A Survey Paper on Storage and Access of Sensitive Information on Cloud

Sachin L S[1], Yashoda S Hubballi[2], Shaila H Koppad[3]

[1] 6th Semester PG Student, Department of MCA, RV College of Engineering®, Karnataka, India
[2] 6th Semester PG Student, Department of MCA, RV College of Engineering®, Karnataka, India
[3] Assistant Professor, Department of MCA, RV College of Engineering®, Karnataka, India

ABSTRACT

As the cloud technology is developing outsourcing the sensitive data is also increasing. To secure the outsourced data encryption should be done to the outsourced data but accessing the encrypted data from cloud is not as simple as accessing an unencrypted data. So, to achieve efficient search on encrypted data and to increase security for outsourced data in this paper searchable schemes are proposed using TF X IDF Model, Vector Space Model and Bloom Filter technique. These are being used to improve the searching efficiency for ranked and multiple keywords searching system. The most powerful encryption algorithm AES is being used for Encryption or Decryption of the data, keywords and index. Two factor authentication technique is being used to improve security of the system. In this work Multiple data owners and multiple data users scenarios are considered in this system unlike the many proposed systems. The proposed work can give ranked top k search results without leaking any private information.

Keyword: - *Cloud, Storage, Access, Outsource, Cloud Security, Keyword Search, Encryption, Encrypted Data*

1. INTRODUCTION

Cloud is the collection of servers which are approached over the internet. The database and software working on those servers is also accessed. Cloud security is associated with the complete infrastructure of cloud computing. It is the set of applications and policies related with cloud computing. It is also associated with directions used to protect data, services and virtualized IP. Cloud computing is associated with the storing of data and acquiring it over the internet. It is also referred to the usage of the services provided by the service providers.

Advantages of cloud computing involve high speed of data access, easy back-up of data and also restoring of data. It also involves savings of cost, portability, trustworthiness, high storage capacity, faster development. Issues involved with cloud security are distributed cloud computing services, denial of service attacks, employee laxity, loss of data, insufficient data backups and system faults.

Methods on storage and access of sensitive information on cloud is to be introduced in this project. Two-factor authentication has to be introduced and thus providing cloud security. Cloud computing has become a famous scheme for financial savings and governance of information over the last few years. A general approach to save the information which is stored in the cloud from confidentiality harm is to encrypt the information before outsourcing. Before deploying to cloud servers the information must be encrypted. This must be done in order to maintain privacy consideration. Usage of plaintext keyword search method which is a traditional information features is impossible for encrypted information. To fix this issue, this paper presents an enquiry system ranked multi-keyword over encrypted cloud information that effectively supports dynamic activities.

