Future Generation Computer Systems 55 (2016) 227–237

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

# The mediator authorization-security model for heterogeneous semantic knowledge bases

Abdullah Alamri [a,*], Peter Bertok [b], James A. Thom [b], Adil Fahad [b,c]

[a] Faculty of Computing and Information Technology, University of Jeddah, Jeddah, Saudi Arabia
[b] School of Computer Science and Information Technology, RMIT University, Melbourne, VIC 3000, Australia
[c] Department of Computer Science, Al-Baha University, Al-Baha, Saudi Arabia

ARTICLE INFO

*Article history:*
Received 14 November 2013
Received in revised form
14 September 2014
Accepted 4 March 2015
Available online 23 March 2015

*Keywords:*
Authorization
Semantic mediator system
Semantic web
RDF
OWL
Semantic repositories

ABSTRACT

Many organizations often need to share semantic knowledge base content with selected members of other organizations. However, sharing semantic knowledge across different organizations is a critical problem. This is because the differences in the vocabulary utilized by the organizations have to be resolved before knowledge can be shared. Also, if semantic repositories are syntactically and schematically heterogeneous, information interoperation becomes a vital challenge. When a system needs to allow unknown entities to access its resources, mechanisms should be in place in order to provide a secure and trusted information-sharing environment and enable users to interact and share information easily and perfectly. To address these challenges, the Mediator Authorization-Security model is proposed to provide secure interoperation among heterogeneous semantic repositories. This paper addresses the issue of interoperability and how to incorporate trust into semantic interoperability. The evaluation showed that, despite the complexity of the mediator system, it still provides acceptable performance.



## 1. Introduction

The major basics of the architecture of the Semantic Web, the currently widely-used semantic models, are utilized to ensure the semantic interoperability of data sources and applications. The need to competently manage these kinds of objects has encouraged the development of specialized repositories, which are normally referenced as semantic ontology databases. Semantic databases and ontologies are expanding and penetrating many areas of information and communication technologies, and most categories of applications. They are gaining attention in many areas including industry, healthcare, content management and life sciences as an efficient means of accomplishing complex information management tasks.

The exchange of information has become a critical factor in many organizations. For example, often organizations which are largely autonomous, distributed and heterogeneous in various aspects including their goals, need to collaborate to better achieve common or compatible goals. However, interoperability problems emerge as these organizations may be heterogeneous. Also, trust appears as the main issue to address in order to achieve secure interoperation among heterogeneous semantic repositories. When a system needs to allow unknown entities to access its resources, mechanisms should be in place in order to provide a secure and trusted information-sharing environment and enable users to interact and share information easily and perfectly across many diverse semantic repositories. In this work, we need to enhance security and interoperability in order to enable two or more semantic store systems to exchange information securely and efficiently. The key aim of this paper is to "share but protect" where the motivation to "protect" is to safeguard the sensitive content from unauthorized disclosure. Hence, we present heterogeneous middleware security policies in order to enhance semantic interoperability and address the heterogeneity gap between semantic databases by identifying the related mapping bridge between semantic systems. Fig. 1 clarifies this concept.

Substantial work on semantic ontology security has been the center of attention in controlling access to a single ontology store. This paper highlights the need to enhance authorization security across semantically heterogeneous repositories. Enabling secure information-sharing among heterogeneous semantic repositories faces different challenges:

– How to interoperate among semantically heterogeneous repositories efficiently?

* Corresponding author.
*E-mail address:* Abdulla-Alamri@hotmail.com (A. Alamri).

http://dx.doi.org/10.1016/j.future.2015.03.004
0167-739X/

- How to design fully supported content-based access control to secure shared semantic knowledge base content?
- How to handle the corresponding confidentiality concerns of the organizations involved in information sharing?
- How to design a mediation system model which ensures flexibility of control and secure knowledge sharing heterogeneity?

Heterogeneity presents challenges in terms of developing a flexible model that works well in different semantic technological contexts. Therefore, a certain flexibility is required when designing an efficient authorization control mechanism across heterogeneous repositories. The efficiency requirement means that, in practice, additional restrictions need to be placed on the models to ensure more controlled reasoning. In fact, mediation security measures need to be implemented that can ensure the security policies of one system are respected by the other, and vice versa. The contributions of the research documented in this paper are:

- Propose a highly-structured multi-layered authorization control for safeguarding semantic data across semantically heterogeneous repositories.
- Design TBox access control which protects the semantic models' concepts and their relationships.
- Design a mediator ABox trust level management which grants trust-level-dependent permissions to the user in order to access ABox facts in the domain knowledge.
- Introduce a mediator TBox bridge rules for semantic mappings which express how to match heterogeneous semantic repositories by means of the mappings.

The mediator semantic authorization-security model has several unique features, that are as follows: it performs semantic TBox access control based on RBAC model in which mandates access to the TBox objects. Requests for users are always directly checked against the access control rules of the local semantic database and not sent or allowed access via the mediator. It also has a middleware-based architecture which performs mapping between semantic ontology resources; and the ABox trust management level which is utilized for restricting access to the ABox individuals at a more fine granular level.

