Confidentiality and efficient query services in Cloud

Sapana Vasave
PG Student, Department of Information Technology,
MITCOE Pune, India
spnvsv@gmail.com

Kishor Kolhe
Associate Professor, Department of Information Technology,
MITCOE Pune, India
krkolhe@gmail.com

Abstract: **Now-a-days increasing users of public cloud computing infrastructures, using clouds to store data with query services is solution which gives more scalability and cost-saving. Hence, most of the data is sensitive that the data owner does not want to move their data to the cloud without user get the data confidentiality and query privacy are guaranteed. On the other hand, a secured query services should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing. Query services reduce the overhead of querying. RASP and kNN query services gives secured data before storing on cloud with the help of order preserving encryption and range query.**

Keywords: Cloud computing, RASP Perturbation, Query services in the cloud, Privacy, range query, kNN query

I. INTRODUCTION

Query services in the cloud are increasingly popular because of the unique advantages in scalability and cost-saving. With the cloud infrastructures, the service owners can conveniently scale up or down the service and only pay for the hours of using the servers. This is an attractive feature because the workloads of query services are highly dynamic, and it will be expensive and inefficient to serve such dynamic workloads with in-house infrastructures. However, because the service providers lose the control over the data in the cloud, data confidentiality and query privacy have become the major concerns. Summarization of these requirements for constructing a practical query service in the cloud as the CPEL criteria: data Confidentiality, query Privacy, Efficient query processing, and Low in-house processing cost. The basic idea is to randomly transform the multidimensional datasets with a combination of order preserving encryption, dimensionality expansion, random noise injection, and random project, so that the utility for processing range queries is preserved.

Driven by lower cost, higher reliability, better performance, and faster deployment, data and computing services 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 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 problem of some delay in the retrieving data from cloud storage. Data privacy and efficiency using file retrieval from Ostrovosky. In this scheme user can retrieves files from an untrusted server. Data Perturbation is to balance privacy protection and data utility. The kNN-R algorithm is designed to work with the RASP range query algorithm to process the kNN queries.

II. LITERATURE SURVEY

In this paper Authors states that, the query services contain the kNN and Random Space Perturbation Method. It also preserves multidimensional ranges, which allows existing indexing techniques to be applied to speedup range query processing. They have carefully analyzed the attacks on data and queries under a precisely defined threat model and realistic security assumptions. Extensive experiments have been conducted to show the advantages of this approach on efficiency and security.

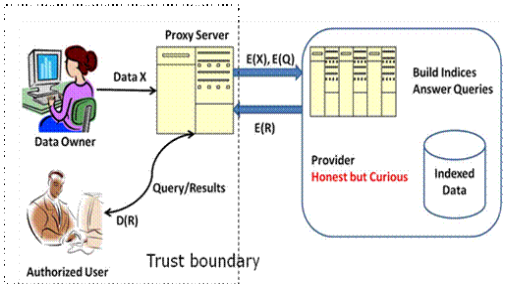

Fig 1: Architecture Diagram[1]

In this searchable encryption focus on single keyword search or Boolean keyword search, and rarely sort the search results. In this paper, the authors solve and describe challenging problem of privacy-preserving multi-keyword ranked search over encrypted data in cloud computing (MRSE).IN that they establish a set of strict privacy requirements for such a secure cloud data utilization system with coordinate matching threat models.[1]

In this paper used file retrieving scheme private keyword-based file retrieval which proposed by Ostrovsky. Three Methods are explain the information retrieval for ranked query (EIRQ) scheme, to reduce querying overhead incurred on the cloud.[2]

Consumer-Centric cloud computing, effective and efficient cloud search service, to find most relevant data, As sensitive data are encrypted before outsourcing to cloud. Authors describes synonym query, the main contribution of this paper is; multi-keyword ranked search to achieve more accurate search results and synonym-based search to support synonym queries.[3] Approximate queries for imbalanced load and space inefficiency among distributed data servers which used for locality-aware in hybrid systems.[4]

In this paper cloud computing as an emerging technology is expected to reshape information technology processes in the near future. Due to the overwhelming merits of cloud computing, e.g., cost-effectiveness, flexibility and scalability, more and more organizations choose to outsource their data for sharing in the cloud. As a typical cloud application, an organization subscribes the cloud services and authorizes its staff to share files in the cloud. Each file is described by a set of keywords, and the staff, as authorized users, can retrieve files of their interests by querying the cloud with certain keywords. In such an environment, how to protect user privacy from the cloud, which is a third party outside the security boundary of the organization, becomes a key problem. User privacy can be classified into search privacy and access privacy. Search privacy means that the cloud knows nothing about what the user is searching for, and access privacy means that the cloud knows nothing about which files are returned to the user. [5]

Range query is one of the most frequently used queries for online data analytics. Providing such a query service could be expensive for the data owner. With the development of services computing and cloud computing, it has become possible to outsource large databases to database service providers and let the providers maintain the range-query service. Without sourced services, the data owner can greatly reduce the cost in maintaining computing infrastructure and data-rich applications. However, the service provider, although honestly processing queries, may be curious about the hosted data and received queries. Most existing encryption based approaches require linear scan over the entire database, which is inappropriate for online data analytics on large databases. While a few encryption solutions are more focused on efficiency side, they are vulnerable to attackers equipped with certain prior knowledge. Random Space Encryption (RASP) approach that allows efficient range search with stronger attack resilience than existing efficiency-focused approaches. [6]

Random Space Perturbation method using kNN and range query is gives more secure data while before saving the data in the cloud. Encrypted data is generated random matrix number generator. Order preserving encryption is gives encrypted data.

