Enhanced Data Security for E-Health Records

T.R.Priya[1], Miss. C.Kanimozhi[2]
[1, 2] Dept of Computer Applications
[1, 2] University College of Engineering(BIT campus), Trichy-620024,
TamilNadu, India.

Abstract- *The file hierarchy based storage is propounded to compute the encryption cost and storage overhead in a protected multi-owner data sharing scheme for dynamic groups in the cloud. Using this technique the user can protect the sensitive clinical data been seen by other admins. The shared data files generally have the characteristic of multilevel hierarchy. The proposed layered access structures are integrated into a single access structure, and then, the hierarchical files are encrypted with the integrated access structure. Therefore, both cipher text storage and time cost of encryption are saved and security level is increased. Also there is Clinical Document Architecture (CDA) through which hospitals are enabled to conveniently generate CDA documents without having to purchase proprietary software and physicians and patients can browse the clinical data in chronological order and historical wise.*

Keywords- File Hierarchy, Clinical Document Architecture, Cipher Text Policy Attribute Based Encryption, Slicing Algorithm.

I. INTRODUCTION

Distributed e-healthcare cloud computing systems have been increasingly adopted worldwide including the European Commission activities, the US Health Insurance Portability and Accountability Act (HIPAA) and many other governments for efficient and high-quality medical treatment. In e-healthcare social networks, the personal health information is always shared among the patients located in respective social communities suffering from the same disease for mutual support, and across distributed healthcare providers equipped with their own cloud servers for medical consultant. However, it also brings about a series of provocations, especially how to fortify the security and privacy of the patients' personal health information from various attacks in the wireless communication channel such as eavesdropping and tampering.

To protect data from leaking, users need to encrypt their data before being shared. Access control is paramount as it is the first line of defence that prevents unauthorized access to the shared data[2]. Recently, attribute-based encryption (ABE)[3] has been attracted much more attentions since it can keep data privacy and realize fine-grained, one-to-many, and non-interactive access control. Cipher text-policy attribute based encryption (CP-ABE) is one of feasible schemes which has much more flexibility and is more suitable for general applications.

Cloud service provider (CSP) is the manager of cloud servers and provides multiple services for client. Data owner encipher and uploads the generated cipher text to CSP. User downloads and decipher the interested cipher text from CSP[4]. The shared files usually have hierarchical structure. That is, a group of files are divided into a number of hierarchy subgroups located at different access levels. If the files in the same hierarchical structure could be encrypted by an integrated access structure, the storage cost of cipher text and time cost of encryption could be saved. Here let us take the personal health record (PHR) for example. To securely share the PHR information in cloud computing, a patient divides his PHR information *M* into two parts: personal information *m*1 that may contain the patient's name, social security number, telephone number, home address, etc. The medical record *m*2 which does not contain sensitive information, such as medical test results, treatment protocols, and operation notes. Then the patient adopts CP-ABE scheme to encrypt the information *m*1 and *m*2 by different access policies based on the actual need.

II. RELATED WORK

The existing work mainly focused on fine-grained static medical text access and analysis, which can hardly afford the dynamic health condition fluctuation and medical image analysis. In that System both storage and computation to the untrusted entity would bring a series of security and privacy issues. Since a user patient has many hierarchical structure, encryption cost and storage cost for cloud is huge. In the Existing System, regenerating-coded data only provide private auditing. Existing System requiring data doctors to always stay online and handle auditing. Homomorphic data aggregation allowed single instance. The data security and data confidentiality is less. Cloud Storage cost is huge and encryption time is huge. So the files in the same hierarchical structure could be encrypted by an integrated access structure, the storage cost of cipher text and time cost of encryption could be saved[5]. In this paper, heteromorphic data aggregation is proposed, which serves the basis for proposed architecture. To share data to the admin, the layered model of access structure to solve the problem of multiple hierarchical files sharing is proposed. The files are encrypted with one integrated access structure which would reduce the encryption cost and increase the storage space [3].Also the CDA document generation and integration service is proposed, through which hospitals are enabled to conveniently generate CDA documents without having to purchase proprietary software. The CDA document integration system integrates multiple CDA documents per patient into a single CDA document and physicians and patients can browse the clinical data in chronological order and historical wise. The formal security proof and extensive performance evaluation demonstrate that the proposed architecture achieves a higher security level[4].Heteromorphic data aggregation allowed more instance. Security requirements of data confidentiality and tag consistency are also achieved by introducing a deterministic secret sharing scheme in distributed storage systems. FH-CP-ABE has low storage cost and computation complexity in terms of encryption and decryption.

