International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 02 | Feb 2019 www.irjet.net p-ISSN: 2395-0072

Keyword Based Retrieval in Secure Cloud Storage

Nikita A. Shambharkar[1], Smita V. Tembhurne[2], Swati A. Tale[3], A. G. Waghade[4]

[1,2,3]Student, Dept. of Computer Science & Engineering, DES'S COET, Dhamangaon Rly
[4]Professor, Dept. of Computer Science & Engineering, DES'S COET, Dhamangaon Rly

---

Abstract - *Cloud Computing becomes more prevalent in data store domain. More and more sensitive to store information into the cloud, Due to the importance the protection of data privacy, to protect sensitive data, should be encrypted before outsourcing, which makes ineffective data utilization and a very challenging task. An effective search that give the better suggestion to the user so that the user can get better choices for the services is also a challenge for cloud computing. The fact that data owners and cloud server are no longer in the same trusted domain may put the outsourced unencrypted data at risk. It follows that sensitive data has to be encrypted prior to outsourcing for data privacy. With enhanced security using encryption techniques and permission to access the data, through accurate security analysis, we show that our proposed solution is secure and privacy-preserving, while correctly realizing the goal of retrieval data.*

Key Words: Keyword search, cloud computing, encryption, data retrieval.

1. INTRODUCTION

As cloud computing is emerging trend in computing the trust gain becomes very important factor. There are two parameters which can help to improve the trust on the cloud services. One is to improve efficiency and another for improving security. To improve the efficiency the keyword search technique is better as it provides two way communications between cloud server and the user. But while deploying security the burden on cloud server gets increased unexpectedly. Thus it is very important to maintain these two factors so that to improve overall efficiency of the cloud services [5]. Also the world is of mobile devices, so everyone wants to use cloud services on their mobile devices and if the computational cost goes to higher then it results into heavy resource consumption, which is not suitable for mobile devices. So there is need of efficient cloud services in the future. Effective information management and retrieval within an enterprise or over internet to a large extent depends heavily on the organizing, searching, browsing and navigating facilities built in to information management systems and their intuitiveness and usability. Information search and document retrieval from a remote database requires submitting the search terms to the database holders. However search terms may contain sensitive information that must be kept secret from the database holder. Moreover the privacy concern is required to apply to the documents retrieved by the user in the later stage because they may contain sensitive information about search terms. Therefore in this paper we will study various secure and fast information retrieval methods by using user query keywords. As a kind of emerging business computational prototype, Cloud Computing distributes computation task on the resource pool which consists of a large number of computers and accordingly the application systems gain the computation working strength, the storage space and software service according to its demand. The working of cloud computing can be viewed by two distinctive features One is the cloud infrastructure which is the building block for the upper layer cloud application. The other is the cloud application. Cloud computing has achieved two important goals for the distributed computing by the means of three technical methods. High Scalability the cloud infrastructure can be expanded to very large scale even to thousands of servers and high Availability so that the services are available even when quite a number of servers fail.

There are a lot of developing countries is an important factor in preparing national development plans. If you use Google documents, there is no need to worry about buying licenses for word processing programs or keep them up to date. There is no reason to worry about viruses that may affect your computer or about backing up files that you create. Google will do it all for you. Basic cloud computing services are the results that you do not need to worry about how to provide the service you are buying. With services on the Internet, the focus will be on every job and leave the problem of providing computing. Available upon request often bought or "pay-as-you-go" cloud services subscription. You can buy cloud computing in the same way that you want to buy Phone services or access to the Internet from the utility company. Sometimes cloud computing is free or paid for other ways. Just like telephone services, you can buy a cloud computing service as you need from one day to the next.

