

사용자의 성향을 고려하는 지능형 음성 웹

권 형 준[†] · 흥 광 식^{††}

요 약

본 논문은 음성 웹의 지능화 및 개인화를 위한 방안을 제시한다. 제안하는 시스템은 한 번의 연결에서 요청된 정보들을 하나의 트랜잭션으로 기록하고, 기록된 트랜잭션들로부터 요청된 정보들 간의 연관 규칙을 탐사하여 빈발하게 요청된 정보들의 집합을 발견하며, 시스템은 이에 기반하여 과거의 사용자들과 유사한 입력 성향을 보이는 사용자에게 발견된 빈발집합에 근거하여 관련 정보를 추천한다. 유효성 검증을 위해서 제안하는 시스템을 구축하고 실험한 결과, 제안하는 시스템이 사용자가 요청한 정보와 빈발하게 요청되었던 정보를 관련 정보로 추천함으로써 사용자의 정보 획득을 도울 수 있음을 확인하였다.

키워드 : 보이스엑스엠엘, 음성 웹, 정보 시스템

Intelligent Speech Web Considering User Inclination

Hyeong-Joon Kwon[†] · Kwang-Seok Hong^{††}

ABSTRACT

In this paper, we propose a method for personalizing and intelligence of speech Web. The proposed system records information that was demanded in the past as a transaction, explores association rules from those transactions, and discovers itemsets from frequent requests. This method is to recommend relevant information, based on frequent itemsets, to users who have similar inclinations to previous users. As a result of experimenting and implementation of proposed system for verification, we confirmed that the proposed system can recommend previously frequently requested information as relevant information.

Key Words : VXML, Speech Web, Information Systems

1. 서 론

1999년 W3C에 의해 국제 표준으로 승인된 VXML(Voice eXtensible Markup Language)은 음성 인식 및 합성 기술을 이용하여 웹 자원에 접근할 수 있는 인터페이스의 제공을 위한 음성 다이얼로그의 정의를 목적으로 한다[1]. VXML의 등장으로 인해서 음성 어플리케이션 개발자들은 음성 처리에 관한 복잡한 문제들로부터 해방되어 사용자를 위한 시나리오의 개발에 집중할 수 있었고, 이는 웹 기반 자동 응답 시스템으로 이용되면서 서비스의 질과 성능을 향상시키는 계기가 되었다.

VXML이 가진 특징 중 하나로써, PC가 없어도 유무선 전화기를 이용하여 웹 자원에 접근할 수 있는 환경을 제공할 수 있다는 것은 사용자의 정보 획득을 위한 시간과 장소의

제약을 크게 완화하였다. 더불어 VXML 기반의 음성 어플리케이션에 관한 연구들은 외형적으로 유사한 시스템인 IVR(Interactive Voice Response) 및 ARS(Automatic Response System)와 차별화 되는 고급 시나리오를 구축할 수 있게 만들었다. 이는 주로 콜 센터의 구축을 목적으로 사용되던 웹 기반 음성 정보 시스템을 음성 웹의 개념으로 확장시켜서 전자 상거래 및 전자 교육과 정보 배급에 적용될 수 있음을 보였다[2-7]. 인간의 삶에 필수품이 된 휴대전화를 통해 시각의 웹 환경에서 얻을 수 있는 이득을 이동 환경에서 얻을 수 있게 만든 것이다. 이렇듯 음성 웹에 관한 연구가 지속적으로 이루어지고 있지만 주로 음성 웹을 어느 분야에 적용할 것인지에 관한 응용 측면에 치중되어 있으며, 현재 IT 기술의 이슈인 지능화 및 개인화에 관한 연구는 미비하였다. 이에 본 논문에서는 사용자의 성향에 따라 사용자가 요구하는 정보에 대한 관련 정보를 추천하는 지능화된 음성 웹 어플리케이션을 가진 시스템을 제안하고 설계 및 구현한다. 제안하는 시스템은 한 번의 연결에서 입력 받은 음성 앵커들의 목록을 트랜잭션으로써 데이터베이스에 기록하고, 누적된 트랜잭션들을 가지고 각 음성 앵커들 사이의 연관

* 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음 (IITA-2008-(C1090-0801-0046))

† 정 회 원 : 성균관대학교 전자전기컴퓨터공학과 박사과정

†† 정 회 원 : 성균관대학교 정보통신공학부 교수

논문접수 : 2008년 1월 22일

수정일 : 1차 2008년 4월 4일, 2차 2008년 4월 29일

심사완료 : 2008년 5월 10일

규칙을 탐사한다. 시스템에 탑재된 에이전트는 빈말하게 등장하는 음성 앵커들의 집합을 추출하여 집합에 포함된 각 음성 앵커들 사이의 연관 규칙을 탐사한 결과를 시스템 관리자에 의해 설정된 주기에 따라서 갱신하고, 음성 웹 어플리케이션은 그 결과에 기반하여 사용자의 입력에 대해 관련 정보로써 제시한다. 이는 조회된 결과를 한눈에 확인할 수 없는 불편함을 가진 유무선 전화를 통해서 음성 웹에 접근하고자 하는 사용자에게 유효할 수 있고, 사용자의 성향을 고려하여 관련 정보를 추천하기 때문에 사용자가 원하는 정보를 얻기 위해 정보를 검색하는 데에 소요되는 시간을 줄일 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 VXML 기반의 음성 웹에 관한 주요 기존 연구들에 대해 논하며, 3장에서는 음성 웹 어플리케이션과 규칙 마이닝 에이전트로 구성되는 제안하는 시스템의 설계 및 구현 방법을 설명한다. 4장에서는 제안하는 시스템의 유효성 검증을 위한 실험과 그 결과를 도출하고, 5장에서 결론을 맺는다.

