基于信息隐藏的外包数据库版权保护系统

朱 勤[1,2] 陆志明[2]
(南通大学计算机科学与技术学院 南通 226019)[1]
(复旦大学公共管理与公共政策研究国家创新基地 上海 200433)[2]

摘 要 在外包数据库运行模式下,由第三方提供的数据库服务器处于非信任域,存在数据文件盗版、数据内容篡改等安全风险。构建了一个基于信息隐藏技术的外包数据库版权保护系统,综合运用数字水印、PKI机制与数字证书、数字签名及可信硬件模块USB Key等技术,设计相应的水印协议,实现对外包数据库的版权保护。与传统的以数据加密、数字签名等方法为主要技术手段的解决方案相比,它具有冗余存储量及网络附加流量小、隐蔽性好、验证信息难以删除等优点。

关键词 外包数据库,信息隐藏,数字水印,版权保护,USB Key
中图法分类号 TP309.2,TP311.13

Copyright Protection System for Outsourced Database Based on Information Hiding

ZHU Qin[1,2] LU Zhi-Ming[2]
(College of Computer Science and Technology, Nantong University, Nantong 226019, China)[1]
(The State Innovative Institute for Public Management and Public Policy Studies, Fudan University, Shanghai 200433, China)[2]

Abstract In outsourced database scheme, since the database server, which is provided by the third part, is not in the trust domain, the data files may be pirated, and the data contents may be tampered with. A copyright protection system for outsourced database using information hiding technique was built, which was based on algorithms of database watermarking, as well as their corresponding protocols, and was combined with technologies of PKI, digital signature and USB Key. Compared with traditional solutions which of data encryption or digital signature, the copyright protection system for outsourced database based on information-hiding possesses advantages of less redundant capacity either for storage or for communication, better imperceptibility, and validation information harder to be removed.

Keywords Outsourced database, Information hiding, Digital watermarking, Copyright protection, USB Key

1 引言

外包数据库(outsourced database)亦称"作为服务的数据库"(database as a service),其概念由美国加利福尼亚大学Irvine分校(UCI)的H. Hacigumus在2002年ICDE会议上首次提出[1]。在外包数据库运行模式中,组织或个体将自己的数据库业务外包给数据库服务提供者运行,外包服务提供者为数据拥有者及数据库用户提供远程的数据库创建、存储、更新、查询及数据库服务器软硬件维护与升级等服务。在某种程度上,外包数据库模式可以实现数据库管理的专业技术对整个社会的共享服务,从而避免组织或个体在数据库管理方面重复的人力和物力投资,并且具有更高的运行维护水平和更好的可扩展性。

由于由第三方提供的数据库服务器处于非信任域,外包数据库系统存在数据文件盗版、数据内容篡改等安全风险,因此,外包数据库系统必须提供对数据库内容的安全保护。归纳起来,外包数据库安全保护的需求主要包括:数据保密与安全存储、数据库访问控制、数据库版权保护、数据完整性验证等[2]。

本文提出了一种基于信息隐藏技术的外包数据库版权保护机制。与传统的以数据加密、数字签名等方法为主要技术手段的解决方案相比,它具有冗余存储量及网络附加流量小、隐蔽性好、验证信息难以删除等优点。

2 数字内容保护技术概述

国外从20世纪80年代开始研究电子版权管理(electronic copyright management, ECM)。90年代,在欧盟的资助下,ECM的研究扩展到了法律、技术和商业等方面,称为数字权益管理(digital rights management, DRM)。DRM技术是对各类数字内容的知识产权进行保护的一系列软硬件技术,用以保证数字内容在整个生命周期内的合法使用与安全传播,平衡数字内容价值链中各个角色的利益和需求,促进整个

到稿日期:2009-02-13 返修日期:2009-05-08 本文受中国博士后科学基金项目(20080440573),江苏省高校"青蓝工程"优秀青年骨干教师培养项目(2008),南通大学博士科研基金(2008)资助。

朱 勤(1971-),男,博士后,副教授,主要研究方向为数据库与信息系统、复杂系统建模,E-mail:zhuqin@fudan.edu.cn;陆志明(1978-),男,博士后,主要研究方向为系统工程。

