Protecting Patients' Electronic Health Records Using Enhanced Active Bundles

Raed M. Salih[1] Leszek T. Lilien[1] Lotfi Ben Othmane[2]

[1] Department of Computer Science
Western Michigan University
Kalamazoo, MI 49008, USA
{raedmahdi.salih, leszek.lilien}@wmich.edu

[2] Department of Mathematics and Computer Science
Eindhoven University of Technology
Eindhoven, The Netherlands
l.ben.othmane@tue.nl

Abstract—**We propose a solution that provides protection for patients' electronic health/medical records disseminated among different authorized healthcare information systems. The solution is known as *Active Bundles using a Trusted Third Party (ABTTP)*. It is based on the use of trusted third parties, and the construct named *active bundles*. The latter keep electronic health/medical records as *sensitive data*; include *metadata* with information describing sensitive data and prescribing their use; and encompass a *virtual machine (VM)*, which controls and manages how its active bundle behaves. An essential task of the VM is enforcement of the privacy and other policies specified by metadata. We also propose enhancements to the ABTTP scheme. They include adding to ABTTP an algorithm finding the degree of privacy policy inclusion between two privacy policies, and a scheme, known as *Agent-Based Active Bundles*, which replaces trusted third parties with intelligent agents.**

Keywords- active bundle; confidentiality; healthcare information systems; intelligent agents; privacy; trust; trusted third party; virtual machine.

I. INTRODUCTION

Health/medical records are an account of patients' long-term health history, which contains doctor's visit summaries, prognoses, laboratory and radiology tests, treatment descriptions, etc. These records have been passing through technological transformation from a physical folder form to a digital form since the end of the last century [10].

The digital form of healthcare information—such as *Electronic Health Records (EHRs)* or *Electronic Medical Records (EMRs)*—are becoming more and more widespread, replacing "paper" medical records.

Use of EHRs[1] has a number of goals: (i) improving safety, quality, and efficiency of patient care; (ii) reducing the cost of healthcare provider delivery; and (iii) enriching the health-services research and public health monitoring [7].

The ownership (or guardianship) of EHRs is a central issue in healthcare because medical information has a commercial value; for example, some companies are making a profit by selling physicians' prescribing routines to pharmaceutical companies [4]. The *ownership* of an EHR, like the ownership of any property, represents the state or fact of exclusive legal rights and control over the property. In turn, a (*legal) guardian* for an EHR is a person or institution who has the legal authority (and the corresponding duty) to care for the EHR in terms of issuing permissions for creating, reading, or modifying the EHR.

According to American Medical Association, healthcare providers own the medical information that they collect [7]. Healthcare provides are not the only parties that access patients' EHRs; health insurance companies, federal or state governments, and researchers are examples of other EHR users.

Exchange of EHRs among *healthcare information systems (HISes)* is necessary for improving the quality of healthcare. However, facilitating data exchange can increase privacy threats—due to easier copying and dissemination of patients' EHRs among more entities. *Users of HISes* (or just "*users*" in this paper) include patients and healthcare providers.

We define *user privacy* as a user's right to protect and control her data. As a special case, *patient privacy* deals with data that are or include a patient's healthcare-related data. *User confidentiality* is the right of a user to keep her data private unless she gives a permission to disclose them to another party (cf. [10]). Hence, user privacy contains user confidentiality.

Privacy/confidentiality of a HIS user is defined as a special case of user privacy/confidentiality, which deals with use of HISes.

The paper is organized as follows. Section II discusses related work. Section III presents motivation and the problem of protecting EHRs disseminated among different authorized HISes. Section IV describes our solution known as *Active Bundles using a Trusted Third Party (ABTTP)*. Section V proposes an enhancement to ABTTP, which is realized by adding to ABTTP an algorithm finding the degree of privacy policy inclusion between two privacy policies. It also sketches a scheme, known as *Agent-Based Active Bundles*, which replaces trusted third parties with intelligent agents. Section VI gives the work status. Section VII concludes the paper and mentions future work.

[1] The term "EHR" in this paper means "EHR or EMR" despite the fact that EMRs are legal records of patients that are created in a single healthcare provider facility, while EHRs are the summaries of EMRs collected from more than one healthcare provider.

