Patient-Controlled Mechanism Using Pseudonymization Technique for Ensuring the Security and Privacy of Electronic Health Records

Bipin Kumar Rai, ABES Institute of Technology, India*
https://orcid.org/0000-0002-9834-8093

ABSTRACT
An internet-based electronic health record (EHR) system allows patients to access their medical history whenever they need it. Access to patient records and transactions related to diagnosis is helpful to patients and the healthcare departments and executives. But this practice may lead to major privacy concerns of patients’ private data. For EHR adaptation, the major elements are laws and regulations, monetary inducement and hurdles, technology state, and corporation effect. In this paper, the author has proposed a patient-controlled mechanism using the pseudonymization technique for ensuring the security and privacy of electronic health records. It is found that most of the potential approaches have used pseudonymization techniques to deal with the issues involved in a healthcare information system. This proposed solution is simple and efficiently ensures the privacy of patient data. Comparative analysis with other existing approaches has been undertaken.

KEYWORDS
EHR, Privacy, Pseudonymization, Security

INTRODUCTION
Currently, events are often covered in the public media, yet there is a lack of care regarding sensitive data. On the other hand, people tend to feel more concerned when their healthcare-related data is at risk, owing to the ease of envisioning reasons for abuse and comprehending the consequences of such misuse. Another obvious example is that almost everyone is presented with loan and insurance applications at some time in their lives. We can no longer dispute that privacy protection directly influences both personal well-being and society as a whole. Privacy is regarded as a fundamental human right.

There are currently no particular entities in India that pay close attention to the necessity of getting informed consent from subjects. As a result, most hospitals and clinics are overly cautious when analyzing their material since they are aware that the implications of the information included are pretty complicated; hence, there is a genuine risk that informed consent is ill-informed consent. Research ethics and security rules compel research units to devote increasing money and effort to privacy and...
identity protection. Yet, restrictive regulations controlling the transfer of medical information may discourage research needlessly. Therefore, a patient-controlled mechanism is required.

EHR systems are extremely craved for the structured unification of all pertinent medical data of an individual and to exhibit the lifelong medical record. Various confidentiality threats of healthcare data are crucial, either from within the institution or outside by some intruder. Each healthcare unit, hospitals and clinic have an information system for maintaining the patient’s data. Therefore, standards for data exchange are required, and electronic health records and data need to be standardized, including semantic interoperability (standards for the exchange of patient’s data among EHR systems). Several solutions are available to create EHR standards, such as openEHR, ANSI, Google Health and Microsoft’s HealthVault, HL7. Most of these services don’t provide complete control to the patients (Al-Hamdani, 2010). Innovative card healthcare systems developed in European countries are not strong privacy-preserving as anyone can access a patient’s information from a health card without their consent. Indivo is the first patient-controlled web-based healthcare system that provides options to own a secure complete medical record, integrating EHRs of different health centers. In Serbia, the architecture of the healthcare system is a hybrid intelligent card-based solution (Vućetić et al., 2011).

The whole patient’s experience of medical care is private. Hence providing confidentiality of medicine prescriptions is an important one (Ateniese & de Medeiros, 2002). In a smart-card-based e-prescriptions system, both patients and doctors have security concerns with this e-prescription data. Other parties are involved, and some parties may use it for their benefits like marketing, etc. (Yang et al., 2004).

Access control mechanisms and applications related to e-prescription systems and other consumer-related healthcare services require a secure mechanism (Rai & Solanki, 2021). Blockchain technology seems to be more appealing in healthcare (Mayer et al., 2020).

The objectives of the paper can be summarized as follows:

- How pseudonymization technique can solve privacy and security issues in the healthcare industry today.
- A patient-controlled architecture is proposed, which will be most suitable for health care information systems.
- An efficient system model of separate storage of patient personal data and pseudonymized health data.
- Furthermore, we show the performance analysis of the proposed scheme with existing solutions.

