A Privacy-Preserving Framework for Personally Controlled Electronic Health Record (PCEHR) System

Mahmuda Begum
*Charles Sturt University, mbegum@csu.edu.au*

Quazi Mamun
*Charles Sturt University, qmamun@csu.edu.au*

Mohammed Kaosar
*Charles Sturt University, kaosar@csu.edu.au*
A PRIVACY-PRESERVING FRAMEWORK FOR PERSONALLY CONTROLLED ELECTRONIC HEALTH RECORD (PCEHR) SYSTEM

Mahmuda Begum¹, Quazi Mamun², and Mohammed Kaosar³
¹³School of Computing and Mathematics, Charles Sturt University, Australia
e-mail: {¹mbegum, ²qmamun, ³kaosar}@csu.edu.au

Abstract
The electronic health record (eHR) system has recently been considered one of the biggest advancements in healthcare services. A personally controlled electronic health record (PCEHR) system is proposed by the Australian government to make the health system more agile, secure, and sustainable. Although the PCEHR system claims the electronic health records can be controlled by the patients, healthcare professionals and database/system operators may assist in disclosing the patients' eHRs for retaliation or other ill purposes. As the conventional methods for preserving the privacy of eHRs solely trust the system operators, these data are vulnerable to be exploited by the authorised personnel in an immoral/unethical way. Furthermore, issues such as the sheer number of eHRs, their sensitive nature, flexible access, and efficient user revocation have remained the most important challenges towards fine-grained, cryptographically enforced data access control. In this paper we propose a patient centric cloud-based PCEHR framework, which employs a homomorphic encryption technique in storing the eHRs. The proposed system ensures the control of both access and privacy of eHRs stored in the cloud database.

Keywords
Ehealth, Electronic Health Record, PCEHR, Homomorphic Encryption.

INTRODUCTION

eHealth has recently been considered a precipitately changing segment of the healthcare industry. eHealth is defined in many ways such as the transfer of health resources and healthcare by electronic means. According to The World Health Organisation (WHO), eHealth is defined as the combined use of electronic communication and information technology in the health sector (Eysenbach, 2001). The recently proposed Australian government’s personally controlled electronic health record (PCEHR) system is one of the best examples of eHealth system implementation (National E-Health Transition Authority Australia, 2011).

The electronic healthcare record (eHR) is the principle aspect of an eHealth system such as PCEHR. eHR is the digitally stored healthcare information about an individual's lifetime with the purpose of supporting continuity of care, education and research, and ensuring confidentiality at all times (Gajanayake, Iannella, & Sahama, 2012). eHRs aid efficient communication of medical data and thus ease organisational disbursements with the help of cloud computing (Muhammad, Zwicker, & Wickramasinghe, 2013, Wu, Ahn, & Hu, 2012).

However, privacy in particular, has always been one of the main concerns in eHealth systems (Petkovic & Ibraimi, 2011, Xanthidis & Aleisa, 2012). According to Richard Rognehaugh, privacy is the right of individuals to keep information about themselves from being disclosed to others (Rognehaugh, 1999). The information that is shared as a result of a clinical relationship is considered confidential and must be protected (Rinehart-Thompson & Harman, 2006).

A patient’s health information may contain sensitive information such as sexual health, mental health, addictions to drug or alcohol, abortions, etc. This makes such a patient demand strong privacy for their eHR system. It is important to note that in the digital environment, disclosed private information cannot be recovered and will last indefinitely. Not giving patients control over their private data might result in patients withholding or trying to delete sensitive medical information from their eHRs in order to preserve their privacy (Klitzman, 2006). Yet again, privacy risks can also arise from health care professionals. Due to the complexity of eHR systems, healthcare professionals can make innocent mistakes to cause disclosure of a patient’s information.

On the other hand, a system operator may intentionally leak out patients’ information for revenge, spite, profit, or other ill purposes. Risks from inadvertent or intentional release of infectious, mental health, chronic disease diagnoses, and genetic information are all well recognized both online and in mass media.
In the conventional privacy preserving techniques, system operators are assumed to be trusted. But in some cases, they may not be reliable. Therefore, we need to construct such a system to eliminate the above assumption.

