

基于 VoiceXML 的图书馆 IVR 系统设计研究

王大伟, 顾卫平

(河北工程大学 图书馆, 河北 邯郸 056038)

[摘要]文章对基于 VoiceXML 的 IVR 系统的关键技术、系统架构模型进行了分析,并对高校图书馆 IVR 系统的主要功能做以描述,最后对部分代码进行了举例说明。

[关键词]IVR VoiceXML 系统架构

[中图分类号]G250 **[文献标识码]**A **[文章编号]**1673-9477(2011)01-0126-03

一、IVR 技术

IVR 即 Interactive Voice Response, 互动式语音应答。它是呼叫中心负责与用户自动交互的部分,是服务提供方所提供的自动式用户服务。一般来说,其所具有的主要功能有:自动应答、数字转语音、语音留言、与数据库连接等。为了方便后台人员对系统进行程序维护,它还具备图形化的操作界面。做为成熟的技术理念,IVR 系统是呼叫中心中的子系统,实际上是一个“自动的业务代表”,其具有能够准确的获取信息、修改个人记录、业务化快的特点,并且能够提供 7*24 小时的全天候服务。因此 IVR 具有广阔的发展前景,对于一个预期呼入用户数量有限或由于成本等因素暂时无法实现人工座的情况下,只应用 IVR 技术就可以满足需求了。

传统的自动交互平台始终是在公共电话网这个网络中做文章,且各公司在开发其业务语音流程时各自为政,都有不同的脚本标准,并且编写难度大。而无法实现不同平台的信息互连,这实际上成为了语音呼叫技术发展的瓶颈。特别是在固话网市场萎缩的背景下,寻求新的空间成为必须实际的目标。在这种情况下, VoiceXML 语音流程语言被提出。

二、VoiceXML 基本架构及关键技术

VoiceXML 是 XML 在 Web 语音上的扩展应用,它是在 W3C 的一次会议上被几家公司提出,在 2000 年互联网联盟批准其成为互联网语音应用的标准。通过 VoiceXML 技术,信息能够在电话网与互联网中被共享。同时,由于它同样是基于 XML 的,因此同样是一种标记语言。编程人员在利用其描述语音事件时,可以只专注语音流程本身,不用了解板卡、交换机等是目前动作的,从而从复杂的底层设计中解放了出来。目前,这种语言最新版本为 VoiceXML2.1,共定义了 43 个功能标记,具有更强的易用性和实用性。

IVR 系统通过对 VoiceXML 文档进行解析的基础上实现与用户进行语音交互,用户也通过解析的放音结果进行收听与进一步的操作选择。系统的基本架构如图 1 所示:

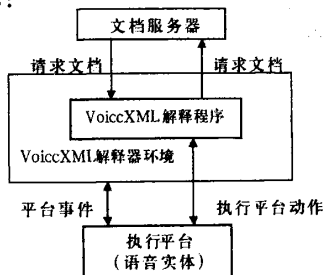


图 1 VoiceXML 系统架构图

从图 1 可以看到,本系统采用 C/S 的结构,即前端为客户机,后端为 VXML 文档服务器,中间件由文档的解析模块和解析环境组成。现将各部分介绍如下:

(1) 执行平台

执行平台是系统的前端架构,由公共电话网和程控交换机所组成,在实验环境下就是一部固定电话和与之连接的交换机。在这一模块接入用户呼叫信号后,会向用户播放一段事先录制好的欢迎词,同时会将信号通过系统信道向中间件传递,由中间件进行解析动作,在一个解析循环完成时,现向用户播放解析结果形成的语音,并等待用户的进一步选择。因此,执行平台实际上由解析器所控制。

(2) 解析器及解释环境

位于系统中间部分的解析器实际上是安装在服务器上的一段计算机程序,被称为 OpenVXI。它是在 Sourceforge 论坛上的一个开源项目,其功能就是对 VXML 文档进行解析,其解析的结果能够被 IVR 系统识别并转化为硬件动作。该程序的功能实现取决于它所提供的各个功能接口,主要有网络接口、电话接口、提示语接口,可以根据系统的设计需求对相应的接口进行编译,写入相应的函数。OpenVXI 的编译环境有 VC 和 VS 两种,目前以 VS 应用较多。

