一种适用于电子病历系统的使用控制模型

王 莹 陈伟鹤 鞠时光

(江苏大学计算机科学与通信工程学院 镇江 212013)

摘 要 通过分析采用传统访问控制的电子病历系统存在的不足,提出了一种使用新一代访问控制模型——使用控制模型(UCON,usage control)思想的电子病历系统,并给出了具体的访问控制策略及其形式化描述。

关键词 电子病历,使用控制,形式化描述

中图法分类号 TP309 **文献标识码** A

Application of UCON Model on Electronic Medical Record

WANG Ying CHEN Wei-he JU Shi-guang

(College of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract By analyzing the disadvantages existing in EMR (Electronic Medical Record) system, this paper applied the idea of the next generation access control model usage control(UCON) model to the EMR system, and access control rules and formal description were given.

Keywords Electronic medical record, Usage control, Formal description

1 引言

电子病历(EMR, Electronic Medical Record),美国国立医学研究所将其定义为:基于一个特定系统的电子化病人记录,该系统提供给用户访问完整准确的数据、警示、提示和临床决策支持系统的能力[1]。电子病历除了具有传统纸质病历所有的功能外,还给我们带来了传统纸质病历所不具有的很多优点。比如,加快了医疗信息传递共享,提高了工作效率;能够辅助医生做出判断,提高医疗工作质量;能够加强环节管理,方便病人信息的异地共享等。

美国联邦政府于2003年4月14日生效的HIPAA法案(Health Insurance Portability and Accountability Act)即"健康保险流通和责任法案"[2]无疑为很多国家的电子病历隐私权立法提供了参考,值得我们借鉴。我国《2006－2020年国家信息化发展战略》也明确提出要建设电子病历系统,出台了《中华人民共和国电子签名法》等相关法律来加强电子病历的安全性保证。而访问控制是一种重要的信息安全技术,是通过某种途径显式地准许或限制主体对客体访问能力及范围的一种方法,通过限制对关键资源的访问,达到防止非法用户侵入或者合法用户的不慎操作造成破坏的目的[3]。在针对访问控制的研究中产生了各种不同的访问控制模型,如DAC, MAC, BLP, RBAC等。

当前,美国国家医疗数据中心和加州大学联合建立了电子病历信息服务实验系统(PCASSO, Patient Centered Access to Secure Systems Online)。在实验系统安全机制方面,PCASSO采用了基于敏感级的保护方法,即根据病历公开后可能对病人生活造成社会影响的程度,按升序将其划分成"低"、"标准"、"拒绝"、"监护人拒绝"和"病人拒绝"五级,信息只允许从低向高或是在相同的敏感级之间流动。这种安全控制方法在本质上等同于BLP模型,其优点在于能够有效地保护信息的机密性,不足在于缺乏完整性控制,授权管理复杂。欧洲的同类系统,如德国不来梅大学和英国南斯塔福德郡建立的实验系统,则采用了基于角色的访问控制。这种安全控制方法的优点在于授权管理灵活,不足之处在于缺乏有效的角色配置工程化和角色动态转换机制。此外,由于没有对病历的敏感级进行划分,不利于在较细的粒度上实现访问控制。

针对上述问题,本文应用使用控制的ABC模型(UCONABC)于电子病历系统来防范非授权用户获取电子病历资料,同时还能有效地控制合法用户对电子病历的使用,并且可以随时授予用户权限,为解决电子病历使用过程中的安全性和可控性提供了一种有效的技术手段和解决方案。

2 UCONABC模型

2004年,Jaehong Park和Ravi Sandu提出了新一代的访问控制模型UCON[4];2005年Xinwen Zhang和Ravi Sandu完成了以时序逻辑为形式化工具对UCON的描述工作[5]。UCON对传统的存取控制进行了扩展,定义了授权(Authorization)、职责(Obligation)和条件(Condition)3个决定性因素,同时提出了存取控制的连续性(Continuity)和可变属性(Mutability)两个重要概念。

到稿日期:2009-12-18 返修日期:2010-03-18 本文受国家自然科学基金(60603041,60773049),省教育厅自然科学基金(09KJB520003),江苏大学高级人才启动基金(07JDG031)资助。

王 莹(1986－),女,硕士生,主要研究方向为访问控制,E-mail:wangying15693@sohu.com;陈伟鹤(1976－),男,博士,副教授,主要研究方向为信息安全、数据库理论;鞠时光(1955－),男,教授,博士生导师,主要研究方向为数据库理论、信息安全、计算机图形学。

