UNIVERSITI PUTRA MALAYSIA

TWO LEVEL SECURITY APPROACHES FOR SECUREXML DATABASE CENTRIC WEB SERVICES AGAINST XPATH INJECTIONS

AZIAH ASMAWI

FSKTM 2016 34
TWO LEVEL SECURITY APPROACHES FOR SECURE XML DATABASE CENTRIC WEB SERVICES AGAINST XPATH INJECTIONS

By

AZIAH ASMAWI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia in fulfilment of the requirements for the Degree of Doctor of Philosophy

May 2016
All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia
DEDICATIONS

I dedicate this thesis to my precious family,

My husband, Kamil Ikhmal Kamarudin
My mother, Sharifah Abdullah
and
My adorable kids, Aqil Iman, Aliya Ariana and Ammar Darwisy

whose love, supports and the prayers has helped me complete this thesis.
Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

TWO LEVEL SECURITY APPROACHES FOR SECURE XML DATABASE CENTRIC WEB SERVICES AGAINST XPATH INJECTIONS

By

AZIAH ASMAWI

May 2016

Chairman : Associate Professor Lilly Suriani Affendey, PhD
Faculty : Computer Science and Information Technology

Web services are deployed using eXtensible Markup Language (XML), which is an independent language for easy transportation and storage. As an important transportation for data, Web services has become increasingly vulnerable to malicious attacks that could affect essential properties of information systems such as confidentiality, integrity, or availability. Like any other application that allows outside user submission data, Web services can be susceptible to code injection attacks, specifically XPath (XML Path Language) injection attacks. This kind of attack can cause serious damage to the database at the backend of Web services as well as the data within it. To cope with this attack, it is necessary to develop effective and efficient secure mechanism from various angles, outsider and insider. This thesis addresses both outsider and insider threats with respect to XPath injections in providing secure mechanism for XML database-centric Web services which yields the following significant contributions.

We propose the two level security approaches for the ultimate solution within XML database-centric Web services. The first approach focuses on preventing malicious XPath input within Web services application. In order to address issues of XPath injections, we propose a model-based validation (XIPS) for XPath injection attack prevention in Web service applications. The second approach focuses on preventing insider threat within XML database. In order to deal with insider threat, we propose a severity-aware trust-based access control model (XTrust) for malicious XPath code in XML database. A prototype of the solution and each approach was designed, implemented and evaluated using synthetic data through experimental research approach to evaluate its security performance. Evidently, result analysis proved that the two level security approaches solution able to provide overall protection for XML database centric Web services environment from outsider and insider threats with respect to XPath injections. Meanwhile, the first approach, XIPS provides alternative solution for Web service applications against malicious XPath input compared to the previous work and the second approach, XTrust provide more secure access control for XML database against malicious XPath code compared to the previous work. As a conclusion, the two
level security approaches solution improved security level in XML database-centric Web services.
Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Doktor Falsafah

PENDEKATAN DUA TAHAP KESELAMATAN BAGI KESELAMATAN LAMAN SESAWANG PERKHIDMATAN BERTERASKAN PANGKALAN DATA XML TERHADAP SUINTIKAN XPATH

Oleh

AZIAH ASMAWI

Mei 2016

Pengerusi : Prof. Madya Lilly Suriani Affendey, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Laman sesawang perkhidmatan dibangunkan menggunakan eXtensible Markup Language (XML) untuk memudahkan pengangkutan dan penyimpanan. Sebagai pengangkutan yang penting untuk data, laman sesawang perkhidmatan telah menjadi semakin terancam dan terdedah kepada serangan jahat yang boleh memberi kesan kepada ciri khas penting di dalam sistem maklumat seperti kadar kerahsiaan, integriti dan kadar ketersediaan. Laman sesawang perkhidmatan juga seperti aplikasi lain yang membenarkan pengguna luar untuk menghantar data, boleh terdedah kepada serangan suntikan kod, khususnya serangan suntikan XPATH (Bahasa Path XML). Serangan seperti ini boleh menyebabkan kerosakan yang serius kepada pangkalan data di sebalik laman sesawang perkhidmatan dan juga data di dalamnya. Untuk mengatasi serangan ini, adalah perlu untuk membangunkan mekanisma keselamatan yang efektif dan efisien dari segenap sudut, luaran dan dalaman. Tesis ini mengalamatkan kedua-dua ancaman luaran dan dalaman yang berkaitan dengan suntikan XPATH di dalam menyediakan mekanisma keselamatan untuk laman sesawang perkhidmatan berasaskan pangkalan data XML.

