

콘텐츠 배급을 위한 RSS 기반의 VoiceXML 다이얼로그 시스템

권 형 준[†] · 김 정 현^{**} · 이 현 구^{***} · 홍 광 석^{****}

요 약

본 논문은 갱신이 잦은 콘텐츠의 배급 및 구독을 위해 등장한 시맨틱 웹의 대표적인 기술인 RSS(RDF Site Summary or Really Simple Syndication)와 인터넷에 존재하는 정보를 음성으로 제공하기 위해 제안된 XML 규격의 W3C 표준 마크업 언어인 VoiceXML을 결합한 시스템의 프로토타입을 제시한다. 제안하는 시스템은 인터넷 기반으로 제공되는 콘텐츠를 유무선 전화망을 통해 음성인식 및 합성기술로 제공하기 위한 것으로서, RSS 서비스를 제공하기 위한 기존의 구성을 수정하지 않고 갱신된 콘텐츠의 구독이 손쉬운 RSS의 장점을 VoiceXML에 적용할 수 있는 특징이 있다. 정보 요청자 측면에서는 RSS로 제공되는 콘텐츠 탐색에 인터넷 환경이 아닌 유무선 전화망을 이용하므로 시간상의 제약을 줄일 수 있으며, 정보 제공자 측면에서는 음성인식 및 합성기술을 이용한 콘텐츠의 배급에 최신 정보의 제공을 위한 별도의 구성요소를 필요로 하지 않는 장점이 있다. 제안하는 시스템의 예로 뉴스 서비스를 구현한 결과, 실제 콘텐츠를 구독하고 탐색할 때에 응답 시간과 음성 인식 측면에서 장애 없이 RSS Feed를 이용해 비주요 환경으로 제공되는 내용과 동일한 내용을 음성으로 제공받을 수 있었다.

키워드 : 보이스엑스엠엘, 알에스에스, 시맨틱웹, 웹서비스

VoiceXML Dialog System Based on RSS for Contents Syndication

Hyeong-Joon Kwon[†] · Jung-Hyun Kim^{**} · Hyon-Gu Lee^{***} · Kwang-Seok Hong^{****}

ABSTRACT

This paper suggests prototype of dialog system combining VXML(VoiceXML) that is the W3C's standard XML format for specifying interactive voice dialogues between human and computer, and RSS(RDF Site Summary or Really Simple Syndication) that is representative technology of semantic web for syndication and subscription of updated web-contents. Merits of the proposed system are as following: 1) It is a new method that recognize spoken contents using wire and wireless telephone networks and then provide contents to user via STT(Speech-to-Text) and TTS(Text-to-Speech) instead of traditional method using web only, 2) It can apply advantage of RSS that subscription of updated contents is converted to VXML without modifying traditional method to provide RSS service, 3) In terms of users, it can reduce restriction on time-space in search of contents that is provided by RSS because it uses wire and wireless telephone networks, not internet environment. 4) In terms of information provider, it does not need special component for syndication of the newest contents using speech recognition and synthesis technology. We implemented a news service system using VXML and RSS for performance evaluation of the proposed system. In experiment results, we estimated the response time and the speech recognition rate in subscription and search of actuality contents, and confirmed that the proposed system can provide contents those are provided using RSS Feed.

Key Words : VoiceXML, RSS, Semantic Web, Web Service

1. 서 론

웹 서비스란 재사용이 가능하고 새로운 서비스를 위해 조합될 수 있는 특정한 작업을 수행하기 위한 웹 기반의 기술

로서, 현재의 웹 서비스는 특정 작업을 수행하는 기능에 대한 구문적 명세를 제공하나 정형화된 의미론을 제공하지는 못하며, 단지 호출 인터페이스와 그 서비스가 어떤 기관에서 개발하였는지, 어떤 작업을 하는지에 대한 메타데이터만을 제공한다. 차세대 웹으로 주목받는 시맨틱 웹은 현재의 웹 서비스와는 다르게 웹 서비스의 구문적 명세와 함께 의미론적 명세를 제공하며, 개방형 환경에서 자동화된 웹 서비스의 조합과 호출을 수행하고, 웹 서비스들의 상호작용을 위한 환경을 제공한다[1]. 이러한 시맨틱 웹의 개념에 따라

※ 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음, IITA-2006-(C1090-0603-0046).
[†] 준 회 원 : 성균관대학교 대학원 정보통신공학부 석사과정
^{**} 준 회 원 : 성균관대학교 대학원 정보통신공학부 박사수료
^{***} 정 회 원 : 서울대학교 정보기술계열 정보통신전공 교수
^{****} 중 심 회 원 : 성균관대학교 정보통신공학부 교수
 논문접수 : 2006년 9월 6일, 심사완료 : 2007년 12월 15일

온톨로지 언어인 RDF(Resource Description Framework)가 연구되고 RDF를 기반으로 등장한 RSS는 웹 서비스를 제공하기 위한 방법을 방문과 수집이 아닌 구독과 편집의 형태로 바꾸었다. 사용자가 원하는 특정한 콘텐츠를 제공받기 위해 인터넷의 이곳저곳을 탐색하여 살펴야 하는 번거로운 작업을 없애주고, 콘텐츠의 갱신 확인을 위하여 일일히 웹 페이지를 방문해 확인하지 않아도 언제나 최신의 정보를 RSS를 통해 구독할 수 있게 되었다. 이러한 연구의 결과는 시맨틱 웹의 개념을 정확히 적용하였고, 그 기능성과 유용함을 인정받고 있다[2, 3].

