XML安全技术分析与应用

顾韵华 傅德胜 王 兴
(南京信息工程大学计算机与软件学院 南京 210044)

摘 要 随着XML的广泛应用,XML的安全性也越来越受到关注。从XML加密、XML数字签名和XML密钥管理等方面分析了XML应用中的安全技术框架与规范。在此基础上通过高校学院信息化平台的设计,给出了综合应用这些技术确保XML安全的模型与实现方法,并指出了在设计中需注意XML解析存在的安全问题及对策。

关键词 XML应用,安全技术,XML加密,XML数字签名,框架与规范,对策
中图法分类号 TP309.2 **文献标识码** A

Analysis and Application of XML Security Technology

GU Yun-hua FU De-sheng WANG Xing
(School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract With the wide application of XML, more and more attention has been paid to the security of XML code. The secure technical frame and the specification in XML application were deeply analyzed according to the XML encryption, XML digital signature and XML cryptographic management. The college information platform was designed based on the research above. Meantime, the model and the implement methods of XML security were ensured by the integrative application of the technology. Furthermore, the security problems existed in the XML analysis and the countermeasures which more attention should be paid to in the designing process were provided out.

Keywords Application of XML, Security technology, XML encryption, XML digital signature, Framework and specification, Strategy

1 引言

可扩展标志语言XML(eXtensible Markup Language)是一种开放型的数据描述语言,具有高度结构性、显示与内容分离、可扩展性等优点,已成为Internet上信息表示和信息交换的标准[1]。现今,XML技术已广泛应用于电子商务、电子政务等诸多领域。然而,由于Internet是公开的网络,任何人都可能截获甚至篡改网络上的数据。因此XML安全是XML应用中一个尤为重要的问题。

与传统的信息安全服务一样,XML安全技术也主要保证XML数据的保密性、完整性、身份鉴别、不可否认性以及对XML的访问控制。通常用加密来保证数据机密性,用数字签名来对信息进行鉴别、保证数据完整性和不可否认。但由于电子商务等应用对XML数据的安全需求出现了新的特点,如:必须以XML形式表现被加密或签名的数据;可以从XML文档中任意选出一部分内容进行加密或签名。因此XML安全既与传统信息安全技术有很多相同之处(如均基于密码学),又有自己的鲜明特点。XML安全技术的基础是由W3C和IETF等机构制定的一系列XML安全规范[2](XML Security Specifications),包括XML加密、XML数字签名、XKMS密钥管理以及Web Service安全等,这些规范目前仍在不断完善。

本文先从XML加密、XML数字签名、XML密钥管理3个方面阐述XML安全技术,然后,详细分析了高校学院信息化平台设计中各项XML安全技术的应用,构建了一个安全XML应用,并指出了在设计中需注意XML解析存在的安全问题。

2 XML安全技术

2.1 XML加密

XML加密的基础是XML加密规范[3](XML Encryption Syntax and Processing),该规范是由W3C制定并于2002年9月公布推荐标准。XML加密规范是加密XML数据、以标准XML格式表示加密结果以及解密器处理过程的一套标准方法。XML加密允许加密任何数据,这些数据可以是一个完整的XML文档或一个XML文档中的指定元素,也可以是从外部引用的任意非XML格式数据。在进行XML加密时,采用标准的XML标记语法来表示相关信息、算法以及实际加密的数据。加密结果被表示为一个XML加密文档。这个XML加密文档可直接含有加密的数据,也可以间接地从外部引用加密的数据。XML加密组合使用了对称密码和非对称密码算法,通常对称密码算法用于XML数据的批量加密,非

到稿日期:2008-12-02 本文受江苏省产业技术研究与开发基金项目(苏发改高技发[2006]1106号)资助。
顾韵华(1965-),女,副教授,CCF高级会员,主要研究方向为信息安全,E-mail:yhgu@nuist.edu.cn;傅德胜(1950-),男,教授,博士生导师,主要研究方向为信息安全;王 兴(1982-),男,硕士研究生,主要研究方向为信息安全。