Kami menyarankan penyelesaian menggunakan pendekatan dua tahap keselamatan untuk penyelesaian yang muktamad di dalam laman sesawang perkhidmatan berasaskan pangkalan data XML. Pendekatan pertama menumpukan kepada pencegahan masukan data XPATH berniat jahat di dalam aplikasi laman sesawang perkhidmatan. Bagi menangani isu ini, kami mencadangkan validasi berasaskan model (XIPS) untuk mencegah serangan suntikan XPATH dalam aplikasi laman sesawang perkhidmatan.

Pendekatan kedua menumpukan kepada pencegahan serangan salahguna pengguna dalaman di dalam pangkalan data XML. Bagi menangani isu ini, kami mencadangkan model kawalan capaian berasaskan kepercayaan kesedaran-keterukan untuk serangan kod XPATH berniat jahat di dalam prosedur simpanan pangkalan data XML. Prototaip untuk cadangan penyelesaian dan setiap pendekatan telah direkabentuk, dilaksanakan dan
dinilai menggunakan data sintetik menerusi pendekatan penyelidikan bereksperimen untuk menilai pencapaian keselamatan setiap cadangan. Ternyata, analisis keputusan membuktikan penyelesaian menggunakan pendekatan dua tahap keselamatan berupaya menyediakan perlindungan kepada laman sesawang perkhidmatan berasaskan pangkalan data XML secara keseluruhannya daripada ancaman luaran dan dalaman yang berkaitan dengan suntikan XPath. Sementara itu, pendekatan pertama, XIPS menyediakan penyelesaian alternatif untuk aplikasi laman sesawang perkhidmatan terhadap masukan data XPath beraniat jahat berbanding penyelidikan terdahulu dan XTrust juga menyediakan kawalan capaian yang lebih selamat ke dalam pangkalan data XML daripada kod XPath beraniat jahat berbanding dengan penyelidikan yang terdahulu.

Sebagai kesimpulan, penyelesaian menggunakan pendekatan dua tahap keselamatan meningkatkan tahap keselamatan di dalam laman sesawang perkhidmatan berteraskan pangkalan data XML.
ACKNOWLEDGEMENTS

Praise to The God, The Most Beneficent, and The Most Merciful, who has given me the opportunity to complete this thesis.

I would like to express my gratitude to my supervisor, Assoc. Prof. Dr. Lilly Suriani Affendey who has supported and guided with me throughout the research years. Also, I thank to Assoc. Prof. Dr. Nur Izura Udzir and Prof. Dr. Ramlan Mahmod for their contributions during my research work. As members of the committees, they have helped to improve the quality of this work with their constructive comments during the progress of the work.

I would also like to thank the Ministry of Higher Education Malaysia for the scholarship, Universiti Putra Malaysia and Faculty of Computer Science and Information Technology for the study leave and allowances they granted me during my studies.

Finally, heartiest thanks to all family members for their encouragement, support and prayers. Without them, it would hard to get through various stages of this life. Last but not least, very special thanks to my husband, Kamil Ikmal for his endless love, patience, sacrifices throughout these years.
I certify that a Thesis Examination Committee has met on 25 May 2016 to conduct the final examination of Aziah binti Asmawi on her thesis entitled "Two Level Security Approaches for Secure XML Database Centric Web Services Against XPath Injections" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Norwati Mustapha, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Hamidah binti Ibrahim, PhD
Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Ali bin Mamat, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Jemal Abawajy, PhD
Professor
Deakin University
Australia
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 27 December 2016
This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Lilly Suriani Affendey, PhD**  
Associate Professor  
Faculty of Computer Science and Information Technology  
Universiti Putra Malaysia  
(Chairman)

