IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Achieving Security and Privacy in Cloud-Based E-Health System Using Encryption and De-duplication

Rushikesh Umak,

Dept. of Computer Science and Engineering,
PRM Institute of Technology and Research, Amravati,India

*Abstract:*E- Healthcare system plays a major role in the society. It monitors the health condition and helps in giving appropriate medical treatments. This system aims at gathering and storing patient’s details and sharing health related information. It also has high legitimate concerns about patient’s privacy and information security. This system minimizes the infrastructural barriers for the developing nations. It also extends healthcare systems to the remote and isolated areas which has limited access to medical technologies, remote health services are provided through telecommunications. Quality of the service and security are added advantage to the system. They collect the real time personal information (PHI) and health problems from patients and transmit them to the healthcare provider for the authorized physicians to decide on the corresponding treatment. They send the PHI in terms of text and image to the cloud, and also the other personal queries related to their medical history. In cloud computing, collected PHI should match the physicians experience to judge the state of the patient and unfortunately, delegating both storage and computation to the untrusted entity would bring a series of security. This is where de-duplication comes into play. It is a technique for eliminating duplicate copies of data, and has been widely used in cloud storage to reduce storage space

.

***Index Terms* - E-Healthcare, PHI** (Personal Health Information)**, cloud computing, de-duplication, ABE** (Attribute Based Encryption**).**

I. INTRODUCTION

Cloud based healthcare computing have changed the face of healthcare in many ways. The main advantages of cloud computing in healthcare are scalability of the required service and the provision to upscale or downsize the data storge, collaborating Artificial Intelligence (AI) and machine learning. The current paper examined various research studies to explore the utilization of intelligent techniques in health systems and mainly focused into the security and privacy issues in the current technologies.E-Healthcare is an emerging field of medical informatics, referring to the delivery of health services and information using the Internet and related technologies.Rendering efficient storage and security for all data is very important for cloud computing. Securing and privacy preserving of data is of high priority when it comes to cloud storage. E-Healthcare is the most important source in the healthcare society. E-healthcare system is now being popularized globally. Implementing the E-healthcare system will have more advantages such as online services for teleconsultation (second medical opinion), e-prescription, e-referral, telemonitoring, telecare etc. E-healthcare system provides high level of security and cost-effective use of patients records, information and communication in support of healthcare and health related issues.

In current healthcare systems, there is a high demand on establishing a framework that minimizes time-consuming work and expensive procedures to retrieve a patient's medical record and integrating this varying set of medical data consistently to deliver it to the healthcare industry. Electronic health records (EHRs) have been widely accepted to allow patients, insurance companies, and healthcare providers to initiate, control and process patients' healthcare information from any place, and at any time. Thus, healthcare providers accept moving their data and operations to the clouds that can perform their operations more efficiently and eliminate the physical distance concern between patients and providers. Cloud service enables different doctors to obtain an access to a patient’s health record even if they are kilometres apart. There is no need for the doctors to make a phone call to ask for a move of the health records; they will just access them in the clouds. Despite all the benefits cloud computing provides for healthcare systems, data privacy and security are among the major concerns, which make healthcare move slowly towards the acceptance of these new technologies. Cloud computing benefits come at a price of the emergence of different risks related to information security that must be cautiously addressed. Risks differ according to the criticality of the data to be processed or stored, and how the specific cloud provider has developed their specific cloud services. In order to be appealing to healthcare community, cloud computing should maintain required guarding to address HIPAA (Health Information Portability and Accountability Act) of U.S. Departmentof Health and Human Services (2013) and other security and privacy requirements. Although Electronic Health Records (EHRs) has been regulated in standards, such as HIPAA, several cloud providers are still not compliant with them. In order to secure healthcare data, the first step to be taken is to categorize the data in the Electronic Health Records (EHRs) in correspondence to its level of security sensitivity. The first category is Personal Identifiable Information (PII), such as patient records, normally saved in a relational database as structured data. The second category is Healthcare data (PHI), which is typically consists of large media files such as radiology, CT scan, x-ray, and other types of video and images that conceal patient's identity. Such files are often stored in distributed storage. A medical record has some components that are classified by both individuals and organizations, such as HIPAA of U.S. Department of Health and Human Services(2005)as highly critical and should be disclosed only to the entities that have an explicit access right to them. This is because revealing such data can lead to unjustly show bias against an individual or refuse them chances that they otherwise entitled to. For example, knowing that a person is diabetic might negatively influence their professional growth, personal relationships, insurance cost, and employment opportunities. Outsourcing the storage of unencrypted information in the cloud, is of a high danger. For a highly sensitive data, such as Personal Health Informations (PHIs), locating them unencrypted, out of site, is considered against the law. However, to access data stored on a distance server, the Cloud providers need to access the primitive, i.e. un-encrypted, data. Most people do not have full confident on the Cloud providers for their sensitive healthcare data because there is no law regulating how they use this data and whether the patients have control over them. On the other hand, data encryption might counteract the advantages of cloud computing, unless the cloud service providers get the secret decryption key. Traditional cryptography is not a solution in this situation (. Patients may only want portions of their record made available to all doctors and specific portions to be available to specific users, i.e. insurance company. Patients can be given maximum control over their data by encrypting each portion of a patient’s record under a different policy. Access control policies should be active to ensure that accessing sensitive information is restricted only to parties that have a valid privilege. This feature can be provided by Attribute-Based encryption.

II. Ease of Use

2.1 Advantages of Cloud-Based E-healthcare Systems

The more a healthcare center connects system information to a global computer network such as the Internet, the more it opens up access from around the world and facilitates data leaks. The need for an electronic health record should be protected from illegal users who may misuse this for a variety of purposes. Identity-based encryption is oneofthebestsecuritysolutionstoprotecte-Healthrecord data. Thealgorithmdeals withproblemsfoundincommoncryptographictechniquesusinganythreadasapublickey. Thesystemcanenhancethesecurityofhealthrecordsbyaddingauthentication procedures to connected servers . In this system, communication between this servers uses encrypted data using ABE, so that each server can perform the encryption and decryption process during the data exchange. Only servers with IDs can access and extract health record data. Currently, test results show performance relative to the speed of the algorithm used in the system . CloudStorage is a computer model that stores data on the Internet or in the cloud. Cloud storage is delivered according to demand and capacity and costs that will leave the customer investing and managing their data storage infrastructure. This provides speed, scale, and durability. Below are some of the general advantages of cloud computing; in our case, we focus on E-health systems.

