

모바일 환경에서 인증기능과 음성인식을 위한 웹 서비스 클라이언트 구현

고유정, 오지영, 김윤중
한밭대학교 컴퓨터공학과

An implementation of the web service client for speech-recognition and certification Module on Mobile environment.

You-Jung Ko, Ji-Young Oh Yoon-Joong Kim
Dept. of Computer Science, Hanba Nat'l University

요약

무선 인터넷 서비스는 언제 어디서나 사용할 수 있는 편리함을 제공한다. 반면 무선 인터넷 서비스에 접속하기 위한 단말기는 저장 공간이 제한적이고 입력하기가 불편하다. 위와 같은 문제점을 개선하기 위해 본 연구에서는 별도로 제공되는 음성인식 시스템을 웹 서비스로 요청하여 음성인식이 가능한 PDA클라이언트 시스템을 구현하였다.

본 시스템은 모바일 클라이언트와 웹 서비스 프로바이더로 구성되어 있다. 모바일 클라이언트는 웹 서비스를 요청하기 위해 사용자정보와 인식할 음성데이터를 웹 서비스 프로바이더로 전송한다. 웹 서비스 프로바이더는 웹 서비스 보호를 위해 인증된 클라이언트에 한하여 음성인식결과와 사용자명을 반환하는 서비스를 제공한다.

1. 서론

무선 인터넷은 이동통신 전화 같은 공중망을 사용하여 데이터 통신을 가능하게 하고 언제 어디서나 사용할 수 있는 편리함을 제공한다[1]. 반면 무선인터넷에 접속하기 위한 단말기는 휴대가 가능하도록 크기는 작고 가벼워야 하므로 프로세서의 처리능력이나 메모리용량이 제한적이고 스크린 크기가 작아 입력하기에 불편함이 따른다. 따라서 PDA(Personal Digital Assistants)와 이용자간의 정보전달 방식으로 음성을 사용하고자 하는 욕구가 증가하면서 음성인터페이스에 대한 연구가 이루어지고 있다.

상용제품으로 D&M Technology의 PDA용 음성인식 소프트웨어인 Pocket Voice가 출시되었고, 관련 연구는 서버와 클라이언트에 음성 인식 모듈을 분산시켜 구축한 'VoiceXML을 이용한 모바일 음성인식 시스템구현'이 있다[2]. 이러한 연구에서는 자원이 제약적인 PDA에 음성인식 소프트웨어를 설치해야 하거나 클라이언트가 시간과 비용을 투자하여 음성인식의 일부 모듈을 구축해야 하는 어려움이 있다.

상기의 문제점을 보완하기 위해 본 연구에서는 별도

로 제공되는 음성인식시스템을 웹 서비스로 요청하여 음성인식이 가능한 PDA 클라이언트 시스템을 구현하였다. 또한 본 연구에서 구현한 음성인식 웹 서비스는 허가되지 않은 사용자로부터 웹 서비스의 보호와 재공격 방지를 위해 사용자를 인증하고 응답시간을 제어하도록 구현하였다.

본 시스템은 (그림1)과 같이 모바일 클라이언트와 인증기능을 지닌 음성인식 웹 서비스 프로바이더로 구성되어 있다.

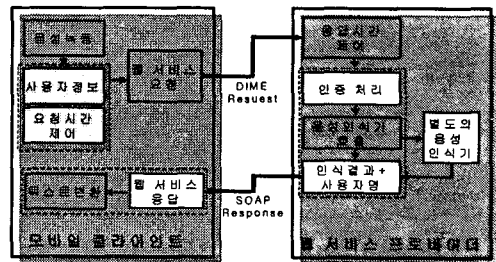


그림 1. 전체 시스템 구성도

모바일 클라이언트는 음성명령을 녹음하고 사용자 정보를 입력받아 DIME 방식으로 웹 서비스를 요청한

다. 이때 요청되는 메시지는 재공격 방지를 위해 요청 시간을 제어한다. 또한 요청된 결과를 응답받고 그 결과를 텍스트로 변환하는 기능을 포함한다.