**Nur Izura Udzir, PhD**  
Associate Professor  
Faculty of Computer Science and Information Technology  
Universiti Putra Malaysia  
(Member)

**Ramlan Mahmod, PhD**  
Professor  
Faculty of Computer Science and Information Technology  
Universiti Putra Malaysia  
(Member)

---

**ROBIAH BINTI YUNUS, PhD**  
Professor and Dean  
School of Graduate Studies  
Universiti Putra Malaysia  

Date:
Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ____________________________ Date: ________________

Name and Matric No.: Aziah Asmawi, GS29708
Declaration by Members of Supervisory Committee

This is to confirm that:
• the research conducted and the writing of this thesis was under our supervision;
• supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: 
Name of Chairman of Supervisory Committee: 

Lilly Suriani Affendey, PhD

Signature: 
Name of Member of Supervisory Committee: 

Nur Izura Udzir, PhD

Signature: 
Name of Member of Supervisory Committee: 

Ramlan Mahmod, PhD
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>APPROVAL</td>
<td>vi</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xvi</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1 INTRODUCTION

1.1 Research Background  
   1.1.1 XPath Injection Issues in Web Service Applications  
   1.1.2 Insider Misuse Issues in XML Database  
1.2 Problem Statement  
1.3 Research Aim  
1.4 Research Objectives  
1.5 Research Scope  
1.6 Research Significance  
1.7 Thesis Structure  

### 2 BACKGROUND AND RELATED WORKS

2.1 Introduction  
2.2 Discussions on Database Security Research  
2.3 XML Databases-Centric Web Services Security Attacks  
   2.3.1 XPath Injection in Web Services  
   2.3.1.1 XPath Language  
   2.3.1.2 XPath Injections  
   2.3.1.3 XPath Injection Attack Categories  
      2.3.1.3.1 Simple XPath Injection  
      2.3.1.3.2 Blind XPath Injection  
   2.3.1.4 XPath Injection vs SQL Injection  
   2.3.1.5 Consequences of XPath Injections  
   2.3.1.6 Existing Works on XPath Injection Attacks  
   2.3.1.7 Schema-based Validation  
      2.3.1.7.1 XQuery Interception  
      2.3.1.7.2 XQuery Analyzer  
      2.3.1.7.3 XQuery Validation  
   2.3.1.8 Static Analysis Approach  
   2.3.1.9 Static and Dynamic Analysis Approach  
   2.3.1.10 Proposed Approach: Model-based Validation
2.3.2 Insider Misuse in XML Database

