# Chapter 13
# A Formal Approach for Securing XML Document

**Yun Bai**
*University of Western Sydney, Australia*

## ABSTRACT

*With the ever increasing demand for the Web-based applications over the Internet, the related security issue has become a great concern. Web document security has been studied by many researchers and various security mechanisms have been proposed. The aim of this paper is to investigate the security issue of the XML documents. We discuss a protection mechanism and investigate a formal approach to ensure the security of Web-based XML documents. Our approach starts by introducing a high level language to specify an XML document and its protection authorizations. We also discuss and investigate the syntax and semantics of the language. The flexible and powerful access control specification can effectively protect the documents from unauthorized attempts.*

## INTRODUCTION

The extensible markup language (XML) is used over the Internet for information exchange. With the increasing demand for the Web-based applications over the Internet by government, financial institution, business and trading, secure Web documents is becoming an essential issue. Web based document security has become a great concern, and has been studied by many researchers. Various security mechanisms have been proposed and investigated since then.

HTML was the initial language used for Web-based information processing. However HTML does not provide a clear structure and semantics, and often its design is just limited for a specific browser. XML was proposed by the World Wide Web Consortium (W3C) to overcome these limitations. XML improves HTML by providing a clear semantics without losing the initial HTML functions and capabilities. With XML, different applications can define and declare their own tags and attributes freely. XML is now widely accepted as a universal language for Web-based information exchange and processing.

DOI: 10.4018/978-1-4666-1580-9.ch013

Since XML is becoming the favourable format for information exchange over the Internet, XML document security has been increasingly studied and has become an active research area. Fine grained access control for XML documents has been investigated by some researchers. Murata et al. (2003) proposed a static analysis for XML access control. Given an access control policy and an access query, they use a static analysis to decide if to grant or deny such an access request. In this way, run-time evaluation is only needed when the static analysis is unable to make such decision. This pre-execution analysis improves the performance of the system response to a query. Damiani et al. (2002) presented a language for specification of access control by exploiting the characteristics of XML to define and enforce access control directly on the structure and content of the document. They provide a flexible security mechanism for protecting XML documents. An authentication approach for XML documents is proposed in (Devanbu, Gertz, Kwong, Martel, Nuckolls, & Stubblebine 2001). The proposal uses signature techniques to ensure the authenticity of the XML documents by having a server processing queries and certifying answers using a digital signature with an on-line private key. This approach allows un-trusted servers to answer certain type of path queries over the Internet without the need for a trusted on-line signing key. It provides the security of XML documents over the Internet by using a signature based document authentication.

Access control or authorization specifications have long been an important issue in computer system security. A variety of authorization specification approaches such as access matrix (Dacier & Deswarte 1994; Denning 1976); role-based access control (Crampton, & Khambhammettu, 2008); access control in database systems (Bertino, Jajodia, & Samarati 1996; Fernandez, Gudes, & Song 1989; Meadows, 1991), authorization delegation (Murray & Grove, 2008); procedural and logical specifications (Bai & Varadharajan, 1997; Bertino, Buccafurri, Ferrari, & Rullo, 2000) have been investigated. Some of the works emphasize on the specification of the access control policies and their functions; others on the access control for specific application areas and their delegations in the mechanism.

Since logic based specification has a clear and precise semantics and powerful expressiveness as stated in (Fagin, Halpern, Moses, & Vardi, 1995), a variety of logic authorization specification approaches have been proposed. A logic language (Jajodia, Samarati, Sapino, & Subrahmanian, 2001) has been proposed for expressing authorizations. They used predicates and rules to specify the authorizations; their work mainly emphasizes the representation and evaluation of authorizations. A formal approach using default logic to represent and evaluate authorizations has also been reported in (Woo & Lam, 1992). However, the constraints of the access control of the system are not considered in their work. Hence it is not clear how to judge whether a policy base is legitimate or not with respect to the system restriction. This approach is not suitable for XML documents security specification since it is hard to capture the hierarchical structure and the constraints of the documents.

A general framework (Bertino, Catania, Ferrari, & Perlasca 2003) on a logic formalism was proposed to model discretionary, mandatory access control and role-based access control models. The syntax and the semantics of the framework are given and also some example applications are presented. This work is mainly used for the analysis and the comparison of some existing access control models and their decidability. Whether this general framework can be used to model access control in XML documents scenario is not clear. The proposed rule-based security policy framework (Bettini, Jajodia, Wang, & Wijesekera, 2002) includes provisions and obligations. It investigated a reasoning mechanism within this framework in a general database scenario. This work investigates authorization policy with a logic framework from management point of view.