- Ease of access using a ‘Web Browser’with integrated Single-Sign-On (SSO). Norequirement for VPN to access Cross Sites or Networks. SimplifiedManagement and On-demand Scalability.

- No Overhead Cost to maintain the physical infrastructure.

- No Hardware post warranty charges for the physical infrastructure.

- No Power Consumption.

One of the major schemes in healthcare systems is attribute-based encryption for data. Encryption provides high-class access control for every user and revocation, scalability, dynamic user management, and traceability

2.2 Challenges in Cloud Computing Cloud Computing

Challenges are always been there. Companies and organizations are aware of the values that cloud computing brings and are taking necessary steps towards the transition to the cloud environment. Like any new technology the adoption of cloud computing is also full of issues and other challenges. Some of them are:

Confidentiality: Confidentiality is a process or mechanism of safeguarding patient health data from unauthorized access from public or internal users. Unauthorized access is dangerousandcanpotentiallyresultindataleakageandcanevencauseseriousdamageto businesses. With respect to the data size, the number of patients on devices increases, and thereisahugepotentialthreattothedatatoexposethesetoexternalparties. Confidentiality is important in the healthcare industry as the patient can be reluctant to give personal details to doctors if they are not confident with the confidentiality. By implementing access control and using encryption techniques, confidentiality can be achieved.

Integrity: Integrity is important factor to make sure that the data are not changed at any single point in time. The HIPAA Security illustrates that covered entities must implement procedures and policies to protect electronic healthcare information from improper destruction or alteration. Integrity can be achieved by a hashing mechanism or checksum for all the data. One of the best and accurate ways is by implementing block chain technology as it is merely impossible to change the hash of the data as it will change the entire chain if any of the hashes are changed.

Availability: The information must be available all the time. Business critical systems should be clustered or must have high availability to have maximum uptime without service interruptions.

Data Violations: Business Impact on Company Dignity and Trust for Customers or Partners. Degradation of intellectual property by competitors can lead to product outsourcing, financial discovery, and the occurrence of events and forensics.

Wrong fix: This is one of the most common cloud challenges. As cloud computing is a shared resource, any misconfiguration of the datacenter will lead to complete exploration of all the customer data hosted within the same datacenter.

Lack of Security Technologies: The biggest challenge during the transition to cloud computing is the implementation of appropriate security architecture to withstand cyberattacks. Unfortunately, this process remains a mystery for several organizations.

Account hijacking: A key feature where attackers gain access to accounts, and serious or sensitive rights are exploited. Criminal attacks on sensitive data, cloud system exploits, or access to stolen signals can put these accounts at risk.

Insider Threat: Circumstances have been identified including malicious servers, employees saving sensitive data on their unprotected devices and programs, employees or other insiders who steal stolen emails exposed by malicious attacks on company assets

.

Unsecured APIs: Cloud computing providers develop a range of user software and APIs to allow customers to manage and interact with cloud services. The security and availability of standard cloud services are linked to the security of those APIs. Poorly designed APIs can lead to misuse or even worse, infringement of information. Exposed, broken, and hacked APIs have serious concerns about data breaches. Healthcare really needs to understand the safety requirements for designing and introducing visible connectors online.

III. THE PROPOSED MODEL

E-Healthcare is an emerging field of medical informatics, referring to the delivery of health services and information using the Internet and related technologies. EHealthcare is the most important revolution in the healthcare society recently. E-healthcare system is now being globalized.

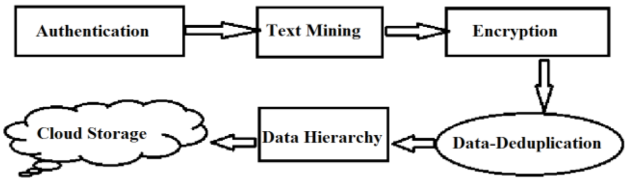


Figure: System Overview

Authentication

Indirectly authorized physicians and unauthorized persons cannot correctly distinguish the identities of the patients from each other. Only the physicians directly authorized by the patients can access the patients' personal health information and authenticate their identities simultaneously. The physicians and research staff indirectly authorized by patients cannot authenticate the patients' identities but recover thepersonal health information.

Text Mining

It is well known that the characteristics of one specific disease would vary as the health deteriorating status and/or the recovering status develops. With a certain dosage during a course of treatment, some specific vital signs such as the body temperature, the blood pressure, the leucocyte count and the blood platelet count possess their own regularities in each time period. Therefore, it is required to compare the dynamically collected personal health information (PHI) from the patient with the experience PHI template for one specific disease, each of which is represented by a vector of multiple elements representing the values of vital signs for each time period, to decide whether the patient's health condition is deteriorating or recovering.

Encryption

Attribute-based encryption (ABE) is a concept of public-key cryptography. 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. Attribute Based Encryption (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 (cipher text-policy ABE - CP-ABE). The key issue is, that someone should only be able to decrypt a cipher text if the person holds a key for "matching attributes" (more below) where user keys are always issued by some trusted party.

Data-Deduplication

Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage. Initially the name of the file is compared with the other available files in the cloud. After which the data inside the file is compared with the data that is already present. If either of which match with the new file, then the file will not be allowed to be saved in the cloud system. The technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. Deduplication can take place at either the file level or the block level. For file level deduplication, it eliminates duplicate copies of the same file. Deduplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.

Cloud Storage

Cloud computing is now the hot spot of computer business and research. Several trends are opening up the era of Cloud Computing, which is an Internet-based development and use of computer technology. Cloud Storage usually contains business-critical data and processes, hence high security is the only solution to retain strong trust relationship between the cloud users and cloud service providers. Thus to overcome the security threats, multiple cloud storage is enhanced. Thus the common forms of data storage such as files and databases of a specific user is split and stored in the various cloud storages (e.g. Cloud A and Cloud B).

Data Hierarchy

The layered model of access structure to solve the problem of multiple hierarchical files sharing is proposed. Layered model also improves the level of security at each layer and the data present will be highly authenticated. 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.

