ISSN:0975-9646

Secured Sharing of Medical Data on Hybrid Cloud Environment

M.Nandhini[#1], K.A.Muthukumar[*2]

[#] *School of engineering and technology-*
Department of computer sciences, Pondicherry University
Pondicherry-12, India

Abstract— **Cloud computing is emerging as a new computing paradigm. It provides lots of service like Software as a Service, Infrastructure as a Service, and Platform as a Service also it provides storing and sharing of data. Sharing of data in cloud platform among multiple applications to be done in secured manner. Many traditional approaches available for sharing of data in secured way but still with some data leakage issues. In this paper, a new approach has been introduced for secured storing and sharing of medical data on Hybrid cloud environment. The data partition technique, encryption and decryption techniques have to be performed on the medical data to store and share them in a secure way. The performance of proposed method is evaluated by comparing them with the information dispersal algorithm.**

Keywords— **cloud computing, storing and sharing algorithms, hybrid cloud data sharing, medical data sharing, secured data sharing..**

I. INTRODUCTION

Cloud Computing [1] is a rapidly growing technology which includes several types of services offered over the Internet. With cloud computing, application deployed can be scaled at any time without having to physically add any sort of hardware, thus the strategy to adopt cloud computing will help the organization to focus on the core activity with much less hassle but greater effectiveness and efficiency. A large number of applications running in multiple virtual machines can be managed with a cloud configured environment, thus reducing the time spent by system administrators and technicians to maintain and monitor data centers. Cloud is a collection of technologies like virtual servers, disk space and networking equipment hosted as Infrastructure as a Service (IaaS), followed by Google App Engine, Windows Azure & Amazon EC2 hosted as Platform as a Service (PaaS).

Hybrid cloud environment [2] composed of two or more clouds, that cloud will be any of private public and community cloud. It offers greater benefits of multiple deployment models. This environment is very flexible and scalable for various services offered by cloud. Because of different cloud environment bounded together, security will be the burning issue during the data sharing. Now-a-days hybrid cloud is starting to use in all industries.

In healthcare industry cloud computing play an important role. Large numbers of health organizations have started shifting to the cloud environment. Introducing the cloud services in the health sector not only facilitates the exchange of medical data among the hospitals and clinics, but also enables the cloud to act as a medical record storage center. Moreover, shifting to the cloud environment relieves the healthcare organizations of the tedious tasks of infrastructure management [3] and also minimizes development and maintenance costs. Nonetheless, storing the patient health data in the third-party servers also entails serious threats to data privacy. Because of probable disclosure of medical records stored and exchanged in the cloud, the patient's privacy concerns should essentially be considered when designing the security and privacy mechanisms. Various approaches have been used to preserve the privacy of the health information in the cloud environments.

In hospital the patient record [4] is very sensitive it should be taken care very secretly. The cloud computing environment to maintain the patient healthcare records is very useful for all doctors, patients, pharmaceuticals, researchers, insurance companies. The people involved in healthcare industry don't need all the information about the patient. It requires only particular details enough for people. For that the patient data shared among the people should be very secure. For example, an insurance company needs only treatment proofs, research needs only treatment details etc.

Hajji et al. [5] Homomorphic algorithm, Rabin et al.[6] Information Dispersal algorithm are widely used technique to secure sharing of data in a cloud environment. Homomorphic algorithm converts the data into cipher text and performs a computational operation. The information dispersal algorithm divides the data and stored in the matrix and performs computational operation. It is not suitable for healthcare industry security policy. These approaches are inefficient due to key management, computational complexity, expensiveness and communication overhead. Therefore a new algorithm is proposed based on data dividing technique to overcome the problems of existing technique. The proposed algorithm shares the data as per the user request; which is more secure and flexible. The whole patient data are divided and encrypted in proposed algorithm. The encrypted data stored in the data center. The user sends a request to the data center. It processes the request and generates the key which is used to decrypt the data. The data center response with the key to the user request. Thus the data is shared between the user and the data center in a secured way. The proposed method reduces

the communication overhead, space complexity and its more flexible to the server to share the data.

