

基于资源整合的手机图书馆系统的设计和实现*

方 玮 张成昱 窦天芳

(清华大学图书馆 北京 100084)

【摘要】以清华大学图书馆提供移动读者服务为例,介绍基于资源整合的手机图书馆 WAP 服务系统的设计与实现,详细讲述在图书馆现有服务系统上,利用图书馆已有资源来拓展读者服务,为用户提供一个移动、综合的读者服务平台。

【关键词】WAP 手机图书馆 移动图书馆 资源整合 读者服务

【分类号】G250

The Design and Realization of a Resource Integration Based Mobile Library System

Fang Wei Zhang Chengyu Dou Tianfang

(Tsinghua University Library, Beijing 100084, China)

【Abstract】Taking the application of mobile services in Tsinghua university library as an example, the paper introduces the design and realization of a resource integration based mobile library WAP service system. Fully utilizing the resources in library to exploit reader services, the system is built upon the existing service systems, and provides an integrated mobile platform for the readers.

【Keywords】WAP Mobile phone library Mobile library Resource integration Reader service

1 引 言

随着移动通信和互联网产业的发展,许多人开始习惯通过手机来获取互联网信息资源,越来越多的互联网服务拓展到手机等移动平台,手机图书馆服务就是这种趋势下的一种应用,也是数字图书馆发展的一个方向。

目前国内外许多图书馆都提供了一种或数种面向手机用户的服务,包括短信、WAP 服务等。清华大学图书馆也积极展开了移动环境下图书馆服务的探索和实践,在 2007 年推出短信催还服务^[1],并于 2008 年开始全面实施 TWIMS (Tsinghua Wireless and Mobile Digital Library System) 项目,旨在为用户提供一种以手机为中介的数字图书馆服务方式,让读者享受到移动互联网技术和传统图书馆服务相结合带来的便利。本文介绍的手机图书馆系统是 TWIMS 项目的一部分,是针对异构电子资源和纸质馆藏资源开展 WAP 服务的探索和实践。文中将详细介绍系统的需求与设计目标、系统架构设计、系统实现等内容。

收稿日期:2009-04-27

收修改稿日期:2009-06-01

* 本文系教育部基金项目“基于无线广域网的移动数字图书馆实现和服务机制的若干关键问题”(项目编号:08JA870010)的研究成果之一。

2 系统需求与设计目标

清华大学图书馆拥有丰富的中外文图书、期刊馆藏和电子资源,并针对这些不同类型资源提供相应的 B/S 结构的 Web 服务系统,用户可通过互联网访问,本系统的设计目标是把相关服务移植到移动通信平台上。为了充分利用 Web 服务系统的硬件资源,需要在分析系统服务流程的基础上,利用“协议网关+页面分析”的机制,把图书馆的不同服务和各种资源整合在一起,构建 WAP 读者服务平台。

根据图书馆的现有条件,手机图书馆系统可以提供面向纸质馆藏资源和电子资源的服务内容,包括馆藏书目查询、个人借阅信息查询、图书预约和续借,以及常用数据库检索等。虽然本系统主要提供的是 WAP 服务,但作为 TWIMS 项目的一部分,系统还应该具有一定的通用性和拓展性,以便将来扩展图书馆手机服务的种类,如增加短信、彩信等相关服务。

手机图书馆系统依赖于移动通信平台,服务终端是手机等手持设备,这些设备种类繁多,软硬件条件差异很大。为了最大化用户群体,要尽量降低系统对服务终端的软硬件要求,使用户可以通过功能最简单的手机获取系统的各项服务。系统还应该具有完善的用户管理和个性化定制机制,提供个性化的用户服务,使用户可以根据不同终端设备的软硬件条件和实际需求控制服务内容。

3 系统架构设计

3.1 系统构建模式

本系统旨在把图书馆现有的各种资源整合起来通过 WAP 服务形式提供给读者,它直接构架在图书馆现有的 Web 服务系统之上,提供的服务内容主要分为电子资源和纸质资源两类。电子资源相关服务构架在图书馆的 MetaLib 系统上,MetaLib 是一个提供异构电子资源统一检索的系统^[2],包含一个用于个性化开发的底层模块 X-Server。X-Server 提供了许多功能 API,通过这些 API 可以比较容易地实现异构电子资源的统一检索功能。X-Server 接受 URL 形式的服务请求并返回 XML 文档,系统从 XML 文档中提取所需信息,具体过程如图 1(左)所示。纸质资源相关的服务构架在 INNOPAC 系统(图书馆馆藏目录系统)上,INNOPAC