SECURITY CHALLENGES IN HEALTHCARE

We need to handle the following security issues properly while accessing EHR (Rai & Srivastava, 2014):

1. **User Authentication**: Only approved users will have the option to get access to the health record.
2. **Confidentiality and Integrity**: It protects medical data from unauthorized access and reliability of healthcare information systems.
3. **Data Ownership**: It is additionally a significant issue associated with the ability to access medical data. Obligations of information possession ought to be handled straightforwardly.
4. **Access Control**: The objectives of the access control are protecting any Information system from unauthorized access and at the same time making it available to authorized users. Electronic Health Records recommends that information systems develop a robust mechanism for protecting the unauthorized access of the data (Byers et al., 2002).

Pseudonymization-Based Mechanisms for Healthcare

Several mechanisms are available for ensuring security and privacy issues related to healthcare by pseudonymization technique (Neubauer & Kolb, 2009).
Pseudonymization is most suitable for a healthcare information system similar to anonymization (Neubauer & Kolb, 2009; Rai & Srivastava, 2016). The only difference is that identifying information is separated from the health records, referenced by pseudonym (a unique random number) but not permanently deleted. Therefore, pseudonymization is a Patient’s controlled reversible process under specified circumstances. It is also applied in several applications like healthcare information systems (Bruland et al., 2018).

Pseudonyms are secret random numbers commonly used as links between patients and health records, with the links only being recoverable when authorized (Amin et al., 2019).

Peterson approach (Peterson, 2003) asks users to register on a service provider’s website. After registration, they are given a unique Global key (GK) and server-side key (SSID). A unique “personal encryption key (PEK)” and password must also be provided. GK is included on the ID card. By entering GK and PEK, the user retrieves data from the database.

(Slamanig & Stingl, 2008) proposed a hull architecture, here instead of storing the relation between patients and their dataset in a centralized manner, all data are stored in two separate databases. One store’s plaintext pseudonyms and related medical datasets in plaintext for performance reasons. Another database is used to store the personal information of users as well as their encrypted pseudonyms.

Electronic health card (eGK) is designed as service-oriented architecture (SOA) having some restrictions like local card access only, supported by the Ministry of Health, Germany (Zhang et al., 2009). Its architecture consists of five layers. The first layer is the Presentation layer which provides a communication interface to the user. The second layer offers different services; the third layer is a Business layer which combines various services. The fourth layer, the application layer, manages data and user rights. The last layer is the infrastructure layer.

(Thielscher et al., 2005) proposed a solution in which “identification data” and medical record anamnesis data are stored in separate databases. In Pommerening (K & M, 2004) approach, different ways for secondary use of healthcare data have been recommended. (Kushida et al., 2012) highlighted preventive steps for patient anonymity and untraceability.

**PATIENT-CONTROLLED PSEUDONYMIZATION BASED EHR**

This Proposed Patient Controlled Pseudonymization Based EHR provides convenient, simple privacy preservation. This security mechanism enables the availability of patients’ health information to any healthcare entity at any time with the consent of the patient.

**The Architecture of the PcPbEHR System**

(Bacelar-Silva et al., 2011) show that different countries have different choices based on community needs, but most popular EHR solutions assert patient-centred because they give the patient total access rights. This PcPbEHR system is patient-centered, which combine pseudonymization techniques with an encryption mechanism to create an efficient mechanism for healthcare information system security and privacy (Rai, n.d.).

In Figure 1, it is demonstrated that a pseudonym is used to protect the patient’s privacy. The patient will create this pseudonym. The patient-controlled access control policy will determine what and to whom the portion of his data may be accessed.

Pseudonymization of the sensitive data records supports both privacy-preserving primary and secondary usage as long as the records are depersonalized.

