

# H.323 交互式语音应答系统的设计与实现

牛国超, 武波, 杜军朝, 张立勇

(西安电子科技大学 软件工程研究所, 陕西 西安 710071)

**摘要:**交互式语音应答(IVR)系统是呼叫中心系统的最前端,它的质量直接影响整个系统的稳定性和可扩展性。传统呼叫中心的 IVR 系统是以电路交换为基础,由语音卡和流程控制软件构成的。基于 H.323 的 IVR 以软交换为基础,系统完全由软件实现,相对于传统的 IVR 来说,成本更低,扩展更容易。Openh323/PWlib 是符合 H.323 协议的开源的开发库,但是它有一个严重缺陷(I/O 非常频繁),利用内存映射机制对这个开发库做了重要改进,从而实现了高效的 IVR 系统。

**关键词:**IVR; 软交换; 呼叫中心; VoIP; H.323

**中图分类号:**TN915

**文献标识码:**A

**文章编号:**1673-629X(2008)04-0192-03

## Design and Implementation of Interactive Voice Response System Based on H.323

NIU Guo-chao, WU Bo, DU Jun-zhao, ZHANG Li-yong

(Software Engineering Institute, Xidian University, Xi'an 710071, China)

**Abstract:** Interactive voice response (IVR) system is the forefront of call center, and its quality affects the stability and scalability of the whole call center system directly. The IVR system of the traditional call center is composed of voice card and flow control software based on circuit switch. The IVR system based on H.323 is completely implemented by software based on soft switch. As compared with the former, the latter's cost is much lower and the scalability is much better. Openh323/PWlib is an open source development kits according with H.323 protocols, but it has a serious drawback (with very frequent I/O). Using memory mapping mechanism, do some important improvements to this development kits, and implement the high efficiency IVR system.

**Key words:** IVR; soft switch; call center; VoIP; H.323

### 0 引言

VoIP技术的发展,带来了呼叫中心的第二次革命<sup>[1]</sup>,它改变了传统呼叫中心的构建模式,基于 VoIP 的呼叫中心实现了以软交换为核心的呼叫中心系统。和传统呼叫中心系统相比,基于 VoIP 的呼叫中心的优点是<sup>[2]</sup>:不需要成本昂贵的程控交换机的支持,取而代之的是软件系统 Soft Switch;不仅仅支持 PSTN 接入,还支持因特网接入;不局限于地理位置的限制,支持分布式座席;不仅仅支持语音传输,还支持视频、即时消息的传输。图 1 是一个 VoIP 呼叫中心的网络结构图。

交互式语音应答系统(IVR)是呼叫中心的核心模块之一<sup>[3]</sup>,几乎所有的客户呼叫都要先接入 IVR,给客户 提供语音导航服务,生活中常见的 IVR 语音服务系

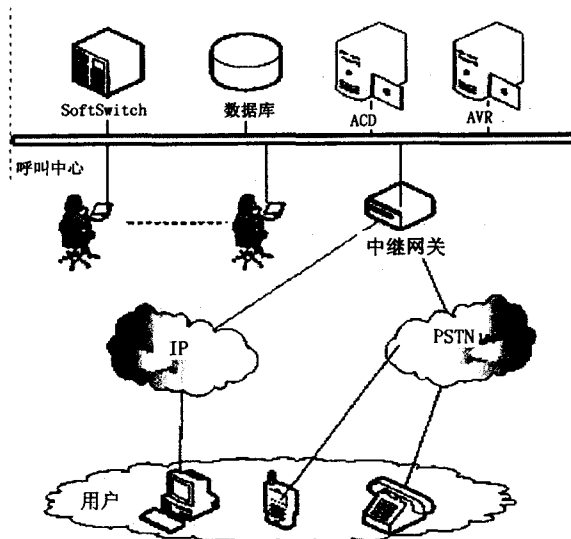


图 1 VoIP 呼叫中心网络结构

统有电话银行、联通移动的自动台、一号通、电话彩票投注等等<sup>[4]</sup>。传统呼叫中心的 IVR 一般采用“语音卡 + 流程控制软件”模式来实现,而一块语音卡一般仅能支持 4 路呼叫,成本高,系统扩充不方便。文中设计并

