Multi-domain and Privacy-aware Role Based Access Control in eHealth

Lorenzo D. Martino
Comp & Info Tech
Purdue University, USA
Email: lmartino@purdue.edu

Qun Ni
Computer Science
Purdue University, USA
Email: ni@cs.purdue.edu

Dan Lin
Computer Science
Purdue University, USA
Email: lindan@cs.purdue.edu

Elisa Bertino
Computer Science
Purdue University, USA
Email: bertino@cs.purdue.edu

***Abstract*—Information Technology-supported Heathcare (eHealth) is crucial in order to reduce healthcare costs, and improve quality of care and patient safety. Among technologies in eHealth, Electronic Medical/Health Records (EMR/EHR) enabling communication of patient data between different healthcare professionals (e.g. specialists, pharmacy), is the most important and sensitive. There are three crucial requirements when accessing EMRs: such access must be both secure and privacy preserving; such access must be allowed to individuals from different organizations; such access should be confined based on meta information about the EMRs. In this paper, we propose a multi-domain privacy-aware role based access control meeting these requirements.**

***Index Terms*—Privacy, Multi-Domain, P-RBAC, Data Profile.**

I. Introduction

Information Technology (IT)-supported healthcare (eHealth) is crucial in order to reduce healthcare costs, improve care quality and patient safety. In the last few years, hospitals and health plan providers have increased their use of eHealth solutions to manage health-related information and to automate administrative and clinical functions. Patient health care data is managed by Electronic Medical/Health Record (EMR/EHR, for short) systems that enable communication of patient data between a variety of healthcare professionals.

Sharing sensitive patient data in a large distributed and heterogeneous environment, however, inherently introduces security and privacy risks. These risks are further increased by the enhanced openness that can be achieved by the use of Web-based applications and pervasive devices in eHealth. The President's Information Technology Advisory Committee (PITAC) identified security and privacy concerns as fundamental obstacles to medical informatics deployment in its 2004 report, *Revolutionizing Health Care Through Information Technology*. The relevance of security and privacy in eHealth is also testified by the activities of regulatory bodies. The Administrative Simplification provisions of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) [1], through the Department of Health and Human Services, established national standards for the security of electronic healthcare information, that constitute the reference framework for security and privacy issues for healthcare.

In order to comply with these regulations, healthcare organizations have to define suitable organizational processes, which are often accompanied by the publication of privacy notices/practices on websites [2] intended to inform patients about the management of their data. Such privacy notices express privacy policies in P3P or incorporate them in some privacy seal programs (e.g. TRUSTe/URAC[3]) by stating general guidelines about:

- The use of medical information about patients.
- Patients' rights on their own medical information.
- The use of patient identifiable information, e.g. name, address, telephone number, e-mail address.
- How to make choice and opt-out, and how to get notification of changes.
- Use of cookies, log files, IP addresses.

High-level privacy policies notices/practices (referred to as *public privacy policies*) have to be translated into privacy policies written in formal languages (referred to as *internal privacy policies*) before being applied for access control at the implementation level. Therefore, there is a strong need for fine-grained access control mechanisms supporting the specification and enforcement of internal privacy policies. Conventional access models, such as Mandatory Access Control and Discretionary Access Control, are not designed to enforce internal privacy policies and barely meet the requirements of privacy protection [4]. To address the shortcomings of existing access control models, a family of Privacy-aware Role Based Access Control (P-RBAC) (see Fig. 1) has been proposed by Ni et al. [5]. Models in such family naturally extend RBAC models [6] to support internal privacy policies.

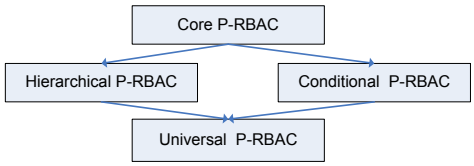

Fig. 1. P-RBAC family models

There are some important advantages in choosing RBAC as a starting point. First, roles, which are an important indirection between users and permissions, directly map onto healthcare organizational positions, such physicians, clinicians, nurses. Second, RBAC is widely accepted by industry and deployed in several products such as the ORACLE DBMS. Last but not the least, privacy extensions to RBAC could be easily

deployed in systems already adopting RBAC, thus allowing one to seamlessly introduce access control policies specialized for privacy enforcement.

The adoption of a model like P-RBAC in a pervasive eHealth environment requires, however, some major additional features. Healthcare is a complex environment which inherently encompasses multiple domains. Classical RBAC, including P-RBAC, does not support "role roaming" among different organizations. Furthermore, not only the content of EMRs but also some meta information about EMRs, e.g., the creators, owners, and validation dates of EMRs, are required for privacy protection.

Therefore, a privacy preserving mechanism for EMRs access control for pervasive eHealth scenarios requires:

- a fine-grained access control model that satisfies both expressiveness and flexibility demanded by privacy policies;
- a system that can be operated smoothly in a multi-domain environment;
- a system that can handle rich meta information.

