Sec-Shield: Security Preserved Distributed Knowledge Management Between Autonomous Domains

Petros Belsis[1], Stefanos Gritzalis[1], Apostolos Malatras[2], Christos Skourlas[3],
and Ioannis Chalaris[3]

[1] Department of Information and Communication Systems Engineering,
University of the Aegean, Karlovasi, Samos, Greece
{pbelsis, sgritz}@aegean.gr
[2] Department of Electronic Engineering, Centre for Communications Systems Research,
University of Surrey, UK
a.malatras@surrey.ac.uk
[3] Department of Informatics, Technological Education Institute, Athens, Greece
{cskourlas, ixalaris}@teiath.gr

Abstract. Knowledge Management (KM) comprises of a variety of distinct technologies and techniques, relative to the uniform treatment of tangible and intangible resources. Attempts to extend the traditional single organizational resource-sharing scheme, confront various challenges, relative to the management of security and heterogeneity issues. In this paper we discuss the various security models, presenting potential limitations - as well as the advantages - relative to their support to extend the single-domain security management framework, to a resilient and robust distributed multi-domain Knowledge Management scheme. We present the architecture of a security enhanced prototype that supports decentralization, while it maintains the autonomic character of the participating domains. We also argue about the implementation dependent choices relative to the alleviation of the multifaceted problems that a collaborative Inter-organizational knowledge asset exchange framework arises.

1 Introduction

Knowledge Management (KM) systems emerged during the last decade and rapidly transformed into a basic business function for many organizations; still though, their flexibility is limited within the borders of a single organization. Among others, a serious challenge is the expansion of the capabilities of such a system to utilize knowledge assets from other organizations according to Nonaka's spiral model [1], with the basic prerequisite that this management of knowledge assets will happen efficiently and through automated, transparent procedures from the user's perspective.

Ordinary KM systems attempt to provide the user with the necessary knowledge to efficiently fulfill her tasks and by doing so, to raise her productivity as well as the overall organization's response to new emerging challenges that demand accurate, constantly updated and on time-fetched knowledge. Still, when it comes to attempt to utilize knowledge from distinct organizations through engagement in a cooperative framework, serious obstacles are posed that retard knowledge exchange and diffusion.

S. Katsikas, J. López, G. Pernul (Eds.): TrustBus 2005, LNCS 3592, pp. 10 – 19, 2005.

The establishment of the necessary pre-coalition procedures and the exact definition of the level of mutual sharing of vital knowledge sources is a long-term and time consuming procedure that poses an important overhead on the overall process. Our system focuses on providing with sufficient solutions towards the alleviation of this problem: first by introducing a scalable and robust solution for correlating roles between different organizations, and second by treating heterogeneity problems which are a commodity between different information systems.

Our goals are: a) to enable the realization of cooperation between autonomous, policy-managed Information Systems and b) to identify the distributed knowledge assets transparently, using agent and ontology technology.

The rest of the paper is organized as follows: after a brief introduction in section 1, section 2 presents the basic concepts and requirements related with distributed KM as well as a review of related work on the area; section 3 presents security architectures supporting the necessary collaborative frameworks; section 4 presents authorization schemes able to support the demands of similar architectures, together with our choices which ensure scalability among other characteristics; section 5 presents the architecture of our developed prototype, while section 6 concludes the paper and provides the directions of our future work.

2 Distributed KM – Related Work

The advent of emerging technologies such as portable devices, which enhance decentralization of resources, directs traditional KM techniques in failure to meet their initial expectations. Users often create knowledge ad-hoc and use their own individual IT infrastructure [2]. Although the need for decentralized KM solutions is obvious, the amplitude of the field of solutions is still very narrow. A number of both technological as well as socio-technical aspects of the problem pose interesting research challenges [3]:

- Heterogeneity (semantic, syntactic).
- Security.
- Network efficiency [4].

