

VoiceXML 기술을 이용한 홈네트워크 정보기기 원격 제어 시스템의 설계 및 구현

이진구, *정문상
경북대학교 대학원 정보통신학과, *경북대학교 대학원 전자공학과
전화 : 053-940-8898 / 핸드폰 : 011-9355-7729

Design and Implementation of Home Network Information Appliance Remote Control System Using Voice XML Technology

Jin-Gu Lee, Moon-Sang Jeong
Dept. of Information and Communication, Kyungpook National University
E-mail : mail@leejingu.pe.kr

Abstract

VoiceXML is designed for creating audio dialogs that feature synthesized speech, digitized audio, recognition for spoken and DTMF key input, recording of spoken input, telephony, and mixed-initiative conversations. Uses the VoiceXML and there is a place objective which does information home appliance machinery and tools control. When it uses the VoiceXML, it will be able to provide a bias characteristic to the user. The XML base the gearing with different civil official system is possible. With studying VoiceXML and OSGi, this paper has designed and implemented the control architecture of information home appliances.

장하였다.

그리고 홈네트워크 기술을 통합을 위하여 OSGi forum에서 OSGi(Open Service Gateway initiative)가 표준화에 주력하고 있다. 현재는 홈네트워크 정보기기를 제어하기 위하여 OSGi 또는 홈서버에서 사용자가 직접 제어명령을 주거나 외부 인터넷에서 정보기기를 제어하는 방법 중심으로 기술이 발전하고 있다. 하지만 제어를 위한 UI(User Interface)를 제조회사별로 다르게 개발하여 사용자 편의성에 문제가 제기되고 있다. 본 논문에서는 이를 해결하기 위하여 VoiceXML 표준을 사용하여 음성으로 가정내부 또는 외부에서도 홈네트워크 가정정보기기 사용의 편의성을 높이고, OSGi와 연동함으로써 단일화된 서비스를 제공할 수 있는 VoiceXML 기반의 홈네트워크 정보기기 원격 제어 시스템을 설계 구현하고자 한다.

II. VoiceXML

I. 서론

앞으로 정보가전기기 분야는 홈네트워크와 인터넷의 연동으로 사용자가 손쉽게 운용가능한 형태로 발전할 것이다. 홈네트워크 플랫폼 기술은 Bluetooth, HomePNA, PLC(Power Line Communication), 통신을 위한 기술은 UPnP(Universal Plug and Play), JINI, HAVi(Home Audio and Video Interoperability) 등이 있다. 이러한 홈네트워크의 다양한 기술과 제품이 등

2.1 VoiceXML

VoiceXML(Voice Extensible Markup Language)은 일반전화기 같이 음성 입출력을 사용하는 디바이스를 통해 사람과 컴퓨터가 소통케 하는 XML 기반의 언어이다. VoiceXML Forum에서 제안하여 World Wide Consortium(W3C)에서 승인을 받았다. 2000년 5월에 VoiceXML 규격 1.0이 발표되었고, 2001년 8월 VoiceXML 2.0 Working Draft가 발표되었다.

2.2 VoiceXML 구조 모델

그림 1 은 VoiceXML 구조모델이다. VoiceXML 구조모델을 보게되면 VoiceXML Interpreter에서 VoiceXML 문서를 해석하고 사용자와의 상호작용을 수행하기 위해 실행 플랫폼을 제어한다. VoiceXML Interpreter context는 Document Server, 즉 웹서버와의 Request/Response를 담당하게 된다.

Implementation Platform은 전화수신기능, 전화 호전달(Call Transfer)기능을 가지고 있다.

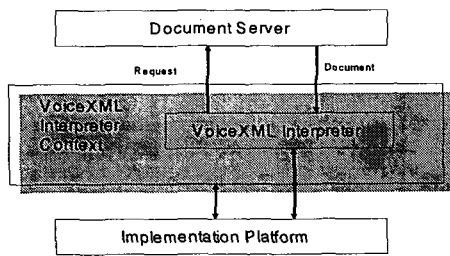


그림 3 VoiceXML 구조 모델

사용자가 전화를 걸게되면 Implementation Platform에서 호제어를 하고, 호(call)를 VoiceXML Interpreter Context에 넘겨준다. 넘겨받은 VoiceXML Interpreter Context가 해당 애플리케이션을 수행하는데 필요한 초기 설정을 하고, Http(Hyper Text Transfer Protocol)를 이용하여 Document Server에 문서를 요청하게 된다. 전달받은 문서는 VoiceXML Interpreter가 해석하고 그 결과를 Implementation Platform에 전달한다.

