Secure Cloud Storage with Multiple Keyword Search Function

Pankaj Vasantrao Manjare[1], Zeeshan I. Khan[2]
[1]*Student of Master of Engineering in (CSE), DRGIT&R, Amravati, Maharashtra*
[2] *Assistant Professor (CSE), DRGIT&R, Amravati, Maharashtra*

***Abstract*- The amount of data and the user's demands of accessing these data both are increasing rapidly. As the cloud computing environment provides great flexibility and availability for the data as and when required, data owners are motivated to outsource their large amount of documents from local system to commercial public cloud. When users wants particular document from the cloud servers they gets readily available, but at the same time this kind of computing model brings challenges to the security and privacy of data stored in cloud servers. So to maintain this privacy this project proposes model with Attribute-based encryption (ABE) technology to design fine-grained access control system, which provides one good method to solve the security issues in cloud setting. Outsourced ABE (OABE) with fine-grained access control system can largely reduce the computation cost for users who want to access encrypted data stored in cloud by outsourcing the heavy computation to cloud service provider (CSP). As their are large number of files, this project present a new cryptographic primitive called attribute-based encryption scheme with outsourcing key-issuing and outsourcing decryption, which can implement keyword search function (KSF-OABE). Instead now a days, there are large number of data users and also large number of documents in cloud, makes it crucial for the search service based on providing single keyword. This demands multiple keyword search query and provide result similarity ranking to meet the effective data retrieval needs. With this demands the proposed system solve the challenging problem of multiple keyword searching over cloud data. Here, the proposed and develop system takes multiple keywords for searching documents stored in cloud in encrypted format along with outsourcing key-issuing facility by third party auditor(TPA), that achieves privacy preservation and gives the matched document properly as a result.**

***IndexTerms*-Attribute-based encryption (ABE), cloud service provider (CSP), Outsourced ABE (OABE), Outsourced Attribute-based encryption with keyword search function (KSF-OABE), Trusted Authority (TA).**

I. INTRODUCTION

In today's information technology with increasing use of internet, clients that need high warehousing and processing power have a tendency to outsource their information, sensitive data and administrations to clouds. Clouds empower clients to remotely store and access their information by bringing down expense of hardware possession while giving strong and quick administrations [1]. The significance and need for the protection for saving pursuit procedures are significantly more claimed in the cloud applications. As the data stored in cloud computing environment is in much large amount, searching from that much data is bit critical [2]. The query that comes in our mind from here is, given an arrangement of keywords, how would we use for secure storage of information in the cloud and how to give the access permission?

Cloud computing is a new computation model in which computing resources is regarded as service to provide computing operations. This kind of computing paradigm enables us to obtain and release computing resources rapidly. So we can access resource-rich, various, and convenient computing resources on demand [3]. The computing paradigm also brings some challenges to the security and privacy of data when a user outsources sensitive data to cloud servers. Many applications use complex access control mechanisms to protect encrypted sensitive information. Sahai and Waters [4] addressed this problem by introducing the concept for ABE. This kind of new public-key cryptographic primitive enables us to implement access control over encrypted files by utilizing access policies associated with ciphertext or private keys. Two types of ABE schemes, namely key-policy ABE (KP-ABE) and ciphertext-policy ABE (CP-ABE) are proposed in this system. And for access the proposed system

concentrates on the arrangement of multi-keyword search over data that is mainly stored in the cloud [5]. For KP-ABE scheme, each ciphertext is related to a set of attributes, and each user's private key is associated with an access policy for attributes. A user is able to decrypt a ciphertext if and only if the attribute set related to the ciphertext satisfies the access policy associated with the user's private key. For CP-ABE scheme, the roles of an attribute set and an access policy are reversed.

To actuate the search for the multiple keyword for successfully using of outsourcing cloud information under the previously stated model, this framework try to proposed by considering the security contemplations too. The framework is required to give the accompanying security with ABE and execution ensures as for Multiple keyword Search: To plan search which permit multiple keyword query and give result similarity ranking to authoritative information recovery [6]. Privacy-Preserving: To keep the cloud server from taking in extra data from the dataset and the record, and to meet the essential security necessities.