2. 관련 연구

VXML 기반의 음성 웹은 프로그램 개발 단계에서 정의된 인식 후보 이외에는 사용자가 입력할 수 없다는 구조적인 단점을 가지고 있어서, 인식 후보가 정의되는 시점에서 음성 시나리오가 확정된다. 즉, 사용자가 자신의 의도대로 음성 시나리오를 수정해 갈 수 없다는 단점을 가지고 있기 때문에, 사용자가 자신의 의도대로 웹에 산재하는 정보를 탐색한다는 것은 사실상 불가능하다. 단지 사용자는 시스템이 안내하는 음성 메시지를 듣고, 그에 대한 응답을 반복하며 서비스 제공자가 제공하고자 하는 서비스를 제공받을 뿐이다. 이러한 이유로 인해서 VXML은 비용 절감과 유지 및 보수 측면의 이점을 취하기 위해 웹 기반 IVR 및 ARS의 대용으로 사용되어 왔다[2]. 이와 같은 문제를 개선하기 위한 방안으로써, Rahul Ram Vankayala는 VXML이 웹 기반으로 동작한다는 사실에 주목하였다. 웹 서버 측면의 스크립트 언어에 의한 동적인 VXML 문서의 생성이 가능함을 보이면서 VXML에서 기본적으로 지원하는 다이얼로그 구조가 가지고 있는 시나리오의 제약을 극복하기 위한 방안을

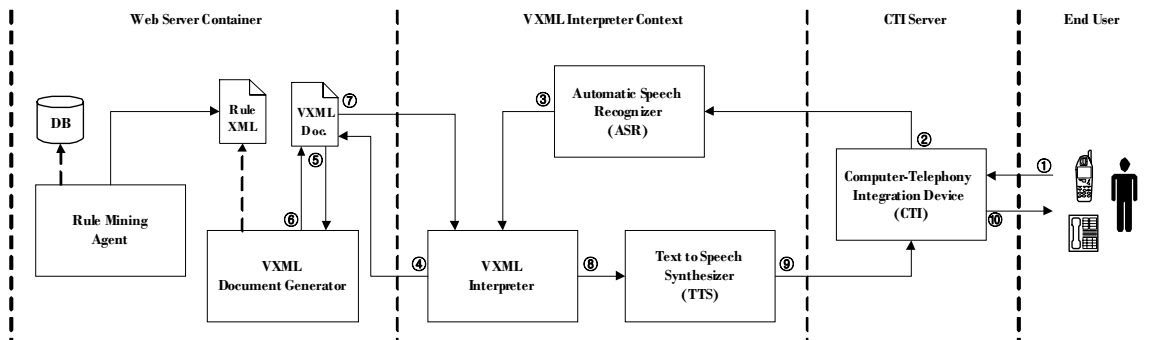
제시하였다[3]. 그것은 DBMS와 서버 측면의 스크립트 언어의 활용을 통해서 사용자의 입력에 따라 인식 후보를 갱신해가며 VXML 문서를 동적으로 생성해내는 방법이다. 이는 음성을 통한 전자 상거래와 전자 학습에 응용된 바 있으며, 이외에도 여러 가지 효과적인 활용 방안들이 꾸준히 제안되고 있다[2-5]. Hemambaradara Reddy에 의해 제안된 Listener-controlled 구조는 기존의 VXML 음성 웹 어플리케이션이 갖는 결점인 시나리오의 제약을 해결하고 음성 인터페이스를 이용한 웹 탐색을 가능하게 하기 위한 방안으로 음성 앵커의 필요성에 대해 논하였다[6]. VXML 기반의 음성 어플리케이션은 사용자의 입력에 의해 인식된 결과에 따라 시나리오가 진행되는데, 음성 앵커는 일반적인 웹에서 사용되는 하이퍼텍스트 구조에서 쓰이는 앵커와 유사한 개념으로써, 사용자의 음성 입력에 따라 새로운 URL을 통해 다른 VXML 문서로 연결시키는 역할을 수행한다. 이에 어플리케이션 개발 단계에서 결정되는 고정된 인식 후보들과 사용자가 요구한 정보로부터 자동으로 검출 및 등록되는 음성 앵커들에 기반한 구조가 제안되었다. 음성 웹 영역에서의 앵커와 노드 개념은 ARS 및 IVR 시스템 구축을 위한 용도로 사용되던 VXML의 활용도를 넓힘과 동시에, 음성 웹에서도 시각 환경의 웹처럼 사용자가 시나리오를 이끌어갈 수 있는 가능성을 열었다. 또한, 시맨틱 웹 개념에서 정보의 배급에 이용되는 RSS(Really Simple Syndication) 기술을 VXML과 접목하여 음성인식 및 합성기술을 통한 정보의 배급 방법도 제안된 바 있다[8].

