Data Integrity and Privacy in Healthcare Management System: A Survey

Ketaki Deshmukh[1], Prof. Pramila M. Chawan[2]

[1]*M.Tech Student, Dept of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India*
[2]*Associate Professor, Dept of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India*

---***---

Abstract – *This article aims to discover the most secure ways to maintain the integrity and privacy in Healthcare Management System. To ensure integrity of the data we will be applying internal database constraints and restrictions. Moreover, Cross Object Resource Sharing (CORS) will be used to filter the devices which can access the project. Data privacy will be ensured by safeguarding sensitive data using International Data Encryption Algorithm (IDEA) along with salt. Moreover, session will be safeguarded using JWT Signature Algorithm HS256. SHA256 Hashing Algorithm along with salt will be used to safeguard passwords. In this article, the focus will be on implementing IDEA cryptographic algorithm for data privacy and understanding its pros and cons over AES-256 encryption algorithm.*

Key Words: Cross Object Resource Sharing (CORS), *International Data Encryption Algorithm (IDEA)*, *Advanced Encryption Standard* (AES), JWT Signature Algorithm HS256, SHA256 Hashing Algorithm, salt.

1. INTRODUCTION

An electronic health record comprises of an electronic version of a medical history of the patient as kept by the health care provider for some time period and it is inclusive of all the vital administrative clinical data that are in line to the care given to an individual by a particular provider such as demographics, progress reports, problems, medications, important signs, medical history, immunization reports, laboratory data and radiology reports. Health related data of individuals falls in the bracket of the most sensitive data. Privacy and security breaches are still common in this field even after having numerous policies and guidelines. Electronic healthcare data is quite prone to all types of attacks from outside as well as within the organization. Patient privacy is best protected by implementing a systematic mix of technologies and best practices such as technical de-identification of data, restrictive data access, and security measures in the underlying technical platforms. To remain effective, electronic health record system must satisfy some requirements such as achieving complete data, resilience to failure, be highly available and be consistent to security policies. . Presently, there are a lot of concerns regarding privacy and security of protected health data and these concerns are the biggest barriers in implementing electronic health records; and hence the need for health organizations to find out strategies that can help them secure electronic health records.

1.1 IDEA

The international data encryption algorithm abbreviated as IDEA is a symmetric block cipher data encryption protocol.

The key size of the block cipher is 128 bits and is regarded as a substantially secure and one of the best public standards. Of the numerous years, this protocol has been in the market, there is no single attack that has been published in spite of the numerous trials to identify them. Typically, the block cipher runs in round blocks. It applies fifty-two subkeys where each has a 16-bit length. Two subkeys are applied for a single round, four subkeys are applied prior to and after every round. Typically, both the plain text and the ciphertext have equal sizes of 16 bytes.

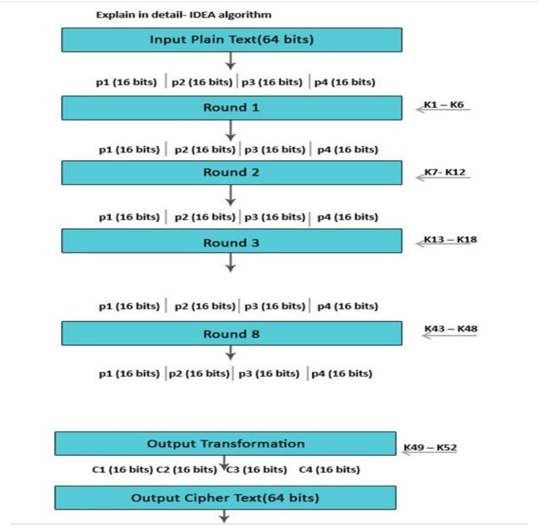

Explain in detail- IDEA algorithm
Input Plain Text(64 bits)
p1 (16 bits) | p2 (16 bits) | p3 (16 bits) | p4 (16 bits)
Round 1
K1 – K6
p1 (16 bits) | p2 (16 bits) | p3 (16 bits) | p4 (16 bits)
Round 2
K7- K12
p1 (16 bits) | p2 (16 bits) | p3 (16 bits) | p4 (16 bits)
Round 3
K13 – K18
p1 (16 bits) | p2 (16 bits) | p3 (16 bits) | p4 (16 bits)
Round 8
K43 – K48
p1 (16 bits) | p2 (16 bits) | p3 (16 bits) | p4 (16 bits)
Output Transformation
K49 – K52
C1 (16 bits) C2 (16 bits) C3 (16 bits) C4 (16 bits)
Output Cipher Text(64 bits)

