ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

A Survey on KNN Query Processing In Cloud and Anonymity with ABE

Kadam Sandip Parashram Kadam[1], Kanchan Doke[2]

M.E. Student, Dept. of Computer Engineering, Bharati Vidyapeeth College of Engineering, Mumbai University, Maharashtra, India[1]

Assistant Professor, Dept. of Computer Engineering, Bharati Vidyapeeth College of Engineering, Mumbai University, Maharashtra, India[2]

ABSTRACT: Users are interested in querying about points of interest (POI) in their physical proximity, such as cafes, ongoing events, etc. this data may be sensitive for data owner due to their contents. User sends their present location points and wants to know about nearest POIs in NN but data Owner or server does not have much storage capacity so we are using cloud service (server). Cloud provides power full storage at low cost but cloud provider may be not fully trusted. So we are using processing of NN queries in an untrusted outsourced environment ie (cloud), whereas at an equivalent time protective the POI and querying users' location positions. We use techniques based on mutable order preserving encoding (mOPE). It is a secure order-preserving encryption and updating database. User identity privacy in existing access control schemes AnonyControl decentralizes the central authority to limit the identity leakage..

KEYWORDS: User, Cloud and Location privacy, Database, Anonymity, Mutable Order Preserving Encoding, attribute-based encryption.

I. INTRODUCTION

Cloud satisfy requirement of data storage and data outsourcing for data owners. Thought cloud service provides such services but security and privacy of owner's data is major concern in cloud storage. Therefore secure data access is critical issue in cloud storage.

Due to mobile having fast Internet connectivity and also geo-positioning capabilities has led to a revolution in different customized location-based services (LBS). Using this location based services user can access information about points of interest (POI) that are relevant to their interests and that is close to their location. So probably the most important type of queries that involve location attributes is represented by nearest-neighbor (NN) queries, where a user wants to retrieve the k POIs (e.g., restaurants, museums, gas stations) that are nearest to the user's current location (kNN).

Another technique which solves the problem of single point failure is Threshold multi-authority CP-ABE access control that is TMACS [1]. In this technique multiple authorities jointly manage the whole attribute set but no one has full control over any specific attribute. In this technique secret sharing key is used among different authorities with (t, n) threshold secret sharing. In TMACS secrete key is known as a Master Key which cannot be obtained by any single authority alone. Personal information (defined by each user's attributes set) is at risk because one's identity is authenticated based on his information for the purpose of access control, As people are becoming more concerned about their identity privacy these days, the identity privacy also needs to be protected before the cloud enters our life. Preferably, any authority or server alone should not know any client's personal information.

II. LITERATURE SURVEY

Protecting location data is an important problem not only in the scenario of outsourced search services, but in a variety of other settings as well, For instance, two approaches for location protection have been investigated in the context of private queries to location-based services (LBS). The objective here is to allow a querying user to retrieve

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

her nearest neighbor among a set of public points of interest without revealing her location to the LBS. The first approach is to use cloaking regions [16-19]. To Most CR based solutions implement the spatial k-anonymity paradigm and assume a three-tier architecture where a trusted anonymizer sits between users and the LBS server and generates rectangular regions that contain at least k user locations. This approach is fast, but not secure in the case of outliers. The second approach uses private information retrieval (PIR) protocols [7, 9]. PIR protocols allow users to retrieve an object from a set $X=\{X, X, \ldots, X\}$ stored by a server, without the server learning the value of i. The work in [7, 9] extends an existing PIR protocol for binary data to the LBS domain and proposes approximate and exact nearest neighbor protocols. The latter approach is provably secure, but it is expensive in terms of computational overhead. The work in [2] uses a secret matrix transformation to hide the data points. The data owner generates randomly a matrix M, and then transforms data points by multiplying them with M. Users transform their query points using multiplication with the inverse matrix M-1, When the server receives the transformed data points from the data owner and a transformed query point from a user, it can determine which data point is nearest to the query point. In contrast with [10], the exact results are returned to the user. However, the matrix transformation is vulnerable to chosen plaintext attacks, as shown in [5].

