Security Middleware Infrastructure for Medical Imaging System Integration

Weina Ma, Kamran Sartipi, Hassan Sharghi

Department of Electrical, Computer and Software Engineering, University of Ontario
Institute of Technology, Oshawa, L1H 7K4 Canada
{Weina.Ma, Kamran.Sartipi, Mohammadhassan.Sharghigoorabi}@uoit.ca

Abstract— **With the increasing demand of electronic medical records sharing, it is a challenge for medical imaging service providers to protect the patient privacy and secure their IT infrastructure in an integrated environment. In this paper, we present a novel security middleware infrastructure for seamlessly and securely linking legacy medical imaging systems, diagnostic imaging web applications as well as mobile applications. Software agent such as user agent and security agent have been integrated into medical imaging domains that can be trained to perform tasks. The proposed security middleware utilizes both online security technologies such as authentication, authorization and accounting, and post security procedures to discover system security vulnerability. By integrating with the proposed security middleware, both legacy system users and Internet users can be uniformly identified and authenticated; access to patient diagnostic images can be controlled based on patient's consent directives and other access control polices defined at a central point; relevant user access activities can be audited at a central repository; user access behaviour patterns are mined to refine existing security policies. A case study is presented based on the proposed infrastructure.**

Keywords— **Security; Middleware; Agent; Medical Imaging; Behaviour Pattern; Access Control**

I. Introduction

Modern Diagnostic Imaging (DI) solutions maintain and manage patient radiology images (e.g., CT scans, Xray, MRI, ultrasound), and corresponding diagnostic reports in digital formats, for the purpose of diagnosis, treatment improvement and medical science research. Over the past decades, Picture Archiving and Communication Systems (PACS) have taken a dominant role in the workflow of DI solutions in a single hospital or radiology department. A federated DI domain allows for a centralized capture, long-term archiving and non-proprietary sharing of radiology information across a large distributed network. A central diagnostic imaging repository (DI-r) provides common services to the participating hospitals. According to the status of DI-r projects across Canada [1], 19 provincial DI-r's have been developed or being developed to reliably maintain, deliver and share DI information to consumers within the electronic health record (EHR) systems. Meanwhile, mobile health information technology (mHealth) is increasingly important in telemedicine, but traditional security infrastructure deployed in PACS and DI-r systems is not ready for accessing DI records through mobile devices.

Integrating the Healthcare Enterprise (IHE) has developed a number of integration profiles [2], [3] that address security requirements to improve the way computer systems in healthcare share information. These security control requirements are achieved through a trusted model where each local medical imaging system is responsible for ensuring that the personal health information is adequately protected. A key challenge with this trusted model is the lack of federated capabilities: i) access control rules are local to each system, which means consistency of access rules across all systems has to be managed manually; ii) patient consent directives and their impact on access control are not communicated automatically to each system; iii) user authentication is local to each system that imposes a significant administrative burden to ensure that individuals are uniformly identified in each system; iv) access to data is audited in each local system which also imposes a significant burden to investigate inappropriate access or monitor security breaches.

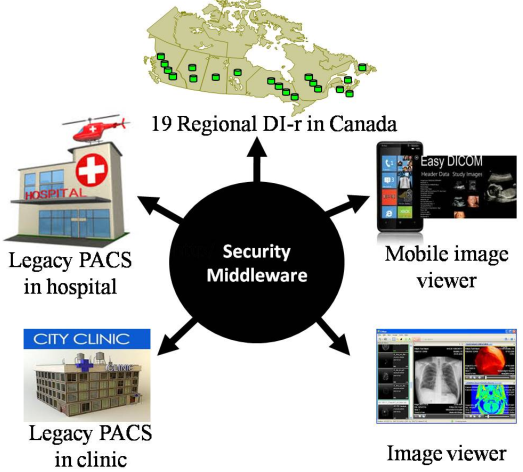

Figure 1. Security middleware integrating with medical imaging service providers and consumers

Middleware is a software layer that lies between service providers and consumers in a distributed computer network. Our proposed security middleware enables secure radiology image sharing among different provincial DI-r's, heterogenous

PACS systems in distributed hospitals, as well as web clients and mobile clients (Figure 1). The main objective of this study is to propose the infrastructure for development of security middleware that provides: i) online security mechanism such as common authentication and authorization method; ii) post security mechanism that assists system administrators discovering user access behavior patterns; and iii) constructing behavior based authorization polices.

