Secure Query Processing Over Encrypted Data via Location Based Service Provider

[1] Dr. R.Rajalakshmi [2] K.L. Nikhitha [3] Rekha. D [4] Vaishali. C
Department of Information Technology,
Valliammai Engineering College, Kanchipuram DT, India
[1] klnikhitha94@gmail.com [2]rekha.dhakshna@gmail.com [3] vaishali1995@gmail.com

***Abstract* — There is a need for secure query processing system due to the explosive growth of Internet- capable and location-aware mobile devices. The existing system consists of an untrusted location based service provider, which is also not unified and no user identity. In order to overcome these problems a secure query processing system is proposed which consists of a certificate authority (CA), data owner, location-based service providers (LBSPs) and system users. The data owner uploads its encrypted details and sends request to the CA for authentication, while LBSP collects and aggregates data sets from the data owner. The user sends query to the LBSP to perform the secure query processing. Based on the query the data is processed. In order to make the application more secure, the encrypted data will be decrypted by the user based on the CA's private key and the user's identity from the user. Thus by using the certificate authority we provide a secure location-based services to the system users.**

***Index Terms* - Location *Based* Service Providers, Certificate Authority, Query processing, Security.**

I. INTRODUCTION

Internet has seen a massive growth in recent times and there are now various applications for providing location based services. The number of mobile users worldwide will surpass 2 billion in 2016, according to new figures from E-marketer— after nearly getting there in 2015. Next year, there will be over 1.91 billion smart phone users across the globe, a figure that will increase another 12.6% to near 2.16 billion in 2016 [1]. As all the mobile phones have Mobile data configuration or Wi-Fi facilities it always acquire accurate location through GPS (Global Positioning System). As a result users will have to browse through various websites to know about the services based on their location. This location-based services can be a query-based and provide the end user with useful information such as "Where is the nearest ATM?". This system primarily focuses on how the user's query is processed securely and reliably.

The existing system is mostly related to data outsourcing in which the individual LBSPs have space constraints i.e. the individual LBSP will have limited amount of memory to store the data sets. The data from the data owner is outsourced to the third- party service provider which in turn answers the queries from the user. In general, the users can share their interest through on-line LBSPs such as Google, Bing etc. For example one may search for nearest. Restaurants, Hospital, banks etc. within their particular radius.

Some of the disabilities of the existing system are observed. First, the third party location based service provider which collects the data sets from the data owner are found to be untrusted. The LBSPs are said to be untrusted as they might modify the data sets and return fake query results to the user. Second, the queries are processed through a number of individual LBSPs which may lead to chaos. Moreover Individual LBSPs consist of small data sets lead to constrained query processing i.e. it does not cover the entire services within the particular radius. Third, all the LBSPs are not unified i.e. it consists of many number of LBSPs. Fourth; the users are assumed to be unauthorized.

The best and efficient solution for the above problems is to build a Certificate Authority (CA) to authenticate the outsourced data to a single trusted LBSP. Many untrusted LBSPs are combined into a single trusted LBSP which collects and aggregates the data owner's details. Certification authority is introduced to enhance the security. The user sends query to the LBSP to perform the secure top k-query processing. Based on the query the data is processed

In order to make the application more secure, the encrypted data will be decrypted by the user based on the CA's private key and the user's identity from the user. Thus

1- http://www.emarketer.com/Article/2- Billion-Consumers-Worldwide- Smartphones-by-2016/1011694 by using the certificate authority we provide secure location-based services to the system users.

The rest of the paper is organized as follows. Section 2 discusses the related work of this system. Section 3 explains about proposed work of the system. Section 4 presents how data is encrypted and stored on the service provider.

II. RELATED WORK:

Our work is most related to secure query processing via LBSP [1]. 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. The data owner assumes that the service provider should be trusted and hence encrypted data sets are sent to the service provider which can be decrypted only by the data owner for total data privacy [2].

To overcome the inference attack service provider within the particular perimeter is said to be trusted. Bucketization approach is introduced to partition the data sets to identify the data using bucket tags (crypto-index). Controlled diffusion algorithm is introduced to fine tune the bucketization approach [3].

Multi-dimensional query processing over encrypted data is introduced were decryption is done at the client side. In order to hide the sensitive information AES algorithm is used for encryption [4].