웹 서비스 프로바이더는 사용자정보를 검사하여 인증된 클라이언트에 한해 별도의 음성인식시스템을 호출하여 인식된 결과와 사용자명을 SOAP 응답 메시지로 반환한다. 응답메시지는 메시지의 만료시간을 설정하여 응답시간을 제어한다.

본 논문의 구성은 다음과 같다. 2장은 웹 서비스를 요청하기 위한 모바일 클라이언트의 구현방법을 설명한다. 3장은 웹 서비스 프로바이더에서 사용자를 인증하고 인식결과와 사용자명을 반환하는 과정을 설명한다. 4장은 구현한 시스템의 실험 결과를 나타내고 마지막으로 5장은 결론을 기술한다.

2. 모바일 클라이언트

화면의 크기가 제한적인 PDA에 스타일러스나 키패드 대신에 웹 서비스의 음성인식 기술을 적용하기 위해서는 음성 녹음이 가능한 클라이언트를 구현해야한다. 클라이언트는 음성녹음기능, 사용자 정보 구성과 요청시간 제어 기능, 웹 서비스 요청 기능, 웹 서비스 응답 처리 기능을 가진다.

2.1 음성녹음기능

본 연구에서는 음성녹음기능을 위해 제공되는 라이브러리(VoiceRecord.dll)의 녹음(Record()), 정지(Stop()), 재생(Play()) 메서드를 호출하여 음성녹음컨트롤을 구현하였다. 관리코드(Managed Code)인 클라이언트 응용프로그램에서 비관리 코드(Unmanaged Code)인 VoiceRecord.dll의 함수를 호출하기 위해 System.Runtime.InteropServices에서 제공하는 DllImport 애트리뷰트를 이용하였다.

2.2 사용자 정보 구성 및 요청시간 제어

본 연구에서 구현한 인증을 지닌 음성인식 웹 서비스는 WS-Security의 <UsernameToken>[3]을 이용하여 사용자를 식별한다. 모바일 클라이언트는 이와 같이 사용자를 인증하는 웹 서비스를 이용하기 위해 사용자 정보를 전달한다. 사용자 정보는 PDA 인터페이스에서 사용자명과 패스워드를 입력받아 <UsernameToken>element로 구성하였다.

송신 메시지는 악의적인 사용자가 정보를 가로채고 저장한 후 재 전송 할 수 있다. 본 연구에서는 이러한 문제점을 방지하기 위하여 웹 서비스 요청시간을 제

어한다. 요청시간 제어는 <TimeStamp>element로 메시지의 생성 시간과 만료시간을 구성하였다. 만료시간은 메시지가 생성된 이후 30초로 설정하였으며 이때 요청 메시지가 만료시간까지 웹 서비스 프로바이더로 전송되지 못하면 오류를 반환한다. 본 논문에서 구현한 메시지 구조는 그림2와 같다.

```
<soap:Header>
  <Security
    xmlns="http://schemas.xmlsoap.org/ws/2002/07/secext">
    <UsernameToken>
      <Username>KYJ </Username>
      <Password>1234</Password>
    </UsernameToken>
    </Security>
  <Timestamp
    xmlns="http://schemas.xmlsoap.org/ws/2002/07/utility">
    <Created>2003-09-23T09:25:28Z</Created>
    <Expires>2003-09-23T09:25:58Z</Expires>
  </Timestamp>
</soap:Header>
```

그림 2. UsernameToken과 Timestamp 메시지 구조

2.3 웹 서비스 요청

본 연구에서는 음성 데이터를 웹 서비스로 송신하기 위한 방법으로 DIME(Direct Internet Message Encapsulation)[4]방식을 사용하였다. DIME 메시지 방식은 바이너리 포맷형식으로 전송되므로 SOAP 메시지 방식에 비해 부피가 가벼운 장점을 가지고 있다. 이러한 DIME을 이용하기 위해서는 마이크로소프트에서 릴리즈된 WSE(Web Service Enhancements)[5]를 사용해야 한다. 그러나 WSE는 서버환경에서는 사용 가능 하지만 모바일 디바이스 환경에서는 아직 사용할 수 없다. 따라서 WSE가 처리해주는 부분 즉, DIME 메시지 생성 및 전송부분을 구현하였다. 그림 3은 본 연구에서 구현한 DIME 메시지 구조이다.

