

## 지능망 기반 음성인식 일기예보 서비스 시스템 개발

박성준, 김재인, 구명완  
KT, 서비스개발연구소

### Development of the Weather Forecasting Service System Based on AIN

Sung-Joon Park, Jae-In Kim, Myoung-Wan Koo  
Service Development Laboratory, KT

**Abstract** - 본 논문에서는 음성인식을 이용한 일기예보 서비스 시스템을 소개한다. 이 서비스는 사용자가 지역명을 말하면 음성인식을 통해 그 지역명을 인식하여 일기예보를 들려주며, 차세대 지능망(AIN: Advanced Intelligent Network)에 구현되었다. 음성인식은 IP(Intelligent Peripheral)에서 이루어지며, 음성인식 실험 결과, 실험실과 시스템 상에서 각각 95.04%와 93.81%의 인식율을 보여주었다.

### 1. 서 론

일기예보는 농업, 여행 및 기타 일상생활에 중요한 정보로서 KT는 전화망을 통한 일기예보 서비스를 제공해 왔다. 현재 전국적으로 16개의 지역번호가 있으며, 각 지역별로 한두 대의 일기예보 시스템이 설치되어 있어서 그 지역의 기상 정보를 담당하고 있다. 사용자들이 지역번호 없이 131 번호를 걸면 사용자가 속한 지역의 일기예보를 들을 수 있으며, 다른 지역의 날씨 정보를 알고 싶을 때는 지역번호도 같이 입력해 주어야 한다.

현재의 서비스가 가지고 있는 문제점으로서 한 번의 전화로 지역 번호가 다른 여러 지역의 일기예보를 들을 수 없다는 점과 지역에 따라 R2 신호방식으로 인한 자동접속지연(PDD: Post-Dialing Delay) 시간이 길다는 점이 있다. 또한 지역별로 시스템이 분산되어 있음으로 인해 집중 관리가 되지 않는다는 단점이 있다. KT에서는 이러한 문제점과 단점을 해결하기 위해 기상청과 협의하여 일기예보 서비스를 지능망에 수용하기로 결정하였으며, 또한 음성인식 기능을 이용하여 다른 지역의 일기예보로 들을 수 있도록 하였다.

음성인식 기능은 사용자가 다른 지역의 일기예보를 들을 때, 여러 지역명을 듣고 DTMF (Dual-Tone Multifrequency)를 눌러서 선택할 필요 없이, 바로 원하는 지역명을 말함으로써 그 지역의 일기예보를 들을 수 있게 해 준다. 음성인식을 위하여 AIN의 IP에 설치되어 있는 음성인식 엔진을 사용하였고, 시나리오는 VoiceXML 형태로 작성되었다.

본 논문의 구성은 다음과 같다. 2장에서는 시스템의 구조와 VoiceXML 인터프리터에 대해서 설명하고, 3장에서는 서비스의 시나리오에 대하여 기술한다. 음성인식 실험 결과는 4장에 나타내었으며, 마지막으로 5장에서 결론을 기술하였다.

### 2. 시스템 개요

현재 지능망 시스템은 여러 종류의 서비스를 제공하고 있으며, VoiceXML 인터프리터 또한 IP 내에 설치되어 있다. 지능망에서 일기예보 서비스를 제공하기 위하여 일기예보 정보를 입력할 수 있는 웹서버를 구축하여 기상청이 연결할 수 있도록 하였으며, 이 서비스를 위한 VoiceXML 문서를 작성하였다.

#### 2.1 시스템 구조

그림 1은 일기예보 서비스와 관련된 지능망의 주요 구성 요소를 나타낸 것이다.

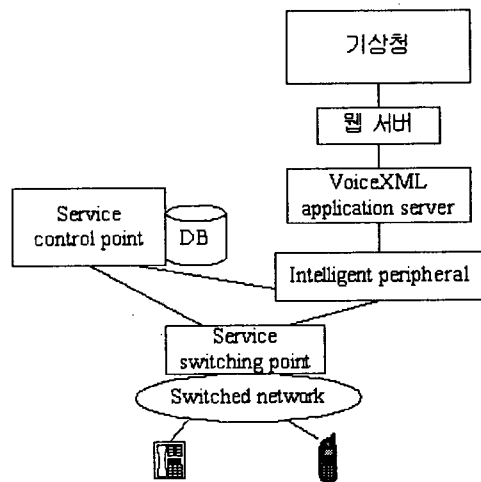


그림 1. 시스템 구성도

서비스 교환기(SSP: Service Switching Point)는 단국 교환기로 사용될 경우에는 사용자가 망으로 접근할 수 있게 하고, 지능망 호 연결을 위하여 필요한 교환 기능을 제공한다. 이를 위하여 SSP는 지능망 서비스를 감지하는 기능과 서비스 제어 시스템(SCP: Service Control System)과 같은 다른 물리 실체와의 통신 능력과, 다른 물리 실체로부터 수신된 명령에 대한 응답 능력을 가진다 [1][2].