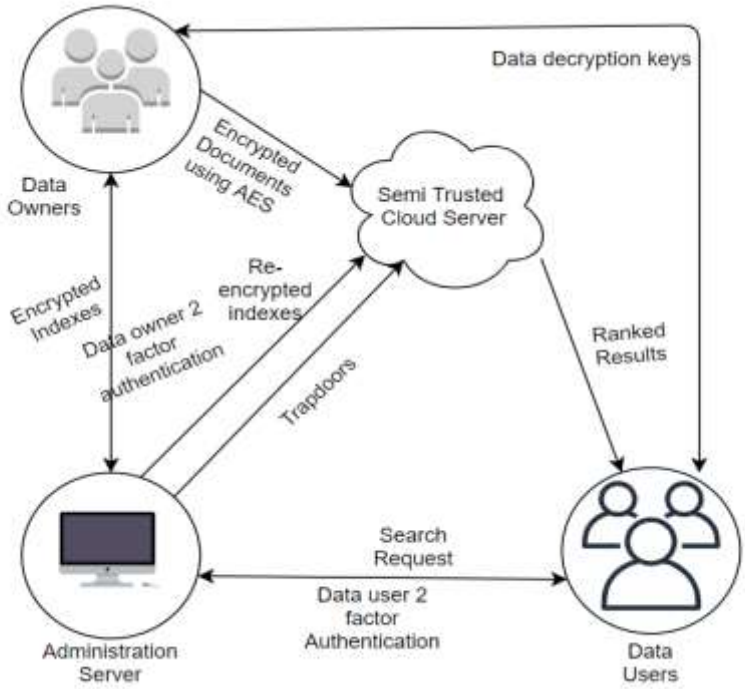

Fig -1: Model of the system

Figure 1 describes multiple data owners having multiple documents to outsource is being considered and each data owners use separate keys to encrypt the outsourcing documents and keywords. Any valid users can search these documents even without knowing their secret keys. User will get the most applicable top k search results without leaking any private details. For encryption and decryption AES algorithm is being used which is the most powerful algorithm as of today and Two Factor authentication method is being used to increase the security of the system.

In this work the introduction section explained about the cloud structure and two factor authentication method. The introduction section is followed by the background work of the techniques used by various authors for searching techniques and providing security on cloud. The next section gaps identified explained about the drawbacks of the techniques used., and how it can be overcome is explained using AES Algorithm followed by that particular section. The Conclusion section provides the information about considering multiusers and multidata owners to propose multi keyword search on encrypted cloud data and security provided at cloud for sensitive data followed by the references used in the work.

2. LITERATURE SURVEY

S.Poonkodi et al. [1] talked about a threshold proxy re-encryption approach. They also explainedabout secure decentralized code integration. It is used to form a distributed storage system which is secure. Servers related to storage and key are introduced in this paper. Data stored and accessed via servers is explained. Methods involved are being explained in this paper. K Govinda et al. [2] has proposed auditing protocol which is secured. RSA algorithm has been used in this paper. Digital signature is being put forth using RAS algorithm. It also deals with encryption and decryption process. Detailed study on encryption and decryption is being done in this paper. RSA algorithm working process and its advantages are also highlighted in this paper.

Sultan and William Allen [3] explained about public key encryption. This is the technique that they have used in their paper. This paper explains about all the security issues that arise with respect to the cloud. There is no solution defined regarding any security issues. All these three papers provide the storage for keys and data,

implementation of RSA algorithm and issues in cloud respectively. AES algorithm is used. It is used to encrypt and decrypt the data. Two-factor authentication helps in providing cloud security.EmanM.Mohmed and Hatem S.Abdelkader [4] explained about data security model in their paper. This model for cloud computing is based on the study of cloud architecture. This has been improved compared to previous versions of it. The software has been developed to enhance the work in that model.

JungwooRyoo et al. [5] talked about all the challenges of cloud security in their paper. Summarization of wide range of standards of cloud security is presented in this paper. Coverage of auditing regarding cloud security is also being done here. Various standards involve different methods. Prashanth Rewagad and Yogita Pawar [6] used AES algorithm in their paper. It is being used to encrypt and decrypt the data. Diffie Hellman Key Exchange algorithm is used in this paper. It is used to exchange the secret key among admin and users. All these three papers explain about data security in the cloud and various algorithms used to access and maintain data in the cloud.

Authors Cong Wang et al. [7] proposed a method to explain the problem persisted with secure and ranked keyword search on the encrypted data in the cloud. Ranked keyword search significantly upsurges system usability by showing the matching documents in a ranked order with a certain relevance criterion like keyword frequency [7]. Authors Xiuxiu Jiang et al. [8] proposed a method which enables users to search over encrypted documents, they initially adopted a structure with name Inverted Matrix (IM) to build search index and the IM includes a number of index vectors, each of which is related to a keyword. Then locating the corresponding index vector by mapping a keyword to a value as an address. Finally, in the need of preserving privacy of users, to obtain an Encrypted Enlarged Inverted Matrix (EEIM) they mask index vectors with pseudo-random bits.