Experimental analysis

To protect patient data confidentiality, privacy preserving techniques are implemented to secure the PHI (Personal Health Information), and also to share the data to the admin. The layered model of access structure to solve the problem of multiple hierarchical files sharing is implemented in order to increase the efficiency of the system and also to improve the security constraints per layer of the structure. Admin plays a major role by handling all of the system data that is stored and retrieved every now and then. Hierarchical file sharing provides more security to the confidential information that is being stored in the system cloud. The files are encrypted with one integrated access structure which would reduce the encryption cost and increase the storage space. For every encrypted file a separate key is generated without which the file cannot be decrypted. The authorized duplicate check for this file can be performed by the user with the public cloud before uploading this file. The results of duplicate check is made by deduplication methodology which performs name check and also data check. Based on the results made from the data check the user uploads this file on the cloud or runs it directly on the system. The encrypted files will be uploaded into the cloud, if the user request match with the image then the file can be decrypted and downloaded. The private keys for privileges will not be issued to users directly, which will be kept and managed by the private cloud server instead. Image based authentication is enhanced.

IV. IMPLEMENTATION

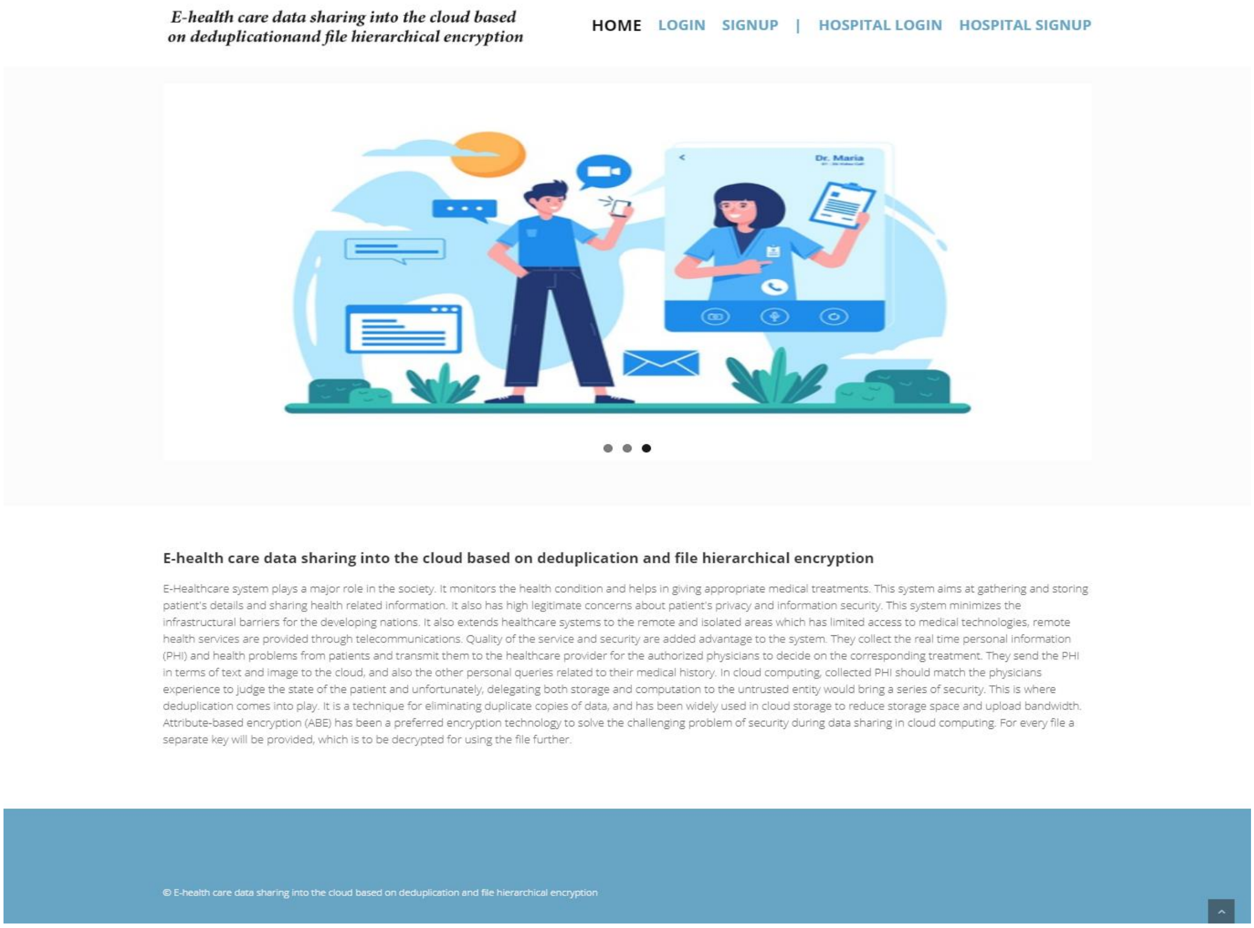
A web base platform implementor creates hypertext markup language (HTML), Common Gateway Interface (CGI) programs, and/or Java scripts and/or applets. The implementation process resembles software development because it involves using a specific syntax for creating hypertext structures in HTML or writing programming language code statements in computer files.

Html: The HyperText Markup Language, or HTML is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript. Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

CGI: stands for Common Gateway Interface and provides an interface between the HTTP server and programs generating web content. These programs are better known as CGI scripts. They are written in a scripting language. The Network Component provides such a scripting language.

Scripts: a computer programming language for adding dynamic capabilities to World Wide Web pages. Web pages marked up with HTML (hypertext markup language) or XML (extensible markup language) are largely static documents. Web scripting can add information to a page as a reader uses it or let the reader enter information that may, for example, be passed on to the order department of an online business. CGI (common gateway interface) provides one mechanism; it transmits requests and responses between the reader's Web browser and the Web server that provides the page. The CGI component on the server contains small programs called scripts that take information from the browser system or provide it for display. A simple script might ask the reader's name, determine the Internet address of the system that the reader uses, and print a greeting. Scripts may be written in any programming language, but, because they are generally simple text-processing routines, computer scripting languages such as PERL are particularly appropriate

V. SYSTEM EXECUTION DETAILS


E-health care data sharing into the cloud based on deduplicationand file hierarchical encryption
HOME
LOGIN
SIGNUP
|
HOSPITAL LOGIN
HOSPITAL SIGNUP
Dr. Maria

