

# 云计算环境图书档案管理系统用户模型研究

李春杰,王晓明,张龙昌

(渤海大学 信息科学与技术学院,辽宁 锦州 121013)

**摘要:**为了提高云计算平台下图书档案管理系统的质量,使其能够更大程度地为用户服务,将用户、终端、网络、模态联系在一起,实现“以用户为中心”的服务理念,文中对云计算环境下图书档案系统的用户模型进行了研究和分析。首先对已有的研究成果进行了分析和研究;其次对用户进行分析和研究,包括用户、终端、网络三者,用户的行为,用户的偏好;再次提出云环境下的用户模型,包括用户基本信息模型、用户行为模型、用户偏好模型;接着探讨一些关键技术问题,如用户模型的初始化与完善等;最后进行了总结。

**关键词:**云;用户模型;系统;图书馆;档案

中图分类号:TP302.1

文献标识码:A

文章编号:1673-629X(2015)05-0233-05

doi:10.3969/j.issn.1673-629X.2015.05.054

## Research on User Model of Cloud Computing Library and Archive Management System

LI Chun-jie, WANG Xiao-ming, ZHANG Long-chang

(College of Information Science and Technology, Bohai University, Jinzhou 121013, China)

**Abstract:** In order to improve the quality of archives management system under cloud computing platform, so that it can be more close to customer service, to achieve the "user centric" service concept, study and analyze the user model of the book archive system in cloud computing environment. Firstly research and analyze the existing research results. Then analyze the user including the user, terminal, network, the user behavior and the user preferences. Next put forward the user model under cloud computing, including the basic user information model, the user behavior model and the user preference model. And then analyze some key techniques, such as the initialization and improvement of user model. Finally carry on the summary.

**Key words:** cloud; user model; system; library; archive

## 0 引言

用户,通常指图书档案数字化服务的使用者或者消费者,其使用图书档案的数字化服务以便获取自己所需要的服务,是图书档案系统的最终受益者。为了使用户更大程度上获得数字化服务带来的优势,人们往往以用户为原型加以研究、优化。

目前国内图书馆方面用户相关的主要研究成果有:文献[1],其从云的原理和概念出发,研究云模式拓展功能与服务的方法,以便实现图书馆 2.0 用户作为中心的理念;文献[2]探讨云数字图书馆环境下保护用户隐私的方法和对策;文献[3]主要针对用户的服务模式和满意度,将云计算与高校图书馆技术相结

合,改善该环境下的用户服务模式,以便提高满意度;文献[4]着眼于站点实现方法和优化模型,通过关注用户访问的时间、访问的路径,利用 Web 的技术和方法挖掘用户的访问信息,以便提高信息获取和访问的效率;文献[5]从用户的需求出发,探索云在信息服务的应用,分析云图书馆中个性化服务的可行性与必要性,再提出个性化服务的模型,将云理念引进到图书馆中。

目前国内档案方面用户相关的主要研究成果有:文献[6]以技术接受模型作依据,提出用户接受模型,并以 SEM 和问卷的方式进行研究<sup>[6]</sup>;文献[7]分析用户感知价值导向下的档案馆核心能力的意义及价值,

收稿日期:2014-07-06

修回日期:2014-10-07

网络出版时间:2015-04-22

基金项目:教育部人文社会科学研究青年基金项目(12YJC870030);辽宁省社会科学规划基金项目(L12CTQ008);辽宁省教育科学“十二五”规划基金项目(JG12DB149)

作者简介:李春杰(1965-),女,教授,硕士生导师,研究方向为多媒体、网络、云。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150422.1008.027.html>

提出构建内容<sup>[7]</sup>;文献[8]用定量分析方法,分析档案馆用户满意的情况,提出提升满意度的手段和方法<sup>[8]</sup>。

从成果上讲,目前国内对云图书档案系统环境下用户的研究可谓不少,也各有千秋,然而依然存在缺欠,如下:

①往往割裂用户、终端、网络之间的联系,缺少对异构终端、异构网络环境下用户的研究。以三网融合为例,笔者于 2014 年 5 月 27 日 21 点 52 分以“三网融合”、“用户”、“模型”为题名精确检索,在万方找到论文 3 篇,均不是图书方面,也都不是档案方面的,已有的研究成果也不能解决异构终端和异构网络环境下的用户模型问题。