解释环境是 OpenVXI 的支持环境,它也是驻存在机器上的一段程序,它与解析器的关系相当于可视界面下 IE 浏览器与操作系统。在交互中它能监测到交换机传输的用户呼叫信号,而解析器也是通过这一环境与执行平台互连。

(3) 文档服务器

文档服务器中所存贮的是 VXML 文档,从运作原理上它与普通的 WEB 服务器没有任何区别,只是所存贮的文件格式不一样。在系统运行时,解析器根据用户选择所指向的 URL 地址,从服务器中调取相应的文档,向用户放音并等待进一步选择。文档服务器与解析器通过网络协议连接,因此访问速度受到网速限制,因此,一般情况下,将其布置在原有的局域网内即可。

以上是 IVR 系统的各组成部分,在运行时,用户通过拨打服务号码与系统取得联系,执行平台在将电话接入,并播放一段欢迎词,同时将呼叫信号传输至中间件即解析器。解析器根据用户对根文档的选择生成 URL 地址,然后向后台服务器调取相关文档。并对文档进行解析,根据解析结果向用户播放合成语音,从而完成一个交互循环。

基于 VoiceXML 的 IVR 系统应用到了语音识别、语音合成等技术来保证系统与用户的交互过程。

(一) 语音识别

语音识别是使机器能够“听”懂自然人的声音,并将声音做为控制信号的技术。VoiceXML 中的语音识别是使系统能够对 DTMF 和自然语音进行识别,在相

关标记语法中也提供了这样的功能标签。能够识别自然语音信号是语音交互系统的目标,因为这样能够为用户提供最大的自由度。但限于目前对于非特定人的识别还受到说话人的发音习惯和环境噪声等因素影响,识别率非常低。因此商用的 IVR 系统都是提供对 DTMF 即按键输入的识别,通过高频和低频的按键组合保证了 100% 的识别率, VoiceXML 系统同样利用了这一成熟的技术。

(二) 语音合成

语音合成较之语音识别技术要成熟一些,一些商用的语音库已经做的非常好,发音自然,如科大讯飞的电话银行系统,已经足以保证用户良好的使用体验。与播放事先录制好的声音相比,从 VXML 文档到合成语音的播放为系统提供了很大的自由度,这保证了 IVR 在人机交互时大部分时间是在对所需要播放的文字进行实时的合成播放输出。

(三) OpenVXI 解析器

OpenVXI 解析器其作用相当于 IE 浏览器,所不同的是 IE 所解析的是 HTML 文档,而 OpenVXI 所解析的是 VXML 文档,并应用于语音平台。OpenVXI 的功能就是将 VXML 文档生成一个对象树,也就是所谓的 DOM 解析,然后根据文档所要实现的交互功能,解析器负责调用 IVR 系统的各个部分,为了实现解析器的功能,需要对它所提供的一系列调用虚拟接口进行编译,写入相应的函数,下面将主要接口介绍如下:

1. 电话接口 (Rec API)

电话接口模块负责所有的关于电话事件的服务,有转移呼叫、挂断事件、电话转接等。

2. 提示语接口 (Prompt API)

这一接口提供文本到语音的接口,包括从 VXML 文档中合成的语音和事先录制好的文件,目前 OpenVXI 只替代对 WAV 后缀的语音文件播放。

3. 识别接口 (Rec API)

识别接口提供语音识别和接入用户的电话呼叫服务,在本系统中只对电话按键音进行识别,而不提供对自然人的发音进行识别。

三、语音平台架构设计

本系统拟利用图书馆现有的服务器做为 WEB 服务器,然后通过网络连接将服务器与解析器以及语音合成服务器等相连。最后再将中间件通过程控交换机与固定电话相连接,组成一个基本的硬件架构。如图 2 所示:

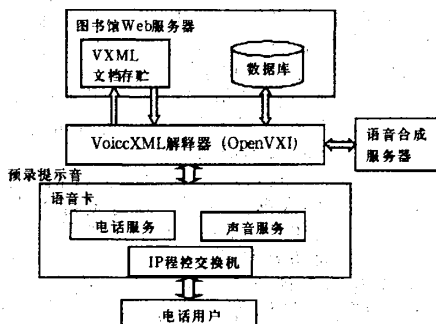


图 2. VoiceXML 语音交互系统设计结构图

系统与用户的交互过程是这样的。在交换机中设置有本系统的接入号码,用户通过拨打这个号码与程控交换机上的 E1 口相连接,交换机在监听到用户呼叫后接入电话并向用户播放一段预先录制好的欢迎词,进入呼叫控制模式。然后,向解析服务器发送用户呼入信号,解析器在接到信号后,会产生第一个 HTTP 的