The system proposed and develop here contains the two more mechanism that helps to perform the task more properly. The Security providing and authentication is done by Third party auditor (TPA), who always performs these task efficiently and after this security checks the user of data or owner of the data is able to uploading and downloading of data in public cloud environment. Along with this the records of large amount of data stored on cloud is available to cloud service provider which also performs the task of proper and secure storage. The remaining paper organized as, section II performs the study of literature and related work done by different authors. Section III contains the architecture of the system that shows storing and accessing of information in secure manner. Section IV shows different techniques proposed here for applying security to cloud storage and access with multiple keyword search function. Finally, section V concludes the paper.

II. LITERATURE REVIEW & RELATED WORK

Bethencourt et al. [7] provided a CP-ABE scheme, ensuring confidentiality of encrypted data stored even if the storage server is untrusted. In order to withstand collusion attack and avoid sensitive information leakage from access structure. Deng et al. constructed a ciphertext-policy hierarchical attribute based encryption (CP-HABE) with short ciphertext, which enables a CP-HABE system to host many users from different organizations by delegating keys. In CPABE scheme, a malicious user maybe shares his attributes with other users, which might leak his decryption privilege as a decryption black box due to financial profits.

The author in Cryptographic Cloud Storage paper [8] said that when the benefits of using a public cloud infrastructure are clear, it introduces significant security and privacy risks. In fact, it seems that the biggest obstacle to the adoption of cloud storage is concern over the confidentiality and integrity of data. In [8], an overview of the benefits of a cryptographic storage service, for example, reducing the legal exposure of both customers and cloud providers, and achieving regulatory compliance is provided. Besides this, cloud services that could be built on top of a cryptographic storage service such as secure backups, archival, health record systems, secure data exchange and e-discovery is stated briefly.

The author in this paper [9], proposed the Division and replication of data in cloud with Attribute based encryption. In the proposed system, a cloud storage security scheme that collectively deals with the security and performance in terms of retrieval time. The data file was fragmented and the fragments are dispersed over multiple nodes. And CP-ABE scheme that provides outsourcing key-issuing, decryption and keyword search function. This scheme is efficient since it need to download the partial decryption ciphertext corresponding to a specific keyword. In the proposed scheme, the time-consuming pairing operation can be outsourced to the cloud service provider, while the slight operations can be done by users. Thus, the computation cost at both users and trusted authority sides is minimized, with help of trapdoor provider work is reduces.

Organizations, companies store more and more valuable information is on cloud to protect their data from virus, hacking. The benefits of the new computing model include but are not limited to: relief of the trouble for storage administration, data access, and avoidance of high expenditure on hardware mechanism, software, etc [10]. Ranked search improves system usability by normal matching files

in a ranked order regarding to certain relevance criteria mainly based on keyword frequency.

The author in paper [10], propose a semantic multi-keyword ranked search scheme over the encrypted cloud data, which simultaneously meets a set of strict privacy requirements. Firstly, we utilize the "Latent Semantic Analysis" to reveal relationship between terms and documents. The relationship between terms is automatically captured. The proposed scheme could return not only the exact matching files, but also the files including the terms latent semantically associated to the query keyword. Here the proposed system meanwhile supports latent semantic search and uses the vectors consisting of TF values as indexes to documents. These vectors constitute a matrix, from which we analyze the latent semantic association between terms and documents by LSA. Taking security and privacy into consideration, author employ a secure splitting k-NN technique to encrypt the index and the queried vector, so that we can obtain the accurate ranked results and protect the confidence of the data well.

The author in [11], proposed a file-level authorized private keyword search (APKS) scheme over encrypted cloud data. However, it incurs additional communication cost, since whenever users want to search, they have to resort to the attribute authority to acquire the search capabilities. Moreover, this scheme is more suitable for the structured database that contains only limited number of keywords. The search time there is proportional to the total number of keywords in the system, which would be inefficient for arbitrarily-structured data search, e.g., free text search, in the case of dynamic file sharing system.