2.3.2.1 Insider Misuse

2.3.2.2 Existing Works on Insider Misuse Attacks

2.3.2.3 Access Control

2.3.2.4 Trust-based Access Control Approach

2.3.2.4.1 Trust Module

2.3.2.4.1.1 Operation Evaluator

2.3.2.4.1.2 Error Detector

2.3.2.4.1.3 Operation Recorder

2.3.2.4.1.4 Trust Calculator

2.3.2.4.2 Access Control Module

2.3.2.4.2.1 Access Manager

2.3.2.4.2.2 Access Decision Maker

2.3.2.5 Proposed Approach: Severity-Aware Trust-based Access Control

2.4 Proposed Solution: Two Level Security Approach

2.5 Summary

3 RESEARCH METHODOLOGY

3.1 Introduction

3.2 Experimental Research Approach

3.3 Experimental Research Approach Stages

3.3.1 Research Problem Identification

3.3.2 Experimental Research Planning

3.3.2.1 Two Level Security Approaches Solution Experimental Design

3.3.2.1.1 Subject

3.3.2.1.2 Performance Metrics

3.3.2.1.3 Input Generation

3.3.2.2 XIPS Experimental Design

3.3.2.3 XTrust Experimental Design

3.3.3 Conducting the Experiment

3.3.3.1 Two Level Security Approaches Solution Experiment

3.3.3.2 XIPS Experiment

3.3.3.3 XTrust Experiment

3.3.4 Data Analysis, Evaluation and Discussion

3.3.5 Report Writing

3.4 Summary

4 MODEL-BASED VALIDATION AND SEVERITY AWARE TRUST BASED ACCESS CONTROL FOR PREVENTING XPATH INJECTIONS

4.1 Introduction

4.2 Two Level Security Approaches

4.2.1 System Architecture

4.3 XIPS

4.3.1 XIPS System Architecture

4.3.1.1 XPath Query Interceptor

4.3.1.2 XPath Model Builder

4.3.1.3 XPath Model Checke

4.3.1.4 XPath Query Validation
5 EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Introduction

5.2 Two Level Security Solution Evaluation and Experimental Result
   5.2.1 Result Analysis Experiment 3.1
   5.2.2 Result Analysis Experiment 3.2
   5.2.3 Discussion

5.3 XIPS Evaluation and Experimental Result
   5.3.1 Result Analysis of Experiment 3.3
   5.3.2 Result Analysis of Experiment 3.4
   5.3.3 Discussion

5.4 XTrust Evaluation and Experimental Result
   5.4.1 Result Analysis of Experiment 3.5
   5.4.2 Result Analysis of Experiment 3.6
   5.4.2 Discussion