The rest of this paper is organized as follows. Section II presented the existing methods and its performance. The Section III introduces the proposed method in detail along with encryption and decryption technique. The Section IV to analyze the performance of the proposed approach with an information dispersal algorithm for secure data sharing. Section V concludes the paper with future work. This document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

II. LITERATURE SURVEY

Various approaches proposed to share the data in the cloud environment based on the access control mechanism, user privacy, key distribution, scalable sharing of data, attribute used in data, proxy sharing. In this section different methods and its performance are presented in detail.

A. Scalable, secure file sharing:

Kallahalla et al. [7] Proposed secure file sharing method. It is one of the cryptographic storage systems. It helps us to store and share the data on untrusted cloud server. It allows the client to handle all the key management. It provides a very limited source to the server because the server is interested. The secure file sharing method divides the files into groups and gives to the data owner to share the file group with unique file back key to protect the data. There are some drawbacks of this model. Large key distribution takes place for large scale file shortly. Each and every time the unique file blocks key need to update and distribute again. The advantage is it provides end to end security for the group strong system.

B. Secure remote untrusted storage:

E. Gosh et al. [8] is designed for multiuser to handle the file in an untrusted network. It is developed using any cryptographic operations like read write access control on file sharing. These have been used for large group file sharing by using NNC key. Key management and permission to the user is simple by the use of out band communication. It provides secure network file system without distributing the file server. The drawback of this method is to provide permission to the dynamic group's private key of every group need to be updated whenever new group is joined.

C. Proxy Re-Encrypted scheme:

Atensie et al. [9] Proposed proxy re encrypted method for access control over the file system and storage of data. The content of the data has been encrypted with a symmetric key of the data owner. That key is encrypted by master public key. The symmetric key and master key is again encrypted by using proxy cryptographic which is useful for access control mechanism. To manage all these encryption, content is stored in the centralized control server. The main advantage of this method is only a limited amount of trust is on the server. There may be a collision occurs the attacks or server try to find the decryption key.

D. Fine grained data access control:

S. Yuwang et al. [10] design this method by combining the Attribute Based Encryption (ABE) and Proxy Re-encryption (PRE) [9] to allow the data owner to perform computation tasks on the server without revealing the content of the data. The files are encrypted with a random key that random key is further encrypted by Key Policy Attribute Based Encrypted (KP-ABE) with a set of attributes. The key group manager handles these keys only authorized user with the key can decrypt the cipher text. The limitation of this method is a concept is not supported and secretes key need to update after every approval.

E. Secure provenance

Lu et al. [11] proposed this method to record the ownership and history of access data. This method is based on bilinear pairing technique and Cipher text Policy Attribute Based Encryption technique (CP-ABE). The basic feature of this method is to allow the user to anonyms' authentication to access the files and document served in the server and then tracks the ownership and identity of the user. The system consists of a single attribute. The user gets two keys after register into the system. The user can encrypt the data by Attribute Based Encryption (ABE). For the decryption attribute key will be used by the further user to preserve the security & traceability of the user the encrypted user need to sign the data using a group signature key.

F. CP-ABE based cryptographic method:

Yong Cheng et al. [12] proposed this scheme to secure, store and share the data in cryptographic method. It uses basic Graphical encryption & a decryption method for security and data confidentiality. Using this there will be some issues. The first issues are to be data owner have to distribute the key to everyone if he wants to share the data. The second issue of access control is very expensive the data owner has to encrypt again the data and republish. It is very tedious processed. The first problem can be overcome by CP-ABE algorithm [12]. To manage the access control the data have to be divided into small parts and stored. When the data access is high data owner have to encrypt again and republish.

G. Knox privacy preserving auditing method:

Wany et al. [13] proposed this method to store and share the cloud data among the large number of users. They developed privacy preserving auditing mechanism to store and share the data. The group signature is constructed over the data by homomorphic authenticator using third party auditor to verify the integrity. The signature identity of the data is kept secret from the third party auditor. The data owner can add new user to be group disclose identity to the new user. The time taken for verification and auditing for an independent number of users in the group.