VoiceXML Gateway는 그림1에서 살펴본 기능에 음성인식기능(ASR : Automatic Speech Recognition), 음성합성기능(TTS: Text To Speech), 음성과 오디오 재생기능, 음성과 오디오 녹음 기능 등을 추가하여 하나의 시스템으로 구성되어있다.[2,3]

위의 과정을 처리단계별로 설명을 하면, 사용자가 음성 전화 서비스 번호로 전화를 걸면, 실행 플랫폼에 있는 전화 인터페이스부가 전화 연결에 필요한 절차에 따라서 교환기와 전화연결을 설정한다. 전화 연결이 설정된 뒤에는 VoiceXML Interpreter Context가 해당 애플리케이션을 수행하는 데 필요한 초기 설정을 마치고, HTTP 통신을 통해 문서 서버인 웹서버에 있는 애플리케이션을 구성하는 VoiceXML 문서들을 로딩하고 VoiceXML 해석기가 해당 문서들을 해석하고 실행하도록 요청한다. VoiceXML 해석기는 사용자와 대화를 진행하면서 필요한 경우에 VoiceXML 문서에 기술된 순서에 따라 실행 플랫폼을 적절히 제어한다.

III. VoiceXML을 이용한 HomeNetwork 가전 정보기기 제어 구조

3.1 HomeNetwork의 기본 구조

재택근무나 가정자동화 시스템의 핵심은 홈 네트워크다. 홈네트워크는 PC에서부터 TV·디지털 카메라·냉장고 등 모든 가전제품을 무선 또는 유선으로 연결시킨 것을 말한다. 이렇게 되면 집 밖에서 냉·난방기기를 작동시키거나 비디오를 제어할 수도 있게 된다.

이러한 홈네트워크 가전기기를 제어하기 위해서는 OSGi(Open Service Gateway initiative)를 이용한다. OSGi는 표현층에 중점을 둔 표준으로 다른 하부계층에 독립적이다. OSGi 표준의 핵심은 바로 서비스 게이트웨이이다. 서비스 게이트웨이는 외부망과 내부망 사이에 위치하게 되는 서버라고 생각하면 쉽다. 서비스 게이트웨이는 가정이나 사무실 또는 어떤 다른 지역으로부터의 또는 그 곳의 음성, 데이터, 인터넷과 멀티미디어 통신을 관리하고 통합하는 역할을 한다.

OSGi는 특정 기능을 수행하는 자바 인터페이스와 실제 구현 객체인 Service, Service를 제공하기 위한 기능적 배포단위인 Bundle, Bundle들의 라이프사이클을 관리하는 Bundle 실행 환경인 Framework으로 구성되어 있다.그림 3은 OSGi의 계층적구조를 보여준다.[4]

OSGi Architecture

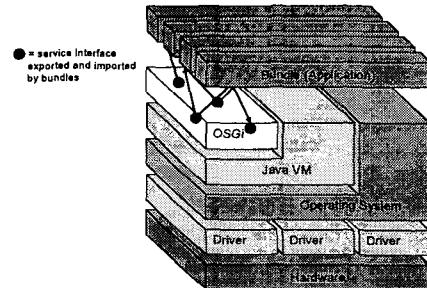


그림 3 OSGi Architecture

3.2 HomeNetwork 가전기기 제어를 위한 VoiceXML Document

홈네트워크의 가전기기를 제어하기 위해서 그림 4와 같은 구조로 동작할수 있다.

