A Security and Privacy Model for Electronic Health Records

Xuezhen Huang 1, Jiqiang Liu 1, and Zhen Han 1
1 School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China

Abstract - The adoption of Information and Communication Technologies (ICT) in healthcare field makes the electronic-healthcare (e-Healthcare) environment and information from different medical organizations collected together, forming shared e-Healthcare information (EHI) named electronic health record (EHR), so that availability and completeness of medical information are improved. The owners of EHI provide them to organizations and drug stores for scientific research and analysis. But they pose threats to security or privacy of patients. Building a secure EHR sharing environment has attracted attention in both healthcare industry and academic community. We propose a privacy preserving security model of EHR for manage the security issues of personal information. Firstly, patients make their privacy policy of medical data according to their own concerns and preferences. Secondly, user can access medical data only if they are authorized. Finally, the adoption of proper technology makes the security model satisfying privacy requirements when data are released for other uses.

Keywords: privacy preserving; security model; EHR; EMR

1 Introduction

Independent medical organizations originally generated and managed patients’ medical information. EHRs are shared among healthcare personnel to provide complete and accurate information and to improve the quality and efficiency of healthcare diagnostics. The ultimate goal is to realize all EHI shared in the network, to end paper based medical records and reduce duplication of inspection. Patients to any hospital, just to provide a medical record number, will be able to achieve all medical records needed and find a variety of test results. EHR systems are run by community, regional emergence, or national wide emergence organizations. They may violate the security and privacy of EHR.

Uses of several purposes of EHR have given rise to long-term challenges for securing EHR. EHR is personal and its primary purpose of uses is to support decision-making of clinical care of an identified individual. Other uses of EHR are involved to as secondary uses. The secondary uses of healthcare information can be regarded as a balance between individual security and society’s necessity to reduce healthcare costs improving quality and efficiency of the healthcare treatment. It is necessary to use the EHRs in medical research, assessment of care quality and healthcare treatment planning and management. Therefore, the secondary uses of EHI ultimately enhance patients’ benefits through a well-managed healthcare treatment environment.

It is required to design a secure environment for operating EHI to protect the privacy and confidentiality of the individual who receives healthcare services that are delivered through e-Health. Advances in security technologies have so far not eliminated the challenge posed by the need to secure EHR.

In [14], a security system for healthcare application clouds was proposed. Their system has three components: EHR secure collection and integration component, EHR secure storage and management component, and EHR secure usage component. Patients’ EHRs are controlled by organizations not considering the special privacy requirements of patient themselves. Different individuals have different degrees of privacy demand in the same privacy case. Even an individual may have different degrees of privacy demand in different privacy cases. So the personal interests and privacy should be reflected in a security system.

In [1], a system providing privacy for EHR is proposed. The system for EHR includes two subsystems- the patient service and the data service. All access control policies are developed by the patients themselves. It fully takes into account personal interests and privacy. However, the reality of the situation will be very complicated. Some solutions of EHI access are not the simple problem of whether or not to release medical data. Some technologies of restrictive release should be considered. What’s more, in the medical organizations, patients aren’t familiar with the rights and responsibilities of doctors. The facts that all privacy preserving policies are made and implemented by patients are infeasible.

The contribution of this study is that we propose a security and privacy model for EHR which is a hybrid model of patients, doctors and organizations co-management through using a combination of means of three methods. Among them, a privacy preserving security model of EHR that offers raw EHI self-determination to the patients including usage control with implicit possibility to trace data flows after sensitive data has been legitimately disclosed. This part is managed by
patient so called patient-centered. Furthermore, our security model involves a sub-model named T-EMR service for access control which is developed by local medical organization where patients get medical treatment to prevent malicious access. This part is managed by doctors and organizations so called organization-centered. As a result, there is no need to interact with the policy service for access to patients’ EHR every time. Third but not the least, we adopt privacy preserving technology based on restrictive release in a sub-model for exploiting to protect patients’ privacy when data are published for secondary uses.