웹 서비스에 대한 연구 중 다른 한편으로 비주얼 환경인 웹의 특징을 탈피하여 보이스 환경으로 웹 서비스를 제공하는 것에 관한 연구가 지속되고 있는데, 웹 서비스를 보이스 환경으로 탈바꿈하기 위한 노력은 세계적인 기업들의 연구 결과로 VoiceXML의 표준화를 이루었다[4]. VoiceXML은 인터넷 환경으로 접근하기 어려워 웹 페이지를 탐색할 수 없는 경우 유선 및 무선 전화로 웹 상의 콘텐츠를 제공하기 위해 제안된 XML의 확장형으로서, 자동화된 서비스를 구성하여 실시간 전화와 같은 유용함 혹은 더 뛰어난 효과를 기대할 수 있어서 서버 측면의 스크립트 언어와 혼용하여 전자상거래나 원격교육 등으로의 적용과 VoIP와 결합된 멀티모달 정보 제공 시스템에 관한 연구들이 진행중이다[5][6].

본 논문에서는 RSS로 배포되는 콘텐츠와 동일한 내용을 VoiceXML로 개발한 어플리케이션을 통해 음성으로 제공하기 위한 시스템의 프로토타입을 제시하고, 그 예로 뉴스 서비스를 구현하여 음성 인식을 및 정보 제공시의 응답 속도를 측정하였다. 제안하는 시스템은 기존 RSS 서비스를 위한 구성을 변경하지 않으며, 언제나 최신의 콘텐츠를 제공하는 RSS를 기반으로 하기 때문에 음성으로 제공되는 콘텐츠가 최근에 갱신된 내용을 유지함에 있어서 그를 위한 별도의 모듈을 필요로 하지 않는다.

본 논문의 구성은 다음과 같다. 2장에서는 RSS와 VoiceXML에 관한 전반적인 내용을 소개하고, 3장에서는 제안하는 시스템의 구조와 특징, 주요 요소와 처리 과정을 설명한다. 4장에서는 실험 방법 및 결과를 기술하고, 마지막으로 5장에서 결론을 맺는다.

2. 관련 연구

2.1 RSS

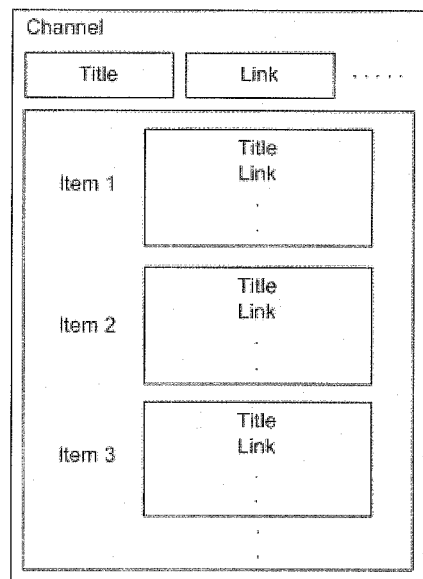
RSS 개념은 넷스케이프(Netscape)사에서 운영하던 인터넷 최대의 포털 사이트인 넷센터(NetCenter)에서 유명 신문의 기사를 제공하기 위하여 등장하였다. 1995년 넷스케이프사가 메타데이터의 기술방법을 XML과 결합시켜 표준 데이터 정의 언어로서 W3C(World Wide Web Consortium)에 표준화 제안된 MCF(Meta Contents Framework)를 거쳐 RDF(Resource Description Framework)와 CDF(Channel Definition Format)로 발전하였으며, 이어서 RSS(RDF Site Summary) 버전 0.9를 제안하게 되는데, 넷스케이프사는 그 이후의 연구를 진행할 수 없게 되면서 두 개의 개발 그룹이

형성되어 연구를 진행하여 현재는 RSS-DEV Working Group의 RSS(Rich Site Summary) 버전 0.92, 1.0과 Userland의 RSS(Really Simple Syndication) 버전 2.0 등 두 종류의 RSS가 제안되어 이용되고 있다.

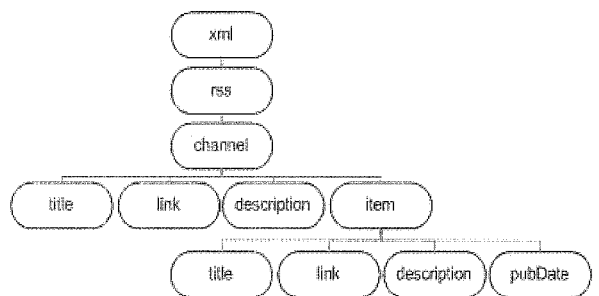
2.1.1 기본 구성

RSS 서비스를 제공할 때에는 RSS 문서에 콘텐츠의 내용을 담아 제공하게 되고, 서비스를 받고자 하는 사용자는 RSS 리더 프로그램을 PC에 설치하여 RSS 문서의 URL을 등록하여 사용한다. 이 때 전달되는 문서를 RSS Feed라 칭하며 문서의 문법 규약은 버전별로 조금씩 다르다. (그림 1)은 RSS 피드의 기본 구성을 나타낸다[7, 8].

본 논문에서 제안하는 시스템에 적용된 RSS 규격은 Userland사의 RSS 2.0으로서, 개발자인 데이브 와이너(Dave Winer)가 미국 하버드대학교의 특별연구원이 되면서 RSS 2.0의 소유권이 하버드대학교로 이전되었다. (그림 2)에 RSS 2.0의 필수 및 주요 엘리먼트를 트리 형태로 나타내었다[9].



(그림 1) RSS의 기본 구성



(그림 2) RSS 2.0의 필수 및 주요 엘리먼트

2.1.2 표준화 및 서비스 현황

RSS의 확산과 더불어 콘텐츠 배포의 중요성에 대한 인식

과 새로운 기능 및 표준화의 필요성이 대두되어 그에 대한 노력이 진행되었으나, 사실상 RSS 규격의 단일화가 어렵다는 결론이 도출되어 새로운 표준화를 위해 컴퓨터 서적 출판사인 O'reilly Media 및 미국 지식 검색 사이트 운영업체인 About의 주도로 Atom 프로젝트가 진행되고 있다[10]. 현재 Atom의 버전은 0.3으로서, RSS 2.0과 Atom이 미래 RSS 서비스를 이끌어 갈 것으로 예상된다.