Similar to [2], the work in [4] also uses a matrix transformation to protect both data and query privacy. Hence, it is also vulnerable to chosen plaintext attacks. In addition, given a kNN query, the server returns more than k data items to the client, and the client must filter out unnecessary data. This additional disclosure is undesirable, as the client who pays for k results should not be allowed access to more data points. In [3], the data owner sends a shadow index to the client. The shadow index is encrypted by the data owner, and the decryption key is given to the server. However, the method requires the entire encrypted index to be transferred to the client. When there are a lot of data points, the size of the index grows large as well, making the method impractical.

Finally, the work in [5] shows how to stage effective attacks against methods such as [2, 3], and that solving the secure nearest neighbor problem is at least as hard as Order Preserving Encryption (OPE) [20]. The proposed method from [5] returns a relevant partition E(G) from the entire encrypted dataset, and E(G) is guaranteed to contain the answer for the NN query. However, the technique from [5] returns significantly more than k data items to the client.

III. METHODOLOGIES

1. System Model

The system model comprises of three distinct entities:

1.1 The data owner.

1.2 The outsourced cloud; service provider (for short cloud server, or simply server).

1.3 The client (Data consumer).

Following diagram shows system model:

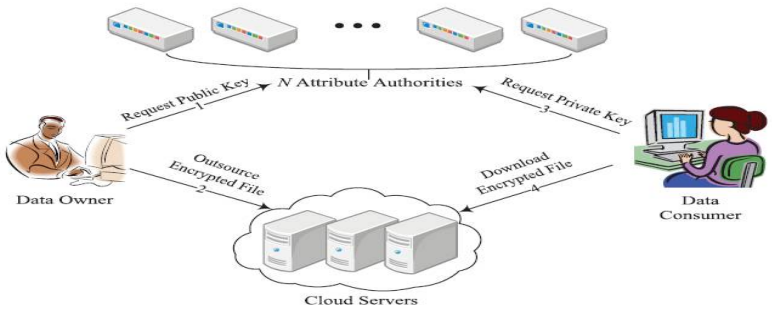

Figure.1. General Flow of system

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Figure. 1.1 System model

CP-ABE access control scheme gives more access control to data owner, where there are multiple authorities co-exist and each authority is able to manage attributes independently. This scheme provides forward and backward security for access control [3].

1.1.The data owner:

The data owner has a dataset with n two-dimensional points of interest, but does not have the necessary infrastructure to run and maintain a system for processing nearest-neighbor queries from a large number of users.

As the dataset of points of interest is a valuable resource to the data owner, the storage and querying must be done in encrypted form.

1.2.Cloud server:

The server receives the dataset of points of interest from the data owner in encrypted format, together with some additional encrypted data structures (e.g., Voronoi diagrams, Delaunay triangulations) needed for query processing. The server receives kNN requests from the clients, processes them and returns the results.

1.3.Client:

The client has a query point Q and wishes to find the point's nearest neighbors. The client sends its encrypted location query to the server, and receives k nearest neighbors as a result.

2 .Query Processing Method:

2.1.Order-preserving encryption:

mOPE allows secure evaluation of range queries, and is the only provably secure order-preserving encoding system (OPES) known to date. The difference between mOPE and previous OPES techniques (e.g., Boldyreva et. al. [11,12]) is that it allows cipher texts to change value over time, hence the mutable attribute. Without mutability, it is shown in [6] that secure OPES is not possible.

Since our methods use both mOPE and conventional symmetric encryption (AES), to avoid confusion we will further refer to mOPE operations on plaintext/cipher texts as encoding and decoding, whereas AES operations are denoted as encryption/decryption.

