数字图书馆访问控制技术

李春旺　　陈瑾

（中国科院文献情报中心　北京 100080）　（中国科院研究生院　北京 100039）

文　摘 文章讨论了数字图书馆安全系统与访问控制技术，着重研究访问控制系统模型及数字资源的控制策略，并对系统实现提出了建议。

关键字　数字图书馆 信息安全 访问控制　模型

Research on Access Control Technology in Digital Library

Li Chunwang

(Documentation and Information Center of the Chinese Academy of Sciences, Beijing, 100080)

Chen Jin

(Graduate School of the Chinese Academy of Sciences, Beijing, 100039)

Abstract： This paper analyzes digital library security and access control, discusses the models of access control and management of digital source, and gives some advice on development and implement.

Keywords： Digital Library, Information Security, Access Control, Model

2001年底，美国通用会计师事务所对美国24家最大联邦代理机构的审计结果表明，这些机构在信息安全管理与访问控制方面都存在严重的漏洞[1]。美国计算机安全协会（Computer Security Institute）与联邦调查局(FBI)的研究认为，数字信息管理中最主要的安全威胁来自未经授权的内部知情人员，占71%[2]。在数字图书馆领域，无论是科科技信息的出版、组织、交流还是永久性保存，安全访问控制技术都扮演着重要角色，加强访问控制技术的研究与应用，对保证数字图书馆的健康、持续发展具有重要意义。

1 数字图书馆安全系统与访问控制策略

（1）数字图书馆安全控制系统

一个完整的数字图书馆安全控制系统包括四部分内容：认证(Authentication)、授权(Authorization)、访问控制（Access Control)、审计(Audit)。认证是确认信息真实性、准确性、完整性的过程，是信息安全管理的基础[3]。授权是指资源的所有者或控制者向用户授予相关访问权限的过程。拥有授权的用户经访问控制系统的验证后，被准许访问指定资源。访问控制就是根据用户身份、授权以及请求访问资源属性，对用户请求做出允许访问或拒绝访问的决策过程。访问控制不但要保障授权用户合理访问信息资源，而且能排除未授权用户对资源的非法访问或越权访问。在复杂的信息系统中，授权与访问控制常常被结合在一起（本文以下所提到的访问控制包括授权管理)。审计是对所有与安全相关的信息活动进行审核、稽查，以便及时发现安全隐患，追查造成安全事故的原因，对安全策略的调整与修改提供支持。

（2）访问控制系统

在功能组成上，访问控制系统主要包括访问控制决策功能（Access Control Decision Function , ADF）与访问控制实施功能(Access Control Enforcement Function, AEF)两部

分。前者响应用户访问请求，依据规则做出允许或禁止访问的判决，后者执行判决结果，批准或拒绝用户的访问请求。在对象构成上，访问控制包括用户对象、资源对象、权限对象[4]。用户对象是访问数字图书馆资源的用户或代表用户执行的程序，通常称为主体(Subject)或发起者(Initiator)。资源对象是指受保护的数字资源，称为客体(Object)或目标(Target)。而权限对象是用户对象与资源对象之间的一种关联，表明一个用户在一个资源对象上的一系列访问权限以及对应的访问控制规则。

（3）访问控制策略

常规的访问控制策略包括：自主访问控制、强访问控制、基于角色的访问控制。自主访问控制(Discretionary Access Control)是在确认主体身份及所属组的基础上，对访问进行限定的一种控制策略[5]。强制访问控制(Mandatory Access Control)是对用户和目标资源实施分级控制的策略，资源的密级可以分为：绝密、秘密、机密、限制、无密级。基于角色的访问控制（Role-Based Access Control, RBAC）是在前两种策略的基础上发展起来的，它根据用户在系统中所属角色进行授权，并根据角色权限实施访问控制，一个用户可以充当多个角色，一个角色也可以由多个用户担任[6]。RBAC包括三个基本组成部分：

①用户-角色分配（User-Role Assignment, UA）。即根据用户的身份和属性赋予其相应的角色。为了简化角色分配操作，可以将用户分组，再将组划分为不同的层次，各个层次的组具有单继承的关系。角色也可以划分为不同的层次，各个层次之间的角色具有单继承或多继承的关系[7]。

②角色-权限分配(Permission-Role Assignment, PA)。指赋予角色相应的权限许可集。权限分配过程中要遵循最小特权原则(the least Privilege principle)，即给每个角色分配足够且仅仅是足够的许可集，最小特权原则对于满足完整性目标是非常重要的[8]。一般的权限操作有授权、权限撤销、权限冻结、权限激活、权限排除、权限扩散、权限修改等[9]。为解决角色权限冲突，防止欺诈行为，RBAC采取静态的或动态的方式实现职责分离[10][11]。