对称密码算法则用于安全地交换对称密钥。

XML 加密的主要特点是：(1) 加密状态持久、保证数据安全：XML 文档一经加密，在解密之前，不论是存储于磁盘空间中，还是在网络的传输过程中，或是在某个网络节点停留时，都处于加密状态，未经授权无法访问到密码信息，能确保数据安全性。(2) 加密粒度可选：不仅可以对整个文档进行加密，还可以对 XML 文档中元素和元素内容进行加密，使得基于 XML 的数据传输有着更加灵活的安全机制。(3) 可实现多方安全会话：由于 XML 加密在网络转发过程中无需加密解密操作，不会造成原始发信人的认证丢失，可实现多方的安全会话。(4) 满足各种应用环境的需求：满足各种应用环境的共性需求，XML 既可以应用于消息传输，也可以应用于文档数据的存储，并支持特定应用的特殊需求，具有扩展能力。

2.2 XML 数字签名

XML 数字签名的基础是 XML 数字签名规范[4]（Decryption Transform for XML Signature），该规范是由 IETF 和 W3C 联合制定并于 2002 年 2 月公布推荐标准。XML 签名规范是对现有数字签名技术的扩展，定义了一套用 XML 表示的数字签名的新方法。为了实现用 XML 来表示数字签名，在 XML 规范中定义了〈Signature〉元素。该元素包括了签名使用的类型、对已签名数据的引用以及验证签名所需的密钥的详细信息。XML 提供了灵活的数字签名机制，既支持对消息整体签名，也支持对 XML 文档或消息的部分进行签名。

2.3 XML 密钥管理

XML 公钥管理规范[5]（XML Key Management Specification，XKMS）是由 W3C 推荐的公钥配置与注册规范。XKMS 可认为是基于 XML 的 PKI，被称作第二代 PKI。XKMS 将传统 PKI 的两层应用模式转化为三层，在 PKI 用户与 PKI 提供者之间加入信任服务中间层。它利用 XML 语法描述密钥和证书信息，通过 XKMS 消息将客户端对密钥和证书的操作部分或全部地委托给基于 Web 的信任服务，从而向客户端屏蔽了底层 PKI 实现的复杂性。XKMS 由两种服务组成：XML 密钥信息服务规范（XML Key Information Service Specification，X-KISS）和 XML 公钥注册服务规范（XML Key Registration Service Specification，X-KRSS），前者包括公钥的定位服务和证实服务，后者包括密钥的注册、重新发行、撤销和恢复。

XML 加密、XML 数字签名和 XML 密钥管理是相对独立又密切相关的技术，XML 加密保证数据的机密性，XML 数字签名保证数据的完整性和不可否认性，PKI＋XKMS 能够为应用透明地提供加密和数字签名等服务所必需的密钥和证书管理功能，为网络信息安全提供数据机密性、完整性、身份认证和数据源不可否认等核心安全服务。

3 学院信息平台中 XML 安全技术的应用

学院信息化平台是一个开放式系统，其功能可扩展，与高校其他信息系统（OA 系统、学生管理系统、教务系统等）需进行大量数据交换。系统的基本功能包括教师管理、课程管理、学生管理、论文管理、公文管理、新闻管理等 12 项。系统运用 XML 相关技术，实现学院与学校以及与各学院之间异构系统和数据库之间的数据交换[6]。系统基于 ASP. NET 2.0＋SQL Server 2005 进行开发，采用了 XML 及 Web Service 来实现学校及各学院内部异构系统间的数据实时传输及信息共享。. NET Framework 2.0 框架对 XML 安全技术提供了良好的支持，其 System. Security. Cryptography 命名空间可实现 DES，3DES，AES，RSA 等加密算法，以及 MD5，SHA1，SHA256 等散列算法，而 System. Security. Cryptography. Xml 命名空间包含支持创建和验证 XML 数字签名的类[7]。

