ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

Healthcare Data Security in Cloud Computing

G. Rathi, Abinaya. M, Deepika. M, Kavyasri. T

Assistant Professor, Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India.

Student, Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India.

Student, Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India.

Student, Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, India.

ABSTRACT: Modern healthcare sectors need to create an environment which reduces time consuming efforts and other costly operations to obtain a patient's complete medical record and uniformly integrates this heterogeneous collection of medical data to deliver it to the healthcare authorities. Electronic health records (EHRs) have been widely adopted to enable healthcare providers and patients to create, manage and access healthcare information from any place, and at any time. Cloud services provide the necessary infrastructure at lower cost and better quality. Cloud computing when used in Healthcare sector reduces the cost of storing, processing and updating with improved efficiency and quality. But the security of data in the cloud is not satisfactory today. The electronic health record consists of images of the patient's record which is highly confidential. The EHRs in the healthcare includes the scan images, DNA reports, X-rays etc., which are considered as the patients private data. Providing security for a huge volume of data with high efficiency is required. This paper introduces a new mechanism in which the EHR images can be secured efficiently and the private data are preserved for later use. Since most of the confidential data are in the form of images, additional care must be taken to secure these images. This is done by reducing the images into pixels and then encrypting those pixels. After the encryption, the single encrypted file is divided into n files and they are stored in the cloud. The original data is obtained by merging the n divided files from the cloud and then decrypting that merged file using the private key. This key is made visible only to the authorised persons as required by the hospital.

KEYWORDS: Cloud; Security; Healthcare Data; Encryption; Decryption; EHR.

I. INTRODUCTION

Cloud Computing is defined as a technology which uses the internet and central remote servers to maintain data and other applications. Cloud computing allows enterprises and consumers to use applications without installation and access their files at any computer with internet access. This technology allows for efficient computing by centralized data storage, processing and bandwidth. Cloud computing has been envisioned as the next-generation information technology (IT) architecture for enterprises, because of its long list of unprecedented advantages in the IT history: on-demand self-service, shared network access, location independent resource pooling, rapid resource elasticity, transference of risk and usage-based pricing. As a disruptive technology with profound implications, cloud computing transforms the nature and the way of how enterprises use information technology. The primary aspect of this paradigm shifting is that data are being centralized or outsourced to the cloud. From the user point of view, storing data remotely to the cloud in a flexible on-demand manner brings appealing benefits such as relief of the burden in case of universal data access with location independence, storage management, and avoidance of capital expenditure on hardware, software components and personnel maintenance, etc., Thus Cloud promises to provide service to users without reference to the infrastructure on which these clouds are hosted.

Though cloud computing has many advantages, there are many risks in it. The major risk is providing security to cloud resources and data from unauthorized access. There are many security issues/concerns in cloud computing. These issues are faced by both cloud providers and cloud users. The issue could be either on data side or network side. For providing data security, many algorithms have been proposed. All the proposed algorithms provide security to the data using the encryption technique. The Pailier cryptosystem and AES algorithm is used to encrypt the image and text files [1]. Other than these two algorithms, a homomorphic encryption algorithm – FHE is also used to secure the data from

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

unauthorized access [2]. Many encryption algorithms are being used to provide security to the data that are stored in cloud.

A. Need for Cloud

- Requires less initial investment
- Needs fewer skilled internal IT resources
- Operating cost is less
- Payments are streamed with usage
- Resources are infinite

B. Cloud Architecture

Cloud Computing architecture can be divided into two sections. They are named as front end and back end. The front end includes the client's computer (or computer network) and the applications required to access and perform operations on the cloud computing system. It is not necessary that all cloud computing systems should have the same user interface as shown in figure 1.1. At the back end of the system there are various systems, servers and several data storage systems which create the "cloud" of computing services.In theory, a cloud computing system can include any computer program, from data processing to video games. In practice, each application will have its own dedicated server.

A central server administers the system, monitors traffic and ensures that everything goes smoothly. It follows a set of rules which are called as protocols and uses a special kind of software called middleware**.** Middleware helps the networked computers to communicate with each other. Mostly, servers don't run at full capacity,which means the processing power gets ruined. It is possible to fool a physical server to make it think that it's actually a multiple server, each running with its independent operating system. This technique is called as server virtualization. Server virtualization reduces the need for more physical machines by maximizing the output of individual servers.