In addition, an eHR system needs to be able to deal effectively with a very high volume of patients’ sensitive data along with ensuring user authentication, role based access control, and patients’ authorisation. Thus, a multi-level security system is required to protect the privacy of eHRs.

To address all the above mentioned issues, in this paper, we propose a framework to access patients’ eHRs. This patient centric framework employs a homomorphic encryption technique in storing and updating the eHRs. The encryption system allows computation on cipher text, thus eliminates the dependency on trusted third parties or system operators. In this framework, the encrypted eHRs residing in the cloud server are accessed by different uses through multi-level security procedures.

RELATED WORK

Researchers have proposed several solutions to solve the security and privacy problems related to eHRs. Existing research work associated with privacy preserving techniques of patient eHRs can be categorized as i) Privacy by access control, and ii) Privacy by cryptographic approaches

Privacy by Access Control

The key objective of access control mechanisms is to permit the authorised users to manipulate data and thus maintain the privacy of data (Barua, Liang, Lu, & Shen, 2011). However, the progresses are not satisfactory enough to fulfill the privacy requirements for eHRs (Santos-Pereira, Augusto, & Cruz-Correia, 2013).

Different access control mechanisms can be found in the literature (Alhaqabani & Fidge, 2007; Chen et al., 2012). Discretionary access control (DAC), mandatory access control (MAC), role based access control (RBAC), and purpose-based access control (PBAC) are the basic models of the access control principles.

DAC restricts access to objects based on the identity of subjects and/or groups to which they belong. However, in DAC granting read access is transitive and the policies are helpless for Trojan Horse Attack (Ferraiolo, Kuhn, & Chandramouli, 2003, Hu, Ferraiolo, & Kuhn, 2006).

MAC policy can prevent the Trojan Horse that occurs in DAC. MAC is based on access control policy decisions, made by a central authority (Ferraiolo et al., 2003, Sandhu & Samarati 1994). In MAC, the individual owner of an object has no right to control the access. Thus, MAC policy fails to preserve the privacy requirement for eHRs of the patients (Motta & Furuie, 2003).

RBAC (Park & Sandhu, 2002) models use consents and rights based on the assigned roles in groups/institutions to limit access. However, RBAC cannot integrate other access parameters or related data that are significant in allowing access to the user (Evered & Bögeholz, 2004).

PBAC is based on the notion of associating data objects with aims (Byun, Bertino & Li, 2005). PBAC has proven the greater privacy preservation by allocating objects with purposes (Naikuo, Howard & Ning 2007; Li, Yu, Ren & Lou 2010).

However, purpose administration creates a great deal of difficulty at the access control level. In Gajanayake et al. (2012), the authors combine three existing access control models and present a novel access control model for eHRs which satisfies the requirements of eHRs but the processes are more complex to implement.

Privacy by Cryptographic Approach

The cryptographic approach is considered one of the safest ways to preserve the security and the privacy of information in distributed settings. To transmit the data safely in cloud computing, cryptographic solutions are suitable enough by practicing the public key structure (Ding & Klein, 2010). Encrypting the private information before sending it to the cloud is an inherent need to a cloud user. But not all settings may allow that to happen. As mentioned in the previous section, in many systems the user has to trust the operator and gives the authority to their data by default. Many cryptographic solutions have now eliminated this requirement and ensure the full authority of the data is in the hands of its owner.

To deal with the potential risks of such privacy exposure, several eHealth systems (Benaloh Chase, Horvitz & Lauter, 2009; Jin Ahn, Hu, Covington & Zhang, 2009; Li et al., 2010) let patients encrypt their health record

All these proposed solutions might preserve some of the privacy issues of a patient. They may require the encrypted data to be downloaded from the cloud to the patients’ local machine when a modification or a computation might be necessary. This unreasonable requirement would ruin the sole purpose of using the cloud system. Therefore, these proposed solutions are impractical in PCEHR settings.

Hence, an encryption-based practical solution for the PCEHR system is extremely important to ensure the full authority of the private data to its owner.

THE PROPOSED MODEL

In this section we describe the proposed cloud-based PCEHR model using homomorphic encryption, which is briefly discussed below.