II. Previous work

IHE is an initiative by healthcare professionals and industry which aims at setting up consolidated healthcare information sharing through standards based approaches [4]. It guides enterprises in using established standards to accomplish interoperability based on existing IT infrastructure. However, IHE suggested trust model in cross-enterprise domains lacks federated capabilities. Also, the small scale medical service providers are not able to make reliable and accurate authorization decision independently, especially in cloud and mobile computing environment. In such context, we introduce a security middleware that provides one common method for integrating various medical service providers.

A software agent is a program that acts on behalf of an agency for different users or other programs. The notion of generic and lightweight agent that resides at client side to be utilized by different service providers is introduced in [5]. Such agents can be customized and trained based on the service provider generated role description and knowledge to perform their assigned tasks. Such technology is an extension of the SOA model that allows for providing personalized services and maintaining client privacy through processing client's data locally. In our proposed framework, we use cooperative-agents that reside at both client side and service side to interact with security middleware and perform their assigned tasks.

In an earlier work [6], we proposed a general secure sharing infrastructure of medical images between PACS and EHR systems. The proposed environment in that work was based on federated authentication and authorization techniques (OpenID and OAuth), and cooperative agents with dedicated tasks to provide both action-based and behaviour-pattern based access control. As for legacy PACS systems, an agent-based approach [7] is proposed allowing for capturing PACS communication messages, identifying PACS users and extracting user actions to feed into an action-based access control mechanism.

Most of the existing access control models deal only with static systems. Behaviour-based access control for distributed healthcare systems is initially introduced in [8]. The proposed access control model captures the dynamic behavior of the user, and determines access rights through comparing with the expected behavior. Ideally, the distance between observed behavior and expected behavior is significant if the user acts abnormally. This model is also applied in security sharing of medical images [6].

Despite the placement of security mechanisms such as authentication, authorization and secure communication in most systems, authorized users, intended or carelessly, exhibit risky behaviours that may cause data leakage or damage to protected resource. Examining human behaviour among authorized users is helpful in assisting security professionals in making decisions. Our proposed security middleware provides: online security services to identity and authorize user access; and post security services to monitor and analyse the authorized user's access behaviour patterns. Such acquired knowledge can be used in security policy enhancement.

We employ data mining technology in user access behaviour discovery. Association mining is firstly introduced in [9], aiming at analysing customer purchase habit by finding association between items in customer shopping baskets. Sequential pattern mining is proposed in [10], detecting frequently occurring ordered events or subsequence as patterns. There are many applications involving sequenced data, such as customer shopping sequences, web click streams, and biological sequences. Both association mining and sequential pattern mining are used in our approach.

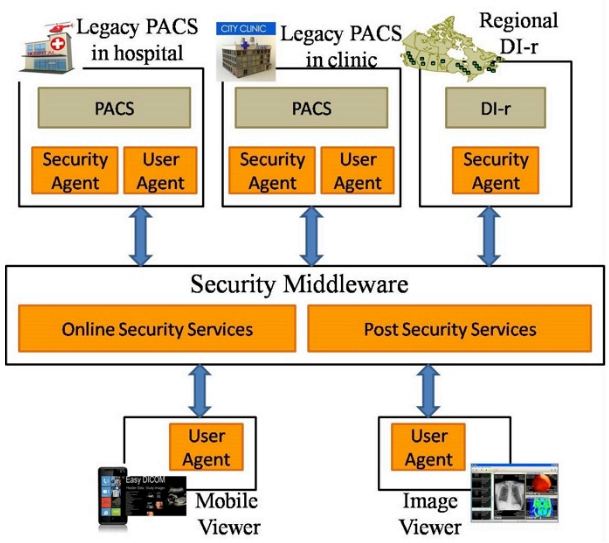

Figure 2. Generic architecture for security middleware integrating with legecy PACS, DI-r's and client applications