The growing availability and ever- increasing accessibility of location technology provides a fertile ground for location-based services. To grow further, security, privacy and assurance of location- based services must be ensured. Distinguishing assurance and privacy from security is relevant, since violations of the former two properties impact the (perceived) quality of the service, not its operation. Neither privacy, nor assurance in location- based services is sufficiently well understood [5].

Call Data Records (CDR) dataset and anonymization techniques are used. A call record is created when a call originates or terminates on the cellular network and it contains various fields of information regarding that call [6].

The system adopts a server-client architecture, where the server and clients communicate through a fixed wireless infrastructure, such as a cellular network. For each query, the server is responsible for identifying the potential clients that could possibly belong to the query result. The potential client needs to periodically check its location to determine if it is actually in some query's region. If it moves into or exits from a query's region, it will notify the server and the server just updates the query result [7].

Casper is a novel framework in which mobile users can entertain location-based services without the need to disclose their private location information. Mobile users register with Casper by a user-specified privacy profile. Casper has two main components, the location anonymizer and the privacy-aware query processor [8].

In order to authenticate data owner's details, RC4 algorithm is user for encryption. This algorithm is used as it is very high execution time when compared to that of other algorithms.

III. PROPOSED SYSTEM:

A proposed system is being implemented to enhance the security of the system and to overcome the disabilities of the existing system. The data owner defines and provides information about the rightful owner of data assets. The certificate authority the ownership of the data assets which certifies consist of two processes, Certificate Verifier and the Certificate Issuer. The location based service provider is employed here which uses real-time geo-data from a mobile device or smartphone to provide information to the user. The user sends query to the service provider based on his interest.

The system consist of consists of a CA, data owner, LBSPs and system users. The solution for the problems is, to build a Certificate Authority (CA) to authenticate the outsourced data to a single trusted LBSP. Encrypted data sets are being collected and aggregated at the LBSP.

The data owner defines and provides information about the rightful owner of data assets. The certificate authority certifies the ownership of the data assets which consist of two processes, Certificate verifier which verifies the user's identity and Certificate Issuer which provides certificate to authenticate the data owner's details.

The data owner sends the data assets to the certificate authority. The certificate authority authenticates the details and sends the certificate to the data owner.

The location based service provider is employed here which uses real-time geo-data from a mobile device or smartphone to provide information to the user. After the authentication of data owner's details, the encrypted data assets along with the certificate are sent to the service provider.

The service provider collects and aggregates the data in its database. The user sends query to the service provider based on his interest. The service provider sends the data (encrypted + unencrypted) along with the certificate to the user. If the user wishes to view the sensitive data (encrypted) then he/she sends his/her identity to the certificate authority.

Based on the certificate and the identity of the user the certificate authority sends the encrypted data to the user. The user decrypts the data in order to view the sensitive field.

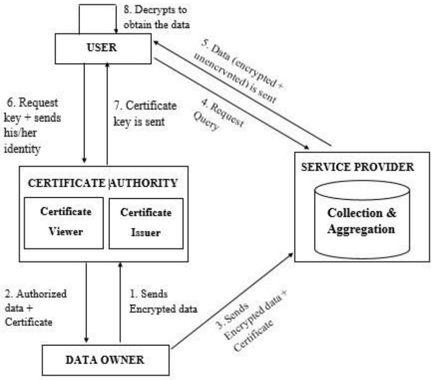
8. Decrypts to obtain the data
USER
5. Data (encrypted + unencrypted) is sent
6. Request key + sends his/her identity
7. Certificate key is sent
4. Request Query
SERVICE PROVIDER
Collection & Aggregation
CERTIFICATE AUTHORITY
Certificate Viewer
Certificate Issuer
2. Authorized data + Certificate
1. Sends Encrypted data
3. Sends Encrypted data + Certificate
DATA OWNER

Fig: Proposed architecture

IV. BACKGROUND – ENCRYPTION RC4

Algorithm is used to encrypt and decrypt the sensitive field in the data sets that is sent by the different data owners. RC4 can only be used one time to maintain its cryptographic security strength.