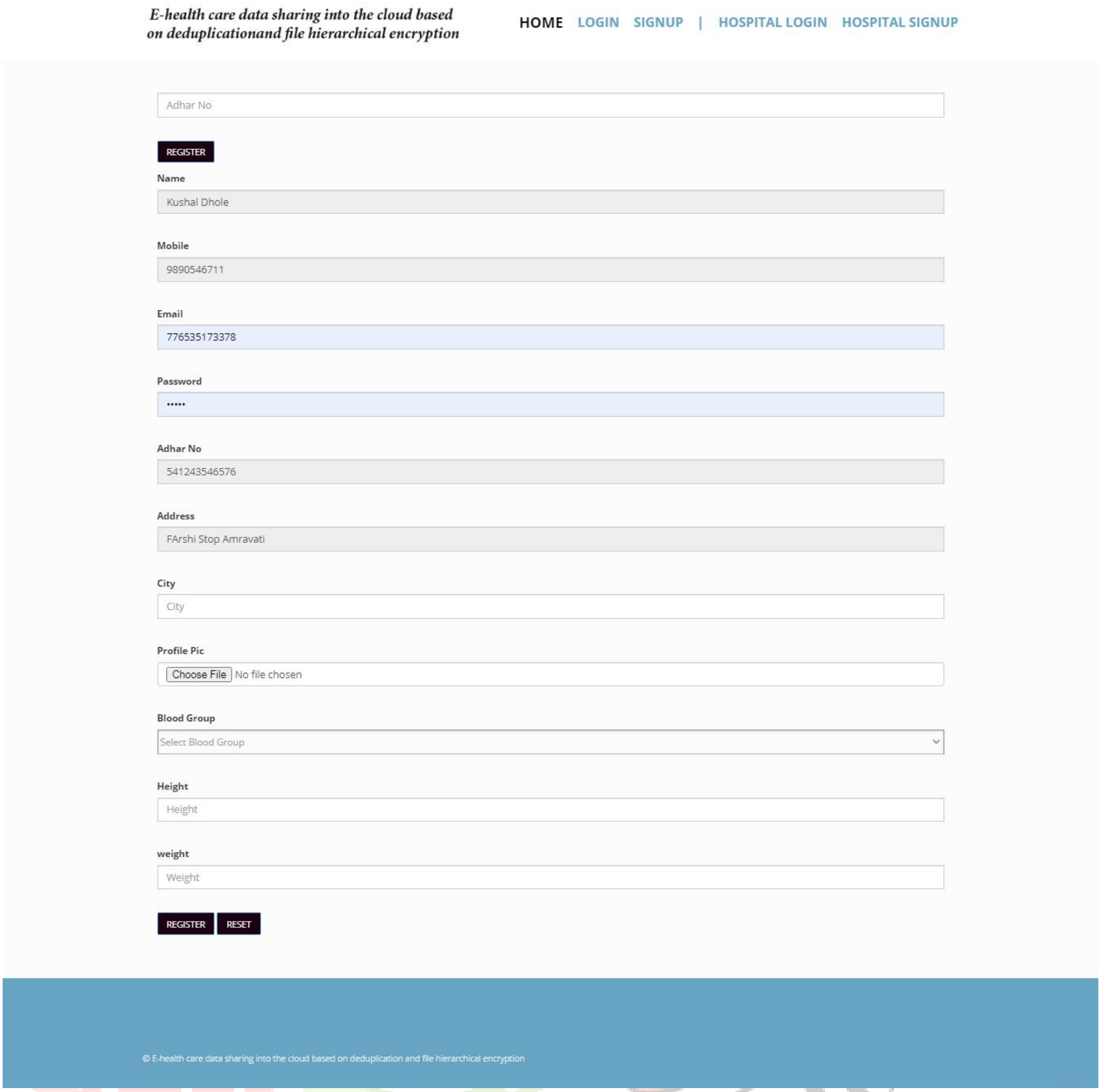
E-health care data sharing into the cloud based on deduplication and file hierarchical encryption

E-Healthcare system plays a major role in the society. It monitors the health condition and helps in giving appropriate medical treatments. This system aims at gathering and storing patient's details and sharing health related information. It also has high legitimate concerns about patient's privacy and information security. This system minimizes the infrastructural barriers for the developing nations. It also extends healthcare systems to the remote and isolated areas which has limited access to medical technologies, remote health services are provided through telecommunications. Quality of the service and security are added advantage to the system. They collect the real time personal information (PHI) and health problems from patients and transmit them to the healthcare provider for the authorized physicians to decide on the corresponding treatment. They send the PHI in terms of text and image to the cloud, and also the other personal queries related to their medical history. In cloud computing, collected PHI should match the physicians experience to judge the state of the patient and unfortunately, delegating both storage and computation to the untrusted entity would bring a series of security. This is where deduplication comes into play. It is a technique for eliminating duplicate copies of data, and has been widely used in cloud storage to reduce storage space and upload bandwidth. Attribute-based encryption (ABE) has been a preferred encryption technology to solve the challenging problem of security during data sharing in cloud computing. For every file a separate key will be provided, which is to be decrypted for using the file further.

© E-health care data sharing into the cloud based on deduplication and file hierarchical encryption

Screenshot 5.1:Home Page

- Screenshot shows home page which is index page of working system
- It consists of links in navigation bar such as login, signup, hospital login, hospital signup.


E-health care data sharing into the cloud based on deduplicationand file hierarchical encryption
HOME
LOGIN
SIGNUP
|
HOSPITAL LOGIN
HOSPITAL SIGNUP
Adhar No
REGISTER
Choose File
REGISTER
RESET
© E-health care data sharing into the cloud based on deduplication and file hierarchical encryption

