ISSN 2395-1621

Master-key based AES Encryption with Enhanced Multi-keyword Ranked Search for Cloud Data

[#1]MaheshM. Sapate

[1]mahesh.sapate@gmail.com

[#1]Department of Computer Engineering, JSPM's ICOER Wagholi, Pune, SavitribaiPhule Pune University, Pune, India

ABSTRACT

The cloud provided scalability, availability, maintainability, security attracting business organization to push their crucial data on cloud servers. Cloud storage is affordable and has great convenience. Mission critical sensitive data need careful handling before putting on cloud server. As cloud is semi trusted domain and security mechanism provide by them might not fulfill organization expectation. Business organizations use their own mechanism to encrypt sensitive data before uploading on cloud server. As organization has multiple users so keeping record of thousands of documents encryption keys is tedious task. Also recovery of sensitive document is very challenging task on revocation of user.A system proposed here is Master key based AES encryption for sensitive data before uploading on cloud storage. In this scheme every document is encrypted with encryption key which is generated runtime by aggregating data owner's private secure key and organization's Master key. This mechanism provides enhanced security mechanism and breach of data owner's private secure key or organizations master key will not have effect on encrypted document. With this approach multiple data owners can contribute in sensitive data and they have freedom to use any secure key for encryption. Data users do not need permission for decryption of document once they are authenticated and authorized. Data user's Multi-keyword ranked search over encrypted documents is efficiently handled by encrypting search query with special encryption mechanism, which supports creation, update, and deletion operations. The secure AES symmetric algorithm is used to encrypt the documents, indexes and search queries.This proposed encryption and search mechanism solves challenges related to multi data owner contribution, user revocation, and search query performance issues.
Extensive experiments are carried out to demonstrate the efficiency of the proposed scheme.

Keywords- Sensitive Data, AES algorithm,Searchable Encryption, Cloud Computing

ARTICLE INFO

Article History

Received :4th January 2016

Received in revised form :

5th January 2016

Accepted : 6th January,2016

Published online :

6th January, 2016

I. INTRODUCTION

The Cloud provided scalability, availability, maintainability, security attracting business organization to push their crucial data on cloud servers. Cloud storage is affordable and has great convenience. Cloud comes to ease for organization as it takes the burden of hardware and software resources investment by providing the data warehouse and its maintenance. Many cloud platforms are available in market like Google Drive, iCloud, SkyDrive, Amazon S3, Dropbox and Microsoft Azure provide storage services [1].

Mission critical sensitive data need careful handling before putting on cloud server. As cloud is semi trusted domain with low degree of transparency so privacy and security mechanism provide by them might not fulfil organization expectation. As organization's data owner, data users and the cloud service provider are in the different trusted domain, the cloud service provider maintained data may be exposed to the unauthorized users. Before storing the valuable sensitive data in cloud, the data needs to be encrypted. Data encryption assures the data confidentiality and integrity [2]. Business organizations use their own mechanism to encrypt sensitive data before uploading on cloud server. As

organization has multiple users so keeping record of thousands of documents encryption keys is tedious task. Also granting permission to many users put addition burden on data owner. Recovery of sensitive document is very challenging task on revocation of user. The cloud server may leak information to unauthorized entities or even be hacked, which puts the outsourced data at risk. Traditionally, sensitive data should be encrypted by data owners before outsourcing, which, however, obsoletes traditional data utilization service like keyword-based information retrieval [4]. To preserve the data privacy we need to design a searchable algorithm that works on encrypted data. Many researchers are actively contributing to searching on encrypted data. The search techniques may be single or multi keyword search. In big database the search may retrieve many documents to be matched with keywords. This causes problem for a cloud user as he need to go through all documents and have most relevant documents [5]. Search based on ranking is one option, wherein the documents are ranked based on their relevancy to the keywords. The researchers combined the rank of documents with multiple keyword searches to come up with efficient economically viable searchable encryption techniques [6]. In searchable encryption related work, computation time and computation overhead are the two most frequently used parameters by the researchers in the domain for analysing the performance of their schemes. Computation time is the amount of time required to perform a retrieval and search process for example entering search keyword then generating its encrypted form generally named as trapdoor. Computation overhead is related to CPU and memory utilization [7].