The mOPE scheme in a client-server setting works as follows: the client has the secret key of a symmetric cryptographic scheme, e.g., AES, and wants to store the dataset of ciphertexts at the server in increasing order of corresponding plaintexts. The client engages with the server in a protocol that builds a B-tree at the server. The server only sees the AES cipher texts, but is guided by the client in building the tree structure. The algorithm starts with the client storing the first value, which becomes the tree root. Every new value stored at the server is accompanied by an insertion in the B-tree. Figure 2.1 shows an example where plaintext values are also illustrated for clarity, although they are not known to the server (for simplicity we show a binary tree in the example).

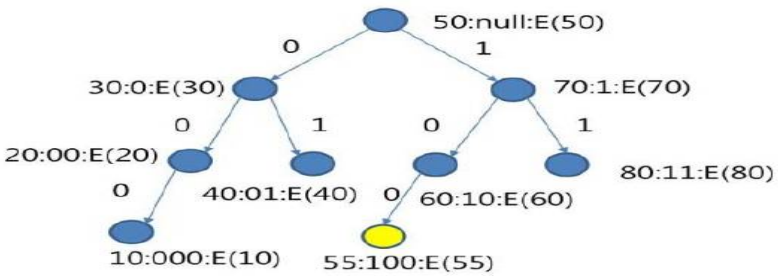

Figure 2.1-mOPE Tree: Inserting node E(55)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

Assume the client wants to store an element with value 55: it first requests the ciphertext of the root node from the server, then decrypts E(50) and learns that the new value 55 should be inserted in the tree to the right hand side of the root. Next, the client requests the right node of the root node and the server sends E(70) to the client. The process repeats recursively until a leaf node is reached, and 55 is inserted in the appropriate position in the sorted B-tree, as the left child of node 60. The client sends the AES cipher ext E(55) to the server which stores it in the tree. The encoding of value 55 in the tree is given by the path followed from the root to that node, where 0 signifies following the left child, and 1 the right child. In addition, the encoding of every value is padded to the same length (in practice 32 or 64 bits) as follows [6]:

The server maintains a mOPE table with the mapping from cipher texts to encodings, as illustrated in Figure 2.2 for a tree with four levels (four-bit encoding). Clearly, mOPE is an order preserving encoding, and it can be used to answer securely range queries without need to decrypt cipher texts. In addition, the mOPE tree is a balanced structure. Using a B-tree, it is possible to keep the height of the tree low, and thus all search operations are efficient. In order to ensure the balanced property, when insertions are performed, it may be necessary to change the encoding of certain cipher texts.

Data Encryption and Upload: The owner encrypts the data based on the sub ACPs in ACPBowner with unique symmetric key, called an ILE keyand uploads it with corresponding public information. The cloud now encrypts the data again gained from owner based on ACPs in ACPBcloud with unique symmetric key, called an OLE key .In this case instead of sharing secret key, users are given secret which combine with public information to obtain actual private keys.

Ciphertext	mOPE Encoding
E(50)	[]1000 = 8
E(30)	[0]100 = 4
E(70)	[1]100 = 12
E(20)	[00]10 = 2
E(40)	[01]10 = 6
E(60)	[10]10 = 10
E(80)	[11]10 = 14
E(10)	[000]1 = 1
E(55)	[100]1 = 9

Figure 2.2 .mOPE Table

3. Voronoi Diagram-based 1NN (VD-1NN):

In this section, we focus on securely finding the 1NN of a query point. We employ Voronoi diagrams [1], which are data structures especially designed to support NN queries. An example of Voronoi diagram is shown in Figure 3. Denote the Euclidean distance between two points p and q by $d(p,q)$, and let $P = \{p_1, p_2, \ldots, p_n\}$ be a set of n distinct points in the plane. The Voronoi diagram (or tesselation) of P is defined as the subdivision of the plane into n convex polygonal regions (called cells) such that a point q lies in the cell corresponding to a point p_i if and only if p_i is the 1NN of q, i.e., for any other point p_j it holds that $dist(q,p_i) < dist(q,p_j)$ [1].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