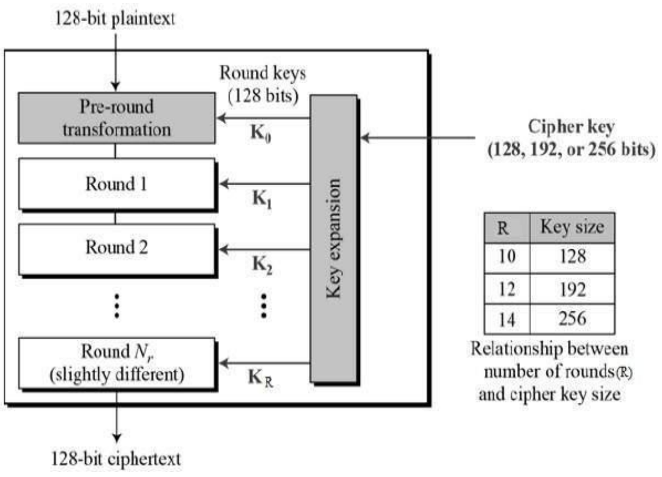
Fig 1: IDEA Algorithm

1.2 AES-256

AES is a symmetric key cipher. This means the same secret key is used for both encryption and decryption, and both the sender and receiver of the data need a copy of the key. By contrast, asymmetric key systems use a different key for each of the two processes. Asymmetric keys are best for external file transfers, whereas symmetric keys are better suited to internal encryption. The advantage of symmetric systems like AES is their speed. Because a symmetric key algorithm requires less computational power than an asymmetric one, it's faster and more efficient to run.

AES is also characterized as a block cipher. In this type of cipher, the information to be encrypted (known as plaintext) is divided into sections called blocks. AES uses a 128-bit block size, in which data is divided into a four-by-four array containing 16 bytes. Since there are eight bits per byte, the total in each block is 128 bits. The size of the encrypted data remains the same: 128 bits of plaintext yields 128 bits of ciphertext.

How does AES work? The basic principle of all encryption is that each unit of data is replaced by a different one according to the security key. More specifically, AES was designed as a substitution-permutation network. AES brings additional security because it uses a key expansion process in which the initial key is used to come up with a series of new keys called round keys. These round keys are generated over multiple rounds of modification, each of which makes it harder to break the encryption.



R	Key size
10	128
12	192
14	256

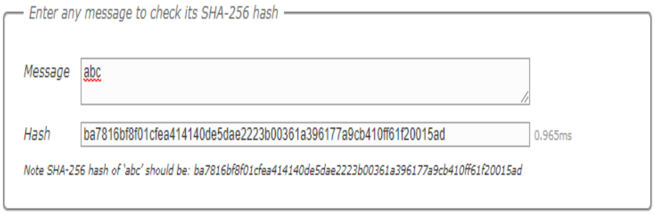
Fig 2: AES Encryption

1.3 Signature Algorithm: HS256

HS256 is a symmetric algorithm, with only one (secret) key that is shared between the two parties. Since the same key is used both to generate the signature and to validate it, care must be taken to ensure that the key is not compromised. If you will be developing the application consuming the JWTs, you can safely use HS256, because you will have control on who uses the secret keys. The issuer appends the JWT header and payload with the secret key, and hashes the result using SHA256, creating a signature. The recipient uses their copies of the secret key, JWT header and payload in the same way to reproduce the signature, checking to see if they match.

1.2 Hashing Algorithm: SHA256

A cryptographic hash (sometimes called 'digest') is a kind of 'signature' for a text or a data file. SHA-256 generates an almost-unique 256-bit (32-byte) signature for a text.



Enter any message to check its SHA-256 hash

Message: abc

Hash: ba7816bf8f01cfea414140de5dae2223b00361a396177a9cb410ff61f20015ad 0.965ms

Note SHA-256 hash of 'abc' should be: ba7816bf8f01cfea414140de5dae2223b00361a396177a9cb410ff61f20015ad

Fig3: Sample Hash Code

A hash is not 'encryption' – it cannot be decrypted back to the original text. SHA-256 is one of the strongest hash functions available.

2. Literature Review

Security and Privacy policies must be developed and coordinated publicly. Despite several privacy and security challenges, many technologies have already been implemented by the medical field. Numerous privacy and security frameworks are already in presence, and we must leverage the existing process as we utilize these standards in the field of healthcare IT. Fig. 4 gives an overview of these approaches. However, the widespread implementation of EHR and clinical data warehouses, as well as the latest attacks on patient data have reinforced the need for a new generation of data security approaches and solutions. The increasing volume of health data from various sources is stored fragmented across different locations and systems. Security flaws in any of these systems could cause the release of data to unauthorized people or organizations, and health information therefore need protection against unauthorized accesses and manipulations [8, 9]. EHRs additionally have issues in keeping up information security [10], to the extent that authoritative staff can for instance access information without patient consent [11]. As security and privacy of these systems has been a vital aspect in the design, implementation and management of the shared care paradigm, we integrated an information model that looks at the clinical applications and the underlying data warehouse of Houston Methodist's METEOR [12] that gives faster alerts to problems and is crucial to timely determination and minimization of the scope of damage.