数字化市场的发展和信息的传播[3-5]。DRM 技术不仅仅局限于数据内容的访问控制，还扩展到对数字资产各种形式的使用进行描述、识别、交易、保护、监控和跟踪等各个过程，对数字内容版权的保护贯穿于数字内容从产生到分发、从销售到使用的整个内容流通过程中，涉及到整个数字内容价值链[5]。目前已有不少基于 DRM 机制的数字内容保护系统（如 Microsoft 公司的 WMRM，IBM 公司的 EMMS，Adobe 公司的 Content Server 等）被开发出来并得到不同程度的应用。

在技术上，目前数字版权保护系统主要基于密码学、数字签名、可信硬件模块和水印技术来实现。密码学技术能阻止对拷贝的非授权访问，但是一旦解密后数字内容完全暴露，不再提供任何保护措施。数字签名技术可以提供对信息来源的可靠性和数字内容真实性与完整性的验证，但是数字签名与数字内容相分离，很容易去除，而且只要数字内容稍作修改，签名就无效。可信硬件模块主要提供防篡改功能，可用来保护秘密密钥等安全性要求较高的数据，其前提是每个合法用户都要有相应硬件支持。数字水印一般用来进行数字内容的版权标识与盗版追踪，但是单纯依靠数字水印进行数字内容保护的程度有限，一般主要用于被动鉴别出非法复制和盗用的数字产品，而不能主动防止对数字内容的非法复制与访问。

综上所述，单独靠某一种技术很难实现全面的数字内容保护，而多种技术相互补充、发挥各自优势的一体化解决方案有可能使系统的安全性与可用性得到统一。

3 外包数据库版权保护系统构成

对于外包数据库版权保护而言，由于数据库的数据容量大、更新频繁等特点，难以直接应用一般多媒体版权保护技术。对数据库直接加密会引起巨大的运算开销，数据查询等常规操作的可用性也大打折扣；对数据库直接进行数字签名同样涉及巨大的运算开销，且数据库的每一次更新都会使原有签名无效。

本文基于数据库信息隐藏技术，综合运用数字水印、PKI（public key infrastructure，公钥基础设施）机制与数字证书、数字签名及可信硬件模块 USB Key 等技术，应用数据库水印算法，设计相应的水印协议，构建外包数据库版权保护系统。

该方案将 PKI 机制引入外包数据库版权保护系统，由认证机构向系统中各实体发放数字证书，各实体以安全存储设备 USB Key 存放数字证书，进行身份认证。利用版权服务中心提供版权登记与仲裁服务，构成基本的可信计算平台，为安全水印协议的实现提供运行环境。系统结构如图 1 所示。

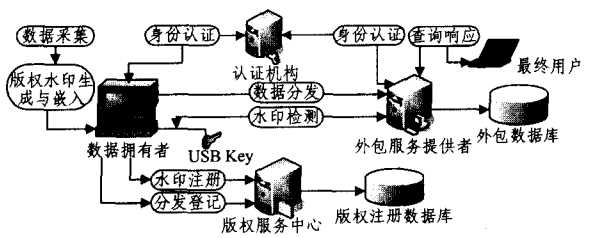

图 1 外包数据库版权保护系统结构

系统的各个组成部分及主要功能如下：

1）认证机构（certification authority，CA）。安全可信的第三方认证机构，受理用户注册，发放数字证书；接受系统中实体要求验证对方身份的请求，为各实体进行身份认证，分配 RSA 密钥对；提供可信时戳服务。

2）版权服务中心。接受版权登记，管理维护版权信息库；记录数据拥有者每一次的数据分发操作；在发生版权纠纷时，担当仲裁者角色，根据争议双方提供的证据作出版权判定。

3）数据拥有者。数据库版权的拥有者，采集数据，生成版权水印并嵌入数据库，分发给外包数据库服务提供者；向版权服务中心注册数据库版权。

4）外包服务提供者。外包数据库服务的提供者，同时也是数据拥有者的协议用户，接受数据拥有者分发与更新的数据，存储与管理数据库；响应最终用户的查询请求。

5）USB Key。可信硬件模块，加密存储用户 RSA 密钥对及设备数字证书，具有全局唯一性，用于身份认证与信息保护。

6）最终用户。外包数据库服务的最终使用者，不纳入外包数据库版权保护系统的安全保护范畴。

4 数据库水印协议

基于信息隐藏的外包数据库版权保护系统的工作流程主要由 5 组水印协议来描述，包括：数字证书签发协议、水印注册协议、水印加载协议、验证协议及仲裁协议。

协议中各实体的符号标识定义如下：

CA 认证机构；

CSC 版权服务中心；

DO 数据拥有者；

OSP 外包服务提供者。

4.1 数字证书签发协议

数字证书的申请和签发是系统中各实体必须完成的初始化工作，是 PKI 体系结构下一切安全交互活动的信任基础。

以数据拥有者 DO 为例，其数字证书签发的协议流程描述如下：

数字证书签发协议

1）CA 在网络上公开以下内容：CA 的公钥 K_{CA_P}；CA 所选用的公钥密码机制的加密算法 D 与解密算法 E；CA 所选用的签名算法 S 与签名验证算法 V；