II. LITERATURE SURVEY

Wast research and implementation has been done for searching techniques in the cloud but it only support exact keyword search. For huge number of documents and large number of data users, it is necessary to allow multiple keywords in the search request and return documents in the order of their relevance to these keywords.The data encryption is an effective way to protect the confidentiality and privacy of data in cloud. But searching, efficiency gets downgraded.Huge work has been proposed in security and privacy preserving multi-keyword search on encrypted data for cloud computing world. A lot of techniques have been put forward that solve the problem of effective secure ranked keyword search over encrypted cloud data. The disadvantages of these techniques are they does not support multi-keyword, does not include IDF (define) for the calculation of scores, does not use advanced crypto techniques [1].Song et al [7], for the first time proposed the practical symmetric searchable method based on cryptography. In this scheme the file is encrypted word by word. To search for a keyword user sends the keyword with same key to the cloud. The drawback of this scheme is that the word frequency will be revealed.Goh et al [8] tried to overcome the drawback of Song's scheme by constructing secure index table using pseudorandom functions and unique document identifier randomized bloom filters.Chang et al [9] proposed scheme, an index is built for each document. The scheme is more secured compared to Goh's scheme since number of words in a file is not disclosed. The limitation of this scheme is that it is less efficient and does not support arbitrary updates with new words. Curtmola et al [3] for the first time proposed the concept of symmetric searchable encryption (SSE), where an inverted index (implemented using linked list) having document identifiers is maintained for each keyword. Every node in the list stores information about the position and the decryption key of the next node. The nodes from all inverted indexes are encrypted with random keys and are randomly inserted into an array. With this, by knowing position and decryption key of the first node of an inverted index, it is possible to find all documents which include the corresponding keyword.Kamara et al [3] proposed an extended version of SSE called dynamic SSE (DSSE), where addition and deletion of documents can be performed in index table.The first public key encryption with keyword search (PEKS) was proposed by Boneh et al. The scheme suffers from inference attack on trapdoor encryption method. Baek et al [4], Rhee et al improved hardness of security of Boneh's scheme. Baek's scheme introduces the concept of conjunction of keyword search. The public key encryption methods are computationally time consuming and complex that makes these algorithms inefficient. In Yang et al [1] scheme the encrypted data is searched by individual users using a unique key allotted to them. The scheme suffers from key management. S.Buyrukbilen et al [3], introduce the first method that provides ranked results from multi-keyword searches on public-key encrypted data. By avoiding a linear scan of the documents and by parallelizing the computations to the possible extent, this method reduces the computational complexity of public key cryptosystem.Wenhai Sun etal [1]. proposed a MRSE scheme that works on similarity based ranking. Here search index is created on the basis of term frequency and vector space. Search index is used for multi keyword search and ranking the search result. Search efficiency is improved by applying tree structure on index.Ruturaj Desai, et al [5]. proposes a new scheme to solve the problem of multi keyword search over encrypted data using trusted third party in cloud computing. User will encrypt their data locally. Before encrypting data, the index will be created. Trusted third party will use all these indexes to search data similar to the search query of user. Using these search results, cloud server will send encrypted document to the user.Ranked search greatly improves system usability by normal matching files in ranked order regarding to certain relevance criteria. Ranked search can also elegantly eliminate unnecessary network traffic by sending back only the most relevant data.

III. PROBLEM STATEMENT AND PROPOSED SYSTEM

A. Problem Statement

The cloud provided scalability, availability, maintainability, security attracting business organization to push their crucial data on cloud servers. Cloud storage is affordable and has great convenience. Mission critical sensitive data need careful handling before putting on cloud server. As cloud is semi trusted domain and security mechanism provide by them might not fulfil organization expectation.For the protection of data privacy, sensitive data has to be encrypted

before outsourcing, which makes effective data utilization a very challenging task.Business organizations use their own mechanism to encrypt sensitive data before uploading on cloud server. As organization has multiple users so keeping record of thousands of documents encryption keys is tedious task. Also recovery of sensitive document is very challenging task on revocation of user.