收稿日期:2007-07-27

**作者简介:**牛国超(1981-),男,河南许昌人,硕士研究生,研究方向为 VoIP 技术与应用;武波,教授,硕士生导师,研究方向为软件工程和软件设计理论与应用。

实现了基于 Openh323 的纯软件的 IVR 系统,该系统运行稳定、成本低、可扩展性好。一个 IVR 软件可以并发支持 30 路呼叫连接,如果把 IVR 分组,即把  $N$  个 IVR 软件聚合为一个 IVR 系统,则可以并发支持  $30 * N$  个话路连接。

H.323 协议是由 ITU-T 制订的基于分组交换网络(PBN)的多媒体会议系统协议族<sup>[5]</sup>,它被普遍认为是目前在分组网上支持语音、图像和数据业务最成熟的协议。采用 H.323 协议,各个不同厂商的多媒体产品和应用可以进行互相操作,用户不必考虑兼容性问题。该协议为商业和个人用户基于 LAN、WAN 的多媒体产品协同开发奠定了基础<sup>[6]</sup>。与 SIP 协议相比,H.323 推出较早,发展的比较成熟,采用传统的实现电话信令的模式,便于与现有的电话网互通<sup>[7]</sup>。

Openh323 项目是澳大利亚的 Equivalence Pty. Ltd. 公司组织开发的,这个项目 1998 年启动,到了 2000 年底基本的 H.323 协议框架已经实现,主要是免费面向所有想从事 VoIP 和网络视频传输的软件开发商使用,这个开发库是完全符合 H.323 协议的,能和任何符合该协议的软件进行视频和语音的通讯,例如 Microsoft 的 NetMeeting。

## 1 IVR 系统设计 与实现

### 1.1 系统设计

交互式语音应答系统一般应有如下功能:自动接听客户电话;根据客户的选择播放事先录制好的不同的音频文件;接收并存储用户的按键信息;检验用户的身份;还要具有把客户电话转接到人工座席的功能。以这些功能为出发点,结合 H323 协议来设计 IVR 的结构,系统由 4 部分构成,如图 2 所示。

- 1) XML 解析器,用以解析语音流程配置文件。
- 2) 状态机,控制呼叫连接的状态转移,用 State 设计模式实现。
- 3) 语音流程配置文件,存放语音流程配置信息。

- 4) H323 终端实体,用于建立跟客户电话间的连接,读取音频文件给客户电话,并接收客户按键信息。

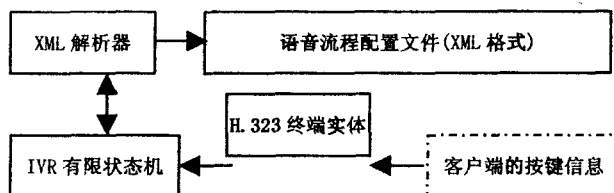


图2 IVR系统逻辑结构

对于每一个呼叫连接,IVR 都建立一个相应的有限状态机对象,系统的 XML 解析器根据用户的按键信息解析 XML 文件(语音流程配置信息被一次性读入内存中的一个 XML 树)把该呼叫连接的状态机转到相应状态,IVR 系统的有限状态机模型如图 3 所示,图中省略了“转人工”、“积分查询”、“留言”等细节。

