International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

PRACTICAL QUERY ACCOMMODATIONS TO SECURE DATA FOR STORING IN CLOUD

[1]Ms. Mustyala Roopa, [2]Mr. S. Sudheer Reddy,
[1,2] Assistant Professor,Dept. of CSE,
Malla Reddy Engineering College (Autonomous), Secunderabad, Telangana State

Abstract

Nowadays incrementing users of public cloudcomputing infrastructures, and utilizing clouds to store data with queryaccommodations are a solution which gives more scalability and cost-preserving.Hence, most of the data is sensitive that the data owner does notwant to move their data to the cloud without utilizer get the dataconfidentiality and query privacy are ensured. On the otherhand, a secured query accommodations should still provide efficientquery processing and significantly reduce the in-house workloadto plenarily realize the benefits of cloud computing. Query accommodationsreduce the overhead of querying. RASP and kNN query accommodationsgives secured data afore storing on cloud with the avail of order-preserving encryption and range query.

Keywords: Target distribution, LBS, bucketization, query privacy, data confidentiality.

1. Introduction

Query accommodations in the cloud are increasingly popular because of the unique advantages in quality and costsaving. With the cloud infrastructures, the accommodation owners can conveniently scale up or down the accommodation and only pay for the hours of utilizing the servers. This is a captivating feature because the workloads of query accommodations are highly dynamic, and it will be extravagant and inefficient to accommodate such dynamical workloads with in-house substructures. However, because the accommodation providers lose the control over the information in the cloud, data privacy and query privacy have become the major concerns. Summarization of these requisites for constructing a practical query accommodation in the cloud as the CPEL criteria: data privacy, query Privacy, effective query working, and Low in-house working cost. The rudimentary conception is to arbitrarily transform the multidimensional datasets with a merger of order preserving encryption, dimensionality expansion, desultory noise injection, and arbitrary project, so that the utility for processing range queries is preserved. Driven by lower cost, higher reliability, better performance, and more expeditious deployment, data and computing accommodations have been increasingly outsourced to clouds such as Amazon EC2 and S3, Microsoft Azure, and Google App Engine. However, privacy has been the key road auction block to cloud computing. On one hand, to leverage the computing and storage capability offered by clouds. While storing data on cloud users have to face the quandary of some delay in the retrieving data from cloud storage. Data privacy and efficiency utilizing file retrieval from Ostrovosky.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

In this scheme utilizer can retrieves files from an untrusted server. Data Perturbation is to balance privacy aegis and data utility. The kNN-R algorithm is designed to function with the RASP range query algorithm to work the kNN queries.

2. Related Work

2.1 EXISTING SYSTEM:

Actually we require the query processing in the cloud to satisfy privacy and efficiency at Low processing cost. But considerably processing complications will be raised. Recently few mechanics were proposed to satisfy these requirements. But they have not reached the actual needs. The crypto index and Order-Preserving-Encryption are weak in prevention of attacks, increases the computational complexity to improve the privacy. The Casper-method uses cloaking boxes to secure information and queries, this reduces query efficiency.

DISADVANTAGES OF EXISTING SYSTEM:

- Malicious users may attack on the data or queries.
- Weak in prevention of attacks.
- Increases the computational complexity to improve the privacy

2.2 PROPOSED SYSTEM:

We propose a secured and efficient Query processing Service in Cloud with Arbitrary Space Decomposition technique for range queries and nearest neighbour queries. The proposed technique satisfies the query processing in the cloud with privacy and efficiency at Low processing cost, without increasing processing complexitiesSeed Block Algorithm supports protection on well-organized range-queries, nearest neighbour search queries. The fundamental inspiration is to arbitrarily change the multi-dimensional information, order preserving encryption, volume extension, adding noise, and protect queries processing. The main mechanisms used are the discomposure, creating secure range queries and creating secured nearest neighbour queries.

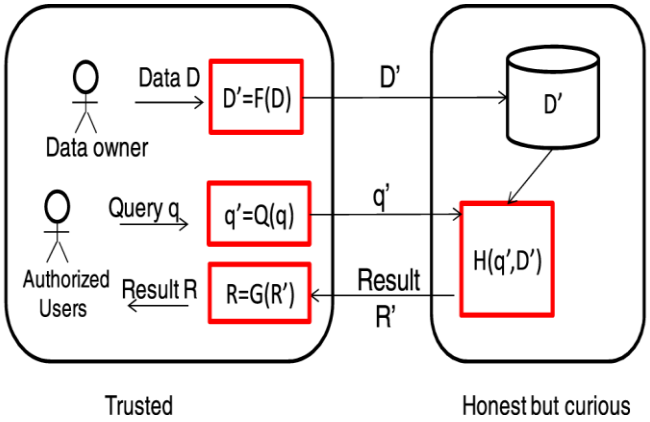

Fig 1: System Architecture for RASP method.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

The untrusted parties consist of the curious cloud provider who hosts the query accommodations and the bulwarked database. The RASP-perturbed data will be acclimated to build indices to maintain query processing.