In [5], a conceptual architecture is presented for a system based on the notion of trust for distributed KM. This system (ADAM), utilizes agent-technique to perform knowledge discovery and authorization. ADAM architecture is based on a pair of agents one responsible for querying for knowledge and the other handling the authorization issues. This system though, manages mainly knowledge about its users and bases the authorization process on grounds of reputation collected for a user from other nodes. Even though it handles scalability issues very efficiently, this system gives the chance to somebody to create a new identity or retain multiple identities concurrently and attempt to enter into relations with the system. The application of these principles on systems such as Internet transactions (e-commerce) where a security failure could direct to financial is not doubted for its validity. ADAM authorizes transactions and not users. Furthermore, it functions on total absence of explicitly stated organizational policy.

XAROP [4], is a peer-to-peer system which utilizes ontologies for handling heterogeneity issues arising from the different conceptualizations among different domains. The notion of security is rather simplistic and cannot be applied to critical environments, comprising of rules manually posed by the user which has to classify for each document separately authorized users or groups. Authentication is based on PKI infrastructure, where root and subordinate certificate authorities are denoted within the XAROP infrastructure.

3 Security Architectures for Collaborative Environments

Security policies emerged in the last decades and have attracted considerable attention in distributed computing, due to their ability to simplify security management and access control enforcement for a large number of heterogeneous components which often span across organizational boundaries [6]. A more complicated situation is related with the attempt to create a policy-managed collaboration scheme between different organizations. In most of the cases, establishing a collaboration access scheme involves negotiation off-line, by extra technological means, and includes complex procedures, such as identification of the negotiating parties, and common agreement every time upon the conditions of sharing a resource, after a new claim has been posed.

Two kinds of systems can be considered under this (collaborating) framework: peer-to-peer networks, and autonomous domains. Peer to peer networks resemble communities with common interests, the terms of bounding though are more loosely coupled than autonomous domains. The second category of systems can be met in many real-life systems, such as e-government environments, or healthcare systems which consist of several cooperating hospitals. In the latter case, sensitivity of the data poses more security restrictions and establishing a common state for knowledge exchange requires both that organizational roles are well defined in terms of access rights and obligations based on the grounds of a well-stated security policy, while a common access state between different organizations is unambiguously allocated.

We can classify these systems according to the access models they adopt, to the following two categories:

1. Trust based systems. The notion of trust is introduced mainly in complex, non-hierarchical or inter-related systems such as the Internet, where unknown totally roles might be interested to enter into relations between them, or to cooperate on basis of financial terms. This situation is often on Internet transaction systems, such as e-commerce etc. The authorization of a transaction is based on the basis of estimating the cost and the substantial loss for a specific role, considering a prerequisite the potential risk, according to the degree of trust that can be associated to his role, for example by questioning his previous activity or users associated with him.
2. Autonomous systems, with well formed security policy and well defined organizational structure.

We will restrict our scenario to the second category of systems, which are characterized by well defined organizational policy and cooperate and on the grounds of a commonly agreed target, such as improvement of efficiency of the governing infrastructure or the reducing of response times for health treatment of patients within the national healthcare system.

Our approach supports the formation of coalitions, based on the idea of establishing mappings between the policies recorded in XML (Extensible Markup Language [14]) type format and by utilizing XML transformation to database techniques and accordingly converting the policy mapping problem to a database mapping problem, gaining on the same time by obtaining policy non-disclosure (due to the easier manipulation of privacy issues based on database technology than through XML files where the policy is recorded).

4 Authorization Schemes

4.1 Trust Based Versus Policy-Managed Autonomous Systems

We can distinguish two approaches [9] concerning authorization schemes: The first can be applied on distributed environments which cooperate on the grounds of a non-formal negotiation scheme. The second applies to more restricted organizational schemes, which cooperate under a formal framework, where most of the access rules are posed by a strict organizational policy and cooperation is substitute to tight rein.

Loosely coupled authorization scheme

Under this framework, we can distinguish two access-control approaches. The first uses predefined set of role mappings. It requires from the constituent systems to indicate the level of sharing they want to allow and to establish a consistent set of mediation rules for inter-domain access. The second relies on bringing together unknown individuals by examining their credentials and mapping assigning a level of trust which corresponds to a specific level of trust.

