
Amit Jain

PhD Dissertation
Center for Information Assurance Engineering
Computer Science and Engineering
University of South Carolina
2008
Abstract

This PhD dissertation proposes a Semantic-Aware Authorization Framework to address some of the problems encountered in sharing semi-structured data securely. The proposed framework ensures the preservation of authorization permissions on data even if the structure of the data is changed during the data exchange. My framework supports data sharing in distributed, and heterogeneous environments by providing syntax independent authorization capabilities for eXtensible Markup Language (XML), the most widely used standard for data format and exchange. Most of the security standards available for XML data security, use the syntax and structure of the XML data to provide different security services, such as access control, encryption and define security policies. In this research work I propose an approach to remove the dependence of security on data syntax and make use of data and application semantics to secure XML documents. In particular, I have developed an Access Control framework for XML that expresses authorization requirements on data semantics. I define a XML to ontology mappings to associate XML data documents with their semantics. First, I show how to map an XML data document to a corresponding ER model. Security requirements are expressed onto ER conceptual models. I use these mappings to derive the security policies for XML document from authorizations expressed on the ER model. This architecture has limitations in the wake of current distributed nature of the web and enterprise application scenarios. Present enterprise applications
and knowledge management systems aim to use ontologies for application integration and for defining domain semantics. Using ontologies enables several advantages over ER model usage such as easier data model integration, relationship modeling, extensibility, and open sharing. So using ontologies for representing the semantics of XML data is more suited for the web environment. But this entails the need for an access control model to secure metadata, such as RDF, against unauthorized access. Hence using semantics for securing XML data presents two requirements, 1) Developing authorization framework for metadata represented in RDF format, 2) Establishing mappings between the XML data and its semantics, presented by RDF ontology, to propagate the RDF authorizations to XML data. Since RDF is the standard language for representing the domain knowledge, by using RDF, the proposed authorization model can ensure wider adaptability for securing large amounts of RDF data and information available on the web. The proposed semantic security architecture, in addition to providing a uniform access control model for XML data, also provides an authorization model for RDF ontological data.

RDF security policies are expressed in the form of RDF policy patterns which map to groups of RDF data. Security labels are generated for mapped RDF triples from the patterns to create a secure materialized view of RDF data. The labels are also consistent and complete such that each RDF data item is assigned one and only one security label. XML to RDF mappings are formally defined between an XML data tree and an RDF ontology. Utilizing these mappings, the RDF authorizations then can be used to generate access control permissions for the mapped XML documents. The resulting XML security policy provides a security label for each of the XML node in the data tree. The security labels are conflict free such that each XML node gets assigned only one security label. The generated security labels for XML data are guaranteed to satisfy the security needs expressed on the meta-data. Also the security labels are complete such that a security label is derived for every XML data
item. This method provides syntax independent authorizations for messages used to exchange and store information by the distributed Web applications in Web Services and Semantic Web settings, increasing the security and interoperability. This method also removes the inconsistent authorizations on structurally different but semantically similar data by mapping them to the same domain ontology and hence deriving its security policies. A Normal Form for XML data representing structural properties of XML semantics, called XML Semantic Normal Form is also proposed.

Formal properties of the proposed model such as completeness, consistency, and default policy for both RDF and mapped XML data are also developed. I have also developed a prototype to implement this authorization framework.
# Table of Contents

Copyright .......................................................... ii

Acknowledgments .................................................. iii

Abstract .................................................................. v

List of Figures ........................................................ x

1 Introduction ......................................................... 1
   1.1 Problem Statement and Dissertation ....................... 3
   1.2 Research Contributions ....................................... 6
   1.3 Dissertation Organization ................................. 7

2 Research Problem Overview ..................................... 9
   2.1 XML Evolution as Data Interchange Standard ............ 9
   2.2 Problem of XML Security Reliance on Syntax ............ 10
   2.3 XML-RDF Mapping Problem .................................. 13
   2.4 RDF Ontology Security Problem ......................... 14
      2.4.1 RDF Inferencing ........................................ 16
   2.5 XML Semantic Normal Form .............................. 16

3 Related Work ...................................................... 18
   3.1 Semantic Annotation ......................................... 18
3.2 XML Security .................................................. 19
3.3 Semantic Web and Security ................................. 22
  3.3.1 Enterprise Solution Space ........................... 23
3.4 Semantic Integration ...................................... 25
3.5 Web Services, Semantics and Security .................. 26
3.6 Database Inferences ....................................... 28

4 System Architecture ........................................ 30
  4.1 Semantically Enhanced XML Data ..................... 33
  4.2 Specification of an Access Control Model for RDF Ontological Data . 34
  4.3 Derivation of XML Access Control Policies from RDF Authorizations 35

5 Semantically Enhanced XML Data .......................... 36
  5.1 Definitions of XML and ER Conceptual Model ........ 38
    5.1.1 Mapping Properties .................................. 43
  5.2 XML Semantic Normal Form ............................... 47

6 Specification of an Access Control Model for RDF ........ 48
  6.1 RDF Access Control Framework ......................... 50
    6.1.1 RDF Database and Security Cover ................ 52
    6.1.2 RDF Security Policies Database ................ 53
    6.1.3 RDF Querying Engine and Security Monitor ........ 53
  6.2 RDF Authorization Model ................................. 54
    6.2.1 RDF Data Model .................................... 55
    6.2.2 RDF Security Policy and Protection Objects .... 57
    6.2.3 RDF Reification .................................... 69
    6.2.4 RDF Entailment ................................... 71
    6.2.5 Conflict Resolution ............................... 80
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>XML to RDF Mappings</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>7.1 Advantages of RDF Ontologies over ER Conceptual Models</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>7.2 Associations between XML and RDF Data</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>7.2.1 XML to RDF Mapping Properties:</td>
<td>89</td>
</tr>
<tr>
<td>8</td>
<td>Derivation of XML Access Control Policies</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>8.1 Derivation of XML Policies Using ER Security Policies</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>8.1.1 Multilevel Security for Relational Database Schema</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>8.2 Derivation of Security Policies from RDF Ontologies to XML Data</td>
<td>97</td>
</tr>
<tr>
<td>9</td>
<td>Prototype</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>9.1 Tools and Technologies</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>9.2 User Interface</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>9.3 Pattern Mapping and Entailment Algorithm Implementation</td>
<td>103</td>
</tr>
<tr>
<td>10</td>
<td>Conclusion and Future Direction</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>111</td>
</tr>
</tbody>
</table>