III. PROPOSED FRAMEWORK

The proposed approach works as follows: a software agent named "User Agent" is deployed on client side for making authentication request against Security Middleware; a software agent named "Security Agent" is deployed on service provider side for making access control decisions and collecting user activities; client's access request is authorized under different access control models in legacy domains but under unified access control polices; Security Middleware monitors and analyses user access behaviour patterns and assists system administrators in consolidating existing access control policies based on acquired knowledge.

A. Proposed Architecture

Figure 2 shows the proposed architecture containing:

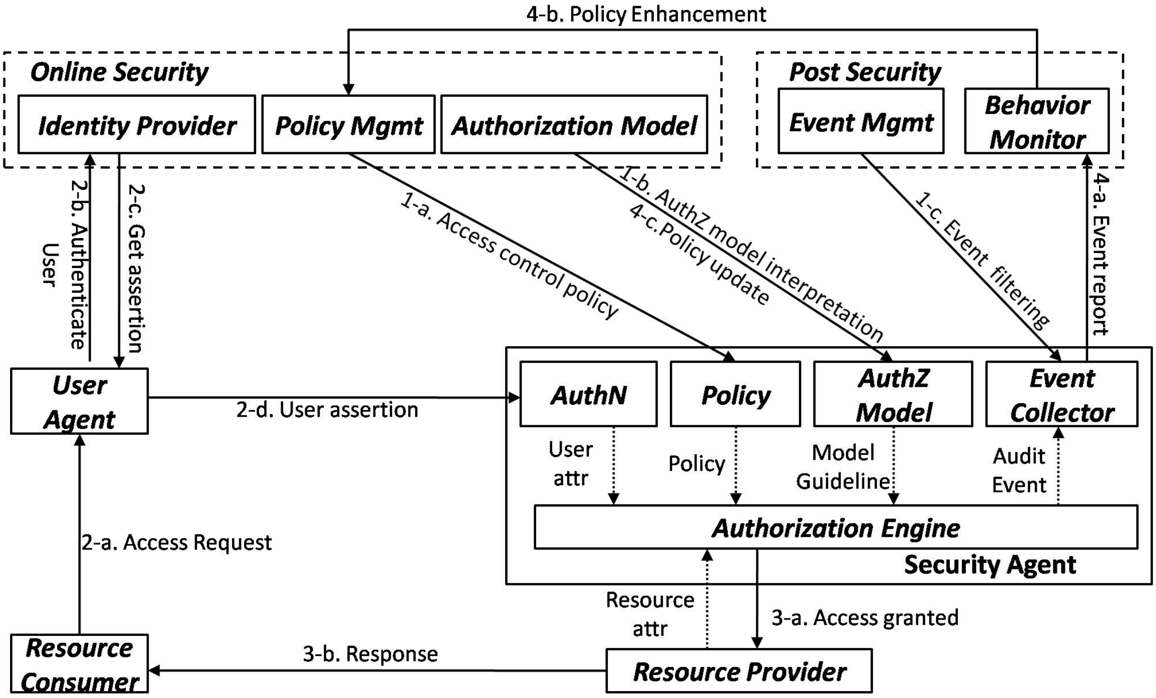

Figure 3. Authentication, authorization and policy enhancemnet workflow of proposed security middleware

User Agent. It is a software agent that acts on behalf of the client application to fulfil authentication flow.

Security Agent. It is a generic agent that is customizable and trainable for different authorization models. The security middleware sends essential information (access control polices), training data (authorization model) and tasks (user activity collection) to customize and train a generic agent. By using the information data and training data, Security Agent acts as local access control system. It also collects local user activities with some filtering criteria for post security services.