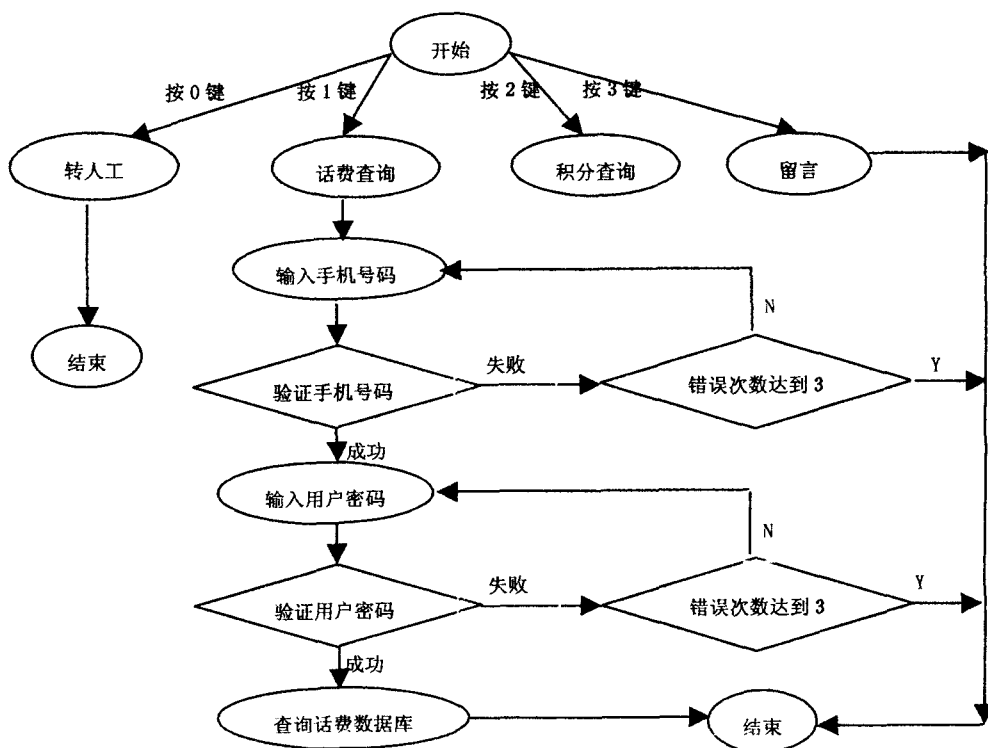


图3 IVR状态机模型示例

IVR 状态机中的节点类型有以下几种:

- (1) 等待用户按键,状态机根据配置文件跳转到相应的状态,给客户播放不同的提示语音;
- (2) 转人工,把用户话路转接到一个人工座席;
- (3) 等待用户输入 ID,如果用户输入 0 到 9 的一个数字则状态不变,如果用户输入 # 号键则跳转到“验证用户 ID”;
- (4) 等待用户输入密码,如果用户输入 0 到 9 的一个数字则状态不变,如果用户输入 # 号键则跳转到“验证用户密码”;
- (5) 验证用户 ID,查询数据库,检验这个 ID 是否

存在,如果验证通过则跳转到下一个状态,否则播放错误提示音;

(6) 验证用户密码,查询数据库,检验 ID 所对应的密码是否正确,如果验证通过则跳转到下一个状态,否则播放错误提示音;

(7) 查询话费数据库,利用 Text To Speech(TTS)技术,把查询到的文本信息以语音形式播放给客户;

(8) 留言录音,录制下客户的留言;