Fully Homomorphic Encryption (FHE)

Homomorphic encryption is a special form of encryption where one can perform a specific algebraic operation on the plain-text by applying the same or different operation on the cipher-text. If \( x \) and \( y \) are two numbers and \( E \) and \( D \) denote encryption and decryption function respectively, then homomorphic encryption holds the following condition for an algebraic operation, such as ‘+’:

\[
D[E(x)+E(y)] = D[E(x+y)]
\]

Most homomorphic encryption system such as RSA, ElGamal, Benaloh, Paillier etc. are capable of performing only one operation. But the fully homomorphic encryption system can be used for many operations (such as addition, multiplication, division etc.) at the same time. In the area of cryptography, the fully homomorphic encryption (FHE) system proposed by Dijk, Gentry, Halevi & Vaikuntanathan (2010) is considered as a breakthrough work which can be used to solve many cryptographic problems. Key generation, encryption and decryption functions of this FHE are as follows:

\[
\text{KeyGen}\; (\lambda) : \text{Choose a random n-bit odd integer } p \text{ as the private key. Using the private key, generate the public key as } x_i = pq_i + 2r_i \text{ where } q_i \text{ and } r_i \text{ are chosen randomly, for } i = 0, 1, ..., \tau. \text{ Rearrange } x - i \text{ such that, } x_0 \text{ is the largest.}
\]

\[
\text{Encrypt}\; (pk, m \in \{0,1\}) : \text{Choose a random subset } S \subseteq \{1,2, ..., \tau\} \text{ and a random integer } r, m \text{ is encrypted to the cipher-text } c = m + 2r + 2 \sum_{i \in S} x_i (mod\; x_0). \text{ Let us denote this operation as } E_{pk}(m).
\]

\[
\text{Decrypt}(sk, c) : \text{The message } m \text{ is recovered simply by performing } m = (c \mod p)\mod 2. \text{ Let us denote this operation as } D_{sk}(c).
\]

Further detail of this FHE can be found in Naehrig, Lauter, & Vaikuntanathan (2011).

In this proposed PCEHR framework, we will use this FHE technique to enable the system to perform computation on encrypted data. The patient will be the owner of the secret key therefore none can decrypt his/her health record; whereas, the user might be able to perform some edit or write operations on the record without knowing the content of the record itself. Figure 1 demonstrates how this FHE can be used in such a secure computation.

Figure 1. A user can update a patient’s record in the secured server without knowing the content of the record
Architecture of the Proposed Model

A simplified architecture of the proposed model is shown in Figure 2. The model consists of several entities. These entities are briefly described below:

a. **User of Patients’ eHR**: In the proposed model, the user refers to any person/organization that needs to access the patients’ eHRs. Thus, the term ‘user’ includes a general practitioner (GP), specialist, pharmacist, nurse, healthcare provider, provider/health insurance company, diagnostic laboratory, hospital, research personnel, family member or relative of the patient. The purpose of the user may differ according to their role, such that a GP might need to access the previous records for making a prescription, whereas a diagnostic lab may need to store a report against a patient only.

b. **Authentication Server**: The authentication server ensures legitimate access in to the network of the model. The authentication process is usually based on passwords. However, different types of information such as biometric information, rather than text based information, can also be used in the authentication process. Usually every user needs to be registered in the system by associated authority. Other algorithms such as the challenge response protocol, Kerberos, and public key encryption, can be used by the authentication server.

c. **Access Control List (ACL) Server**: The main purpose of this server is to verify when users want to access a specific sub-profile of eHRs of the patients. After a user is authenticated in the system, the access control list server applies access control policies/rules correlated with the authenticated users. Access policies can be
defined as relationships between subjects, objects and actions. For example, a pathology lab technician usually does not need to access a patient’s mental health eHRs, or an insurance company personnel does not need to modify patient’s disease history. The ACL list also specifies how a user can access an object class of a patients’ eHRs, in other words, the actions that the user can perform on a sub-profile, e.g. read, write, etc.

d. **Authorisation Server:** After passing through the authentication and ACL servers, users need to be authorised by a patient through an authorisation server to access specific eHRs of the patient. The ACL server confirms the eHR class (known as sub-profile) accessibility, while authorisation confirms a particular object of that eHR class. If a patient provides permission to a user, the authorisation server will issue a token using the encrypted data which can be retrieved from the database server. The encrypted data can be decrypted by the patient’s private key only.