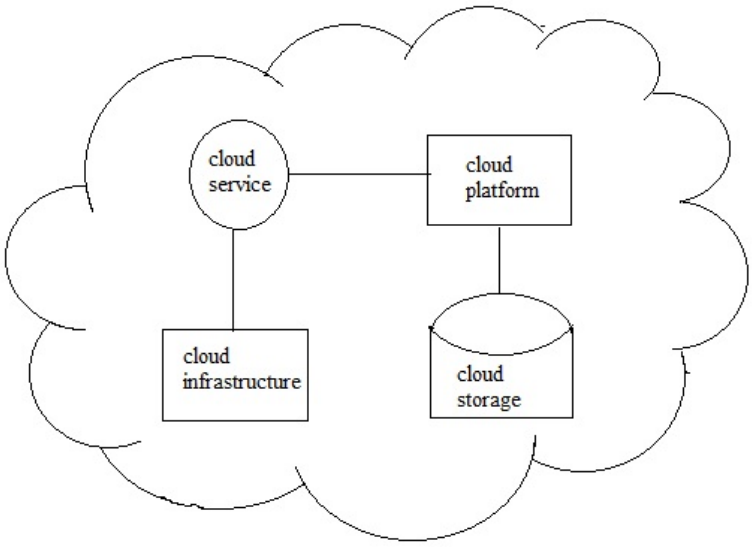

Figure 1.1 Cloud Basic Architecture

C. Cloud Deployment Models

- **Public Cloud:**In a public cloud, IT resources are made available to the public organizations and are owned by the Cloud service provider as shown in figure 1.2. The cloud services are made accessible to everyone via standard internet connection. In a public cloud, a service provider makes IT resources such as applications, storage capacities available to any consumer. This model is considered as an "on-demand" and "pay-per use" environment, where there are no on-site infrastructure or management requirements. These benefits come with certain risks such as no control over the resources, data security, network performance and interoperability.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

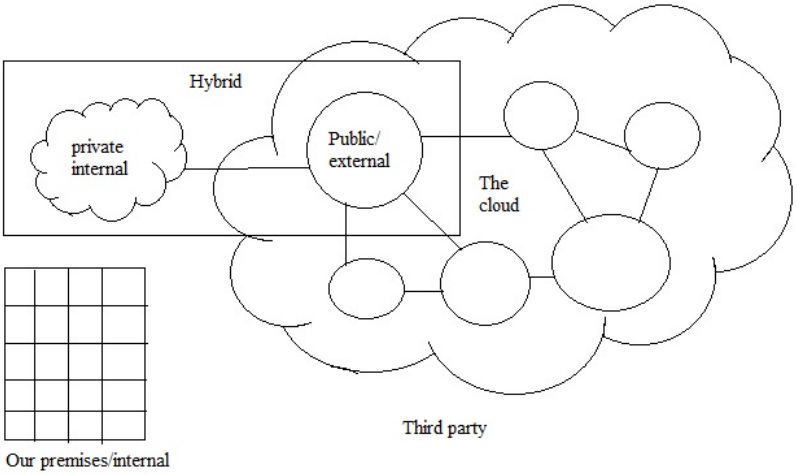

Figure 1.2 Cloud Computing Deployment Models

- **Private Cloud**:In a private cloud, the cloud infrastructure operates separately for each organization and is not shared with any other organizations. This cloud model offers the greatest level of security and control. The two variations are as follows,
  - **On-premise private cloud**: This is also known as internal clouds and are hosted by an organization within their own data centers. This model provides a more standardized process, but is limited in terms of size and scalability.Thisis best suited for applications which require complete control of the infrastructure and security.
  - **Externally-hosted private cloud:** This type of private cloud is hosted externally with a cloud provider, where the provider facilitates an exclusive cloud environment for a specific organization with full guarantee of privacy or confidentiality.
- **Hybrid Cloud**: In a hybrid cloud environment, the organization consumes resources from both private and public clouds.For the maintenance of service levels, the public cloud resources are imbibed with the private cloud resources. Organizations use their computing resources on a private cloud for normal usage, but access the public cloud for peak load/high requirements. This ensures that a sudden increase in computing requirement is handled gracefully.

