ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 2, Issue 4, April 2016

Semantically Encryption Data in a Cloud Based On Symmetric Key Algorithm

Sangeetha.M[1], Premalatha.J[2]

PG Scholar, Department of CSE, Vandayar Engineering College, Thanjavur, India [1]

Assistant Professor, Department of CSE, Vandayar Engineering College, Thanjavur, India [2]

***Abstract*—Current approaches to enforce fine-grained access control on confidential data hosted in the cloud are based on fine-grained encryption of the data. Under such approaches, data owners are in charge of encrypting the data before uploading them on the cloud and re-encrypting the data whenever user credentials or authorization policies change. Data owners thus incur high communication and computation costs. A better approach should delegate the enforcement of fine-grained access control to the cloud, so to minimize the overhead at the data owners, while assuring data confidentiality from the cloud. We propose an approach, based on two layers of encryption Enhanced Access control Scheme. The cloud performs a fine-grained encryption on top of the owner encrypted data. Enhanced Access control Scheme is an efficient group key management scheme that supports Semantically Secure Encrypted Relational Data. Our system assures the confidentiality of the data and preserves the privacy of users from the cloud while delegating most of the access control enforcement to the cloud.**

***Index Terms*— Security, k-NN classifier, outsourced databases, encryption.**

I. INTRODUCTION

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. When the files are stored in the clear forms, a naive solution to protect user privacy is for the user to request all of the files from the cloud; this way, the cloud cannot know which files the user is really interested in. While this does provide the necessary privacy, the communication cost is high.

Private searching was proposed by Ostrovsky et al. (referred to as the Ostrovsky scheme in this paper), which allows a user to retrieve files of interest from an un-trusted server without leaking any information. However, the Ostrovsky scheme has a high computational cost, since it requires the cloud to process the query (perform homo-morphic encryption) on every file in a collection. Otherwise, the cloud will learn that certain files, without processing, are of no interest to the user.

It will quickly become a performance bottleneck when the cloud needs to process thousands of queries over a collection of hundreds of thousands of files. I argue that subsequently proposed improvements, like also have the same drawback. Commercial clouds follow a pay-as-you-go model, where the customer is billed for different operations such as bandwidth, CPU time, and so on. Solutions that incur excessive computation and communication costs are unacceptable to customers.

To make private searching applicable in a cloud environment, our previous work designed a Co-operate Private Searching protocol (COPS), where a proxy server, called the Aggregation and Distribution Layer (ADL), is introduced between the users and the cloud. The ADL deployed inside an organization has two main functionalities: aggregating user queries and distributing search results. Under the ADL, the computation cost incurred on the cloud can be largely reduced, since the cloud only needs to execute a combined query once, no matter how many users are executing queries. Furthermore, the communication cost incurred on the cloud will also be reduced, since files shared by the users need to be returned only once. Most importantly, by using a series of secure functions, COPS can protect user privacy from the ADL, the cloud, and other users.

In this paper, I introduce a novel concept, differential query services, to COPS, where the users are allowed to personally decide how many matched files will be returned. This is motivated by the fact that under certain cases, there are a lot of files matching a user's query, but the user is interested in only a certain percentage of matched files. To illustrate, let us assume that Alice wants to retrieve 2% of the files that contain keywords "A, B", and Bob wants to retrieve 20% of the files that contain keywords "A, C". The cloud holds 1,000 files,

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 2, Issue 4, April 2016

where {F1, … F500} and {F501,.. F1000} are described by keywords "A, B" and "A, C", respectively. In the Ostrovsky scheme, the cloud will have to return 2, 000 files. In the COPS scheme, the cloud will have to return 1, 000 files. In our scheme, the cloud only needs to return 200 files. Therefore, by allowing the users to retrieve matched files on demand, the bandwidth consumed in the cloud can be largely reduced.

Motivated by this goal, I propose a scheme, termed Charge Effective Enquiry Facilities Complete Combination Then Delivery Coat (EIRGE), in which each user can choose the rank of his query to determine the percentage of matched files to be returned. The basic idea of EIRGE is to construct a privacy preserving mask matrix that allows the cloud to filter out a certain percentage of matched files before returning to the ADL. This is not a trivial work, since the cloud needs to correctly filter out files according to the rank of queries without knowing anything about user privacy.

