A Decisive Way to Secure Data Leakage from Intruders

[1]Amrapali Paul, [2]Rutuja Medankar, [3]Kajal Bambale, [4]Harshali Gosavi, [5]Prof. Mrs. S. R. Bhujbal

[1]Student, [2]Student, [3]Student, [4]Student, [5]Professor
Department of Computer Engineering
PK Technical Campus, Chakan, Pune, Maharashtra, India

***Abstract*: Statistics from security firms, research institutions, healthcare, and government organizations show that the number of data-leak instances has grown rapidly in recent years. Among various data leak cases, human mistakes are one of the main causes of data loss. There exist solutions detecting inadvertent sensitive data leaks caused by human mistakes and providing alerts for organizations. A common approach is to screen content in storage and transmission for exposed sensitive information. Such an approach usually requires the detection operation to be conducted in secrecy. A systematic and comprehensive review of security and privacy-preserving challenges in e-health solutions indicates various privacy-preserving approaches to ensure the privacy and security of electronic health records (EHRs). This paper highlights the research challenges and directions concerning cyber security to build a comprehensive security model for EHR. We surveyed, investigated, and reviewed various aspects of several articles and identified the following tasks: 1) EHR security and privacy; 2) security and privacy requirements of e-health data; 3) HER architecture, and; 4) diverse EHR cryptographic and non-cryptographic approaches. We also discuss some crucial issues and ample opportunities for advanced research related to the security and privacy of EHRs. Since big data provide a great mine of information and knowledge in e-Health applications, serious privacy and security challenges that require immediate attention exist. Studies must focus on efficient comprehensive security mechanisms for EHR and also explore techniques to maintain the integrity and confidentiality of patients' information.**

***Keywords*: Electronic Medical Records (EMR), Electronic Health Records (EHR), Personal Health Records (PHR), Denial of Service attacks (DoS), Electronic Health Data (EHD), Distributed Denial of Service (DDoS), Protected Health Information (PHI), Health Insurance Probability and Accountability Act (HIPAA), Data Privacy Preservation (DPP), Health Information System (HIS).**

I. Introduction

The beginning of the 21st century has witnessed great leaps in digital technology that are changing the landscape of the healthcare system across the world. There is a gradual and systematic transformation in healthcare systems from paper-based records to electronic records ushering in a revolution in the healthcare industry. This evolution converts paper-based records into digitalized electronic records such as Electronic Medical Records (EMR), Electronic Health Records (EHR), Personal Health Records (PHR), and Electronic Health Data (EHD). EHR and EMR are health records of patients handled by healthcare professionals, whereas PHR carries personal data which is handled and monitored either by the patient or their relatives regularly. EHD as electronic health records or computerized patient records is a systematized collection of smart health records of patients. The advantages of EHRs include easier and swift clinical data access, the ability to maintain effective clinical workflows, mitigation of medical errors, enhanced patient safety, reduced medical costs, and better and stronger support for clinical decision-making. Realizing the benefits offered by EHD systems more than 90 percent of healthcare institutions in the world have adopted this system to facilitate effective medical resource allocation and efficient healthcare. The large-scale proliferation of health information in the age of big data necessitates the burgeoning role of web networks for hosting unlimited amounts of data and easy access across the Internet. Moreover, the majority of medical data is highly sensitive and strictly confidential, its storage on third-party servers naturally increases these vulnerabilities. Generally, a patient may have several healthcare providers viz primary care physicians, therapists, specialists, and several insurer providers for medical, dental, vision, etc. According to the Health Insurance Portability and Accountability Act (HIPAA), it is the responsibility of healthcare providers to maintain the confidentiality of the health data.

II. LITERATURE SURVEY

1. Unwilling to reveal the sensitive data: The data owner may need to outsource the data-leak detection to providers but may be unwilling to reveal the sensitive data to them. Therefore, one needs new data leak detection solutions that allow the providers to scan content for leaks without learning the sensitive information. 2. Lack of transparency on the service provider: Users are not aware of the massive amount of data stored with the web service provider. It is difficult to be aware of where, how, and when the data is processed, making it difficult to trust the service provider, who can also be a reason for huge data loss. 3. Insider attacks: Web computing is a centralized mainframe computing paradigm owned by the provider which is less patient-centric and is prone to insider attacks that make the health records more vulnerable. 4. Do not support dynamic user management: Most efficient among encryption techniques and provide fine-grained, well-formed access to health records, it is still impractical for proper execution on EHRs due to its expensive computation, key management complexity, and challenges in managing access control policies when attributes in the access structure grow.