III. SYSTEM MODEL

System model is woks on the three model that is encryption, decryptions and random number generator before saving the data in the cloud. New approaches are preserve data confidentiality and query privacy; the efficiency of query services and the benefits of using the clouds should also be preserved. The modified RASP perturbation is designed in such a way that the queried ranges are securely transformed into polyhedral in the RASP-perturbed data space, which can be efficiently processed with the support of indexing structures in the perturbed space. The RASP kNN query service (kNN-R) uses the RASP range query service to process kNN queries.

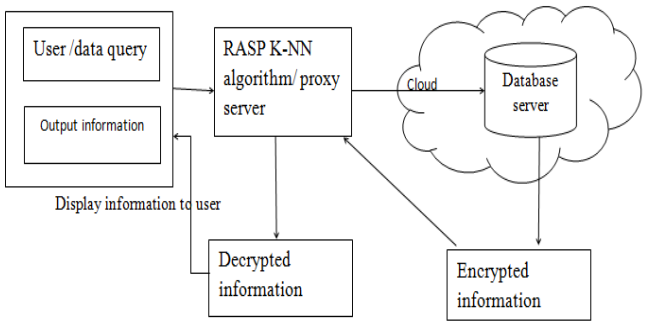

Fig 2: Procedure of RASP method

KNN algorithm is particularly sensitive to outliers and noise contained in the training data set. In this paper, we use the reverse kNN algorithm to map the training samples into clouds. New KNN algorithm based on Cloud Model and compares our algorithm with classic KNN algorithms and other well-known improved KNN algorithms using some data sets. Our approach could achieve a better or at least a comparable classification accuracy with other algorithms.

Hosting data query services in public clouds is an attractive solution for its great scalability and significant cost savings. However, data owners also have concerns on data privacy due to the lost control of the infrastructure. The modified RASP approach provides a privacy guarantee practical to the setting of cloud based computing, while enabling much faster query processing compared to the encryption-based approach. This will allow users to more intuitively understand the technical merits of the RASP approach via inter active exploration of the visual interface.

This helps users understand the unique advantages and possible limitations of the proposed approach. This demonstration system will be highly interactive and visual, allowing the users to easily understand the technical details.To provide secured query service with efficient query processing and significantly reduce the in-house workload:

Order Preserving Encryption (OPE), The enhanced crypto-index approach, New Casper approach.

IV. CONCLUSION

The RASP method for encryption technique to protect the data using query services in cloud infrastructure. The RASP perturbation approach to hosting query services in the cloud, which satisfies the CPEL criteria: data Confidentiality, query Privacy, Efficient query processing, and Low in-house workload. The requirement on low in-house workload is a critical feature to fully realize the benefits of cloud computing, and efficient query processing is a key measure of the quality of query services.

REFERENCE

[1] Huiqi Xu, Shumin Guo, Keke Chen et al. "Building Confidential and Efficient Query Services in the Cloud with RASP Data Perturbation", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING VOL: 26 NO: 2 YEAR 2014

[2] Qin Liu, Chiu C. Tan, Member, IEEE, Jie Wu, Fellow, IEEE, and Guojun Wang, Member, IEEE, "Towards Differential Query Services in Cost-Efficient Clouds," IEEE, TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL.25, NO.6 JUNE 2014

[3] Ning Cao, Cong Wang, Ming Li, Member, IEEE, Kui Ren and Wenjing Lou, Senior Member, IEEE, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS,VOL. 25, NO.1, JANUARY 2014

[4] K. Chen, R. Kavuluru, and S. Guo, "RASP:Efficient multidimensional range query on attack- Resilient encrypted databases," in ACM Conference on Data and Application Security and Privacy, 2011

[5] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing private queries over untrusted data cloud through privacy homomorphism," Proceedings of IEEE International Conference on Data Engineering (ICDE), pp. 601–612, 2011.

[6] Zohreh Alavi, Lu Zhou, James Powers, Keke Chen Data Intensive Analysis and Computing (DIAC) Lab, Kno.e.sis Center Department of Computer Science and Engineering Wright State University, Dayton, Ohio 45435, USA" RASP QS: Efficient and Confidential Query Services in the Cloud"

[7] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing private queries over untrusted data cloud through privacy homomorphism," Proceedings of IEEE International Conference on Data Engineering (ICDE), pp. 601–612, 2011

[8] Kishor R. Kolhe and Sapana S. Vasave, "Survey on: Query services in Cloud security," International Journal Of Science and Research (IJSR) ISSN (Online):2319-7064 Index Copernicus value (2013):6.14| Impact factor(2013):4.438