Focusing on different design goals, I provide two extensions: the first extension emphasizes simplicity by requiring the least amount of modifications from the Ostrovsky scheme, and the second extension emphasizes privacy by leaking the least amount of information to the cloud.

II. Related Works

First review some Enhancement technique, user creation, vector calculations, decrypt the data, and then briefly introduce some classification over enhanced access control scheme.

A. *Daniel J. Abadi Yale University New Haven, CT, USA dna@cs.yale.edu*

Recently the cloud computing paradigm has been receiving significant excitement and attention in the media and blogosphere. To some, cloud computing seems to be little more than a marketing umbrella, encompassing topics such as distributed computing, grid computing, utility computing, and software as- a-service, that have already received significant research focus and commercial implementation. Nonetheless, there exist an increasing number of large companies that are offering cloud computing infrastructure products and services that do not entirely resemble the visions of these individual component topics. In this article we discuss the limitations and opportunities of deploying data management issues on these emerging cloud computing platforms (e.g., Amazon Web Services). We speculate that large scale data analysis tasks, decision support systems, and application specific data marts are more likely to take advantage of cloud computing platforms than operational, transactional database systems (at least initially). We present a list of features that a DBMS designed for large scale data analysis tasks running on an Amazon-style offering should contain. We then discuss some currently available open source and commercial database options that can be used to perform such analysis tasks, and conclude that none of these options, as presently architected, match the requisite features. We thus express the need for a new DBMS, designed specifically for cloud computing environments.

B. *Hasan Omar Al-Sakran*

Department of Management Information Systems, King Saud University, Riyadh, Saudi Arabia

Number of businesses using cloud computing has increased dramatically over the last few years due to the attractive features such as scalability, flexibility, fast start-up and low costs. Services provided over the web are ranging from using provider's software and hardware to managing security and other issues. Some of the biggest challenges at this point are providing privacy and data security to subscribers of public cloud servers. An efficient encryption technique presented in this paper can be used for secure access to and storage of data on public cloud server, moving and searching encrypted data through communication channels while protecting data confidentiality. This method ensures data protection against both external and internal intruders. Data can be decrypted only with the provided by the data owner key, while public cloud server is unable to read encrypted data or queries. Answering a query does not depend on it size and done in a constant time. Data access is managed by the data owner. The proposed schema allows unauthorized modifications detection.

C. *Ming LiAffiliated withDepartment of ECE, Worcester Polytechnic Institute, Shucheng Yu, Kui Ren, Wenjing Lou*

Online personal health record (PHR) enables patients to manage their own medical records in a centralized way, which greatly facilitates the storage, access and sharing of personal health data. With the emergence of cloud computing, it is attractive for the PHR service providers to shift their PHR applications and storage into the cloud, in order to enjoy the elastic resources and reduce the operational cost. However, by storing PHRs in the cloud, the patients lose physical control to their personal health data, which makes it necessary for each patient to encrypt her PHR data before uploading to the cloud servers. Under encryption, it is challenging to achieve fine-grained access control to PHR data in a scalable and efficient way. For each patient, the PHR data should be encrypted so that it is scalable with the number of users having access. Also, since there are multiple owners (patients) in a PHR system and every owner would encrypt her PHR files using a different set of cryptographic keys, it is important to reduce the key distribution complexity in such multi-owner settings. Existing cryptographic enforced access control schemes are mostly designed for the single-owner scenarios. In this paper, we propose a novel framework for access control to PHRs within cloud computing environment. To enable fine-grained and scalable access control for PHRs, we leverage attribute based encryption (ABE) techniques to encrypt each patients' PHR data. To reduce the key

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 2, Issue 4, April 2016

distribution complexity, we divide the system into multiple security domains, where each domain manages only a subset of the users. In this way, each patient has full control over her own privacy, and the key management complexity is reduced dramatically. Our proposed scheme is also flexible, in that it supports efficient and on-demand revocation of user access rights, and break-glass access under emergency scenarios.