③角色-角色分配(Role-Role Assignment, RA)。指对角色层次关系的管理。角色层次关系主要包括权限继承关系、角色继承关系、管理层次关系和约束继承关系。角色层次关系具有自反性、传递性和反对称性，是一种偏序关系[12]。

2 访问控制模型

2.1 分布式环境下RBAC系统参考模型

分布式资源组织与服务是数字图书馆最大的特点。与集中式环境相比较，分布式环境下的角色管理要复杂的多。根据系统事件、策略、能力以及角色分类方式的不同，分布式环境下RBAC系统具有多种不同的结构模型。

（1）按UA管理方式分，RBAC系统有4种类型。①推送方式（Push Architecture），即UA信息以事件通知方式被推送到所有子系统中。②拉取方式(Pull Architecture)，每个子系统建立与中心系统的联系，在本地维护一个角色信息缓冲，定期更新。③查询方式(Lookup Architecture)，与拉取方式类似，但本地不保留角色信息缓冲，需要时时查询中心系统数据库。④基于信任证书方式(Credential based Architecture)，即通过给用户签发信任证书，支持分布式服务器信任证书的验证，典型的信任证书是Kerberos票据、X.509证书。

（2） 依据RH的不同，RBAC可分为3种体系结构：①同构体系（Homogeneous Architecture），即有一个单一的、全局的层次结构模型，所有系统都使用该层次结构，遵循统一的定义方法，本地系统不能引进新的角色继承关系。例如：定义角色x比角色y级别高，则该关系在整个分布式系统中将保持一致性。②异构体系（Heterogeneous Architecture），本地系统可以对没有限制的角色定义新的继承关系，并作为全局层次模型的一部分，但它不能修改全局模型，也不能定义受限制的角色。③联邦体系（Federated Architecture），在联邦框架下，不同系统间角色继承关系是不同的，如在一个系统中角色x比角色y级别高，而在另一个系统中可能是角色y比角色x级别高。联邦结构框架没有单一的全局性角色层次模型。

（3）根据系统能力不同，可以将RBAC分为支持角色层次模型的系统与不支持角色层次模型的系统，后者可以模拟一个层次模型，即当一个用户被分配一个高级角色时，明确将该用户分配给所有低级角色。

（4）根据用户和角色在系统中出现的情况不同，RBAC包括四种结构类型，其中，所有用户或所有角色都存在的状况是不希望出现的，它违背了最小特权原则；同样，所有的角色存在于所有系统中的状况也是不希望的[13]。

2.2 基于内容的访问控制

Nabil R. Adam等人提出一个基于内容的数字图书馆授权模型（Digital Library Authorization Model, DLAM）[14,15]。它提供基于用户资格与特征的肯定性或否定性授权，对数字图书馆资源对象实行内容依赖与内容独立两种方式的访问控制。受控对象包括多个控制级别，如图书馆对象集、对象的组成部件等。DLAM包括四部分内容：代表各种媒体类型的数字图书馆对象(DL Object，DLO)、信任证书、权限、访问授权。DLAM具有以下特点：①授权可以依照一定规则沿着概念和信任证书层级结构进行传播，从而精确表示授权之间的相关性；②可以管理肯定授权与否定授权中产生的冲突；③既可以对信息对象整体内容进行访问控制，也可以对其组成部件进行控制；④由于该模型适用于传统的图书馆，使纸质图书馆向数字图书馆过渡变得十分容易。该模型在全球法律信息网络系统(Global Legal Information Network, GLIN)[16]项目中得到应用。

2.3 GTRBAC

B. M. Thuraisingham等人的研究[17]表明，实现Web应用安全所需要的基本条件是：支持访问控制策略描述与实施的工具与机制；基于时间限制条件的安全工作流；安全的联邦协作组织。针对Web信息安全，James B. D. Joshi在他的博士论文[18]中提出了GTRBAC（Generalized Temporal Access Control）模型。GTRBAC是一个基于事件驱动的模型系统，它通过引进一组复杂的临时约束，实现对RBAC的扩展。这些临时约束包括：在角色使用与分配上的周期性和持久性约束、角色激活的持久性约束与基本约束等。

GTRBAC模型的特点是：①GTRBAC引入多种类型的角色层次模型，提供针对各种临时约束的权限继承与角色激活场景，定义了层级关系派生规则。一个安全管理工具利用这些规则，可以发现应用系统中存在的安全漏洞。②GTRBAC包含一种复合层次模型，它由不同类型的角色关系构成。在复合层次模型中，可以产生多个角色集，一个用户在同一个会话（session）内可以同时激活这些角色，以满足不同的控制需求。作为一个动态变化的系统，GTRBAC还研