PervasiveHealth 2012, May 21-24, San Diego, United States

DOI 10.4108/icst.pervasivehealth.2012.248719

II. Related Work

Due to space limitation, we present related work very concisely.

Bhattacharya *et al*. [5] proposes middleware architecture called Privacy Broker to enforce the legal privacy requirements. It uses: (i) a unique key to encrypt and decrypt data that it retrieves from a database and (ii) capability certificates that verify and evaluate all users' requests and enforce policies to access patient's data.

Benaloh *et al*. [4] propose an encryption system, called Patient Controlled Encryption (PCE), to protect patient privacy. Using the PCE, the patient controls and shares his EHR with other authorized entities through generating and distributing a set of sub keys.

Akinyele *et al*. [1] suggest using a self-protecting EHR inside and outside of the hospital environment. The solution uses attribute-based encryption and a cloud system. The solution uses a set of policies and encryption/decryption keys that allow a patient to read, write, and manage his EHR through his mobile device that interfaces with a cloud system such as Google Health.

The current solutions for protecting EHRs have two main limitations: (i) they require not only an extensive exchange of user messages between caregivers to protect data, but also exchange of numerous control messages among caregivers' systems; and (ii) they depend only on encryption (in which data decryption keys must be provided to specific caregivers).

III. Motivation and Problem Statement

Protecting patient privacy is considered the main problem in HISes.

Figure 1 illustrates EHR dissemination. The hospital represents the main guardian for a patient's EHR. The hospital might send a copy of the patient's EHR to other guardians. For example, a clinic (*Guardian 4* in Figure 1) receives from the hospital (*Guardian 1*) a copy of a patient's EHR before, during, or after his visit. In turn, the hospital (*Guardian 1*) may distribute the patient's EHR to multiple guardians, as shown in Figure 1.

Such EHR dissemination increases the risk of disclosing (or leaking) of private information to unauthorized parties.

IV. The Initial Solution: Using Active Bundles

An *active bundle (AB)* [2, 3] is a software construct, which—as illustrated in Figure 2— bundles together the following three components: (i) *sensitive data*, which can contain a patient's EHR, so it is protected from privacy violations; (ii) *metadata*, which contain information describing sensitive data and prescribing their use; they include a privacy policy for the sensitive data (which control accesses to sensitive data or their portions), as well as the rules for AB dissemination; and (iii) a *virtual machine* (VM), which

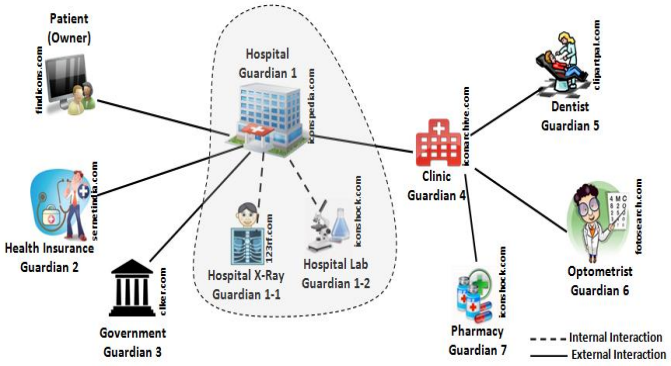

Figure 1. EHR dissemination example.

controls and manages how its AB behaves, thus making the AB active; the essential task of the VM is enforcement of the privacy and other policies specified by metadata.

For the problem presented above, we propose a solution named *Active Bundles using a Trusted Third Party (ABTTP)* [2], which is the current implementation of the AB scheme. ABTTP provides protection for the patients' EHRs disseminated among different authorized HISes. It uses a *trusted third party (TTP)*, which maintains and provides to ABs the trust levels of *visited hosts (VHs)*.

AB's lifecycle in its ABTTP realization consist of two phases: *AB creation* and *AB enabling*, discussed next in turn. The AB lifecycle for a patient's EHR in the *absence* of an attack on the AB is shown in Figure 2, and the AB lifecycle for a patient's EHR in the *presence* of an attack on the AB is shown in Figure 3. (Due the space limitation we do not include a complete scenario here; more details are in Refs. [10, 11].)

A. AB creation

Figures 2 and 3 show creation of an AB in a *hospital EHR system (HES)*, which received a request for a patient's EHR from a *clinic EHR system (CES)*.