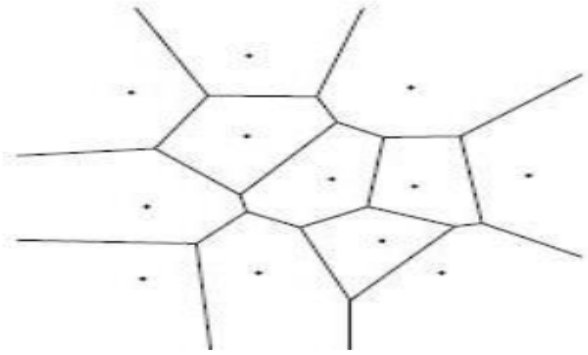

Figure3. Voronoi Diagram

Answering a 1NN query boils down to checking which Voronoi cell contains the query point. In our system model, both the data points and the query must be encrypted. Therefore, we need to check the enclosure of a point within a Voronoi cell securely. Next, we propose such a secure enclosure evaluation scheme.

IV. CONCLUSION

We proposed two methods to support secure kNN query processing: VD-kNN which is based on Voronoi diagrams, TkNN which relies on Delaunay triangulations. These both are uses mutable order preserving encoding (mOPE) for building block.VD-kNN fetch exact results, but this methods performance overhead may be high.TkNN only offers approximate NN results, but with better performance.Anonymity withAttribute-based encryption effectively for users identity (Location) and secure content (POI) in the cloud

REFERENCES

1. Sunoh Choi, Gabriel Ghinita, Hyo-Sang Lim and Elisa Bertino,'Secure kNN Query Processing in Untrusted Cloud Environments',ieee transactions on knowledge dta engineering, Vol.26, pp. 4-11june, 2014.
2. Taeho jung, xiang-yang li, senior member, ieee, zhiguo wan, and meng wan,control cloud data access privilege and and anonymity with fully anonyous attribute-based encryption',ieee transactions on information forensics and security, vol. 10, no. 1,pp.2-8, january 2015.
3. Haibo Hu, Jianliang Xu, Chushi Ren, Byron Choi,'Processing Private Queries over Untrusted Data Cloud through Privacy Homomorphism',pp.4-9,2005.
4. Gabriel Ghinita, Panos Kalnis, Murat Kantarcioglu, and Elisa Bertino, 'Approximate and exact hybrid algorithms for private
5. nearest-neighbor queries with database protection',pp.1-5 dec 2015.
6. W. K. Wong, David W. Cheung, Ben Kao, and Nikos Mamoulis,'Secure kNN Computation on Encrypted Databases',June 29–July 2009.
7. Vladimir Kolesnikov and Abdullatif Shikfa ,'On The Limits of Privacy Provided byorder- Preserving Encryption ,'Bell Labs Technical Journal 17(3),pp, 135–146 ,2012.
8. Alexandra Boldyreva, Nathan Chenette, Younho Lee and Adam O'Neill,'Order-Preserving Symmetric Encryption',2009.
9. Adi Shamir,'IDENTITY-BASED CRYPTOSYSTEMS AND SIGNATURE SCHEMES', pp. 47-53, 1985.
10. Amit Sahai and Brent Waters,'Fuzzy Identity-Based Encryption',pp. 457–473, 2005.
11. Kan Yang and Bo Zhang,' Effective Data Access Control for Multiauthority Cloud Storage Systems',2013.

BIOGRAPHY

Sandip Parashram Kadam is Student of M.E. Computer Engineering, Bharati Vidyapeeth College of Engineering, Mumbai University, Navi Mumbai, Maharashtra, India. He is an software Engineer in I.T company, Mumbai, Thane, Maharashtra, India. His area of interest is Cloud computing and Programming languages.

Kanchan Doke is a Professorin the Department of I.T. and computer science, Bharati Vidyapeeth College of Engineering, Mumbai University, Navi Mumbai, Maharashtra, India. Her area of interest is cloud computing, algorithms and programming.