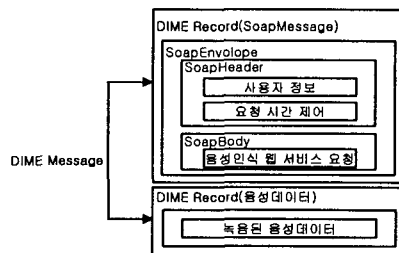


그림 3. 음성데이터를 첨부하는 DIME메시지 구조

DIME 메시지 구조는 SOAP 메시지 형식으로 구성된 다임 레코드와 음성데이터를 첨부하는 다임 레코드로 구성된다. 첫 번째 SOAP 메시지 형식으로 구성된 다임 레코드는 SOAP 헤더와 SOAP 바디를 포함

한다. SOAP Header에는 2.2절에서 구성한 사용자정보와 요청시간 제어를 정의하고 SOAP Body에는 음성인식 웹 서비스에서 제공되는 Rec()메서드를 호출하였다. 이와 같이 구성된 SOAP 메시지는 DimeAttachment 객체를 이용하여 다임 레코드에 첨부한다. 두 번째 다임 레코드에는 3.1 절에서 녹음한 음성데이터를 첨부한다. 본 연구에서는 음성 데이터를 다임 레코드에 첨부하기 위해 음성 데이터 파일을 바이트 배열로 변환하고 DimeAttachment 객체를 이용하였다. 이 때 파일포맷형식은 audio/wav로 지정한다.

이러한 다임 레코드들은 하나의 다임 메시지를 이루어 웹 서비스를 요청한다. 생성된 DIME 메시지를 웹 서비스로 전송하기 위해 HttpRequest 객체를 사용한다.

2.4 웹 서비스 응답처리

웹 서비스 요청이 성공적으로 이루어지면 웹 서비스 프로바이더는 인식결과와 사용자명을 SOAP 응답 메시지로 반환한다. 클라이언트는 HttpResponse 객체를 이용하여 응답메시지 정보를 수신 받는다. 응답메시지는 Xml 문서 형태로 구성되어 있다. 클라이언트는 전달받은 Xml 문서에서 결과 값을 포함하고 있는 RecResult 엘리먼트 값을 읽고 텍스트로 변환하여 PDA 화면에 사용자명과 인식결과를 나타낸다.

3. 웹 서비스 프로바이더

웹 서비스 프로바이더는 웹 서비스 사용자를 인증하고 응답 시간을 제어하고 음성데이터를 인식하여 그 결과와 사용자명을 반환한다. 웹 서비스 프로바이더에서 제공하는 웹 메서드인 Rec()은 먼저 HttpContext.RequestContext 객체를 사용하여 모바일 클라이언트가 요청한 메시지를 수신 받는다.

3.1 응답 시간 제어

웹 서비스 프로바이더는 모바일 클라이언트에게 SOAP 메시지 형태로 응답한다. 응답메시지는 60초 이내에 완료되게 지정하여 그 이후에 도착한 메시지는 오류를 반환한다.

3.2 인증처리 및 결과 반환

웹 서비스 프로바이더는 허가되지 않는 사용자로부터 웹 서비스를 보호하기 위해 인증을 필요로 한다. 본 연구에서는 사용자를 인증하기 위하여 WS-Security의 <UsernameToken>을 이용하였다.