III. SYSTEM DESIGN

3.1 System Architecture

This scheme encrypts a single access structure in which the trusted authority (operator) will issue public and private key pairs. The proposed system developed as per patient and his disease details, operator and doctor can register and create their own user id and password. Operator also able to add, edit or delete disease details along with report attachments. As per patient's personal and disease details using Fuzzy fingerprint and Symmetric algorithm data saved in database in encrypted format.

System set access control and timestamp to protect hacking data form intruders. The access policy and time period is set by the information owner before outsourcing the data to the database. Fuzzy23keyword searches enrich system utility by providing matching files or nearest possible matching files for the user input with the predefined keywords based on keyword similarity semantics

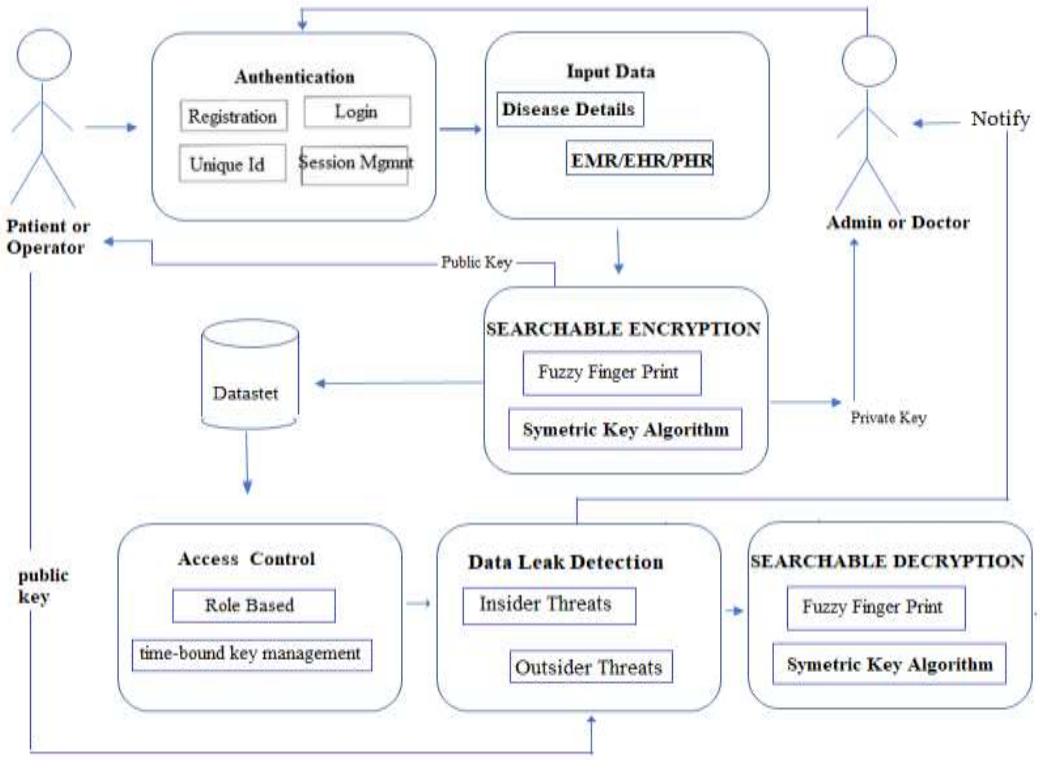

Figure 4.1: System Overview

Fig-1: Hardware Representation of Proposed System

1) MATHEMATICAL MODEL

• System Description:

Let q-grams on an input binary string. For example,
the 3-gram shingle set of string abcdefgh consists of six elements {abc, bcd, cde, def, efg, fgh}