이와 같이 여러 가지 관점에서 VXML에 기반한 음성 웹에 관한 연구들로 인해 음성 웹이 일반적인 시각 환경의 웹과 비슷한 수준의 활용도를 갖게 되었으나 지능화 및 개인화에 대한 연구는 미비하였다. 이에 본 논문에서는 과거의 사용자가 요청했던 정보들의 연관 규칙 마이닝을 통해서 현재 및 미래의 사용자가 흥미를 보일만한 정보와 연결되는 음성 앵커를 제시하는 지능을 가진 음성 웹 시스템을 제안한다.

3. 제안하는 지능형 음성 웹

3.1 시스템 아키텍처

제안하는 시스템의 구동을 위한 하드웨어 구조와 어플리



(그림 1) 시스템 아키텍처

케이션에 포함된 주요 소프트웨어 모듈 및 상호작용의 순서를 (그림 1)에 나타내었다. 사용자가 유무선 전화기를 통해 시스템에 연결하고 음성 입력을 시도하면(①), CTI 장비가 장착된 서버는 유무선 전화망을 통해서 아날로그로 전달받은 사용자의 음성 입력을 디지털로 변환하여 ASR에 전달한다(②). ASR은 음성 입력을 인식하고, 그 결과를 VXML 해석기에 전달하여(③) VXML 문서에 정의된 시나리오에 따라 다이얼로그를 실행한다(④). 제안하는 음성 웹 어플리케이션은 VXML 문서 생성기에 의해 사용자의 음성 입력을 수신할 때 마다 음성 앵커를 갱신하고, 시나리오를 변경한다(⑤). 주기적으로 실행되는 규칙 마이닝 에이전트에 의해 생성된 규칙 XML 문서를 탐색하여 사용자가 입력한 음성 앵커에 해당하는 정보 및 그와 관련 있는 음성 앵커들을 인식 후보들을 포함하는 새로운 VXML 문서를 만들어낸다(⑥). VXML 해석기는 새롭게 만들어진 VXML 문서에 정의된 다이얼로그 및 정보를 TTS와 CTI 서버를 통해 사용자에게 전달하면서 사용자의 입력에 대한 출력의 한 번의 주기를 마친다(⑦, ⑧, ⑨, ⑩).

3.2 음성 웹 어플리케이션

웹 서버 컨테이너에서 처리되는 절차를 (그림 2)에 나타내었다. 사용자가 연결을 요청하면, 웹 서버는 해당 연결에 대한 세션 ID와 함께 그 세션 ID를 통해서 접근할 수 있는 세션 메모리를 생성한다. 시각 환경의 웹에서는 세션 메모리 공간을 주로 페이지 간 이동이 발생할 때에 사용자 인증

정보를 유지하기 위한 용도로 사용하지만, 제안하는 어플리케이션에서는 한 번의 연결에서 입력 받은 음성 앵커들을 기억하고 연결 종료 시점까지 유지하기 위해 사용한다.

사용자가 입력할 수 있는 음성 앵커는 두 종류로 구분된다. 첫째는 어플리케이션의 사용을 중지하기 위해서 연결을 종료하거나, 직전에 청취한 내용을 다시 듣는 등의 시나리오 흐름을 제어하기 위한 정적인 음성 앵커이고, 둘째는 음성 어플리케이션에 의해 동적으로 등록된 음성 앵커들이다. 사용자가 동적으로 생성된 음성 앵커를 입력한 경우에만 해당 음성 앵커를 세션 메모리에 저장하며, 이와 같은 처리는 연결이 종료되어 세션이 소멸할 때까지 반복된다. 예를 들어, 어떠한 사용자가 연결하여 연결이 유지되는 동안 3번의 음성 앵커를 입력했다면 연결 종료 시점에는 3개의 음성 앵커가 세션 메모리에 저장되어 있을 것이다. 연결이 종료되기 직전에 세션 메모리에 저장되어 있는 사용자의 음성 앵커 목록은 데이터베이스에 하나의 트랜잭션으로써 저장된다. 즉, 한 번의 연결이 시작되고 종료되기까지 입력된 음성 앵커들은 데이터베이스에 존재하는 하나의 트랜잭션이 되어서 주기적으로 동작하는 규칙 마이닝 에이전트가 음성 앵커들 간의 연관 규칙을 발견하기 위해 사용된다.

3.3 규칙 마이닝 에이전트

제안하는 시스템에 탑재된 규칙 마이닝 에이전트는 음성 웹 어플리케이션에서 수집한 사용자의 음성 입력의 집합들을 가지고 각 음성 앵커들 사이의 연관 규칙을 일정 주기에 맞추어 탐사한다. 하나의 연관 규칙에 대한 X와 Y의 관계를 다음과 같이 나타낸다.