Since Core P-RBAC provides a reasonable basis to describe flexible and complex internal privacy policies, in this paper we extend Core P-RBAC to support multiple domains and meta information. The following extensions are introduced: (i) automatic user-to-role assignments driven by preconditions on roles, that is, *role provisioning*; (ii) a flexible *data specification* through the use of data profiles.

The paper is organized as follows. Section 2 reviews related work. Section 3 introduces some use cases to illustrate our approach. Section 4 presents the extended Core P-RBAC model. Section 5 concludes the paper.

II. Related work

In this section, we compare our proposal against well known standards: P3P[7], EPAL[8] and XACML [9]. P3P enables websites to express privacy practices in a standard format that can be automatically retrieved and interpreted by agents. However, P3P policies are not sufficiently fine-grained and expressive to handle the description of privacy policies at the implementation level [10]. EPAL is proposed to encode enterprise's privacy-related data-handling policies and practices, which can be imported and enforced by a privacy-enforcement system. XACML is a widely adopted access control model based on XML. Both EPAL and XACML aim at providing flexible policy languages, but leave the policy analysis task to policy analyzers. Unlike such approaches, P-RBAC achieves a balance between expressiveness and tractability, and also guarantees that the insertion of a new policy will not affect the consistency of existing policies [5]. Concerning the more challenging pervasive eHealth scenarios, only our access control model supports multiple domains and meta information.

III. Case study

Before delving into technical details, we discuss two scenarios. The first one is derived from a policy privacy notice [2], and the second one describes a real need for multiple domain support in an access control system.

A. Scenario 1: privacy policies

For treatment purposes, patients' medical information can be accessed by physicians, nurses, technicians, medical students, or others who are involved in the patients' care or by other departments of the healthcare organization for the care/therapy coordination or by contracted physician services, such as emergency department physicians, pathologists, anesthesiologists, radiologists. For example, a physician treating a broken leg may need to know if the patient has diabetes because diabetes may slow the healing process. In addition, the dieticians should know if the patient has diabetes so that appropriate meals can be arranged.

B. Scenario 2: multi-domain privilege management

Alice is a medical student of Purdue University who needs to access patient and therapy data from Home hospital where Alice is doing internship. To grant proper permissions to Alice, Home hospital checks the hospital privacy policies and finds a policy stating that senior medical students from Purdue University can be assigned to an external medical student role which allows the access to unsensitive information about patients and therapies. Then Home hospital asks Purdue University to verify Alice's status. After Purdue University confirms that Alice is a senior medical student, Home hospital assigns the external medical student role to Alice.

Both those scenarios show the need for the role provisioning mechanisms able to describe the connection between roles in different organizations. Also, we can observe the use of a data profile, like specifying patient data that is not very sensitive.

IV. Multi-domain P-RBAC

We now present our solution towards a multi-domain privacy-aware access control for pervasive eHealth.

A. Core P-RBAC

Core P-RBAC [5] includes seven sets of entities: Users(U), Roles(R), Data(D), Actions(A), Purposes(P), Obligations(O), and Conditions (C) expressed by a customized language, referred to as LC_0. A user in the Core P-RBAC model is a human being, and a role represents a job function or job title within the organization with some associated semantics regarding the authority and responsibility conferred on a member of the role. Data in P-RBAC means any information or meta information relating to an identified or identifiable individual. An action is an executable image of a program, which upon invocation executes some function for the user. The types of action and data objects that P-RBAC controls depend on the type of system in which they are deployed. Purposes which are bound to actions on data in Core P-RBAC directly reflect the OECD [11] *Data Quality Principle*, *Purpose Specification Principle*, and *Use Limitation Principle*. Obligations, that is, actions to be performed after an action has been executed on data objects, are also part of many privacy policies. Conditions, that is, prerequisites to be met before any action can be executed, are frequent components of privacy policies too.

In Core P-RBAC, as in classical RBAC, permissions are assigned to roles and users obtain such permissions by being assigned to roles. The distinctive feature of Core P-RBAC lies in the complex structure of privacy permissions, which reflects the highly structured ways of expressing privacy rules to represent the essences of OECD principles. Hence, aside from the data and the action to be performed on it, a privacy permission explicitly states the intended purpose of the action, along with the conditions under which the permission can be granted and the obligations that are to be finally performed. In this paper we extend the definition of Core P-RBAC by introducing an additional language component, referred to as LD_0, supporting the flexible specifications of data items inside privacy permissions. Extended Core P-RBAC is defined as follows.