Figure 3 shows how the ACL and the authorisation servers allow users to access patients’ sub-profiles and eHRs. The figure shows four patients P1, P2, P3 and P4. Each patient owns a profile. A patient’s profile is divided into several sub-profiles, such as mental health, sexual health, physical health, personal information etc. The privacy policy dictates which professionals (known as user in the proposed model) are permitted to access which sub-profiles of the patients. For example, Figure 3 shows that the trainer is authorised via the ACL server to deal with patients’ personal information and physical health. However, the trainer does not have any priviledge to access patients’ sexual health or mental health. After accessing through the ACL server, the user waits for a patient’s exclusive permission to access the eHR of the patient. Figure 3 also shows that the psychiatrist receives permissions to access P1’s mental health, the trainer receives permission to access P2’s physical health. Although verified by the ACL server to access physical health sub-profiles of patients, the trainer can not access P1’s physical Health.

![Figure 3. An overview of the function of the authorisation server](image)

e. **Patient:** The patient may be defined as the owner of his/her medical data and therefore holds full rights on the access control of his/her data. In our proposed model, only the patient has the key to decrypt the data and hence without the patient’s consent no-one can access the eHRs. This means that the patient must be able to access the information about all types of data transfer, its purpose and its user. In addition, the patient must receive all the notifications whenever someone uses or access the records.

f. **Cloud Storage:** Scalability would be one of the challenges of eHR system, because the system needs to handle millions of patients’ electronic health records. For example, body are networks (BANs) are recently been proposed to monitor patients’ health. BANs sense, generate, and send monitoring data to the healthcare system. Indeed, the sampling is performed at high frequency, which increases the amount of collected data. In addition, the frequency of sampling is often increased if the condition of patients being monitored gets worse. The
amount, size, and heterogeneity of data drive a need for an increasing storage and processing capacities. Besides scalability issues, medical data could be life saving and must be accessible at any time and from everywhere. Existing solutions rely on a centralized paradigm to store and process sensed data thus cannot tackle the aforementioned challenges. Thus, we need new innovative solutions to meet the great challenges of handling the exponential growth in data. We leverage cloud computing technology to dynamically scale storage resources via on demand provisioning. Cloud service providers can be any type of internet provider or application that lives in the cloud and is accessed online. As encrypted data is used in our model, the providers do not know the original records.

g. Encrypted Database Server: A database server can be referred to as the back-up system of a database usage client/server structure. A database server accomplishes several tasks such as data analysis, storage, data handling, archiving, and other non-user specific tasks. In the cloud environment, eHRs are always vulnerable to attacks. Encryption of the database server helps us prevent unauthorised access to information database. Applying the homomorphic encryption technique, we can ensure the privacy of the eHRs of the patient.

h. System Operators: According to the PCEHR system, the system operator is the entity that is responsible for generating and operating the PCEHR system. The system operator must respect the instructions and recommendations (if any) during their duties given by the PCEHR Jurisdictional Advisory Committee and the PCEHR Independent Advisory Council (2013).

In general, system operators are either persons or machines who oversee the operation of a large computer network. So, system operators enable access to the database as it is assumed that they are always trusted authority. In our proposed model, we use cryptographic solutions to encrypt the central database. This is done by the encryption mechanism which can support addition and multiplication of the encrypted data. One instance of data access method is depicted in Figure 4.

![Figure 4. Accessing patient’s eHR by a user (doctor)](image)

Many encryption techniques support one of these operations, either addition or multiplication like the encryption schemes in Rivest, Shamir & Adleman, 1978; El Gamal, 1985; Goldwasser & Micali, 1984; Paillier, 1999. A cryptosystem which supports both addition and multiplication (homomorphic) can be successful for data security, and supports the creation of programs that accept encrypted input and generate encrypted output. As a result, system operators cannot know the original records of the patient.

**ANALYSIS OF THE PROPOSED MODEL**

Below we discuss different possibilities of attacks against patients’ eHRs and demonstrate how these attacks would be handled by the proposed model.
Case I: An Intruder wants to access eHRs

The Intruder could be controlled by the authentication server. The Remote Authentication Dial-In User Service (RADIUS) protocol can be used in the authentication server. In this client/server based protocol, the client passes user information to designated RADIUS servers and acts on the response that is returned. RADIUS servers receive user connection requests, authenticate the user, and then return the configuration information necessary for the client to deliver service to the user.