究了层次模型的演变规则。③GTRBAC利用一个通用框架来描述宽泛的基于时间的基本约束，它能够描述GTRBAC所有的状态类型。④GTRBAC开发了一个触发器，可以在系统环境与事件中捕获复杂的依赖关系，为基于上下文的访问控制提供一个基本的RBAC模型。

2.5 其他模型

（1）基于系统先决条件的授权模型

在用户角色授权模型URA97[19]中，由角色指派和撤销机制定义的先决条件(prerequisite Condition，PC)只约束被授权用户已经拥有的角色，而不约束系统中其他用户拥有的角色，因而会引起授权冲突，存在严重的缺陷。基于系统先决条件的授权模型(SPC-based authorization model,SBAM)定义了新的角色指派和撤销机制,从而克服了URA97的缺陷，能够正确地实施安全策略，更好地满足实际的安全需求[20]。

（2）基于工作流状态的动态访问控制模型

基于工作流状态的动态访问控制机制是指在确定用户对资源的访问权限时,需要考虑工作流的当前状态，这样，传统的访问控制矩阵将由二维(用户，信息资源)变成三维(用户、信息资源和工作流状态)。通过实施动态的访问控制，未授权资源被访问的可能性会减少，从而使工作流的执行更安全[21]。

3 数字资源访问控制管理

（1）访问控制粒度

理论上，可以对数字图书馆系统中的所有资源进行访问控制，包括网络中的所有对象，如：目录文件、数据库记录、Web网页信息等。实际应用中，不同的系统对资源对象的访问控制存在不同的详细程度，即访问控制粒度，如资源主机、目录、文件等。访问控制粒度可分为粗粒度(coarse grained)、中粒度(medium grained)和细粒度(fine grained)，这些访问控制级别之间并没有严格的区分标准。一般来说，能够控制到文件或数据库记录级别的可以称为细粒度访问控制，只能控制到主机对象级别的是粗粒度访问控制。资源控制粒度可以根据系统的安全级别来决定，在数字图书馆系统中，每一份资料、文献都受到版权许可和其他外部承诺的限制，对这些文献的访问控制一般都应该达到细粒度访问控制。

（2）数字资源的属性管理

访问控制实施的依据是数字对象的属性描述，属性信息通常作为管理元数据和信息对象一起存放。在数字图书馆系统中，通常采取以下方式对信息对象进行属性划分：

①公开访问和受限访问。例如，将全文信息与概要、索引、文摘、元数据等信息分离开来，允许用户不受限制地访问索引、概要、文摘、元数据信息，但要获得全文信息必须通过身份认证与授权。

②版权许可。需要出版商授权的信息以及带有访问条款和条件的信息，如电子期刊数据库系统为订购用户提供IP授权。

③时间和物理特性。数字图书馆根据数字信息的时间或物理特征来制定访问策略，如斯坦福大学的HighWire出版社在对其出版的全部期刊实行开放访问服务时做了一些限定,即期刊出版一段时间后才向公众开放自由访问，一般期限是一年[22]。

④媒体类型。根据数据的存放格式和媒体类型制定访问策略。例如将声音文件、图像

文件和可执行程序区别对待，分别设定不同的策略[23]。

数字图书馆需要对整个馆藏系统、分馆馆藏、单个信息对象以及对象组成元素属性给予分别定义，以实现细粒度的访问控制。同时，对访问控制属性赋值也应该有相应的粒度，将具有相同属性的资源归结为一个集合，赋予整体控制属性。

4 系统实现

由于人们的信息安全意识比较淡薄，造成在法律制定以及技术开发上的相对滞后，特别是基于访问控制的知识产权保护技术没有得到很好地解决，在一定程度上影响了数字图书馆的正常发展。当前，国内数字图书馆在进行访问控制系统建设时应注意以下问题：

（1）将访问控制功能实现纳入到数字图书馆总体建设之中。只有将访问控制功能与数字图书馆安全系统中的其他部分以及数字图书馆总体功能构成进行统一考虑，才能保证各安全部件之间协调工作，兼顾知识产权人、用户以及图书馆中介等各方利益，保证数字图书馆从资源组织到用户服务各项目标的实现。

（2）在访问控制模型设计上，要充分考虑特定系统资源组成以及服务需求特点。访问控制技术的研究不仅包括相关计算机技术，对数字图书馆来说，同时还包括数字资源的组织、描述与分布技术。只有采取良好的策略，在对资源内容、属性、层次结构进行全面揭示的基础之上，才能实现任意粒度的资源控制目标。