The author in this paper [12], proposes a different yet more challenging scenario where the outsourced dataset can be contributed from multiple owners and are searchable by multiple users, i.e. multi-user multi-contributor case. Inspired by attribute-based encryption (ABE), we present the first attribute-based keyword search scheme with efficient user revocation (ABKS-UR) that enables scalable fine-grained (i.e. file-level) search authorization. Compared with existing public key authorized keyword search scheme, this scheme could achieve system scalability and fine-grained ness at the same time. Different from search scheme with predicate encryption, this scheme enables a flexible authorized keyword search over arbitrarily-structured data.

III. ARCHITECTURE OF SYSTEM

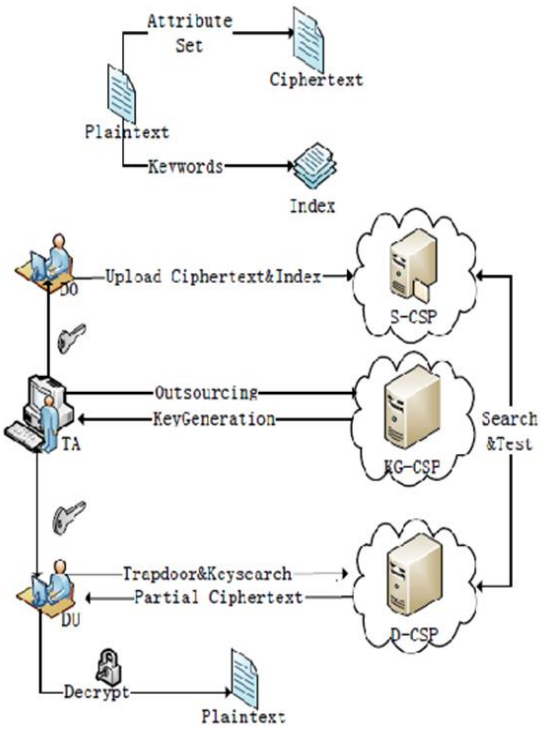

Figure 1: System Architecture

The architecture of the proposed system is shown in Figure 1 above. The architecture consist of following participants along with their functionality.

Data Owner (DO): This is a participant of our proposed system who intends to upload, outsource and share his data files on the cloud storage system in a secure way. For the mechanism of sharing, the encrypted ciphertext will be shared with intended receivers whose access structure will be satisfied by attribute set embedded in ciphertext. The responsibility of data owner is to generate indexes for some keywords and upload encrypted data with the indexes.

Trusted Authority (TA):

Trusted Authority is work as authority center, which is responsible for the authentication of system users. It also performs the task of providing attribute private keys to only authorized persons based on their demands.

Cloud Service Provider (CSP): The cloud service provider is one of the main component of the cloud

computing environment. CSP performs the task of outsourcing computing service for Trusted Authority and Users. The different task performs by CSP are as follows

KG_CSP: Could Service Provider perform the task of key generation (KG) and outsourced it, which is allocated by TA. Which is one of the most helpful task for TA done by CSP.

S-CSP: It is a participant that supplies outsourcing data storage service for users who want to share their files in cloud. The storage activity is done with the help of TA and by KG-CSP functionality.

D-CSP: It is a participant that supplies outsourcing computing service through accomplishing partial decryption for ciphertext and keyword search service. It is done on the partially decrypted ciphertext for data users who want to access the Data from the file. This activity is also done with the help of TA who checks the secret key of the file should be made available to the user or not.

Data User (DU): This is a participant who access the data by performing decryption of the encrypted data stored in S-CSP with the help of D-CSP. If the attribute set given by data user satisfies the access structures, DU is able to access the encrypted files and recover the original files from it. DU can also downloads intended data files with the help of getting key from TA of that file and with appointed keyword. Data user can choose by providing multiple keywords to get decrypting data and files.