②虽然考虑了用户的偏好问题,但是却极少考虑数据的不同模态。

③部分研究成果忽视了用户模型数据的可变性,降低了模型的实用性。

为此文中首先研究用户,针对用户、终端、网络、行为和偏好进行研究;然后推出用户模型,包括用户基本信息模型、用户行为模型、用户偏好模型;再对一些关键问题如用户模型的应用、用户模型的初始化和完善进行了探讨;最后进行了总结。

## 1 用户研究

将用户、终端、网络、模态联系在一起,需要贴近复杂的环境,贴近异构终端、异构网络环境,尤其是能够符合三网融合环境的需要,而在这种环境下,用户首先通过终端,经过各种异构网络接入三网,再统一连接到图书档案管理系统,用户才能使用系统所提供的服务、海量的信息资源<sup>[9]</sup>;学术界一直研发新一代信息网络,国外的参考文献主要包括文献[10-13];同时现有技术中能够连接终端、三网和系统的主要是张宏科教授及其团队的一体化标识网,该技术通过接入标识代表终端的身份,通过交换路由标识表示终端的位置信息,通过连接标识作为服务连接和用户身份的标识,服务标识表示业务的各种服务。该技术对流媒体、话音、数据等给予支持,相关的参考文献主要包括文献[14-17];基于此进行研究。

### 1.1 用户、终端、网络的研究

为了研究用户模型,首先研究用户如何使用云平台下的图书档案管理系统,用户接入到图书档案管理系统,必须使用终端,通过网络接入。在这个过程中,终端呈现异构化、复杂化、多样化的特征。

①终端目前可能使用电视网、电信网和计算机网中的任何一网,可以通过 WLAN、WWAN、Internet、PLMN、HFC、传统 CATV、Ethernet 和 OAN 中任何一网,随着时代和科技的发展也可能使用其他网络,其网络

方面的特征呈现异构化、多样化;

②终端类型和型号都很多,其功能、用途等明显不同,造成构造的不同和复杂化;终端和网络的异构化、复杂化、多样化,造成了其效果差异化很大,主要体现在:显示效果、音效效果、网络、计算复杂度、可用空间等方面。

从显示效果的角度来看,需要注意以下几点:

①终端不同,其显示屏幕的大小不同。从这个角度来说,手机和电视的差距是最为显著的,笔记本、PC 等终端设备屏幕往往适中,显示屏幕大小的不同一方面和用户持有终端是否便利有关,另一方面关系着用户浏览信息资源时屏幕的尺寸和效果。

②终端不同,其显示所使用的分辨率不同,体现为显示时的效果明显不同,如显示时的清晰细腻程度、拍照取景和视频播放的效果等等,而且部分终端的分辨率是固定不变的,无法进行调节。

③文件的格式不同,其显示效果可能不同,例如常见的视频格式有 AVI、MPEG、DIVX、MOV、ASF、WMV、RM 等,不同的格式记录相同的视频信息,其显示效果也可能完全不同。

④显示效果的程度决定了用户视觉角度质量的高低,其对于文本、图片、图像、单一的视频是有决定性作用的,同时对多媒体信息资源的效果也是有一定影响的。

⑤显示效果的程度所带来的影响,对于不同的用户是不同的,其高度影响依赖视觉的用户,但当用户不依赖于视觉的时候影响则不大。

从音效效果方面来看,需要注意以下几点:

①终端硬件不同,其发声所用的效果不同,即使终端硬件相同其发声的效果也可能不同。例如一台笔记本电脑,如果其声卡是集成的则往往没有独立的声卡效果好,同样不同厂商所生产的声卡效果往往也有一定的差别,不同时间、不同技术所生产的声卡效果也往往不同,但音效效果在设备、设置、播放内容等完全相同的前提下是一致的。

②不同的音频文件,其效果可能不同。首先数字音频技术指标,采样率、压缩率、比特率、量化级,分别控制单位时间内波形采样的数据数量、音乐文件压缩前后大小的比值、每记录音频数据一秒钟所消耗的比特值的平均值、用多少位二进制的数字描述的声音波形;其次,不同的音频文件格式不同,其音效的效果也可能不同,常见的音频文件格式有 WAV、MP3、WMA、OGG、APE 等。