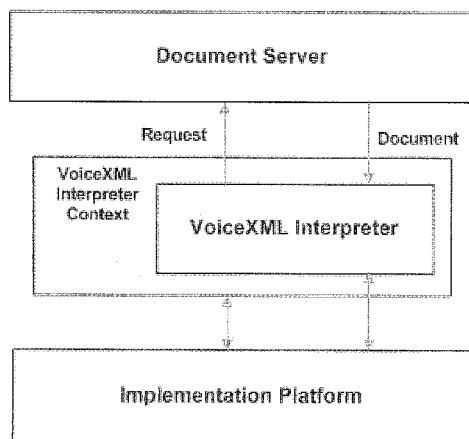
2.2 VoiceXML

VoiceXML은 Motorola, AT&T, IBM, Lucent Technology 등 정보통신 분야의 4개 대기업에 의해 설립된 VoiceXML Forum에서 제안한 것으로서, ASR(Automatic Speech Recognition) 및 DTMF(Dual Tone Multi Frequency)를 이용한 입력과 오디오 및 TTS(Text to Speech) 등을 이용한 출력으로 음성 기반의 사용자 인터페이스 개발을 위해 고안된 XML의 확장형 마크업 언어이다. 1999년 VoiceXML Forum에서 버전 0.9를 발표한 후, 2000년 3월에 버전 0.9를 크게 보완한 버전 1.0을 제안해 W3C의 승인을 받았고, 현재 표준화된 최신 버전은 2004년 3월에 승인된 2.0이다[11].

2.2.1 VoiceXML의 실행 구조

VoiceXML로 개발된 음성 어플리케이션은 VoiceXML Gateway를 통해 전화와 인터넷 및 웹 사이트 콘텐츠에 연결을 할 수 있게 하며, 보이스 유저 인터페이스를 실현한다. 또한, XML 기반의 확장형 언어이므로 ASP, JSP, PHP 등의 서버 측면의 스크립트 언어와 혼용하여 VoiceXML 문서를 사용자의 요구에 따라 적절히 생성할 수 있어서 정적인 시나리오와 함께 동적인 시나리오를 갖는 음성 어플리케이션을 구현할 수 있다[12]. VoiceXML로 작성된 문서 및 음성 어플리케이션은 (그림 3)에 도시한 바와 같이 VoiceXML 해석기에 의해 실행된다[11].

Document Server는 VoiceXML Interpreter인 클라이언트 어플리케이션으로부터 요청을 받는다. 서버는 VXML Interpreter에게 응답으로서 VoiceXML Document를 제공한다. VoiceXML Interpreter Context는 VoiceXML Interpreter



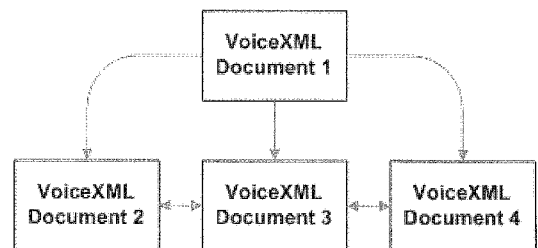
(그림 3) VoiceXML 아키텍처 모델

와 동등하게 사용자 입력을 모니터링 한다. Implementation Platform은 Document Server와 VoiceXML Interpreter의 상호 작용으로 이루어진 결과물을 최종적으로 사용자에게 전달한다.

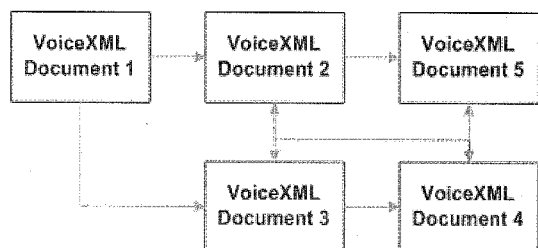
2.2.2 VoiceXML의 문서 구조

VoiceXML의 문서 구조는 두 가지 기준을 두고 분류할 수 있는데, 첫째로는 시나리오 구성 방식을 기준으로 분류된다. 단일 문서(Single Document Application)와 다중 문서(Multi Document Application)로 분류되는데, 단일 문서는 서비스 시나리오의 구성이 하나의 문서로 이루어져 있는 형태이며, 다중 문서는 (그림 4)에 도시한 모습과 같이 루트 문서에 여러 개의 하위 문서들이 서로 정보를 공유하여 하나의 시나리오를 구성하는 형태이다.

둘째로 실행 순서에 따라 분류하면 Computer Directed Forms과 Mixed Initiative Forms로 분류되는데, Computer Directed Forms는 다이얼로그에 있는 Form Item이 미리 정의된 순서에 따라 실행되는 형태이며, Mixed Initiative Forms는 컴퓨터와 사용자 모두의 진행을 능동적으로 변경할 수 있는 다이얼로그 모델로서, Form 레벨의 Grammar를 필요로 한다. (그림 5)에 도시한 바와 같이 사용자의 응답에 따라 다이얼로그가 실행되는 순서가 변경될 수 있으며, 여러 개의 다이얼로그 Form Item을 한 번에 처리할 수 있다[11].