Name: Kushal Dhole

Mobile: 9890546711

Email: 776535173378

Password: •••••

Adhar No: 541243546576

Address: FArshi Stop Amravati

City: City

Profile Pic: No file chosen

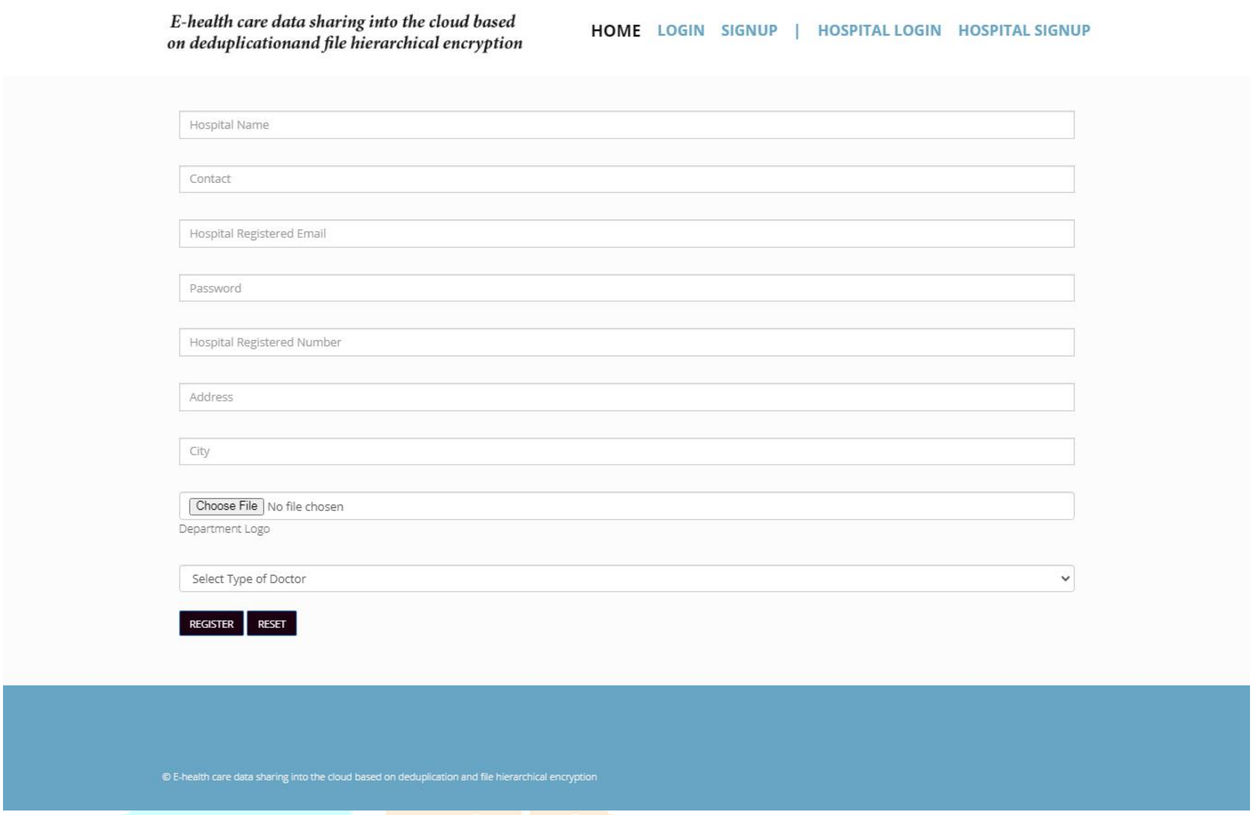
Blood Group: Select Blood Group

Height: Height

weight: Weight

Screenshot 5.2: Patient Register using Adhar Details

- Screenshot 5.2 shows registration form for new patients.
- only new patients are allowed for registration.


E-health care data sharing into the cloud based
on deduplicationand file hierarchical encryption
HOME
LOGIN
SIGNUP
|
HOSPITAL LOGIN
HOSPITAL SIGNUP
Choose File
REGISTER
RESET

Hospital Name

Contact

Hospital Registered Email

Password

Hospital Registered Number

Address

City

No file chosen

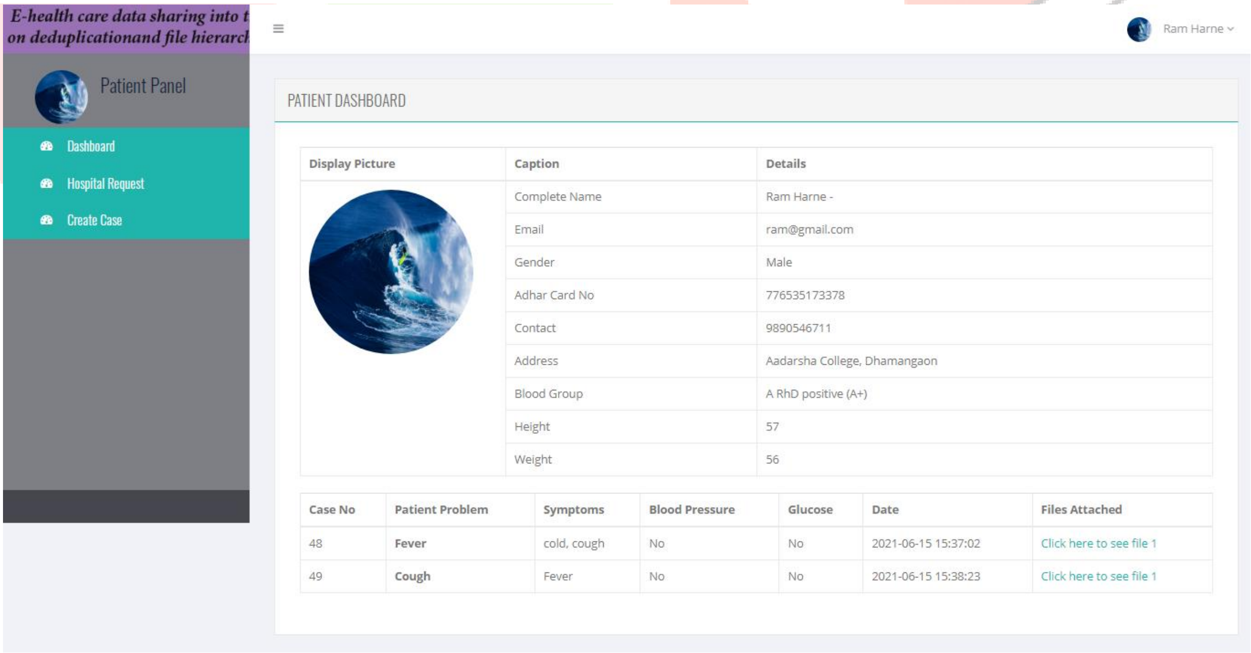
Department Logo

Select Type of Doctor

© E-health care data sharing into the cloud based on deduplication and file hierarchical encryption

Screenshot 5.3: Hospital registration

- Registration is required for those hospital which are new to the system new hospital.
- Registration details required hospital name, contact, email, hospital registered number and so on for registration of hospital to the system.