$$R: X \rightarrow Y$$

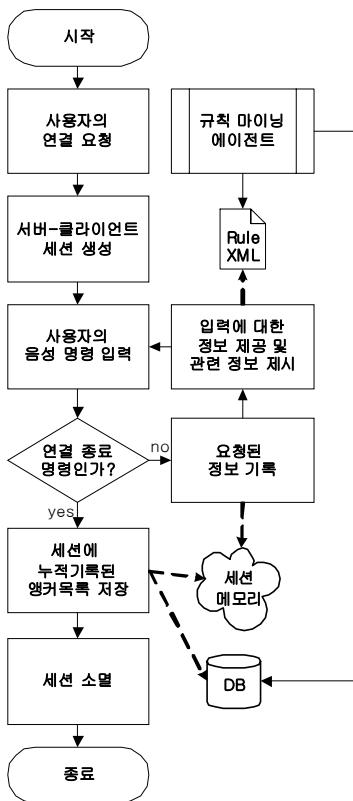
즉, 연관 규칙을 탐사한다는 것은 “X가 발생하면 Y도 발생한다” 라는 명제에서 X와 Y를 찾아내는 것이다. 데이터 마이닝 기법 중 하나인 연관 규칙의 탐사에서는 X와 Y를 찾아내기 위해서 지지도(Support), 신뢰도(Confidence) 및 개선도(Lift)를 고려한다. X에 대한 지지도는 전체 트랜잭션에서 X를 포함하고 있는 트랜잭션의 개수에 대한 비율이다. 신뢰도는 조건부 확률의 개념으로써 집합 X가 발생했을 때에 집합 Y도 동시에 발생할 확률을 의미한다. 즉, 트랜잭션이 항목 X를 포함할 경우 항목 Y를 동시에 포함할 확률을 나타내며, 신뢰도가 큰 규칙일수록 믿을 만한 규칙이라 말할 수 있다.

$$conf(R) = \frac{supp(X \cup Y)}{supp(X)}$$

개선도는 결과가 단독으로 발생할 빈도에 대한 조건과 연계하여 결과가 발생할 가능성의 빈도의 비로 정의된다. 개선도가 1이 된다는 것은 다음과 같은 규칙이 성립된다.

$$supp(X \cup Y) = supp(X)supp(Y)$$

이는 항목집합 X와 Y의 발생이 독립적임을 의미한다. 즉,



(그림 2) 제안하는 시스템의 음성 웹 어플리케이션 실행 순서도

개선도가 1보다 큰 규칙은 우연한 관계가 아닌 필연적 관계를 나타내며, 발견한 규칙의 검증을 위해서 사용된다. 규칙 R 에 대한 개선도는 다음과 같이 나타낸다[9].

$$lift(R) = \frac{conf(R)}{supp(Y)} = \frac{supp(X \cup Y)}{supp(X)supp(Y)}$$

이상의 3가지 파라미터를 고려하는 연관 규칙의 탐사 알고리즘은 여러 가지가 있으나, 본 논문에서는 처리 과정이 비교적 간편하면서도 뛰어난 성능으로 인해 여러 가지 변형된 알고리즘이 제안된 바 있는 Apriori 알고리즘을 적용하여 사용자들이 입력했던 음성 앵커들 간의 연관 규칙을 탐사하였다. Apriori 알고리즘에서는 데이터베이스의 반복적인 스캔에 의해 추출된 빈발 집합이 가진 모든 부분집합을 대상으로 X 와 Y 에 대입하여 지지도와 개선도를 계산하는 방법을 사용하여 연관 규칙을 도출한다[10].

(그림 3)은 1부터 5까지의 항목으로 구성된 4개의 트랜잭션으로부터 최소지지도가 0.3인 경우에서 Apriori 알고리즘이 빈발 집합을 탐색하는 과정을 보인다. 데이터베이스가 포함하고 있는 모든 항목들에 대한 지지도를 계산하여 1항목 빈발집합 후보인 $C1$ 을 생성하고, $C1$ 으로부터 최소지지도 이상의 지지도를 갖는 항목들로 1항목 빈발집합 $L1$ 을 생성한다. 다음으로, $L1$ 의 항목을 가지고 서로간의 곱집합을 수행하여 2항목 빈발집합 후보인 $C2$ 를 생성하고, $C2$ 가 가진 항목들의 지지도를 계산하여 2항목 빈발집합인 $L2$ 를 생성한다. Apriori 알고리즘에서는 이러한 방법으로 빈발 집합을 찾아내는 작업을 `apriori_gen` 함수라 칭하며, 더 이상 항목의 수를 증가시킬 수 없을 때까지 반복하여 최종 빈발 집합을 찾아낸다. 즉, (그림 3)에서 3항목 빈발집합 후보인 {2, 3, 5}의 지지도를 계산하면 0.5가 되어 최소지지도를 만족하므로 4항목 빈발집합 $L4$ 를 생성해야 하나, 생성할 수 없으므로 빈발 집합의 탐사가 종료되고 {2, 3, 5}는 빈발 집합으로 확정된다[10].

빈발집합을 찾아내기 위한 최소지지도의 적정 수치는 수 집된 트랜잭션들이 포함하고 있는 항목들이 어떻게 구성되어 있는지에 따라서 달라진다. 최소지지도가 지나치게 낮으면 n 항목 빈발집합 후보의 개수가 계속 증가하다가 결국 빈발집합을 찾을 수 없게 되는 경우가 있다. 즉, 트랜잭션 및

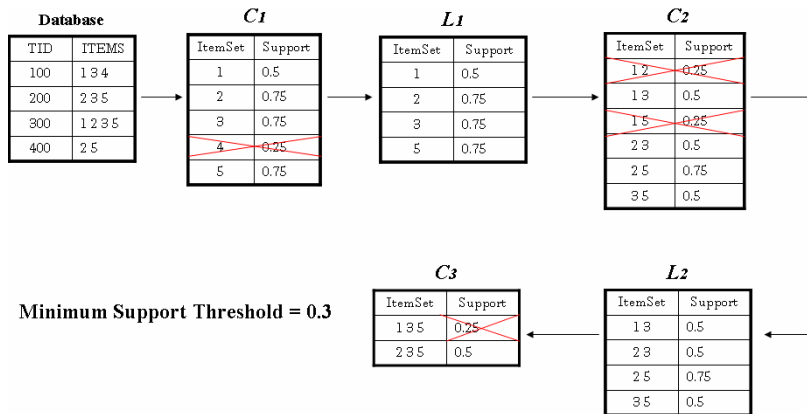
전체 항목집합의 개수에 따라서 적절한 최소지지도를 설정하는 것이 중요하다.