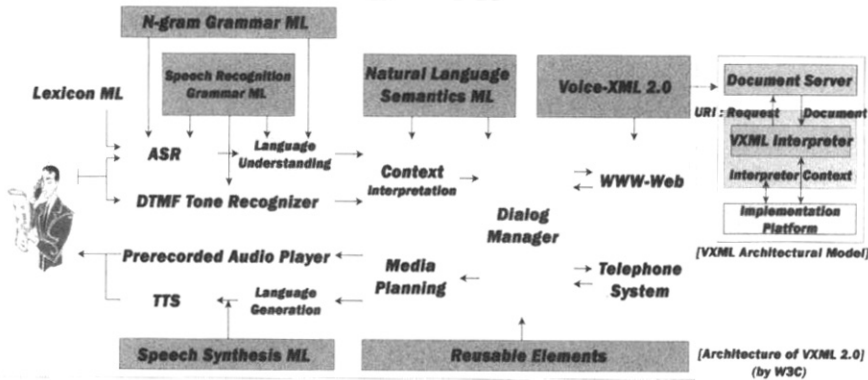
(그림 4) Multi Document Application



(그림 5) Mixed Initiative Forms

2.2.3 음성 어플리케이션 서비스 아키텍처

(그림 6)은 W3C의 표준화에 따른 VoiceXML 2.0을 이용해 구축한 음성 어플리케이션의 전체적인 서비스 아키텍처를 나타낸다[11]. 본 논문에서는 KT 서비스 개발연구소에서 제공하는 VoiceXML 기반의 음성 어플리케이션 통합 솔루션인 휴보이스(HUVOIS)를 사용한다. 휴보이스는 VoiceXML 기반의 음성 어플리케이션을 개발하고 서비스할 수 있도록 ASR, TTS, VoiceXML Interpreter등을 내장하고 있다[13].



(그림 6) VoiceXML 2.0 어플리케이션 서비스 아키텍처

3. RSS 기반의 VoiceXML 다이얼로그 시스템

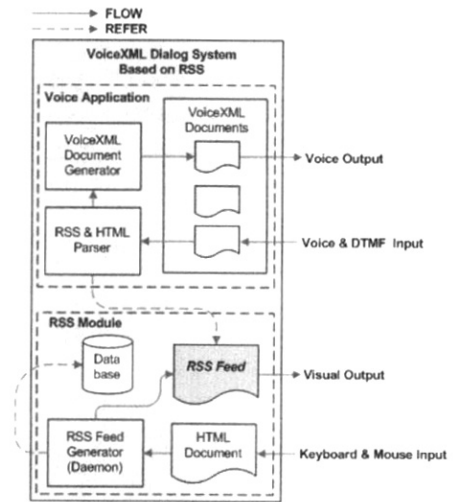
XML 문서를 여러 형태의 문서로 변환하기 위해 등장한 XSLT(Extensible Style Language Transformation)는 XML 문서간의 상호 변환도 가능하기 때문에 RSS 피드와 VoiceXML 문서를 서로 변환할 수 있다[14]. 하지만 제안하는 시스템은 XSLT를 사용하지 않는다. 그 이유는 RSS 피드가 완전한 콘텐츠를 담고 있지 않은 경우의 음성 서비스를 위해서 HTML 파서를 이용해야 하기 때문에 서버 측면의 스크립트 언어로 모든 것을 처리하며, 본 장에서 그 과정을 보인다. RSS 피드에 콘텐츠의 모든 것을 기록하면 HTML 파서가 필요하지 않으나 콘텐츠의 내용을 요약하여 전달하는 RSS 서비스의 특징을 그대로 유지하면서 보이스 환경의 서비스를 추가하는 것에 목적을 둔다.

3.1 시스템의 구조 및 특징

제안하는 시스템은 RSS를 기반으로 한 VoiceXML 다이얼로그 시스템으로, 기존 RSS 서비스의 구성을 수정하지 않고 VoiceXML 음성 어플리케이션을 첨가하여 RSS 피드로 제공되는 콘텐츠와 동일한 내용을 전화망을 통해 유선 및 무선 전화로 제공하는 기능을 수행한다. 하나의 시스템으로 비주얼 환경의 RSS 서비스와 보이스 환경의 VoiceXML 음성 서비스를 동시에 제공할 수 있는 특징이 있고, RSS Feed를 이용하기 때문에 VoiceXML 음성 어플리케이션으로 제공될 정보를 위한 별도의 데이터베이스를 필요로 하지 않는 장점이 있다.

(그림 7)은 제안하는 시스템의 주요 구성요소들 사이의 관계를 나타낸다. 본 시스템은 비주얼 환경 및 보이스 환경 기반의 접근이 가능하다. 비주얼 환경에서는 콘텐츠 탐색을 위해 RSS Reader 프로그램을 PC에 설치하여 연결하며, 보이스 환경에서는 유무선 전화로 연결한다. PC의 RSS Reader를 통한 콘텐츠의 탐색은 기존에 공개된 방법과 같다.

유무선 전화로 본 시스템을 이용하는 과정은 다음과 같다. 제안하는 시스템에 연결을 시도하여 세션이 생성되면 정보 요청자는 음성 혹은 DTMF로 입력을 수행하고, 입력을 받아들인 VoiceXML 문서는 비주얼 환경에서 RSS 서



(그림 7) 제안하는 시스템의 구성요소들의 관계

스 제공을 위해 생성된 RSS 피드를 파싱한다. 파싱한 결과가 콘텐츠의 본문이 하이퍼링크 주소인 경우는 해당 주소의 HTML 문서를 획득하여 태그를 삭제하여 실제 콘텐츠만을 얻는다. 이렇게 정보 요청자가 원하는 정보를 추출한 후, VoiceXML 문서의 동적인 생성을 수행하여 TTS를 통한 보이스 출력으로 정보 요청자에게 콘텐츠를 전달한다. 이러한 과정을 거치며 기존 RSS 서비스를 음성으로 제공할 수 있게 된다.

음성 서비스의 제공을 위해, RSS Feed를 생성할 목적으로 사용되는 데이터베이스에 직접 액세스할 수 있지만 비주얼 환경 및 보이스 환경의 서비스를 동시 제공하는 본 시스템에서는 데이터베이스로의 접근을 최대한 줄일 필요가 있다고 판단하여, 상대적으로 부하가 적도록 RSS Feed에서 데이터를 추출해 보이스 환경의 서비스에 이용한다. 이와 같은 특징은 최신 정보를 제공하는 RSS 서비스의 장점을 음성 서비스에 적용할 수 있게 만든다.

