top-k query imply the relative ranking of various objects. To address this, a formal security model founded on the computational is introduced. Second, the major cryptographic challenge of this problem is comparing the rank values of 2 objects without disclosing their locations or scores.

2. Related Work

Our work is most related to data outsourcing [4], for which we can only review representative schemes due to space constraints. The framework of data outsourcing was first introduced in [4], in which a data owner outsources its data to a third-party service provider who is responsible for answering the data queries from either the data owner or other users. In general, there are two security concerns in data outsourcing: data privacy and query integrity [5].Ensuring data privacy requires the data owner to outsource encrypted data to the service provider, and efficient techniques are needed to support querying encrypted data.

A bucketization approach was proposed in [6], [7] to enable efficient range queries over encrypted data, which was recently improved in [8]. Shi et al. presented novel methods for multi-dimensional range queries over encrypted data[9]. Some most recent proposals aim at secure ranked keyword search [10], [11] or fine-grained access control [12]over encrypted data. This line of work is orthogonal to our work, as we focus on publicly accessible location-based data without need for privacy protection. Another line of research has been devoted to ensuring query integrity, i.e., that a query result is indeed generated from the outsourced data (the authenticity requirement)and contains all the data satisfying the query (the correctness requirement). In these schemes, the data owner outsources both its data and also its signatures over the data to the service provider which returns both the query result and verification object (VO) computed from the signatures for the querying user to verify query integrity. Many techniques were proposed for signature and VO generations, such as those [13], [14], [15] based on signature chaining and those [5], [16], [17], [18] based on the Merkle hash tree [19]or its variants. None of these schemes consider spatial top-k queries and are directly applicable to our intended scenario, as spatial top-k queries exhibit unique feature in that whether a POI is among the top-k is jointly determined by all the other POIs in the query region and that the query region cannot be predicted in practice. Secure remote query processing in tiered sensor networks[20], [21], [22], [23], [24] is also loosely related to our work here. These schemes assume that some master nodes are in charge of storing data from regular sensor nodes and answering the queries from the remote network owner. Various techniques were proposed in [20], [21], [22], [23] to ensure data privacy against master nodes and also enable the network owner to verify range-query integrity. Moreover, Zhang et al. [24] proposed efficient techniques for the network owner to validate the integrity of top-k queries.

3. System Model Construction

Construct a distributed system comprising a data collector, data contributors, LBSPs, and top-k query users. Common people constitute a data contributor, who will submit POI reviews to the data collector's website. The data collector normally stimulates the review submission and has necessary counter measures on malicious data contributors who provide a fake review. POI reviews based on the location is aggregated and sold to individual LBSPs by the data collector. For users LBSP operates to on a website to perform top-k queries over the purchased data set and may add some likeable functionalities to the query result such as street maps and photos. One pair of data collector and LBSPs is used for study. Based on POI categories the data set is classified as restaurants, bars, and coffee shops. There will be a unique record for all POI in each category. Here top k query consisting of single category is considered. The data collector covers a geographic area, which is partitioned into M >= 1 equally-sized non-overlapping zones. For every zone i, let ni denote the number of POIs and $POI_{i,j}$ and $D_{i,j}$ denote the j^{th} POI and its corresponding data record, respectively. Assumed that n_{ij} data contributors provide a review about $POI_{i,j}$ to the data collector. A rating on every attribute and text comments are included in every review. The data record di,j for POIi,j includes its name, location $l_{i,j,}$ review sn_{ij}, and possibly other information.

4. Creating a snapshot of Top- K Query Processing

Creation of Snapshot involves three phases. The first phase is data pre-processing in which, the data set is pre-processed by data collector before selling it to LBSP's. The data collector sorts Di according to the attribute-q rating to generate a systematiser list, the data collector chains the ordered POI using cryptographic hash function. This is done to ensure the correct order among them. Finally, a Merkle hash tree is built by the data

collector to enabling effective authentication of query results. Query processing is the second phase. In this, the data sets of interested POI categories are purchased by LBSP from the data collector. For every POI category selected by the LBSP, the data collector returns the original data set D, the signatures on Merkle root hashes, and all the intermediate results for constructing the Merkle hash tree. The processing of snapshot top-k query includes the desired POI category, the interested attribute q, for ranking POIs the query region R, and k are needed. k POIs in R with the highest k attribute-q is denoted by k, and the lowest attribute q rating is denoted by g. In addition, each zone either completely or partially covered by the query region is called as a candidate zone. A right and veritable query result needs to satisfy two conditions. The rightness condition needs the query result to contain at least the following information: (1) the accomplished data records for k POI;(2) the data indexes for entire the POIs in each candidate zone but not in R whose attribute-q rating is more prominent than g; and (3) approximate information is needed to prove that the query solution includes either the data record or index of every POI in every candidate zone with attribute-q rating not smaller than g. In addition, the variability condition requires that the query result include the ancillary set for every candidate zone for the computation and confirmation of the q^{th} Merkle root hash. The Query result verification is third phase. A small plug-in developed by the data collector which is installed in the web browser is used for authentication and query result. This is done by user.