3.1 安全数据传输

为确保 XML 文档在网络中的传输安全，系统综合运用包括 XML 加密和签名以及密钥管理等技术，构建一套以数字信封为主流程的数据安全传输框架，如图 1 所示。数字信封的功能类似于普通信封，采用混合密码体制保证只有规定的接收人才能阅读信息的内容。

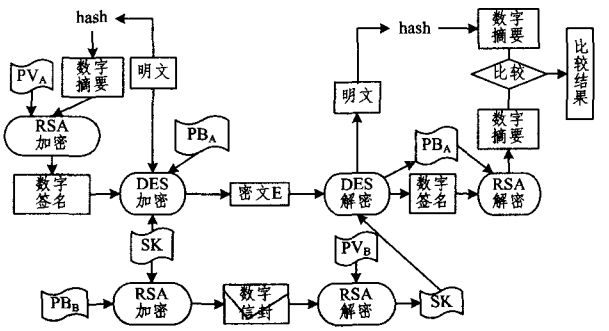

图 1 数据传输模型

设 A，B 分别为发方和收方，数据交换流程为：

(1) A 将明文进行 hash 运算，得数字摘要 MD；

(2) A 用私钥 PV_A，采用 RSA 算法对 MD 进行加密，即得数字签名 DS；

(3) A 用对称密钥 SK 对明文、SD 及 A 的公钥 PB_A 采用对称算法加密，得加密信息 E；

(4) A 用收方 B 的公钥 PB_B，采用 RSA 算法对对称密钥 SK 加密，形成数字信封 DE；

(5) A 将加密信息 E 和数字信封 DE 一起发送给 B；

(6) B 接收到数字信封后，用私钥 PV_B 解密数字信封，取出对称密钥 SK；

(7) B 用对称密钥解密加密信息，还原出明文、数字签名 SD 及 A 的公钥 PB_A；

(8) B 验证数字签名，用 A 的公钥 PB_A 解密 MD；同时将明文用同样的 hash 算法得哈希值 MD′；比较 MD 和 MD′，若两者相等，则接收该签名；否则，拒绝该签名。

3.2 XML 文档元素及密钥的加密/解密

XML 加密规范定义了〈EncryptedData〉元素以及〈EncryptionMethod〉，〈KeyInfo〉，〈CipherData〉，〈EncryptionProperties〉等子元素来具体描述 XML 文档的加密信息。发送者对 XML 文档元素加密时，需创建符合以上结构的〈EncryptedData〉元素；接收者根据从〈EncryptedData〉元素中得到的解密所需的加密算法、参数和密钥信息，进行解密。

本系统基于 XmlDocument 类设计了 XML 文档元素加密算法，以下是利用该加密算法对 XML 文档加密的实例。被加密文档如下：

〈? xml version＝"1.0" encoding＝"UTF-8" standalone＝"yes" ?〉

```
〈补考信息〉
  〈学号〉200526001〈/学号〉
  〈姓名〉Zhangshan〈/姓名〉
  〈补考科目〉计算机组成原理〈/补考科目〉
  〈补考成绩〉60〈/补考成绩〉
〈/补考信息〉
```

对其中补考成绩元素值“60”进行加密，结果如下：

```
〈EncryptedData Type = " http://www. w3. org/2001/04/xmlenc#Element" xmlns="http:// www. w3. org/2001/04/ xmlenc#"〉
  〈EncryptionMethod Algorithm = " http://www. w3. org/2001/04/ xmlenc#tripledes-cbc"/〉
  〈KeyInfo xmlns="http://www. w3. org/2000/09/xmldsig#"〉〈KeyName〉tDESKey〈/KeyName〉
  〈/KeyInfo〉
  〈CipherData〉
    〈CipherValue〉3g9Em9BhDQJ …
kSP2bVS0=〈/CipherValue〉
  〈/CipherData〉
〈/EncryptedData〉
```