3.2 주요 요소

제안하는 시스템은 Mixed Initiative Forms 다이얼로그 구조를 기반으로 서버 측면의 스크립트 언어를 이용해 동적

으로 동작하도록 개발되었다. 또한, 기존의 RSS를 위한 웹 서비스 및 VoiceXML 음성 어플리케이션에 포함되지 않는 모듈인 XML파서와 HTML파서가 새롭게 추가되었다. 본 절에서는 RSS 피드를 생성하는 방법을 소개하고, 아울러 HTML과 XML문서를 연결해주는 프로그래밍적 인터페이스인 DOM(Document Object Model)을 이용한 RSS 피드를 파싱하는 모듈, HTML에서 원하는 내용을 추출하는 모듈 및 적용한 알고리즘들을 소개한다. 본 알고리즘들은 RSS 2.0을 기준으로 작성되었으며, 다른 버전의 RSS 에서는 다소 변경이 필요하다.

3.2.1 RSS 피드 생성기

RSS 피드 생성기는 ASP, JSP, PHP등의 서버 측면의 스크립트 언어로 개발하는 것이 적당하다. RSS 피드의 각 엘리먼트에 담을 내용을 가져오기 위해 데이터베이스에 질의하여 *item* 엘리먼트에 넣을 질의 결과를 얻고, RSS 버전에 맞는 엘리먼트들과 함께 각각의 레코드가 포함하고 있는 필드의 내용을 텍스트형식으로 저장한다. *xml*, *rss*, *channel*, *title*, *link*, *description* 엘리먼트는 특별한 작업 없이 정형화된 내용으로 채워 넣으며, *item* 엘리먼트는 반복문을 사용하여 알고리즘 1과 같은 방법으로 코드를 작성한다. 알고리즘 1은 RSS 2.0이 지원하는 엘리먼트들의 이름과 *item* 엘리먼트가 포함할 내용이 담긴 데이터베이스 테이블의 컬럼명이 같다는 것과 가장 마지막 필드의 이름이 *pubDate* 임을 가정하고 작성되었다.

[알고리즘 1] RSS 피드 생성 알고리즘

```

for (each record in resultSet)
  for (each field in currentRecord)
    if ([column_name] != "pubDate")
      writeline "<[column_name]>"
      writeline "<[column_content]>"
      writeline "</[column_name]>"
    endif
  endfor
endfor

```

[알고리즘 1]은 데이터베이스에 질의한 결과에서 레코드가 있는지 확인하고, 현재 레코드의 각 필드를 순회하며 엘리먼트명과 함께 엘리먼트가 가진 내용을 저장한다. 현재 레코드의 가장 마지막 필드의 내용을 저장한 후에는 데이터베이스에 질의한 결과에서 다음 레코드가 있는지 확인하는 작업부터 다시 반복하고, 마지막 레코드의 마지막 필드까지 순회한 후 실행을 종료한다.

3.2.2 XML 파서

XML은 각 엘리먼트의 내용을 읽어낼 수 있도록 하기 위해 문서를 파싱하는 인터페이스를 제공한다. 본 논문에서는 XML 규격의 RSS 피드를 DOM을 활용해서 파싱하여 각각의 엘리먼트가 가진 내용을 이용해 VoiceXML 문서를 생성한다. [알고리즘 2]는 RSS 피드를 파싱하여 각 엘리먼트의

내용을 가져오는 방법이다.

[알고리즘 2]는 RSS 피드의 내용을 줄 단위로 탐색하여 엘리먼트명을 확인하고, 필요한 엘리먼트인 경우, 그 엘리먼트명과 엘리먼트가 가진 내용을 저장한다. 이러한 방법으로 RSS 피드의 모든 엘리먼트들을 탐색하고 추출하는 파싱 방법은 경우에 따라 효율적이지 못할 수 있다. (그림 8)은 조건선택에서 제공하는 인기기사 RSS 피드에 알고리즘 2를 실행한 후의 결과 화면이다.

[알고리즘 2] DOM을 이용한 XML 파싱 알고리즘

```

xmlDomVariable ← RSS_FEED_URL
for (each element in xmlDomVariable)
  if ([element_name] = requiredElements)
    writeline "[element_name]"
    writeline "[element_content]"
  endif
endfor

```



(그림 8) 알고리즘 2를 실행한 후의 결과

[알고리즘 3] DOM을 이용한 XML 파싱 알고리즘 2

```

xmlDomVariable ← RSS_FEED_URL
currentItemCount ← 0
maxItemCount ← Output Number

for (each element in xmlDomVariable)
  if ([element_name] = requiredElements)
    if ([element_name] = [item] and
        currentItemCount < maxItemCount)
      currentItemCount ++
      writeline "[element_name]"
      writeline "[element_content]"
    endif
  endif
endfor

```

[알고리즘 3]은 RSS 피드가 *item* 엘리먼트를 많이 가지고 있을 경우 파싱에 소요하는 시간의 증가로 인하여 VoiceXML 문서의 생성이 늦어져 음성으로 콘텐츠를 제공하는 데 있어서

지장을 초래할 수 있으므로, 알고리즘 2를 개량하여 원하는 개수만큼 *item* 엘리먼트를 추출하는 방법에 관한 것이다.

제안하는 시스템은 데이터베이스로 접근하지 않는 대신 앞서 설명한 과정을 통해 RSS 피드를 파싱하여 데이터를 얻는다. 그러나 RSS 피드를 서비스의 요청이 생길 때 마다 생성하면 직접 데이터베이스에 접근하는 것 보다 처리 과정 및 소요 시간에서 효율적이지 못하다. 그러므로 일정 시간이 지날 때 마다 자동으로 RSS 피드를 생성하도록 개발하고 음성 서비스를 제공하는 시점에서 최근에 생성된 RSS 피드를 파싱하는 것이 바람직하다.