E-health care data sharing into t
on deduplicationand file hierarch
≡
Ram Harne
Patient Panel
Dashboard
Hospital Request
Create Case

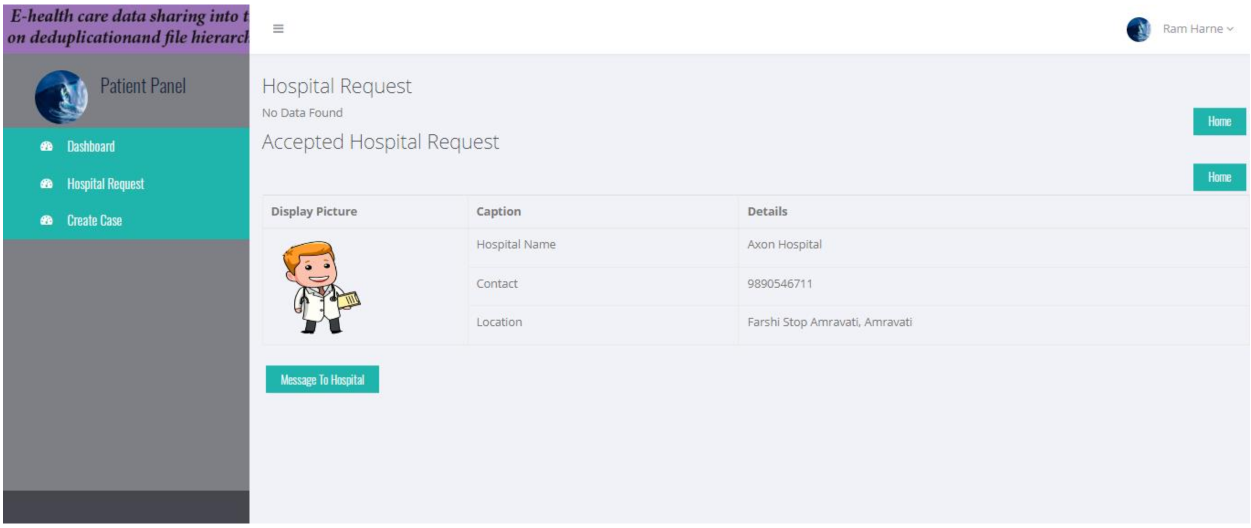
PATIENT DASHBOARD

Display Picture	Caption	Details
	Complete Name	Ram Harne -
	Email	ram@gmail.com
	Gender	Male
	Adhar Card No	776535173378
	Contact	9890546711
	Address	Aadarsha College, Dhamangaon
	Blood Group	A RhD positive (A+)
	Height	57
	Weight	56

Case No	Patient Problem	Symptoms	Blood Pressure	Glucose	Date	Files Attached
48	**Fever**	cold, cough	No	No	2021-06-15 15:37:02	Click here to see file 1
49	**Cough**	Fever	No	No	2021-06-15 15:38:23	Click here to see file 1

Screenshot 5.4: Patient dashboard

- Screenshot 5.4 shows Patient dashboard.
- It consists of patient personal details with patient's problem, symptoms and other factors.



Hospital Request

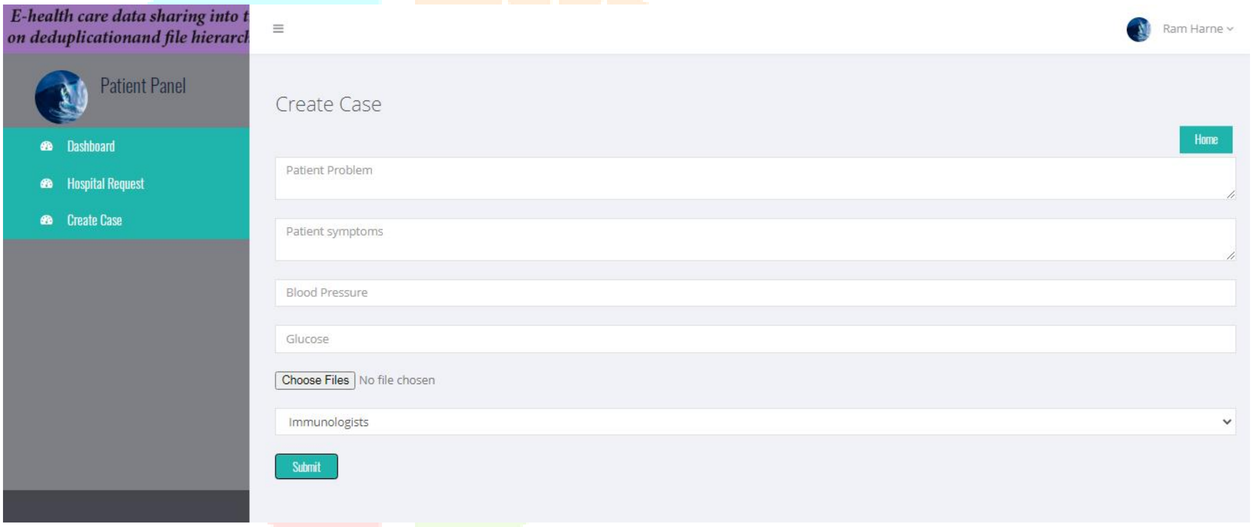
No Data Found

Accepted Hospital Request

Display Picture	Caption	Details
	Hospital Name	Axon Hospital
	Contact	9890546711
	Location	Farshi Stop Amravati, Amravati

Screenshot 5.5: Patient’s dashboard Hospital Request comes because of patient create own case

- When patient clicks on hospital request this dashboard gets opened.
- Patients send message to hospital by click on message to hospital.