系统提供了馆藏书目查询、个人借阅信息查询、图书预约与续借等功能,但 INNOPAC 没有开放实现这些服务的 API,因而需要按照 Web 服务的流程构造 URL 请求,然后抓取返回的 Web 页面,通过 HTML 源代码分析来提取信息,具体过程如图 1(右)所示。

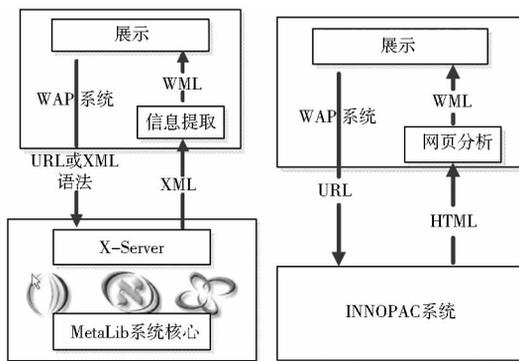


图 1 WAP 系统与图书馆现有系统的信息交互示意图

3.2 系统整体架构

系统按业务功能可分为三个基本模块:用户管理模块、电子资源检索模块和纸质资源相关服务模块。其中,用户管理模块实现用户注册、身份认证和登录等功能;电子资源检索模块完成异构电子资源和数据库的跨库检索功能;纸质资源相关服务模块实现馆藏书目查询、个人借阅信息查询、图书预约和续借等功能。

系统的各服务模块架构在不同的 Web 系统平台上,实现流程有较大差异。为了方便系统的移植和扩展,本系统使用 WebService 技术对各业务功能模块进行封装,在上层设计开发 WAP 展示层来进行交互和展示。整个系统按服务流程顺序基本可分为三层:业务逻辑层、WebService 封装层以及 WAP 展示层,系统架构如图 2 所示。

各层实现的功能如下:

(1) 业务逻辑层直接与 INNOPAC 系统、MetaLib 系统进行交互以实现具体业务功能,向这些服务系统发送服务请求和接收数据(HTML 或 XML 文档),同时解析返回的数据,从中提取所需信息。

(2) WebService 层对业务功能进行封装,生成通用功能接口,以便上层模块通过调用来传递业务流程。

(3) WAP 展示层直接面向用户,生成用户与系统交互的界面,在用户输入相关信息并提交服务请求后,

系统通过 WebService 接口向业务逻辑层发送请求并接收返回的数据,然后按照 WML 语言规范把返回的信息动态生成 WML 文档返回给用户。

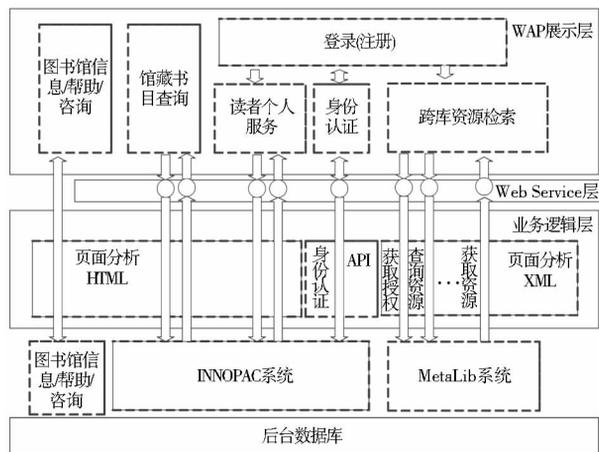


图2 WAP图书馆服务系统开发架构图

4 系统部分功能的实现

笔者开发完成了包含电子资源服务和纸质资源服务的手机图书馆系统。系统采用 Windows 系统下的 Java 开发环境,安装了 JDK1.5,并使用了集成开发工具 Eclipse。系统的 WAP 展示层使用 WML 语言和 JSP 技术。目前针对手机的网页开发语言一般有 WML 和 XHTML 两种,WML 相对简单、功能单一,XHTML 功能强大、色彩也较为丰富,考虑到图书馆服务系统功能性强、娱乐性弱的特点,笔者选择了 WML,从而避免部分较老款式手机用户受到限制。