Advantages of Proposed System:

- The Discomposure integrates order conserving encryption, volume extension, and arbitrary noise additions.
- Provides strong confidentiality guarantee.
- computational complexity is less
- High precision on results

2.3 Algorithm Definition:

SEED BLOCK ALGORITHM

Simple but power full cryptography algorithm called the Seed Block Algorithm is proposed to reduce the computational overhead. Any transaction between the data owner, user, data Server is encrypted using the Seed Block algorithm.

SBA ALGORITHM

Algorithm 1:

Initialization	:	Main Cloud: Mc; Remote Server: Rs;
		Customers of Main Cloud: Ci; Files: a1 and a1';
Seed block	:	Si;Random Number: r';
Customer's ID	:	Client_Idi
Input	:	a1 created by ci; is generated at; Mc
Output	:	Recovered file a1 after deletion at ci
Given	:	Authenticated clients could allow uploading,
		Downloading and do modification on its own the files only.
Step 1	:	Generate a random number.
		Int r=rand ( );
		Int r=rand ( );
Step 2	:	Create a seed Block Si for eachCi and Store
		Si at Rs
		Si=r xorClient_Idi / / (Repeat step 2 for all clients)
Step 3	:	If Ci /admin creates/modifies an a1and stores at
		Mc, then create as
		a1'= a1 xor Si
Step 4	:	Store a 'at Rs.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Step 5 : If server crashes a1 deleted from Mc,

Then, we do EXOR to retrieve the original a1 as:A1=a1' xor Si

Step 6 : Return a1 to Ci.

Step 7 : END.

In cloud computing, data is generated in that form of an electronic are large in amount. In cloud we have to maintain any data confidentiality or efficiency necessary of a data recovery services. Here seed block algorithm can also be mainly used to develop a smart remote data backup algorithm; this algorithm is also called as a seed block algorithm. The objective of seed block algorithm is double in the first place is clients to gather any data from any remote area without system integration or disappointments just and second to recuperate the documents in the event of the record erasure or if the cloud gets crushed because of any reason The SBA is mainly focused on the storing of the backup files in the remote server only, without using the existing encryption techniques.

Advantages of this algorithm is

- Recover same size data
- Low host
- Privacy
- Data security
- Remote Data Backup Server

Here data security can Giving full security to the client's data and this data also present in the utmost priority for the remote backup server either intentionally or unintentionally, here third party user or any un authorized users cannot access the data in the remote backup server only. The main advantages of remote backup server are we can access any data from remote backup server, this server is not connected to the network connectivity also, we can get the data from the remote backup server only.

The work flow of the above algorithm is:

Here we have to maintain one of the remote back-up server this is mainly used to storing of the user data, when the user login in to the cloud it can automatically generates the random id Int r=**rand ( );**and user id also, based on this user id and random id it has automatically generated by the seed block key

Create a seed Block **Si** for each**Ci** and Store

Si at **Rs**

Si=r xorClient_Idi

After the completion of this thing whenever user upload the data in to the remote back-up server with the help of user generated seed block key only the uploaded can be decrypted. For assumption we have to hack a remote back-up server that is not user understandable, here index format also encrypted using that of the md5 algorithm.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

3. Implementation

Three modules are utilized. They are RASP, range query and kNN query.

3.1 RASP:

RASP denotes Desultory Space Perturbation. It withal amalgamates OPE, desultory projection and desultory noise injection. Here OPE denotes Order Preserving Encryption is utilized for data that sanctions any comparison. And that comparison will be applied for the encrypted information; this will be done without decryption. Desultory projection is mainly used to process the high dimensional information into low dimensional information representations. It contains features like good scaling potential and good performances. Arbitrary noise injection is mainly used to integrating noise to the input to get felicitous output when we compare it to the estimated potency. The RASP method and its coalescence provide confidentiality of data and this approach is mainly used to bulwark the multidimensional range of queries in secure manner and additionally with indexing plus efficient query working will be done. RASP has some consequential features.