③对于部分终端来说,音效效果是可以调节的,但依然存在音效无法调节的终端。

④音频效果的程度所带来的影响,对于不同的用

户是不同的,其影响高度依赖听觉的用户,但当用户不依赖于听觉的时候影响则不大。

从网络的角度来看,目前大部分的电视网依然是单向的,也就是说存在单向和双向的问题,虽然随着时代的发展和技术的进步,单向和双向的问题最终会被解决,但是目前来说其依然是问题的一部分;网络不同登陆方式可能不同,例如如果用户借助电信网登陆和使用服务有两种模式,其一是通过以点播为基础的短信息服务,需要利用短信服务的平台来传输信息,由于其平台经常是第三方提供的,因而实时性往往较差;第二种则是通过 GPRS 网络,这种方案的缺点是覆盖率较差、传输速度低、费用高,但可以实现实时。

从计算复杂度方面来看,终端之间的差异很大,以电视网的机顶盒和电视机来说,即使运行一个非常简单的动画也会非常困难,而相同的困扰在计算机上往往不存在,换句话说计算复杂度高的资源无法在部分终端上运行,但同时部分终端上如果使用计算复杂度低的文件则也会浪费资源。从某种程度上来说,不同的终端适合使用的信息资源在内容、格式等方面是完全不同的。

从可用空间的角度看,如果可用空间的大小无法支持对应的操作,则后续的操作无法进行,同时可用空间较小也会在一定程度上降低运算速度。对于可用空间的大小,不同的用户可能会有不同的定义,部分用户可能对终端的最小可用空间有自己的要求;还有一部分人对最小可用空间没有要求,但是最小可用空间不够用的时候会有自己期望的行为和操作。

## 1.2 用户行为研究

用户的行为,指用户使用云平台下图书档案管理系统的行为,其根据用户的身份、目的、习惯、兴趣等不同而不同,用户可能是进行检索、获得、修改、保存信息资源,也可能是对图书档案管理系统的管理,如计费等;既包括其行为的种类、参数,同时还应该包括其行为所导致的结果;用户的行为受到其本身的制约,既要受到其身份的影响,又受到其兴趣和目的的影响,但最终通过行为表现出来,可以通过用户的行为获得用户的信息,并做出调整,以便更精确地认识和了解用户。

## 1.3 用户偏好研究

用户偏好,指用户在考虑服务或者商品的时候按照其自身意愿所做出的带有倾向性的符合理性的选择,是用户理性、认知和心理感受权衡后的综合结论,是用户个性化的体现。在云计算平台下的图书档案管理系统中,其可以作为个性化服务推荐的基础,主要包括以下几个方面的内容。

①由终端、网络等带来的与视频、音频、可用空间等有关的偏好信息,此偏好信息用于确认用户偏好的

信息资源的模态类信息,在此类偏好中用户往往偏爱一种或者几种终端,同时对于不同终端参数设置有不同的偏好;

②用户偏好的信息资源内容方面的信息,包括学科、方向、难度等,这种偏好主要来源于用户原有的背景,如职业、专业、知识层次等,并随着用户的发展而改变;

③其他偏好,主要包括终端、所途经的网络、所在的位置等;

④用户 ID,用以唯一地标识和区别用户。

## 2 用户模型研究

在研究用户、终端、网络、行为、偏好的基础上研究云计算平台下图书档案管理系统用户模型,其由三个子模型组成,分别是用户基本信息模型、用户行为模型和用户偏好模型。

### 2.1 用户基本信息模型

根据对用户、终端、网络的研究,建立用户基本信息模型,包括三类信息,分别是用户信息、终端信息和网络信息(详见图1)。

(1)用户信息,如用户 ID、姓名、身份、年龄、知识层次、密码、密码提示问题、组别。其中用户 ID、密码是不可以省略的,其需要在用户登陆的时候确定用户的基础信息,同时用户 ID 对于整个系统来说是唯一的,也就是说所有用户的 ID 是没有重复的,是“独一无二”的;身份、年龄、知识层次三项可以用于辅助确定用户的身份,以便在后续用户偏好模型未获得具体信息之前初步判断和获得用户的偏好信息,但鉴于部分用户的特殊性某些情况下是允许被省略的;用户信息可以通过组别确定用户的权限的确认。