2）数据拥有者 DO 向 CA 申请数字证书；

3）CA 审核 DO 的请求，验证 DO 的身份；

4）若 DO 的请求审核成功，则：

5）CA 生成 DO 的数字证书 DC_{DO} 并写入 USB Key；

6）CA 将含 DC_{DO} 的 USB Key 分发给 DO；

7）DO 公开自己的公钥 K_{DO_P}。

外包服务提供者 OSP 及版权服务中心 CSC 的数字证书申请和签发过程与此类似。数字证书签发完毕后，OSP 与 CSC 分别公开自己的公钥 K_{OSP_P} 与 K_{CSC_P}。

4.2 水印注册协议

数据拥有者在对外发布一个数据库之前，必须先到版权服务中注册水印信息。水印注册协议使数据拥有者建立起与版权服务中心之间的安全信道，安全可靠地完成数据库水印注册。

数据库水印信息 DBWI 的数据结构定义为：

```
DBWI {
    String DBName; // 数据库名称
    String DBVersion;// 数据库版本号
```

String Columns; // 数据库列名列表
String DBOwner; // 数据拥有者
String Watermark;// 水印标识
String WMKey; // 水印密钥
DateTime DT; // 发布时间
}

水印注册协议的工作流程如图 2 所示。

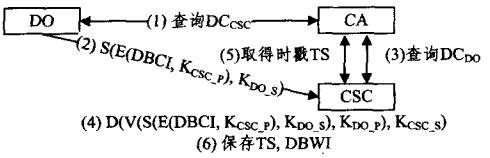

图 2 水印注册协议工作流程

4.3 水印加载协议

水印加载协议使数据拥有者与外包服务提供者即数据库的协议用户之间建立起安全信道。由数据拥有者进行水印嵌入运算，并将含水印的数据提交给外包数据库存储。

水印加载协议的工作流程如图 3 所示。

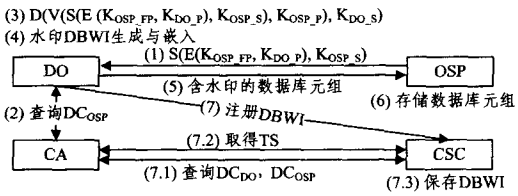

图 3 水印加载协议工作流程

水印加载协议实现了版权水印的数据库嵌入，并在版权服务中心注册备案。对于已完成水印注册的数据库，其在同一个水印密钥下进行的数据库后续更新无须再到版权服务中心进行注册。以数据库元组为水印嵌入的基本运算单位，满足了数据库动态更新的水印同步要求。

4.4 验证协议

数据拥有者通过验证协议检测数据库中的水印来验证水印是否被正确添加，利用提取出来的版权水印可以证明数据库版权。

验证协议的工作流程描述如下：

验证协议

1)数据拥有者 DO 向被检测数据库提交投影－选择类查询；

2)OSP 向 DO 返回查询结果；

3)DO 运行水印检测检测算法，提取版权水印；

4)DO 生成数据库版权信息 DBWI，用版权服务中心 CSC 的公钥 K_{CSC_P} 加密，并用自己的私有密钥 K_{DO_S} 签名后，将消息[S(E(DBWI,K_{CSC_P}),K_{DO_S})]发送给 CSC；

5)CSC 用 DO 的公钥 K_{DO_S} 验证消息签名，验证通过后用自己的私钥 K_{CSC_S} 解密消息：D(V(S(E(DBWI, K_{CSC_P}), K_{DO_S}), K_{DO_P}), K_{CSC_S})，获得水印信息 DBWI；否则丢弃消息并结束；