(9) 其它类型节点。

下面以客户查询手机话费为例来说明客户端电话跟 IVR 系统的交互过程,如图 4 所示。

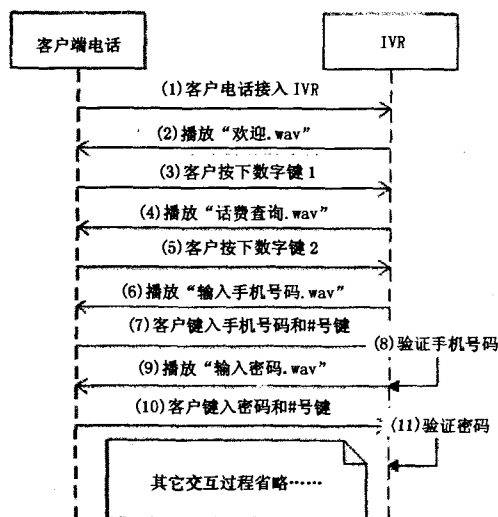


图 4 客户端电话与 IVR 的交互序列图

## 1.2 系统实现(基于 Openh323/PWlib)

### 1.2.1 改进底层开发库

IVR 系统和客户之间的通话连接实际上是由两个逻辑通道构成的,如图 5 所示,IVR 通过 OUT 通道给客户播放音频提示文件。在 Openh323 中,OUT 通道就是一个 PDelayChannel 对象 objOutChannel,这个对象和一个 PWaveFile 音频文件绑定在一起,即 IVR 按照一定的速率读取音频文件中的数据并写入 objOutChannel 中,例如对于 16 位 8k 赫兹采样频率的音频文件,IVR 每次仅仅读取 480 个字节,这就造成系统的 I/O 操作非常频繁,从而系统效率相当低,对硬盘寿命也有一定影响,文中提出了一种改进策略:从 PWaveFile 派生一个类 CMyWaveFile,重载它的 Read 函数,在系统中构建一个映射 map,每次调用 Read 函数读取音频数据的时候,先从 map 中查找相应的文件名,如果找到了,就从得到的内存区域读取数据,如果没找到,就把相应的音频文件一次性读入内存,并把相关信息加到 map 中,以备下次读取。

数据结构如下:

```
map <PString, BYTE * > g_mapWaveBuffer;
```

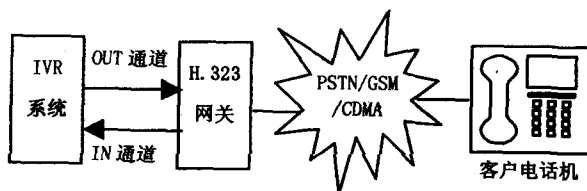


图 5 IVR 系统物理结构

### 1.2.2 系统扩展

对于一个小型的呼叫中心来说,并发支持 30 路呼叫的 IVR 已经足够了,但是对于大型的呼叫中心显然不行。笔者提出 IVR 分组的方法,即在 VoIP 呼叫中心系统中加入一个负载均衡器(Load Balancer),网络结构如图 6 所示,客户电话接入某个 IVR 之前,软交换模块 Soft Switch 询问负载均衡器(Load Balancer),负载均衡器查看所有 IVR 的状态,把负载最小的那个 IVR 告诉 Soft Switch,客户电话被接入这个 IVR。

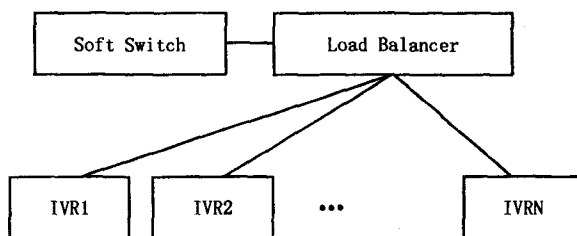


图 6 可扩展 IVR 系统结构

## 2 系统测试

为了检验系统的效率及可扩展性,笔者做了大量测试,测试环境如下:

软件:操作系统 Windows XP SP2, IVR 系统 MyIVR.exe,呼叫发生器 callgen323.exe;

硬件:IBM Workstation, Xeon 双 CPU 3.60GHz, 4GB DDR RAM;

测试用例:用呼叫发生器 callgen323.exe 并发的向 IVR 发出  $n$  路呼叫,然后主线程 Sleep 一分钟,再发起  $n$  路呼叫,如此循环往复,连续运行 10000 次,共建立  $n * 10000$  个呼叫连接。改进前后系统性能对比见表 1,并发呼叫量如图 7 所示。

表 1 改进前后系统性能对比

并发呼叫量	10	20	30	40	50	60
改进前 I/O 读取次数	91693569	182186698	273114830	364565312	455101182	546378603
改进后 I/O 读取次数	2815715	5676492	8442373	11201519	14371689	16832016
前后对比	降低 96.93%	降低 96.88%	降低 96.91%	降低 96.93%	降低 96.84%	降低 96.92%

(下转第 198 页)