(2)终端信息,包括终端 ID、终端参数等。其中终端 ID 用于确定接入和登陆到图书档案系统的终端的身份,通过检索对应于终端的数据表可以获得包括其使用者、终端的类型、终端的基本参数及范围、终端的可能配置等的信息,可以初步确定终端的信息;终端参数包括的是显示参数、音效参数、计算复杂度参数、可用空间参数,显示参数和音效参数分别是用户提供的其习惯使用的显示、音效参数的设置,而计算复杂度参数需要通过终端 ID 判断用户终端的具体计算复杂度获得,可用空间则需要在图书档案系统运行的时候通过其与终端之间的通讯获得;终端信息中终端 ID 是不可以缺省的,但终端参数是可以缺省的,在缺省的状态下,系统依然可以通过终端 ID 获得终端的类型及基础的参数范围,再根据参数范围进行其他的判断,同时由于计算复杂度参数和可用空间参数不需要用户提供,因而也可以在缺省的情况下获得。

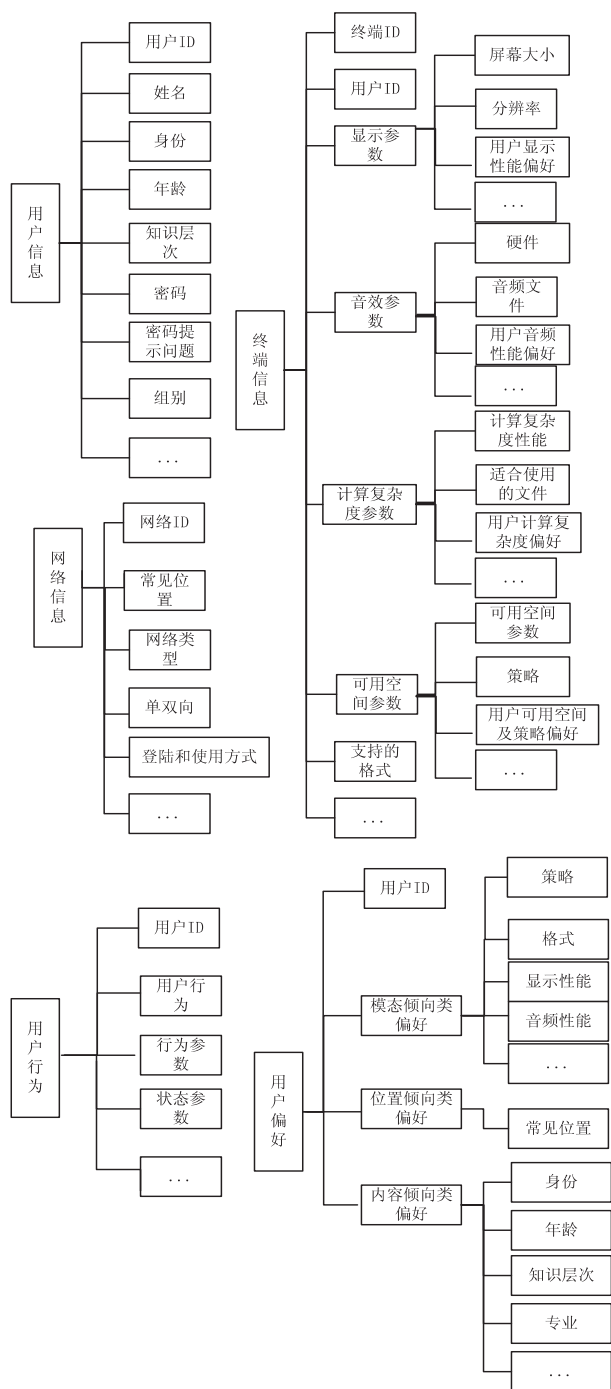


图1 用户模型

(3)网络信息,如网络类型、常见位置等。其中,电视网的单双向问题可能会带来新的问题,需要给予重点关注;常见位置,是用户通常使用终端接入和使用图书档案管理系统服务的位置;网络ID,用以唯一地标识和区别网络。

关于用户基本信息模型,需要注意的有:

(1)用户基本信息模型虽然是由三个部分组成的,但其是一个整体,可以通过数据库表示出来,其中的每一条记录都只说明的是在某一种情况下用户使用某种终端通过某种网络接入和使用图书档案管理系统的具体设置以及由此而来的各种参数,每个用户的具