<UsernameToken>은 SOAP 메시지의 헤더로 들어오는 사용자명과 패스워드 확인이 가능하다.

본 연구에서 구현한 인증처리 및 결과 값 반환 과정은 다음과 같이 이루어진다. 전달 받은 SOAP 메시지 에서 보안토큰을 열거하고 UsernameToken이 있는지 확인한다. UsernameToken의 패스워드 확인은 암호 공급자인 IPasswordProvider 인터페이스를 통해 검사한다. 인터페이스의 GetPassword 함수는 UsernameToken의 사용자명과 SQL데이터베이스의 사용자명을 비교하여 매핑되는 패스워드를 반환한다. 사용자 패스워드와 DB의 패스워드가 일치하면 Microsoft.Web.Services.Dime.DimeAttachment 객체를 이용하여 녹음한 음성데이터를 수신 받고 그렇지 않으면 오류를 반환한다. 별도로 제공되는 음성인식기 [6]는 수신 받은 음성데이터를 매개변수로 전달받아 인식과정을 수행하고 인식결과 값을 반환한다. 마지막으로 웹 서비스 프로바이더는 인식결과와 사용자명을 SOAP 응답 문자열로 작성하여 모바일 클라이언트에게 전송한다.

4. 실험 및 결과

4.1 실험환경

모바일 클라이언트 응용프로그램 개발은 Pocket PC 2002 OS와 .NET기반의 Microsoft Compact Framework를 이용하였다. 웹 서비스 프로바이더는 Microsoft .NET Framework 1.1과 Microsoft WSE 1.0을 이용하여 구현하였다.

음성인식 시스템은 기존의 연구에서 구현된 것으로 0~9까지의 숫자를 인식한다. 음성녹음 환경은 샘플링 값을 11kHz로, 비트 샘플링수는 16Bit, 모노로 설정하였다. 실험은 HP IPAQ 5420기종을 이용하였다.

4.2 실험 및 분석

본 연구에서 구현한 모바일 환경에서 인증을 지닌 음성인식 웹 서비스 클라이언트의 성능을 분석하기 위해 다음과 같은 실험을 하였다. 사용자는 그림4의 모바일 클라이언트 인터페이스 화면을 통해 회원 가입된 사용자명과 패스워드 정보를 입력한 후 음성을 녹음하여 웹 서비스를 요청한다. 요청 결과는 그림 5와 같이 사용자가 인증되고 클라이언트에서 녹음된 음성이 인식되어 텍스트로 변환된 결과를 알 수 있다. 이에 따라 모바일 클라이언트의 웹 서비스 요청과 인증기능을 지닌 음성인식 웹 서비스 프로바이더가 정상적으로 동작되고 있음을 확인 할 수 있다.

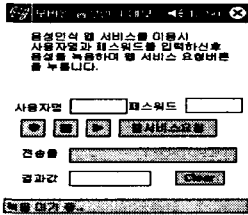


그림 4. 모바일 클라이언트 인터페이스

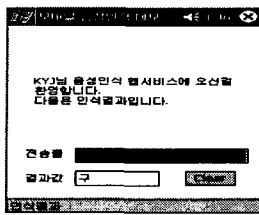


그림 5. 사용자인증 및 음성인식 결과

그림 6은 모바일 클라이언트가 DIME방식으로 웹 서비스 프로바이더에게 요청한 메시지이다. 첫 번째 <DimeRecord> 엘리먼트는 SOAP 메시지를 첨부하고 SOAP 메시지에 대한 정보를 나타내고 있다. 두 번째 <DimeRecord> 엘리먼트는 녹음된 음성데이터를 첨부하고 음성데이터에 관한 정보를 나타내고 있다.

```
<DimePayload>
<DimeRecord >
<Recordinfo Version="1" MB="1" ME="0" CF="0"
IDLength="41" />
<Typefield TNF="2" TypeLength="41" />
<Options O="0" OptionLength="0" />
<Datalength length="738" />
<ID value="uid:79c451d3-da11-4272-86e7-54d029aed7ee" />
<Type value="http://schemas.xmlsoap.org/soap/envelope/" />
</DimeRecord>
<DimeRecord >
<Recordinfo Version="1" MB="0" ME="1" CF="0"
IDLength="41" />
<Typefield TNF="1" TypeLength="9" />
<Options O="0" OptionLength="0" />
<Datalength length="21044" />
<ID value="uid:f47a625f-cfdb-4e7c-87c4-e27d97768ddd" />
<Type value="audio/wav" />
</DimeRecord>
</DimePayload>
```

