

# VoiceXML 语音平台性能指标研究<sup>①</sup>

## On Performance Index of VoiceXML - Based Voice Platform

庞亭亭 廖建新 朱晓民 (北京邮电大学网络与交换技术国家重点实验室 北京 100876)  
 (东信北邮信息技术有限公司 北京 100083)  
 吕文锋 (东信北邮信息技术有限公司 北京 100083)

**摘要:**提出了 VoiceXML (Voice eXtensible Markup Language) 语音平台的一套性能指标,为找出系统的瓶颈和优化系统性能提供参考,并且通过一个对系统进行优化的例子说明了对应性能指标的有效性。

**关键词:**VoiceXML 性能指标 文本语音转换 自动语音识别

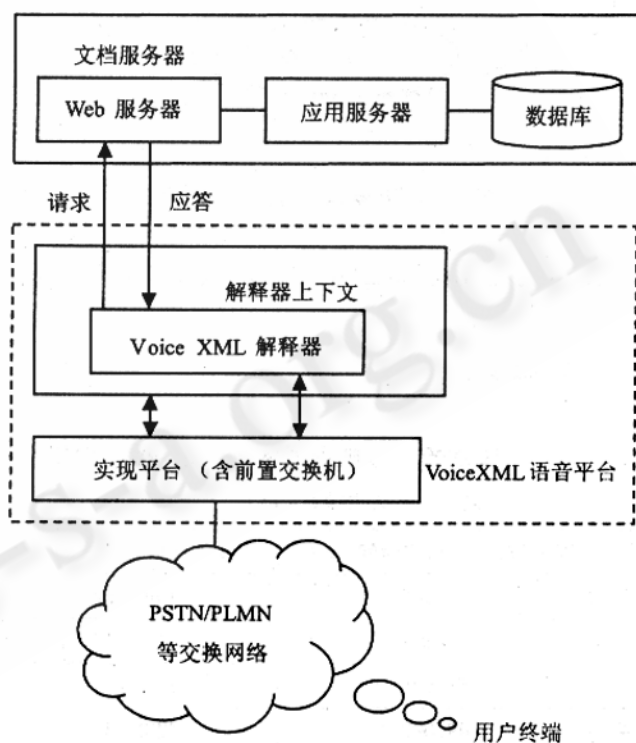
### 1 VoiceXML 语音平台及其性能研究现状

VoiceXML(语音扩展标记语言, Voice eXtensible Markup Language) 是一种基于 XML(可扩展标记语言, eXtensible Markup Language) 的因特网标记语言, 它的设计目的是实现类似于 HTML(超文本链接标记语言, Hypertext Markup Language) 的人机对话。但是 HTML 是通过图形浏览器并且使用显示终端、键盘、鼠标等设备来实现人机的交流, 而 VoiceXML 则是通过语音浏览器使用语音输入和输出来实现更人性化的人机会话, 达到访问因特网的目的<sup>[1]</sup>。

VoiceXML 的应用环境如图 1 所示。

该应用环境通过浏览器/服务器结构来集成语音服务和数据服务。文档服务器 (Document Server) 处理来自 VoiceXML 解释器的页面请求, 并以 VoiceXML 文档作为回应, 返回给解释器。解释器执行文档, 通过底层平台向用户放音, 进行 DTMF (Dual Tone Multi Frequency) 收号、语音识别, 执行呼叫前转等功能<sup>[2][3]</sup>。

从图 1 可以看出, 用户与 VoiceXML 语音平台的交互过程比传统的 IVR (交互式语音应答, Interactive Voice Response) 平台更加复杂, 其中任何环节出问题都可能会中断用户的通话, 低效的系统实现可能给运营商造成巨大的损失, 同时也给用户带来不便。



PSTN: Public Switched Telephone Network, 公共交换电话网络  
 PLMN: Public Land Mobile Network, 公众陆地移动网络

图 1 VoiceXML 的应用环境

VoiceXML 语音平台作为电信级系统, 必须具有高可靠、高效等特点。有很多基于 VoiceXML 的语音平台实

① 基金项目: 国家杰出青年科学基金 (No. 60525110); 国家 973 计划项目 (No. 2007CB307100, 2007CB307103); 新世纪优秀人才支持计划 (No. NCET-04-0111); 电子信息产业发展基金项目 (基于 3G 的移动业务应用系统); 电子信息产业发展基金重点项目 (下一代网络核心业务平台)。

现,但系统地研究 VoiceXML 语音平台性能的文献却很少,特别是没有衡量 VoiceXML 语音平台性能的统一标准。

目前对 VoiceXML 语音平台性能的研究,主要还是根据一般的通信性能指标来衡量的,这样只能从总体上对 VoiceXML 语音平台进行评价,并不能有针对性地说明 VoiceXML 语音平台各方面的优劣,更不能为系统细节方面的改进提出参考。

## 2 VoiceXML 语音平台性能指标

要想对 VoiceXML 语音平台的性能做有效的研究,只有通信系统的一般性能指标是不够的。本文提出一套 VoiceXML 语音平台性能指标,它们与通信系统的一般性能指标一起,为系统优化提供理论上的支持,文中如果没有特别说明,则语音平台指的是 VoiceXML 语音平台。