D. PriyankaKorde, Vijay Panwar, SnehaKalse

Cloud computing facilitates efficient management of Personal Health records. When the health records are uploaded on the cloud there is an advantage of easy accessibility. At the same time there is the risk of privacy and security. The data has to be encrypted and the patient should be able to control the data access. The managing of the Personal Health Records should be both scalable and secure. This paper addresses the issue of scalability and security. It proposes to use the concept of attributes for maintaining the data and corresponding keys are given to the users. This data is encrypted with Advanced Encryption Standard and then uploaded on the cloud.

E. Archana Sharma

The Cloud has become a new vehicle for delivering resources such as computing and storage to customers on demand. Rather than being a new technology in itself, the cloud is a new business model wrapped around new technologies such as server Virtualization that take advantage of economies of scale and multi-tenancy to reduce the cost of using information technology resources. From one perspective, cloud computing is nothing new because it uses approaches, concepts, and best practices that have already been established. From another perspective, everything is new because cloud computing changes how we invent, develop, deploy, scale, update, maintains, and pay for applications and the infrastructure on which they run. Nonetheless, there exist an increasing number of large companies that are offering cloud computing infrastructure products and services that do not entirely resemble the visions of these individual component topics. The challenge of building consistent, available and scalable data management systems capable of serving peta bytes of data for millions of users has confronted the data management research community as well as large internet enterprises. Financial institutions are not strangers to cloud computing adoption. One of the earlier cloud uses in banks and financial institutions were for SaaS deployments, which allowed for more social media banking. However, now FI.s faces the issue of security due to the increased number of data leaks. As a result, cloud within IT strategies and architecture for FIs will increase the risk of a security breach among servers and networks unless there is an adoption of a multiyear cloud strategy to keep data protected. This paper highlights the data management in cloud applications and deployments of various services of cloud computing in Financial Institutions with the study of risk factors in the deployment of transaction data of Financial Institutions on clouds.

III. PROPOSED SYSTEM

In the proposed system, the process like updating, deleting of the data can be having verification process can be done. In the system, the data can be having security based on the cryptographic primitives for the data integrity production. The data can be having the dynamic security for the erasable data in the cloud and it can be having verification process. The security can be performed based on the two layer encryption in the cloud data. This paper concentrates on executing the two encryption based Enhanced Access control Scheme algorithm classification method over encrypted data in the cloud computing environment.

A. User Creation for SSE

The user creates their own login credential when they want entering in cloud. The keys are automatically generated when the user register in cloud. Considering the large number of data users and documents in the cloud, it is necessary to allow multi-keyword in the search query and return documents in the order of their relevancy with the queried keywords. Scoring is a natural way to weight the relevance. Based on the relevance score, files can then be ranked in either ascending or descending. Several models have been proposed to score and rank files in information retrieval (IR) community.

B. Vector Calculation

Although all data files, indices and requests are in encrypted form before being outsourced onto cloud, the cloud server can still obtain additional information through statistical analysis. We denote the possible information leakage with statistic leakage. There are two possible statistic leakages, including term distribution and inter distribution. The term distribution of term t is t's frequency distribution of scores on each file .The inter distribution of file f is file f's frequency distribution of scores of each term. Term distribution and inter distribution are specific. They can be deduced either directly from cipher text or indirectly via statistical analysis over access and search pattern. The vector calculation is performed to the term frequency user querying file. Here access pattern refers to which keywords and the corresponding files have been retrieved during each search request, and search pattern refers to whether the keywords retrieved between two requests are the same.

C. Enhanced Access control Scheme algorithm

Existing SSE schemes employ server-side ranking based on order preserving encryption to improve the efficiency of retrieval over encrypted cloud data. However, server-side ranking based on order-preserving encryption violates the privacy of sensitive information, which is considered un-comprisable in the security-oriented third-party cloud

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 2, Issue 4, April 2016

computing scenario, to achieve data privacy, ranking, Privilege Access Control has to be left to the user side. The user including searchable index return and ranking score calculation. In the proposed scheme, the data owner encrypts the searchable index with Enhanced Access control Scheme algorithm. When the cloud server receives query consisting of multi-keyword, it computes the scores from the encrypted index stored on cloud, and then returns the encrypted scores of files to the data user. The retrieval takes a Multi-Keyword communication between the cloud server and the data user. We thus name the scheme as Enhanced Access control Scheme algorithm in which ranking is done at the user side while scoring calculation is done at the server side.