IV. IMPORTANT TECHNIQUES

Following are some important techniques that are used in the proposed system to perform secure cloud storage and access with multiple keywords.

A. Attribute Based Encryption

The attribute based encryption performs fine-grained access control. With this property ABE is used for Encryption operation in cloud environment. The first key policy attribute-based encryption (KP-ABE) scheme was invented, where ciphertext can be decrypted only if the attributes that are used for encryption satisfy the access structure on the user private key [13]. The second policy for the reverse situation, CP-ABE allows user private key to be associated with a set of attributes and ciphertext associated with an access structure. For the broadcast environment CP-ABE is a preferred choice when designing an access control mechanism. Since the first construction of CP-ABE, many works have been proposed for more expressive, flexible and practical versions of this technique. Cheung et al. proposed a selectively secure CP-ABE construction in the standard model using the simple Boolean function, i.e. AND gate. By adopting proxy re-encryption and lazy re-encryption techniques, Yu et al. also devised a selectively secure CP-ABE scheme with the ability of attribute revocation, which is perfectly suitable for the data-outsourced cloud model.

B. Secure Data Storage in Cloud Environment

The concept of "Cloud storage" can be defined as data storage that is made available as a service via a network [14]. Advances in networking technology and for the need of computing resources have prompted many organizations to outsource their storage and computing services. Now cloud storage is not remains as new technology, as experts have already developed some best practices for getting the most from moving to cloud storage. This new economic and computing model is commonly referred to as *cloud computing* and includes various types of services that are important, serves as base and mostly known to us as: infrastructure as a service (IaaS), where a customer makes use of a service provider's computing, storage infrastructure; platform as a service (PaaS), where a customer leverages the provider's resources that run some custom applications; and finally software as a service (SaaS), where customers use software that is run on the provider's infrastructure By moving the data of the users into the cloud, customers can lessen the costs of building and also maintaining a private storage infrastructure, and have to pay according to their use. For the customers, it provides several benefits including 'availability' of data that is being able to access data without the concern of location and 'reliability' of our data for not having to worry about backups at a relatively low cost. CSP demands some basic cost per gigabyte of cloud storage. That helps to figure out how much it will cost you per month depending on amount of data user need to store.

C. Cryptographic Approach

The cryptography is serves as the important tools for providing security to the data in the cloud computing platform. In this technique, the information is protected by transforming it or encrypting it which is

not readable by humans called as cipher text. An important aspect of this cryptographic storage in the cloud service is that the security properties described above are achieved that is based on some strong cryptographic guarantees but opposed to legal, physical and access control mechanisms. Only those users who have a secret key can decrypt the encrypted message or download files [15]. As the electronic communication become more prevalent it serves as increasingly important need.

Cryptography is used to protect most important and personal information which customer stored in the cloud environment while uploading and downloading as for both the time, when the data present in the network it will be in the encrypted format. Cryptography system can be classified into two main categories as symmetric-key system and public-key or asymmetric-key system. In the symmetric-key system, one single key is used by both sender and receiver. And in public-key system, as the name suggest the public key is known to everyone and private key to only the recipient users. For our purpose the asymmetric cryptography technique is most suitable for encrypting the customer's information which have great importance from different application point of view. The system uses the ABE algorithm for providing the security to our data as explain above. Asymmetric key cryptography is a class of cryptographic algorithms, mainly require two different keys one is secret or private and other is public key according to their different uses. The public key is used to encrypt the plaintext and private key is used to decrypt the cipher text to plain text. The public key is published without compromising security but private key used only after authentication and authorization phase only by the authorized people.

D. Multiple Keyword Search

Among various multi-keyword semantics, we choose the efficient principle of "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents. Specifically, we use "inner product similarity", i.e., the number of query keywords appearing related to the document, to quantitatively evaluate the similarity of that document to the search query in "coordinate matching" principle. The search query is also described as a binary vector where each bit means whether corresponding keyword appears in this search request, so the similarity could be exactly measured by inner product of query vector with data vector. To design search schemes which allow multiple keyword query and provide result similarity ranking for effective data retrieval, instead of returning undifferentiated results.