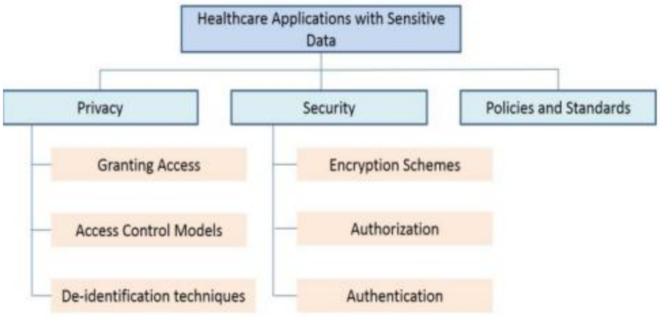


Fig 4: Categories of privacy and security approaches.

2.1 Methods:

The framework of Methodist Environment for Translational and Outcomes Research (METEOR) consists of two components: the enterprise data warehouse (EDW) and a software intelligence and analytics (SIA) layer that enables a wide range of clinical decision support (CDS) systems. Most of the fundamental requirements for protecting privacy and security are distinguished, and apply similarly to an EDW as to CDS systems: the application must stop unauthorized users from accessing or changing data; the applications and underlying information must not be subject to data-theft by hackers; the system must keep a record of actions performed by its users; and the data must be made accessible to the right users at the right time.

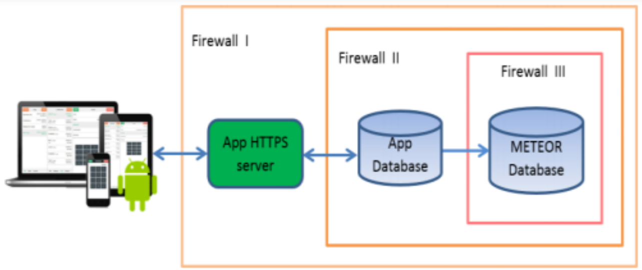


Fig 5: METEOR privacy and security information architecture

3. Proposed System

3.1 Problem Statement:

"A huge amount of sensitive data like mobile numbers, email ids and critical medical data of the patients is at stake once it's out on the web. Our project intends to safeguard the integrity and privacy of patient data using internal database security by implementing IDEA and compare it with the security and privacy that AES256 provides."

3.2 Problem Elaboration

Internal database constraints and restrictions. For eg: Admin even after having all the privileges won't be able access the personal medical reports of the patients. This features will be achieved by the use of a public and a private key for patient database. CORS(Cross Object Resource Sharing) configuration for filtering the devices accessing the project. Patient sensitive data will be safeguarded using International Data Encryption Standard Algorithm (IDEA) with salt. Session will be safeguarded using JSON Web Tokens which internally uses Signature Algorithm HS256. Passwords will be safeguarded using SHA256 Algorithm with salt. All the research done till date intends to safeguard Electronic Medical Data (EMD) from external theft. Providing Integrity and Privacy for sensitive data do not apply only for external intruders but also for the data leakage which can happen internally. For eg: Admin can get access of the medical reports of patients or a doctor can get access of the medical reports of a patient who is no longer under him. All these privacy breaches can also result in major breaches which were not taken care of in earlier researches. That's where our research will prove helpful.

3.3 Proposed System Architecture

The Model-View-Controller (MVC) is an architectural pattern that separates an application into three main logical components: the model, the view, and the controller. Each of these components are built to handle specific development aspects of an application. MVC is one of the most frequently used industry-standard web development framework to create scalable and extensible projects.

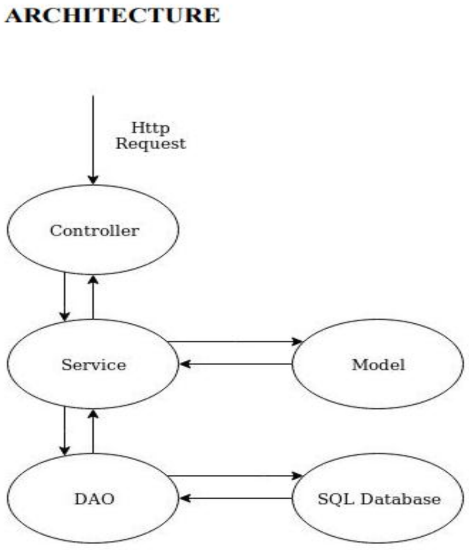


Fig 6: Proposed System Architecture: MVC

With DAO design pattern, we have the following components on which our design depends:

The model which is transferred from one layer to the other. A controller class consisting of the APIs. The service class consisting of business logic. The DAO class consisting of database queries.

