International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 08 | Aug 2019 www.irjet.net p-ISSN: 2395-0072

Efficient Geometric Range Search on RTREE Occupying Encrypted Spatial Data

Maria Mathews

Maria Mathews, Govt. Engineering College, Idukki

Abstract — Over past decades, necessity for cloud had in- creased drastically. Cloud computing lends a cinch move to access servers, storage, databases and a broad set of application services over the Internet.Need for storage capacity increases with increase in amount of data. Cloud environment is in- troduced to increase this storage space. Storing confidential informations like patient record, employee record etc in cloud need secure methods as cloud data is insecure. RTREE is implemented to provide cache for such data. Searching those informations from RTREE obligate geometrical coordinates (latitude, longitude) as it stocks spatial data. To cater more security for stored data in RTREE, trusted third party is selected for uploading the encrypted data (using DES) to RTREE by obscuring data owners identity from search user. By the end, the confidentiality of data, data owner's privacy is theoretically turn out.

Keywords: Cloud computing, RTREE, Spatial Data, Geo- metric range search, Trusted Third Party

I. INTRODUCTION

The cloud is customarily adopted to imitate the internet. Cloud computing is a domain used to enable the carting of software, infrastructure and storage services over the network mainly internet. Users of the cloud can perk from other organizations posting services consorted with their data, software and other computing needs on their behalf, without the need to own or run the usual physical hardware and software themselves. Generally, cloud is of 3 sorts, Private cloud cater Services that are owned on-site by you and your company, with your data behind your organizations own firewall. Public cloud provide services that may be staked with other organizations, with data security lend by the cloud dealer. Hybrid cloud contribute for a single organization delivered over a combination of private and public cloud. Accord security for data in cloud is an important concern as it is easily retrieved by a malicious user.

A. Cloud Computing

Cloud computing accomplish system resources, especially storage and computing power, available on demand without direct active management by the user. Cloud computing is the next stage in the evolution of the internet, it provides the means through which everything from computing power to computing infrastructure, applications and business pro- cesses can be delivered to you as a service wherever and whenever you need them. Cloud model envisages a world where components can be rapidly orchestrated, provisioned, implemented and decommissioned, and scaled up or down to provide an on-demand utility-like model of allocation and consumption. From an architectural perspective, there is much confusion surrounding how cloud is both similar to and different from existing models of computing and how these similarities and differences impact the organizational, operational, and technological approaches to network and information security practices. There is a thin line between conventional computing and cloud computing. However, cloud computing will impact the organizational, operational and technological approaches to data security, network se- curity, and information security good practice. There are many definitions today that attempt to address cloud from the perspective of academicians, architects, engineers, de- velopers, managers, and consumers. This document focuses on a definition that is specifically tailored to the unique perspectives of IT network and security professionals.

II. RELATED WORK

Security for cloud data is becoming an important concern in the many fields. Research studies are also conducted in the field of cloud In this section, we will open the research and analyze the advantages and weakness of them. In the paper by Hongwei Li[1] spatial data is stored in R-tree. Attacker cannot find the data in leaf node. Data owner will draft the data and generate the secret key. Data encryption is done by using DES (Data Encryption Standard). Data owner upload data to cloud with index for each data record in R-tree. When search user wants data,he/she give request to data owner. Find the intersection point by traversing through R-tree to gain

encrypted data and Extract the leaf node containing the encrypted data and send to the user. Finally, user decrypt the data.

Another paper by Hongwei Li[2]propose a method for circular range search on encrypted spatial data. Helps to find points within a circular range. Initially, Circle is divided in to 2 outer and inner rectangle and again rectangle is further divided whenever necessary. Data points within circle is identified.

In the paper by Boyang Wang[3], he proposes a method which converts the Coordinates of geometrical range in to quality vector form to avoid large data decryption. Data will be encrypted and decrypted using AES algorithm. Data will be stored as linked list. Firstly search for each x-sub query in the linked list. Once nd a match in terms of x, we continue to evaluate an inner product of its y-sub query with a node in a corresponding link list.