Authors Sonu Pratap Singh Gurjar, Syam Kumar Pasupuleti [9] proposed a confidentiality preserving multiple keyword ranked search scheme for encrypted data stored in the cloud. It was done considering data integrity using a data structure MIR-tree which is newly authenticated. The index based on MIR-tree with the combination of the vector space model and TF x IDF model in the construction of index and generation of query [9]. In this paper threat model, they referred to is the Known Ciphertext model and known Background model. In all the above three papers [7] [8] [9], they discussed about the model single owner and single user. Sonu Pratap Singh Gurjar et al. [9] proposed the scheme with a ranked multi-keyword search functionality that was more efficient compared with others but it was also limited to a single owner and single user.Authors Zhiguo Wan and Robert H. Deng [10] proposed the planning for a Confirmable Privacy preserving keyword Search scheme. They proposed that by combining an altered homomorphic Message Authentication Code (MAC) method with a privacy preserving multiple keyword search scheme. The projected method in this paper enables to authenticate search results efficiently by client and there is no need for storing the outsourced data as a local copy [10]. This paper discusses on multiple keyword search but doesn't discuss the ranked multiple keyword search and also it is limited to a single owner and single user [10].

Authors Wenhai Sun et al. [11] presented confirmable privacy keeping multiple keyword text search method with ranking based on similarity. In the same paper to help multi-keyword search and query output ranking, they proposed to construct the search index dependent on term recurrence and the vector space model with cosine likeness measure to accomplish higher item exactness [11]. Authors Jinguo Li et al. [12] proposed a Privacy-safeguarding Similarity Search method named PSS. They make use of "n-grams technique and counting bloom filters" to characterize and calculate the keyword order and dependent on this order, all indexing components could be sorted out in a chord-ring to help multiple cloud comparability search with high effectiveness [12].

3. GAPS IDENTIFIED

- From the background work it is identified that most of the authors worked on the single owner or single user model and the thing going to be suggested as best in this paper is multiple owner and multiple user model
- This scheme uses ranked multiple keyword search functionality which is better than the single keyword or multiple keyword search functionality without ranked scheme used in few of the early papers identified
- Advanced Encryption Standard (AES) algorithm has been suggested to use which is one of the most secured encryption methods than other algorithms used in other papers
- In order to improve searching efficiency this scheme uses Bloom Filter structure which reduces the searching time

- It is one of the first paper to support Dynamic operations like updating and deletion of documents
- Two-factor authentication has also been used in this scheme at the user end when user wants to download or decrypt the documents which is missing in many papers already published

4. AES ALGORITHM

AES acronym stands for Advanced Encryption Standard Algorithm. It is published by National Institute of Technology. AES falls into three areas. The areas include Cost, Security and Implementation. This algorithm was developed by Jaen Daemen and Vincent Rijmen, the two Belgian cryptographers. It's very much useful when one needs to decrypt a confidential encrypted format. AES is a symmetric type of algorithm. It uses the same 128, 192, or 256 bit key for both encryption and decryption of the data. The security of the system using the AES algorithm exponentially increases with its key length increases. Even with a 128-bit key, the process of cracking the AES algorithm by testing each of the 2128 possible keys (brute force attack) is impossible that even the fastest supercomputer may require on average more than 100 trillion years to crack it and AES has never been cracked. It shows resistance against various attacks like key attack, square attack, key recovery attack and differential attack. Thus, it is a highly secure encryption method. Data can also be protected against future attacks like smash attacks.

This is the most popular algorithm. It is widely used symmetric encryption algorithm. It is six times faster than triple DES. The operation of AES includes encryption and decryption process. Encryption process includes various steps. Steps included are Byte Substitution (Sub Bytes), Shift rows, Mix columns and Add round key. Decryption process also involves the same steps but in reverse order. AES is widely adopted. It is supported in both hardware and software in current day cryptography. No cryptanalytic attacks have been discovered against AES algorithm till date.

