



Top-K Query Dispensation In Secure Spatial Environment Via Untrusted Location-Based Service Providers

V. JAYALAKSHMI

PG Scholar, Department of CSE
G.K.C.E, Sullurpet, Andhrapradesh, INDIA.

Y. MADHU SEKHAR

Assistant Professor, Department of CSE
G.K.C.E, Sullurpet, Andhrapradesh, INDIA.

Abstract: IN Collaborative location-based information generation and sharing is considered, which become increasingly popular due to the explosive growth of Internet-capable and location-aware mobile devices. The system consists of a data collector, data contributors, location-based service providers (LBSPs), and system users. The data collector gathers reviews about points-of-interest (POIs) from data contributors, while LBSPs purchase POI data sets from the data collector and allow users to perform spatial top-k queries which ask for the POIs in a certain region and with the highest k ratings for an interested POI attribute. In practice, LBSPs are untrusted and may return fake query results for various bad motives, e.g., in favor of POIs willing to pay. Three novel schemes is used for users to detect fake spatial snapshot and moving top-k query results as an effort to foster the practical deployment and use of the proposed system.

Keywords: Location-Based; Querying; Collector;

I. INTRODUCTION

The explosive growth of Internet-capable and location aware mobile devices and the surge in social network usage are fostering collaborative information generation and sharing on an unprecedented scale. Almost all smart phones have cellular/ Wi-Fi Internet access and can always acquire their precise locations via pre-installed positioning software. Also owing to the growing popularity of social networks, it is more and more convenient and motivating for mobile users to share with others their experience with all kinds of points of interests (POIs) such as bars, restaurants, grocery stores, coffee shops, and hotels. Meanwhile, it becomes commonplace for people to perform various spatial POI queries at online location-based service providers (LBSPs) such as Google and Yelp. As probably the most familiar type of spatial queries, a spatial (or location-based) top-k query asks for the POIs in a certain region and with the highest k ratings for a given POI attribute. For example, one may search for the best 10 Italian restaurants with the highest food ratings within five miles of his current location.

II. RELATED PROBLEM

There are two essential drawbacks with current top-k query services. First, individual LBSPs often have very small data sets comprising POI reviews. This would largely affect the usefulness and eventually hinder the more prevalent use of spatial top-k query services. The data sets at individual LBSPs may not cover all the Italian restaurants within a search radius. Additionally, the same restaurant may receive diverse ratings at different LBSPs, so users may get confused by very different query results from different LBSPs for the same

query. A leading reason for limited data sets at individual LBSPs is that people tend to leave reviews for the same POI at one or at most only a few LBSPs's websites which they often visit. Second, LBSPs may modify their data sets by deleting some reviews or adding fake reviews and return tailored query results in favor of the restaurants that are willing to pay or against those that refuse to pay. Even if LBSPs are not malicious, they may return unfaithful query results under the influence of various attacks such as the Sybil attack whereby the same attacker can submit many fake reviews for the same POI. In either case, top-k query users may be misled by the query results to make unwise decisions. To introduce some trusted data collectors as the central hubs for collecting POI reviews is a good solution. In particular, data collectors can offer various incentives, such as free coffee coupons, for stimulating review submissions and then profit by selling the review data to individual LBSPs. Instead of submitting POI reviews to individual LBSPs, people can now submit them to a few data collectors to earn rewards. The data sets maintained by data collectors can thus be considered the union of the small data sets currently at individual LBSPs. Such centralized data collection also makes it much easier and feasible for data collectors to employ sophisticated defenses, to filter out fake reviews from malicious entities like Sybil attackers. Data collectors can be either new service providers or more preferably existing ones with a large user base, such as Google, Yahoo, Facebook, Twitter, and MSN. Many of these service providers have already been collecting reviews from their users and offered open APIs for exporting selected data from their systems.

The above system model is also highly beneficial for LBSPs. In particular, they no longer need struggle to solicit faithful user reviews, which is often a daunting task especially for small/medium-scale LBSPs. Instead, they can focus their limited resources on developing appealing functionalities (such as driving directions and aerial photos) combined with the high-quality review data purchased from data collectors. The query results they can provide will be much more trustworthy, which would in turn help them attract more and more users. This system model thus can greatly help lower the entrance bar for new LBSPs without sufficient funding and thus foster the prosperity of location-based services and applications.