SCP는 지능망 서비스 처리에 필요한 서비스 로직 프로그램과 데이터를 가지고 있으며, 신호망을

통하여 SSP, IP 등과 연결된다.

IP는 사용자와 망과의 특수한 상호작용을 처리하게 되는데, 안내음성과 음성인식 엔진이 설치되어 있다.

SCP가 SSP로 하여금 호를 IP로 넘기면, IP는 안내음성을 내보내고 DTMF를 받아들이는 것과 같은 기능을 수행한다. 본 시스템에서는 IP 내에 VoiceXML 인터프리터가 설치되어 있어서 응용 서버(AS: Application Server)로부터 VoiceXML 문서를 가져오고, 문서에 기술된 시나리오에 따라 서비스를 제공한다. AS는 일기예보 서비스뿐만 아니라, 1582 음성다이얼 서비스나 멤버링 서비스와 같은 다른 지능망 서비스를 위한 문서도 저장하고 있다.

전국에 분산되어 있는 여러 기상청에서는 본 시스템에서 제공하는 웹 서버로 접속하여 매 시간 일기예보 정보를 갱신한다.

### 2.2 VoiceXML 인터프리터

VoiceXML 인터프리터는 VoiceXML 언어 표준에 의해 정의된 VoiceXML 문서를 읽고 처리하는 소프트웨어다[3][4].

VoiceXML은 상호 작용을 지원하는 markup 언어로서, 음성과 전화를 통하여 웹 정보를 검색하고 응용 프로그램을 개발하는 환경을 제공한다. 즉, 음성 합성, 음성 인식, DTMF 입력 등을 통해 음성 대화를 할 수 있도록 설계되어 있다. VoiceXML 구조는 구현 플랫폼, VoiceXML 인터프리터 환경, VoiceXML 인터프리터, 문서 서버로 이루어져 있으며, 구성요소들간의 관계를 그림 2에 나타내었다.

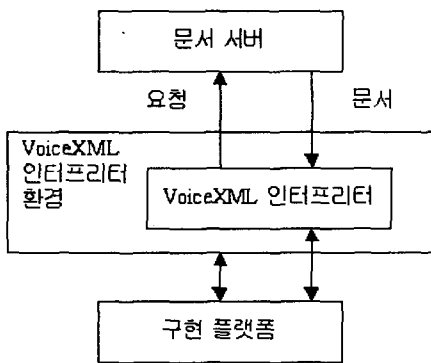


그림 2. VoiceXML 구성 요소

구현 플랫폼은 들어오는 호와 나가는 호를 체크하고, 사용자의 입력과 시스템의 상태에 따라 이벤트를 발생시킨다. 이벤트에 따라 VoiceXML 인터프리터가 작동되거나 VoiceXML 인터프리터 환경이 변하게 된다. VoiceXML 인터프리터 환경은 사용자의 입력과 세션에 관련된 환경 변수를 관리한다.

VoiceXML 인터프리터는 문서 서버로부터 받은 VoiceXML 문서를 파싱하고 대화에 따라 전체 로직을 제어한다.

문서 서버는 VoiceXML 인터프리터로부터의 요구를 처리하게 되는데, 결과물로 VoiceXML 문서

를 제공한다. VoiceXML 인터프리터는 CGI 스크립트나 오디오 파일과 같은 형태의 문서도 처리할 수 있다.

구현 플랫폼, VoiceXML 인터프리터 환경, VoiceXML 인터프리터는 IP내에 포함되며, 문서 서버는 AS에 구현되었다.

KT는 VoiceXML 인터프리터를 개발하여 음성 대화 시스템에 적용하였으며[5], 이를 개선하여 지능망에 도입하였다.