3.2.3 HTML 파서

제안하는 시스템의 새로운 구성요소 중 하나인 HTML 파서는 RSS 피드의 *item* 엘리먼트의 하위 엘리먼트인 *description* 엘리먼트의 내용이 완전하지 않을 시에 사용된다. *description* 엘리먼트의 내용이 많으면 RSS 피드의 문서의 용량이 방대해져서 서버의 트래픽을 불필요하게 증가시키기 때문에, 일반적으로 *description* 엘리먼트에 실제 내용의 일부분만 기록하며, *link* 엘리먼트에 콘텐츠의 전체가 담긴 하이퍼링크 주소를 기술한다. 제안하는 시스템에 포함된 HTML 파서는 하이퍼링크 문서에서 필요한 내용을 추출하는 역할을 수행한다. 알고리즘 4는 HTML 파서가 HTML 문서에서 내용을 추출하는 방법을 보인다.

[알고리즘 4]는 HTML 문서 안의 핵심 내용인 추출할 문자열의 시작점과 끝점을 나타내는 주석문이 첨가되어 있어야만 완전한 동작을 보장한다. HTML문서의 모든 텍스트를 메모리에 저장한 후, 시작점 문자열을 기준으로 전체 문자열을 두 부분으로 나누어 열거형 변수 혹은 배열형 변수에 저장한다. 다음으로, 저장된 열거형 변수 혹은 배열형 변수의 두번째 요소에서 끝점 문자열을 기준으로 두 부분으로 다시 나누어 각각 열거형 변수 혹은 배열형 변수에 저장하고, 첫번째 요소에서 HTML 태그를 삭제하여 원하는 내용을 추출한다. 결론적으로, 시작점과 끝점의 사이에 있는 텍스트 중 음성으로 제공하기 위해 필요한 내용만을 추출해낸다. (그림 9)는 알고리즘 4를 적용하여 조선일보에서 제공하는 RSS 피드의 콘텐츠에 수록된 하이퍼링크 주소가 가리키는 HTML 문서에서 음성으로 제공하기 위해 필요한 부분만

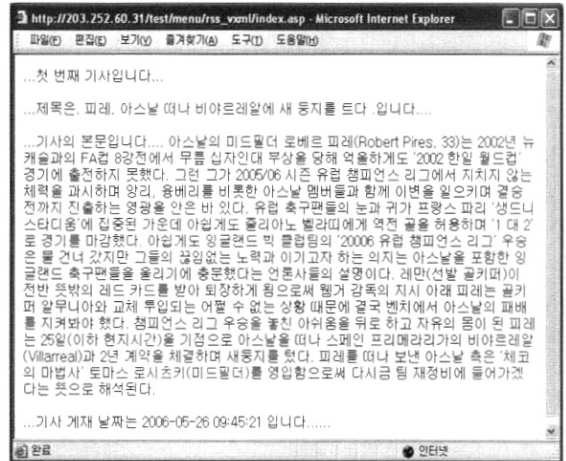
[알고리즘 4] 주석문을 이용한 HTML 문서 파싱 알고리즘

```

textString ← Text in html document.
stHeadString ← Insert startpoint string.
stTailString ← Insert endpoint string.

stringArray_1st ← Enumerate textString
                  with stHeadString
stringArray_2st ← Enumerate stringArray_1st(1)
                  with stTailString
stringResult ← Delete HTMLTAG in stringArray_2nd(0)

Return StringResult
    
```

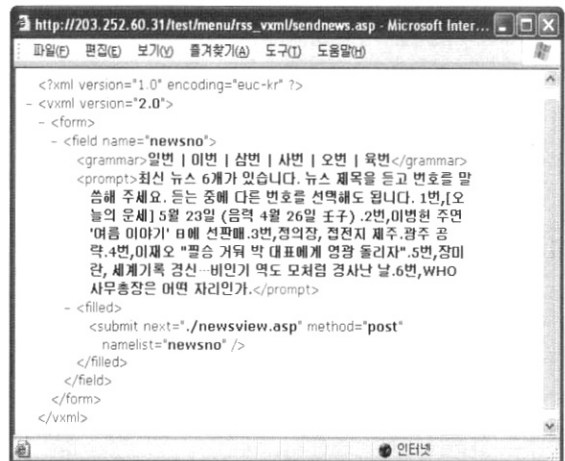


(그림 9) HTML 문서에서 필요한 내용만을 추출한 결과

을 추출한 결과 화면이다. 사용자가 요청한 정보가 담긴 HTML 문서에서 태그들이 제거되고 실제 콘텐츠만 추출되었음을 확인할 수 있다.

3.2.4 VoiceXML 문서 생성기

VoiceXML 문서 생성기는 RSS 피드를 파싱하고 필요한 경우 HTML 문서까지 파싱하여 실제 서비스할 내용들을 추출한 결과물을 가지고 VoiceXML 문서를 동적으로 생성하는 역할을 수행한다. RSS 피드 엘리먼트들은 각자 고유한 목적이 있으므로 그 목적에 맞추어 VoiceXML 엘리먼트에 삽입한다. (그림 10)은 동적으로 생성된 VoiceXML 문서의 예이다.



(그림 10) 동적으로 생성된 VoiceXML 문서의 예

4. 실험 및 결과

본 논문에서는 제안하는 시스템의 예로서 RSS 2.0 규약을 만족하는 조선닷컴의 조선일보 전체기사 RSS 피드 (<http://www.chosun.com/rss/rss.xml>)를 기반으로 뉴스 서비스를 구현하였다. 성능 평가 방법은 정보 제공시의 응답 속도와 음성 인식률을 측정하여 실제 서비스 제공 및 사용에 있어서 장애가 없는지 판단하였다.

4.1 동작 환경

서버의 중앙처리장치는 안정성을 위하여 듀얼프로세서를 탑재한 Intel Xeon을 사용하였고, 전화와 컴퓨터의 상호작용을 위한 CTI 보드는 한 번에 12채널을 수용할 수 있는 Intel Dialogic D/120JCT-LS를 사용하였으며, 음성인식 및 합성 엔진은 KT의 휴보이스를 사용한다. <표 1>은 제안한 서비스에 이용한 서버의 하드웨어 및 소프트웨어의 사양을 나타낸다.