6)CSC 将 DBWI 与版权注册数据库中的记录进行比对，若匹配，则对 DO 出具版权证明，否则结束；

7)若被检测数据库为 OSP 数据库，则 DO 版权得到验证；否则，DO 确认该数据库为盗版。

4.5 仲裁协议

当系统中的实体间出现版权纠纷时，由版权服务中心按照仲裁协议进行版权的最终判定。

假设实体 A 和 B 同时声称自己对同一个数据库拥有版权，版权服务中心 CSC 启用的仲裁协议工作流程描述如下：

仲裁协议

1)实体 A,B 分别向版权服务中心 CSC 提交含各自水印密钥的数据库水印信息 DBWI_A,DBWI_B 及各自的水印检测算法；

2)CSC 到版权注册数据库中分别比对 DBWI_A,DBWI_B,若均无匹配，则判定 A,B 均为非版权所有者，结束仲裁；

3)若 DBWI_A,DBWI_B 在版权注册数据库仅有一个匹配，CSC 运行对应实体提供的水印检测算法提取水印；若水印匹配，则判定对应实体为版权所有者，结束仲裁；否则判定 A,B 均非版权所有者，结束仲裁；

4)若 DBWI_A,DBWI_B 在版权注册数据库均有相应匹配(必有一方盗取了数据拥有者的水印密钥并虚假注册)，且与水印检测算法提取出的水印相匹配，则 CSC 查询对应的水印时戳 TS_A,TS_B,判定时戳早于对手者为版权所有者，结束仲裁。

由于版权注册数据库中的水印时戳是由 CA 认证中心提供，并由版权服务中心与水印信息绑定后保存的，具有全局唯一性与权威性，因此仲裁协议有效抵抗了对外包数据库的重复水印攻击与水印逆向攻击。

上述协议中凡涉及数字证书的查询验证，均需要各实体 USB Key 的参与才能完成。协议中涉及的数据库水印嵌入与验证算法[6]为本文的前期工作，此处不再赘述。

4.6 协议安全性分析

本方案所述数据库水印协议的安全性主要体现在以下几个方面：

1)基于 PKI 体系结构，构建了系统各实体间的安全信道；

2)基于数字证书的权威性、全局唯一性与不可伪造性，实现了实体身份的可验证性与操作的不可抵赖性，有效抵抗了中间人攻击；

3)应用数据加密及数字签名技术，无明文密钥在实体间传输，降低了安全风险；同时，数字签名实现了消息传递的真实性、完整性及不可抵赖性；

4)应用可信硬件模块 USB Key 作为数字证书与秘密密钥的安全存储设备，加强了对关键信息的保护；

5)应用系统时戳有效抵抗了重复水印攻击与水印逆向攻击。

由于协议的设计均假设认证机构与版权服务中心提供的服务是安全可靠的，因此协议的安全性建立在可信计算平台安全性的基础之上。

5 系统实现

在 Windows 2003 Server 操作系统环境下，以 Borland C++ Builder 6.0为编程工具，建立了基于信息隐藏的外包数据库版权保护原型系统。

5.1 CA 服务器配置与数字证书签发

Windows 2003 Server 系统内建了对 PKI 应用的支持，本实验通过配置实现了 CA 服务器的构建与数字证书的签发，实现步骤的要点简述如下：

1)在服务器端安装并配置 Windows 2003 Server 系统自带的 CA 认证组件，即数字证书服务器；

2)在客户端从浏览器访问数字证书服务器，获取并安装 CA 根证书；

3)在服务器端 IIS 服务管理器中配置 SSL(secure sockets layer，安全套接字层)通讯协议；

4)在客户端从浏览器访问数字证书服务器，申请并安装数字证书。

由于本方案采用可信硬件模块 USB Key 来保存数字证书，因此，步骤 4)在实际操作中是通过配置 USB Key 来申请并保存数字证书的。

5.2 USB Key 应用

原型系统使用北京飞天诚信公司的 USB Key 产品 ePass1000ND。ePass1000ND 采用 8 位处理器，提供 8kB 安全数据存储空间，无需驱动。每个 ePass1000ND 硬件都有一个 64 位序列号，作为全局唯一标识[7]。

使用 ePass1000ND 访问 CA 服务器 SSL 加密站点，申请并保存数字证书与秘密密钥。对系统中各实体的身份认证均通过验证 USB Key 中保存的数字证书来完成。在 USB Key 的应用中，通过以下设置加强 USB Key 应用的安全性：