A main challenge for realizing the appealing system above is how to deal with untrusted and possibly malicious LBSPs. Specifically, malicious LBSPs may still modify the data sets from data collectors and provide biased top-k query results in favor of POIs willing to pay. Even worse, they may falsely claim generating query results based on the review data from trusted data collectors which they actually did not purchase. Moreover, nonmalicious LBSPs may be compromised to return fake top-k query results.

III. PROBLEM ANALYSIS

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 to enable efficient range queries over encrypted data, which was recently improved.

Shi et al. presented novel methods for multi-dimensional range queries over encrypted data. Some most recent proposals aim at secure ranked keyword search or fine-grained access control over encrypted data.

Ensuring query integrity is studied, i.e., that a query result is indeed generated from the outsourced data and contains all the data satisfying the query. 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 a verification object (VO) computed from the signatures for the querying user to verify query integrity.

Many techniques were proposed for signature and VO generations, based on signature chaining and based on the Merkle hash tree. Secure remote query processing in tiered sensor networks is also studied. 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. None of these schemes consider spatial top-k queries

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.

To propose three novel schemes to tackle the special top k query processing challenge for fostering the practical deployment and wide use of the envisioned system. The key idea is that the data collector pre-computes and authenticates some auxiliary information about its data set, which will be sold along with its data set to LBSPs. To faithfully answer a top-k query, a LBSP need return the correct top-k POI data records as well as proper authenticity and correctness proofs constructed from authenticated hints. The authenticity proof allows the query user to confirm that the query result only consists of authentic data records from the trusted data collector's data set, and the correctness proof enables the user to verify that the returned top-k POIs are the true ones satisfying the query.

The first two schemes both target snapshot top-k queries but differ in how authenticated hints are precomputed and how authenticity and correctness proofs are constructed and verified as well as the related communication and computation overhead. The third scheme, built upon the first scheme, realizes efficient and verifiable moving top-k queries.

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.

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 to enable efficient range queries over encrypted data, which was recently improved.

Shi et al. presented novel methods for multi-dimensional range queries over encrypted data. Some most recent proposals aim at secure ranked keyword search or fine-grained access control over encrypted data.

Ensuring query integrity is studied, i.e., that a query result is indeed generated from the outsourced data and contains all the data satisfying the query. 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 a verification object (VO) computed from the signatures for the querying user to verify query integrity.

Many techniques were proposed for signature and VO generations, based on signature chaining and based on the Merkle hash tree.

None of these schemes consider spatial top-k queries and, 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 is also studied. 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.

IV. IMPLEMENTATION

Data contributor

Data contributors are common people who submit POI reviews to the data collector's website. Data contributor can submit the POI information with complete details of the object found nearby.

Data collector

Data collectors are considered to be trusted and as the central hubs for collecting POI reviews. Instead of submitting POI reviews to individual LBSPs, people can now submit them to a few data collectors to earn rewards. The data sets maintained by data collectors can thus be considered the union of the small data sets currently at individual LBSPs. Such centralized data collection also makes it much easier and feasible for data collectors to employ sophisticated defenses, to filter out fake reviews from malicious entities like Sybil attackers. Data collectors can be either new service providers or more preferably existing ones with a large user base.

Location based server

The data collector sells aggregated POI reviews in the form of a location-based data set to individual LBSPs. Every LBSP operates a website for users to perform top-k queries over the purchased data set and may add some appealing functionalities to the query result.