Definition 4.1: The extended Core P-RBAC model is composed of the following components:

- A set U of *users*, a set R of *roles*, a data specification language LD_0, a set P of *purposes*, a set A of *actions*, a set O of *obligations*, and a condition language LC_0.
- The set of *Privacy-sensitive Data Permission* $PDP = \{(a, d, p, c, o)\ a \in A,\ d$ is a valid data object specified by LD_0, $p \in P$, c is an expression of LC_0, $o \in \mathcal{P}(O)\}$, where $\mathcal{P}(O)$ denotes the powerset of O.
- *User Assignment* $UA \subseteq U \times R$, a many-to-many mapping user to role assignment relation.
- *Privacy-sensitive Data Permission Assignment* $PDPA \subseteq R \times PDP$, a many-to-many mapping privacy-sensitive data permission to role assignment relation. □

LD_0 is the extended P-RBAC language component for the specification of data. LD_0 supports three types of data specifications; by data identifiers; by conditions against data contents; and by conditions against data profiles. Such data profiles are the mechanism we provide in extended P-RBAC to store and manage meta data information. LC_0 is another language, part of both P-RBAC and extended P-RBAC, for expressing conditions like conjunctions of equality constraints over *context variables* which record privacy-relevant requirements taken into account when enforcing privacy permissions. For details, please refer to [5].

The permission assignments deriving from the previous scenarios can be directly expressed in Core P-RBAC. For example, in scenario 1) the permission:

```
(physician, read, patient.EMR.raw, treatment,
 subject = patient.duty_physician, ∅)
```

specifies that the physician role can read patient EMR content for treatment purpose only if the data user[1] is the patient's on duty physician, where `patient.EMR.raw` is a data object specified according to LD_0 and `subject = patient.duty_physician` is a LC_0 condition.

B. Role Provisioning

Role provisioning refers to the automatic process of creating user accounts and assigning roles to enable access to all needed applications and services for valid end users from multiple domains. Our approach to role provisioning is based on the notion of *role precondition*, which expresses the fact that a user can be assigned to a certain role provided that the user is associated to one or more specific roles in his/her home organization. We make a distinction between the organization where the patient's data are generated and maintained, referred to as *owner organization*, and the home organization of a user, denoted as *external organization*. In the pervasive environments, which we address in this paper, very often the home organization of user wishing to access the data is different from the owner organization of the data.

Definition 4.2: Let r_1 be a role in organization O_1, and let r_{21},..., r_{2n} be roles in organization O_2. A role precondition, denoted as ($\{r_{21}, ..., r_{2n}\}$,O_2,r_1), specifies that users with one of roles r_{21},..., r_{2n} in O_2 can be assigned role r_1 in O_1. □

Such role precondition allows an owner organization to specify access control policies based on limited information about external organizations, that is, the roles (or a subset of them) defined in the external organization. Such approach closely mirrors organizational practices, where access authorization is based not only on the identification and the authentication of the user but also on the roles that user plays.

To implement the role preconditions in our system, we use the notion of *Role Mapping Table* (RMT). There is one such table for each organization[2]. Each row of an RMT is a 3-tuple (r_e,e, r), where r_e is a role in the external organization e, and r is a role in the owner organization. The semantic of this tuple is that role r_e in an organization e maps to a role r in the owner organization. Fig. 2 shows an example RMT for Home hospital in scenario 2.

External Role	External Organization	Local Role
...	...	...
Senior Medical Student	Purdue University	External Medical Student
Senior Medical Student	Indiana University	External Medical Student
Senior Investigator	State Farm Insurance	External Investigator
...	...	...

Fig. 2. Home hospital role mapping table

Based on the example RMT, we describe the execution flow that highlights the salient features of our architecture:

1) Alice inputs her login name, password and organization "Purdue University" in a console at Home hospital[3].
2) Since Alice belongs to an external organization, Home hospital skips the local authentication procedure and checks its own RMT.
3) In the RMT, Home hospital finds an organization which matches "Purdue University". An authentication request is then sent to Purdue University to verify Alice's status.

[1]By data user we mean the user accessing the data to perform some tasks. In the example, the data user is a user assigned to the physician role.

[2]In practice an organization may have several such tables if the organization comprises several sectors that are independently administered and have different role systems. We do not elaborate on this aspect in the paper for lack of space.

[3]We assume that an approach like Shibboleth [12] is used for user authentication.

4) Purdue university confirms that Alice's role at Purdue University is senior medical student.
5) After receiving confirmation from Purdue University, Home hospital automatically assigns an external medical student role to Alice according to the RMT and stores the login information for Alice in the system.