2. RELATED WORK

We present a general methodology for constructing privacy-assured searchable schemes based on several building blocks, including recently developed efficient symmetric-key encryption primitives (e.g., symmetric searchable encryption [SSE]). For each of the proposed usable search functionalities, we survey recent research advances, and give insights on the advantages and limitations of each approach. Ending with a set of future challenges, this article intends to bring attention to and motivate further research on enabling privacy-assured searchable cloud storage a reality. Secure top-k retrieval from Database Community from database community are the most related work to our proposed RSSE. Theidea of uniformly distributing posting elements using an order-preserving cryptographic function. However, the order-

preserving mapping function proposed does not support score dynamics, i.e., any insertion and updates of the scores in the index will result in the posting list completely rebuilt. Zerr-et-al. use a different order-preserving mapping based on pre-sampling and training of the relevance scores to be outsourced, which is not as efficient as our proposed schemes. Besides, when scores following different distributions need to be inserted, their score transformation function still needs to be rebuilt.

In fuzzy keyword search [1] a trapdoor request is generated to search for the files. It uses three steps as setup, encryption and decryption. In setup phase the index is created from the files and then the files in encrypted formats sent to the database. When user wants to retrieve the file, a trapdoor request is generated to search for the specific file and then the keyword is matched with the index entries from the database and set of matching file entries are sent to that user. This scheme works well but it also compromises security as the index is sent to the server is not in encrypted format and only files are sent in encrypted format. Thus it is not the secure method as it does not guarantee the privacy of the user. In ranked searchable symmetric encryption scheme a trapdoor request is generated to search the files. This keyword searches files through the index generated from the files and searches for the similar matching files through the database. The files stored on database are encrypted by using SHA-1 160 bit, and then these files were sent to the database along with the index files. Here the index files are not encrypted. This scheme gives better results for the single keyword search but the security is again compromised by letting the index files in non-encrypted format on the database. In secure rank-ordered multi-keyword search, both the index and files are encrypted before sending to the database. The security is well guaranteed here. The index is encrypted with SHA-1 160 bits while the files are encrypted with the AES algorithm containing CTS functions. The index created for the particular file takes much more time to build and the size of index file is much higher yet it doesn't support the misspelled words. Thus the overall efficiency of the scheme decreased with improved load on the system to handle and search through this large set of index files.

In [2] proposed an idea of hierarchical attribute based encryption. Cipher Policy Hierarchical Attribute Based Encryption scheme, attributes of the user are arranged in a matrix format where users with high level attributes can grant the access rights to the lower level users. This scheme includes multiple users from various organizations. In [6] proposed an idea of searchable encryption. Searchable encryption is a well-organized technique for data retrieval which uses attribute-based encryption. Various facets include Privacy of information: High level privacy is assured. None of them can access communication information about data content such as response, query and also regarding the ciphertext. Data holder's Privacy: True Identification of the data owner cannot be found from the encrypted data. End user's Privacy: Original Identities of the receivers cannot be acquired from the encrypted content

3. SYSTEM OVERVIEW

In Our application whenever we are storing any documents to cloud we need to secure it because there may be a malicious or third party attackers can hack data which in stored in cloud so that we need to create dynamic secure storage system. Our application represents file upload, file download and search keyword process. First text file and grade is selected from the local system. From the file all the unnecessary words, special characters and whitespaces are removed. The remaining keywords are extracted from the file. Remaining keywords occurrence is checked how many times the keywords are repeated. The repeated keywords weightage is noted in index array table. After weightage calculation, keyword ranking is checked. Keyword ranking is the range of word frequency divided by the entire number of words in particular files (Term frequency). Finally keyword ranking is stored in index array. After completing this process all the keywords are converted into hashed index with the grade access control key using MD5 algorithm. This converted hash key is inserted into index array using sql queries. After keywords extraction process; file is uploaded to cloud by giving access permission. After that encrypted file is compressed and uploaded to the hybrid cloud by using the cloud account details. User gives the keyword, that keywords will be converted into hash code using md5 algorithm, the converted hash code is checked with server generated hash code, if it is matched files retrieved from the cloud, decompressed then decrypted using Decryption. Finally original file is downloaded to local system.