1)设置 USB Key 只允许单进程访问，确保不被跟踪；

2)对 USB Key 访问设置 PIN 码(个人识别码)保护；没有 PIN 码与管理员密码无法访问某些文件；

3)设置 USB Key 的 PIN 码的最大可重试次数。当 PIN 码连续输入错误达到最大可重试次数时，USB Key 自动锁死，可防范穷举攻击。

5.3 MS CryptoAPI 编程接口应用

目前业界实际使用的 PKI 编程标准主要有两种：RSA Lab 的 PKCS # 11 和 Microsoft 的 CryptoAPI。MS CryptoAPI 是 Win32 应用程序的通用加密接口，在本方案中被用来实现 USB Key 证书存储、RSA 密钥对生成及对消息传输的数字签名。MS CryptoAPI 的体系结构如图 4 所示。

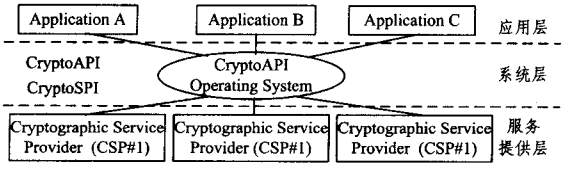

图 4 MS CryptoAPI 的体系结构

MS CryptoAPI 环境中实际提供加密服务的模块是 CSP (Cryptographic Service Provider，加密服务提供者)，应用程序不直接与 CSP 打交道，而是通过 CryptoAPI 来协调应用程序与 CSP 之间的交互。由于 ePass1000ND 已经实现了一个 PROV_RSA_FULL 类型的 CSP，因此本系统在 USB Key 应用中只需编程调用 CryptoAPI 来建立与 ePass1000ND 之间的通讯。

结束语 基于信息隐藏的外包数据库版权保护系统构建在 PKI 基础之上，借助 PKI 机制较为容易地实现了身份认证、密钥分发、可信时戳等必需的安全性要求和功能。水印技术与加密、数字签名等技术的结合弥补了水印在主动安全保护方面的不足。设计的数据库水印协议满足了数据库数据容量大、更新频繁的要求，具有抵抗多种水印攻击的能力。可信硬件模块 USB Key 的引入，增强了系统各实体间安全交互的能力。

该系统的主要不足在于：系统的整体安全性建立在第三方可信计算平台基础上，存在一定的安全失效风险。

参 考 文 献

[1] Hacigumus H, Iyer B, Mehrotra S. Providing Database as a Service[C]// ICDE 2002. San Jose, California, USA, 2002

[2] 朱 勤，于守健，乐嘉锦，等. 外包数据库系统安全机制研究[J]. 计算机科学，2007，34(2)：152-156，195

[3] Rosenblatt W, Trippe W, Mooney S. Digital Rights Management: Business and Technology[M]. New York: M &T Books, 2002

[4] Garnett N. Digital Right's Management, Copyright, and Napster [J]. ACM SIGecom Exchanges, 2001, 2(2): 1-5

[5] 俞银燕，汤帜. 数字版权保护技术研究综述[J]. 计算机学报，2005，28(12)：1957-1967

[6] 朱勤，刘良旭，乐嘉锦. 一种基于 m 序列的关系数据库鲁棒水印算法[J]. 小型微型计算机系统，2008，29(8)：1486-1490

[7] Epass1000ND[EB/OL]. [2007-12-02]. http://www.ftsafe.com.cn/products/viewproduct.php? p=epass1knd, 2008

(上接第 98 页)

[4] Ratnasamy S, et al. A Scalable Content - Addressable Network [C]// Proceeding of ACM SIGCOMM. NewYork: ACM Press, 2001: 161-172

[5] Rowstron A, Druschel P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems[C]// IFIP/ACM International Conference on Distributed Systems Platforms. Kluwer Academic Press, 2001: 329-350

[6] Cooper B F, Garcia-Molina H. Ad hoc, self-supervising peer-to-peer search networks[R]. Stanford University, 2003

[7] Liu Y, Zhuang Z H, Xiao L, et al. AOTO: Adaptive overlay topology optimization in unstructured P2P systems[C]// Proc. of the IEEE GLOBECOM 2003. San Francisco, 2003

[8] Xiao L, Liu Y, Ni L M. Improving unstructured peer - to - peer systems by adaptive connection establishment[J]. IEEE Trans. on Computers, 2005, 54(9): 1091-1103

[9] Condie T, Kamvar S D, Garcia-Molina H. Adaptive peer-to-peer topologies[C]// Lambrix P, Duma C. eds. Proc. of the 4th Int'l Conf. on Peer-to-Peer Computing. New York: IEEE Press, 2004: 53-62