Apriori 알고리즘은 이러한 반복적인 계산에 따라 발견된 빈발 집합들의 모든 부분집합을 X 와 Y 에 대입하여 신뢰도를 계산한다. 계산된 신뢰도가 최소신뢰도의 값보다 큰 경우에 연관규칙 후보로 결정하고 개선도를 계산하여 개선도가 1보다 크면 해당 규칙을 연관 규칙으로 확정한다. (그림 3)에서 최소신뢰도를 0.6으로 설정하는 경우, 추출된 빈발 집합 {2, 3, 5} 중 $R: \{2\} \Rightarrow \{3, 5\}$ 를 고려하면 규칙에 대한 신뢰도는 0.66, 개선도는 1.32가 되어 연관 규칙이라 할 수 있지만, $R: \{3\} \Rightarrow \{2, 5\}$ 를 고려하면 규칙에 대한 신뢰도는 0.66, 개선도는 0.88이 되어 연관 규칙이라 할 수 없다[9-10]. 최소신뢰도 임계치를 높게 설정할 수록 더 믿음만한 규칙을 발견할 수 있지만 과도하게 높은 경우에는 규칙을 찾지 못하게 될 것이고, 과도하게 낮은 경우에는 많은 규칙을 찾아내지만 유용하지 못한 규칙들을 포함하게 될 수 있다.

제안하는 시스템에서는 과거 사용자들의 입력이 X 일 때 빈발하게 발생했던 Y 를 찾아내어 X 를 입력한 현재와 미래의 사용자에게 Y 를 추천하기 위해 사용되었다. 탐사된 결과는 (그림 4)에 나타낸 DTD(Document Type Definition)에 따라 XML 문서로 출력되며, VXML 문서 생성기가 VXML 문서를 생성할 때 사용자의 입력에 대한 관련 정보를 선별하기 위해 사용된다[11]. 4장에서 제안하는 DTD 문서의 실제 사용 예를 보인다.

```
<? xml version="1.0" ?>
<! DOCUMENT result [
<! ELEMENT result (transcount, threshold, rule)>
  <! ELEMENT transcount (#PCDATA)>
  <! ELEMENT threshold (#PCDATA)>
<! ELEMENT rule (xvalue, yvalue, confidence, lift)>
  <! ELEMENT xvalue (#PCDATA)>
  <! ELEMENT yvalue (#PCDATA)>
  <! ELEMENT confidence (#PCDATA)>
  <! ELEMENT lift (#PCDATA)>
]>
```

(그림 4) 연관규칙 탐사 결과가 기록되는 XML의 DTD



(그림 3) Apriori 알고리즘의 빈발 집합 탐색 과정

4. 실험 및 결과

제안하는 시스템의 유효성을 검증하기 위해 가상의 증권 정보 시스템을 구축하여 실험하였다. 사용자는 시스템에 접속하여 정보 조회를 원하는 종목 명을 입력할 수 있으며, 입력한 종목의 정보를 알려주고, 과거에 그 종목을 입력했던 사용자들이 빈발하게 조회한 종목을 추천하는 시스템이다.

(그림 1)에 보인 Web Server Container와 VXML Interpreter Context를 <표 1>에 나타난 것과 같은 사양으로 구축하였고, CTI 서버는 같은 사양의 또 다른 서버 컴퓨터를 사용하여 총 2대의 서버로 구성하였다[12-13]. 실험에 사용된 데이터베이스는 나스닥에 상장된 종목 중 임의의 17개 종목을 가지고 8명의 인원이 일정 기간 동안 시스템에 접근하여 이용한 결과를 수집한 결과인 161개의 트랜잭션으로 구성되었다. 규칙 마이닝 에이전트는 플랫폼에 구매 받지 않고 탑재될 수 있도록 Java 2 Standard Edition 1.6.0_03으로 콘솔 환경에서 동작하도록 구현하였다.