Online Security Services. It supports a set of centralized user directories and provides a common service that handles all user authentication requests. It also provides centralized access control policy management and a set of authorization models. The existing IT infrastructure in legacy domains is operating based on different technologies, procedures and models. It is not necessary to employ exactly the same access control mechanism across these domains, but it is reasonable that they will agree at the policy level.

Post Security Services. It defines the types of activities that must be controlled from legacy systems. A data mining engine and knowledge-based decision support system are employed to assist system administrators obtain deep insight into the user access behaviour patterns.

A typical PACS system contains: image acquisition devices namely modalities (e.g., CT scan, MRI system); image archives where the acquired images are stored; and workstations where radiologists view the images. Both User Agent (serving workstations and modalities) and Security Agent (serving image archives) are deployed at each PACS system. The DI-r provides registry services for querying patient's medical images from legacy systems, and repository service for storing and retrieving medical images. Security Agent is deployed to each DI-r system serving such services.

B. Workflow Model

The overall workflow model is shown in Figure 3. The steps of the model's operations are as follows.

1) Security Agent customization (1-a to 1-c). Security Middleware generates the required training knowledge to train the generic Security Agent. The training knowledge is defined as a set of: 1-a) rule based access control polices that are applicable to the protected resources; 1-b) authorization model interpretation that defines the access control procedure, and information provider such as user attribute provider and resource attribute provider; 1-c) event filtering criteria to collect user access activities. Security Agent receives the provided knowledge as well as the relevant resource-provider's context, and then modifies the general authorization process and event collection task.

2) Authentication (2-a to 2-d). Resource Consumer sends an "access request" message to the Resource Provider; and the Resource Provider employs User Agent to fulfil the authentication flow (2-a). Identity Provider is an identity authentication server that is capable of authenticating end users (2-b) and provides "assertions" containing authentication statement and user attribute statement (2-c). A security assertion is transferred between identity provider and service provider for exchanging authentication and authorization data. Authentication statement confirms that the user has been identified and authenticated with the authentication server; the attribute statement asserts that the user is associated with certain attributes (2-d). These asserted attributes feed Security Agent to make access control decisions.

3) Authorization (3-a to 3-b). Resource Provider sends instructions to Security Agent to perform authorization. Security Agent constitutes the following components: Authorization Engine that evaluates applicable policies and renders an access control decision; AuthN that provides the user's associated attributes; Policy that contains security middleware assigned policies and sends available policies to Authorization Engine for a specified target; AuthZ Model that guides authorization engine to fulfil the agreed-on authorization procedure; Event Collector that records authorization decisions. If this request is granted, Security Agent sends an access request to Resource Provider (3-a). Resource Provider serves the request and returns its response to service consumer (3-b).

4) Behaviour pattern mining and policy enhancement (4-a to 4-c). Behaviour pattern is defined as consistent observations of a sequence of actions performed by the same user, under certain environment as well as during a specific time interval. Event Collector sends the collected data to Behaviour Monitor component after filtering out the uninterested events (4-a). A knowledge driven behaviour pattern discovery process is applied to orchestrate user's common behaviour and abnormal behaviour. Finally, the system administrators explore the opportunities to refine existing security policies by means of analysing salient features and characteristics of the discovered behaviour patterns (4-b). Then the consolidated polices are dispatched to corresponding Security Agent to take effect (4-c).

C. *Behaviour monitor*

We apply an association mining engine on collected events to extract sharing attributes among events. The number of sharing attributes as well as the number of shared events measures the similarity between events. Such an association-based similarity metric is used for clustering highly related events in the system. After a clustering phase based on this association-based similarity metric, sequential pattern mining is applied on each cluster to extract frequent behaviour patterns. The recommendation system guides system administrators to investigate the properties of the recovered behaviour patterns, such as the actor who issues the behaviour, the order of actions that represent the behaviour, the context under which the behaviour is performed, and the time window of the behaviour occurrence. Based on the characteristics of the recovered common or outlier behaviour patterns, there are huge possibilities to identify anomaly behaviour of some users and to profile the expected behaviour. Finally, the system administrators are capable of refining existing authorization polices through comparing and analysing the gap between the knowledge acquired from recovered behaviour patterns and the existing polices.