可变性属性是指主客体属性不再只能被管理员进行修改，它可以作为主体对客体的访问行为结果而改变。存取控制的连续性是指在整个主体访问客体的过程中，主体的访问权利被重复检查，当权利被撤消时，系统可及时终止其访问行为。这两个重要特性是开放式网络环境中访问控制所必不可少的。

一次授权行为的执行要综合考虑主客体属性及3大决策因素：授权(A)、义务(B)和条件(C)。其中，义务是指主体获得访问客体权利所必须完成的操作。条件是指主体获得对客体行使权利前或过程中所必须满足的系统或执行环境的约束条件。根据这3大决策因素和属性更新的时机，ABC模型可细化为16种基础模型[4]。通过对这些基础模型进行组合，使用控制模型能够实现不同的访问控制机制，以满足不同的应用系统需求，如表1所列。

表1 UCON基础模型

	0(属性不可变)	1(属性在访问前更新)	2(属性在访问过程中更新)	3(属性在访问后更新)
预先授权 preA	Y	Y	N	Y
过程授权 onA	Y	Y	Y	Y
预先责任 preB	Y	Y	N	Y
过程责任 onB	Y	Y	Y	Y
预先条件 preC	Y	N	N	N
过程条件 onC	Y	N	N	N

对于实际不存在的情况，表中标记为N，存在的表决为Y。例如用户在购买一本书提交订单时，先判断账户这个客体的余额属性是不是满足要求(余额大于书的价格)，如果满足就执行订单提交并修改账户余额。这就是预先授权，属性访问后改变的例子，如UCONpreA3。正因为这种随意的组合和其本身的新特性，体现了UCONABC模型的灵活性和动态性。UCONABC模型的各种组合，不仅可以而且超越了诸如自主访问控制模型(DAC)、强制访问控制模型(MAC)、基于角色访问控制模型(RBAC)、信任管理模型(TM)、数字版权管理模型(DRM)等传统模型的实现，因此UCONABC模型能满足现今数字资源访问的需要，在现实生活中的应用将更广泛，是新一代的访问控制模型，是目前最为适合全面解决Web系统访问控制问题的技术方法[4-7]。

3 基于UCONABC模型的电子病历系统

电子病历的安全性问题最直接的是电子病历的真实性和准确性遭到质疑，下面我们将从电子病历的保存、修改、隐私保护这3方面出现的问题进行研究和改进[8]。

3.1 电子病历的保存

惠州市某医院2006年5月至2008年2月的病历档案填写存在如下部分问题：身份证号码、住址、单位、联系方式等基本信息不填或不详；医生漏签名；疾病诊断太笼统等一些问题。

(1)问题分析与UCON模型选择

为了保证电子病历信息的完整性，我们可以限制在病历提交保存时必须保证病人身份证号码、联系方式、主治医生姓名等必要信息的填写[9]。此为典型的UCONpreB3模型(即在授权前判断是否完成其义务操作，执行结束后改变此病人病历客体的相关属性)。

(2)约束条件

当医生此主体提交保存填写的病历时，必须保证必填项的填写，否则撤销主体操作权限。

(3)形式化描述

S，主体

O，客体

R，权限

idCard：o －＞ idNum，病人身份证号

contact：o －＞ con，病人联系方式

docName：o－＞ N，主治医生姓名

ATT(o)：{idCard，con，docName }

OBS，代表职责主体

OBO，职责客体

OB，职责操作

preB，表示预先职责的判定

preOBL，表示预先的各个职责内容

preFullfilled：OBS×OBO×OB －＞{true，false}

getPreOBL：S×O×R －＞ 2preOBL，为执行请求选择的预先职责的函数

$preB(s,o,r)=preFulfilled(s,idCard(o),notNull) \wedge preFulfilled(s,contact(o),notNull) \wedge preFulfilled(s,docName(o),notNull)$

$preB(s,o,r)=true \text{ if } getPreOBL(s,o,r)=\emptyset$

预先职责为病历客体的身份证号码、联系方式、主治医生姓名都已经填写，且该做的义务都已做。

Allowed(s，o，r) ＝＞preB(s，o，r)，请求操作的允许取决于预先义务判断函数返回的值，真即允许请求执行，假即不允许。

postUpdate(ATT(o))，请求操作执行完成以后修改病历客体属性。

3.2 电子病历的修改

某一纠纷案中被告不承认原告给出的《医疗事故鉴定报告》，其原因就在于该医学会鉴定所采信的病历资料都是用电子文本生成打印出来的，这些通用软件具有可修改性，且修改后无法查询原始记录[8]。