Federal environments

In this case, criticality can arise as a key concept. Federal environments have a role common security policy organizational scheme and local roles correspond to a generic

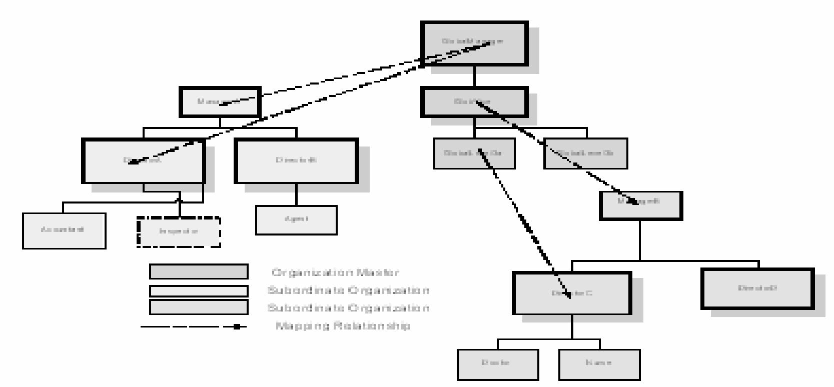

Fig. 1. Role mapping across different domains

representation scheme. These systems can be e-government environments, for example ministries that participate in the federal government infrastructure and share a common policy interpretation model. Typically the roles of a designated system have to map to the generic role representation scheme. In Figure 1 a typical representation of organizational structure and relative roles from federal environments, are mapped through mapping relations (arrows) to a global organizational schema. We utilize security clearance levels which are frequently used for critical environments and reflect typical situations within military or governmental systems. This organizational structure and the interrelated role correlation and security attributes can be – together with other organizational assets – represented in XML files (policy files), something that facilitates -in terms of interoperability- the effective and automated management of resources and simplifies administration tasks as already mentioned.

4.2 Automating the Authorization Process

Mapping between roles has been proposed as a potential solution for multi-domain environments [9] [8]; many issues still remain to be resolved. Recording roles and their attributes needs an appropriate, both human and machine interpretable format, which can be easily integrated in the security policy requirements and codified in means of duties, obligations and permissions as security policy languages demand. For interoperability issues, the usage of languages that export their rules in XML format is highly required.

The mapping process between roles needs additive handling on grounds of interrelating similar documents that record each organization's structure (policy files). In order to automate the authorization process, we apply a direct mapping between security levels (clearances) of different organizations. More specifically, the global role scheme to which all the subordinate organizations have to comply, establishes several security levels. Mapping of domain roles and their security levels to the global role scheme and to the corresponding global level of security clearance is handled by the administrator of the global domain. This is a typical necessity in order to reduce complexity. In our approach we have adopted a general mapping scheme, to which the cooperating domains have to confront, while maintaining their local autonomy in policy declaration. This is a typical practice in real scenarios, such as e-government environments, where establishment of rules is mainly directed centrally, without affecting the establishment of procedures in the interior of subordinate organizations. The mapping process which at the lower level reflects to mapping between XML files is mainly performed by administrators in each domain, who are aware of the legal, ethical implications and the consequences of an incorrect mapping between roles among different domains, while they are also technically capable to handle the mapping details.

4.3 Sec – Shield's Scalable Approach for Mapping Between Roles

We have utilized the XACML [7] policy framework for enabling distributed management of resources. XACML is an XML based framework for specifying and applying access control for Web-based resources. XACML specification supports both identical and role based access control and incorporates contextual information such as location and time and under several extensions XACML can be applied also to secure Web-services environments [10].

The administrator for each domain is editing the local policy and classifies the access rights to the resources for each role within the organizational borders. For each domain an XML file is retained, determining the access rights and the security clearance level for each person. A mapping between the general role scheme and the role schemes for the local domains is an administrator's task. An appropriate ontology which is an essential part of the system can maintain information about the security clearance level of certain roles or individuals within an organization.

5 Sec-shield's System Architecture

We will refer to the key concepts of our implemented prototype, the functionality of which continuously arises; namely the design of agents in our framework, the role of ontology and techniques utilized, and we will describe the overall architecture of the platform.