2 Related research

Now we introduce the concept of Electronic Medical Record (EMR), Electronic Health Record (EHR), and Personal Health Record (PHR). EMR is the legal record of what happened to the patient during their encounter at a Care Delivery Organization (CDO) across inpatient and outpatient environments and is owned by the CDO. EMR is created, used and maintained by healthcare practitioners to document, monitor, and manage health care delivery within a CDO [14]. The electronic records sharing between different EMR systems are called electronic health records (EHRs). The EHR systems are created, maintained by community, state, or regional emergence, or national wide emergence organizations. The electronic personal healthcare records (EPHRs) [15] are online web-based healthcare records that are created, maintained and managed by the owner of the information. They have arisen as free services managed by IT enterprises, such as Google Health and Microsoft Health Vault. In fact, there is no clear boundary between EHR and EPHR considering the widespread cooperation between them.

We present some privacy preserving technologies which can be carried out on the EHI field.

2.1 Data Anonymization

Data anonymization is the current hot research. Anonymity is used to remove or dilute a data set of all a patient’s identity information, making compromise between the risk of privacy disclosure and data accuracy, which take availability, security and privacy of data into account.

EHR typically contain three types of attributes. Explicit identifier is able to uniquely identify a single individual attribute, such as name, ID number and phone number. Quasi-identifiers are multiple attributes which combine to uniquely identify a person, such as zip code, gender, birthday and other co-expression. Sensitive attribute contains sensitive data, particularly in relation to the details of individual privacy, such as disease, illness records, personal salary etc. If the data sheets are published with only simple identifiers removed, private information may still be obtained through co-locating quasi-identifiers.

The principle of k-anonymity [2] requires that the release of the data table for each record cannot be distinguished from other records. Equivalence class is known as k records which cannot be distinguished. Generally the larger the value k, the better the protection of privacy, but lost the more information. There are many anonymity principles based on k-anonymity. The l-anonymity [3] requires the release of data sheet for each equivalence class has at least l different sensitive attribute values. The t-closeness anonymity [4] based on the of l-anonymity consider the sensitive attribute distributions within an equivalence class, requiring differences in the sensitive attribute distributions within an equivalence class and their values no more than t. Identity-keeping anonymity [5] is another anonymity approach. It is aimed at individuals in a data sheet, rather than data records. This principle protects privacy better in case of a single individual corresponding to multiple records. The above principles can redefine based on identity-keeping anonymity.

The above anonymity principles are for static data release, the following anonymity principles are for dynamic data. Under dynamic condition, we should ensure that not only every release satisfies the data anonymization, but also even the collaborative attack to multiple data releases can still protect privacy. The m-invariance anonymity [6] requires that in dynamic data release, there must be m pieces of records having different sensitive attribute values, and in the repeated releases of a record, sets of sensitive attribute values of the record’s equivalence class must be equal. The l-scarcity anonymity [7] requires that at any time after the release of anonymization data, disclosure risk of any sensitive attribution values of any individual is no more than l.

There are different anonymity algorithms depending on the anonymity principles. We should pay attention to weigh carefully the security, privacy and accuracy of release data to select an anonymity principle of and an anonymous algorithm.

2.2 Data Pseudonymization

Anonymization removes explicit of the individual from EHRs mainly because the individual identity is unnecessary for secondary uses. However, situations exist where it may be required to re-create the link between the EHR and the owner of EHR [12]. Such situations include handling follow-up data, individual’s request to withdraw their information, further treatment of a patient in light of new discoveries and quality control. Neubauer and Riedl [13] define the concept of pseudonymization as: a technique where identification data is transformed into, and afterwards replaced by, a specifier, which cannot be associated with the identification data without knowing a certain secret. Iacono[12] identifies two pseudonymization schemes based on the reversibility. The first is the one-way pseudonymization scheme, which generates pseudonyms which are impossible to be used to re-identify the patients. This type of scheme requires the maintenance of a mapping database to store associations between pseudonyms and explicit identifiers. The second is the reversible pseudonymization scheme, which allows the patient to be re-identified through the use of cryptographic
mechanisms applied to the pseudonyms. The latter does not require a mapping database. Neubauer et al. [16] provides an methodology that combines primary and secondary use in one system and guarantees data privacy. The security analysis showed that the methodology is secure and protected against common intruder scenarios.