(1)问题分析与UCON模型选择

为保证电子病历信息的安全性，我们规定各级医生有不同的修改权限。一旦下级医生将病历提交于上级医生，下级医生即丧失了对病历的修改权限，而上级医生可以修改下级医生的记录，但任何改动内容以及改动时间、修改人都是被准确记录在日志里面的，以方便查询。可以采用模型UCONpreA3。

(2)约束条件

病历的修改只能由病历提交者的同一部门的上级进行修改，并形成日志。

(3)形式化描述

ROLE，一组无序的角色集合

DEPARTMENT，一组部门的集合

docid，识别医生的唯一标识

DOC:S －> docid,一一对应

sRole:S －> 2ROLE,医生对应的角色

sDep:S －> DEPARTMENT,医生和部门之间的映射关系

oDocid:O －> docid,每一张病历上都对应一个主治医生

allowed(s,o,r) => sDep(s)=sDep(DOC(oDocid(o))) ∧ sRole(s)> sRole(DOC(oDocid(o)))

如果访问主体和访问客体病历上的主治医生是同一部门的医生,而且访问权限在主治医生之上,那么此次访问修改操作将被允许。

postUpdate(ATT(o)),请求操作执行完成以后修改病历客体属性。与上面不同的地方是,对于病历的修改系统会生成日志文件,记录本次的修改。把此日志文件称为衍生客体(derivative object)。在 UCON 中,把由对原始客体执行的操作而产生的另一客体叫做衍生客体[4],它也有它自己的属性并与其它客体之间有联系。所以这些日志也是系统安全保护的客体之一,我们规定任何日志文件一旦生成就只能往里面添加对病历的修改记录,而不能修改或删除任何日志记录。

3.3 隐私保护

在局域网电子病历数据库模式下,数据库里所有的电子病历将在网络上有条件、有偿地让潜在的客户检索、查阅。人们担心的事也就随之发生了,这样医疗保险的推销员也可能查到你的病历资料和电话号码,从而不厌其烦地向你推销一种疾病保险;疾病慈善基金的募捐人也可能知道你的信息然后游说你;你的老板也能通过电子病历数据库得知你的身体状况,从而以身体不适合为由解雇你。那么哪些人可以调用患者的电子健康记录呢?

(1)问题分析与 UCON 模型选择

HIPAA 法案对那些非治疗目的而需要使用患者电子健康信息的机构设置了一个"最低必须"要求的限制,这是一种权限的划分,这些机构的员工只能看到他们权限等级所允许看到的信息,以保证他们完成自己权限范围内的工作。这是一种基于权限的访问控制,我们采用模型 UCONpreA0。

(2)约束条件

我们规定护士只可以查看病人现在的病症和用药情况;病人的主治医生及医生所在部门的上级医生能查看其所治病人的病历详细信息;普通医生不能查看到病人的姓名和治疗方案等敏感信息;一般医院以外的个人或机构只能看到病症和病史,不能得知病人和医生等其它信息。

由于病历系统自身的特殊性,当我们把某一病历看成是一客体时,其客体的属性(姓名,身份证号码,主治医师,医疗方案等)又具有不同的密级,此时我们可以把病历中的每个属性看成是一个客体。对应于本病历系统我们角色和权限控制如表 2、表 3 所列。

表 2 系统角色等级

角色	等级
普通访客	1
护士,实习生	2
普通医生	3
某病历的主治医生或其直属上级医生,所属病人	4
法律追究部门	5

表 3 病历属性密级

病历属性	密级
病种名称	1
用药,病症	2
病史	3
主治医师,姓名,身份证,联系方式,治疗方案等隐私信息	4
绝密,不予公布	5

我们规定角色可以访问密级小于等于其角色等级的病历属性,如角色 5,可以查看所有信息,包括绝密的病历属性;角色 3 可以查看到密级为 3、2、1 的所有属性。但有一个问题,普通医生在注册时等级都为 3,当他访问普通病历时的访问权限是 3,而当他访问到自己病人病历时他的访问权限就相当于等级为 4 的权限了,这就涉及到了动态授权的问题。

(3)形式化描述

S,主体

Oi,病历中的各个属性作为访问客体

Num:病历中所有属性的个数

ROLE,访问者的角色

secret,病历的各个属性的密级

Sec:oi －> secret

Role:s －> ROLE

docid,识别医生的唯一标识

DOC:S －> docid,主体和医生唯一标识一一对应

sRole:S －> 2ROLE,医生对应的角色

Odoc:o －> docid,主治医生的唯一 id 号

$$\text{allowed}(s,o,r) => \sum_{i=1}^{\text{NumSec}(oi)} \sum_{1} \text{Role}(s) \geqslant \text{Sec}(oi) \text{ if}$$