H. Broadcast encryption:

Fiat et al. [14] is developed for secure transmit the data from the user and the new method is introduced qualitative and quantitative assessment encryption is used to broadcast the data with minimum key management. The broadcast encryption transmits the data to all memory with privileged subset. The new user is added to the group existing user need to decide and great the privileged access to the new user.

I. Collusion secure method:

Collusion secure method [15] is high security is a standard model. The new user can join the group at any time without any modification of the decryption key. The new user has great access without any modification to be already existing keys. This method is optimal bound for cipher text or a decryption key. It also improves efficiency in private key settings.

J. Homomorphic algorithm:

Homomorphic algorithm [16] is the conversion of data into cipher text that can be analysed and worked with as if it were still in its original form. It allows complex mathematical operations to be performed on encrypted data without compromising the encryption. In mathematics, homomorphic describes the transformation of one data set into another while preserving relationships between elements in both sets. It is one of the promising techniques used in many cloud applications to ensure the security. It shows more efficient and accurate result. The homomorphic encryption is a more computationally complex to perform operation. It totally depends on public key cryptographic system, difficult to implement.

K. Private information retrieval

The private information retrieval [17] is one secure method in the database operation. The main aim of this method is to provide security for the user operation. It hides all the user operations in the database. It provides privacy for the user operation of the service provider. A lot of private information retrieval technique has been developed on the basis of user privacy concern. Symmetric private information retrieval, this method provides high security to the user data. It is totally infeasible on single server computation.

L. Information dispersal algorithm

The Information dispersal algorithm is one of the data sharing technique [18] widely used in many network. Divide the data D into n pieces only the whole data can be retrieved k<n, k is the threshold value. Each split data i⩽n is of a size |D|/k where |D| is the size of the data. The total size of all split data will be (n/k) ×|s|. It is one of data sharing methods it will overcome storage complexity, but it will form a frequent pattern of data so the intruders can attack the data easily. It is vulnerable to attack.

TABLE I
COMPARISON OF DATA SHARING ALGORITHMS

S. No	***Method***	***Merits***	***Demerits***
1	Scalable file sharing [7]	End to end security	Heavy key distribution
2	Secure remote sharing [8]	Read, write access control for the group user	Access control for dynamic group member
3	Proxy Re-encryption [9]	Full control access to the data owner	Too much of the encryption key used
4	Fine gradient data access control [10]	The combination of ABE, proxies, encryption and KP-ABE algorithm	It is not flexible for the data owner
5	Secure provenance [11]	Confidentiality of the data	User access control permission change is not allowed
6	CP-ABE cryptography cloud storage [12]	Confidentiality and security	Too much operation performed
7	Knox [13]	Monitoring the user and their operation	The third party auditor is not trusted
8	Broadcast encryption [14]	Secure data transmission with encryption	Chance of collusion between the data
9	Collusion secure broadcast encryption [15]	High security collision never occurs	Difficult to implement
10	Homomorphic algorithm [16]	Suitable for all cloud applications	Computationally complex
11	Private information retrieval [17]	Privacy preserving of user operation	Space complexity
12	Information dispersal algorithm [18]	Reduced space complexity	Vulnerable to attack

III. HYBRID CLOUD DATA SHARING MODEL

The Hybrid cloud environment consists of two or more public private clouds connected together. Each cloud has their own security policy so the privacy of the data is very important. For data sharing in a hybrid cloud environment a common data center is used to manage all the data through this data can be shared among the different cloud user. In the health care industry all hospitals, research, development centers [19] all connected together to form a hybrid cloud environment. The user from one cloud need to access the data of another cloud need to send requests to the common data center. Based on the request data provided to the user. The Data center has a large database [20] it contains all cloud data in the network It include patient's details,

disease details, treatment details, medicine details. Data center doesn't share the whole data with the cloud user it shares only required data for a user for example, researchers need only disease details and treatment details others details are kept secure.