* f= c1x ^(k – 1) + c2x^(k – 2) + ... + c(k−1)x + ck mod p(x)

Where,

* ci (0 <i< k), is one bit in the shingle
* q(x) = Each shingle is treated as a polynomial
* p(x) = q(x) is mod by a selected irreducible polynomial
* k = bit shingle
* f = bit fingerprint
* degree of p(x) is p^(f + 1).
* User
* Administrator/ Operator
* Searchable Symmetric Key

* Data Set
* Reports

3.2 EXTERNAL INTERFACE REQUIREMENTS

i. User Interface

The requirements section of hardware includes a minimum of 100 GB hard disk and 4 GB RAM with 1 GHz or higher speed.The primary requirements include a memory of 1 GB for the application of PHP and MySQL.The user interface of this program is the common web interface, nothing additional is required. The System user interface should be intuitive, such that 99.9% of all new system users are able to use the Proposed System application without any assistance. User registered into the system and log in then input to the system as a gesture and wait for a response for the system. A user interacts with the application.

2) Hardware Interface

To improve the quality of an image or to extract the required information from it, digital image processing can be used. The hardware should have the following specifications:

Ability to exchange data over the network. Keyboard for convenience. Continuous power supply. Ability to connect to the network. Ability to take input from the user. Ability to provide the public key to the user. Ability to validate the user.

3) Software Interface

The computer this software going to be installed on needs to have NetBeans IDE equal or above, Windows 7. On that Windows platform java, java version 1.8.* will be installed and that will be the platform the particular software will be run. There will be a java Xampp Server data transmission with the MySQL Server.

4) Communication Interface

Communication architecture defines the frequency and fidelity of information flow between individuals in your organization. It helps structure how and when you communicate, both within a team and cross-functionally. The specific tactics are unique to each organization, but it requires proactive thought and investment.