5.1 The Role of Agent and Ontology Technologies

The presence of agents was decided in order to enable transparent identification of assets and to provide automated authorization for users. For each domain an agent performing the knowledge assets identification, while another one carries the user credentials and according to the security policy and the role the user is assigned to, provides her with access to the resource or in the opposite case denies access to the resource. The agents in our system were implemented with the aid of the JADE platform [11]. The agents exchange communication messages based on the agent specific FIPA-ACL language [15].

In order to handle semantic heterogeneity issues between the different domains [3], the Resource Description Framework (RDF) [13] technology is being utilized. The RDF ontology enables upon querying, provision of semantically enhanced answers.

For example, a user upon providing a query for an expert who is specialized on certain field, the system will look only to return experts who are specialized in the specific area and will regard other knowledge sources that are related to the specific subject. In order to enable performance optimization and in order to avoid information disclosure to the agents such as the security clearance level for a specific role, a transformation of the RDF file to relational database is performed.

This enables better optimization in terms of avoiding network congestion when performing queries on the central ontology. For the transformation process of RDF to relational schema, the hybrid inlining algorithm [12] has been utilized. Conceptually, we consider a Document Type Definition (DTD) of an XML (or alternatively an RDF) file as similar to a schema in relational databases. Having this in mind, the process of storing and querying RDF files to a relational database, consists of transforming basically the DTD to relational tables. A sample of DTD used for our purposes is represented in Table 1.

<?xml version="1.0" encoding="UTF-8"?>		
<! ELEMENT	SecurityFileXL	(Security*)>
<!ELEMENT	Security	(Field+,ComplexField+,Field+,Researchers+, Field+)>
<!ELEMENT	Field	(title, Resource+)>
<!ELEMENT	Title	(#PCDATA)>
<!ELEMENT	Resource	(title, url, description)>
<!ELEMENT	url	(#PCDATA)>
<!ELEMENT	description	(#PCDATA)>
<!ELEMENT	ComplexField	(title, Field+)>
<!ELEMENT	Researchers	(Researcher+)>
<!ELEMENT	Researcher	(Name,Email,PersonalPage,ClearanceLevel>
<!ELEMENT	Name	(#PCDATA)>
<!ELEMENT	Email	(#PCDATA)>
<!ELEMENT	PersonalPage	(#PCDATA)>
<!ELEMENT	ClearanceLevel	(#PCDATA)>

Now the basic idea is to transform first the DTD to a graph, and accordingly for the major concepts to create relational tables. The graph of the DTD of Table 1 is presented on Fig 2.