3 Security and Privacy Requirements

In practice, EHRs exist in a dynamic cycle in some form. The dynamic cycle includes collection, access, creation and publication of EHRs. In every link, there are specific security and privacy challenges. We list requirements for a privacy preserving EHR:

1) A patient may have his or her EHI in different EMR systems. To increase efficiency in medical services distributed EHI should centralized to EHR accurately, securely and fast.

2) When a patient sees a doctor, the medical organization where the patient is served should get patient’s EHR accurately, securely and fast.

3) The medical organization should assure that doctors who provide the medical serve can obtain patient’s EHR and unauthorized individual who is even at the same medical organization can’t get the patient’s EHR.

4) We need to address the authenticity of EHR with respect to both content authentication and source verifiability.

5) When a patient finishes treatment at a hospital, after completing the EHI this time, doctors can’t access the EHR any more, unless the patient allows. If it’s necessary to get raw information of patients for any uses, it must be allowed by patients and trails and logs can be audit.

6) If some organizations apply for EHI for secondary uses, information provider can’t disclose patients’ sensitive information.

4 A Security Model of EHR

We propose a privacy preserving security model of EHR including four sub-models: the EHR service, the temporary EMR (T-EMR) service, the EMR service and the patient service. Fig.1 shows the schematic architecture of these components and their resulting data flows.

The EHR service integrates distributed EMRs and controls the storage and disclosure of EHRs. Other medical organizations use the EHR service for storing or retrieving EHI. The T-EMR service is managed by an organization. A medical organization obtains the EHRs of patients, only when patients in hospital here. After authorized medical organization reaching EHR of a registered patient, the T-EMR service implement locally access control and generate new EHI. At this point, the organization is responsible for security and privacy of EHRs. That is to say, the patient is out of control of her or his own EHR which is more efficient and feasible, considering the critical moments. The EMR service controls the storage and disclosure of EMRs which is generated locally, after patients discharged from organization. Patients, other organizations or enterprises can obtain some local statistical information about any doctors or diseases by the EMR service. The patient service offers administrative communication and is an interface for patients to the system where they can develop privacy policies on the release of their data and check whether their enforcements are run as intended.

Figure 1. A security model of EHR.

5 EHR Service

The EHR service (Fig. 2) is one of the most important building blocks in the security model of EHR.

5.1 EHR Secure Collection and Integration

EHRs in the EHR Service are integrated from the distributed EMRs in the EMR Service. EMRs between the interaction entities meet cooperation and provide strong mutual authentication and responsibilities. Therefore, EHR system administrators verify authenticity, confidentiality, integrity and trust of EMRs from different medical institutions. EHR system administrators combine and integrate the successfully verified EMRs into a new composite EHR with a security certificate signed by the integrator. EHRs are encrypted and stored at EHR point.

5.2 Access Control of Raw EHRs

When there is an applicant request for raw EHRs, we implement following process. The policy enforcement point I (PEPI) enforces access control policies. Answer to an access request and storage of access control policies are done by the policy service, which combines a policy repository and the policy access point. When there is an applicant applies for
access EHR, the PEPI asks the policy service whether the request is admitted. The policy service decides according to the agreed-upon privacy policy $D_{\text{Policy}}$ made by the owner of EHR and returns the answer including relevant obligations. Then it sends $D_{\text{log}}$ which records the access request and the resulting decision to the patient service.