基于 RSACryptoServiceProvider 类设计了产生公/私密钥对和对对称密钥的加解密算法，分别采用了 RSACryptoServiceProvider 类的 ToXMLString，Encrypt 和 Decrypt 方法。对对称加密密钥加密后的数据是 Byte 数组格式，可将其转换为 Base-64 格式而保存于 XML 文档中。XML 加密规范使用〈ds:KeyInfo〉元素存放交换密钥的信息，该元素在〈EncryptedData〉元素中是可选的，即 XML 加密规范并不一定要求将加密后的密钥信息放入〈EncryptedData〉元素内，密钥信息也可以使用其他方式交换。

3.3 XML 文档数字签名

基于 SignedXml 类设计了对 XML 文档签名和签名文档验证的算法，分别采用了 SignedXml 类的 ComputeSignature 和 CheckSignature 方法实现。对 3.2 节中所列实例加密后的 XML 文档的数字签名结果如下：

```
〈SignatureValue〉kxqQhCUDt5YQ5GBfJL1V
giVsChPc0ZiNZRmktQw7Nay0fKTsavGlHS9ISE7g
WDJ7AW/oUOslj+KZGF9pDSDtv0wMSkD3xRAY
WQG5vc95/Mi1Z0WTNQDu48T4/WYPiRMgZbzbE
hL2d5JRcljFIcj9z3Y9NlP1V7xfAmoxnYR0/BE=〈/SignatureValue〉
```

要注意的是在接收方对数字签名进行有效性验证时，在对 XML 文档进行摘要计算之前需对其规范化，这样才能确保逻辑上相等的 XML 文档具有同样的物理表示形式。

3.4 XKMS 的应用

本系统利用已有的 PKI 设施，设计了基于 XKMS 的密钥管理方案。XKMS 的应用包括两个方面：通过 X-KRSS 的注册服务进行公钥的注册，通过 X-KISS 的搜寻服务对公钥进行查找定位以及使用 X-KISS 的验证服务对公钥进行验证。

3.4.1 公钥注册

客户端通过 Web 服务发出一个注册请求，请求中包含需要注册的公钥信息，请求是以 SOAP 格式封装的。在向 XKMS 服务注册时需要提供客户端请求私钥的证明，注册模型如图 2 所示。

图 2 X-KRSS 注册服务

3.4.2 公钥验证

公钥注册成功后，如果需要对公钥信息进行查询，可以使用 X-KISS 的搜寻服务。搜寻服务的目的是查找在 XML 数字签名和加密中所使用的公钥信息，它能够解析 XML 数字签名标准的元素〈ds:KeyInfo〉，并且可以寻找到客户端所需要的信息，信息中可以包含公钥的值、公钥的名称、X.509 凭证，以及任何关于公钥属性信息。搜寻服务模型如图 3 所示。

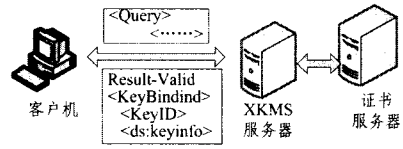

图 3 X-KISS 验证服务

3.5 XML 文档解析的安全性

在设计 XML 应用时，还需注意 XML 文档解析中存在的安全问题。在此简要分析一下基于 DOM(文档对象模型)的 XML 解析器的安全问题及防范措施。

3.5.1 DTD 攻击及防范

由于对于 XML 文档的有效性验证，可同时使用内部和外部 DTD。若一个元素或属性同时由外部和内部 DTD 定义，则优先采用内部 DTD 的定义[8]。正是由于内部 DTD 对外部 DTD 的覆盖带来了安全隐患。攻击者有可能编写出特定 DTD 的 XML，用于查询甚至修改他本没有权限访问的信息。此外，利用嵌入的 DTD 还能够编写出“合法”的超大数据量 XML 文档，将此文档上传到后台系统，就可能造成 DoS(拒绝服务攻击)。