Data center process the user request query in the database extracts the data and encrypt it. Before processing the query [21] in the database user need to be authenticated [22] like a researcher or doctor, etc. Using the cloud id. After authenticating the request query will start to process encryption operation on the processed data. During the encryption the threshold value is assigned randomly to the requested user. The threshold value is not unique it will generate random [23] for the user request. The Threshold value is used as a key to decrypt the data. Data center response the request with the threshold value to the user. The user uses the threshold value and can decrypt the data.

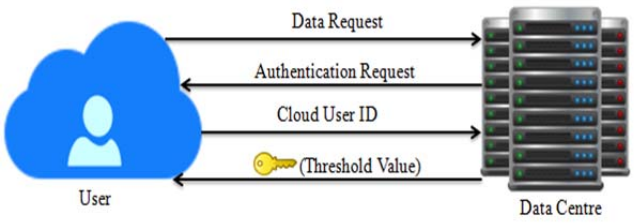

Fig1.Authentication

After the requested query processed the data are divided into n pieces and then the threshold value K is assigned after that K-1 value need to encrypt the data in polynomial form. Once the data is represented in polynomial form apply the value in and make the polynomial form of data into fully encrypted data. By splitting the data, space complexity and transmission overhead is reduced.

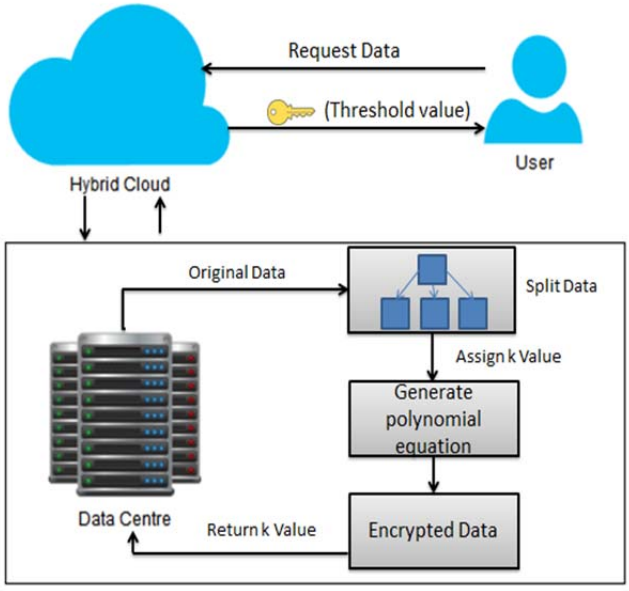

Fig 2: Encryption

Algorithm (Encryption)

Step 1: Divide the data into n parts.
Step 2: Assign the value k i.e. It required to reconstruct the encrypted data.
Step 3: Select random number k-1 to produce a polynomial form of data.
Step 4: Apply n to the polynomial and divide the data
Step 5: Data is encrypted

In decryption operation the user uses the obtained k value of the data center to decrypt the data. The k value applied in the Lagrange polynomial equation [24]. Multiply the k value with the polynomial equation and perform summation. After this original data is retrieved. By applying Lagrange polynomial equation computational complexity is reduced.

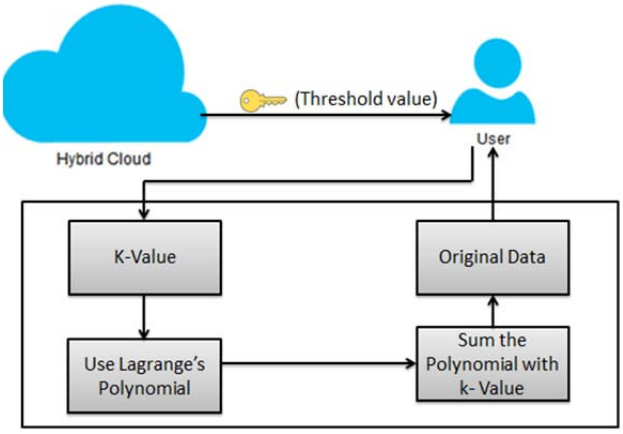

Fig 3: Decryption

Algorithm (Decryption)