IV. IMPLEMENTATION

4.1 Results:

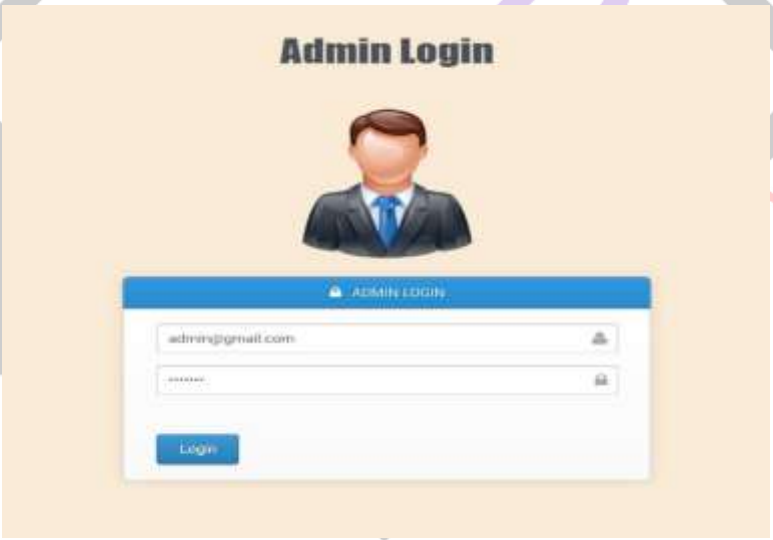

Fig-5: Home Page

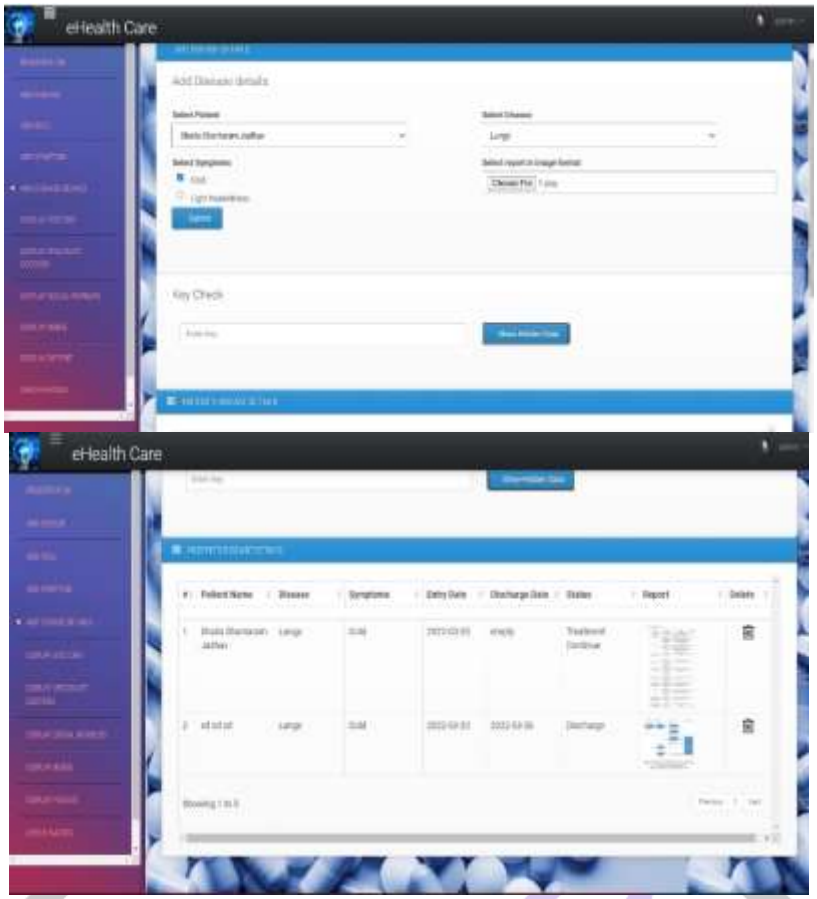

Fig-6: Add disease details to the patient

Fig-7: Display Intruders details

Fig-8: Registration

V. CONCLUSIONS

Smart health care services are a great boon and are dominantly used by patients, doctors, and other healthcare. Security and Privacy-Preserving Challenges of e-Health Solutions in Web Computing providers nowadays. Since the majority of data is stored in web servers, which are highly susceptible to threats and breaches, there is an imminent need to safeguard them from unauthorized access. Existing smart health solutions provide a certain level of immunity but are not a foolproof mechanism. In this context, a breakthrough in research to sustain the confidence and credibility of patients is essential for the wide-scale usage and success of digital health care. This review highlights a comprehensive study of existing e-health web preserving cryptographic and non-cryptographic mechanisms to secure privacy aspects in the web and their vulnerabilities in the fast-changing digital era. Moreover, our work also provides and identifies key research areas with diverse aspects viz architecture, encryption techniques, access control mechanisms and has also identified some remarkable research issues and future research directions to bring deliberate action for ensuring foolproof privacy in smart health solutions. The evolution of a holistic security mechanism as suggested by this work can make health care data more secure and sustainable.

REFERENCES

[1] Xiaokui Shu, Danfeng Yao, Member, IEEE, and Elisa Bertino, Fellow, IEEE, "Privacy-Preserving Detection of Sensitive Data Exposure" VOL. 10, NO. 5, MAY 2020.

[2] H. Cui, R. H. Deng, and Y. Li, ''Attribute-based storage with secure provenance over encrypted data,'' Future Gener. Comput. Syst., vol. 79, no. 2, pp. 461–472, Feb. 2018.

[3] Zhang, R. Xue, and L. Liu, ''Searchable encryption for health care: A survey,'' IEEE Trans. Services Comput., vol. 11, no. 6, pp. 978–996, Nov./Dec. 2017.

[4] C. S. Kruse, M. Mileski, A. G. Vijaykumar, S. V. Viswanathan, U. Suskandla, and Y. Chidambaram, ''Impact of electronic health records on long-term care facilities: Systematic review,'' JMIR Med. Inform., vol. 5, no. 3, p. e35, 2017.

[5] L. Griebel, H.-U. Prokosch, and F. Köpcke, D. Toddenroth, J. Christoph, I. Leb, I. Engel, and M. Sedlmayr, ''A scoping review of cloud computing in healthcare,'' BMC Med. Inform. Decis. Making, vol. 15, no. 1, p. 17, Mar. 2015.