系统运行环境采用 Windows Server 2003 操作系统、Tomcat 6.0 服务器以及 MySQL 5.0 数据库的组合,测试工具采用 Nokia 手机模拟器。

4.1 用户管理功能的实现

系统的许多功能需要用户注册并登录后才能使用,而注册用户主要分为校内和校外两大类。为了区分不同的用户类别,控制用户权限并实现用户的个性化功能定制,笔者建立了一张用户基本信息表(User_Info)用来保存用户的个人信息。用户在注册时需要选择注册类别,注册“校外用户”的必需信息只有用户名和密码,而注册“校内用户”还需增加用户真实姓名、学号/工作证号、密码(Pin 码)以供系统验证注册用户是否为真正的校内用户。在“校内用户”注册过程中,

系统通过 Java 类 pinTest 中的方法把个人信息发送给 INNOPAC 系统的身份认证 API,并通过返回的信息判断个人信息是否有效^[3]。只有通过身份认证的用户才能注册为“校内用户”,否则只能注册为“校外用户”。当注册成功后系统会把用户的用户名、密码、用户组别(group_ID)以及其他个人信息存入数据库的用户基本信息表中,其中校外用户的 group_ID 值为 1,校内用户为 2。

系统的一些服务,例如用户注册和个性化功能定制需要用户输入比较多的信息,而且只需用户在初次使用系统时进行设置。由于受到手机等移动设备的硬件条件限制,用户直接在手机上进行这些操作比较麻烦,笔者设计开发一个基于 Web 的服务系统来帮助用户进行相关操作,WAP 图书馆服务系统和 Web 服务系统通过底层数据库连接,组合成一个有机整体。这样的系统设计为用户提供了两种注册渠道,可供用户灵活选择。

用户管理模块的架构图如图 3 所示:

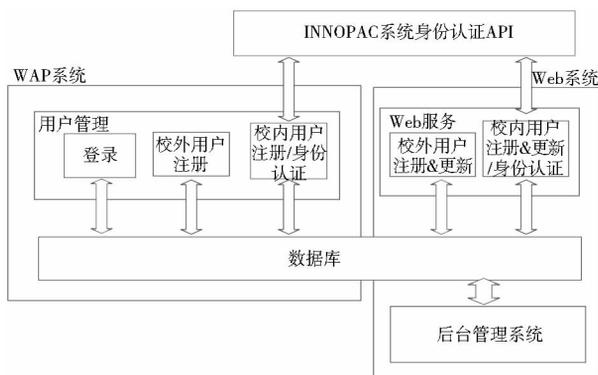


图3 用户管理模块架构图

4.2 电子资源检索功能的实现

笔者主要利用 X-Server 提供的 find 和 present 两个 API 设计了跨数据库统一检索功能。当用户从数据库列表中选择多个目标数据库进行检索时,系统按照 find API 的规则构造 URL 请求,然后提交给 X-Server,检索完毕后 X-Server 把各数据库结果集的标号返回给系统,系统再通过 present API 从结果集中获取包含各数据库检索记录信息的 XML 文档,从 XML 文档中提取所需信息并通过 WAP 展示层返回给用户。在资源检索过程中,系统还要把一些缓存数据和日志信息写入相应数据表中。电子资源检索功能模块的框架

图如图 4 所示:

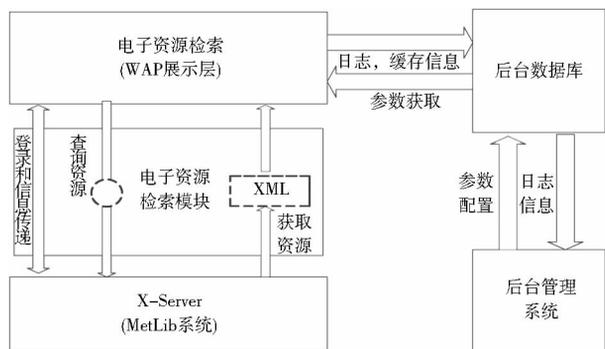


图 4 电子资源检索模块架构图

X - Server 返回的包含数据库检索记录信息的 XML 文档符合 MARC21 规范, 部分信息对照如表 1 所示:

表 1 XML 文档信息格式提取对照表

系统要提取的信息	XML 文档格式
题名	245 \$ a
第一作者	700 \$ a
第二作者	100 \$ a
出版年	DAT \$ a
出版日期	YR \$ a
来源刊	773 \$ t
原记录链接	856 \$ u
摘要	520 \$ a

XML 文档示例如下:

```

< datafield tag = "245" ind1 = " " ind2 = " " >
  < subfield code = " a " > 题名 < /subfield >
< /datafield >
< datafield tag = " DAT " ind1 = " E " ind2 = " " >
  < subfield code = " a " > 出版日期 < /subfield >
< /datafield >
< datafield tag = " 856 " ind1 = " " ind2 = " " >
  < subfield code = " u " > 源记录链接 < /subfield >
< /datafield >
< datafield tag = " 700 " ind1 = " " ind2 = " " >
  < subfield code = " a " > 作者 < /subfield >
< /datafield >
< datafield tag = " 100 " ind1 = " " ind2 = " " >
  < subfield code = " a " > 作者 < /subfield >
< /datafield >
< datafield tag = " YR " ind1 = " " ind2 = " " >
  < subfield code = " a " > 出版年 < /subfield >
< /datafield >
< datafield tag = " 773 " ind1 = " " ind2 = " " >

```

```

< subfield code = " g " > . . . < /subfield >
< subfield code = " t " > 来源 (刊物) < /subfield >
< /datafield >

```

本系统使用 Java 类库提供的 XML 处理类来解析和提取 XML 文档中的信息, 主要使用了 javax.xml.parsers 和 import org.xml.sax 类包中的方法, 系统主要按照 < datafield > 标记中 tag 参数的不同来识别 < datafield > 的下级标记 < subfield > < /subfield > 中所包含的信息类别。

4.3 纸质资源相关服务的实现

纸质资源相关服务主要包括馆藏书目查询、个人借阅信息查询、图书预约和续借等, 业务流程通过传递 URL 和提取 INNOPAC 系统返回的 HTML 信息来实现。

馆藏书目查询对所有注册和未注册用户开放。在用户提交书目查询请求后, 系统业务层首先按 INNOPAC 系统规则生成 URL 形式的字符串, 然后向 INNOPAC 服务器发送 URL 请求并接收返回的 HTML 网页, 接着通过网页分析来提取需要的书目信息并保存到缓存数据库中, 最后系统的 WAP 展示层向用户分页展示缓存数据库中的信息。本系统使用 Java 开源项目 HTML Parser 工具包来处理 HTML 网页, 可以比较高效地提取 HTML 源码中的目标信息。

纸质资源相关服务模块的架构图如图 5 所示:

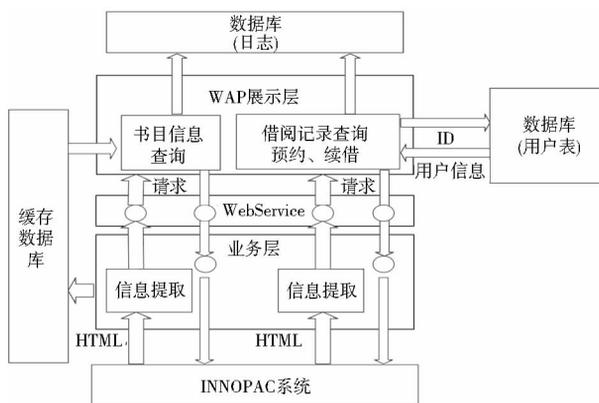


图 5 纸质资源相关服务模块架构图

个人借阅信息查询服务只对校内用户开放。校内用户登录后, 系统会自动从用户基本信息表中读取用户的全名、学号/工作证号和 Pin 码, 当用户提交个人借阅信息查询请求时, 系统把这些信息自动提交给 INNOPAC 服务器, 获取包含个人借阅信息的 HTML 网页。系统分析 HTML 网页、信息缓存以及结果显示的