<표 1> 시스템의 구축 및 실험을 위한 사양

항목	내용
CPU	Intel XEON 3.0Mhz Dual
RAM	2GB
CTI Board	D / 120JCT-LS PCI (12CH)
OS	Windows 2000 Professional
IIS Ver.	5.0
DBMS	Microsoft MS-SQL 2003
ASR, TTS and VXML Interpreter	KT Huvois R1.1

4.1 최소지지도 임계치의 변화에 대한 탐사 결과의 관측

제안하는 시스템에 의해 수집된 트랜잭션을 바탕으로, 최소지지도 임계치의 변화에 대한 연관규칙의 탐사 결과를 관측하였다. 최소신뢰도를 0.1로 고정하고, 최소지지도 임계치를 0.1부터 0.05씩 증가시켜서 연관규칙의 탐사한 결과, 최소지지도가 낮으면 많은 연관규칙 후보를 발견하지만 그 신뢰도가 낮음을 확인할 수 있었다. 제안하는 시스템이 수집한 트랜잭션에 대해서, 최소지지도 임계치의 변화에 따라 발견된 연관규칙 후보의 개수와 그에 대한 신뢰도 범위 및 탐사에 소요된 시간을 <표 2>에 나타내었다. 소요시간은 트랜잭션의 개수와 트랜잭션들이 포함하고 있는 항목의 수에 따라 다를 수 있다. 특히, 항목집합의 개수를 늘려가며 지지도를 계산하는 빈발집합 추출 단계에서 L_n 및 C_n 의 개수가 증가할수록 오랜 시간을 필요로 한다. <표 2>에 나타난 소요시간을 살펴보면, 최소 지지도가 낮을수록 빈발집합의 추출을 위해서 n항목 빈발집합의 많은 반복계산을 수행하기 때문에 오랜 시간이 소모됨을 알 수 있다.

최소신뢰도 임계치가 낮을수록 많은 연관규칙을 발견할 수 있으나, 지나치게 낮은 경우에는 유용하지 못한 연관규칙을 포함할 수 있다. (그림 5)는 제안하는 시스템에서 최소

<표 2> 최소지지도의 변화에 따른 신뢰도 범위

최소지지도	신뢰도 범위	발견된 후보 개수	소요시간
0.10	-	0	9.58 초
0.15	0.28 ~ 0.42	5	5.12 초
0.20	0.28 ~ 0.42	5	4.32 초
0.25	0.44 ~ 0.64	4	3.97 초
0.30	0.44 ~ 0.64	4	4.15 초
0.35	0.53 ~ 0.68	3	3.23 초
0.40	0.53 ~ 0.68	3	3.08 초

지지도와 최소신뢰도를 각각 0.25와 0.5로 설정한 경우의 연관규칙 탐사 결과를 보인다. <표 2>에 따르면 최소지지도가 0.25인 경우에는 4개의 연관규칙 후보를 발견할 수 있으나, 최소신뢰도를 0.5로 설정함에 따라서 한 개의 연관규칙 후보가 탈락되었다. (그림 5)에서, 최소신뢰도 임계치를 낮춘다면 최대 4개의 규칙을 발견할 수 있을 것이며, 임계치를 높일수록 발견된 규칙들 중 일부가 연관규칙에서 배제될 것이다.

```

** Print C4: [|Apple|BEA Systems|Dell|Ebay| ,
** Print L4: [|Apple|Dell|Ebay|Nvidia|]

** I found 1 frequently Itemsets. **

Rule: |Apple| -> |Dell|Ebay|Nvidia|
conf(R) 0.44 // lift(r) 1.415

Rule: |Dell| -> |Apple|Ebay|Nvidia|
conf(R) 0.648 // lift(r) 1.8 (discovered)

Rule: |Ebay| -> |Apple|Dell|Nvidia|
conf(R) 0.564 // lift(r) 1.814 (discovered)

Rule: |Nvidia| -> |Apple|Dell|Ebay|
conf(R) 0.558 // lift(r) 1.696 (discovered)
    
```

(그림 5) 규칙 마이닝 에이전트의 실행 결과

4.2 탐사 결과의 출력 및 기록여부 확인

(그림 6)은 제안하는 시스템이 (그림 5)에서 나타난 탐사 결과를 (그림 4)에 정의한 DTD에 따라서 연관규칙의 탐사 결과를 기록한 모습을 출력한 화면을 보인다. 에이전트는 (그림 5)에서 보는 것과 같이 3개의 연관 규칙을 찾아냈고, 음성 웹 시스템은 사용자의 입력이 <rule> 엘리먼트의 하위 엘리먼트인 <xvalue>의 속성들과 일치하는 항목이 있다면, 그 <xvalue>와 동일한 레벨의 <yvalue>를 관련 정보로써 사용자에게 제시해줄 것이다.

(그림 7)은 사용자의 입력이 “Nvidia”일 때에 동적으로 생성된 VXML 문서를 시각적으로 나타낸 모습이다. <grammar> 엘리먼트의 속성을 살펴보면 규칙 마이닝 에이전트에 의해 발견된 음성 앵커들이 인식 후보로써 기술되었음을 확인할 수 있고, <prompt> 엘리먼트는 사용자에게 TTS 메시지로 제공될 내용을 담고 있다.

```

<?xml version="1.0" encoding="euc-kr" ?>
- <result>
  <transactioncount>161</transactioncount>
  <supportthreshold>0.25</supportthreshold>
  <confidencethreshold>0.5</confidencethreshold>
  - <rule>
    <xvalue>Dell</xvalue>
    <yvalue>Apple, Ebay, Nvidia</yvalue>
    <confidence>0.648</confidence>
    <lift>1.8</lift>
  </rule>
  - <rule>
    <xvalue>Ebay</xvalue>
    <yvalue>Apple, Dell, Nvidia</yvalue>
    <confidence>0.564</confidence>
    <lift>1.814</lift>
  </rule>
  - <rule>
    <xvalue>Nvidia</xvalue>
    <yvalue>Apple, Ebay, Dell</yvalue>
    <confidence>0.558</confidence>
    <lift>1.696</lift>
  </rule>
</result>