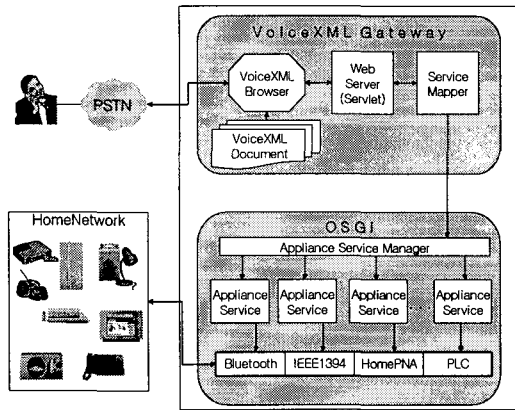


그림 4 VoiceXML 홈네트워크가전기기 제어구조

사용자로부터의 음성이 VoiceXML Gateway에 들어 오면 VoiceXML 문서를 분석하여 명령을 URL(uniform resource locator)형태로 웹서버에 요청을 하게 된다. 웹서버에서 URL의 ServiceMapper를 호출하게 되는데 ServiceMapper는 Servlet, JSP, ASP 등으로 작성되어 있다. ServiceMapper는 입력된 명령을 분석하여 OSGi의 Appliance Service Manager로 보낸다. 이 ServiceMapper는 VoiceXML 문서를 만들어 낼뿐만 아니라 HTML, WML(Wireless Markup Language)등으로 변환이 가능한 기능을 제공한다. 이는 음성시스템 이외에도 다른 시스템과의 호환성을 가질수 있는 기능을 한다.

Appliance Service Manager는 각각의 해당되는 Appliance Service에 데이터를 전달한다. OSGi에서는 각각의 가전정보기기들을 서비스 단위로 관리를 하는데 ServiceMapper에서 전달받은 명령을 분석하여 해당되는 Service로 전달하게 된다.

Appliance Service는 해당되는 하부 네트워크로 명령을 보내어 가전 정보기기에 전달하게 된다. 각각의 정보기기에 전달된 데이터는 수신된 명령을 분석하여 현재 장비의 정보를 사용자에게 알리거나 명령에 맞는 동작을 취한후 결과를 사용자에게 보고한다.

가전정보기기가 명령을 처리한후 사용자에게 보고는 위의 과정의 역순으로 돌아가는데 데이터가 ServiceMapper에 오게되면 결과 메시지를 분석하여 Servlet에서 VoiceXML 문서를 만들어 VoiceXML Browser에게 보낸다. VoiceXML Browser는 받은 문서를 해석하여 사용자에게 음성으로 결과를 통보한다.

위의 절차로 진행되는 과정에 하나의 에러라도 발생하게되면 사용자에게 에러의 원인을 음성으로 보고할

수 있다. 특정 시스템이 동작하지 않을 경우 해당시스템에서 에러가 있음을 보고 할수 있다.

본 시스템에서 홈네트워크의 가전정보기기들의 로그(log) 상태를 저장한후 사용자가 검색하여 보고받을수도 있다. OSGi에 Log Service를 두어서 각각의 정보기기들의 상태 및 에러를 저장하여 사용자의 요청이 있을 경우 보고할수도 있다.

사용자의 인증문제가 야기될수 있다. 이를 해결하기 위해서 사용자로부터의 전화가 걸려오면 인증을 하고 다음 단계로 진행을 한다. 인증을 위한 방법은 크게 두가지로 나누어 설명할수 있다. 첫째, VoiceXML 문서에 인증기능을 넣어 처리하는 방법이 있다. 사용자가 이름과 비밀번호를 넣으면 VoiceXML 문서에서 사용자의 이름과 비밀번호 일치여부로 구분가능하다. 둘째, VoiceXML Browser에서 입력된 이름과 비밀번호를 웹서버로 보내어 별도의 인증프로그램을 두어서 사용자를 구분할수 있다. 첫 번째 방법은 간편하며 속도가 빠른반면 VoiceXML 문서가 노출될 경우 보안의 위험이 존재하고 다른 인증 시스템과의 연동에 어려움이 존재할수 있다.

IV. 구현

본 논문에서는 VoiceXML 기술을 이용한 홈네트워크 원격제어시스템의 구현을 위하여 IBM WebSphere Voice Server SDK 2.0, Apache Tomcat 4.0.1 그리고 BlueTooth Toolkit을 사용하여 구현해 보았다. IBM WebSphere Voice Server SDK 2.0은 자체에 VoiceXML Browser, 음성인식모듈, 음성합성모듈 등 일부 제한적인 VoiceXML Gateway기능을 공개하고 있다.