V. CONCLUSION

In this paper, the mechanism is given for secure storage of large amount of data in cloud computing environment. For applying security different schemes of Attribute based encryption mechanism are used. Here the CP-ABE scheme that provides outsourcing key-issuing, decryption and keyword search function is proposed. The proposed scheme here is efficient since we only need to download the partial decryption ciphertext corresponding to a specific keyword. In our scheme, the time-consuming pairing operation can be outsourced to the cloud service provider, while the slight operations can be done by users. Thus, the computation cost at both users and trusted authority sides is minimized. Furthermore, the proposed scheme of multiple keywords search greatly improve communication efficiency and access guarantee for the user instead he knows less about the files from the cloud environment. The complete scheme helps to protect the privacy of users and helps easy access to the files stored in cloud environment.

REFERENCES

[1] Jiguo Li, Xiaonan Lin, Yichen Zhang and Jinguang Han, "KSF-OABE: Outsourced Attribute-Based Encryption with Keyword Search Function for Cloud Storage", IEEE Transactions on Services Computing, (Volume:PP , Issue: 99),16 March 2016.

[2] Ning Caoy, Cong Wang, Ming Li, Kui Ren, and Wenjing Lou, "Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data", Department of ECE, Illinois Institute of Technology.

[3] S. Pearson, Y. Shen and M. Mowbray, "A Privacy Manager for Cloud Computing,"Proc.First International ConferenceCloud Computing (CloudCom'09), M. Gilje-Jaatun, G. Zhao and C. Rong, eds.,

LNCS 5931, Berlin: Springer-Verlag, pp. 90-106, 2009.

[4] A. Sahai and B. Waters,"Fuzzy Identity-Based Encryption,"EUROCRYPT'05, LNCS, vol. 3494, pp. 457-473, 2005.

[5] Li Chen, Xingming Sun, Zhihua Xiaand Qi Liu, "An Efficient and Privacy-Preserving Semantic Multi-Keyword Ranked Search over Encrypted Cloud Data", International Journal of Security and Its Applications, Vol.8, No.2 (2014).

[6] D. Boneh, "Public key encryption with keyword search", Advances in Cryptology-Eurocrypt 2004, Springer, (2004).

[7] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy Attribute-Based Encryption," Proc. IEEE Symp. Security and Privacy (SP '07), pp. 321-334, May. 2007, doi:10.1109/SP.2007.11.

[8] J.T.Ning, Z.F.Cao, X.L.Dong, L.F. Wei and X.D.Lin, "Large Universe Ciphertext-Policy Attribute-Based Encryption with White-Box Traceability," ESORICS'14, LNCS 8713, Berlin: Springer-Verlag, pp. 55-72, 2014.

[9] Kollipara Anil Kumar, Shaik Akbarm, "Outsourced Attribute Based Encryption (OABE) Watchword Search Function for Cloud Computing", International Journal for Modern Trends in Science and Technology, Volume: 03, Issue No: 09, September 2017, ISSN: 2455-3778.

[10] Mikhail Strizhov and Indrajit Ray, "Multi-keyword Similarity Search Over Encrypted Cloud Data", Colorado State University, Fort Collins, CO, 80523, USA.

[11] M. Li, S. Yu, N. Cao, and W. Lou, "Authorized private keyword search over encrypted data in cloud computing," in Proc. of ICDCS. IEEE, 2011, pp. 383–392.

[12] Wenhai Sun, Shucheng Yu, Wenjing Lou, Y. Thomas Hou, and Hui Li, "Protecting Your Right: Attribute-based Keyword Search with Fine-grained Owner-enforced Search Authorization in the Cloud", University of Arkansas at Little Rock, Little Rock, AR, USA, IEEE, 2014.

[13] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in Proc. of CCS. ACM, 2006, pp. 89–98.

[14] T. Moataz and A. Shikfa. Boolean symmetric searchable encryption. In Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, Hangzhou, China, May 2013.

[15] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.