Create Case

Patient Problem

Patient symptoms

Blood Pressure

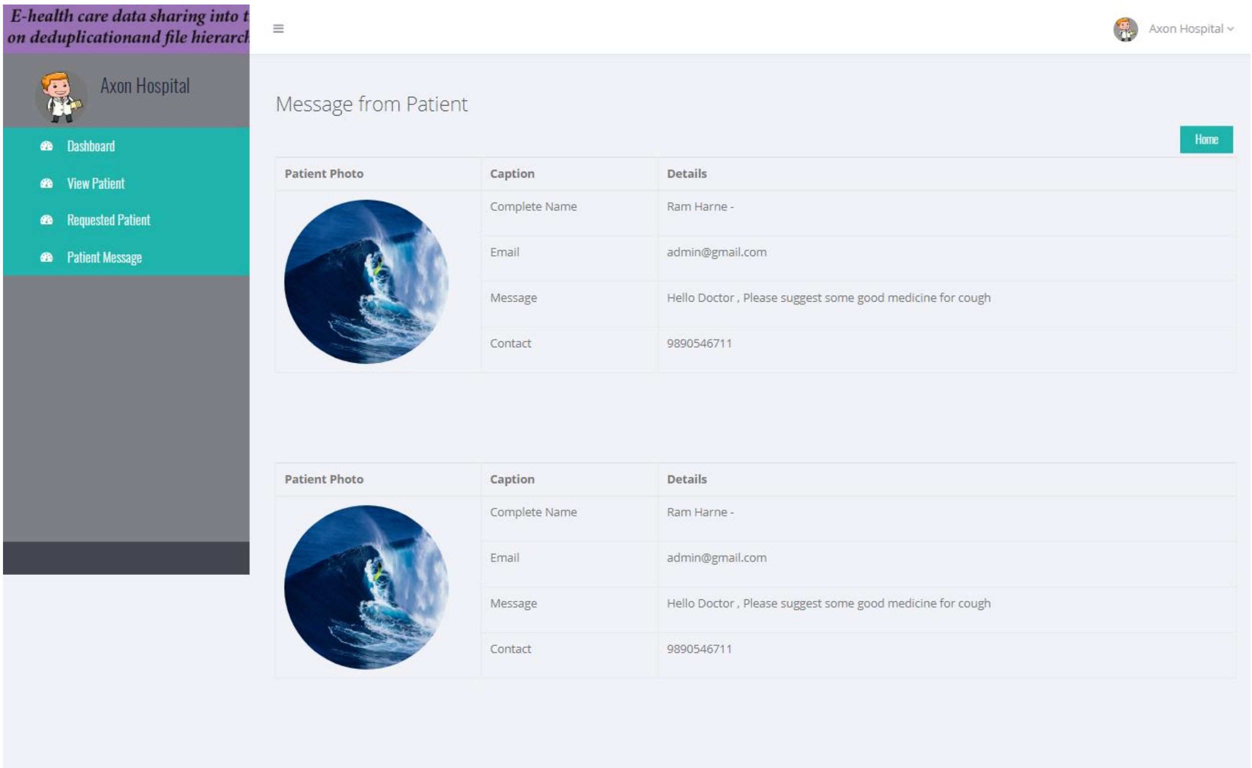
Glucose

No file chosen

Immunologists

Screenshot 5.6: Patient Case form here to generate encrypted file

- When patient clicks on create case, this form gets opened.
- It includes patient problems, symptoms, Bp and so on.
- When patient clicks on submit button , the information gets encrypted and stored in database.



Message from Patient

Patient Photo	Caption	Details
	Complete Name	Ram Harne -
	Email	admin@gmail.com
	Message	Hello Doctor , Please suggest some good medicine for cough
	Contact	9890546711

Patient Photo	Caption	Details
	Complete Name	Ram Harne -
	Email	admin@gmail.com
	Message	Hello Doctor , Please suggest some good medicine for cough
	Contact	9890546711

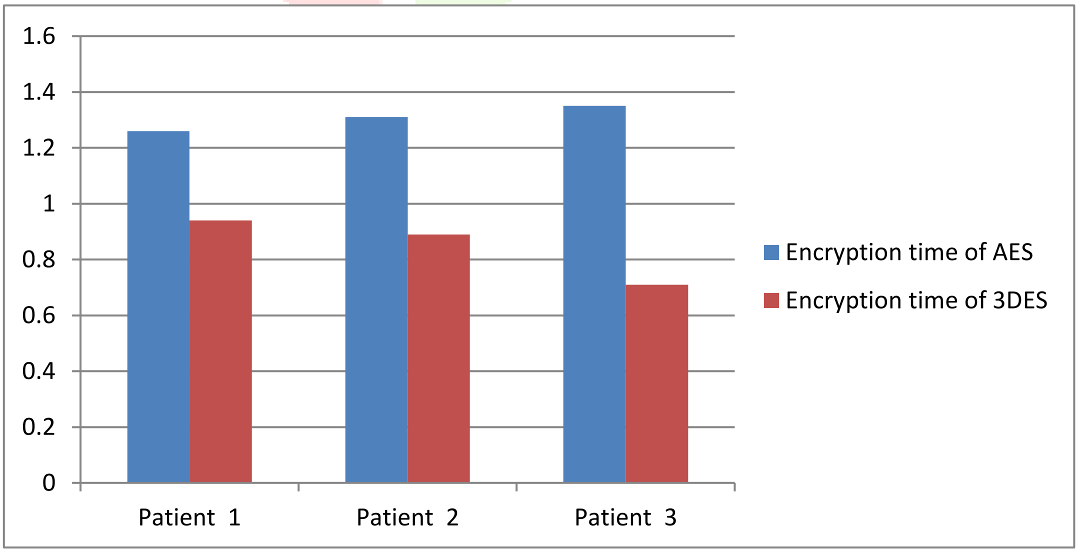
Screenshot 5.7: Patient Message to Hospital

- When hospital clicks on patient message this page gets opened.
- It consists of messages from patients.

VI . RESULT ANALYSIS

Table 6.1 Encryption times of AES and 3DES

patients	Encryption time of AES	Encryption time of 3DES
Patient 1	1.26 sec	0.94 sec
Patient 2	1.31 sec	0.89 sec
Patient 3	1.34 sec	0.71 sec



Graph 6.1: Analysis of AES and 3DES

The bar graph shows the performance of AES algorithm and 3DES algorithm which are implemented in the existing system of this dissertation report. The performances of both the algorithms are in terms of time required to encrypt the patient details. To compare the performance of both algorithms, we uploaded patient details of three patients i.e. by using AES and 3DES algorithm. According to above result AES takes more time than 3DES algorithm.

VII. ACKNOWLEDGMENT

It is my proud privilege and duty to acknowledge the kind of help and guidance received from several people in preparation of this report. It would not have been possible to prepare this report in this form without their valuable help, cooperation and guidance.

My sincere thanks to our **Dr. A. P. Bodkhe,** Principal, Prof. Ram Meghe Institute of Technology & Research, Badnera .

My sincere thanks to **Prof. Dr. G. R. Bamnote,** Head of Department, Computer Science & Engineering, for his valuable suggestions and guidance throughout the period of this report.

I express my sincere gratitude to my guide **Dr. S. R. Gupta,** Department of Computer Science and Engineering, for guiding me in investigations for this seminar. My numerous discussions with him was extremely helpful.

Last but not the least, I would like to thank all teaching and non-teaching staff and my colleagues for helping me directly and indirectly in preparation of this seminar report.