B. Proposed System

To allow efficient multi-keyword search on the cloud server, the data user must specify a set of keywords and submit the search request to the server. To preserve the privacy, the input keywords should not be exposed to the adversary. We first propose a search scheme using polynomial function to enable ranked multi-keyword search and hide input keywords in cloud computing system, and then show how to improve it to be privacy-preserving against different levels of threat models in our framework.To enable effective, efficient and secure multi-keyword ranked search over encrypted cloud data under the aforementioned models, our mechanism is aiming to achieve the following:AES Algorithm for Document, Index encryption and decryption:AES algorithm is widely used for protecting sensitive data. AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data. This symmetric behaviour helps in solving shared key issue. AES algorithm provides better security over other well-known algorithm.Cloud Server based index update:In Current implementation data owner is responsible for the update operation of his documents stored in the cloud server. Data owner generates the update information locally and sends it to the server. We propose a scheme where cloud server will update information and re-construct or modify a tree-based index structure accordingly.Multi Owner data search support:All previous techniques partially support multi owner data search feature. To decrypt documents data user need secret keys from multiple data owners.We propose a method where authenticated and authorized data user can search for data from multiple data owners and he/she does not need secret keys for decryption of documents.Remove secure key issue in user revocation:User revocation is big challenge in searchable encryption as it needs to rebuild the tree based index and need to distribute secure keys to all authorized users.Our implementation will be AES symmetric algorithm based so it will rebuild tree based index internally. Data users do not need secure keys.User defined relevance for query keyword:Naïve user provided relevance in search query helps to find relevant documents easily.We propose a mechanism for data users to explicitly set relevance for search query words and its relevance will be considered while retrieving documents.Space and Efficiency improvement by dynamic vector space model:We propose a scheme of dynamic vector space model. Any node in balanced tree based index will have vector, whose size will be dynamically updated as per the keywords in its document.

IV. ARCHITECTURE OF PROPOSED SYSTEM MODEL

Architecture has three main entities:

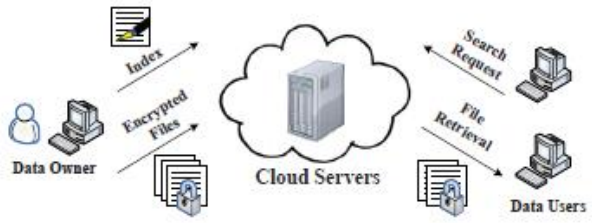

Fig1. Architecture of keyword search over encrypted data in cloud.

1. Data owner: Data owner is responsible for uploading encrypted sensitive data on cloud server. The encryption algorithm is chosen by either organization or Data owner if he is working individually.
2. Data user: Data user is end user who wants to see sensitive information for business operations. Data user wants to search the required files he enters a keyword and these keywords are encrypted before sending to cloud server or application server.
3. Cloud server: It is the place of hardware and software resources where a pool of data files and different applications can store.

Encryption Process:
Before uploading sensitive data file it needs to be encrypted the encryption of data file is done by combining data owner's secure private key and organizations secure private master key. The encrypted file then uploaded on cloud server.

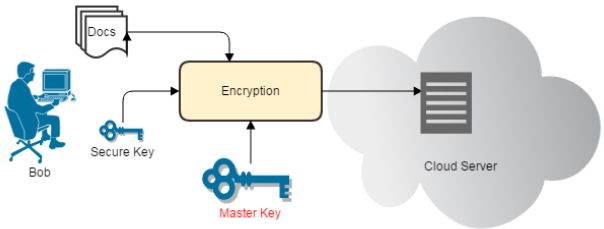

Fig2.Encryption of data file

Decryption Process:
Data user needs to retrieve sensitive data for business purpose so he or she search over encrypted file and ask for decryption of interested file. Cloud system then use stored data owners' secure private key and master key of organization to decrypt the file. Decrypted file is then given to data user.