Specially, when a patient registered in a hospital, the hospital will apply for raw EHRs of the patient with registration information signed by the patient. Of course, this application complies with privacy policy stated by the patient.

5.3 Access Control of EHRs after Handling of Privacy Preserving Technology

EHRs have an influential value for secondary uses. Centers for Disease Control need to collect EHRs for disease prevention and control. Research institutions for the good use of EHRs could benefit the patient. However, in this process, if the patient raw data release, which will inevitably expose sensitive data, the owners of sensitive data do not want to be exposed. So, whether it is data mining or data release, the raw data must take appropriate protective measures. On the one hand to protect personal privacy from being compromised, the other is to have sufficient information for analysis applies. Therefore, the privacy preserving technology I (PPTI) is necessary. PPTI tries to retain more validity of information while protecting the privacy of individual information. The verification service I verify the resulting data corresponding with the privacy preserving objectives and principles. Then pass EHR to the applicant. Privacy preserving technology based on restrictive release is to selectively release the raw data, not to release or to release of sensitive data of lower accuracy, in order to achieve privacy preserving.

Anonymization and pseudonymization especially anonymization are widely used in privacy technologies. Data anonymization and data pseudonymization attempt to address this problem by de-personalization of e-Healthcare information. However, these two methods ensure the privacy of the disclosure of the probability no more than a particular threshold. Data anonymization is a chronic problem. Therefore, before technologies taken into application, a thorough threat analysis is needed. Related methods and algorithms in anonymization and pseudonymization fields have been mentioned in section 2.

6 T-EMR Service

The temporary-EMR (T-EMR) service implements access control to protect the security and privacy of EHI locally. After a patient’s registration in a medical organization, the organization applies for patient’s raw EHRs. When the organization gets patient’s EHRs, the encryption service encrypts them. The access control point (ACP) carries out access control according to the local doctor's roles and responsibilities based on the patient's treatment department and attending physician. The attending physician is probably
in a medical stuff which is in charge of the patient. After the physician (or the stuff) completing the treatment and decision, the physician (or the stuff) generates new EMR with the signature of the physician (or the stuff). The signature of new EMRs generated locally need to be verified in verification service II. T-EMR point passes the successfully verified data to the EHR service.

7 EMR Service

The architecture of the EMR service is as Fig. 2. After the EMR point obtaining data of EMRs from the T-EMR service, it encrypts and storages them. The approach and the process of raw data on EHR access control and data privacy preserving technology handled on release process are both as same as the EHR service, except for the fact that supervisor of the EMR service and the data itself are different from the EHR service. The other difference is that the Policy Service is located in the EHR service. So every accessing to raw EMR, PEPII need to interactive with Policy Service in the EHR service in order to obtain the patient’s security policy.

8 Patient Service

The patient service is the communication interface where patients can realize policy management and check data access log. It is divided into three components. The policy management is a point which patients use to view, present and modify access control policies DPolicy. Every access request creates an event that is delivered from the policy service to the logging service of the patient service. The verification service IV offers a list of access details to a verifier whether the system runs as intended and the agreed-upon policy has been enforced[1].

9 Conclusion

We have concluded security and privacy requirements for e-Healthcare application. Then we present an EHR security and privacy model for managing security issues in e-Healthcare information, which highlights four important core components. Patients are given right to decide on the usage and disclosure of personal raw EHRs. For the secondary use of EHRs, we adopt privacy preserving technology such as data anonymization to protect patients’ privacy. Access control during patient treatment is managed by the medical organization where patient’s treatment is implemented. The combination of the above three patterns of data management support security and privacy of e-Healthcare information. This hybrid model makes the management of EHI security, privacy, autonomy, quickness and efficiency.

10 Acknowledgment

This paper is based upon work partially supported by the National Natural Science Foundation of China (Grant No. 60973112) and Innovative Research Team in University and Science and Technology Foundation of Beijing Jiaotong University (No. 2007xm009).

11 References