D. Cloud Service Models

- **Infrastructure as a Service [IaaS]:**IaaS provides virtual storage, virtual machine, virtual infrastructure, and other hardware components as resources that clients can provision. The IaaS service provider manages all the infrastructure, while the client is responsible for all other aspects of the deployment. This can include the applications, operating system and user interactions with the system as shown in figure 1.3.
- **Platform as a Service [PaaS]:** PaaS provides virtual machines, operating systems, services, applications, transactions, development frameworks and control structures. The client can deploy his/her application on the cloud infrastructure or use applications which are programmed using languages and tools that are supported by the PaaS service provider. The service provider controls the cloud infrastructure, the operating systems, and the enabled software. The client is responsible for managing and installing the application which it deploys.
- **Software as a Service [SaaS]:** SaaS is the top most layer of the cloud computing stack, which is directly consumed by the end user. The consumer can make use of the service provider's application that runs on a cloud infrastructure. It is accessible from various client devices through a thin client interface such as web browser. They offer many advantages such as reducing the need for infrastructure because they provide storage and compute powers remotely which also reduces the need for manual updates as it could perform those tasks automatically.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

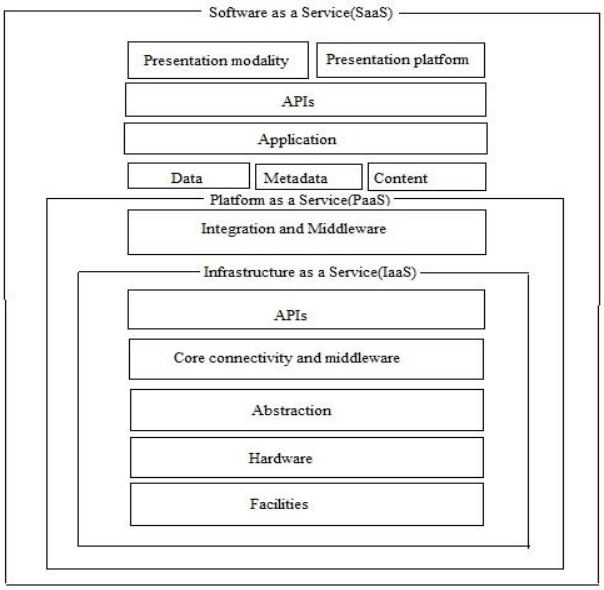

Figure 1.3 Cloud Computiing Service Models

II. RELATED WORK

In [1], the authors used 2 different encryption techniques to protect the data from unauthorized access. They used Pailier cryptosystem to encrypt the images and to encrypt the text files, AES algorithm is used. In [2], the authors used a homomorphic encryption technique to secure the data. They analyzed how cloud homomorphic encryption with splitting key and key delegation can help in securing the healthcare data. They proposed an FHE algorithm with key delegation to ensure data privacy in multilevel hierarchical order. In [3],the authorsproposed a secure cloud storage system supporting privacy-preserving public auditing.To provide security, third party auditors are used. The third-party auditors perform audits for multiple users simultaneously and efficiently. In [4], the authors focused on data integration using Data Integrity protection scheme for preserving its intrinsic properties of fault tolerance**.**In [5], the authorsfocused on data access control to enhance data security. The authors used Ciphertext-policy Attribute based encryption to encrypt the data.

III. CLOUD FOR HEALTHCARE SECTORS

In modern healthcare environments, there is a strong need for an infrastructure which reduces time consuming efforts and costly operations to obtain a patient's complete medical record and uniformly integrates this heterogeneous collection of medical data to deliver it to the healthcare professionals. Electronic health records have been widely adopted to enable healthcare providers, insurance companies and patients to create, manage and access healthcare information at any circumstances.

All the healthcare industries need to handle more requests with the available resources. The main objective of all the healthcare organization is to increase the number of people getting access to healthcare services. Therefore day by day the amount of data that need to be stored, processed and updated is increasing exponentially. The healthcare industries demand more computation ability so that the quality of the service increases. Cloud computing improves patient care by providing faster, better, secure and ubiquitous services at a lower cost and which meets the requirements of the healthcare sector.

As a result, healthcare providers are more willing to shift their systems to clouds that can remove the geographical distance barriers among providers and patients. With cloud computing, different doctors can access a patient's health records even if they're miles apart. These physicians need not have a direct communication to request for a transfer of health records. They can just access them through clouds.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

A. Risks in Cloud Computing

Cloud computing has many risks like data security, data confidentiality and overhead. The data processed in the cloud is highly confidential, such as business records, patient records, military records etc. Therefore a proper encryption standard must be applied in order to secure the sensitive data from being tampered. And most of the time data that is being stored or processed in cloud are in large numbers and the cloud servers sometimes become lazy while computing and affects the correctness of the end result. Therefore the computation has to be made transparent.