The Architecture of the PcPbEHR system consists of two separate databases PcPbEHR Health Records database and the PcPbEHR Patient Profile database, a Pseudonymization module, an Access control Module for Patients and Access control Module for Doctor/healthcare authority, and data access policy and sharing management as shown in Figure 1.
PcPbEHR Databases

If any intruder successfully accessed the database, then the patient’s privacy may be compromised; hence to ensure strong privacy PcPbEHR maintain two separate databases, one for identifiable information in encrypted form and another for pseudonymized health records: PcPbEHR Health Records database, which contains patient’s health records after pseudonymization. PcPbEHR Patient Profile database, which includes encrypted patient’s profiles and encrypted patients’ pseudonyms.

Pseudonymization Module

Before storing health records from the patient/healthcare center into the PcPbEHR Health Records database, the pseudonymization module removes all identifiers and quasi-identifiers from the patient’s health record. If the intruder gets access to the database, he will not determine a particular health record owner.

Each patient can create a unique pseudonym (digital long random number) using the Pseudonym Generation mechanism. A pseudonym can be generated locally in her environment without any information exchange between PcPbEHR and the Patient.

Pseudonyms cannot be guessed from patients’ information and need not remember. A pseudonym is stored in encrypted form. Patients decrypt it when a new record is added. A pseudonym is appended to all patients’ health records before storing the PcPbEHR Health Records database, as shown in Figure 2.

Patient's Profile and Pseudonym Encryption

His public key encrypts the patient’s pseudonym (using the public-key cryptography technique). Several fields have been identified as personally identifiable information, e.g. name, date of birth,
age, mobile number, AADHAR Number, email, etc. This identifiable information is needed when the patient visits a new healthcare center first time. All this identifiable information are encrypted by a shared key (using symmetric essential cryptography technique) shown in Figure 3.

Encrypted profiles and encrypted pseudonyms are stored in a secure PcPbEHR Patient Profile database. A pseudonym is decrypted by the patient’s private key known to patients only for the addition of a new record.

**PcPbEHR_Pseudonym_Generator()**

The patient will use the following algorithm to generate a unique pseudonym. The AADHAR, a unique identification number in India, will be used in this process. In the case of another country, any unique identification number can be used in place of the AADHAR. The patient’s AADHAR will be verified firstly, as shown in Figure 4.

**Pseudonym_Generator()**

1. Select two random prime no. p, q
2. by using RSA algorithm, compute \{e, n\} public key and \{d, n\} private key.
3. Compute C = E((AADHAR No ||DOB|| Pad), (d, n)) where E stands for encryption
4. Compute H=SHA-1(C). where SHA-1 is the standard hash function
5. Pseudonym P =ENCODING-64(H)

Return P

PcPbEHR system will accept a new pseudonym only when that is not in use by other patients. A pseudonym is visible to the patient and used by the system only.
Access Control Module

Several access controls models have been discussed. RBAC is the most frequently used model with some variations. Every entity of the healthcare system has different access requirements that the access control module should handle without compromising patient privacy. Each entity of the healthcare system, i.e., patient, doctor, and health centers/health authority, will register in the PcPbEHR system. AADHAR number will be used for verification of each entity. Hence the identity of the accessing applicant is verified. The patient can access all the health records containing his pseudonym. The
patient needs to decrypt their pseudonym using their private key, known only to the patient. Hence privacy is maintained. Figure 5 shows this operation.

Doctors/healthcare authorities have limited access rights. They can access only those health records or some fields of the patients’ health records.

Firstly, a patient needs to decrypt their pseudonym using their private key. This private key is known only to the patient. Then the doctor/healthcare authority will pass his pseudonym by following the same step of decryption. With both pseudonyms of patient and pseudonym of doctor/healthcare authority, the access control module will permit to access those health records. Hence it will be validated who accessed the health records. Figure 6 shows this operation.