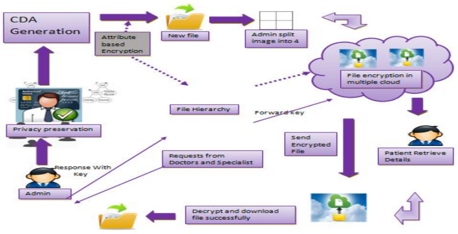


Fig1: Architecture of Proposed System

The figure 1 shows the overall architecture of the system. The proposed system has the registration of patients detail, doctors detail ,hospitals details by the admin. Then the corresponding patients ,doctors and hospitals and their respective details are mapped accordingly. Then the patient's personal history and medical history are entered. With the help of the data entered the CDA document is generated dynamically. While generating CDA document File Hierarchy is adopted where the multiple layered access structure are integrated into single integrated access structure. To improve security of the documents the cipher text policy attribute based encryption is adopted and the document is encrypted. Further if there is any image uploaded it is partitioned into 4 parts and the privacy of the data is preserved by means of usage of slicing algorithm and the privacy access right is given to user. Since there are multiple documents it is stored in cloud (cloud is used for storage) and patient can retrieve details by means of decrypting the secret key and the document is directly downloaded into their respective folder. For privacy preservation slicing algorithm is adopted which is different from conventional key generation method and dynamic medical text mining is adopted for ensuring privacy and security.

III. IDENTITY AND AUTHORITY

Indirectly authorized admins and unauthorized admins cannot correctly distinguish the identities of the user from each other. Only the admins directly authorized by the users can only access the user's information and authenticate their identities simultaneously. The various admins indirectly authorized by user cannot authenticate the user's identities but recover the health information. Unauthorized persons can obtain neither.

IV. PERSONAL HEALTH INFORMATION

In real time personal health information (PHI) in terms of both text and image, which is further aggregated and transmitted to the healthcare provider for the authorized physicians to access and decide corresponding treatment. In smart e-healthcare systems, collected PHI is required to match kinds of medical templates from physicians' experience in the cloud based on specific similarity metrics, to judge the state of the patient suffering/recovering from certain diseases. A patient divides his PHI information M into two parts: personal information $m1$ that may contain the patient's name, social security number, telephone number , home address, etc. The medical record $m2$ which does not contain sensitive personal information, such as medical test results, treatment protocols, and operation notes.

V. FILE HIERARCHY

The layered model of access structure is proposed to solve the problem of multiple hierarchical files sharing. The files are encrypted with one integrated access structure. we also formally prove the security of FH-CP-ABE scheme that can successfully resist chosen plaintext attacks (CPA)An attending physician needs to access both the patient's name and his medical record in order to make a diagnosis, and medical researcher only needs to access some medical test results for academic purpose in the related area, where a doctor must be a medical researcher, and the converse is not necessarily true. Suppose that the patient sets the access structure of $m1$ as: $T1$ {("Cardiology" AND "Researcher") AND "Attending Physician"}. Similarly, $m2$ is termed as: $T2$ {"Cardiology" AND "Researcher"} the information needs to be encrypted twice if $m1$ and $m2$ are encrypted with access structures $T1$ and $T2$, respectively. The two structures could be integrated into one structure T. the computation complexity of encryption and storage overhead of cipher text can be reduced greatly.

IV. DYNAMIC MEDICAL TEXT MINING

The authorized physicians only can successfully decipher the computation result as the evidence of medical treatment. In this subsection an efficient privacy preserving dynamic medical text data mining scheme PPDM1 is proposed by outsourcing the correlation function computation to the cloud on the basis of proposed privacy preserving data aggregation.

VII. PRIVACY PRESERVING- SLICING ALGORITHM

Slicing partitions the data both horizontally and vertically. Slicing preserves better data utility than generalization and can be used for membership disclosure protection. Another important advantage of slicing is that it can handle high-dimensional data.

```
IMPORT  Math,In,Out;
VAR x,y: REAL;
op: ARRAY 10 OF CHAR
In. Open;
In. String(op);
In. Real(x);
IF op=“sin” THEN
y:=Math.sin(x);
Else
y:=Math.cos(x);
End
Out.REAL(y);
```

VIII. CIPHER TEXT POLICY ATTRIBUTE BASED ENCRYPTION

In traditional public-key cryptography, a message is encrypted for a specific receiver using the receiver’s public-key. Identity-based cryptography and in particular identity-based encryption (IBE) changed the traditional understanding of public-key cryptography by allowing the public-key to be an arbitrary string, e.g., the email address of the receiver. ABE goes one step further and defines the identity not atomic but as a set of attributes, e.g., roles, and messages can be encrypted with respect to subsets of attributes (key-policy ABE – KP-ABE) or policies defined over a set of attributes (ciphertext-policy ABE – CP-ABE). The key issue is, that someone should only be able to decrypt a ciphertext if the person holds a key for “matching attributes” (more below) where user keys are always issued by some trusted party.