地址,并根据这一地址向后台的服务器调用 VXML 文档,接下来解析服务器会对这个文档进行解析,并根据需要调用语音合成服务器进行语音合成,并将返回的结果向用户播放,同时等待用户选择,然后根据选择重复刚才的解析动作。在整个交互过程中,始终是由解析器通过对存储在 WEB 服务器中 VXML 文档进行解析来与用户实现人机对话的。此外, VoiceXML 还提供了容错机制,如用户按了错误的按键、长时间无人响应等,系统会产生一个事件,并将其抛出。

以上是对语音平台架构的一个设计思路。但由于条件所限,在实验中,交换机与语音卡是利用插在计算机上的模拟卡进行模拟。TTS 服务器和 OpenVXI 解析程序被集成在这台计算机上,固定电话做为用户端口连接在模拟卡上。用 oracle 做为数据库,计算机的操作系统是 Windows2000 Server,在用户呼入被监测到时,解析器会自动进行文档的调用处理。

四、高校图书馆语音服务流程

在经过对国内几家公共图书馆的电话调查,如国家图书馆、厦门市图书馆、北京市公共图书馆等,发现图书馆语音服务功能主要是馆情介绍、办证信息、读者卡信息以及读者图书续借、挂失服务、密码修改,特别是国家图书馆的呼叫中心功能非常全。

但考虑到高校图书馆读者群情况有所不同,相比公共图书馆读者群体狭窄,以教师和学生为主。因此,笔者通过对读者问卷调查的方法进行语音系统功能评估。以知已知彼网(www.zhijizhibi.com)为平台发送网络调查问卷 50 份,对图书馆到馆读者现场发放问卷 150 份,共回收有效答卷 183 份。经过对各个服务功能的选择比例进行统计得出:选择图书续借和图书证挂失服务的读者占到 80% 以上,而选择馆情介绍服务的则只有 23%。考虑到图书馆在使用“校园一卡通”之后,读者的超期欠费可以直接从卡上扣除,而丢失一卡通对于读者来说损失更大。因此,去掉一些华而不实的功能后,“图书证挂失”和“图书续借”就是语音服务平台的主要功能,接下来的工作是编写相关代码。

五、代码设计举例

根据 VoiceXML 语言的设计规范,笔者编写了接入电话和图书馆挂失两个 VXML 文档。举例如下:

```

<? xml version="1.0"? >
< vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" >
< menu dtmf="true" >
< prompt >
用户您好,请选择图书馆服务:
1 借阅证挂失 2 图书续借
< choice next="weather.vxml" > 借阅证挂失 </choice >
< choice next="stocks.vxml" > 图书续借 </choice >
</prompt >
</menu >
</vxml >
  
```

在这个文档中为用户提供了 < menu > 输入形式,当 dtmf="true" 时,前两个 < choice > 自动分配为电话按键的“1”和“2”按键,用户根据提示按键后,系统将转向另一个 VoiceXML 脚本文件,目前可以转向定义借阅证挂失的对话流程文档。实现系统自动挂失借阅证的脚本是非常复杂的,同时在系统环境配置实现方面还需要进一步实践研究。

(下转封三)

(三) 体育场地及设施

体育场地、设施建设少,质量又无法保证,又没有专业人的管理及维修,无法满足体育锻炼的需要。据调查很多学校的体育场地都是学校的自用,几乎不对外开放。如果想去可能要交一定的费用。私人经营又要收费,还有的农村体育设施常常被占用、破坏,很多设施陈旧不堪,存在安全隐患,由此算来,真正农民群众使用的体育场地少而又少。可见,农村体育场地设施匮乏的现状 & 城乡体育资源的不平衡性,在短时间内从根本上是难以解决。