A widely used real world industry which requires a continuous innovation in storage, access and computation of data in the form of records is the healthcare domain. Healthcare data mainly comprises of large media files such as radiology, X-ray, CT scans, and other type of images and videos. Such files are called as the Electronic Health Records (EHRs) and are often stored in distributed storage. These EHR holds the healthcare sector people and the patients from all forms of misinterpretation like doctor's handwriting, losing prescriptions etc. For example, a healthcare sector may have large number of records like 50, 00,000 and they can be preserved through this cloud efficiently. In modern healthcare domain, EHRs have been widely adopted to enable healthcare providers and the patients to create, manage and access patient's healthcare information from anywhere and at any time. As a result, a patient's EHR can be found scattered throughout the entire healthcare sector. Typically, an EHR contains sensitive information that is personal information concerning a person's health disorders, their pictures, and other health information. These images are the most confidential ones and needs to be protected.

To put everything online "in the cloud", unencrypted, is a big risk. Healthcare data security has been in exist for a long time, but since cloud computing gains more and more attention, healthcare providers are aiming at utilizing cloud's advantages to their benefit. However, these advantages come at a cost of various information security risks that need to be carefully considered. Risks vary depending on the sensitivity of the data to be stored or processed, and how the chosen cloud vendor has implemented their specific cloud services.

IV. EXISTING SYSTEM

The existing system secures the EHR using the encryption technique. The images and other data are encrypted before they are stored in the cloud. For encrypting these data, the Pailier cryptosystem is used. The Pailier cryptosystem converts the images into pixels. The array of pixels is then converted into matrix of pixels according to the dimension of the image. Using the homomorphic encryption, this matrix is encrypted. For encrypting the text files, the Advanced Encryption Standard (AES) technique is used. The decryption of the encrypted files is done after the retrieval from the cloud. The decryption is done using the private key which is present with the doctors and other physicians who use the EHRs to get the information of the patients.

A. Drawbacks in the Existing System

In the existing system, if the decryption key is known, then the data can be retrieved easily from the cloud. There are several techniques available to hack the decryption key. Some of the techniques are Key Search technique, Brute Force attack, Crypt Analysis and Systems-Based attack. The existing system provides single layer protection to the EHRs. If that single layer is passed, then the data can be easily retrieved from the cloud.

V. PROPOSED SYSTEM

The proposed system provides two-layer protection to the EHRs. In the first layer, the images and the text files are encrypted using Advanced Encryption Standard (AES). In the second layer, the encrypted files are divided into n files. These n files are then stored in the cloud. The original EHR can be decrypted only if the n files are merged. For splitting and merging the ciphertexts, a sequence key will be used.
The overall architecture of the proposed system is depicted in Figure 4.1.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

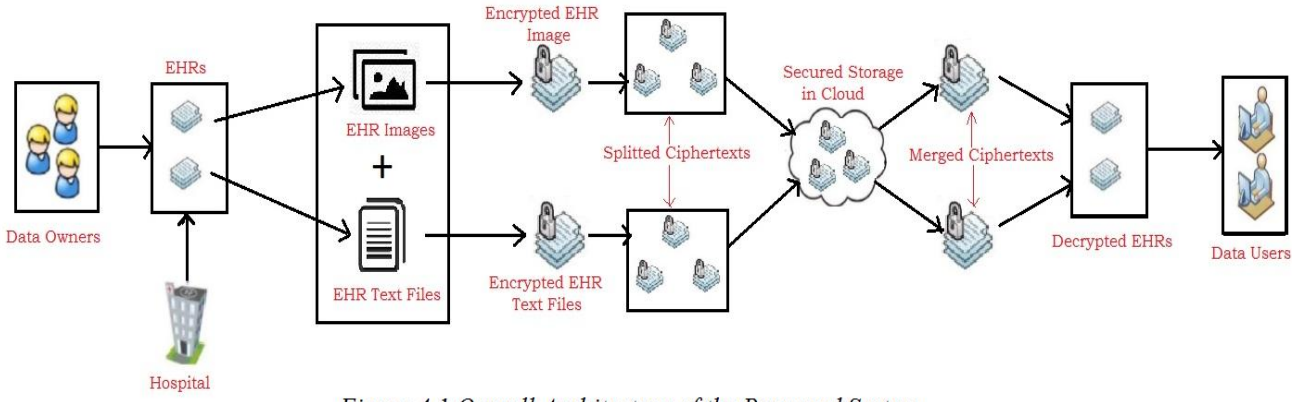

Figure 4.1 Overall Architecture of the Proposed System

The proposed system includes the following components,

A. Data Owners

The data owners are the patients whose information is stored in the EHRs. When a patient is in need of a treatment and if the hospital maintains EHRs of the patients, the patient need not spend time and effort in explaining his medical history. He does not have to maintain paper health reports as well. This is because these EHRs when stored in cloud can be updated and can be retrieved frequently by the hospitals maintaining them. They can also be shared with other healthcare institutions when required, with prior consent from the patients.