5.5 Summary

6 CONCLUSION

6.1 Introduction

6.2 Research Contributions

6.3 Future Work

REFERENCES

APPENDIX

BIODATA OF STUDENT

LIST OF PUBLICATIONS
LIST OF TABLES

Table | Page
--- | ---
2.1 Existing work in XPath injection Attacks | 18
2.2 Existing Works on Insider Misuse Attacks | 27
2.3 Bad Transaction Factor value based on BTNum (Farooqi & North, 2011) | 33
2.4 Table 2.4: Error Factor value based on ENum (Farooqi & North, 2011) | 33
3.1 Sample XPath injection | 43
3.2 VALID and ATTACK test calls for subject | 44
4.1 Severity value based on Bad Transaction Factor | 60
4.2 TV Equations | 61
5.1 Outsider and insider threats distribution | 63
5.2 Effectiveness comparison between two level security approaches and schema-based validation | 64
5.3 Efficiency comparison | 65
5.4 XPath injection distribution | 67
5.5 Effectiveness comparison between model-based validation and schema-based validation | 68
5.6 Efficiency comparison | 69
5.7 Define access permission result based on severity | 71
5.8 Define access permission result based on BTNum | 71
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Categories of XPath injection attack</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>XML document store user information</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>XPath query to read username and password</td>
<td>12</td>
</tr>
<tr>
<td>2.4</td>
<td>XPath query with valid input</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>XPath query with injection input</td>
<td>12</td>
</tr>
<tr>
<td>2.6</td>
<td>XPath query which can count the number of nodes</td>
<td>13</td>
</tr>
<tr>
<td>2.7</td>
<td>XPath query which can be used to guess the number of character in node</td>
<td>13</td>
</tr>
<tr>
<td>2.8</td>
<td>XPath query which can be used to guess the exact character in node</td>
<td>13</td>
</tr>
<tr>
<td>2.9</td>
<td>XPath query which accept username and password input</td>
<td>14</td>
</tr>
<tr>
<td>2.10</td>
<td>XPath query with injection input</td>
<td>14</td>
</tr>
<tr>
<td>2.11</td>
<td>XPath Injection Vulnerability Detector Architecture</td>
<td>19</td>
</tr>
<tr>
<td>2.12</td>
<td>User provide input for username and password</td>
<td>20</td>
</tr>
<tr>
<td>2.13</td>
<td>Sample XML file for the above query</td>
<td>20</td>
</tr>
<tr>
<td>2.14</td>
<td>Partial XML schema definition</td>
<td>21</td>
</tr>
<tr>
<td>2.15</td>
<td>Sample log file generated for different attack inputs</td>
<td>22</td>
</tr>
<tr>
<td>2.16</td>
<td>Trust-based access control’s process</td>
<td>29</td>
</tr>
<tr>
<td>2.17</td>
<td>Operation Policy file (Farooqi &amp; North, 2011)</td>
<td>30</td>
</tr>
<tr>
<td>2.18</td>
<td>Error Policy file</td>
<td>31</td>
</tr>
<tr>
<td>2.19</td>
<td>XLog file</td>
<td>32</td>
</tr>
<tr>
<td>2.20</td>
<td>The User’s Access Permission Policy file</td>
<td>34</td>
</tr>
<tr>
<td>2.21</td>
<td>The XML database’s Access Permission Policy file</td>
<td>35</td>
</tr>
<tr>
<td>2.22</td>
<td>Decision Process</td>
<td>36</td>
</tr>
<tr>
<td>2.23</td>
<td>Trust-based access control’s process (Farooqi &amp; North, 2011)</td>
<td>36</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>Research Methodology by Leedy and Ormrod (2001)</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>XIPS Architecture</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>System Flow</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>System Architecture</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Example servlet code</td>
<td>53</td>
</tr>
<tr>
<td>4.5</td>
<td>Sinkpoint query</td>
<td>53</td>
</tr>
<tr>
<td>4.6</td>
<td>Code which call to the XPath Model Checker</td>
<td>54</td>
</tr>
<tr>
<td>4.7</td>
<td>System Architecture</td>
<td>56</td>
</tr>
<tr>
<td>4.8</td>
<td>Severity Policy File</td>
<td>58</td>
</tr>
<tr>
<td>4.9</td>
<td>XLog file</td>
<td>59</td>
</tr>
<tr>
<td>4.10</td>
<td>Trust Policy file</td>
<td>61</td>
</tr>
<tr>
<td>5.1</td>
<td>Effectiveness analysis graph</td>
<td>65</td>
</tr>
<tr>
<td>5.2</td>
<td>Efficiency analysis graph</td>
<td>66</td>
</tr>
<tr>
<td>5.3</td>
<td>Effectiveness analysis graph</td>
<td>68</td>
</tr>
<tr>
<td>5.4</td>
<td>Efficiency analysis graph</td>
<td>70</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary of research contributions</td>
<td>75</td>
</tr>
</tbody>
</table>
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTF</td>
<td>Bad Transaction Factor</td>
</tr>
<tr>
<td>BTFW</td>
<td>Bad Transaction Factor Weight</td>
</tr>
<tr>
<td>BTNum</td>
<td>Bad Transaction Number</td>
</tr>
<tr>
<td>DAC</td>
<td>Discretionary Access Control</td>
</tr>
<tr>
<td>DOM</td>
<td>Data Object Model</td>
</tr>
<tr>
<td>DTD</td>
<td>Document Type Definition</td>
</tr>
<tr>
<td>EF</td>
<td>Error Factor</td>
</tr>
<tr>
<td>EFW</td>
<td>Error Factor Weight</td>
</tr>
<tr>
<td>ENum</td>
<td>Error Number</td>
</tr>
<tr>
<td>ETV</td>
<td>Existing Trust Value</td>
</tr>
<tr>
<td>ETVW</td>
<td>Existing Trust Value Weight</td>
</tr>
<tr>
<td>MAC</td>
<td>Mandatory Access Control</td>
</tr>
<tr>
<td>SDL</td>
<td>Secure Development Lifecycle</td>
</tr>
<tr>
<td>SDLC</td>
<td>Software Development Life Cycle</td>
</tr>
<tr>
<td>RBAC</td>
<td>Role Based Access Control</td>
</tr>
<tr>
<td>RDBMS</td>
<td>Relational Database Management System</td>
</tr>
<tr>
<td>SAX</td>
<td>Simple API for XML</td>
</tr>
<tr>
<td>SGML</td>
<td>Standard Generalised Mark-up Language</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>TBAC</td>
<td>Trust Based Access Control</td>
</tr>
<tr>
<td>TV</td>
<td>Trust Value</td>
</tr>
<tr>
<td>XCure</td>
<td>XPath Secure Coding</td>
</tr>
<tr>
<td>XIPS</td>
<td>XPath Injection Prevention System</td>
</tr>
<tr>
<td>XLog</td>
<td>XML Log File for Security Rather Than Recovery</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Mark-Up Language</td>
</tr>
<tr>
<td>XPath</td>
<td>Xml Path Language</td>
</tr>
<tr>
<td>XQuery</td>
<td>Xml Query Language</td>
</tr>
<tr>
<td>XTrust</td>
<td>Severity-aware Trust-based Access Control</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Research Background