为解决 DTD 攻击的安全隐患，应将 DTD 的验证过程置于处理 XML 文档的后台系统。由后台系统负责将接收到的 XML 文档进行解析，与公共的 DTD 进行比较验证。这样即使原 XML 文档中没有指定 DTD 或定义了非法的 DTD，后台系统也能作出准确有效的验证。另外对于此类 DoS 攻击，可通过对 HTTP 协议中内容长度(Content-Length)参数的设定，限制每一个 PUT Request 的最大文本长度。

3.5.2 XML 文档注释攻击及防范

解析器在创建 DOM 树形结构时一般都假设注释不可能出现在有意义的数据项中，而这一假设将可能被恶意利用。攻击者利用这一安全隐患，可以在不违反 XML 完整性等规则的情况下，恶意地破坏 XML 文档中的数据结构。特别对于 XML 加密及 XML 签名文档，由于这类文档都建立在一个良好的 XML 文档结构基础上，一旦遭到此类攻击，后台系统将无法解码被注释隔断的 XML 元素，应用程序也不能正确获得所需的数据。针对此类攻击的防御，应在编写用于处理 DOM 树形结构中数据的程序是保证能够从任一个元素结点转换子元素列表，并连接所有包含了 CDATA 结点数据项的数据值，并使用该数据值作为其父结点的元素值。

3.5.3 结点攻击

结点攻击的目标是处理 XML 文档的后台系统。通过上

(下转第 141 页)

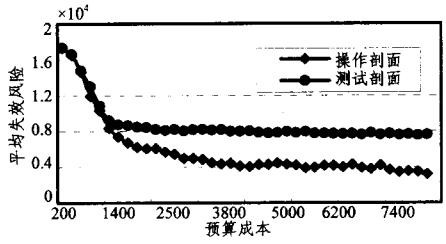

图 3　不同测试预算成本下平均失效风险比较

从图 3 可以看出，当测试预算成本 B 较小时，采用最优测试剖面和操作剖面使得平均失效风险降低幅度差不多，随着测试预算成本的增加，采用最优测试剖面的平均失效风险下降幅度要远远大于操作剖面，并且采用最优测试剖面使平均失效风险一直有下降趋势，而操作剖面则当测试预算成本达到一定值时，平均失效风险不再下降，而是保持不变。

结束语　本文针对安全关键软件系统利用交叉熵方法通过一种修正机制调节操作剖面，增加使用概率小的关键操作的测试机会，加速软件测试，在提高软件系统质量的同时降低软件测试成本。该方法是为实现测试目标而采用的一种导向性测试数据集生成方法，将整个测试数据集作为优化对象，利用“功效”值最好的测试数据集所包含的状态间转移信息，把修正测试剖面归结为一个“伴随”随机优化问题的迭代求解，用以指导整个测试数据集的生成，加速软件测试。

参 考 文 献

[1] Whittacker J A, Poore J H. Markov analysis of software specifications[J]. ACM Transaction on Software Engineering and Method, 1994, 2(1): 93-106

[2] Whittaker J A, Thomason M G. A Markov chain model for statistical software testing[J]. IEEE Transaction on Software Engineering, 1994, 20: 812-824

[3] Walton G H, Poore J H, Trammell J. Statistical testing of software based on a usage model[J]. Software-Practice and Experience, 1995, 25(1): 97-108

[4] Walton G H, Poore J H. Measuring complexity and coverage of software specifications[J]. Information and Software Technology, 2000, 42: 859-872

[5] 冯华，徐锡山，王戟. 统计测试中测试链与使用链的相似性判别[J]. 计算机工程与科学，2003，25(1)：17-19

[6] Gutjahr W J. Failure risk estimation via Markov software usage models[C]// E. Schoitsch, ed. SAFECOMP 96, Proc. of the 15th International Conference on Computer Safety. reliability and security. Springer, 1997: 183-192