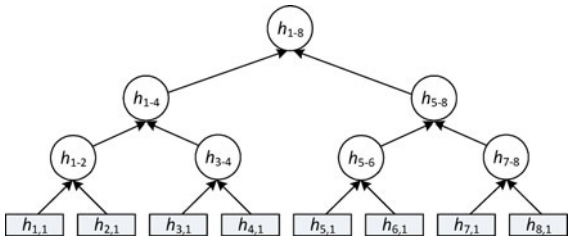

Fig. 1. An example of constructing the Merkle hash tree over.

The user checks that every piece of information in the query result will lead to the same Merkle root hash which matches the data collector's signature. This is done for authentication purpose. If the query result is authentic, the user can deduce the same root hash for each i to I, by which further verification is done to check whether the data collector's signature in the query result is a valid signature on the derived root hash. If so, it is considered that the query result is authentic.

5. Moving Top-k query processing

For inciting a top-k query, the user consequences a sequence of snapshot top-k queries according to a query schedule. As a reception to Q a, k POI is received from LBSP, the usersets a timer of length ~t. Then Qaþ1 is issued. When the timer started or when the 1st POI in k POI is no longer in the user moving query region, which comes first. To alleviate query treating at the LBSP, it admits both an integer id which uniquely identifying this moving top-k query and a one-bit flag pointing whether Qa is the last snapshot query for this moving top-k query. On experiencing query Q1, the LBSP lays kPOI1 in the query region R1 and so brings back a complete query outcome constructed as in Scheme 1. Along with this the LBSP records id, k POI1and R1, to relieve the working of a subsequent snapshot top-k queries with the similar going to top-k identifier id. Then it processes any subsequent query Qb (b > a)as follows. Without loss of abstraction, it is considered that the last finish query result the LBSP brought back is are ply to Q a (a >=1), which comprises k PO I a in Ra. The LBSP places the top-k POIs (i.e., k P OI b) in the query region R b which has attribute-q evaluations are amongst the most prominent k. The recorded query region is retrieved by LBSP. Next, the LBSP calculates a confirmation region to find the set of zones either completely or partially covered by R b µ V a->b, denoted by I a->b. For all snapshot top-k query Qb of the similar moving top-k query, the LBSP (if benign) will return a complete query consequence if b > 1 or there has been any alteration in the top-k POIs, or an ACK is returned if b > 1 and the antecedent brought back top-k POIs are still formalized. First, user can immediately say that the result is incorrect, if the user experiences an ACK when Q b is the final snapshot query. Second, User marks this query result as unverified and waits for the adjacent accomplished query result, if receiving an ACK when Q b is not the final snapshot query.

Performance Evaluation:

The user's computation overhead for k = 5 showing an impact of d have been analysed. Here the single signature verification is not included for brevity.

Results closely match under specified schemes. Along with this, the user's computation overhead increases with d nether earlier scheme, while it increases initially as d go from 1to 10 and then are relatively stable under our scheme. This is because the earlier scheme requires the LBSP to return information for every zone in R for the user to verify. This shows larger d, the higher the user's computation overhead in Scheme 1. But our scheme needs the LBSP to return information only for the zones that have at least one POI among the top-k POIs beneath the simulation settings, and there are all most k such zones in R. This shows that our scheme has lower calculation overhead on the user for small k and large d.

Conclusion:

For collaborative location-based information generation and sharing, a novel distributed system is used. By this approach, the queries are retrieved based on POI. The fake POI values and fake top most queries are removed by using secure feature like authentication. The key value which is used to construct the tree is also encrypted so that it is not possible to add a fake query to the list. Thus this approach provides a secure and correct query being placed on top k positions.