Let us now discuss how the data sets are encrypted and stored at the service providers. For example consider the following table,

It is only necessary to create an index for attributes involve in search by the user. The table consists of title, place, address and price. Here the price attribute alone is being encrypted to hide some sensitive information. The price attribute is encrypted using rc4 algorithm. The strengths of RC4 algorithm are, the difficulty of knowing where any value is in the table, the difficulty of knowing which location in the table is used to select each value in the sequence, a particular RC4 Algorithm key can be used only once and encryption is about 10 times faster than DES. The symmetric key algorithm is used identically for encryption and decryption such that the data stream is simply XORed with the generated key sequence. The algorithm is serial as it requires successive exchanges of state entries based on the key sequence. Hence implementations can be very computationally intensive.

The algorithm uses a variable length key from 1 to 256 bytes to initialize a 256- byte state table. The state table is used for subsequent generation of pseudo-random bytes and then to generate a pseudo-random stream which is XORed with the plaintext to give the cipher text. Each element in the state table is swapped at least once. The key is often limited to 40 bits, because of export restrictions but it is sometimes used as a 128 bit key. It has the capability of using keys between 1 and 2048 bits.

V. EXPERIMENTAL RESULTS:

The query processing over encrypted data is carried out using NetBeans V8.0.1. While encryption technique is carried out using internet based applications. On using number of individual LBSPs the outsourced data are scattered and has less memory space to store them. While using single LBSPs the data from the data owner is collected and aggregated after the authorization of the certificate authority for query processing. This results in efficient results. On comparing different encryption algorithms like AES, DES and RC4, the RC4 encryption algorithm is found to be more efficient [5]. The performance time is completely dependent on the processor speed and RAM.

Title	Place	Address	Price	Encrypted Price
A2b	Potheri	GST Road	500	vjhuFssdf
Papa Johns	Vandalur	GST Road	750	gdfgH5re
Pizza Hut	Kilpauk	Taylor’s Road	800	cBxy7e3r

File Size	RC4 (ms)	AES (ms)	DES (ms)	3DES (ms)
File 1	0.0787	0.0970	0.1464	0.1856

Fig: Comparative Execution time between RC4, AES, DES and DES The performance is analyzed using Intel(R) Core(TM) I5-4210U CPU @ 1.70GHz 1.70GHz with a memory of 500GB and a RAM of 6GB.

VI. CONCLUSION

This paper consists of an interesting distributed system for collaborative location based information generation and sharing. It enables the users to verify the authenticity and correctness of any query results. We have studied the problem of securely executing the range queries over outsourced data. We presented the RC4 encryption algortihm in order to provide enhanced security and to fine – tune the details of the data outsourcer. The service provider retains the responsibility of managing the data. The effectiveness of our proposed algorithms is validated by our experiments that show promising results on different datasets. Thus the proposed architecture ensures complete reliable and secure query processing. The efficacy and efficiency of our system are thoroughly analyzed and evaluated through detailed simulation studies.

REFERENCES:

[1] Rui Zhang, Jingchao Sun, Yanchao Zhang, and Chi Zhang, "Secure Spatial Top- k Query Processing via Untrusted Location- Based Service Providers", IEEE Transactions on Dependable and Secure Computing, Vol. 12, No. 1, January/

[2] H. Hacigumus, 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.

[3] B. Hore, S. Mehrotra, M. Canim, and M. Kantarcioglu, "Secure Multidimensional Range Queries over Outsourced Data," The VLDB J., vol. 21, no. 3, pp. 333-358, 2012.

[4] 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, May 2007.

[5] Hugo JONKER, 1, Sjouke MAUW and Jun PANG, "Location-Based Services: Privacy, Security and Assurance", Digital Enlightenment Yearbook 2012 J. Bus et al. (Eds.) IOS Press, 2012 © 2012.

[6] Prerana Deokar, Praveen Barapatre, "Location Based Query Processing For Providing Privacy to Exploit Service Similarity: A Survey", International Journal of Science and Research (IJSR), 2013.

[7] Fuyu Liu, "Query Processing in Location-Based Services", Thesis 2010.

[8] Mohamed F. Mokbel1 Chi-Yin Chow1 Walid G. Aref2, "The New Casper: Query Processing for Location Services without Compromising Privacy", VLDB '06, September 12-15, 2006, Seoul, Korea. Copyright 2006 VLDB Endowment, ACM 1-59593-385-9/06/09.