VIII. Conclusion

In this Paper, we have implemented E- Healthcare system which plays a major role in the society, monitors the health condition and helps in giving appropriate medical treatments. This system aims at gathering and storing patient's details and sharing health related information. It also extends healthcare systems to the remote and isolated areas which has limited access to medical technologies, remote health services are provided through telecommunications. Quality of the service and security are added advantage to the system. system collect the real time personal information (PHI) and health problems from patients and transmit them to the healthcare provider for the authorized physicians to decide on the corresponding treatment. In cloud computing, collected PHI should match the physicians experience to judge the state of the patient and unfortunately, delegating both storage and computation to the untrusted entity would bring a series of security. This is where deduplication comes into play. To overcome this problem we have implementedAES algorithm for securing the PHI such as Layered model of access structure which solves the problem of multiple hierarchical files sharing. Deduplication is implemented which allows only a single instance of a file to be save which saves memory wastage and time.

REFERENCES

[1]. "R. Shiny Sharon, Dr. R Joseph Manoj, "E-HEALTH CARE DATA SHARING INTO THE CLOUD BASED ON DEDUPLICATION AND FILE HIERARCHICAL ENCRYPTION", INTERNATIONAL CONFERENCE ON INFORMATION,COMMUNICATION & EMBEDDED SYSTEMS,2017.

[2]. "Private Data Deduplication Protocols in Cloud Storage" Wee Keong Ng SCE, Yonggang Wen SCE, Huafei Zhu

[3]. "DupLESS: Server-Aided Encryption for Deduplicated Storage" Mihir Bellare and Sriram Keelveedhi, University of California, San Diego; Thomas Ristenpart, University of Wisconsin—Madison.

[4]. "RevDedup: A Reverse Deduplication Storage System Optimized for Reads to Latest Backups" Chun-Ho Ng and Patrick P. C. Lee The Chinese University of Hong Kong, Hong Kong Technical Report June 28, 2013.

[5]. "A secure data deduplication scheme for cloud storage" Jan Stanek Alessandro Sorniotti, Elli Andreoulaki, Lukas Kencl.

[6]. "Dynamic Data Deduplication in Cloud Storage" Waraporn Leesakul, Paul Townend, Jie Xu School of Computing University of Leeds, Leeds, LS2 9JT United Kingdom.

[7]. Mrs. Suchitra Shelke, Prof. Babita Bhagat, "Techniques for Privacy Preservation in Data Mining", International Journal of Engineering Research & Technology, Vol. 4 Issue 10, October-2015.

[8] “Hierarchical Attribute-Based Encryption for FineGrained Access Control in Cloud Storage Services”Guojun Wang, Qin Liu School of Information Science and Engineering Central South University Changsha, Hunan Province, P. R. China, 410083 Jie Wu Dept. of Computer and Information Sciences Temple University Philadelphia, PA 19122, USA.

[9] “Scalable and Secure Sharing of Personal Health Records in Cloud Computing using Attribute-based Encryption” Ming Li Member, IEEE, Shucheng Yu, Member, IEEE, Yao Zheng, Student Member, IEEE, Kui Ren, Senior Member, IEEE, and Wenjing Lou, Senior Member, IEEE.

[10] “Privacy Preserving EHR SystemUsing Attribute-based Infrastructure” Shivaramakrishnan Narayan, Martin Gagné and Reihaneh Safavi-Naini Department of Computer Science University of Calgary, Alberta, Canada {snarayan,mgagne,rei}@ucalgary.ca

[11] R. Agrawal, R. Srikant, “Privacy-Preserving Data Mining”, ACM SIGMOD Record, New York, vol.29, no.2, pp.439-450,2000.

[12] H. Kargupta, S. Datta, Q. Wang, K. Sivakumar, “On the Privacy Preserving Properties of Random Data Perturbation Techniques”, In Proceedings of the 3rd International Conference on Data Mining, pp.99-106, 2003.

[13] Z. Huang, W. Du, B. Chen, “Deriving Private Information from Randomized Data”, In Proceedings of the ACM SIGMOD Conference on Management of Data, Baltimore, Maryland, USA, pp.37-48, 2005.

[14] D. Agrawal, C.C. Aggarwal, “On the Design and Quantification of Privacy Preserving Data Mining Algorithms”, In Proceedings of the 20th ACM SIGMOD-SIGACTSIGART Symposium on Principles of Database Systems, pp.247-255, 2001.

[15] S. L. Warner, “Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias”, J. Am. Stat. Assoc., vol.60, no.309, pp.63-69,1965.

[16] S.J. Rizvi, J.R. Haritsa, “Maintaining Data Privacy in Association Rule Mining”, In Proceedings the 28th VLDB conference, pp.1-12, 2002.

[17] W. Du, Z. Zhan, “Using Randomized Response Techniques for Privacy Preserving Data Mining”, In Proceedings 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.505-510, 2003.

[18] L. Guo, S. Guo, X. Wu, “Privacy Preserving Market Basket Data Analysis”, In Proceedings the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases.

[19]L. Sweeney, “K-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems”, 10 (5), 2002

[20] Pingshui WANG,” Survey on Privacy Preserving Data Mining”, International Journal of Digital Content Technology andits Applications, Volume 4, Number 9, December 2010.

[21] L.Sweeney, “Achieving k-Anonymity Privacy Protection Using Generalization and Suppression”, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, vol.10, no.5.

[22] A. Machanavajjhala, J.Gehrke, D. Kifer and M. Venkitasubramaniam, "IDiversity: Privacy Beyond kAnonymity", Proc. Int'l Con! Data Eng. (ICDE), p. 24, 2006.

[23] Slava Kisilevich, Lior Rokach, Yuval Elovici, Bracha Shapira, “Efficient Multi-Dimensional Suppression for K-Anonymity”, in proceedings of IEEE Transactions on Knowledge and Data Engineering, Vol. 22, No. 3. (March 2010), pp. 334-347, IEEE 2010.

[24]Martin Beck and Michael Marh¨ofer,” Privacy-Preserving Data Mining Demonstrator”, in proceedings of 16th International Conference on Intelligence in Next Generation Networks, IEEE 2012.

[25] A.S.Shanthi, , Dr. M. Karthikeyan” A Review on Privacy Preserving Data Mining “IEEE International Conference on Computational Intelligence and Computing, 2012 .

[26] Jianghong Wei; Wenfen Liu; Xuexian Hu” Secure Data Sharing in Cloud Computing Using Revocable-Storage Identity-Based Encryption” IEEE Transactions on Cloud Computing