Rather than using DBMS for storing data, large number of data can be stored while usin RDBMS. This was proposed by Venus Bron Lima[4]. Data points are stored as tuples. Geometric range search can be performed with this tuples. Data is encrypted using AES and will be stored with in the table. This data can be retrieved by the user by searching through this table.

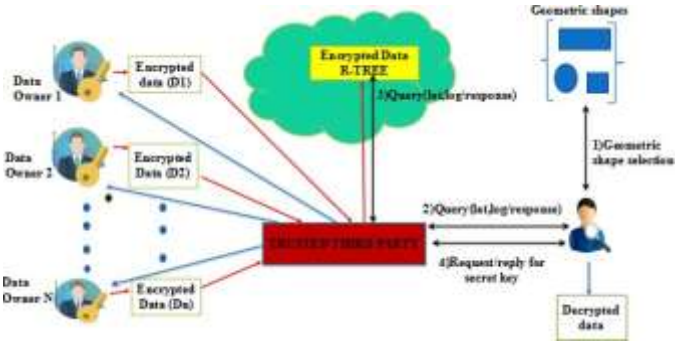

Fig. 1. System Design

III. PROPOSED WORK

A. Basic Design

Here we pick up data owner, cloud service provider, search user to frame the system.

1) Data Owner:- Data owner initially sync their data to cache in the cloud by announcing RTREE. After loading their data, encrypt data using DES encryption and upload data to RTREE in cloud.

2) Cloud Service Provider:- CSP gather data from data owner and backlog data in RTREE.

3) Search User:- Whenever search user need a data in a geometrical range, he or she requests range query to csp. Search user gets response in the form of encrypted data. To decrypt the data, search user requests for key from data owner. If, data owner is willing to provide it, search user decrypt data using same DES algorithm.

B. System Design

Cloud data is always insecure and may be destroyed by malicious attack. Cloud generally consist of CSP, Data owner and Search user. Cloud service provide and search user itself may be malicious. Data uploaded to cloud need security measures to improve integrity and confidentiality. Here, data owner upload his spatial data to cloud by introducing RTREE in cloud. To give more security for data in RTREE, Trusted Third Party is added to the system design. Trusted third party improves the trust level by affording secure communication between two end parties.

C. Encryption and Decryption

Data owner use DES [5][7][8] algorithm to encrypt the data. Data Encryption Standard is a symmetric-key algorithm for the encryption and decryption. DES algorithm generally consist of 4 steps,

1) Data Expansion: The 32-bit half-block is expanded to 48 bits using the expansion permutation, denoted E in the dia- gram, by duplicating half of the bits. The output consists of eight 6-bit pieces, each containing a copy of 4 corresponding input bits, plus a copy of the immediately adjacent bit from each of the input pieces to either side.

2) Key mixing: The result is combined with a sub key using an XOR operation. Sixteen 48-bit sub keysone for each roundare derived from the main key using the key schedule

3) Substitution: After mixing in the sub key, the block is divided into eight 6-bit pieces before processing by the S-boxes, or substitution boxes. Each of the eight S-boxes replaces its six input bits with four output bits according to a non-linear transformation, provided in the form of a look up table. The S-boxes provide the core of the security of DES without them, the cipher would be linear, and trivially breakable.

4) Permutation: finally, the 32 outputs from the S-boxes are rearranged according to a fixed permutation, the P-box. This is designed so that, after permutation, the bits

from the output of each S-box in this round are spread across four different S-boxes in the next round.

To decrypt the message, algorithm works similar to encryp- tion where ciphertext is provide to above steps for process- ing, Output will be the plain text.

D. Trusted Third Party

Trusted Third Party[5][9] is established with in a system in order to afford a web of trust. Sometimes, attacker may modify the owner's data by malicious move. Data owner and cloud service provider will be unaware of it. Hence we add a malicious modification detection mechanism in our scheme. Message-Digest Algorithm 5 (MD5)[6] is a popular algorithm often used in consistency check. TTP will be authenticated by using the ID generated from MD5. Takes as input a message of arbitrary length and produces as output a 128 bit fingerprint or message digest of the input. It consists of 5 steps,