C. Data Profile

The notion of data profile is a key in our approach in that it records all information about data that is relevant for privacy enforcement. A data profile is organized as a record storing attributes, that is, properties about a data item or a set of data items. We refer to such items or set of data items as "raw" data. Values of attributes in such profiles can be used when specifying permissions, allowing for different levels of granularity in data specification. In the healthcare domain, the content of a Electronic Health record can be considered as containing groups of related information. Several standardization organizations, such CHI, AHIMA and ASTM International [13], are working to define standards for both the content and the structure of EMRs. For example, ASTM E1384-02a identifies the content and logical structure of a EMR, and ASTM E2369-05 is an example of standard categorization of the disparate data contained in EMRs, with the purpose of aggregating, summarizing and transferring them among healthcare systems in a standard way. Examples of such groups, or data categories, are patients' identification data, therapy data, medications, prescriptions, surgical data, and so forth. When defining permissions, conditions against attributes in data profiles can be used to concisely denote data categories. For example a read permission can be issued that applies to surgical data only. Specifying such permission requires recording for each data item a *category* for the data; in our example, the category would be surgery. The permission could then automatically apply to all the raw data that have the value of category equal to surgery. A data profile can also contain context related information and other information that can be used to make an access decision. Attributes that are currently included in data profiles for extended P-RBAC are: Data-category; Creator-name; Creator-affiliation; Date-of-creation and Valid-to; Privacy-sensitive (Y/N).

Denoting data by their profile allows one to define privacy-enhanced access control policies at a level close to high-level privacy policies, regardless of the specific data management system. Moreover, the binding of the profiles to raw data, their storage and access can be dealth with by taking advantage of the features provided by database engines. Recent extensions of relational database engines support the notion of INFORMATION_SCHEMA allowing applications to define and manage their own meta data.

V. Conclusions

In P-RBAC the Privacy-Sensitive Data Permission (PDP) binds together the access' purpose and the obligation consequential to the access. Such an approach has several advantages. It helps in verifying that the access control policies of the healthcare organization are compliant with privacy regulations, since the obligation may include data logging actions. The availability of such logs may facilitate subsequent audits. Second, our approach enhances security in that it relies upon organizational control processes and not on the end-user claims. In particular, with respect to external users, role precondition provides a further control in addition to user identification and authentication. Our approach to role mapping assumes that: a) there is a trust relationship between the owner organization and the users' home organization, and b) the users' home organization itself adopt a controlled process before declaring that its users plays a certain role.

A prototype of the Core P-RBAC management console has been implemented [14]. Policy writers can use this console to maintain the underlying system data used in privacy policies, such as user, role, context variable, action, purpose, obligation, to assign users to roles, and to assign permissions to roles. Future work includes the development of consistency analysis techniques on privacy permissions w.r.t. data profile. We also plan to investigate different role provisioning strategies in multi-domain environments.

Acknowledgment

The work reported here has been supported by the IBM OCR project "Privacy and Security Policy Management" and the NSF grant 0712846 "IPS: Security Services for Healthcare Applications".

References

[1] United State Department of Health, "Health insurance portability and accountability act of 1996," available at http://www.hhs.gov/ocr/hipaa/.
[2] O. R. Heathcare, "Orlando regional heathcare privacy practices," available at http://www.orlandoregional.org/OrlandoRegional/AboutUs/ PrivacyPractices.
[3] TRUSTe.org, "An independent, nonprofit enabling trust based on privacy for personal information on the internet," available at http://www.truste.org/.
[4] S. Fischer-Hubner, *IT-security and privacy: design and use of privacy-enhancing security mechanisms*. Springer-Verlag New York, Inc., 2001.
[5] Q. Ni, A. Trombetta, E. Bertino, and J. Lobo, "Privcy aware role based access control," in *SACMAT '07: Proceedings of the 12th ACM symposium on Access control models and technologies*. New York, NY, USA: ACM Press, 2007.
[6] D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. R. Kuhn, and R. Chandramouli, "Proposed nist standard for role-based access control," *ACM Trans. Inf. Syst. Secur.*, vol. 4, no. 3, pp. 224–274, 2001.
[7] W3C, "Platform for privacy preferences (p3p) project," available at http://www.w3.org/P3P.
[8] IBM Zurich Research Laboratory,Switzerland, "The enterprise privacy authorization language(epal 1.1)," available at http://www.zurich.ibm.com/security/enterprise-privacy/epal/.
[9] OASIS, "extensible access control markup language (xacml) 2.0," available at http://www.oasis-open.org/.
[10] A. H. Anderson, "A comparison of two privacy policy languages: Epal and xacml," in *SWS '06: Proceedings of the 3rd ACM workshop on Secure web services*. ACM Press, 2006.
[11] Organisation for Economic Co-operation and Development, "Oecd guidelines on the protection of privacy and transborder flows of personal data of 1980," available at http://www.oecd.org/.
[12] T. Scavo and S. Cantor, "Shibboleth architecture - technical overview, 8-june-2005," avaliable at http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-tech-overview-latest.pdf.
[13] ASTM International, "Astm international," available at http://www.astm.org.
[14] Q. Ni, "P-RBAC," http://www.cs.purdue.edu/homes/ni/prbac.html.