In RASP the utilization of matrix multiplication does not bulwark the dimensional values so no desideratum to suffer from the distribution predicated attack. RASP obviates the data that are perturbed from distance predicated attacks; it does not bulwark the distances that are occurred between the records. And additionally it won't forfend more arduous structures it may be a matrix and other components. The range queries can be send to the RASP perturbed data and this range query describes open bounds in the multidimensional space.

In arbitrary space perturbation, the word perturbation is utilized to do collapsing this process will transpire according to the key value that is given by the owner. In this module the data owner have to register as owner and have to afford owner name plus key value. And then the utilizer have register and get the key value and information owner name from the owner to do get at in the cloud. Here utilizer can submit their query as range query or kNN query and get their response. We analyze and show the result with encrypted and additionally in decrypted format of the data for the query construct by the utilizer.

3.2 RANGE QUERY

Range query is the query used to retrieve the information from the database. It will retrieve the information value that is between the upper bound and lower bound. The range query is not customary because utilizer won't ken in advance about the result for the query, how much ingressions will come as answer for the query. For example

```
SELECT id
FROM table name
WHERE id (
SELECT top 10*
FROM United States
```

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

WHERE age >50

);

The above example expresses the sample query for range query. Here the example query is to retrieve the ingressions from Amalgamated States it will retrieve the persons who are above 50 years in the top 10 list from the record of Amalgamated States. The range search is mainly used to return the values that are present between the two designated values given in the query. For example database designation is AAAworkers2012 then

Go

SELECT product id

FROM AAAworkers2012.production

WHERE price BETWEEN 40 and 60

The above example will show another example of range query probe it will provide the ingressions of what are product id that are present in engenderment database with price above 40 and within 60. So by utilizing range query utilizer can facilely retrieve the data's from records and this query treat will be done in secure mode plus the accelerate of the query process will withal increase.

3.3 kNN QUERY

kNN query represents k-Most proximate Neighbor query. This query is mainly used to retrieve the most proximate neighbor values of k. here k used to denote positive integer value. kNN algorithm is mainly utilized for relegation and regression. In this it utilizes kNN-R algorithm to process the range query to kNN query. This algorithm consists of two methods. That is utilized to make interaction between the client and the server. The client will send the query to the server with initial upper bound and lower bound. This upper bound range has to be more than the k points and the lower bound range have to be less than the k points.

The above process is utilized to give the inner range of the database by the server. With that inner range the client will estimate the outer range and send this outer range to the server. Then the server will probe and find the records in the outer range from the database and send it to client and then the client will decrypt the record and detect the top k files to provide the final result. This algorithm is utilized to find the compact inner square range for supplying high precision and it has two arduous processes in it. They are to find the number of points that are present in the square range and updating of the boundary (i.e) upper bound and lower bound is arduous because range queries are well secured by utilizing arbitrary space perturbation. The security of kNN query and range query is equipollent.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

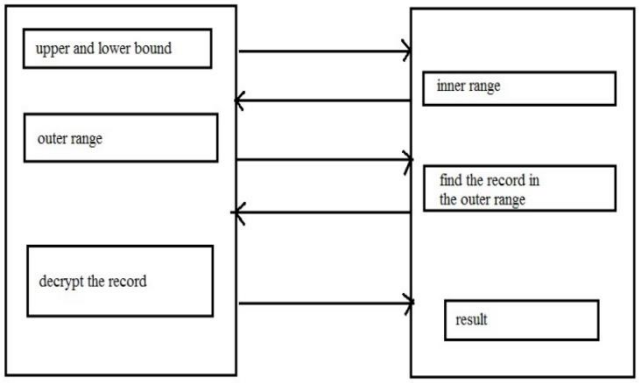

Fig 2: kNN query process.

The above diagram shows the process of k-nearestneighbor query.

4. Experimental Work

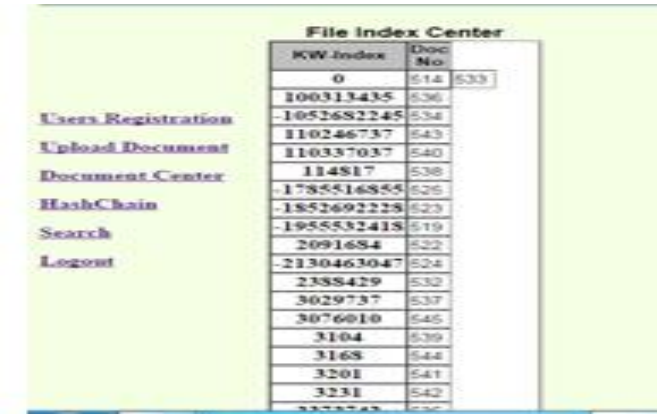

Fig 3: shows the index hash key value.