（3）在系统实现上，将访问控制技术与 XML 技术相结合。XML 文档包含广泛的语义信息与结构信息，特别适合用来描述复杂的数字图书馆用户、资源、角色属性以及各种映射关系、约束条件信息等。在对资源实施细粒度访问控制时，采用 XML 技术具有突出的优势，它可以增加系统的扩展性、兼容性以及可持续发展性。

参考文献

1 Critical Infrastructure Protection: Significant Homeland Security Challenges Need to Be Addressed. GAO-02-918T, 2002,7. http://www.gao.gov/new.items/d02918t.pdf.(检索日期：2003-10-17)

2 A. K. Ghosh, E-Commerce Security: No Silver Bullet. Proceedings of the Twelfth IFIP WG 11.3 Working Conference on Database Security, Greece, 1998,7. http://citeseer.nj.nec.com/95582.html.(检索日期：2003-10-17)

3 李春旺.网络环境下信息安全认证.图书馆杂志, 2003,(2):25～27,32

4 Ravi Sandhu and Jaehong Park. Usage Control: A Vision for Next Generation Access Control. 2003. http://www.list.gmu.edu/confrnc/misconf/2003_MMS_UCON.pdf.（检索时间：2003-10-18）

5 B. Lampson. Protection. In the Princeton Symposium on Information Sciences and Systems. 1971,3. http://citeseer.nj.nec.com/287804.html.检索时间：2003-10-18）

6 David F. Ferraiolo,Ravi Sandhu 等. Proposed NIST Standard for Role-Based Access Control. 2001,8.http://csrc.nist.gov/rbac/rbacSTD-ACM.pdf. (检索时间：2003-10-20)

7 余晖,刘亚军.基于角色访问控制的研究与实现.微机发展.2003,13(1):13~15

8 王广慧.基于角色的访问控制.网络安全技术与应用.2002,(9):21~22

9 梅苏文,高县明等. 基于角色权限管理模型的设计与实现. 现代计算机：下半月

刊.2002,(11):10~13,38
10 David F. Ferraiolo,Ravi Sandhu 等. Proposed NIST Standard for Role-Based Access Control. 2001,8. http://csrc.nist.gov/rbac/rbacSTD-ACM.pdf. (检索时间：2003-10-20)
11 付志峰,张焕国.RBAC 系统中职责分离的实现.计算机工程.2003,29(6):61~63
12 吴和生,伍卫民,蔡圣闻,黄皓,谢俊元.分布式环境下 RBAC 的高效实现.计算机工程.2003,29(6):134~136
13 Venkata B.and Ravi S. Push Architectures for User Role Assignment. 2000. http://csrc.nist.gov/nissc/2000/proceedings/papers/018.pdf. (检索时间：2003-10-20)
14 Nabil R. Adam.A content-based authorization model for digital libraries.2000. http://cimic.rutgers.edu/~atluri/diglib.pdf. (检索时间：2003-10-20)
15 E. Ferrari,N.R. Adam,V. Atluri, E. Bertino, U. Capuozzo.An authorization system for digital libraries.2002. http://www.iasi-tekno.com/discussion_material/IT/s007780200063.pdf. (检索时间：2003-10-18)
16 Global Legal Information Network . http://www.loc.gov/law/glin/ (检索时间：2003-10-18)
17 B. M. Thuraisingham, C. Clifton, A. Gupta, E. Bertino, E. Ferrari. Directions for Web and E-Commerce Applications Security. WETICE 2001: 200-204. http://ce.sejong.ac.kr/~shindk/031_gia/xml/IEEE/Directions for Web and e-commerce applications security.pdf. (检索日期：2003-10-17)
18 James B. D. Joshi.A GENERALIZED TEMPORAL ROLE BASED ACCESS MODEL FOR DEVELOPING SECURE SYSTEMS.2003,8. https://www.cerias.purdue.edu/tools_and_resources/bibtex_archive/archive/2003-23.pdf. (检索时间：2003-10-20)
19 Ravi Sandhu, Venkata Bhamidipati. The URA97 Model for Role-Based User-Role Assignment. http://citeseer.nj.nec.com/sandhu97ura.html. (检索时间：2003-11-18)
20 赵庆松 孙玉芳 等. 基于系统先决条件的授权模型研究 . 计算机研究与发展. 2003, 40(3):406~412
21 刘道斌 白硕. 基于工作流状态的动态访问控制. 计算机研究与发展.2003,40(3).-417-421
22 http://highwire.stanford.edu/lists/freeart.dtl.(检索时间：2003-11-18)
23 李爽.数字图书馆资源的访问控制问题.情报科学.2002,20(12):1292~1294,1297

作者简介

李春旺 1966 年生，副教授，博士研究生。
陈 瑾 1966 年生，馆员。