D. *User Query for SSE*

To alleviate the computational burden on user side, computing work should be at the server side, so we need an encryption scheme to guarantee the operability and security at the same time on server side. Enhanced Access control Scheme algorithm allows specific types of computations to be carried out on the corresponding cipher text. The result is the cipher text of the result of the same operations performed on the plaintext. That is, Enhanced Access control Scheme algorithm allows computation of cipher text without knowing anything about the plaintext to get the correct encrypted result. Although it has such a fine property, original fully homo-morphic encryption scheme, which employs ideal lattices over a polynomial ring, is too complicated and inefficient for practical utilization.

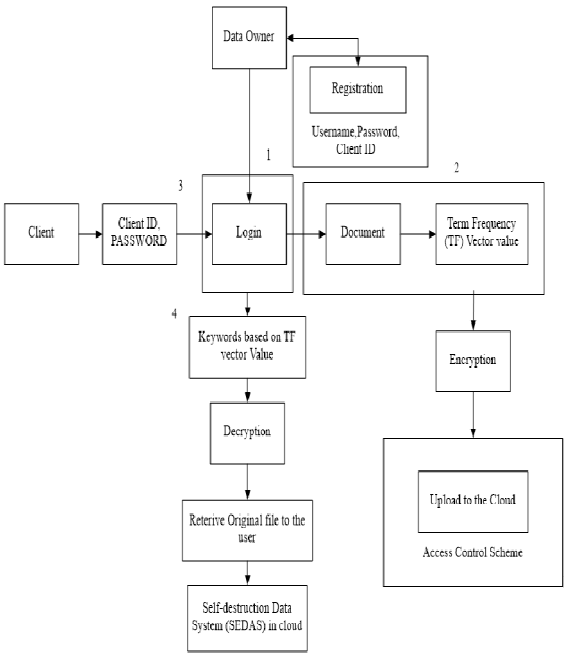

Fig 1 System Architecture

IV. EXPERIMENTAL ANALYSIS

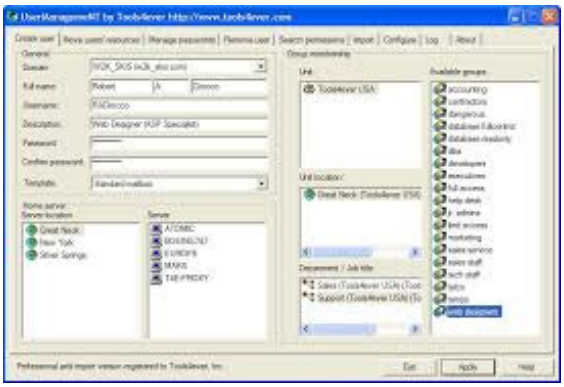

Fig 2 User Creation

tf x idf

$$w_{ik} = tf_{ik} * \log(N / n_k)$$

T_k = term k in document D_i

tf_{ik} = frequency of term T_k in document D_i

idf_k = inverse document frequency of term T_k in C

N = total number of documents in the collection C

n_k = the number of documents in C that contain T_k

$$idf_k = \log\left(\frac{N}{n_k}\right)$$

Fig 3 Vector Calculation

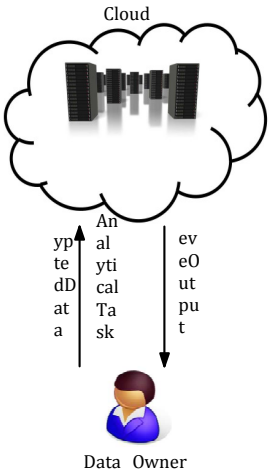

Fig 4 Encrypted Data

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 2, Issue 4, April 2016

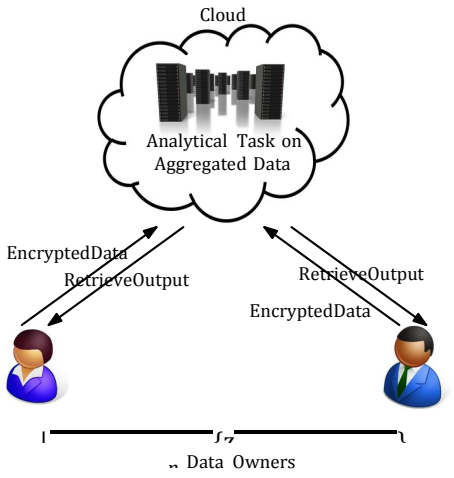

Fig 5 Retrieve Data

Fig 6 Self Destruction

V. CONCLUSION

In this paper, we motivate and solve the problem of secure multi-encryption retrieval over encrypted cloud data. Our protocol protects the confidentiality of the data, user's input query, and hides the data access patterns. We devise a server-side ranking SSE scheme. We then propose a SSE scheme employing the fully homo morphic encryption, which fulfills the security requirements of multi-keyword over the encrypted cloud data. Future work we design a spatial database encryption scheme that produces a transformed database from the original database by using network distances and plan to study on a pruning technique to improve the performance of our method by reducing the size of the returned candidate set.

REFERENCES