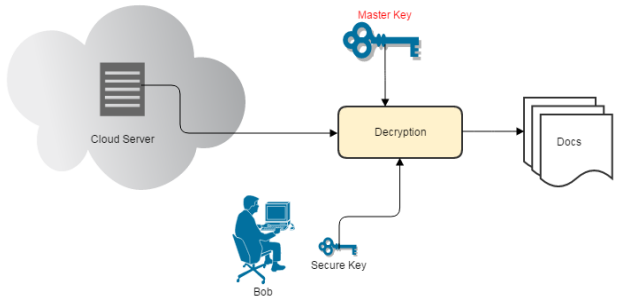

Fig 3.Decryption of data file

V. MATHEMATICAL MODEL

Notations and Preliminaries:
•W :Set of keywords W = {w1,w2,…..wm}.
• m :Count of keywords in W.
• Wq:Keywordsin the query.Wq subset of W.
• F :Sensitive documents F = {f1, f2,….fn}.
• n :Count of documents in F.
• C :Encrypted documentson cloud C = {c1,c2,….cn}.
• Q :Data user query for keyword set Wq.
• TD :Encrypted side of Qtrapdoor search request.
•DO : Set of Data owners DO = {DO1, DO2,….. DOk}
•DU: Set of Data users DU= {DU1, DU2,….. DUk}
•MK: Master key of organization
• AES:AES algorithm

Encryption Key:

EK = MasterKey + DO (private key)

Document Encryption:

C = AES(F, EK)

These encrypted documents C are uploaded on cloud server
Data User request for data
Request→query (keyword) →Re-ranking
After data user's proper successful authentication and authorization decryption process is started

Decryption Key:

DK = Master Key + DO (private key)

This decryption key is applied to encrypted cloud document to decrypt the sensitive data

Document Decryption:

F = AES(C, DK)

This way document is encrypted and decrypted.

VI. IMPLEMENTATION STRATEGY AND EXPERIMENTAL SETUP

To implement master key based AES algorithm following steps need to follow: Setup, EncryptionMulti-keyword search and Decryption.
•Setup: Data owners and data users are created in system and AES 128 bit algorithm is selected for encryption and decryption process. 128 bit Master key is finalized and kept secret, means this master key will never get shared with data owners and data users. Data owner have sensitive documents to be encrypted.
•Encryption: Before encryption all keywords fromdocuments are extracted and used to build index tree. Encryption processuses master key and data owners secure private key to create encryption key at runtime. Encryption key will be used to encrypt the document. This algorithm generates ciphertext of sensitive document and uploaded to cloud server.
•Multi-keyword search: Data user builds search query of multiple keywords. This search query is encrypted with master key and sent to cloud server for document retrieval.
•Decryption: Retrieved documents are decrypted by decryption key, generated by combining master key and data owners secure private key. Inputs for decryptionprocess are cipher text and decryption key. This algorithm decrypts the ciphertext and generates plain text. This plain text is nothing but the original message.

By following above steps a proposed system can be implemented.

VII. CONCLUSION

In this paper, we propose master key based AES encryption with enhanced multi-keyword ranked search for encrypted cloud data,which supports dynamic update operations. Eventuallyexperiments on the real-world dataset demonstrate the effectiveness and efficiency of our scheme.

REFRENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829-837, Apr, 2011.

[2] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition," ACM SIGCOMM Comput. Commun.Rev., vol. 39, no. 1, pp. 50-55, 2009.

[3] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.

[4] S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.

[5] A. Singhal, "Modern Information Retrieval: A Brief Overview," IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, Mar. 2001.

[6] I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Publishing, May 1999.

[7] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.

[8] E.-J. Goh, "Secure Indexes," Cryptology ePrint Archive, http:// eprint.iacr.org/2003/216. 2003.