Step 1: The value of k parts required to decrypt the data.
Step 2: Apply the k value in Lagrange polynomial.
Step 3: $I_0, I_1, \ldots I_n$.

$$I_0 = \frac{(a-a_1)}{(a_0-a_1)}.\frac{(a-a_2)}{(a_0-a_2)}. \ \ldots$$

$$I_1 = \frac{(a-a_0)}{(a_1-a_0)}.\frac{(a-a_2)}{(a_1-a_2)}. \ \ldots$$

$$I_n = \frac{(a-a_{n-1})}{(a_n-a_{n-1})}.\frac{(a-a_{n+1})}{(a_n-a_{n+1})}. \ \ldots$$

Step 4:

$$\sum_{j=0}^{n} \text{bn}.\, In$$

Step 5: Get the original data

The performance of the algorithm was measured by comparing with the already existing information dispersal algorithm [25]. The metrics used to measure the performance of the algorithm [25] is the size of the data, different threshold value and the time required to perform the operation. According to the different threshold value the time to perform the encryption decryption operation differs.

It depends on the size of the data requested by the user. The evaluation of the algorithms is based on the threshold value and time. In proposing algorithm the cryptographic key distribution is reduced by generating random threshold value. And then no need to configure the access control mechanism for the user involved, it will be authenticated during the request provided using the cloud Id in the hybrid cloud environment. The computational complexity is reduced in this method using simple polynomial computation on the data. It reduces the communication overhead and transmission overhead between the user and server by sending only parts of the data. It's comparatively easy to implement to the cloud server compare with the already existing method.

IV. EXPERIMENTAL ANALYSIS.

The open stack cloud framework [26] is used to set up the hybrid cloud environment. Two or more public and private clouds are developing using open stack. Linux machine is used to setup compute node and controller node on the virtual machine. Virtual machine provides the required hardware resource. Cloud server run on Linux operating system in a virtual machine. A Client can have any operating system. Maria DB [27] server is used to set up a database to handle the medical records. The data center in the cloud environment contains database, it is managed by the Maria DB server. Java programming language is used to implement the proposed algorithm and information dispersal algorithm. After implementing the algorithms need to compare using the threshold value and the time taken to perform the encryption and decryption operation. The time is also depends on the size of data requested by the user (i.e.) query obtains from the user. The algorithm is evaluated by comparing the threshold value, time, the size of the data.

V. CONCLUSION AND FURTHER WORK

In this paper existing methods available for data sharing in a hybrid cloud environment are presented from this a new approach is derived to overcome the disadvantage in the existing method. It is very secure for data sharing in Hybrid cloud environment and it reduces space complexity, computational complexity and transmission overhead. Further work for implementation of the proposed method and the performance is evaluated by comparing proposed algorithm with the information dispersal algorithm.

REFERENCES

[1] Cloud computing [online]. Available: http://www.nist.gov/itl/cloud/.

[2] Hybrid Cloud [online]. Available: http://www.interoute.com/cloud-article/what-hybrid-cloud.

[3] Onik F. A., Salman-Al-Musawi S. S, Anam K. and Rashid N., "A Secured Cloud based Health Care Data Management System," International Journal of Computer Applications, 49 (12), 0975 – 8887, 2012.

[4] Benaloh J, Chase M., Horvitz E. and Lauter K., "Patient controlled encryption: ensuring privacy of electronic medical records," in CCSW '09, 103–114, 2009.

[5] Soubhagya B, Venifa M. G, Jeya A. and Celin J, "A Homomorphic Encryption Technique for Scalable and Secure Sharing of Personal Health Record in Cloud Computing," International Journal of Computer Applications67 (11), 0975 – 8887, 2013.