**Figure 5. Accessing Patient’s health record mechanism**

**Figure 6. Granting Permission to Doctor/ Healthcare Authority to access Patients’ health record**
Addition of Patients’ New Record

When a patient visits a healthcare center or in some other situation and new health records are generated, he can store those records either by directly entering the data or can send to the system. Henceforth system then removes all identifiable information from new arriving health records and is intimate to the owner. The patient checks the source of the data, the validity of data. If the patient accepts, then the system asks his private to decrypt his pseudonym. A patient pseudonym will be appended to the arrival document, and then pseudonymized data will be stored in the PcPbEHR Health Records database. Figure 7 shows this whole process.

Implementation of the PcPbEHR System

In this prototype implementation, we use AADHAR NUMBER, a unique identification number in India, to generate an individual pseudonym. It can be replaced by another national level Identification Number in other countries. HTML, CSS, BOOTSTRAP, PHP and MySQL have been used to develop this prototype. Following are the steps involved in this implementation.

- **Registration:** Any patient can register by entering personal and medical information. The patient will enter their login id and password, which will be used for future activities. The UIDAI portal will verify the AADHAR number. The patient will fill several required fields of information at the time of registration shown in figure 8.

---

Figure 7. Addition of new records in PcPbEHR Health Records Database
Several fields have been identified as personally identifiable information, e.g., name, date of birth, age, mobile number, AADHAR Number, email, etc. All this identifiable information is encrypted. Encrypted profiles and encrypted pseudonyms are stored in a secure PcPbEHR Patient Profile database.

A pseudonym is decrypted by the patient’s private key known to patients only for the addition of a new record.

- **Login**: Any registered user can log in to availing of health services. Smart health IC cards are highly recommended for the deployment of this proposed solution.

But for simplicity, only user name and password-based authentication are implemented. It is shown in Figure 9.
- **Pseudonymization Module:** Before storing health records from the patient/healthcare center into the PcPbEHR Health Records database, the pseudonymization module removes all identifiers and quasi-identifiers from the patient’s health record. Patients can create a unique pseudonym (digital long random number) using the Pseudonymization module. A pseudonym can be generated locally in her environment without any information exchange between PcPbEHR and the Patient, shown in Figure 10.

- **Access control Module:** Patient enters their private key to decrypt the pseudonym. This private key is known only to the patient. The patient can access all the health records containing his pseudonym shown in Figures 11 and 12.

---

**Figure 10. Screenshot of Pseudonymization**

**Figure 11. Screenshot of Access control module (for the patient)**
Doctor/healthcare authority can access only those health records or some fields of the health records allowed by patients, shown in Figure 13.

After decryption of the patient pseudonym, the doctor/healthcare authority will pass his pseudonym by following the same decryption step. With both pseudonyms of patient and pseudonym of doctor/healthcare authority, the access control module will permit to access those health records to doctor/healthcare authority.

ANALYSIS OF THE PROPOSED PCPB EHRS SYSTEM

Significance of Proposed Solution

This prototype solution can help for the development of a national-level healthcare information system. AADHAR database can be used for the authorization of the patient. Furthermore, a separate study
of social and geographical scenarios will be needed for actual implementation, which is beyond the scope of this work. It can be used for other countries as well.

**Strong Authentication Mechanism**

Each entity of the healthcare system, i.e., patient, doctor, and health centers/health authority, will register in the PcPbEHR system. AADHAR number will be used for verification of each entity. Hence the identity of the accessing applicant is verified.

The unique pseudonym of the patient is encrypted by their private key known to him only. Patients’ health records, consisting of his pseudonym, are stored in a separate PcPbEHR Health Records database.

Hence only the patient can view his health records. If another malicious intruder duplicates a patient’s identity, he cannot view health records without a private key known to the patient only. If any intruder gets full access to the PcPbEHR health records database, the relationship between pseudonyms and the secret message is known only to the patient. Hence the intruder cannot know the actual owner of the pseudonym.