Modern web applications today are database-driven. Most web applications support features such as login, registration, online payment, money transfer and billing address. In order to access these features, the client must submit personal and confidential information such as one’s name, username, bank account number, social security number, password, credit card number, and address, which are stored in the database of the application. Attacks on these kinds of application cost not only losing credentials, but also misuse of them.

Data is a very critical asset in any organizations. Over the last few decades, it has become an organization’s most precious asset and everything an organization does involves using data in some way or other. It is important that partners and customers have access to the data. For that purpose, the data cannot simply be hidden behind a firewall because the partner and customer need to access the data for business transactions or data sharing as well. Therefore, employing secure mechanisms to protect data in a database system from being exposed to outsiders and misused by unauthorized users is vital.

Korth and Silberschatz (1997) stated that the Web is, in effect, a large distributed database, though with a query language and access mechanism quite different from those traditionally included in a database system. Almost all computer system applications have the database at its back-ends as the main information sources. This information is considered as an organization’s most important asset, and thus needs to be protected from rival companies or malicious attacks. Web services that are connected to the Internet expose the databases that lie at the backend to attacks. One needs to secure the databases in order to protect the information stored in them.

In this section, the motivations behind this research are highlighted. The discussion starts by explaining the Web services’ security issues which motivate the real need to prevent XPath injection attacks in Web service applications. Furthermore, the discussion on XML database security issues motivates the need to improve access control in database systems to prevent against insider misuse attacks in XML database stored-procedures. Overall, the motivation for this research is to provide secure mechanisms to increase the level of security in XML database-centric Web services.
1.1.1 XPath Injection Issues in Web Service Applications

A report by McKinsey indicates Web services as one of the most important trends in modern software development (McKinsey, 2008). However, the wide use and exposure of Web services results in any existing security vulnerability being most probably uncovered and exploited by hackers. In fact, XPath injection is one of the most frequent types of attacks in the web environment (Stuttard, 2007).

A recent report by Torrid Networks (2015) explains that their application security team encountered a few XPath injections during a recent engagement to audit an application for a large telecom provider, Telecom X-Factor. XPath injection added some interest to their assignment since the application was using XML to store data and used XPath to query the data (Torrid Networks, 2015).

These attacks take advantage of improperly coded applications to change queries sent to a database, enabling, for instance, access to critical data. If a website uses an XML (eXtensible Markup Language) document to store data and user input is included in an XPath query against that document, the user may be vulnerable to an XPath injection. Vulnerabilities allowing XPath injection attacks are particularly relevant in Web services (Antunes et al. 2009), as their exposure is high and they frequently use a data persistence solution based either in a traditional relational database or in a XML database. Currently major database vendors and several open source efforts provide XML databases, and typically, access to these types of databases uses XPath expressions. Therefore, the goal of XPath injection is to maliciously explore any existing vulnerabilities in XPath expressions used by an application to access an XML database.

There are two versions of XPath queries: which are Version 1.0 (released in 1999) and Version 2.0 (released in 2010). The XPath query can be altered to achieve authentication bypass, business logic bypass and extraction of arbitrary data from the XML database (Siddharth, 2012). Upon noticing XPath injection vulnerability in an XPath-based application, an attacker does not need to fully understand the structure of the application. In fact, the attacker can generate a data query template that can be used for Blind XPath Injection within a few injection attack attempts (Klein, 2004).