[6] Information Dispersal Algorithm [online]. Available: http://searchnetworking.techtarget.com/Information-dispersal algorithms - Data-parsing-for-network-security

[7] M. Kallahalla, E. Riedel, R. Swami Nathan, Q. Wang, and K.Fu,"Plutus: Scalable Secure File Sharing on Untrusted Storage," Proc.USENIX Conf. File and Storage Technologies, pp. 29-42, 2003.

[8] E. Goh, H. Shacham, N. Modadugu, and D. Boneh,b"Sirius:Securing Remote Untrusted Storage," Proc. Network and Distributed Systems Security Symp. (NDSS), pp. 131-145, 2003.

[9] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved Proxy-Encryption Schemes with Applications to Secure Distributed Storage," Proc. Network and Distributed Systems Security Symp. (NDSS), pp. 29-43, 2005.

[10] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Secure, Scalable, and Fine-Grained Data Access Control in Cloud Computing," Proc. IEEE INFOCOM, pp. 534-542, 2010.

[11] R. Lu, X. Lin, X. Liang, and X. Shen, "Secure Provenance: The Essential of Bread and Butter of Data Forensics in Cloud Computing," Proc. ACM Symp. Information, Computer and Comm. Security, pp. 282-292, 2010.

[12] Yong CHENG, Jun MA and Zhi-ying "Efficient revocation in cipher text-policy attribute-based encryption based cryptographic cloud storage" Zhejiang University and Springer-Verlag Berlin 2013.

[13] B. Wang, B. Li, and H. Li, "Knox: Privacy-Preserving Auditing for Shared Data with Large Groups in the Cloud," Proc. 10th Int'l Conf. Applied Cryptography and Network Security, pp. 507-525, 2012

[14] Fiat and M. Naor, "Broadcast Encryption," Proc. Int'l CryptologyConf. Advances in Cryptology (CRYPTO), pp. 480-491, 1993.

[15] C. Delerablee, P. Paillier, and D. Pointcheval, "Fully CollusionSecure Dynamic Broadcast Encryption with Constant-SizeCiphertexts or Decryption Keys," Proc. First Int'l Conf. Pairing-Based Cryptography, pp. 39-59, 2007.

[16] Maha TEBAA, Said EL HAJJI, Abdellatif EL GHAZI., "Homomorphic Encryption Applied to the Cloud Computing Security", Proceedings of the World Congress on Engineering 2012, Vol I, July 2012.

[17] Sion, R., and Carbunar, B.: On the computational practicality of private information retrieval. In: Proc. of the Networks and Distributed Systems Security, 2007.

[18] Rabin, M.O.: Efficient dispersal of information for security, load balancing, and fault tolerance. In: Journal of the ACM 36 (2), pp. 335–348, 1989.

[19] Li, M., Yu, S., Ren K. and Lou W., "Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings," in SecureComm'10, 89–106, 2009.

[20] B. Hore, S. Mehrotra, and G. Tsudik, "A privacy-preserving index for range queries," in Proc. Of theVLDB Conf., Pp. 720–731, 2004.

[21] H. Hacigumus, B. R. Iyer, C. Li, and S. Mehrotra, "Executing SQL over encrypted data in the database service provider model," in Proc of the ACM SIGMOD Conf., 2002.

[22] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order is preserved encryption for numeric data," inProc. Of the ACM SIGMOD Conf., pp. 563–574,2004.

[23] Z. Yang, S. Zhong, and R. Wright, "Privacy-preserving queries on encrypted data," in Proc. Of the 11European Symposium on Research In Computer Security, 2006.

[24] Lagranges Theorem [online]. Available: http://mathworld.wolfram.com/LagrangesGroupTheorem.html

[25] Resch, Jason; Plank, James. "AONT-RS: Blending Security and Performance in Dispersed Storage Systems". Usenix FAST'11, 2011.

[26] Openstack [online]. Available: https://www.openstack.org

[27] MariaDB [online]. Available: https://mariadb.org/