**Privacy-Preserving**

The patient personal profile is stored in a separate PcPbEHR Patient Profile database in encrypted form. The unique pseudonym of the patient is also encrypted by their private key known to him only. Patients’ health records, consisting of his pseudonym, are stored in a separate PcPbEHR Health Records database. No editing in pseudonyms is possible by the health authority. Even a patient cannot modify his pseudonym once it is stored. Doctor/healthcare authority cannot access patients’ health records until and unless patients enter their private key known to them only. Even if some basic information is known about the patient, one cannot find the pseudonym.

**Confidentiality**

If any intruder becomes successful in accessing the healthcare database, then the patient’s trust is compromised. It is the most significant bottleneck for the adoption of EHR in several countries. Hence to ensure the confidence of the Patient, PcPbEHR maintains two separate databases. PcPbEHR Health Records database contains patient’s health records after pseudonymization and PCPbEHR Patient Profile database, which includes encrypted patient’s profile and encrypted patient’s pseudonym.

Some identifiable information is needed when a patient visits a new healthcare center for the first time. A shared key encrypts all this identifiable information (using symmetric essential cryptography technique). Their public key encrypts the patient’s pseudonym (using the public-key cryptography technique). Encrypted profiles and encrypted pseudonyms are stored in a secure PCPbEHR Patient Profile database.

No intruder can know the pseudonym of the patient and their profile data even if s/he gets full access to the PcPbEHR health records database because the relationship between pseudonym and the secret message is known to the patient only. Hence intruder can not know the actual owner of the pseudonym. For the addition of a new record, a pseudonym is decrypted by the patient’s private key known to the patient only.

**Secure and Control Accessing of Health Records**

Patients can access all the health records containing their pseudonyms. Firstly, the patient needs to decrypt their pseudonym using their private key. This private key is known only to the patient. Doctor/healthcare authority has limited access rights. They can access only those health records or some fields of the health records allowed by the patient by entering their private key.

Both pseudonyms of patient and pseudonym of doctor/healthcare authority are needed to access those health records. Hence it will be validated as to who accessed the health records.
Comparison with other Solutions

Several solutions have been proposed for categories of patient-controlled mechanisms. I found that most of the potential approaches have used pseudonymization techniques to deal with issues in the healthcare information system. This proposed solution is simple and efficiently ensures the privacy of patient data. Comparative analysis with other existing approaches has been undertaken and shown in table 1.

CONCLUSION

This paper proposes a patient-centered system to provide a convenient, simple, privacy-preserving, secure mechanism. This approach ensures the privacy of patient data and the patient’s complete control over their data. Instead of using unnecessary encryption/decryption, this method employed a unique pseudonymized strategy for efficient data access and privacy. This will also limit internal misuse because pseudonymized data would be kept, ensuring that no administrator entity may access a patient’s health data. It will enable the availability of a patient’s health information to any healthcare entity at any time with the consent of the patient. This proposed solution provides a considerable possibility for medical research on a particular disease using anonymous health data. It can give anonymous health data which will not have any identifiable information of the patient to researchers. Hence without compromising privacy, anonymous health data will be available for the researchers.

Blockchain technology provides a secure way to share health data and also offers improved healthcare transactions. At present, blockchain technology is not considered much crucial in the healthcare industry; however, it will soon come from numerous points of view as it is a definitive resource tracker (Shahnaz et al., 2019). This blockchain technology-based way to deal with the healthcare industry would allow users to securely transfer medical data (Khezr et al., 2019). This also provides users with a way to share their medical data for medical research anonymously. (Chenthara et al., 2020; Gordon & Catalini, 2018) Those systems empowered by blockchain technology can significantly decrease the expense and the grating of current intermediates.
REFERENCES


---

*Bipin Kumar Rai is Professor in the Information Technology Department, ABES Institute of Technology, Ghaziabad (India). He is B. Tech.(CSE) from UPTU Lucknow, India, M. Tech(CSE) from RGPV Bhopal, India and Ph.D. from Banasthali University, Rajasthan, India. He has more than 16 years of teaching experience. He has published three books, two patents and authored many research papers in the reputed Journals/Conferences.*