Among the issues that arise from XPath injection, the most serious of these is the lack of secured protection in Web service applications against Blind XPath Injection attacks. Therefore, our research proposed a preventive mechanism that employed model-based validation to prevent Blind XPath input in Web service applications.

1.1.2 Insider Misuse Issues in XML Database

A large quantity of information is presented in XML format on the web for easy transportation. Due to the increased use of XML databases over the web, the need to secure these databases has increased. In a multiuser system, where the information is being shared across users who have different permissions, the need to implement a
security model which gives controlled access to the authorized users is vital. XML access control was introduced to suit this purpose. XML access control is a security mechanism which restricts the access of the XML data to authorized users only. Many access control models and enforcement mechanisms have been proposed to prevent the unauthorized disclosure of XML data.

Different models for XML database access control have been proposed and developed. Access control systems for XML databases can be categorized into three core approaches: discretionary access control (DAC), mandatory access control (MAC) and role-based access control (RBAC) (Zhu et al., 2007). Most traditional access control models protect data from malicious activities of outside users but cannot protect the data from insiders (Chagarlamudi et al., 2009). Research has suggested that damage caused by insiders is more harmful than that of outsiders (Park et al., 2006).

Commonly, when there are reports on cybercrime against computer systems, the first thing that comes across the readers’ mind is that a hacker or attacker is involved. It is less likely that they will think the crime involves employees or insiders in the organization. In reality, however, employees or insiders often cause the most significant and costly security incidents. Research also suggested that damage caused by insiders is more harmful than that of outsiders (Richardson, 2011). Indeed, the fact that insiders are already within the organization often puts them in an ideal position to misuse a system if they intend to do so. The greater an individual’s knowledge of an organization’s internal resources, the greater the potential threat from that person. In fact, insiders are not interested in damaging systems or applications, but focus on obtaining critical information and accessing the internal level of resources for their personal advantage and gain.

Insider misuse can threaten personal data, national security, and economic prosperity. The 2008 CSI Computer Crime and Security Survey ranks insider abuse second only to viruses in terms of attack types experienced by respondents (Richardson, 2008). Furthermore, 87.1 percent of respondents said that 20 percent or less of their losses should be attributed to malicious insiders (Richardson, 2011). The damage is difficult to quantify, because it can extend far beyond the actual cost of the items stolen or corrupted.

In this research, we focus on mitigating malicious XPath code in XML database-stored procedures. Stored procedures are a part of a database that allows the programmer to set an extra abstraction level in the database. By using stored procedure a user can store its own function according to its needs (Wei, 2006). In stored procedures, a collection of XPath queries are included. As stored procedures could be coded by programmers, so too is this one of the causes of XPath injection attacks. XPath injection is one type of insider threat when it occurs in the stored procedure database level. This type of insider threat is a huge topic in data security and many methods have been proposed to identify and prevent misuse.
1.2 Problem Statement

Many Web services applications are used to distribute information from organizations to different users over a network. Most often, these applications that accept interactions from users, and perform some accesses to databases (DBs), are based on assumptions about legitimate input and legitimate code that are used to build XPath queries. These Web services applications are possibly vulnerable to XPath injection attacks (Vieira et al., 2009), which rely on some weak validation of the textual input that is somehow used to compose XPath queries. Attacker can maliciously crafted input, that contains XPath instructions or fragments of these, produces queries whose semantics is different from the one meant by the designers and may threaten the security policies of the underlying databases. Also, regardless of input validation, insiders may introduce malicious code in an application that, when triggered by some specific input, for example would violate designer’s intention regarding security and accesses.

Researchers have started to contribute in the area of XPath injection and its possible liabilities (Blasco, 2007; Jinghua and Sven, 2008; Mitropoulos 2009; Antunes et al., 2009; Shanmughaeneethi et al. 2011; Karumanchi and Aquicciarini 2015; Thome et al. 2015). Above all existing works in XPath injection; they focus on the solution for outsider threats without considering the insider threats caused by XPath injection.