그림 6. 모바일 클라이언트의 DIME 요청 메시지

그림 7은 웹 서비스 프로바이더가 음성 인식시스템을 수행한 후 생성된 결과와 사용자명을 클라이언트로 응답한 SOAP 메시지의 일부이다. 결과 값은 SOAP Body 안에 <RecResult> 엘리먼트에 포함하고 있다.

```
<soap:Body>
<RecResponse
xmlns="http://203.230.96.148/RecWebService">
<RecResult>KYJ-구</RecResult>
</RecResponse>
</soap:Body>
```

그림7. 웹 서비스 응답 SOAP 메시지

5. 결론

무선 인터넷 단말기의 저장 공간의 제약성과 입력 기능의 불편함을 개선하기 위해 본 논문에서는 별도

로 제공되는 음성인식시스템을 웹 서비스로 요청하여 음성인식이 가능한 PDA 클라이언트 시스템을 구현하였다. 따라서 PDA 클라이언트는 시간과 비용을 투자하여 음성인식 시스템을 구축하거나 설치할 필요 없이 언제 어디서나 웹 서비스를 요청함으로써 음성인식 서비스를 이용할 수 있다. 클라이언트는 웹 서비스 요청 시 무선인터넷을 사용한다는 점을 고려하여 음성데이터를 DIME방식으로 전송하므로 빠른 응답결과를 받을 수 있다. 또한 본 연구에서 구현한 웹 서비스 프로바이더는 허가되지 않은 사용자로부터 웹 서비스의 보호와 재공격을 방지하고 음성인식결과와 사용자명을 반환하는 서비스를 제공하였다.

향후에는 본 연구를 통해 PDA의 각종 프로그램 실행·인터넷 접속·웹 서핑 등을 음성으로 작동할 수 있는 인터페이스를 구현하고 모바일 환경에서 보다 안전한 웹 서비스 보호를 위해 메시지 기밀성, 디지털 서명 그리고 무결성 등을 고려하여 웹 서비스 보안 알고리즘을 연구하고자 한다.

[참고문헌]

- [1]윤세미, 조익성, 임재홍, “모바일 환경에서 vCard 서비스를 위한 웹 에이전트의 설계 및 구현”, 제9-D 권 제3호, 2002.
- [2]하안돌이, “VoiceXML을 이용한 모바일 음성인식 시스템 구현”, 동신대학교 석사학위 논문, 2002
- [3]Web Services Security: Username Token Profile ,<http://www.oasis-open.org/committees/download.php/3154/WSS-Username-04-081103-merged.pdf>,2003.08
- [4]Specification: Direct Internet Message Encapsulation (DIME), <http://www-106.ibm.com/developerworks/webservices/library/ws-dime/?dwzone=webservices#DATA>, 2002.06
- [5]Web Services Enhancements를 사용하여 SOAP 메시지 암호화, <http://www.microsoft.com/korea/msdn/library/dnwebsrv/html/wseencryption.asp>
- [6]오신영, “웹 서비스를 위한 음성인식시스템 구현”, 한밭대학교, 석사학위논문, 2002.
- [8]casey chesnut, WSE and the Compact Freamwork , “<http://www.learnmobile.net/MobileClient/Tutorials/cf/WSE/>”, 2002.