In the AB creation phase, an AB is created either automatically or interactively with help of user-friendly AB creator software. An active bundle encapsulates its sensitive data, metadata, and the VM. The VM encrypts data and metadata of the AB to protect their confidentiality, using a single encryption key. (Note that these encryptions are only a supplemental privacy protection mechanism for ABs.) The VM itself is protected by code obfuscation [2, 9].[2] Then, the VM computes the hash value of AB's encrypted data and metadata (to protect their integrity), and encrypts it using another encryption key.

The two decryption keys, corresponding to encryption keys used by VM (for encrypting AB's data and metadata, and for encrypting the AB's hash value) plus the hash value (protected by encryption) are sent by VM to the designated TTP.

Now, a created AB becomes ready to be sent to a CES.

[2] We are also investigating use of homomorphic computing [2].

B. *AB enabling*

Figures 2 and 3 show the enabling process in the *absence* of an attack, and in the *presence* of an attack on the AB, resp.

When an AB reaches a visited host (VH) in CES, the *AB enabling* process starts. Enabling includes AB's verification activities and AB's enforcement activities.

1) AB Verification Activities

a) Checking VH's Trust Level (Items 7/7 in Figures 2/3)

When AB reaches a VH, VH sends request to AB to access the sensitive data. AB's VM (not encrypted, only obfuscated) sends request to TTP to certify the VH's trust level. TTP sends back the VH's trust level. There are two cases: (i) either host's trust level is *Not sufficient*, and the VM *apoptosizes* the AB (see Item 2b below); or (ii) the host's trust level is *Sufficient*, and the VM performs the next AB's verification activity, integrity checking (Item 1b next).

b) Checking AB's Integrity (Items 10/10 in Figures 2/3)

The AB's VM verifies integrity of the entire AB, by calculating the hash value for the AB, and comparing it with the hash value retrieved from the TTP and decrypted (with the key also received from TTP). If the two hash values are identical (Item 11, *Success*, in Figure 2), the VM starts the AB enforcement activities (Item 2 below). Otherwise (if the check result is *Failure*; cf. Item *End 2* in Figure 3), AB's VM apoptosizes [3] (cf. *End 1* in Figure 3).

2) AB Enforcement Activities

a) *Evaporation (Item End 3 in Figure 3)*

AB's VM might decide to destroy irretrievably portions of the sensitive data that the VH is not authorized to access (that is, all AB's data with the "required trust threshold" exceeding the trust level of the VH). This would happen, among others, in a situation when the VM feels threatened by unauthorized VH's access to some portions of the AB's EHR (VH is authorized to access other portions of the EHR in this case).

b) Apoptosis (Items End 1 and End 2 in Figure 3)

AB's VM destroys irretrievably the *entire* AB (including its EHR, metadata, and itself) in cases when: (i) a VH's trust level is lower than the required trust threshold specified by the AB's privacy policy; or (ii) the integrity check fails (cf. Item 1a above).

c) *Full or partial data disclosure*

If the VH meets the trust threshold requirements defined in the AB's privacy policies for the entire AB's EHR, the entire EHR is released to the VH (Item 13 in Figure 2).

If the VH meets these trust requirements for only *some* portions of EHR, these portions (after evaporation of all more sensitive portions) are released to the VH (Item *End 3* in Figure 3).

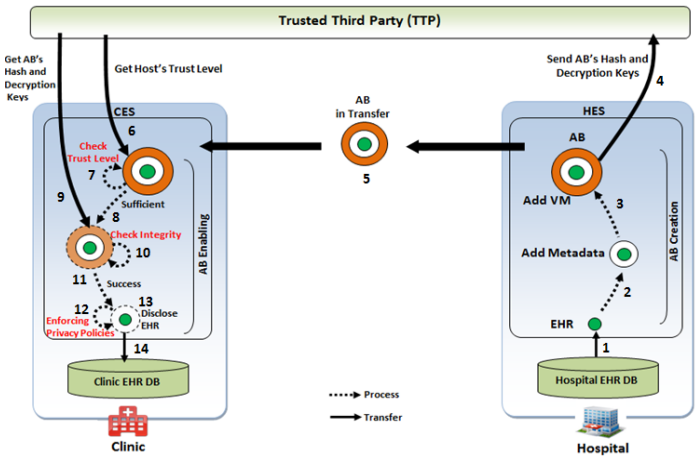

Figure 2. The AB lifecycle for a patient's EHR in the absence of an attack.