B. EHRs

EHRs are the patient information that are stored digitally. Information like the patient's medical history, their current medications, scan reports, X-rays, etc. that require high privacy are present in the EHRs. These EHRs can be updated and processed whenever and wherever necessary when stored in the cloud. They also become readily accessible to the specified users and doctors. Since these EHRs can be in the form of text data and multimedia data like the images, different techniques have been adopted to secure the different formats of data in the EHRs.

C. Encryption

The encryption of EHRs is done using the Advanced Encryption Standard (AES). It is a symmetric encryption algorithm in which the same key is used to both encrypt and decrypt files. The key size used to encrypt the plain text is 128 bits. This algorithm helps in encrypting both the text and image files.

D. Sequence Key

The sequence key is used to split and merge the ciphertexts. After encryption, the ciphertext will be splitted into 'n' ciphertexts. Before decryption, the splitted ciphertexts will be merged to get the original ciphertext. The encrypted files are divided into 'n' files based on the sequence key entered by the physician or other authorized persons. There should be at least 8 characters in the sequence key. The value of 'n' depends on the length of the sequence key. For example, if the sequence number is entered as adk21937 (Length = 8), then the encrypted files will be divided into 8 files.

The sequence key will be considered as valid only if it satisfies the following conditions,

i. The sequence key should not contain any space.
ii. The frequency of each character in the sequence key should be 1.
iii. The length of the sequence key must be at least 8.

Sample invalid sequence keys:

i. Adk 1234 – Space is used.
ii. Adk21224 – The character '2' is used more than once.
iii. Adk12 – Length is lesser than 8.

The sequence key that is used to split the ciphertext will be used to merge the splitted files as well. The first and the last character in the splitted file would be the characters in the sequence key. The first character indicates the index of the current file and the last character indicates the index of the next file that need to be merged with the current file.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

E. Cloud Storage and Retrieval

Once the ciphertext is splitted, the splitted files are stored in the cloud. Cloud provides services as required by the clients. It is also easy for the clients to avail the services of the cloud. It provides the required resources to store the large amount of EHRs at low price. 24X7 service is also available for the clients. The services provided are of good quality and suitable for the healthcare sector. The updated EHRs are stored in the cloud and then retrieved whenever necessary by the physicians and doctors. The storage and retrieval time is also less when cloud resources are utilized. The EHRs can beretrieved from the cloud and then decrypted after merging the ciphertexts.

F. Decryption

In decryption, the splitted ciphertexts will be retrieved from the cloud and then those ciphertexts will be merged to get the original ciphertext. The decryption of the encrypted files will be done after merging the splitted ciphertexts. The decryption is done with the help of a private key which is made available to the doctors and other users who need the healthcare data. The key is generated from the Advanced Encryption Standard. For decrypting the text files, AES algorithm can be used.

VI. RESULT

The following snapshots exhibit the original EHR and its encrypted form, which is followed by the user defined sequence key and also the splitted ciphertexts, which is generated as a result of the predefined sequence key.