Role(s)=3 and DOC(s)=Odoc (o) then Role(s)+1

依次显示属性客体的密级比访问者权限级别低的所有病历属性,当访问者为医生而且访问的病历自己为主治医生时,其访问权限临时加一级。

结束语 UCON 是近年来计算机安全领域一个新的研究方向,但 UCON 技术在国内的研究还处于起步阶段。本文设计的基于使用控制模型的电子病历系统能够根据分布式开放网络环境中的属性信息,基于 authorization,obligation 和 condition 3 种决定策略来检查访问控制决策,实施动态的 Web 服务访问控制。它吸收了 MAC 的安全等级保护机制,预先制定了医生和病历各属性项的安全级别,增强了 RBAC 的可管理性和易维护性,按角色把医生分为不同级别,还引入了主客体的非安全等级类别(如医生所在部门),它将不同的策略综合应用,对于不同的功能模块选择不同的访问控制模型,形成了综合的访问控制策略,可以获得更灵活、更好、更安全的系统保护。但电子病历系统的安全性问题不仅仅是访问控制问题,还涉及到比如数字签名、病历或登录账号信息的加密解密、数据传送管理协议、黑客、病毒等防御技术的使用。限于本文的篇幅没有列出医疗病历系统中诸如账号登录、复制等访问控制问题,另外将在后续工作中使用 Petri 网实现对应用于病历系统的 UCON 使用控制模型进行安全策略的分析和验证。

参 考 文 献

[1] Institute of Medicine. The Computer-Based Patient Record[M]. National Academy Press,1997:55-56

[2] United States Department of Health and Human Services. Health Insurance Portability and Accountability Act of 1996 [S/OL]. (1996208221) http://aspe.hhs.gov/admnsimp/pl104191.html#261

[3] 沈海波，洪帆. 访问控制模型研究综述[J]. 计算机应用研究，2005(6)：9-11

[4] Jaehong P, Ravi S. The UCONABC Usage Control Model [J]. ACM Transactions on Information and System Security (TISSEC), 2004(10): 1-47

[5] Zhang Xin-wen, Parisi-Presicce, Ravi S. Formal model and policy specification of usage control [J]. ACM Transactions on information and System Security, 2005, 8(4): 351-387

[6] Osborn S, Sandhu R, Munawer Q. Configuring Role-based Access Control to Enforce Mandatory and Discretionary Access Control Policies [J]. ACM Transactions on Information and Systems Security(TISSEC), 2000, 3(2): 85-106

[7] Pretschner A, Hilty M, Basin D. Distributed usage control [J]. Communications of the ACM, 2006, 49(9): 39-44

[8] 黄云，周敏. 从电子病历的发展历程谈医疗安全管理中的风险管理[J]. 中国医院管理，2007，27(10)

[9] 林彩霞. 电子病历档案首页存在的问题及改进意见[R]. 广东档案业务探讨，2009

(上接第 159 页)

对于主题 11(见图 9)，可以看出主要关注的是阎崇年被打事件的相关讨论，并在 8 个时间片内都保持了相当高的热度，而且从历史文化相关的主题中也有一定量的阎崇年被打的帖子。可以分析出关于此事件的讨论有往历史文化方面扩展的趋势，发现该主题讨论的不同侧面。根据主题 11 的主题在 8 个时间片内的差异变化柱形图，可以看出其十分稳定，在 2008 年 11 月到 12 月之间是个热点话题。

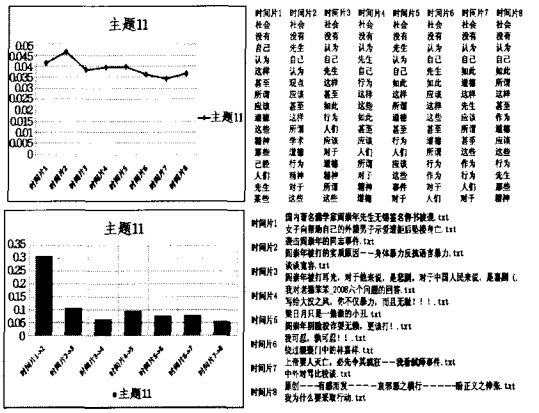

图 9 主题 11 主题演化趋势图

结束语 本文首先分析并定义了主题演化的概念和蕴含在其中的演化模式，然后建立并实现了在线 LDA 模型，最后通过实验表明此模型在主题演化分析上有着良好的效果，能够发现主题传播过程中的“主题遗传”和“主题变异”现象，基于更深层次的文本语义分析的概率主题模型具有坚实的理论基础，更加客观和准确。但本文只考虑了文本的内容和时间属性，没有考虑文本的作者、出处、链接、回帖等结构属性信息。如何把这些属性信息融合进去，改进模型的效果；如何采用基于云的 Gibbs 算法改进模型的效率，是我们下一步要努力的方向。