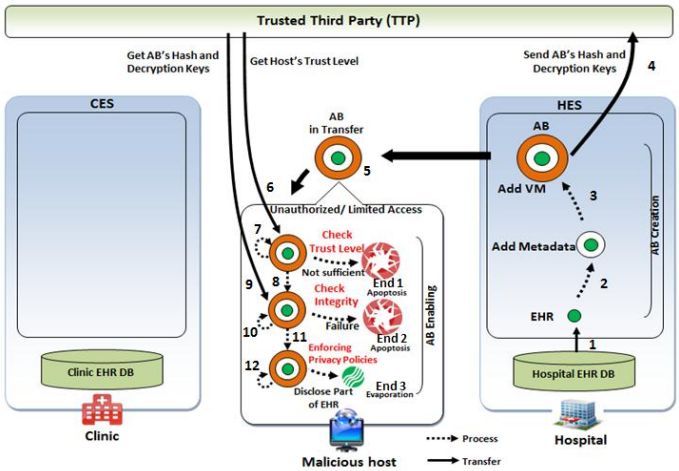

Figure 3. The AB lifecycle for a patient's EHR in the presence of an attack.

We believe that the AB scheme eliminates both limitations described in Section II (an extensive exchange of user and control messages between caregivers and their systems; and dependence on encryption alone). First, it does not require so many control messages between the AB and the VH in order to deliver an EHR from a source to a destination. Second, the AB scheme protects EHR privacy not only by encryption but also by enforcement of AB privacy policies.

Additionally, our approach does not need to distribute decryption keys (which are typically used in other privacy solutions) for all authorized healthcare providers; instead AB relies on TTP that provides the decryption keys. Moreover, ABs provide tools able to protect owners' (or guardians') privacy rights for arbitrary EHR fragments (down to a single record level). Also, an AB can protect a patient's EHR even if different records within the EHR are owned by different guardians who have privacy policies of differing strength [10].

V. An Improved Solution: Enhanced Active Bundles

We propose two enhancements to the current AB scheme.

A. Adding to ABTTP the Algorithm for Finding Degree of Privacy Policy Inclusion

Privacy policy PP1 is weaker than privacy policy PP2 (denoted $PP1 \subset PP2$) iff the closure[3] of all *PP1rules* is strictly (properly) included in the closure of all *PP2* rules. *Stronger* and *equal* relationships between two privacy policies are defined analogously.

If $PP1 \subset PP2$, we can calculate the *degree of inclusion of PP1 in PP2*, denoted *DI(PP1, PP2)*, using statistical similarity computation [8].

AB verifies VH's privacy policies before starting enforcement of its own privacy policies. First, an AB computes *DI(ABPP, VHPP)*, i.e., the degree of inclusion of its ABPP in VHPP. Then, the AB considers the following cases:

1) If $DI(ABPP, VHPP) \leq thr^A$, then the AB's VM *apoptosizes* the whole AB; thr^A is the *apoptosis threshold* for the AB.
2) If $thr^A < DI(ABPP, VHPP) < 100\%$, then the AB's VM *evaporates* a portion of the EHR for which VHPP has insufficient privacy policy rules.
3) If $DI(ABPP, VHPP) = 100\%$, then the AB's VM *discloses* the entire EHR to the VH.

We use eXtensible Access Control Markup Language (XACML) to specify privacy policies for ABs and VHs [11]. This new verification step is added as the third step of AB verification activities, following the integrity check (Step 1b).

B. Proposed Agent-Based Active Bundle (ABAB) Scheme

We propose ABAB to implement ABs as multi-agent objects, with data-carrying agents, a trust-verification agent, and an audit agent. We are investigating the use of distributed hash tables (DHTs) to implement fully distributed directory services for trust-verification, data-carrying, and audit agents (e.g., using the prefix labeling approach [6]).

VI. Work Status

The conceptual ABAB model is being designed. Presently, we are working on validating this approach via a simulation. For a pilot system, we will use Java Agent DEvelopment Framework (JADE) middleware for implementing and deploying multi-agent systems.

We are also investigating whether our solution can be retrofitted into the legacy EHR-processing software.