[7] Gutjahr W J. Importance sampling of test cases in Markovian software usage models[J]. Probability in the Engineering and Informational Sciences, 1997, 11: 19-36

[8] Doerner K, Laure E. High performance computing in the optimization of software test plans[J]. Optimization and Engineering, 2002, 3: 67-87

[9] 颜炯，王戟，陈火旺. 基于重要抽样的软件统计测试加速[J]. 计算机工程与科学，2005，27(3)：64-66

[10] Doerner K, Gutjahr W J. Extracting test sequences from a Markov software usage model by ACO[J]. LNCS, Springer Verlag, 2003, 2724: 2465-2476

[11] Boer D P-T, Kroese D P, Mannor S, et al. A Tutorial on the Cross-Entropy Method [J]. Annals of Operations Research, 2005, 134: 19-67

[12] Margolin L. On the convergence of the cross-entropy method [J]. Annals of Operations Research, 2005, 134: 201-214

（上接第 120 页）

传恶意 XML 文档，达到发布欺骗性信息、破坏数据库系统等目的[9]。防御结点攻击的一种有效途径是对 XML 中每个元素的类型进行定义，包括数据类型定义、元素可能出现次数的最大/最小值及元素值的取值范围等。由于 XML/DTD 缺乏对文档结构、属性、数据类型等约束的足够描述，而 XML Schema 在此方面更具优势，所以也可以采用 XML Schema 作为 XML 文档模式描述语言，代替 XML DTD。对于提取并置入 SQL 语句中的数据及 SQL 语句本身，也需进行提炼并作安全性分析，以防止 SQL 注入攻击。如，所提取出的 XML 数据不允许含有 select, delete, update 等特定含义的字符。此外，还可以将数据编码成特定的码制（如十六进制），再通过后台系统提取数据并解码。

结束语　XML 正发挥着越来越重要的作用，如何通过有效的安全技术实现 XML 的安全应用已经成为人们研究的热点。虽然已经制定了多个有关的 XML 安全技术规范，但如何应用好这些规范来保证 XML 的安全性，仍是一个具有很大挑战性的课题。本文的工作力图通过一个应用系统的设计，探讨综合应用 XML 加密、签名等安全技术确保 XML 安全的模型以及实现方法。从实践来看，该系统已经建成并投入使用，并已在管理工作中取得了较好效果。

参 考 文 献

[1] W3C XML Core Working Group. Extensible Markup Language (XML) 1.0 (Fourth Edition) W3C Recommendation [EB/OL]. http://www.w3.org/TR/xml/, 2006, 8

[2] BlakeDournaee. XML安全基础[M]. 北京：清华大学出版社，2003

[3] Eastlake D, Reagle J. XML Encryption Syntax and Processing W3C Recommendation [EB/OL]. http://www.w3.org/TR/xmlenc-core, 2002, 12

[4] Eastlake D, Reagle J, Solo D. XML-Signature Syntax and Processing W3C Recommendation [EB/OL]. http://www.w3.org/TR/xmldsig-core, 2002

[5] Ford W, Hallam-Baker P, Fox B, et al. XML Key Management Specification (XKMS) [EB/OL]. http://www.w3.org/TR/xkms/

[6] 顾天竺，沈洁，陈晓红，等. 基于 XML 的异构数据集成模式的研究[J]. 计算机应用研究，2007，24(4)：94-96

[7] Donald M M, Johansson E. C# 数据安全手册 [M]. 北京：清华大学出版社，2006

[8] 丘威，张立臣. XML DTD 规范化处理研究[J]. 计算机科学，2006，33(7)：63-67

[9] Buehrer G, Weide B W, Paolo A, et al. Using parse tree validation to prevent SQL injection attacks [EB/OL]. http://portal.acm.org/citation.cfm? id=1108496, 2007