```

(그림 6) 제안하는 시스템에 의해 출력된 연관규칙 탐사 결과

```

<?xml version="1.0" encoding="euc-kr" ?>
- <vxml version="2.0">
  - <form>
    - <field name="userinput">
      <!-- Voice-anchors are described in property of <grammar> -->
      <!-- <prompt> is provided to user via TTS module as selected information by user. -->
      <grammar>Apple | Ebay | Dell</grammar>
      <!-- <prompt> is provided to user via TTS module as selected information by user. -->
      <prompt>You choose a Nvidia. Current actual price of Nvidia is 5,145. Today trading volume is 51,327,231. Relevant issues are Apple, Ebay, Dell. If you want to refer to them, state desired issue.</prompt>
    - <filled>
      <submit next="navigate.asp" namelist="userinput" />
    </filled>
  </field>
</form>
</vxml>

```

(그림 7) 제안하는 시스템이 동적으로 생성한 VXML 문서

4.3 적중률 측정

제안하는 시스템이 에이전트에 의해 추출된 정보를 사용자에게 추천했을 때, 사용자가 시스템에 의해 추천된 정보에 관심을 가지고 그 정보를 요청한 횟수를 측정하여 <표 3>에 나타내었다. 연관규칙 마이닝 에이전트는 트랜잭션이 추가될 때 마다 자동적으로 실행되도록 설정하였다. 이와 같은 실험은 제안하는 시스템이 사용자에게 도움이 될 수 있는지 판단할 수 있는 척도가 된다.

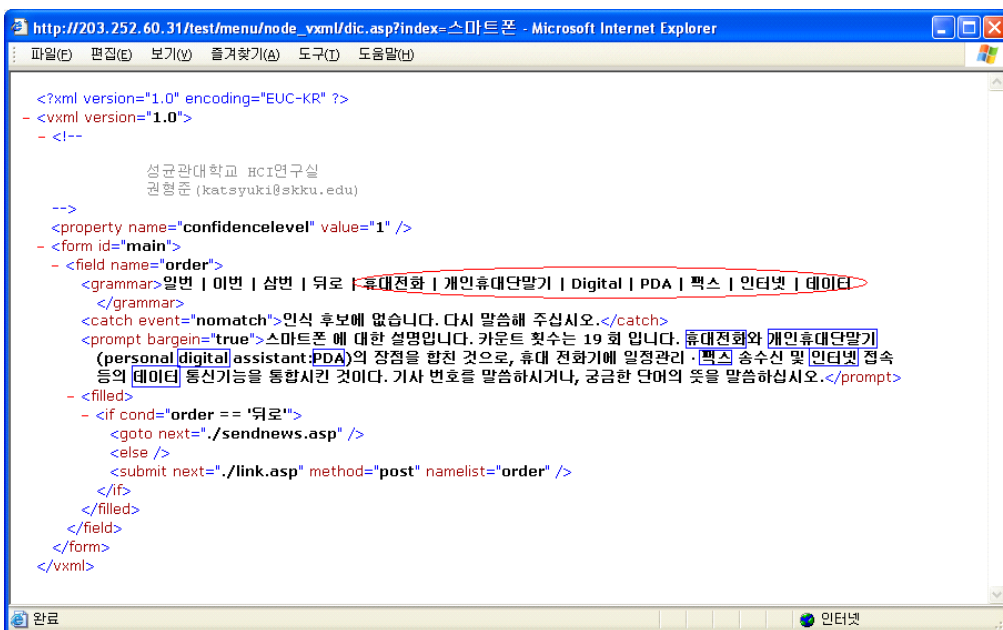
총 161개의 트랜잭션을 가지고 적중률을 측정한 결과, 8명의 인원에 대한 추천 횟수는 76회였고 요청된 횟수는 54회로써 71.05%의 적중률을 보였다. 트랜잭션이 채워지지 않은 경우에는 마이닝의 수행이 불가능하여 관련 정보를 추천하지 못했음을 고려했을 때, 트랜잭션이 증가할수록 적중률의 상승을 기대할 수 있다.

<표 3> 적중률 측정 결과

사용자	연결횟수	추천횟수	요청횟수	적중률(%)
A1	27	16	13	81.25
A2	19	8	6	75.00
A3	24	6	5	83.33
A4	18	12	9	75.00
A5	17	5	3	60.00
A6	14	9	6	66.66
A7	23	7	4	57.14
A8	19	11	8	72.72
8명	161회	76회	54회	71.05

4.4 기존 연구와의 비교 및 통합 방안

(그림 8)은 사용자가 요청한 정보가 포함하고 있는 명사



(그림 8) 기존의 사용자 주도적 폼 다이얼로그 구조

형 단어들 중에서 시스템이 가진 단어 사전에 존재하는 모든 키워드들을 인식 후보에 기술하는 방법을 통해서 시나리오를 자동적으로 확장해주는 기존의 연구이다[14]. 사용자에게 TTS로 들려주는 <prompt> 엘리먼트의 내용으로부터 IT용어들을 검출하여 인식 후보에 추가시킨 모습이 (그림 8)에 나타나 있다.