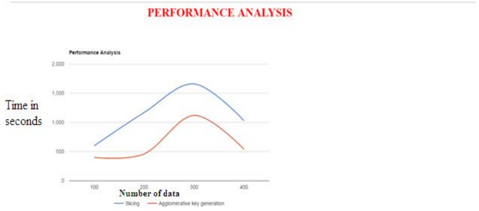
IX. CLINICAL DOCUMENT ARCHITECTURE

CDA (Clinical Document Architecture) is a document markup standard that specifies the structure and semantics of ‘clinical documents’ for the purpose of exchange. The CDA document pertaining to a patient is generated at the clinic where the patient is diagnosed. The generated CDA document can be sent to other clinics after patient’s consent is acquired. The exchange of CDA document is triggered in the following cases: when a physician needs to refer to the patient’s medical history; when referral and reply letters are needed for a patient who is being taken care of by multiple clinics; when the patient is in an emergency and the medical history needs to be reviewed.

X. CLOUD STORAGE

CloudMe is a file storage service operated by *CloudMe AB* that offers cloud storage, file synchronization and client software. It features a blue folder that appears on all devices with the same content, all files are synchronized between devices. The CloudMe service is offered with a freemium business model and provides encrypted ssl connection with SSL Extended Validation Certificate. *CloudMe* provides client software for Microsoft Windows, macos, Linux, Android, iOS, Google TV, Samsung Smart TV, WD TV, Windows Storage Server for NAS and web browsers. Now data stored in cloud can be retrieved safely with the help of slicing algorithm which divides the data so cannot be easily identified also cipher test to encrypt

XI. PERFORMANCE ANALYSIS



Thus these helped to make more efficient by performance.

XII. CONCLUSION

This can be made more time saver with the help of more data mining techniques.

REFERENCES

[1] “ClubCF:A Clustering-based Collaborative Filtering Approach for Big Data Application” IEEE Transactions on Emerging Topics in Computing 2014

[2] G. Thilagavathi, D. Srivaishnavi, N. Aparna, et al., “A Survey on Efficient Hierarchical Algorithm used in Clustering,” International Journal of Engineering, vol. 2, no. 9, September 2013.

[3] V. Gupta, G. S. Lehal, “A Survey of Common Stemming Techniques and Existing Stemmers for Indian Languages,” Journal of Emerging Technologies in Web Intelligence, vol. 5, no. 2, pp. 157-161, May 2013

[4] R. S. Sandeep, C. Vinay, S. M. Hemant, “Strength and Accuracy Analysis of Affix Removal Stemming Algorithms,” International Journal of Computer Science and Information Technologies, vol. 4, no. 2, pp. 265-269, April 2013.

[5] A. Bellogín, I. Cantador, F. Díez, et al., “An empirical comparison of social, collaborative filtering, and hybrid recommenders,” ACM Trans. on Intelligent Systems and Technology, vol. 4, no. 1, pp. 1-37, January 2013.

[6] M. A. Beyer and D. Laney, “The importance of ‟big data‟: A definition,” Gartner, Tech. Rep., 2012.

[7] F. Chang, J. Dean, S. mawat, et al., “Bigtable: A distributed storage system for structured data,” ACM Trans. on Computer Systems, vol. 26, no. 2, pp. 1-39, June 2008.

[8] X. Wu, X. Zhu, G. Q. Wu, et al., “Data mining with big data,” IEEE Trans. on Knowledge and Data Engineering, vol. 26, no. 1, pp. 97-107, January 2014.

[9] Z. Liu, P. Li, Y. Zheng, et al., “Clustering to find exemplar terms for keyphrase extraction,” in Proc. 2009 Conf. on Empirical Methods in Natural Language Processing, pp. 257-266, May 2009.

[10] M. R. Catherine, and E. B. Edwin, “A Survey on Recent Trends in Cloud Computing and its Application for Multimedia,” International Journal of Advanced Research in Computer Engineering & Technology, vol. 2, no. 1, pp. 304-309, January-February 2013.