[1] P. Williams, R. Sion, and B. Carbunar, "Building castles out of mud: Practical access pattern privacy and correctness on untrusted storage," in Proc. 15th ACM Conf. Comput. Commun. Security, 2008, pp. 139–148.

[2] P. Paillier, "Public key cryptosystems based on composite degree residuosity classes," in Proc. 17th Int. Conf. Theory Appl. Cryptographic Techn., 1999, pp. 223–238.

[3] B. K. Samanthula, Y. Elmehdwi, and W. Jiang, "k-nearest neighbor classification over semantically secure encrypted relational data," eprint arXiv:1403.5001, 2014

[4] C. Gentry, "Fully homomorphic encryption using ideal lattices," in Proc. 41st Annu. ACM Sympos. Theory Comput., 2009, pp. 169– 178

[5] C. Gentry and S. Halevi, "Implementing gentry's fully-homomorphic encryption scheme," in Proc. 30th Annu. Int. Conf. Theory Appl. Cryptographic Techn.: Adv. Cryptol., 2011, pp. 129–148.

[6] A. Shamir, "How to share a secret," Commun. ACM, vol. 22, pp. 612–613, 1979.

[7] D. Bogdanov, S. Laur, and J. Willemson, "Sharemind: A framework for fast privacy-preserving computations," in Proc. 13th Eur. Symp. Res. Comput. Security: Comput. Security, 2008, pp. 192–206.

[8] R. Agrawal and R. Srikant, "Privacy-preserving data mining," ACM Sigmod Rec., vol. 29, pp. 439–450, 2000.

[9] Y. Lindell and B. Pinkas, "Privacy preserving data mining," in Proc. 20th Annu. Int. Cryptol. Conf. Adv. Cryptol., 2000, pp. 36–54.

[10] P. Zhang, Y. Tong, S. Tang, and D. Yang, "Privacy preserving Naive Bayes classification," in Proc. 1st Int. Conf. Adv. Data Mining Appl., 2005, pp. 744–752.

[11] A. Evfimievski, R. Srikant, R. Agrawal, and J. Gehrke, "Privacy preserving mining of association rules," Inf. Syst., vol. 29, no. 4, pp. 343–364, 2004.

[12] R. J. Bayardo and R. Agrawal, "Data privacy through optimal kanonymization," in Proc. IEEE 21st Int. Conf. Data Eng., 2005, pp. 217–228.

[13] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing private queries over untrusted data cloud through privacy homomorphism," in Proc. IEEE 27th Int. Conf. Data Eng., 2011, pp. 601–612.

[14] M. Kantarcioglu and C. Clifton, "Privately computing a distributed k-nn classifier," in Proc. 8th Eur. Conf. Principles Practice Knowl. Discovery Databases, 2004, pp. 279–290.

[15] L. Xiong, S. Chitti, and L. Liu, "K nearest neighbor classification across multiple private databases," in Proc. 15th ACM Int. Conf. Inform. Knowl. Manage. 2006, pp. 840–841.