<표 1> 서버 사양

항목	설명
CPU	Intel Xeon 3.0Ghz Dual
RAM	2GB
CTI Board	Intel Dialogic D/120JCT-LS PCI(12CH)
OS	Windows 2000 Professional Service Pack 4
IIS Ver.	5.0
ASR	KT Huvois
TTS	KT Huvois

4.2 서비스 시나리오

(그림 11)은 제안하는 시스템의 실험을 위해 구현된 어플리케이션 시나리오의 예이다. 차례대로 RSS 피드에 수록된 뉴스 기사들의 제목을 들려주고, 정보 요청자는 기사 제목에 매겨진 번호를 발성하여 원하는 뉴스 기사의 원문을 청취하는 시나리오이다.

System	안녕하세요? HCI 연구실의 RSS뉴스 서비스입니다. 뉴스의 제목을 듣고, 청취를 원하시는 기사에 매겨진 번호를 발성하시면 됩니다. 본 기사는, 조선닷컴 RSS 피드를 이용하였습니다.
System	이번. 브라질, 2007년부터 10년간 5%대 성장 지속 이번. 출국 명령 한국인 수백 명 카불 출발 삼번, 국내 5개 은행, LA 중앙은행 상대 제스.....
User	이번
System	이번을 선택하셨습니다. 기사의 제목은 출국 명령 한국인 수백 명 카불 출발입니다. 기사가 작성된 시각은 2006년 8월 5일 4시 46분 50초입니다.
System	한국 기독교인 약 400명이 아프가니스탄 정부의 출국 명령에 따라 4일 삼엄한 경비 속에 항공편으로 카불을 떠났다. 아프간 정부는 한국인 출국자들을
User	삼번
System	삼번을 선택하셨습니다. 기사의 제목은 국내 5개 은행, LA 중앙은행 상대 제스입니다.
System	국민, 외환은행 등 국내 5개 은행이 로스앤젤레스의 한국계 중앙은행을 상대로 거액의 손해배상 소송을 제기했다. 중앙은행 관계자는
User	종료
System	이용해 주셔서 감사합니다.

(그림 11) 서비스 시나리오

4.3 성능 평가

4.3.1 음성으로 콘텐츠 제공 시 응답 속도

앞에서 설명한 바와 같이, 제안하는 시스템에서 음성으로 콘텐츠를 제공하기 위해서는 여러 과정을 거치면서 많은 문자열 데이터의 처리와 함께 인터넷을 통한 정보의 전송이 이루어진다. 이러한 요소들이 서비스를 제공하는 것에 연결 지연 현상을 가져올 가능성이 있다고 판단되어, 음성으로 콘텐츠를 제공할 시의 응답 속도를 측정하였다. 인식 결과에 관한 정확도를 고려하지 않고 서버가 정보 요청자의 음성을 인식한 시

간부터 다시 정보 요청자에게 출력이 시작되는 시점의 시간차를 총 15회 반복하여 관찰한 결과를 <표 2>에 나타냈다.

<표 2> 응답속도 측정 결과

회차	응답속도	발성음
1	1초이내	일번
2	1초이내	이번
3	1초이내	삼번
4	1초이내	사번
5	1초이내	오번
6	2초이내	일번
7	1초이내	이번
8	1초이내	삼번
9	2초이내	사번
10	2초이내	오번
11	1초이내	일번
12	1초이내	이번
13	1초이내	삼번
14	1초이내	사번
15	2초이내	오번

정보 요청자가 발성한 음성이 인식된 시점은 ASR이 동작한 시간이 초 단위로 기록되는 휴보이스의 로그 데이터로 측정하였고, 응답이 이루어진 시점은 정보 요청자의 요청에 대한 응답으로서 TTS 발성음을 정보 요청자가 청취하기 시작하는 시점의 시간을 출력하여 시간차를 구하였다. 실험 결과, 최대 2초 이내로 응답함으로써 응답 속도로 인한 서비스 제공 및 이용에 문제가 없음을 확인하였다. 응답 시간을 더 줄이기 위해서는 서버의 소프트웨어와 하드웨어의 최적화가 관건이다.

4.3.2 인식을 측정

제안하는 시스템의 음성 인식 성공률 실험을 위하여, 음성 어플리케이션 서비스를 이용해 본 경험이 있는 피험자 5명과, 음성 어플리케이션 서비스를 이용해 본 경험이 없는 피험자 5명으로 이루어진 총 10명의 피험자로 일상 소음이 존재하는 업무용 사무실에서 유선 전화를 이용하여 1인당 10회 발성하여 인식률을 측정한 결과를 <표 3>과 <표 4>에 나타내었다.

<표 3> 음성 인식 경험자의 인식률

피험자	오인식수/시도횟수	인식률(%)
A	1 / 10	90
B	2 / 10	80
C	1 / 10	90
D	1 / 10	90
E	2 / 10	80

<표 4> 음성 인식 무경험자의 인식률

피험자	오인식수/시도횟수	인식률(%)
F	2 / 10	80
G	3 / 10	70
H	4 / 10	60
I	2 / 10	80
J	3 / 10	70

음성 어플리케이션 서비스를 경험해 본 피험자는 최소 80% 이상의 인식률을 보이며 서비스를 이용하는 데 있어서 큰 어려움이 없었으나, 경험해 보지 않은 피험자는 최소 60%, 최고 80%의 인식률을 보였다. 본 실험에서 도출된 결

과를 살펴보면, 무경험자들이 음성 어플리케이션을 이용하는 것에서 경험 미숙으로 인한 발성의 부정확함이 인식률의 저하를 초래하였다. 또한, 의도되지 않은 노이즈 및 음성 인식 엔진의 성능에 따른 필연적인 인식률의 저하로 인한 오인식을 제외하면 음성 인식 경험자가 서비스를 이용하는 것에는 장애가 없을 것이라 예상할 수 있다.