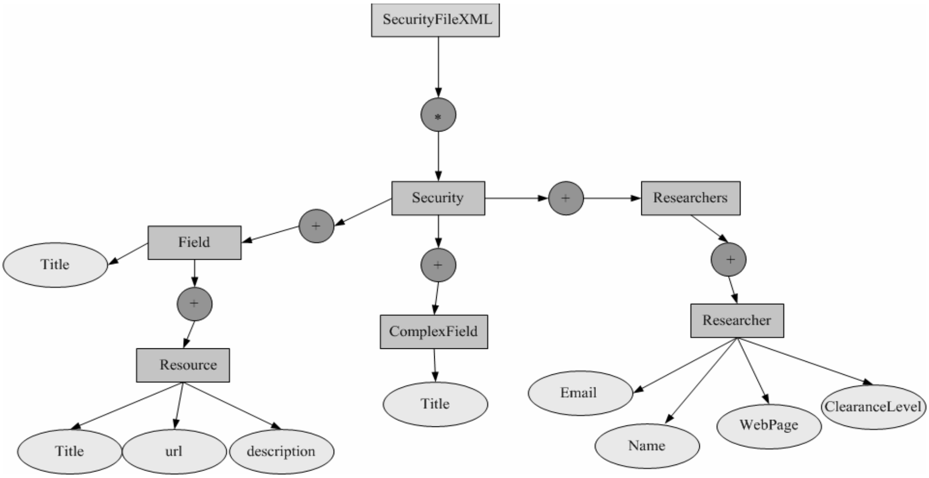

Fig. 2. DTD to relational schema transformation

Accordingly, for concepts like the researcher (domain expert), which represents a role within one domain, a relational table can be defined. One of the major attributes defined for our scenario, is the clearance level. The value stored in this attribute can be easily hidden for non-disclosure purposes to all non authorized users, with the facilities most DBMS's provide and can be easily retrieved for authorized purposes with easy to form SQL queries. Therefore, we edit security policies in XML format, accordingly this XML-type policy transformed and stored in relational schema, which can be further queried for policy mappings and can be efficiently protected against non-authorized disclosure.

5.2 Overall *Sec-shield* Architecture

Our prototype implementation consists of an organizational memory, consisting of the organization's past experience codified in semi-structured documents, while at the same time we correlate the document-based information with each domain's human network of experts. Support is provided also for multimedia files (images, videos) through a special purpose repository implemented in Java and Oracle 9i. For multimedia file retrieval purposes, a set of meta-data is stored in the organizational memory module. This architecture is deployed in different domains, each one maintaining its own autonomy. For each domain there is a policy decision point (PDP) which directs the policy enforcement point (PEP) to provide -or not- access to distributed resources of the system upon a user's request.

Upon a user query for a topic of his/her interest, initially the local document management module is utilized and accordingly, the knowledge discovery agent queries the other domains for similar knowledge sources. The messages exchanged between the domain specific agents are based on FIPA protocols [15], and the content embodied is based on the RDF ontology, which plays also a key role relative to the facilitation of heterogeneous assets knowledge discovery. After resources identification is performed, the next step is related with the authorization process activation. For transparency reasons, the authorization process will be treated through the authorization agent (Auth-agent, fig3). The authorization agent provides the user

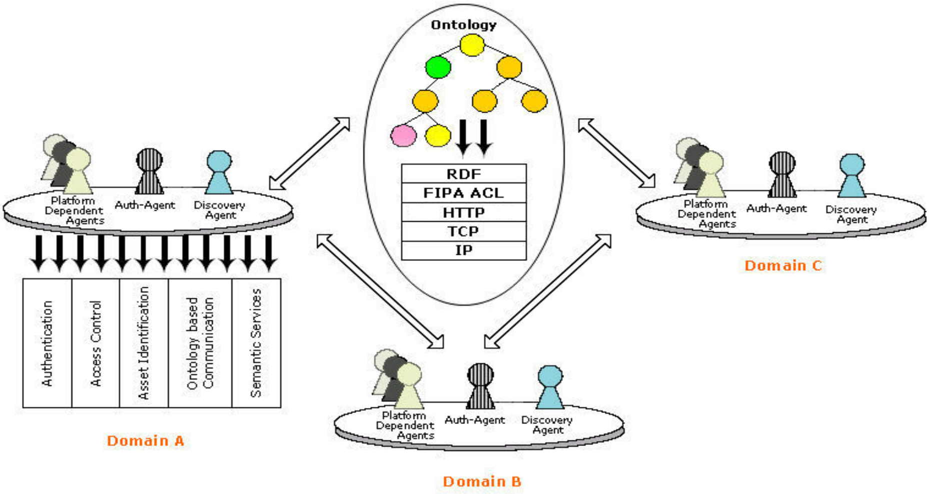

Fig. 3. Overall system architecture

credentials, and the security clearance associated with the user's role on the local domain, is exalted to the global security role mapping-scheme, as mentioned in section 4.1. Accordingly, a role on the remote domain is assigned to the user and on the basis of the remote security policy, the PDP authorizes or not the user upon the requested resources. Therefore, the user is provided with the chance to utilize knowledge from multiple domains transparently, where all the asset discovery procedures and authorization between the domains, are treated by the system, through the use of the pair of Auth-Agent and Discovery-Agent, assigned to each domain.