### 3. 시나리오

사용자가 131을 걸어서 시스템과 연결될 때, 지능망 시스템은 사용자의 전화번호를 이용하여 어느 지역에서 전화를 거는지를 알아낸다. 시스템은 우선 사용자가 속한 지역의 일기예보를 제공하며 다른 지역에 대한 선택 메뉴도 제공한다.

다른 지역의 일기예보를 듣고 싶을 때는 '0'을 누르고 원하는 지역명을 말한다. 시스템은 약 340개의 시,군 명칭에서 사용자가 말한 지역을 선택하여 그 지역의 일기예보 정보를 들려준다. 그림 3에 전체 시나리오를 나타내었다.

본 시스템에서 음성인식이 사용되는 것은 사용자가 다른 지역의 정보를 듣기 위하여 '0'을 누를 때로서, 음성인식 기능이 제한적으로 사용되었다. 그 이유는 기존의 131 서비스에서 6개월간의 통화유형을 분석한 결과, 동일 통화권 내의 통화 점유가 타통화권의 통화 점유보다 훨씬 많았기 때문이다. 표1에 통화유형 분석 결과를 나타내었다.

구분	점유율	평균통화시간	연결방법
동일통화권	96.8%	57초	131
타 통화권	3.2%	62초	지역번호 +131

표 1. 통화 유형

### 4. 음성인식 시험

KT에서 개발한 음성인식 엔진은 HMM(hidden Markov model)에 기반하고 있으며[6], 크게 전처리 부분과 디코더 부분으로 나눌 수 있다. 전처리 단계에서는 8kHz로 샘플링되는 16bit PCM 음성 데이터로부터 12차의 LPC(Linear Predictive Coding) 켈스트럼 계수와 에너지 값을 계산해 낸다. 샘플링되는 음성은 20ms 단위의 프레임으로 생성되며, 각 프레임들은 10ms씩 중첩되고, Hamming 윈도우가 사용된다.

각 프레임별로 생성된 12차 켈스트럼 계수와 에너지 값을 이용하여 delta 켈스트럼 계수, acceleration 켈스트럼 계수, delta 에너지, acceleration 에너지 등을 계산해 내고, 처음에 구한 에너지 값을 제외한 총 38차의 특징 벡터를 각 프레임별로 생성한다. 또한 채널의 영향을 없애기 위하여 켈스트럴 평균 차감법을 적용하였다.

실험을 간단히 하기 위하여, 시, 군 명칭을 위한

새로운 HMM 파라미터는 생성하지 않았으며, 현재 AIN에서 저장되어 있는 HMM 파라미터를 사용하여 실험을 수행하였다.

실험은 2회에 걸쳐 이루어졌으며, 한 번은 시스템에 녹음된 음성을 이용하여 실험실에서 수행되었고, 두 번째는 실제 시스템에 전화를 걸어서 인식 결과를 수집하였다.

처음 실험에서는 12명으로부터 총 847개의 음성을 수집하였고, 두 번째 실험에서는 12명이 총 339개의 단어를 발화하였다. 인식 결과는 표2에 나타내었다.

	실험실	시스템
인식율(%)	95.04	93.81

표 2. 음성 인식율

### 3. 결 론

본 논문에서는 음성인식을 이용한 일기예보 서비스를 소개하였다. 서비스가 구현된 AIN의 구조와 각각의 구성 요소를 기술하였으며, 실험실과 시스템에서 수행된 음성인식 실험 결과를 나타내었다. 각각의 인식율은 95.04%와 93.81%이다.

현재, 인식율 향상을 위하여 HMM을 갱신하는 작업을 진행하고 있으며, 대량의 호가 발생할 때 안정적으로 서비스를 제공할 수 있도록 시험을 진행 중이다. 시험이 끝나면 기존의 131 번호를 AIN에 연결시킴으로써 사용자들은 새로운 시스템을 사용할 수 있을 것이다.