4. CONCLUSIONS

In this paper, we have successfully found means to make the Electronic Medical Data (EDM) more secure on the web. We have learned the IDEA, AES256 and SHA256 algorithms to secure the sensitive data and passwords. We have seen the technologies which can ensure us to maintain the integrity and privacy of patient data on the web.

REFERENCES

[1] M. A. Rothstein, "Health privacy in the electronic age". The Journal of legal medicine, vol. 28, no. 4, pp. 487-501, 2007.

[2] M. Farzandipour, F. Sadoughi, M. Ahmadi, and I. Karimi, Security requirements and solutions in electronic health records: lessons learned from a comparative study. Journal of medical systems, vol. 34, no. 4, pp. 629-642, 2010.

[3] S. Haas, S. Wohlgemuth, I. Echizen, N. Sonehara, and G. Müller, Aspects of privacy for electronic health records. International journal of medical informatics, vol. 80, no. 2, pp. e26-e31, 2011.

[4] Olivier, M.S., Database privacy: balancing confidentiality, integrity and availability. ACM SIGKDD Explorations Newsletter, vol. 4, no. 2, pp.20-27, 2002.

[5] Fernández-Alemán, J.L., Señor, I.C., Lozoya, P.Á.O. and Toval, A., Security and privacy in electronic health records: A systematic literature review. Journal of biomedical informatics, vol. 46, no. 3, pp. 541-562, 2013.

[6] C. A. Gunter, Building a smarter health and wellness future: Privacy and security challenges. ICTs and the Health Sector: Towards Smarter Health and Wellness Models, pp. 141-157, 2013.

[7] A. F. Westin, Privacy and freedom. Washington and Lee Law Review, vol. 25, no. 1, pp. 166, 1968. [8] The New Threat: Attackers That Target Healthcare Organizations (And what you can do about it). http://www.infosecwriters.com/text_resources/pdf/New_Threat_Briga de.pdf

[9] D. Mellado, E. Fernández-Medina, and M. Piattini, Security requirements engineering framework for software product lines. Information and Software Technology, vol. 52, no. 10, pp. 1094- 1117, 2010. [10] L. S. Liu, P. C. Shih, and G. R. Hayes, Barriers to the adoption and use of personal health record systems. In Proceedings of the 2011 iConference. pp. 363-370, Feb 2011.

[11] R. Anderson, I. Brown, T. Dowty, P. Inglesant, W. Heath, and A. Sasse, Database state. Joseph Rowntree Reform Trust, York, 2009.

[12] M. Puppala, T. He, S. Chen, R. Ogunti, X. Yu, F. Li, et al., "METEOR: An Enterprise Health Informatics Environment to Support Evidencebased Medicine," IEEE Trans Biomed Eng, Jun 26 2015.

[13] Understanding Holistic Database Security: 8 Steps to Successfully Securing Enterprise Data Sources, IBM Corporation, 2012. http://www.ascent.co.za/documents/ibm/guardium/IMW14277USEN. pdf

[14] A. P. Deshmukh, and R. Qureshi. "Transparent Data Encryption-- Solution for Security of Database Contents." arXiv preprint arXiv:1303.0418, Mar 2013.

[15] S. Arora, J. Yttri, and W. Nilsen, Privacy and Security in Mobile Health (mHealth) Research. Alcohol Research: Current Reviews, vol. 36, no. 1, pp. 143-150, 2015.

[16] P. Varchol, D. Levický, and J. Juhar, Multimodal biometric authentication using speech and hand geometry fusion. In Systems, Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference on, pp. 57-60, Jun 2008.

BIOGRAPHIES

Ketaki Deshmukh, M.Tech Student, Dept of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India.

Prof. Pramila M. Chawan, is working as an Associate Professor in the Computer Engineering Department of VJTI, Mumbai. She has done her B.E. (Computer Engg.) and M.E (Computer Engineering) from VJTI COE, Mumbai University. She has 28 years of teaching experience and has guided 75+ M. Tech. projects and 100+ B. Tech. projects. She has published 99 papers in the International Journals, 21 papers in the National/ International conferences/ symposiums. She has worked as an Organizing Committee member for 13 International Conferences, one National Conference and 4 AICTE workshops. She has worked as NBA coordinator of Computer Engineering Department of VJTI for 6 years. She had written proposal for VJTI under TEQIP-I in June 2004 for creating Central Computing Facility at VJTI. Rs. Eight Crore (Rs. 8,00,00,000/-) were sanctioned by the World Bank on this proposal.