VII. Conclusions and Future Work

We propose a solution that provides protection for the patients' EHR during entire EHR lifetime, including its dissemination among different HISes. We argue that this can be achieved through the use of the active bundle (AB) scheme.

There are many issues left for longer-term research, including the following: (i) enhancing the current AB scheme with an automatic trust-privacy negotiation; (ii) using ABTTP/ABAB to protect patients' privacy in public and private healthcare cloud computing [12];[4] and (iii) protecting patients' privacy in electronic prescription transfer (EPT) [11].

[3] The *PP closure* is the set of all rules that can be inferred from the PP rules.

[4] The classic definition of security is: "*security = confidentiality + integrity + availability (CIA)*." A healthcare cloud must provide proper data availability, while the ABs stored in the cloud must assure confidentiality and integrity of data they encompass.

References

[1] J.A. Akinyele, M.W. Pagano, M.D. Green, C.U. Lehmann, Z.N.J. Peterson, and A.D. Rubin, "Securing Electronic Medical Records Using Attribute-Based Encryption on Mobile Devices," *Proc. 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM'11)*, New York, NY, October 2011, pp. 75-86.

[2] L. Ben Othmane, "Active Bundles for Protecting Confidentiality of Sensitive Data throughout Their Lifecycle," Ph.D. Dissertation, Department of Computer Science, Western Michigan University, Kalamazoo, MI, December 2010.

[3] L. Ben Othmane and L. Lilien, "Protecting Privacy in Sensitive Data Dissemination with Active Bundles," *Proc. 7th Annual Conf. on Privacy, Security and Trust; 2009 World Congress on Privacy, Trust and Management of e-Business*, Saint John, New Brunswick, Canada, August 2009, pp. 202-213.

[4] J. Benaloh, M. Chase, E. Horvitz, and K. Lauter, "Patient Controlled Encryption: Ensuring Privacy of Electronic Medical Records," *Proc. 2009 ACM Workshop on Cloud Computing Security*, New York, NY, November 2009, pp. 103-114.

[5] J. Bhattacharya, S.K. Gupta, and B. Agrawal, "Protecting Privacy of Health Information through Privacy Broker," *Proc. 39th Annual Hawaii Intl. Conf. (HICSS)*, vol. 9, Kauai, Hawaii, January 2006, pp. 1-10.

[6] J.J. Garcia-Luna-Aceves and D. Sampath, "Scalable Integrated Routing using Prefix Labels and Distributed Hash Tables for MANETs," *Proc. IEEE 6th Intl. Conf. on Mobile Adhoc and Sensor Systems (MASS 2009)*, Vancouver, Canada, October 2009, pp. 188-198.

[7] M.A. Hall, "Property, Privacy, and the Pursuit of Interconnected Electronic Medical Records," *Iowa Law Review*, vol. 95(2), Iowa City, Iowa, February 2010, pp. 631-663.

[8] D. Lin, P. Rao, E. Bertino, and J. Lobo, "An approach to evaluate policy similarity," *Proc. 12th ACM Symposium on Access Control Models and Technologies (SACMAT 20007)*, New York, NY, June 2007, pp. 1-10.

[9] C. Linn and S. Debray, "Obfuscation of Executable Code to Improve Resistance to Static Disassembly," *Proc. 10th ACM Conf. on Computer and Communic. Security*, New York, NY, October 2003, pp. 290-299.

[10] R.M. Salih and L.T. Lilien, "The Active Bundle Scheme for Protecting Electronic Medical Records," *Trans. of the Intl. Conf. on Health Information Technology Advancement*, Vol. 1(1), Western Michigan University, Kalamazoo, MI, October 2011, pp. 109-115.

[11] R.M. Salih, "Using Agent-based Active Bundle Scheme for Protecting Privacy in Healthcare," Ph.D. Proposal, Department of Computer Science, Western Michigan University, Kalamazoo, Michigan, May 2012 (in preparation).

[12] R.M. Salih, L.T. Lilien, and L. Ben Othmane, "Protecting Users' Privacy in Healthcare Cloud Computing with ABTTP," submitted for publication.

[13] O. Shimrat, "Cloud computing and healthcare: Bad weather or sunny forecast citation," *San Diego Physician Magazine*, San Diego, California April 2009, pp. 26-29.