5. CONCLUSION

This paper considers a multi-users and multiple data owners model for cloud and proposes an efficient enquiry system ranked multi-keyword search on encrypted cloud data that effectively provisions dynamic activities such as insertion or deletion of a document. First, the data owners can encrypt documents and also indexes with the help of algorithm Advanced Encryption Standard (AES). AES algorithm has been used because AES algorithm successively encrypts each 128 bit block of data and the computational requirements for this approach are low and because of many more advantages as mentioned earlier in this paper. In order to obtain the Multi-Keyword ranked search, this scheme uses the vector space model including TF x IDF model and cosine similarity measure. Traditional alternatives, however, need to beer heavy computational expenses. Our Scheme presents Bloom filter to construct an enquiry index tree to accomplish the sub-linear search time. Moreover, the proposed scheme can correctly and efficiently promote dynamic operation due to the Bloom filter, which implies our systems updating price is smaller than other schemes. In this scheme Two-factor authentication has also to be used at the user end when user wants to decrypt or download documents thus increasing the security. Real-world information set tests demonstrate that the performance of the suggested scheme is satisfactory.

6. REFERENCES

[1] S.Poonkodi, V.Kavitha, K.Suresh, "Providing A Secure Data Forwarding In Cloud Storage System Using Threshold Proxy Re-Encryption Scheme", International conference on Information Systems and Computing (ISISC-2013)

[2] K.Govinda, V.Gurunathaprasad, H.Sathishkumar, "Third Party Auditing for Secure Data Storage in Cloud Through Digital Signature Using RSA", International journal of advanced scientific and technical research, Issue 2, Volume 4- August 2012

[3] Sultan Aldossary, William Allen, "Data Security, Privacy, Availability and Integrity in Cloud Computing : Issues and Current Solutions", International journal of advanced computer science and applications, Volume 7, No 4, 2016

[4] EmanM.Mohmed, Hatem S.Abdelkader, "Enhanced Data Cloud Security Model for Cloud computing", The 8 th International Conference on INFOrmatics and systems(INFOS2012)-14-16-May

[5] JungwooRyoo, Syed S Rizvi, William Alken, "Cloud Security Auditing : Challenges and Emerging Approaches", IEEE Security and Privacy Magazine, November 2014

[6] Prashanth Rewagad, Yogita Pawar, "Use of Digital Signature with Diffie Hellman Key Exchange and AES Encryption Algorithm to Enhance Data Security in Cloud Computing", International Conference on Communication Systems and Network Technologies, 2013

[7] Cong Wang, Ning Cao, Jin Li, Kui Ren, and Wenjing Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data", International Conference on Distributed Computing Systems, 2010

[8] [Xiuxiu Jiang, Jia Yu, Fanyu Kong, Xiangguo Cheng, Rong Hao "A Novel Privacy Preserving Keyword Search Scheme over Encrypted Cloud Data", 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 2015

[9] Sonu Pratap Singh Gurjar, Syam Kumar Pasupuleti "A Privacy-Preserving Multi-keyword Ranked Search Scheme over Encrypted Cloud Data using MIR-tree", 2016 International Conference on Computing, Analytics and Security Trends (CAST) College of Engineering Pune, India. Dec 19-21, 2016

[10] Zhiguo Wan and Robert H. Deng, "VPSearch: Achieving Verifiability for Privacy-Preserving Multi-Keyword Search over Encrypted Cloud Data", IEEE transactions on Dependable and Secure Computing, VOL. 15, NO. 6, November/December 2018

[11] Wenhai Sun, Bing Wang, Ning Cao, Ming Li, Wenjing Lou, Y. Thomas Hou, Hui Li, "Verifiable Privacy-Preserving Multi-Keyword Text Search in the Cloud Supporting Similarity-Based Ranking", IEEE Transactions on parallel and distributed systems, VOL. 25, NO. 11, November 2014

[12]. Jinguo Li, Mi Wen, ChunhuaGuand Hongwei Li, "PSS: Achieving High-efficiency and Privacy-preserving Similarity Search in Multiple Clouds", IEEE ICC 2016 Communication and Information Systems Security Symposium, 2016