#### (참 고 문 헌)

- [1] Roger K. Berman and John H. Brewster, "Perspectives on the AIN Architecture", IEEE Communications Magazine, pp. 27-32, February 1992.
- [2] C.D. Sharp and K. Clegg, "Advanced intelligent networks - now a reality", Electronics & Communication Engineering Journal, pp. 153-162, June 1994.
- [3] VoiceXML Forum, VoiceXML (Voice extensible Markup Language) version 1.0, "http://www.w3.org/TR/2000/NOTE-voicexml-20000505/", 2000.
- [4] Bob Edgar, *The VoiceXML Handbook - Understanding and Building the Phone-Enabled Web*, CMP Books, New York, 2001.
- [5] Hak-Gyoon Kim, Myong-Wan Koo, and Yong-Ki Park, "TeleGateway: A Spoken Dialogue System Based on VoiceXML", KT Journal, vol. 6, no. 3, pp. 56-63, December 2001.
- [6] L. Rabiner and B. H. Juang, *Fundamentals of speech recognition*, Prentice-Hall, NJ, 1993.

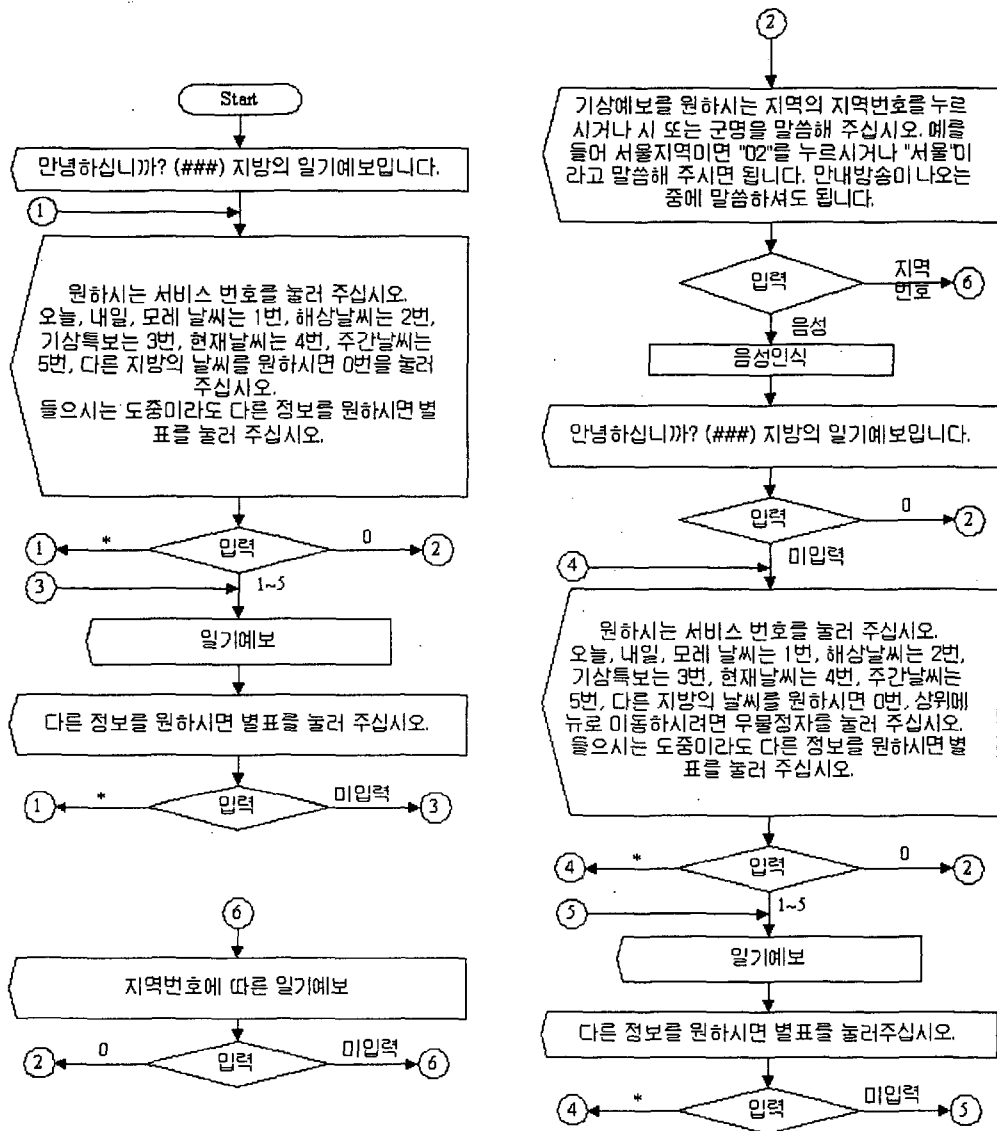


그림 3. 음성인식 일기예보 서비스 시나리오