이와 같은 시스템은 사용자에게 제공되는 내용이 많아질 수록 인식 후보의 개수가 증가하는 특징이 있는데, 인식 후보를 한눈에 살펴 볼 수 없고 집중해서 청취해야 하는 음성 웹의 특성상 인식후보가 기하급수적으로 증가할 수 있는 기존의 시스템과의 결합은 더욱 지능적이고 효율적인 시스템을 만들 수 있다. 제안하는 시스템에 의해서, 수많은 인식 후보들 중 사용자의 성향과 유사한 성향을 보였던 사용자들에 의해 요청되었던 정보들을 시스템이 사용자에게 추천 함으로써 정보 획득에 직접적인 도움을 줄 수 있을 것이다.

5. 결 론

본 논문에서는 VXML 기반의 음성 웹 시스템의 지능화에 대한 방법을 제안하였다. 기존의 음성 웹 시스템들은 사용자의 의사 결정을 직접적으로 돕지 못하였으나, 본 논문에서는 문제의 해결을 위해서 연관규칙 마이닝을 수행하여 사용자가 관심을 보일 만한 정보를 추천하는 지능형 에이전트를 탑재하여 음성 웹의 지능화를 이루었다. 제안하는 시스템은 과거의 사용자들에 의해 요청된 정보들의 패턴으로부터 연관 규칙을 탐사하고, 탐사 결과에 기반하여 현재 및 미래의 사용자가 정보를 요청할 시에 요청된 정보와 함께 그에 관련된 음성 앵커를 제시함으로써 사용자의 정보 획득을 돕는다.

참 고 문 헌

[1] S. McGlashan, D. C. Burrnet, J. Carter, P. Danielsen, J. Ferrans, A. Hunt, B. Lucas, B. Porter, K. Rehor and S. Tryphonas, "Voice Extensible Markup Language Version 2.0," W3C Recommendation, 2004.

[2] J. H. Shin and K. S. Hong, "Implementation of New CTI Service Platform Using Voice XML," Lecture Notes in Computer Science, Vol.3046, pp.754-762, 2004.

[3] R. R. Vankayala and H. Shi, "Dynamic Voice User Interface Using VoiceXML and Active Server Pages," Lecture Notes in Computer Science Vol.3841, pp.1181-1184, 2006.

[4] M. J. Tsai, "The VoiceXML Dialog System for the E-Commerce Ordering Service," IEEE Proceedings of the Ninth International Conference, pp.65-100, 2005.

[5] J. H. Shin and K. S. Hong, "Simple and Powerful Interactive E-Learning System Using VXML: Design and Implementation

of Web and PSTN Linked Efficient Learning System," Lecture Notes in Computer Science, Vol.3980, pp.354-363, 2006.

[6] H. Reddy, N. Annamalai and G. Gupta, "Listener-Controlled Dynamic Navigation of VoiceXML Documents," Lecture Notes in Computer Science, Vol.3118, pp.347-354, 2004.

[7] H. J. Kwon and K. S. Hong, "A User-Controlled VoiceXML Application Based on Dynamic Voice Anchor and Node," Lecture Note in Computer Science, Vol.4743, pp.265-272, 2007.

[8] H. J. Kwon, J. H. Shin and K. S. Hong, "Design and Implementation of Enhanced Real Time News Service Using RSS and VoiceXML," Lecture Notes in Computer Science, Vol.4557, pp.677-686, 2007.

[9] R. Agrawal, T. Imielinski and A. Swami, "Mining association rules between sets of items in large databases," Proceedings of the 1993 ACM SIGMOD international conference on Management of Data, pp.207-216, 1993.

[10] G. I. Webb, "Efficient search for association rules," Proceedings of the Sixth ACM SIGKDD international conference, pp. 99-107, 2000.

[11] W3C, "XML Specification DTD," <http://www.w3.org/XML/1998/06/xmlspec-report-19980910.htm>

[12] Intel Dialogic Homepage, <http://www.dialogic.com/manuals/sr60winpci/IssuesByProduct.pdf>

[13] KT Service Develop Laboratory, "HUVOIS Solution Manual R1.1," 2003.

[14] 권형준, 홍광석, "사용자 주도 폼 다이얼로그 시스템의 VoiceXML 어플리케이션에 관한 연구", 정보처리학회 논문지, 제14-B권 제3호, pp.183-190, 2007.



권 형 준

e-mail : katsyuki@skku.edu
 2005년 서울보건대학 전산정보처리과 (학사)
 2005년 ㈜블루엠 소프트웨어개발팀
 2005년 ㈜티니아텍 자바누리(<http://www.javanuri.com>) 운영진
 2007년 성균관대학교 전자전기컴퓨터공학과 (공학석사)

2008년~현 재 성균관대학교 전자전기컴퓨터공학과 박사과정
 관심분야 : 지능시스템, 오감정보처리, HCI



홍 광 석

e-mail : kshong@yurim.skku.ac.kr

1985년 성균관대학교 전자공학과(학사)

1988년 성균관대학교 전자공학과(공학석사)

1992년 성균관대학교 전자공학과(공학박사)

1990년~1993년 서울보건대학 전산정보처리과
전임강사

1993년~1995년 제주대학교 정보공학과 전임강사

1996년~현재 성균관대학교 정보통신공학부 교수

관심분야 : 오감인식, 융합 및 재현, HCI