- Step 1 - Append padded bits: The message is padded so that its length is congruent to 448, modulo 512. A single 1 bit is appended to the message, and then 0 bits are appended so that the length in bits equals 448 modulo 512.
- Step 2 - Append length: A 64 bit representation of b is appended to the result of the previous step. The resulting message has a length that is an exact multiple of 512 bits.
- Step 3 - Initialize MD Buffer. A four-word buffer (A,B,C,D) is used to compute the message digest. Here each of A,B,C,D, is a 32 bit register. These registers are initialized to the following values in hexadecimal: word A: 01 23 45 67 word B: 89 ab cd ef word C: fe dc ba 98 word D:76 54 32 10
- Step 4 - Process message in 16-word blocks. Four auxiliary functions that take as input three 32-bit words and produce as output one 32- bit word. F(X,Y,Z) = XY v not(X) Z G(X,Y,Z) = XZ v Y not(Z) H(X,Y,Z) = X xor Y xor Z I(X,Y,Z) = Y xor (X v not(Z))
- Step 5 - output the message digest produced as output is A, B, C, D. That is, output begins with the low-order byte of A, and end with the high-order byte of D.

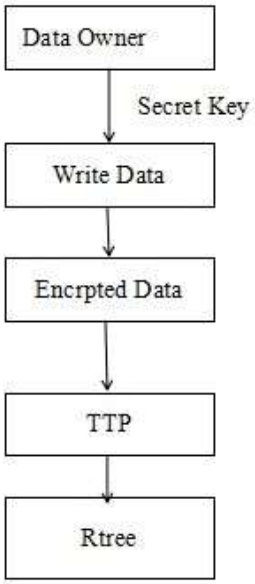

Fig. 2. Data Uploading

E. RSA Algorithm

RSA algorithm is used to to provide security for key. Encryption and decryption are carried out using two different keys. The two keys in such a key pair are referred to as the public key and the private key.

- Choose two distinct prime numbers.
- Calculate n=p*q. For strong unbreakable encryption, let n be a large number, typically a minimum of 512 bits.
- Number e must be greater than 1 and less than (p 1)(q 1).
- The pair of numbers (n, e) form the RSA public key and is made public.
- Interestingly, though n is part of the public key, diffi- culty in factorizing a large prime number ensures that attacker cannot find in finite time the two primes (p q) used to obtain n. This is strength of RSA.
- Private Key d is calculated from p, q, and e. For given n and e, there is unique number d.
- Number d is the inverse of e modulo (p - 1)(q 1). This means that d is the number less than (p - 1)(q - 1) such that when multiplied by e, it is equal to 1 modulo (p - 1)(q - 1).

F. Data Uploading

If a data owner wants to upload his/her data, first load the data. After loading the data, data is encrypted using DES algorithm. After encryption, data is passed to trusted third party. TTP upload the data to RTREE(fig 2) in cloud. To authenticate TTP, MD5 cryptographic hash function is used.

G. Data Retrieval

For retrieving a confidential data, search user first provide geometric range request query with latitude and longitude