그림 5는 텔레비전의 VoiceXML 문서이다. 사용자가 메뉴를 듣고 메뉴를 전체 혹은 일부를 말하게되면 음성인식 모듈이 사용자의 음성을 인식한 후에 해당되는 명령을 수행한다. 음성이 외에도 전화의 숫자를 누르면 DTMF 신호로 전달되어 명령을 수행한다.

```
<!-- 생략 -->
<option dtmf="6" value="ChangeCh">Number 6, Change Volume</option>
<filled> <assign name="command" expr="menu"/>
<if cond="menu == 'ChangeVol'/"> <goto next="#inputVolume"/>
<!-- 생략 -->
<form id="InputVolume"> <field name="value" type="digits?length=2">
<prompt> press two digit Volume.</prompt>
<filled> <submit next="http://localhost/ServiceMapper" nameList="machine
command value"/> </filled> </field> </form>
<!-- 생략 -->
```

그림 5 Television.vxml

메뉴가 선택이 되면 ServiceMapper가 호출이 되는데 이때 가전제품의 이름과 명령 그리고 값이 전달된다. 전달된 값을 OSGi에 전달하고 최종적으로 가전정보보기기에 명령이 전달되어 가전정보보기가 명령을 수행한다.

가전기기에서 명령을 수행한 후 역으로 메시지를 보내면 ServiceMapper에서 그림 6과 같은 VoiceXML 문서를 생성하여 VoiceXML Browser에 보낸다.

```
<?xml version="1.0" encoding="iso-8859-1"?>
<vxml version="1.0" application="root.vxml">
  <form>
    <block>
      Television Turn On Success
    </block>
  </form>
</vxml>
```

그림 6 ServiceMapper에서 생성된 VoiceXML

그림8의 문서를 VoiceXML Browser에서 해석하여 사용자에게 결과를 통보한다.

V. 결론

홈네트워크의 정보 가전기기를 좀더 쉽게 제어하기 위하여 많은 방법들이 논의되고 있다. 많은 방법들의 공통점은 사용자의 편의성에 초점이 맞춰진다. 본 논문에서는 사용자의 편의성을 도모하면서 XML을 사용하여 정보가전기기들의 제어를 표준화하고 단일화에 관하여 논하였다.

VoiceXML의 성능은 음성인식모듈의 인식성공율과 직결되는 문제이며 데이터의 처리시간에 민감한 문제이다. 인식성공율이 낮을 경우 사용자는 만족스럽게 느끼지 못할수도 있으나 음성인식분야는 전 세계에서 활발히 연구, 논의되고있는 분야이다. 이는 차후 기술이 발전하면서 가까운 시간안에 극복될수 있는 문제이다.

그리고 처리시간에 관한 문제를 고려하여야 한다. 만일 사용자가 음성명령을 내린후 기다리는 시간이 2-3초가 된다면 사용자는 지루하게 느끼거나 사용함에 편의성을 느끼지 못할수도 있다. 이는 현재까지 사용되어온 음성정보 서비스의 처리단계보다 몇단계 더 거침으로써 나타날 수 있는 문제이다. 하지만 기반기술들의 발달과 소프트웨어의 발전에 힘입어 이또한 문제를 줄일수 있다.

차후 확장성을 고려하여 음성 뿐만아니라 WML(Wireless Markup Language)을 이용하여 무선 인터넷환경에서도 제어가 가능할 뿐만아니라 일반적인

인터넷을 통하여 제어도 가능할것이다. 그리고 정보기만을 대상으로 하지 않고 개인정보관리시스템, 자동차 항법안내장치, 일반적인 음성안내 시스템, Voice-Commerce, UMS(Unified Messaging Service) 등과의 연동도 함께 고려할수 있다.

참고문헌

- [1] 박성협, "음성 웹 애플리케이션 구축을 위한 VoiceXML", 한빛미디어, 2001. 12
- [2] Eve Andersson, Stephen Breitenbach, Tyler Burd, Nirmal Chidambaram, Paul Houle, Daniel Newsome, Xiaofei Tang, Srdjan Vujosevic, Xiaolan Zhu, "Early adopter VoiceXML", WROX press, August 2001
- [3] <http://www.voicexml.org>
- [4] <http://www.osgi.org>
- [5] <http://www.w3c.org>