过程与馆藏书目查询类似。本系统中图书的预约和续借服务业也是面向校内用户的,当用户在查询馆藏书目或个人借阅信息后,可对某一特定图书进行预约或续借操作。

4.4 系统运行

目前,本手机图书馆服务系统已初步整合了多种类型馆藏资源,具备了电子资源、纸质资源检索和相关服务功能,并已开通测试。具体实现功能如下:

(1) 用户注册:包括校内用户和校外用户注册,其中校内用户注册需要通过图书馆 INNOPAC 系统的身份认证,以确保用户的个人信息真实有效。

(2) 校内身份认证:当用户注册为校外用户后,可以通过校内身份认证功能来更改注册用户类别,认证成功即可获得校内用户权限。

(3) 电子资源检索:用户可以进行跨库统一检索,同时在多个数据库中查询电子资源,系统可以控制用户查询权限,不同类别用户可以查看的资源信息量不同。

(4) 馆藏书目查询:用户(包括注册和未注册用户)可以查询图书馆的馆藏书目信息,查询所需图书是否存在以及具体信息,当目标图书不存在时系统还会自动给出具有一定关联性的图书。

(5) 个人借阅信息查询:校内用户可以查询个人的借阅信息,包括所借图书的基本信息、借阅日期以及归还日期等。

(6) 图书预约:校内用户可以预约已借出的图书。

(7) 图书续借:校内用户可以进行图书续借操作。

(8) 图书馆信息:用户可以查看图书馆的开放时间等信息。

(9) 留言反馈:用户可以提交留言信息,以供管理员查看。

本系统是清华大学 TWIMS 项目的一部分,主要实现了 WAP 服务功能,系统还提供了面向 Web 管理和服务的功能接口,包括用户管理、电子资源/数据库信息的维护、不同类别用户资源访问权限的控制、用户的个性化定制等。

5 结 语

本系统在图书馆现有的系统平台基础上进行设计和开发,不仅实现了读者查询纸质资源及其相关服务

的功能,还提供了面向手机用户的异构电子资源统一检索服务,拓展了读者获取图书馆各种服务的渠道。这种基于资源整合的手机图书馆服务可以提高图书馆资源的利用率,也将提升图书馆的整体服务品质。

除了已实现的功能外,系统还存在一些有待完善和扩展的地方,主要包括以下方面:

(1) 完善业务功能的封装。笔者通过 WebService 对系统的业务功能进行封装,目前各模块的独立性不够,封装 API 的设计有待优化。功能明晰参数传递简单的 API 可以增加系统的重用性,方便日后系统功能的扩展。通过这种方式,只要改变最顶层的展示模块,就可以把系统移植到其他服务类型系统上,如短信服务、彩信服务等,或者把展示层语言从 WML 更换为 XHTML。

(2) 改进跨库电子资源检索中的响应时间控制机制。如何综合考虑异构电子资源之间的访问速度差异和手机网络的延时给用户带来的影响是一个值得研究的问题。

(3) 尝试通过个性化管理来实现用户在不同信息载体之间的“漫游”。如在 WAP 系统中增加全文下载并转发 E-mail 功能,以提高用户的资源获取能力和效率。

(4) 增加面向 3G 手机的图书馆服务。3G 手机与 2G 手机在很长一段时间内将并存。与 2G 相比,3G 的网速更快,3G 手机的上网功能也更为强大,可以支持更多类型资源的服务,如流媒体播放等。今后还需扩展 WAP 服务系统,兼顾 2G 手机和 3G 手机,为用户提供更全面的 service。

参考文献:

- [1] 张蓓,周虹,张成显,等. 清华大学图书馆流通短信发送服务的设计与实现[J]. 现代图书情报技术,2009(1):93-97.
- [2] 姜爱蓉,黄美君,窦天芳. 数字资源整合与信息门户建设——清华大学图书馆的探索与实践[J]. 现代图书情报技术,2006(11):2-6.
- [3] 陈武,窦天芳,林容. INNOPAC 读者 PIN 认证 API 应用实例[J]. 现代图书情报技术,2005(10):84-87.

(作者 E-mail: willfw@gmail.com)