The RADIUS server supports a variety of methods to authenticate a user. When it is provided with the username and original password given by the user, it can support PPP, PAP or CHAP, UNIX login, and other authentication mechanisms. Typically, a user login consists of a query (Access-Request) from the NAS to the RADIUS server and a corresponding response (Access-Accept or Access-Reject) from the server. When the RADIUS server receives the Access-Request from the NAS, it searches a database for the username listed. If the username does not exist in the database, either a default profile is loaded or the RADIUS server immediately sends an Access-Reject message. This Access-Reject message can be accompanied by a text message indicating the reason for the refusal.

Case II: X tries to masquerade as Y

The ACL server maintains a list of permissions attached to each object class (for example, sexual health information, neonatal information, drug related information etc.) of a patient’s eHRs. An ACL specifies which users or system processes are granted access to object classes, as well as what operations are allowed on given object classes. Each entry in a typical ACL specifies a subject and an operation. For instance, if a file has an ACL that contains (Alice, read), this would give Alice permission to read a particular type of eHR. When a subject X requests an operation on a eHR type in the proposed model, the operating system checks the ACL for an applicable entry to decide whether the requested operation is legitimate. A key issue in the definition of any ACL-based security model is determining how access control lists are edited, namely which users and processes are granted ACL modification access. If the subject X is not entitled to access the object class, the system will generate an alert and inform appropriate level of personnel.

Case III: The subject X manages to pass through the ACL server and tries to access information of a specific patient’s eHR

Passing through the authentication and ACL server, the subject X needs authorisation from the patient to access his/her eHRs. The authorisation server will generate a key to retrieve a patient’s a particular type of eHR. Thus, if healthcare personnel try to access an eHR, for which she/he is not authorised, the authorisation server will not allow the illegitimate access/retrieval of the patients’ eHRs.

Case IV: System operators try to abuse patients’ eHRs

This is a crucial part in the proposed model. Homomorphic encryption would be used in the database server of the proposed mode. All eHRs of patients are encrypted in the database server using homomorphic encryption. This encryption technique allows the patients to update their eHRs without letting the system operator know about the content of the modification. Even the system operators will not be able to identify in which sub-profile the modification occurred.

Case V: One patient tries to access/permit other patients’ eHRs

Each patient, whenever they enter the system, is authenticated by the authentication server. After successful authentication, if the patient wants to hack another patient’s eHRs, he/she will not be successful because only the corresponding patient has the private key to decrypt the eHRs.

Case VI: Man in the Middle Attack

In our proposed framework, man in the middle attack is impossible. Let us consider the instances when te e-HR of a patient is updated or used:

(i) Update in the cloud server: When any e-HR user e.g. lab, wants to update the profile of the patient then the authentication and access right is ensured by authentication server, ACL server and the patient’s consent. These server would maintain a session freshness mechanism to protect any man in the middle attack.

(ii) Display and update at the doctor’s end: When a doctor or specialist will need to access any e-HR of a patient, we assume, the patient will be present in the session. The doctor will download the encrypted
relevant part of the record from the server to his local machine. Then he only can decrypt locally with the help of the key provided by the patient. Again at the end of the session if the doctor needs to update the e-HR, he will encrypt the relevant part of the profile on his machine locally with the key provided by the patient and upload to the cloud server. Therefore, man in the middle attack cannot have access to the e-HR since it is encrypted.

CONCLUSION AND FUTURE WORK

In this paper we present a PCEHR model to protect the privacy of patients’ eHRs using a cryptographic technique. Many studies have been performed to ensure the privacy of the system, where the system operators are able to access patient’s eHRs. In our proposed model, system operators cannot learn about a patients’ eHRs. Highest priority is given to the patients to control their eHRs to ensure the highest level of privacy.

According to the proposed model, only the patient has the key to decrypt the data. As a result, when a patient is disabled or intellilectually impaired or in the case of emergency, it is infeasible for any medical service to retrieve the eHRs. Our future work will include the access control policy using which patient’s eHR can be accessed during an emergency while still preserving their privacy. We also want to implement the proposed model which can be compared with other existing solutions in terms of efficiency and privacy.

REFERENCES