体信息可以不只一条而是若干条记录的组合。

(2)用户基本信息模型的三个子模型,可以通过一个数据表体现,也可以分成用户信息表、终端信息表、网络信息表三个数据表的组合来实现。

## 2.2 用户行为模型

用户行为模型,主要用于记录用户的行为、状态等信息,其由用户ID、用户行为、行为参数、状态参数三个部分组成。其中,用户的行为用于记录用户状态变化的原因,同时可以从中获得用户基本信息模型的部分信息,同时也可以完善用户偏好模型;用户行为参数用于说明用户行为的细节,是用户行为有机的补充;状态参数,用于表明经过用户的行为之后用户的变化。

## 2.3 用户偏好模型

用户偏好模型是用户服务推荐的基础,其主要包括3个方向:其一是对所获取资源的模态倾向性,是与终端相关的,包括格式、显示性能、音频性能等,其大多数属性是范围;其二则是对信息资源内容的倾向性,主要包括用户的身份、年龄、知识层次、专业等;其三则是用户通常所在的位置区域,可以用于以就近原则对其提供服务。

# 3 关键技术分析

## 3.1 用户模型应用研究

从用户登陆并使用系统服务的角度看用户模型,用户通过终端接入计算机网、电信网和电视网三网中的一网,可以利用一体化标识网络技术连接到系统,通过终端ID确定终端的身份(终端信息模型),通过终端所在的位置、IP等可以获得对应的网络信息(网络信息模型),通过用户ID和密码登陆系统进而系统可以确定登陆的用户(用户信息模型),如果需要调用用户偏好模型可以根据用户ID在用户偏好模型中检索相应的记录则找到对应用户偏好的信息(用户偏好模型),用户使用过程中会有所操作记录下用户ID、相关操作及对应参数,则可以将信息添加到用户行为模型,根据用户ID可以从用户行为模型获取相关的参数,则可以获得对应信息据此修改和完善其他模型。

对于用户模型的应用,文中认为:

①对于信息资源的检索和使用,既涉及到用户对信息资源模态的偏好,又涉及到用户对信息资源内容的偏好,还涉及到用户所使用的终端;在检索的时候首先要根据终端信息模型和网络信息模型判断终端和网络的特性,获得支持的模态的信息,获得显示、音效、网络、计算复杂度、可用空间等信息,再根据这些信息和模态结合用户的偏好以集合的形式进行交的运算,进而缩小可以接受的信息资源的模态范围并将其作为用户即时信息资源模态的偏好,再通过用户偏好模型内

容倾向类偏好获得用户对信息资源内容的偏好。

②显示、音效、网络、计算复杂度、可用空间等的变化会带来很多不同的情况,需要形成对应的策略,该策略既要符合用户的意愿,又要符合运营商等商家的利益,同时必须符合法律法规并受到相关部门的监管。在实践的过程中,可以设置相应的阈值来划分对应的范围,根据阈值与当前值差别的组合来判断当前所处的情况,当差距过大或者达到阈值的时候可以按照对应的既定策略采取行动,如在许可的情况下重新生成对应的信息资源、只发送部分信息资源、压缩等。

③用户基本信息模型中用户信息模型、终端信息模型和网络信息模型需要相互结合,是一个有机的整体,在使用的时候相互辅助;同时用户基本信息模型、用户行为模型和用户偏好模型也是一个有机整体,其通过用户 ID 彼此关联。用户 ID 在整个模型中是唯一的,其作用是唯一地标识一个用户,以便与区别于其他用户。

### 3.2 用户模型的初始化与完善

文中认为用户模型的初始化信息可以从以下几个方面获得:用户注册的信息;利用调查研究等方式从侧面获得的用户信息;利用公众的大众化特征获得,其实是首先获得用户基础信息模型,然后再从用户基础信息模型获取数据,以基础信息模型数据为基础结合概率归类和整理,去生成用户偏好模型和用户行为模型的初始数据。