(1) 语音平台接通率。指呼叫到达语音平台前置交换机后,语音平台正常接通此呼叫的概率,它是由一段时间内对语音平台的统计值计算得出的:

语音平台接通率 = 语音平台成功接通呼叫数 / 成功到达前置交换机的呼叫 \* 100%

语音平台接通率与瞬时用户量、系统负载、语音平台容量等因素有关。

(2) 语音平台可用性。代表语音平台向用户提供服务的能力,其基础是语音平台接通率的门限值  $c$ 。在实际测量中  $c$  可以根据具体情况进行设置,用户向语音平台发起呼叫,若在规定时间内接通率不低于  $c$ ,则语音平台可用性值为“真”,否则为“假”。

(3) 语音平台容量。代表语音平台同时向多用户提供服务的的能力,它是指系统在满负荷运行时能同时服务的用户数。它通常是前置交换机时隙数和语音资源数量两者的最小值。

(4) 语音平台平均无故障率。代表语音平台的稳定性,它是指系统无故障时间与系统运行时间的比率。

(5) 语音平台接通时延。代表语音平台对入呼叫的响应速度,它是指从用户完整输入并拨打语音平台接入号码开始,到听到语音平台的第一条提示音为止所经过的时间。语音平台接通时延主要由两部分组成:呼叫信令通过交换网络到达语音平台前置交换机的时延(网络交换时延)与语音平台响应时延。

(6) 语音平台响应时延。代表语音平台对正常操作的响应速度,它是指从系统收到入呼叫信令或者用户的输入信息开始,到系统挂机或用户听到语音平台的提示音为止所经过的时间。语音平台响应时延通常由四部分组成:获取 VoiceXML 文档所需时间,解析所获得的 VoiceXML 文档的时间,执行必要的自动语音识别(Automatic Speech Recognition)语法加载和文本语音转换(Text-To-Speech)的时间,语音平台输入/输出时延。

(7) VoiceXML 文档获取时间。代表系统获取 VoiceXML 文档的速度,它是 VoiceXML 解析器从发出文档获取请求,到获得所需文档所经过的时间。文档获取时间跟要获取的文档大小、是否缓存、缓存策略与获取顺序等因素都有关系。

(8) VoiceXML 文档解析时间。代表系统处理 VoiceXML 文档的速度,它是系统从获得 VoiceXML 文档开始,到将该文档解析成一棵 XML 文档树所经历的时间。

(9) 文本语音转换时间。VoiceXML 支持文字形式的输出信息,它是通过文本语音转换(语音合成)来实现的。文本语音转换时间是指,从系统开始文本语音转换,到获得文本语音转换结果所需要的时间。

(10) 语音平台输入/输出时延。语音平台的输入包括用户的语音输入和 DTMF 按键输入。语音平台的输入时延是指从系统接收到用户的输入信息开始,到系统对语音信息或 DTMF 信息进行识别,并生成具有语义的文本信息所需要的时间。语音平台的输出时延是指从系统发出放音请求,到放音设备获取音频文件并播放给用户听到所经历的时间。

这些性能指标相互之间既有区别又有联系,有些可以直接测量得出,而有些则可以通过对已有的测量数据统计或计算得出,不但涵盖了用户与语音平台交互过程的细节,又有语音平台接通率等基于统计数据的指标,形成 VoiceXML 语音平台的多层次性能指标体系,如图 2 所示,其中基于测量的性能指标用虚线框表示,基于统计或计算的指标用实线框表示。