IV. Case study

In this section, we present an end-to-end case study to examine our proposed approach.

A. *Implementation*

We developed a prototype and applied on a simulated legacy PACS system and DI-r. ClearCanvas [11] is an open source implementation of a PACS viewer. A User Agent is deployed on the workstation to assist the ClearCanvas viewer to render the authentication flow. Health information exchange open source (HIEOS) [12] is an open source implementation that is used to simulate a set of DI-r web service interfaces to retrieve images. A generic Security Agent is deployed in front of HIEOS to perform authorization flow. Security middleware and DI-r make an agreement about applicable authorization policies, authorization model, and event filtering criteria. Security Agent is trained based on the security middleware generated training knowledge to perform its tasks.

B. *Online security services*

Let us consider a scenario where a user intends to use a PACS viewer application to display a patient's diagnostic report that is stored at the DI-r. One applicable authorization policy in this case is "*Only physicians are allowed to view and change a patient's diagnostic reports; other healthcare staff only has the privilege of viewing the patient's diagnostic reports.*" Identity Provider issues an assertion including the statement of user's role "physician" after authenticating the end user. Resource Provider supplies the resource type as "diagnostic report" and the resource owner as "patient". Authorization engine grants this access request after evaluating the applicable policies with attribute values.

C. *Post security services*

The system kept running over one month and the Behaviour Monitor component totally collected 3000 user access events from the DI-r. These events are parsed and converted to attributed events. Each event is described by following attributes: "User(U), Role(R), Location(L), Operation(O), Resource owner(W), Resource(E), Date(D), Time(T)". Each attribute value is represented by a quantitative value (e.g., L-1 means location "Oshawa"; L-2 means location "Toronto"; R-1 means role "physician"; R-2 means role "nurse").

The Apriori algorithm [9] is applied on the attributed events for discovering highly associated groups of events, where all events in one group share the same set of attribute values. We refer to the group of events as *basketset* and the shared set of attribute values as *itemset*. We define an association-based similarity metric between two events, which encodes both the size of basketset and the length of itemset. Figure 4 is a visualization of the relationship among events. This graph is generated by Gephi [13], an open source network analysis and visualization software package. The undirected graph illustrates the association between events according to our defined similarity metric. Each node represents an event, and each weighted edge represents the similarity value between two events. The events are grouped into a couple of clusters. Our approach allows an event being assigned to multiple clusters.

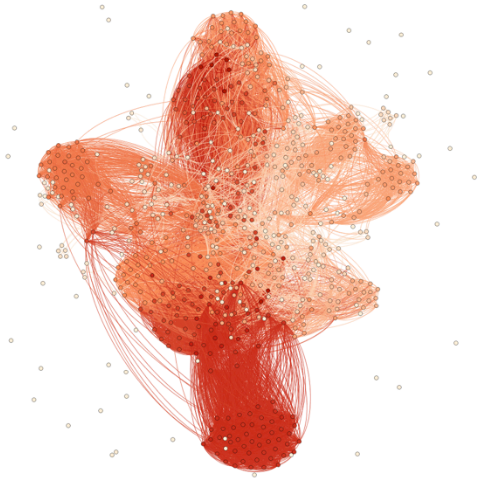

Figure 4. Visualization of assocation between events

Sequential pattern mining algorithm CloSpan [14] is employed to discover user's daily behaviour in each cluster. Firstly, we convert event database to sequence database where each sequence is a set of ordered events performed by the same user within one day. So that the discovered frequent sequence patterns can be viewed as such user's daily behaviour. In a post-analysis phase, we investigate the characteristics of the discovered sequence patterns in each cluster. For example, what is common among the users who accessed the system around the rush hour? What is the frequent behaviour pattern of a specific user in the system? Through analysing the common attribute values in each item of sequence patterns, context attributes are extracted to describe the circumstances of the complete sequence. Followings are some discovered behaviour patterns in the experiment:

- 50% of users have access requests at most 6 times during rush hour "10:00am".
- 80% of access requests from user "U-22" at location "L-6" are at time "1:00pm".