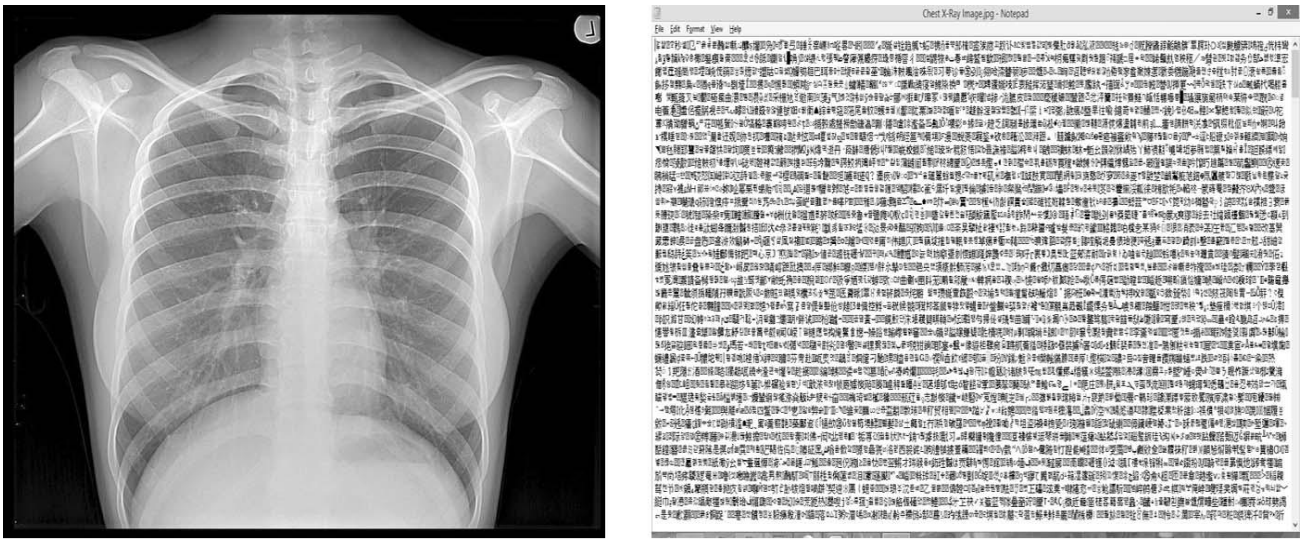

Figure 5.1 Original and Encrypted EHR

Figure 5.1 shows how data appear before and after encrypting the data. The left side image is the original EHR (Plaintext) and the right side image is the encrypted EHR (Ciphertext).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

Figure 5.2 User defined SEQUENCE KEY to split the encrypted EHR

Figure 5.2 shows the browser screen that asks for the user defined sequence key to split the encrypted EHR. The same sequence key should be used while merging the splitted EHRs.

Figure 5.3 Splitted Electronic Health Records

Figure 5.3 shows a set of splitted files that are splitted using the above entered sequence key. The characters in the sequence key are used as the index of the splitted files, which will be removed in the merging process.

VII. CONCLUSION

In this paper, a new mechanism is proposed to protect the healthcare data in the cloud. This system has a double layer protection in which the EHRs are stored in the cloud. Encryption/ Decryption will be done in one layer and in the other layer, Splitting/ Merging of the ciphertext will be done. Thus, data security can be improved in cloud computing. As the proposed system is in the development stage, the actual results will be shared in future publication.

REFERENCES

1. Aishwarya.R, Divya.R, Sangeetha.D and Vaidehi.V, 'Harnessing healthcare data security in cloud', International Conference on Recent Trends in Information Technology (ICRTIT), Pages 482 – 488, 2013.
2. Elmogazy.H and Bamasak.O, 'Towards healthcare data security in cloud computing', The 8th International Conference for Internet Technology and Secured Transactions (ICITST), Pages 363 – 368, 2013.
3. Cong Wang, Chow S.S.M. Qian Wang, Kui Ren and Wenjing Lou, 'Privacy preserving public auditing for secure cloud storage', IEEE Transactions on Computers, VOL. 62, NO. 2, February 2013.
4. Henry C.H. Chen and Patrick P.C. Lee, 'Enabling data integrity protection in regenerating – coding based cloud storage', IEEE Transactions on Parallel and Distributed Systems, VOL. 25, NO. 2, February 2014.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

5. Kan Yang and Xiaohua Jia, 'Expressive, Efficient and revocable data access control for multi authority cloud storage', IEEE Transactions on Parallel and Distributed Systems, VOL. 25, NO. 7, July 2014.

BIOGRAPHY

Rathi. G is working as an assistant professor in Sri Ramakrishna Engineering College, Coimbatore, India. She has teaching experience of about 9 years. Her areas of interest include Cloud Computing and Internet of Things (IoT).

Abinaya. M is a final year Computer Science and Engineering student of Sri Ramakrishna Engineering College, Coimbatore, India. Her areas of interest are Cloud Computing and Computer Networks.

Deepika. M is a final year Computer Science and Engineering student of Sri Ramakrishna Engineering College, Coimbatore, India. Her areas of interest are Cryptography and Computer Networks.

Kavyasri. T is a final year Computer Science and Engineering student of Sri Ramakrishna Engineering College, Coimbatore, India. Her areas of interest are Database Management System and Data Mining.