文中认为用户模型的信息完善与补充,是一个不断学习和修正的动态反馈过程,通过机器学习对参数的重新估计和修正提高了预测精度,进而提高了服务的质量,可以为个性化服务推荐打下基础。其实质是通过训练来逐渐完善用户行为模型,再根据用户行为模型逐步完善用户基础信息模型和用户偏好模型。具体的实现上,用户行为模型可以作为执行单元,学习单元通过用户行为模型提供的信息根据用户基本信息模型和用户偏好模型建立并改进知识库,执行单元格局知识库中的知识执行任务,再将执行后的信息反馈到用户行为模型作为下一步学习的资源;可以用命令序列作为最小单元描述用户行为,通过定义两个序列、两个状态之间的相似度来代表和判断行为模式之间、状态之间的相似程度;模型工作的时候,计算序列相似度来判断行为和状态的变化。

## 4 结束语

文中首先从目前国内对用户的研究出发,分析用

户,然后研究用户模型,既包括用户基本信息模型,又包括用户行为模型和用户偏好模型,最后对一些关键问题进行了分析,但没有研究海量数据量环境下信息资源的检索方法,在以后的研究中将逐步完善。

### 参考文献:

- [1] 周 舒,张岚岚. 云计算改善数字图书馆用户体验初探[J]. 图书馆学研究,2009(4):28-30.
- [2] 潘 辉. 数字图书馆用户隐私问题研究及其对云计算服务的启示[J]. 情报理论与实践,2011,34(4):44-47.
- [3] 郭红英. 云计算环境下高校图书馆用户服务模式发展研究[J]. 图书馆理论与实践,2011(2):84-85.
- [4] 卢 云. 云图书馆用户访问效益的配送路径开发和优化处理[J]. 新世纪图书馆,2011(11):72-74.
- [5] 江 涛. 面向用户需求的图书馆云平台个性化服务系统模式研究[J]. 国家图书馆学刊,2013,22(3):30-35.
- [6] 戴 畅,李 晶,谢 笑. 技术接受视角下的网络教学资源库用户使用意愿分析—以档案学教学资源库为例[J]. 现代教育技术,2012,22(10):27-32.
- [7] 程结晶,彭小芹. 用户感知价值导向的公共档案馆核心能力的建构[J]. 档案学研究,2011(2):24-28.
- [8] 刘金霞. 档案馆用户满意度定量分析—兼论“两个体系”的建设[J]. 档案学研究,2010(1):52-56.
- [9] 王晓明. 三网融合环境区域云数字图书馆构建研究[J]. 计算机技术与发展,2014,24(5):227-230.
- [10] Rangonathan K. Trustworthy pervasive computing; the hard security problems[C]//Proc of the 2nd annual conference on pervasive computing and communications workshops. Orlando, USA:IEEE,2004:117-121.
- [11] Crovella M, Kolaczky E. Graph wavelets for spatial traffic analysis[C]//Proc of INFOCOM 2003. San Francisco, USA:IEEE,2003:1848-1857.
- [12] Gkantsidis C, Rodriguez P R. Network coding for large scale content distribution[C]//Proceeding of INFOCOM. Miami, FL, USA:IEEE,2005:2235-2245.
- [13] Clark D. Making the world (of communication) a different place[J]. ACM SIGCOMM Computer Communication Review,2005,35(2):91-96.
- [14] 张宏科,苏 伟. 新网络体系基础研究——一体化网络与普适服务[J]. 电子学报,2007,35(4):593-598.
- [15] 董 平,秦雅娟,张宏科. 支持普适服务的一体化网络研究[J]. 电子学报,2007,35(4):599-606.
- [16] 杨 冬,周华春,张宏科. 基于一体化网络的普适服务研究[J]. 电子学报,2007,35(4):607-613.
- [17] 王浩学,汪斌强,于 婧,等. 一体化承载网络体系架构研究[J]. 计算机学报,2009,32(3):371-376.

# 云计算环境图书档案管理系统用户模型研究

作者：[李春杰](#)，[王晓明](#)，[张龙昌](#)，[LI Chun-jie](#)，[WANG Xiao-ming](#)，[ZHANG Long-chang](#)

作者单位：[渤海大学 信息科学与技术学院, 辽宁 锦州, 121013](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2015(5)

引用本文格式：[李春杰](#). [王晓明](#). [张龙昌](#). [LI Chun-jie](#). [WANG Xiao-ming](#). [ZHANG Long-chang](#) [云计算环境图书档案管](#)  
[理系统用户模型研究](#) [期刊论文] - [计算机技术与发展](#) 2015(5)