The rest of the paper is organized as follows: Section 2 summarizes the components of semantic DL knowledge bases. Section 3 examines and discusses the strengths and weaknesses of several current semantic data access control mechanisms. Section 4 presents the proposed model: the mediator authorization-security Model. It describes the main components used to secure access across semantically heterogeneous repositories. Section 4.1 shows the design and architecture of semantic payload caching which help to achieve performance improvement in the mediator model. Section 5 shows a motivation example. Section 6 describes the test setting for the mediator model.

## 2. Semantic knowledge representation

The Semantic Web community implicitly adopted description logics (DL) as a core technology for the ontology layer. One of the reasons behind this is that this logic has been heavily analyzed in order to understand how constructors interact and combine to affect tractable reasoning [1]. Thus, description logics are useful and efficient for knowledge representation, and reasoning about structured knowledge, fitting into the structural provision of RDF, RDF schema and OWL technologies. Typically, a DL knowledge base is comprised of two parts.

- The terminological part that describes conceptualization, i.e. a set of concepts and properties for these concepts, and captures the concept hierarchies (i.e. relations between concepts).
- The assertional part that captures the facts in an application domain [2,3].

The semantics of DL is defined by interpretations. Every interpretation is a pair $(\mathcal{D}, \mathcal{I})$, where $\mathcal{D}$ is a non-empty set of individual objects, called the domain, and $\mathcal{I}$ is an interpretation function that maps:

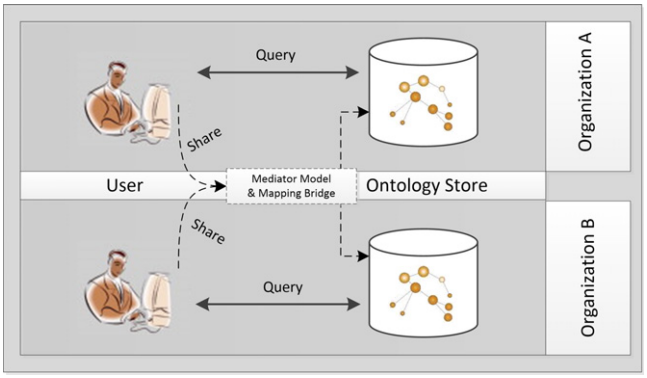


**Fig. 1.** Interoperation semantic system.

- every concept (class) C is a subset of $\mathcal{D}$,
- every property P is a binary relation on $\mathcal{D} \times \mathcal{D}$,
- every individual (instance) $x$ is an element $x \in \mathcal{D}$.

For example, in OWL, concepts are explained as sets of objects that represent the individuals in the domain of discourse. A property, that is a binary relation connecting concepts, can be further distinguished as an object property or data property. An object property represents the relation between individuals (instances) of two concepts, whereas a data property represents the relation between an individual concept and the literal value. In this paper, classes are described as concepts, while instances of a class are described as individuals. For example, as shown in Fig. 2, we have concepts (such as Organization, Country, Person), properties (such as HasBranchOffice, HasHeadOffice, HasEmployee), and individuals (such as WS, Smith, Bob, India, USA). In the following, we clarify the components of DL knowledge bases.

**Definition 1.** A DL knowledge base $(\mathcal{KB})$ is an ordered pair $(\mathcal{T}, \mathcal{A})$ where

- $\mathcal{T}$ is a set of `terminological` axioms (the TBox).
- $\mathcal{A}$ is a set of `assertional` axioms (the ABox).

**Definition 2** (*TBox* $\mathcal{T}$)**.** Terminological axioms represented by a collection of definitions of concepts, and their properties (relations), and may contain inclusion relations between concepts (concept hierarchy).

**Definition 3** (*Concept Hierarchy*)**.** Let $\mathcal{H} = \{H_1, \ldots, H_n\}$ be a finite set of concept hierarchies. Each concept hierarchy $H_i$ is a set of concepts partially ordered by the subsumption relation: $H_i = (C_i, \sqsubseteq)$ where $C_i$ is a subset of $\mathcal{C}$, the set of all concepts, and $\sqsubseteq$ is the subsumption relation which is used to create a hierarchy of concepts.

For example, Fig. 3, $(C_{i_2} \sqsubseteq C_{i_1})$ means that $C_{i_1}$ subsumes a concept $C_{i_2}$ in the concept hierarchy.

**Definition 4** (*ABox* $\mathcal{A}$*: Assertional Knowledge Base in an Ontology*)**.** Given a set $\mathcal{C}$ of concepts and a set $\mathcal{P}$ of relations in an ontology, and a set $\mathcal{O}$ of individual objects, the assertional knowledge base in the ontology is represented by:

- $C(x)$ where $C \in \mathcal{C}$ and $x \in \mathcal{O}$.
- $P(x; y)$ where $P \in \mathcal{P}$ and $x; y \in \mathcal{O}$.