References:

[1] H. Yu, P. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks," IEEE/ACM Trans. Networking, vol. 18, no. 3, pp. 885-898, June 2010.
[2] H. Hacig€um€us, S. Mehrotra, and B. Iyer, "Providing Database as a Service," Proc. IEEE 18th Int'l Conf. Data Eng. (ICDE), Feb. 2002.
[3] W.-S. Ku, L. Hu, C. Shahabi, and H. Wang, "Query Integrity Assurance of Location-Based Services Accessing Outsourced Spatial Databases," Proc. Int'l Symp. Advances in Spatial and Temporal Databases, July 2009.
[4] H. Hacig€um€u¸s, B. Iyer, C. Li, and S. Mehrotra, "Executing SQL over Encrypted Data in the Database-Service-Provider Model," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD'02), pp. 216-227, 2002.
[5] B. Hore, S. Mehrotra, and G. Tsudik, "A Privacy-Preserving Index for Range Queries," Proc. 30th Int'l Conf. Very Large Data Bases (VLDB'04), pp. 720-731, Aug. 2004.
[6] B. Hore, S. Mehrotra, M. Canim, and M. Kantarcioglu, "Secure Multidimensional Range Queries over Outsourced Data," TheVLDB J., vol. 21, no. 3, pp. 333-358, 2012.
[7] E. Shi, J. Bethencourt, H. Chan, D. Song, and A. Perrig,"Multi-Dimensional Range Query over Encrypted Data," Proc. IEEE Symp. Security and Privacy (S&P'07), pp. 350-364, May2007.
[8] N. Cao, Z. Yang, C. Wang, K. Ren, and W. Lou, "SecureRanked Keyword Search over Encrypted Cloud Data, Proc. IEEE 30th Int'l Conf. Distributed Computing Systems (ICDCS'11),June 2011.
[9] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-PreservingMulti-Keyword Ranked Search over Encrypted Cloud Data,"Proc. IEEE INFOCOM, Apr. 2011.
[10] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Secure, Scalable,and Fine-Grained Access Control in Cloud Computing," Proc.IEEE INFOCOM'10, Mar. 2010.
[11] H. Pang and K.-L. Tan, "Verifying Completeness of RelationalQuery Answers from Online Servers," ACM Trans. Information andSystem Security, vol. 11, no. 2, pp. 1-50, Mar. 2008.
[12] M. Narasimha and G. Tsudik, "Authentication of Outsourced Databases Using Signature Aggregation and Chaining," Proc. 11thInt'l Conf. Database Systems for Advanced Applications (DASFAA'06),pp. 420-436, Apr. 2006
[13] H. Pang, J. Zhang, and K. Mouratidis, "Scalable Verification for Outsourced Dynamic Databases," Proc. VLDB Endowment, vol. 2,no. 1, pp. 802-813, 2009.
[14] Y. Yang, S. Papadopoulos, D. Papadias, and G. Kollios, "Spatial Outsourcing for Location-Mased Services," Proc. IEEE 24th Int'lConf. Data Eng. (ICDE), pp. 1082-1091, Apr. 2008.
[15] M. Yiu, Y. Lin, and K. Mouratidis, "Efficient Verification of Shortest Path Search via Authenticated Hints," Proc. IEEE 26th Int'l Conf. Data Eng. (ICDE), pp. 237-248, Mar. 2010.
[16] M. Yiu, E. Lo, and D. Yung, "Authentication of Moving kNN Queries," Proc. IEEE 27th Int'l Conf. Data Eng. (ICDE), pp. 565-576, Apr. 2011.
[17] R. Merkle, "A Certified Digital Signature," Proc. Ninth Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO), pp. 218-238, Aug. 1989.
[18] B. Sheng and Q. Li, "Verifiable Privacy-Preserving Range Query in Sensor Networks," Proc. IEEE INFOCOM'08, pp. 46-50, Apr. 2008.

[19] J. Shi, R. Zhang, and Y. Zhang, "Secure Range Queries in Tiered Sensor Networks," Proc. IEEE INFOCOM'09, Apr. 2009.
[20] R. Zhang, J. Shi, and Y. Zhang, "Secure Multidimensional Range Queries in Sensor Networks," Proc. ACM MobiHoc'09, pp. 197-206, May 2009.
[21] F. Chen and A. Liu, "Safe Q: Secure and Efficient Query Processing in Sensor Networks," Proc. IEEE INFOCOM'10, pp. 1-9, Mar. 2010.
[22] R. Zhang, J. Shi, Y. Liu, and Y. Zhang, "Verifiable Fine-Grained Top-K Queries in Tiered Sensor Networks," Proc. IEEE INFOCOM'10, Mar. 2010.