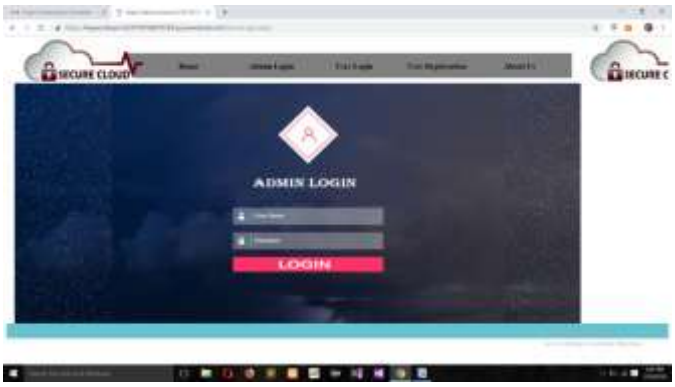

Figure 1 Admin Login Page

Figure 2 User File Upload

Figure 3 User Search File using Keyword

Figure 4 Search Result

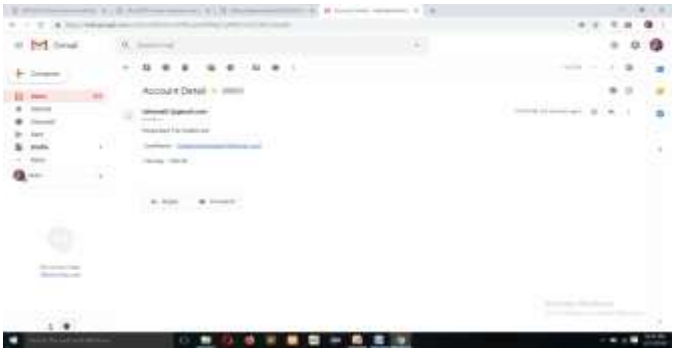

Figure 5 Key Send on Registered Mail Id

Figure 6 Download File

4. CONCLUSION

Efforts have been made to address the issues of storing the data on cloud and the drawbacks or problems that follow it that is security. Noteworthy progress has been made in 3 main directions, mainly dealing with query effectiveness, security and efficiency. This system developed in dot net architecture implemented and tested in Microsoft azure cloud platform. The experimental results show the system meet the all designed constraints which are discussed in system analysis. The system automatically extract keywords from uploading file and the keywords are converted into hash key with respective access control while users are searching for a file by providing keywords based on users grade hash key will be generated as well as searched and corresponding files are retrieved for security and efficiency.

REFERENCES

[1] Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute-based encryption. IEEE Symposium on Security and Privacy, 2007 SP'07. 2007; IEEE.

[2] Liu Q, Wang G, Wu J. Time-based proxy re-encryption scheme for secure data sharing in a cloud environment. Information Sciences. 2014; 258:355–70.

[3] Hur J, Koo D, Hwang SO, Kang K. Removing escrow from ciphertext policy attribute-based encryption. Computers and Mathematics with Applications. 2013; 65(9):1310–7.

[4] Ren K, Wang C, Wang Q. Security challenges for the public cloud. IEEE Internet Computing. 2012 Jan 1;16(1):69.

[5] Awad A, Matthews A, Lee B. Secure cloud storage and search scheme for mobile devices. InMELECON 2014-2014 17th IEEE Mediterranean Electrotechnical Conference 2014 Apr 13 (pp. 144- 150). IEEE.

[6] Abir Awad, Adrian Matthews, Yuansong Qiao, Brian Lee," Chaotic Searchable Encryption for Mobile Cloud Storage", IEEE Transactions on Cloud Computing,, no. 1, pp. 1, July2015.

[7] Xiong AP, Gan QX, He XX, Zhao Q. A searchable encryption of CPABE scheme in cloud storage. In Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2013 10th International.

[8] Zhang B, Zhang F. An efficient public key encryption with conjunctive-subset keywords search. Journal of Network and Computer Applications. 2011 Jan 31;34(1):262-7.