参 考 文 献

[1] CaliforniaAllan J, Carbonell J, Doddington G, et al. Topic detection and tracking pilot study: Final report[C]// Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop. February 1998

[2] The 2002 Topic Detection and Tracking (TDT2002) Task Definition and Evaluation Plan[EB/OL]. ftp://jaguar. ncsl. nist. gov//tdt/tdt2002/evalplans/TDT02. Eval. Plan. v1. 1. pdf

[3] Gruhl D, Guha R V, Liben-Nowell D, et al. Information diffusion through blogspace[C]// WWW. 2004: 491-501

[4] Zhou Ding, Ji Xiang, Zha Hongyuan, et al. Topic evolution and social interactions: how authors effect research[C]// CIKM. 2006: 248-257

[5] Mei Qiaozhu, Zhai Chengxiang. Discovering Evolutionary Theme Patterns from Text-An Exploration of Temporal Text Mining [C]// Proceedings of KDD'05. 2005

[6] Mei Qiaozhu, Zhai Chengxiang. A Mixture Model for Contextual Text Mining[C]// Proceedings of KDD'06. 2006

[7] 于满泉，骆卫华，许洪波，等. 话题识别与跟踪中的层次化话题识别技术研究[J]. 计算机研究与发展，2006，43(3)：489-495

[8] 王永恒，贾焰，杨树强. 面向汉语短文的话题识别系统研究[C]// 第 21 届全国数据库年会(NDBC2004). 计算机科学，31(10 增刊)

[9] 张立，刘云. 网络舆论传播的无标度特性及其衰减模型的研究[J]. 北京交通大学学报，2008(4)

[10] 胡勇，张羽斌，王祯学，等. 网络舆论形成过程中意见领袖形成模型研究[J]. 四川大学学报：自然科学版，2008(4)

[11] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation[J]. Journal of Machine Learning Research, 2003(3): 993-1022

[12] Blei D, Lafferty J. A correlated topic model of science[J]. Annals of Applied Statistics. 1: 1 17-35

[13] Blei D, Lafferty J. Dynamic topic models[C]// Proceedings of the 23rd International Conference on Machine Learning. 2006

[14] Wang C, Blei D, Heckerman D. Continuous time dynamic topic models[C]// Uncertainty in Artificial Intelligence(UAI). 2008

[15] Rosen-Zvi M, Griffiths T, Steyvers M, et al. The author-topic model for authors and documents[C]// Proceedings of the 20th International Conference on Uncertainty in AI. July 2000

[16] McCallum A, Wang A, Corrada-Emmanuel. Topic and role discovery in social networks[C]// Proceedings of 19th Joint Conference on Artificial Intelligence. 2005

[17] Alsumaitm L, Barbara D, Domeniconi C. On-line LDA: Adaptive Topic Models for Mining Text Streams with Applications to Topic Detection and Tracking[C]// IEEE International Conference on Data Mining. 2008: 3-12

[18] 石晶，胡明，石鑫，等. 基于 PLSA 模型的文本分割[J]. 计算机研究与发展，2007(2)

[19] 石晶，胡明，石鑫，等. 基于 LDA 模型的文本分割[J]. 计算机学报，2008(10)

[20] 俞辉. 基于 LSA 和 pLSA 的多文档自动文摘[J]. 计算机工程与科学，2009，31(9)

[21] 俞辉. 基于 PLSA 模型的 Web 用户聚类算法研究[J]. 计算机工程与科学，2008，30(7)

[22] imdict-chinese-analyzer[EB/OL]. http://code. google. com/p/imdict-chinese-analyzer/

[23] Lavrenko V, Allan J, DeGuzman E, et al. Relevance models for topic detection and tracking[C]// Proceedings of the Second International Conference on Human Language Technology Research. March 2002: 24-27

[24] NIPS dataset[EB/OL]. http://nips. djvuzone. org/txt. html

[25] Griffiths T L, Steyvers M. Finding scientific topics[C]// Proceeding of the National Academy of Sciences. 2004: 5228-5235