The problem of insider threats is one of the most challenging to the organizations and research community since a long time. It is well proved that the damage done by insiders is more severe than that of external attackers (Shatnawi et al., 2011). Database at the backend of every Web services is vulnerable to insider threat (Farooqi & North, 2011). For that reason, researchers start to give attention to this area (Magklaras et al., 2006; Kandias et al., 2010; Farooqi & North, 2011; Greitzer et al., 2012; Brdiczka et al., 2012; Eldardiry et al., 2013; Legg et al., 2015; Hashem et al., 2015). Based on the previous works on insider threat, none concentrates on such threat in XML database stored procedure. Furthermore, above all existing works in insider threat, none of them consider on the XPath malicious code as their research issue.

Mitigating both insider and outsider are very important issue but yet, there are lack of researches provide the solution (Naidu, S., 2013). Only several researchers discussed and tackled both insider and outsider threats in their work (Nayak & Rao, 2014; Lindvall, J., & Rueda, D., 2014; Merlo et al., 2006). The problems of the existing works are none of them focus on the solution for XPath injections. Therefore, this research will consider both insider and outsider threats with respect to XPath injections. We come out with an effective way for securing the XML database from these threats by proposing two level security approaches solution. In this solution, we will combine the model based validation prevention system and severity-aware trust-based access control in order to improve the effectiveness and the efficiency of the propose approach compared to previous work by Shanmughaeneethi et al., (2011) which use schema-based validation. The implementation of two level security approaches solution should overcome the XPath injections and hence, introducing the XML database-centric Web service for better security.
1.3 Research Aim

The aim of this research is to provide a secure mechanism for XML database-centric Web services. It must not only be capable of protecting against malicious XPath input in Web services application, but also manage to protect against malicious XPath code in XML database.

1.4 Research Objectives

In order to achieve the aim of this research, several research objectives have been identified as follows:

i) To propose two level security approaches solution which consist both model–based validation and severity-aware trust-based access control for XML database-centric Web services.

ii) To design system architecture for two levels security approaches solution which consist both model–based validation and severity-aware trust-based access control for XML database-centric Web services.

iii) To design system architecture for intrusion prevention system which employs model-based validation for Web services application.

iv) To design system architecture for access control which employs severity-aware trust-based access control for XML database stored-procedure.

1.5 Research Scope

This research was conducted within the research scope as described below:

i) This research focuses on security and performance.

ii) The two level security approaches solution would focus on protecting XML database-centric Web services from both outsider and insider threats with respect to XPath injections.

i) The intrusion prevention system would focus on preventing Web services from malicious XPath input.

ii) The severity-aware trust-based access control would focus on protecting XML database stored-procedures from malicious XPath code.

1.6 Research Significance

The purpose of this research is to provide a secure mechanism for XML database-centric Web services from outsider and insider threats with respect to XPath injections. The aim of the research can be achieved by providing two level security approaches solution for
XML database Web services using model-based validation and severity-aware trust-based access control against malicious XPath input and malicious XPath code.

1.7 Thesis Structure

This thesis presents the general issues in database security, the specific issues in both XPath injections upon database-centric Web services. This chapter has introduced the basic concepts in database security as a platform for the understanding of the research. The following is the outline of the thesis.

- Chapter 2: A detailed review regarding the study and work that relates to this research is given. The review covers topics related to the research issues focusing on XPath injections in Web services and XML database systems.
- Chapter 3: We discussed the research methodology employed in this research in this chapter. We elaborated each stages consists in this chapter, Research Problem Identification, Experimental Research Planning, Conducting the Experiment, Data Analysis, Evaluation and Discussion and Report Writing.
- Chapter 4: The implementation of the two level security approaches solution, XIPS, and XTrust is given in this chapter. This chapter presents the system architecture and description of each module.
- Chapter 5: Experimental results and discussion for each experiment are discussed in this chapter.
- Chapter 6: The discussion on research contributions and conclusion of this thesis are given in this chapter. Recommendations and suggestions for future improvements are covered.
REFERENCES


79