## 3 性能指标的有效性

由于性能指标比较多,限于文章篇幅,这里不能一一证明这些指标的有效性。我们通过针对语音平台某

一个性能指标进行相应优化的测试数据,来说明性能指标的有效性。

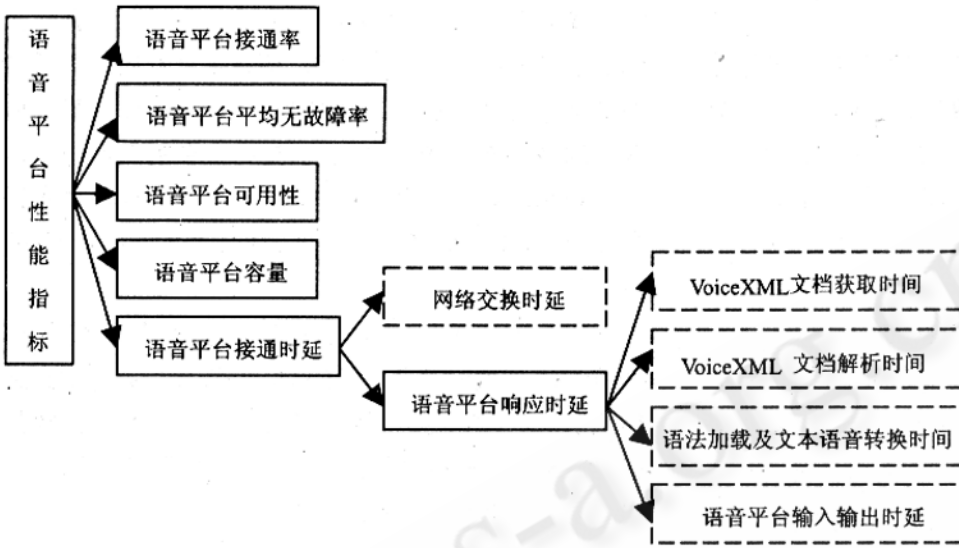


图 2 VoiceXML 语音平台性能指标层次结构

以 VoiceXML 文档获取时间为例。根据定义,它是 VoiceXML 解析器从发出文档获取请求,到获得所需文档所经过的时间。为了测试数据的可比性,我们在测试时保证除了 VoiceXML 文档的存放位置不同外,其余软硬件条件都相同。我们认为,从定性的角度来说,一般情况下 VoiceXML 通过计算机网络比通过本地磁盘需要更长的获取时间,同一个语音平台的 VoiceXML 文档放在远程服务器上和本地磁盘时的文档获取时间是有差别的<sup>[4]</sup>,而 VoiceXML 文档获取时间这一性能指标正好反映了这种差别。如果 VoiceXML 文档放在本地时具有更短的获取时间,使得语音平台的响应速度更快,那么就能够证明针对此性能指标所做的优化对语音平台的整体性能提高是有帮助的,即证明了其有效性。

下面通过测试数据来说明该指标的有效性。测试时使用一台 HP DL380 服务器作为语音平台服务器,互联网中的一台 PC 机作为远程服务器, VoiceXML 文档的大小都为 100KB。在语音平台程序及呼叫模拟程序中采集以下数据:用户呼叫信令到达前置交换机的时刻、发出文档请求的时刻、获取到文档的时刻、用户听到提示音的时刻。经过对采集到的数据进行计算,能够得到每次呼叫的语音平台响应时延和 VoiceXML 文档获取时间,进而获得两种情况下(从远程服务器获取

或者从语音平台服务器本地获取 VoiceXML 文档)的平均语音平台响应时延和平均 VoiceXML 文档获取时间。

我们对两种不同的文档位置分别做了 10 次单呼叫串行模拟测试,测试结果如表 1 与表 2。

从测试数据可以看出, VoiceXML 文档获取时间优化后的平均文档获取时间比优化前少了 30.246 毫秒,减少了 35.97%,而平均语音平台响应时延也比优化前减少了 3.98%,说明针对 VoiceXML 文档获取时间进行的优化能够提高语音平台性能, VoiceXML 文档获取时间这一性能指标是有效的。从另一个角度还可以看出,对 VoiceXML 文档获取时间的优化并没有为语音平台的执行效率带来显著提高,说明语音平台的瓶颈并不在文档的获取上,还需要优化其他性能指标才能更好地提高整个语音平台的性能。

表 1 从远程服务器获取 VoiceXML 文档的情况

	VoiceXML 文档获取时间(毫秒)	语音平台响应时延(毫秒)
1	84.763	1269.03
2	82.136	1264.415
3	83.869	1264.069
4	89.039	1263.033
5	76.625	1261.656
6	80.855	1262.375
7	85.321	1259.053
8	88.584	1259.265
9	87.216	1259.233
10	82.518	1260.361
平均	84.093	1262.249

#### 4 总结

本文通过分析 VoiceXML 语音平台的执行流程及其性能指标的研究现状,提出了衡量 VoiceXML 语音平

(下转第 19 页)

表 2 从语音平台服务器本地获取 VoiceXML 文档的情况

	VoiceXML 文档获取时间(毫秒)	语音平台响应时延(毫秒)
1	55.865	1189.952
2	62.186	1220.057
3	53.826	1220.069
4	49.099	1210.033
5	56.716	1210.086
6	50.056	1210.053
7	48.928	1210.053
8	51.574	1210.065
9	51.684	1220.062
10	58.531	1220.068
平均	53.847	1212.050

台的一整套性能指标,为找出系统瓶颈和提出优

化方案提供了参考,并针对某性能指标对系统性能进行优化举例说明了性能指标的有效性。下一步工作是研究如何通过本文提出的这套性能指标找到系统优化的重点,及如何优化系统各模块的执行效率,进而提高系统的整体性能。

### 参考文献

- 1 舒挺、张国焯,基于 VoiceXML 技术的信息服务集成,计算机应用,2003,23(6):1~3.
- 2 李明华、徐良贤,VoiceXML 语音浏览器的研究,计算机工程,2002 年 10 月,28(10):1~3.
- 3 郭志伟、武家春、刘岩、廖建新,应用于 VoiceXML 系统的缓冲机制的研究,信息网络,2005 年第 11 期:1~3.
- 4 詹营、武家春、蔡斌、廖建新,通过改写缓冲文件提高 VoiceXML 性能的研究,中国数据通信,2005,7(4):1~3.