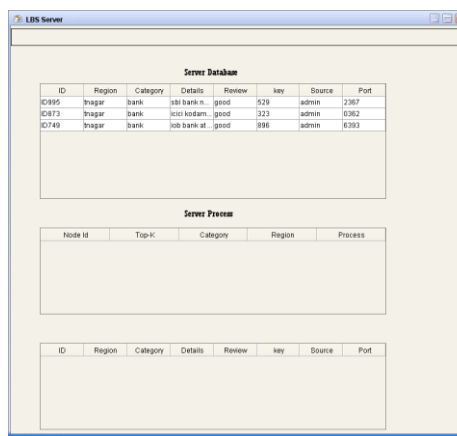
User query

User rise a query to location based server. This model enables the user to verify the authenticity and correctness of the query result returned by the LBSP. The query result is considered authentic if all its k POI records exist in the data collector's data set and have not been tampered with, and it is called correct if it contains the true top-k POI records in the query region.

Top-k query processing

Top- k results will be given for user. The LBSP purchases the data sets of interested POI categories from the data collector. For every POI category selected by the LBSP, the data collector returns the original data set D, the signatures on root hashes, and all the intermediate results for constructing the Merkle hash tree. Alternatively, the data collector can just return the first two pieces of information and let the LBSP itself perform a onetime process to derive the third piece in the same way as the date collector.

V. RESULT ANALYSIS



Server Database

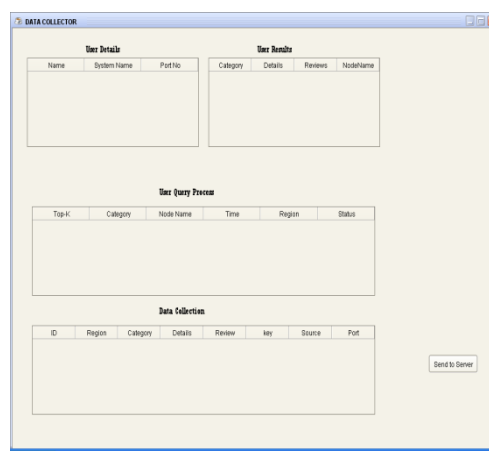
ID	Region	Category	Details	Review	key	Source	Port
ID955	Tnagar	bank	ubi bank n.	good	529	admin	2387
ID473	Tnagar	bank	icici kodam.	good	323	admin	8362
ID749	Tnagar	bank	job bank st.	good	886	admin	6393

Server Process

Node ID	Top-K	Category	Region	Process

Server Database (Bottom)

ID	Region	Category	Details	Review	key	Source	Port



User Details

Name	System Name	Profile

User Results

Category	Details	Reviews	NodeName

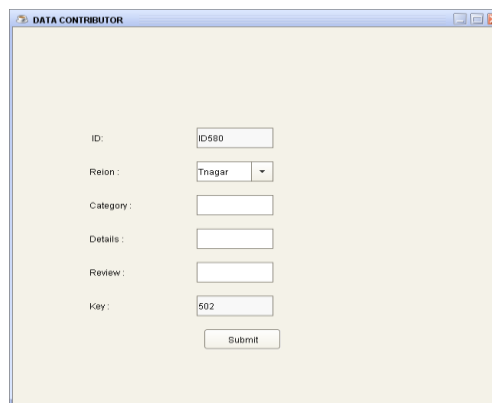
User Query Process

Top-K	Category	Node Name	Time	Region	Status

Data Collection

ID	Region	Category	Details	Review	key	Source	Port

Send to Server



DATA CONTRIBUTOR

ID:

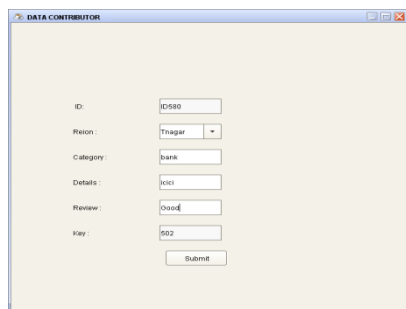
Region:

Category:

Details:

Review:

Key:



DATA CONTRIBUTOR

ID:

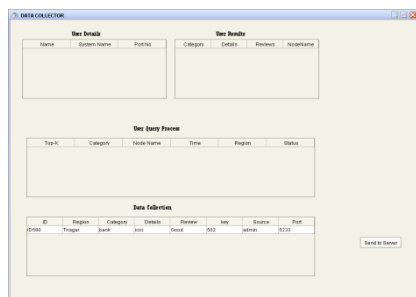
Region:

Category:

Details:

Review:

Key:



DATA COLLECTOR

User Details

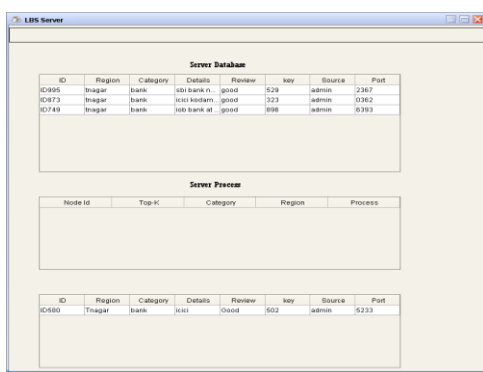
Name	System Name	Profile	Category	Details	Review	Note/Name

User Query Process

Top-K	Category	Note/Name	Time	Region	Status

Data Collection

ID	Region	Category	Details	Review	Key	Source	Port
IC580	Tr Nagar	bank	ic51	Good	502	admin	5233



LBS Server

Server Database

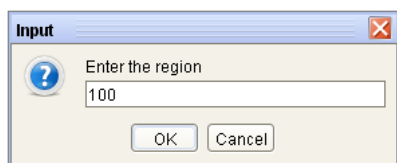
ID	Region	Category	Details	Review	Key	Source	Port
IC585	Tr Nagar	bank	ic51 bank n.	good	528	admin	2387
IC673	Tr Nagar	bank	ic51 bank n.	good	523	admin	0362
IC749	Tr Nagar	bank	ic51 bank n.	good	898	admin	0393

Server Process

Node ID	Top-K	Category	Region	Process

Server Data

ID	Region	Category	Details	Review	Key	Source	Port
IC580	Tr Nagar	bank	ic51	Good	502	admin	5233



Input

Enter the region

VI. CONCLUSION

Considered a novel distributed system for collaborative location-based information generation and sharing. Three novel schemes to enable secure top-k query processing via untrusted LBSPs for fostering the practical deployment and wide use of the envisioned system. Proposed schemes support both snapshot and moving top-k queries, which enable users to verify the authenticity and correctness of any top-k query result. The efficacy and efficiency of schemes are thoroughly analyzed and evaluated through detailed simulation studies.

VII. REFERENCES

- [1] R. Zhang, Y. Zhang, and C. Zhang, "Secure Top-k Query Processing via Untrusted Location-Based Service Providers," Proc. IEEE INFOCOM '12, Mar. 2012.
- [2] H. Yu, M. Kaminsky, P. Gibbons, and A. Flaxman, "SybilGuard: Defending against

Sybil Attacks via Social Networks," IEEE/ACM Trans. Networking, vol. 16, no. 3, pp. 576-589, June 2008.

- [3] 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.
- [4] H. Hacigümüş, S. Mehrotra, and B. Iyer, "Providing Database as a Service," Proc. IEEE 18th Int'l Conf. Data Eng. (ICDE), Feb. 2002.
- [5] 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.
- [6] H. Hacigümüş, 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.
- [7] 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.

AUTHOR'S PROFILE



V. Jayalakshmi has received her B.Tech degree in Computer Science Engineering from Narayana engineering college Nellore, JNTU Hyderabad in 2007 & she is now pursuing the

M.Tech Degree at Gokula Krishna College of Engineering at Sullurpet, Nellore (dist), Andhra Pradesh, India. Her areas of research include Secure Spatial Top-K Query.



Y. Madhu Sekhar has received his B.Tech degree Computer Science & Engineering from Priyadarshini college of Engineering from J.N.T.U Hyderabad in 2005 & M.Tech degree in Computer science & engineering from

Chadalawada Ramanamma Engineering College from J.N.T.U, Ananthapur in 2010. He is a life member of MISTE, he is dedicated to teaching field from the last 8 years, his last research areas included Data mining and warehousing and also software testing, at present he is working as an assistant professor in Gokula Krishna College of Engineering at Sullurpet, Nellore (dist), A.P., India.