实用性。

## 5 仿真结果

图 3 是在引入杂波的环境下的某一火控雷达显示器仿真,图 4 为被捕获目标飞行参数的实时显示。

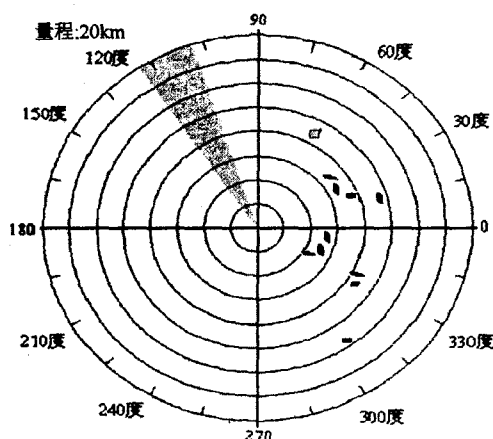


图 3 PPI 显示器仿真图

## 6 结束语

本设计通过交互式的操作来改变飞行目标的航行轨迹,并采用多线程技术很好地实现了多目标的火控雷达终端显示器的仿真。此软件在雷达手模拟操作训练和雷达课程教学以及火控雷达仿真研究中有很好的

目标编号:	000111
距离:	10657
速度:	114.93
加速度:	0
高低角:	43.95
水平角:	62.23

图 4 捕获目标跟踪数据显示

## 参考文献:

- [1] 丁鹭飞,耿富录.雷达原理[M].西安:西安电子科技大学出版社,2003.
- [2] 刘伟,杨万海.用VB实现PPI显示器的仿真[J].计算机仿真,2004,21(5):191-196.
- [3] 倪国旗,李如年.火控雷达的仿真目标建模及实现[J].火力与指挥控制,2001,26(3):57-59.
- [4] Skolnik M I.雷达系统导论[M].左群生,徐良国,马林,王德纯,等译.北京:电子工业出版社,2001.
- [5] 张万里,陈战林.C++ Builder 5.0 高级开发技巧与范例[M].北京:电子工业出版社,2001.
- [6] Hughes C, Hughes T. C++ 面向对象多线程编程[M].周良忠译.北京:人民邮电出版社,2002.

(上接第 194 页)

从表 1 可以看出,改进后 I/O 读取次数大幅度下降,约降低了 97%,这样既保护了硬盘,又提高了系统反映速度。

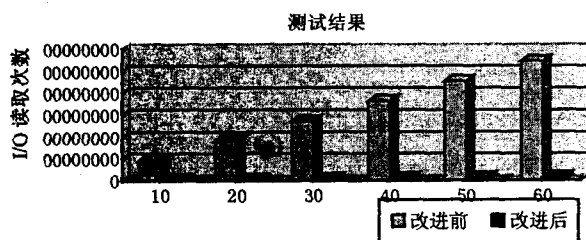


图 7 并发呼叫量

## 3 结束语

文中实现的 IVR 系统所依赖的开发库 Openh323/PWlib 是跨平台的,使得这个 IVR 具有良好的可移植性,已经应用于福州市邮政 11185 呼叫中心系统中(Linux 环境下),其运行状况良好。目前基于 VoIP 的 IVR 系统都是以 Voice XML 为基础的,它们遵从 Voice XML 1.0 标准,文中实现的 IVR 系统同样遵从

这个规范,但是系统中的 XML 解析器完全是笔者独立自主实现的,并开发了相应的语音流程配置工具,用户可以用这个图形化的工具自动生成语音流程配置文件(Voice XML 格式)。

## 参考文献:

- [1] 张登银,孙精科.VoIP 技术与系统设计[M].北京:人民邮电出版社,2003.
- [2] Collins D. VoIP 技术与应用[M].舒华英,等译.北京:人民邮电出版社,2003.
- [3] 周宽久,王宏玲,曲义飞.交互式语音专家系统研究[J].哈尔滨工业大学学报,2006(11):1981-1985.
- [4] 张臻,张仕军,黄本雄.可灵活扩展的 IVR 软件设计思想[J].计算机工程与设计,2002(10):47-57.
- [5] ITU-T. Proposed Revision of Recommendation H.323[S]. [s.l.]:[s.n.],1997.
- [6] Ransome J F, Rittinghouse J W. Amsterdam VoIP security[M]. Boston: Elsevier Digital Press, 2005.
- [7] 张荣,武波.SIP 协议的应用研究[J].计算机技术与发展,2006,16(6):71-73.