6 Conclusions

Contemporary KM approaches suffice to utilize knowledge residing in different organizational domains, limiting the resource sharing potential of the developed KM systems. *Sec-shield* pays special emphasis on covering this aspect. *Sec-shield* is characterized by its extended functionalities relative to multi-domain knowledge utilization, such as heterogeneous files management (images, text), transparency relative to knowledge asset identification, user authorization and access control enforcement. Our solution while it maintains its scalability potential, it is characterized by its robustness and supports well defined policy frameworks in comparison with other approaches [4] that put more emphasis on knowledge sharing rather than on access control enforcement. Based on its policy dependent security management, it can support large scale infrastructures with frequent changes in the policy specification or the number of participating users. In the future, we plan to expand our framework to integrate the identification and authorization of knowledge assets through the creation of Web-services running independently for each domain.

Acknowledgments

This work was co-funded by 75% from E.E. and 25% from the Greek Government under the framework of the Education and Initial Vocational Training Program – Archimedes. The authors would like to thank John Varnas for aiding at the improvement of the aesthetics of the paper, by providing assistance with the drawings.

References

1. Nonaka I., Takeuchi H. (1995). "The knowledge Creating Company", Oxford University Press.
2. Bonifacio M., Bouquet P., Danieli A., Dona A., Mameli G., Nori M.: "Keex: A peer-to-peer solution for distributed Knowledge Management". In Tochtermann K., Maurer H. eds.: Proceedings on the 4th International Conference on Knowledge Management Graz Austria, 2004

3. Belsis P., Gritzalis S.: ‘‘Distributed autonomous Knowledge Acquisition and Dissemination ontology based framework’’, in Proceedings of PAKM 04 5th International Conference on Practical Aspects of Knowledge Management – Workshop on Enterprise Modeling and Ontology: Ingredients for Interoperability H. Kuhn (ed.) Dec. 2004 Vienna Austria, Univ. of Vienna.
4. Tempich C., Ehrig M., Fluit C., Haase P., Marti E.L., Plechawski M., Staab S. “XAROP: A Midterm Report on Introducing a Decentralized Semantics based Application, Proceedings of Practical Aspects of Knowledge Management (PAKM) 2004, Vienna Austria, LNAI 3336 Kluwer Academic publishers, pp. 259-270.
5. Seleznyov A., Mohamed A., Hailes S. “ADAM: An agent-based Middleware Architecture for Distributed Access Control” Twenty-Second International Multi-Conference on Applied Informatics: Artificial Intelligence and Applications, 2004.
6. Damianou, N., N. Dulay, E. Lupu and M. Sloman . Managing Security in Object-based Distributed Systems using Ponder. In Proceedings of the 6th Open European Summer School (Eunice 2000), Enchede, The Netherlands, 13-15 September 2000.
7. Organization for the Advancement of Structured Information Standards (OASIS), ‘‘XACML Extensible access control markup language specification 2.0’’, OASIS Standard, (available at http://www.oasis-open.org
8. Belokosztolski A., “Role based access control for policy administration”, available at http://www.cl.cam.ac.uk/ as technical report No 586, university of Cambridge, UK.
9. Joshi J.B.D., Bhatti R., Bertino E., Ghafoor A., “Access Control Language for Multi-Domain Environments”, IEEE Internet Computing, Nov. 2004
10. Bhatti R., Bertino E., Ghafoor A., Joshi J.B.D., XML-based Specification for Web services Document Security. IEEE Computer, April 2004, pp. 41-50.
11. The JADE agent development kit. Available at http://jade.tilab.com/
12. Lee Dongwon, Chu Wesley, (2001) CPI: Constraints- Preserving Inlining algorithm for mapping XML DTD to relational schema, Data and Knowledge Engineering, 39, pp. 3-25.
13. S. Decker, S. Melnik, F. van Harmelen, D. Fensel, M. Klein, J. Broekstra, M. Erdmann, I. Horrocks, The semantic web: the roles of XML and RDF, IEEE Internet Comput. 4 (5) (2000) 63–74.
14. Extensible Markup Language Specification (XML), http://www.w3.org/XML/.
15. FIPA standard status specifications www.fipa.org/repository/standardspecs.html