We can see the busiest time of user "U-22" is different from other users: "U-22" has more access request at 1:00pm but the normal rush hour is 10:00am. The system administers may limit the maximum access request number during rush hour with differentiated policies.

V. CONCLUSION AND FUTURE WORK

This paper contributes to provide a common method for secure sharing medical images among legacy PACS systems and DI-r's. We propose a novel security middleware that replaces the existing trusted model for cross-domain integration. Customizable and trainable software agents are deployed at the legacy systems to fulfil authentication flow, to make authorization decisions as well as to collect user activities. Besides online security services, the security middleware also provides post security services to recover the dynamic user access behaviour patterns. Finally, the recommendation system helps system administrators to explore the opportunities to refine existing security policies though analysing salient features and characteristics of discovered behaviour patterns.

We plan to extend our work to provide step-by-step guidance tool throughout the whole policy enhancement process such as: i) investigating the characteristics of the extracted behaviour patterns and committing recommendations to identify common behaviour and abnormal behaviour; ii) detecting system security policy vulnerabilities and providing reasonable advice on policy consolidation.

ACKNOWLEDGMENT

This research was conducted with collaboration of Dr. David Koff and Dr. Peter Bak at MIIRCAM Centre of McMaster University. This research was funded by an ORF grant for the project "Secure Intelligent Content Delivery System for Timely Delivery of Large Data Sets in a Regional/National Electronic Health Record."

REFERENCES

[1] Dossier Santé du Québec, "Diagnostic imaging group, Status of Diagnostic Imaging Repository (DI-r) projects across Canada". Available: http://www.camrt.ca/

[2] IHE IT Infrastructure Technical Framework Integration Profiles Volume 1. Available: http://www.ihe.net/, 2012

[3] IHE IT Infrastructure White Paper for Access Control, Available: http://www.ihe.net/, 2009

[4] Integration the Healthcare Enterprise website. Available: http://www.ihe.net

[5] N. Mehran, and K. Sartipi, "Modeling service representatives in enterprise systems using generic agents," *Service Oriented Computing and Applications (SOCA),* vol. 5, pp. 245-264, Dec. 2011.

[6] K. Sartipi, K. Kuriakose, and W. Ma, "An Infrastructure for Secure Sharing of Medical Images between PACS and EHR Systems," *International Conference on Computer Science and Software Engineering (CASCON),* pp. 245-259, 2013

[7] W. Ma, and K. Sartipi, "An Agent-Based Infrastructure for Secure Medical Imaging System Integration," *in Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on,* pp. 72-77. IEEE, 2014.

[8] M. H. Yarmand, and K. Sartipi, and D. G. Down, "Behavior-based access control for distributed healthcare systems," *Journal of Computer Security,* 21.1, pp. 1-39, 2013

[9] A. Rakesh, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *in ACM SIGMOD Record,* vol. 22, no. 2, pp. 207-216. ACM, 1993.

[10] A. Rakesh, and R. Srikant, "Mining sequential patterns," *in Proceedings of the Eleventh International Conference on. IEEE,* pp. 3-14. IEEE, 1995.

[11] Open Source ClearCanvas PACS Website. [online]. Available: http://www.clearcanvas.ca/

[12] Open Source HIEOS Website. [Online]. Available: http://sourceforge.net/projects/hieos/

[13] Gephi - The Open Graph Viz Platform Website. [Online]. Available: http://gephi.github.io/

[14] X. Yuan, J. Han, and R. Afshar, "CloSpan: Mining closed sequential patterns in large datasets,"i*In Proceedings of SIAM International Conference on Data Mining,* pp. 166-177, 2003