5. 결 론

본 논문에서는 인터넷상에 존재하는 콘텐츠의 갱신 사항을 즉시 전달하는 RSS 서비스의 RSS Feed를 활용하여 기존의 아키텍처를 수정하지 않으면서 VoiceXML 음성 어플리케이션을 통합한 시스템의 프로토타입을 제시하고 그 예로 뉴스 서비스를 구현하였다. 시스템의 신뢰성을 위해 음성 인식률 및 대용량 문자열 처리 시 문제점 발생 확인을 위해 정보 요청시의 응답 속도와 음성 인식률을 측정하였다. 실험 결과, 서비스 제공 및 이용에 문제가 없음을 확인하였다.

제안하는 시스템은 인터넷을 통한 비주얼 환경에서의 구동이 가능하면서도 인터넷 환경을 접하기 어려운 곳에서 RSS 파드로 제공되는 콘텐츠와 동일한 콘텐츠를 음성으로 제공받을 수 있도록 하여 시공간의 제약을 완화하였다. 또한, 보이스 및 비주얼 환경 기반의 서비스를 동시에 제공하면서, 데이터베이스의 접근 횟수를 줄여 서버의 부하를 최소화 하였다.

RSS는 시맨틱 웹 시대를 개척할 선두 기술이자 이미 대부분의 포털에서 지원하고 있어서 인터넷 기술에 특별한 관심을 두는 사람들이 먼저 활용하기 시작하여 점차 대중화되고 있는 서비스이다. 또한 VoiceXML로 개발된 음성 어플리케이션은 기존 음성 어플리케이션 시장을 대체할 유력한 대안으로 주목받고 있다. 향후 인터넷 시장을 이끌어갈 두 기술을 통합한 제안된 시스템이 웹 기술 및 아키텍처에 관한 연구의 발전과 더 좋은 웹 서비스를 구축하려는 사람들에게 도움이 될길 기대한다.

참 고 문 헌

[1] 김인철, 김현식, "시맨틱 웹 서비스 기술 동향", 한국인터넷 정보학회지, 제6권 제2호, pp.7-15, 2005.
 [2] 조부현, 민영근, 이복주, "시맨틱 웹을 이용한 웹 변경 탐지 시스템", 정보처리학회논문지B, 제13-B권 제1호, pp.21-26, 2006.
 [3] 김중태, "시대의 기회 시맨틱 웹", 디지털미디어리서치, 2006.
 [4] Eve Astrid Andersson, Stephen Breitenbach, Tyler Burd, Nirmal Chidambaram, Paul Houle, Daniel Newsome, Xiaofei Tang, Xiaolan Zhu, "Early Adopter VoiceXML", Wrox, 2002.
 [5] Min-Jen Tsai, "The VoiceXML Dialog System for the E-Commerce Ordering Service", IEEE Proceedings of the Ninth International Conference, pp.95-100, 2005.
 [6] Jeong-Hoon Shin, Kwang-Seok Hong, "Simple and Powerful Interface E-Learning System Using VXML : Design and Implementation of Web and PSTN Linked Efficient Learning System", LNCS 3980, pp.354-363, 2006.
 [7] De Sutter R, Lerouge, S, De Schrijver D, Van de Walle R, "Enhancing RSS Feeds : Eliminating Overheard through Binary Encoding", IEEE Proceedings of the Third International Conference, pp.520-525, 2005

[8] RSS Specifications, <http://www.rss-specifications.com>
 [9] Harvard University, <http://blogs.law.harvard.edu/tech/rss>
 [10] Atomenabled, <http://www.atomenabled.org>
 [11] Scott McGlashan, Daniel C. Burnett, Jerry Carter, Peter Danielsen, Jim Ferrans, Andrew Hunt, Bruce Lucas, Brad Porter, Ken Rehor, Steph Tryphonas, "Voice Extensible Markup Language Version 2.0 Specification", <http://www.w3c.org/TR/voicexml20>, 2004.
 [12] Rahul Ram Vankayala, Hao Shi, "Dynamic Voice User Interface Using VoiceXML and Active Server Pages", LNCS 3841, pp.1181-1184, 2006.
 [13] KT 음성 인식 서비스 개발팀, "VXML 기반 HUIVOIS Solution Manual R1.1", KT 서비스 개발 연구소, 2003.
 [14] VOX POPLUI, http://www.voiceingov.org/blog?page_id=8

권 형 준



email : katsyuki@skku.edu
 2005년 서울보건대학 전산정보처리과(이학사)
 2005년 (주)블루엠
 2006년~현재 성균관대학교 대학원
 정보통신공학부 (석사과정)
 관심분야 : 시스템 통합, 웹 서비스, HCI

김 정 현



email : kjh0328@skku.edu
 2000년 대전대학교 전자공학과(공학사)
 2002년 대전대학교 대학원 전자공학과
 (공학석사)
 2002년~2003년 순천향정보전문학교
 정보통신공학과 전임강사

2004년~현재 성균관대학교 대학원 정보통신공학부(박사수로)
 관심분야 : 제스처 인식 및 표현, 차세대PC

이 현 구



email : lhg@seoil.ac.kr
 1988년 성균관대학교 전자공학과(공학사)
 1991년 성균관대학교 대학원 전자공학과(공학석사)
 1999년 성균관대학교 대학원 전자공학과(공학박사)
 1991년~1996년 대영전자기술 연구소
 1996년~1998년 부일이동통신 중앙연구소

1988년~현재 서일대학 정보기술계열 정보통신전공 교수
 관심분야 : 통신 및 신호처리, 멀티미디어 통신

홍 광 석



email : kshong@yurim.skku.ac.kr
 1985년 성균관대학교 대학원 전자공학과(공학사)
 1988년 성균관대학교 대학원 전자공학과(공학석사)
 1992년 성균관대학교 대학원 전자공학과(공학박사)
 1990년~1993년 서울보건대학 전산정보처리과
 전임강사

1993년~1995년 제주대학교 정보공학과 전임강사
 1996년~현재 성균관대학교 정보통신공학부 교수
 관심분야 : 오감인식, 융합 및 재현, HCI