PROCESSING PRIVATE KNN QUERY OVER UNTRUSTED DATA CLOUD

Miss Monika D.Rokade[1], Mr.S.A.Kahate[2]

[1]*M.E. Student, Computer Engineering Department, SPCOE , Dumberwadi, Pune Maharashtra, India*

[2] *Head of Computer Engineering Department, SPCOE , Dumberwadi, Pune Maharashtra, India*

ABSTRACT

Cloud computing and data outsourcing helps with more convenient ways of working .Cloud, data owner and customers/clients. Query of user and privacy of the data owners are most importance part of cloud computing.Many people researched on cloud computing and cloud security because query processing preserves data privacy of the owner as well as clients. To provide more security features a PH technique is used. Privacy Homomorphism (PH) is emphasizes to resolve the privacy of query processing from client side and cloud with the kNN query.

Keyword*: - Privacy kNN Query, privacy preserving process.*

1. INTRODUCTION

In cloud computing, client use data and querying services for outsourcing on the cloud data. During this process, data is the separate and private asset of the data owner, hence data must be protected in cloud and querying client. Query is send by the client may disclose the sensitive details/information of the client. They should be secure in cloud and data owner.In cloud computing one major problem is to protect both, data privacy, data owner and the client.The cloud refer Fig- 1. The social networking is one of the rising sectors facing such type of privacy problem [2].The various types of storage deploying,managing and provide in cloud computing using internet-based infrastructure. Goggle Docs, Amazon EC2,Microsoft Azure, and Online file storage etc are the services provided by cloud computing and they are used by many people in the world.

However, it is very sensitive issue to upload our personal data on the cloud because data privacy is the big issue and major problem of security. Private information has encrypted before outsourcing and creates the effective data utilization services and it is big challenging task.Symmetric Searchable Encryption (SSE) technique are used for encrypted data on the cloud ,there is leakage of data privacy.Order-preserving Encryption (OPE) are includes the similarity relevance and robustness [3].

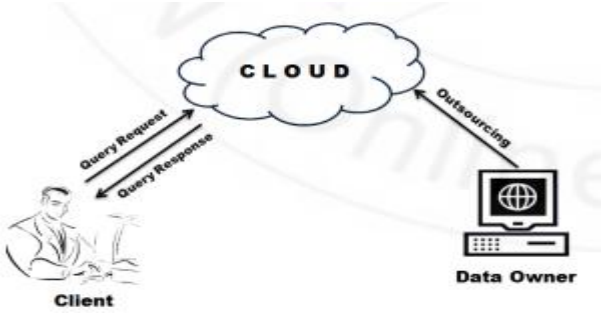


Figure 1: General Model for query processing in Cloud

2. RELETED WORK

A lot of studies were done prior to this which provides a secure framework and substitute encryption schemes, both are imperative. Here, we wish to present a systematic and result oriented solution based on Privacy Homomorphism (PH). It is an encrypted transformations mapping set of operations on plain text to another set of operations on Ciphertext[8]. There are three basic steps to solve problems of outsourced data processing query in cloud and for client.

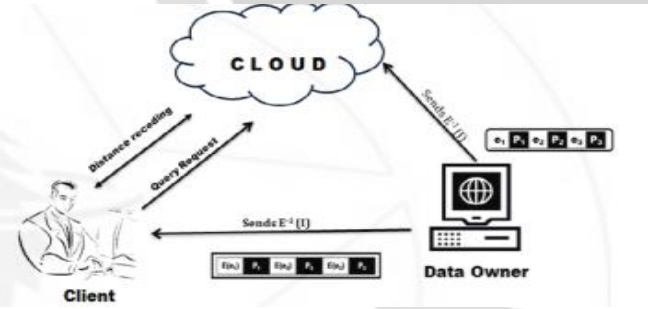
1. An index consists of multiple nodes which are used for processing queries including traversing the nodes. Data owner and cloud may not be able to trace the access pattern hence hardly get any clue of the query. Here, a client-lead processing terms excludes the display of query to third party.