三、邯郸市城乡体育产业发展的建设

建议加大对城乡的经济投入,提高农民的体育文化素质,建立体育锻炼场所,因地制宜发展特色体育,加大宣传和经营的投入,创编农村体育项目建立配备的体育场所。

(一) 摆正城乡体育的地位,改变农民对体育的价值观

让农民对体育有所了解有所认识,激发他们对体育参与意识及体育的经济投入,使农村体育持续、健康的发展,加强农村经济建设的体制改革,把发展农村经济,增加农民收入作为重点,农民收入提高,物质上去了,才有时间、精力了解体育、认识体育的重要性。

(二) 加强政府部门的职能,提高领导认识

要想发展城乡体育事业是各级人民政府的基本职责。保证一定的体育经费和基本体育建设经费。把提高领导认识、争取领导认识为重要任务来抓。邀请领导参与体育活动、体育竞赛和代表队,争取领导对体育的支持和关心。让领导感受体育的魅力。

(三) 发挥体育志愿者的作用

加快培训农村体育社会指导员队伍的建设,农

村体育活动的开展需要建立一支有一定的组和技术水平的社会体育指导员队伍,引导农民科学锻炼,可借助中小学力量,以体育教师和学生体育骨干为一批体育志愿者,为农民健身提供指导和帮助,同时,乡镇要注意发现和培养体育积极分子以他们为骨干,把农民组织起来,使农民参与到体育活动中去,如果条件允许可派一些有较高素质,一定能力的人员到高校去参加培训、学习回来担负起城乡体育发展的重任。

(四) 因地制宜地开展体育特色项目

城乡体育的锻炼要从客观上考虑、解决农民实际问题。从体育项目少、时间难统一的实际问题来着想。因此,可以利用节假日(元旦、春节、端午、中秋)农闲季节,开展有意义娱乐性的体育项目和比赛活动。项目上可选一些集体娱乐、观赏性的。例:永年的太极拳,文化历史悠久,可以很好的推广与带动周边县市。群众参与了才能调动他们的积极性,才能达到促进改善身体形态和机能,以邯郸市城乡体育产业的管理体制及运行机制改革,建立科学有效的发展策略,特别是在依托地方邯郸市城乡体育产业发展的有针对性的对策。将会大大促进其快速发展的进程。因而大力发展邯郸城乡体育产业是在必行,不可耽搁的。

[参考文献]

- [1] 舒钧.连云港市体育产业发展现状及对策的研究[D].苏州大学,2009.
- [2] 刘兆春,申振川.调结构 增效益 统筹发展城乡经济——以河北省邯郸市为例[J].中国城市化电子杂志,2004,(4): 63.
- [3] 苏光华.上饶市体育产业发展现状与对策设想[J].体育经营与管理,2009,3(5):19-20.

[责任编辑:王云江]

The current development situation of Handan urban and rural P.E. industry and its strategy research

YAN Gui-fang, LI Da-yong, HAN Qiu-yu

(P.E. Department, Hebei University of Engineering, Handan 056038, China)

Abstract: By investigating and analyzing the current development situation of Handan urban and rural P.E. industry, problems are pointed out and the corresponding strategies and suggestions are put forward in this paper.

Key words: Handan city; urban and rural; P.E. industry; current situation; strategy; research

(上接第127页)

[参考文献]

- [1] 房国志,何丽,赵云志.基于组件技术IVR系统的设计与实现[J].情报检索,2008,(1): 5-6.
- [2] 于周锋,蒋泽军,王丽芳.动态加载TTS的IVR平台设计与实现[J].航空计算技术,2009,(6): 89-90.
- [3] 李远志,李浮滨.语音合成技术在信息服务应用的前景分析[J].现代情报,2002,(2):106-107.
- [4] Beasley R, Farley KM. VoiceXML语音应用程序开发[M].王建华,王卫峰译.北京:机械工业出版社,2002.

[责任编辑:王云江]

Design and research of Library's IVR system based on VoiceXML

WANG Da-wei, GU Wei-ping

(Library, Hebei University of Engineering, Handan 056038, China)

Abstract: In the paper, a new voice center platform is designed and implemented based on VoiceXML. This system's main skills includes the realization of DTMF, TTS and OpenVXI. The article build up a Voice Service system which can be used on the library and define the main functions of the system, service flow chart and VXML codes according to the practical situation.

Key words: IVR; VoiceXML; system configuration