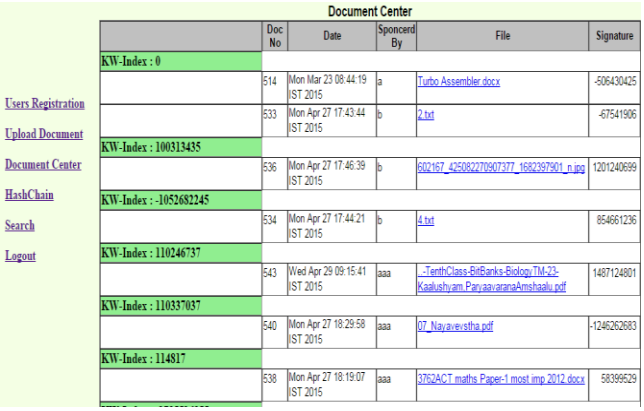

Fig 4: list of uploaded documents with his encrypted key.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

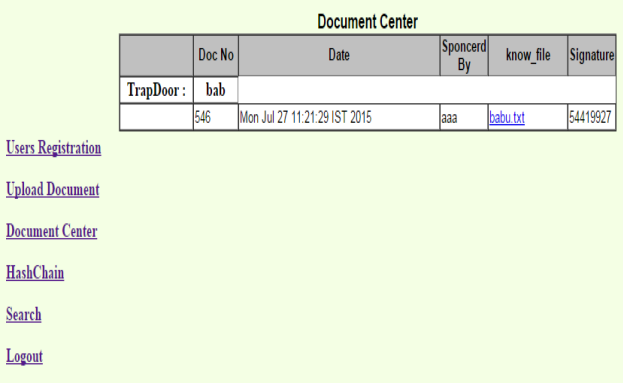

Fig 5: Based on the index The documents will be open.

5. Conclusion

Here we propose the desultory space perturbation approach to perform hosting query accommodations in the cloud, which mainly slakes the data confidentiality, query Privacy, Efficient query processing, and Low in-house work-load. Here we propose Seed Block Algorithm fortifies bulwark on well-organized range-queries, most proximate neighbour search queries. The fundamental inspiration is to arbitrarily transmute the multi-dimensional information, order preserving encryption, volume extension, integrating noise, and forfend queries processing. The main mechanisms used are the discomposure, engendering secure range queries and engendering secured most proximate neighbour queries.

6. References

[1] Huiqi Xu, Shumin Geo, Keke Chen, ”Building confidential and Efficient Query services in the Cloud with RASPData Parturbation” IEEE Transaction on knowledge and data Engineering vol:26 no:2,2014.

[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. K. andAndyKonwinski, G. Lee, D. Patterson, A. Rabkin, I.Stoica, and M. Zaharia, “Above the clouds: A berkeley view of cloud computing,” Technical Report, University ofBerkerley, 2009.

[3] H. Hacigumus, B. Iyer, C. Li, and S. Mehrotra, “Executing sql over encrypted data in the database-service-providermodel,”in Proceedings of ACM SIGMOD Conference, 2002.

[4] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, “Order preserving encryption for numeric data,” in Proceedings ofACM SIGMOD Conference, 2004.

[5] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, “Private information retrieval,” ACM Computer Survey, vol.45, no. 6,pp. 965–981, 1998.

[6] B. Hore, S. Mehrotra, and G. Tsudik, “A privacy-preserving index for range queries,” in Proceedings of Very LargeDatabases Conference (VLDB), 2004.

[7] J. Bau and J. C. Mitchell, “Security modeling and analysis,” IEEE Security and Privacy, vol. 9, no. 3, pp. 18–25,2011.

[8] K. Chen, L. Liu, and G. Sun, “Towards attack-resilient geometric data perturbation,” in SIAM Data MiningConference, 2007.

[9] M. F. Mokbel, C. yin Chow, and W. G. Aref, “The new casper: Query processing for location services withoutcompromisingprivacy,” in Proceedings of Very Large Databases Conference (VLDB), 2006, pp. 763–774.

[10] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, “Private information retrieval,” ACM Computer Survey, vol.45, no. 6, pp. 965–981, 1998.