2. To assess various types of complex queries such as kNN and other distance-based queries, an inclusionary set of client-cloud protocols must be organized to work together with a PH to supports most arithmetic operations.

3. IMPLEMENTATION DETAILS

A. Query Processing Framework

Distance-based queries processing a multidimensional index can be treated as tree nodes.They can divided into two alternate processes i.e. node traversal and distance access. The distances computed from the current node and query point. Query and data privacy, both procedures must remain secure in the outsourcing model of three parties i.e. when query is being processing not only data owner but the cloud can identify the traversed nodes also or may obtain any information that may point out the query point as the exact distances to the query point. The client should not access to the actual node contents during distance access and node traversal. Here, in fig-2, showing the framework of secure query processing. Whereas, other part is to protect data privacy, the client has only access to an encrypted version of the index, and must go ahead to process their query together with the cloud, which will decrypt the distances it, computes locally. The distance access is a collective procedure of the client and data cloud, in which not a single party has access to the actual distances [2].



The detailed process flow of this framework is as follows:

1. Send query to cloud by client

2.Data owner sends an encrypted variant of index . In each index node, the key entry .

3.The index has common topology as the basic index but each key value is encrypted. The index is to be saved at the client side for future connections.

4. The data owner sends decryption scheme to the data cloud for future distance decryption. It does not require that data owner should get involved in initial stage and reduce their involvement by handing over the task of decrypted indexing to the cloud.

5. Index in the cloud encrypted by the owner's private key through any public key cryptography. Then owner needs to forward their public key to the client and they recollects and decrypts the index from cloud.

6. The client is required to go for index node which results node that computes the local distances, and are sent to the data cloud which decrypts and re-codes them for the client

7. Only client can receive an encrypted version of the actual distances that acceptable and tolerable for the query processing. Whereas additionally to prevent the cloud from accessing the actual distances after decryption, the client is required to scramble local distances prior to forwarding them to the cloud from accessing the actual distance after decryption.

8. The traversal begins at the root node, and the node access process repeats until the query is completed.

4. CONCLUSIONS

As per the process mentioned herewith a study is conducted on processing problems of private queries on indexed data in a cloud. A secure traversal framework in indexed environment is given to secure protocols for such classic queries. The assumptions and approached mentioned in this paper are thoroughly useful, efficient to perform and effectively used under settings of different parameters. It has been summarized that the process mentioned here, on privacy homomorphism, is used to protect processing queries on cloud is high scalable

ACKNOWLEDGEMENT

We thankful to Teacher, Friends and Department of Computer Engineering for their constant guidelines and support.We also thankful college staff for providing the required infrastructure and support.

REFERENCES

[1] Guo, Yubin, et al. "A solution for privacy- preserving data manipulation and query on nosql database." Journal of Computers 8.6 (2013): 1427-1432.

[2] Hu, Haibo, et al. "Processing private queries over untrusted data cloud through privacy homomorphism." Data Engineering (ICDE), 2011 IEEE 27th International Conference on. IEEE, 2011.

[3] Nandhini, N., and P. G. Kathiravan. "An Efficient Retrieval of Encrypted Data In Cloud Computing."

[4] Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and YirongXu. Order preserving encryption for numeric data. In Proceedings of the 2004 ACM SIGMOD international conference on Management of data, SIGMOD '04,pages 563–574, New York, NY, USA, 2004. ACM.

[5] P. Samarati and L. Sweeney. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical report, 1998.

[6] TingjianGe, Stanley B. Zdonik, and Stanley B. Zdonik. Answering aggregation queries in a secure system model. In VLDB, pages 519–530, 2007.

[7] Haibo Hu and JianliangXu. Non-exposure location anonymity. In Yannis E. Ioannidis, DikLun Lee, and Raymond T. Ng, editors, ICDE, pages 1120–1131. IEEE,2009.

[8] Yonghong Yu and Wenyang Bai. Enforcing data privacy and user privacy over outsourced database service. JSW, 6(3):404–412,2011