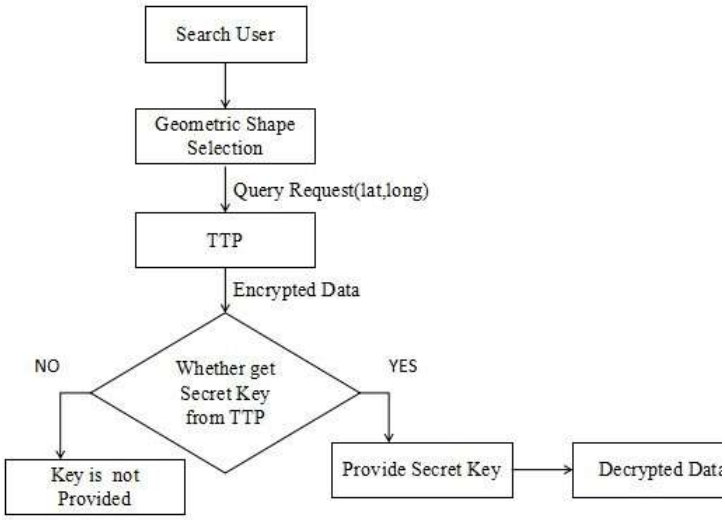

Fig. 3. Data Retrieval

Shape	Starting latitude	Starting longitude	Ending latitude	Ending longitude	Total data	True positive	Hit Rate	True negative	accuracy
Rectangle 1	51.564004	-0.243982	51.504602 01	-0.14752233	5	5	5	0	(5/(5+0))* 100= 100%
Rectangle 2	51.574347	-0.255393	51.522138 23	-0.20060462	2	2	2	0	(2/(2+0))* 100= 100%
Rectangle 3	51.574347	-0.255393	51.504602 01	-0.14752233	7	7	7	0	(7/(7+0))* 100= 100%
Rectangle 4	52.875295	-0.133476	51.504602 01	0.14752233	11	11	11	0	(11/(11+0))*100= 100%

Fig. 4. Rectangle Hit

parameters. Those values are compared with latitude and longitude value in the index. If matches, search continues in the RTREE[10] to retrieve the encrypted data. To decrypt the data, search user ask for key from data owner. But request is forwarded to the Trusted Third Party. Trusted third party forward the request to the data owner if he/she is authentic.

IV. EVALUATION PARAMETERS

A. Rectangle Hit Rate

Data points within rectangular range can be retrieved by range search. Table below shows the result for rectangular range search.

B. Circle Hit Rate

Data points within Circular range can be retrieved by range search. Table below shows the result for Circular range search.

shape	Starting latitude	Starting longitude	distance	Total data	True positive	True negative	accuracy
Circle 1	51.564004	-0.243982	10km	5	5	0	(5/(5+0))*10 0=100%
Circle 2	51.574347	-0.255393	6km	2	2	0	(2/(2+0))*10 0=100%
Circle 3	51.574347	-0.255393	100km	7	7	0	(7/(7+0))*10 0=100%
Circle 4	52.875295	-0.133476	400km	11	11	0	(11/(11+0))* 100=100%

Fig. 5. Circle Hit

C. Accuracy

Proposed system for both rectangular and circular range search within rtree is 100 percent accurate. To find the accu- racy, Accuracy=(TP/(TP+TN))*100, where Tp is the no of positive value retrieved correctly and Tn is the no ofnegative value retrieved.

V. CONCLUSION

Cloud computing is a scalable and distributive architecture. But it faces data loss and modifications for confidential data also. So, we need some security measures to provide more security for such data. Embedding trusted third party within the design ensure data security by hiding data owner identity. Whenever search user need a geometrically ranged data, he/she provide a request to cloud. Request to cloud is directly passed to TTP and verify the authentic user and forward request to data owner. Thereby data owner identity is concealed from search user. Without knowing the data owner identity, search user as an attacker cannot find the source of data. Hence, security is provided up to a great extend.

REFERENCES

[1] HongweiLi, Enabling Efficient andGeometricRangeQuery with Access Control overEncrypted Spatial Data, 2018.

[2] Hao ren, Hongwei li, Efficient privacy-preserving circular range search on outsourced spatial data, 2016.

[3] Boyang wang, Mingi li, Fastgeo:Efficient Geometric range queries on encrypted spatial data, 2018.

[4] Venus bron lima, Geometric location finder based on encrypted spatial data using geometric range queries, 2018.

[5] Susmita JA Nair, Anitha K.L, Rosita F Kamala,Trusted Third Party Authentication in Cloud Computing, 2013.

[6] Adviti Chauhan, Jyoti Gupta, A novel technique of cloud security based on hybrid encryption by Blowfish and MD5, 2017.

[7] Jian Zhang, Xuling Jin, Encryption System Design Based on DES and SHA-1, 2012.

[8] Akash Kumar Mandal, Chandra Parakash, Archana Tiwari, Perfor- mance evaluation of cryptographic algorithms: DES and AES, 2012.

[9] D.K.Aarthy, M.Aarthi, K.Afritha Farhath, S. Lakshana ; V. Lavanya, Reputation-based trust management in cloud using a trusted third party, 2017.

[10] Guowen Xu, Hongwei Li, Yuanshun Dai, Jian Bai, Xiaodong Lin, EFRS:Enabling Efficient and Fine-Grained Range Search on En- crypted Spatial Data, 2018.