

# Web을 이용한 음성메일 전송 시스템 설계 및 구현\*

임종철, 홍의경  
서울시립대학교 전산통계학과  
{jcyim92, ekhong}@venus.uos.ac.kr

## Design and Implementation of Voice Mail Transmission Systems on WWW

Jong-Choul Yim, Eui Kyeong Hong  
Dept. of Computer Science and Statistics, University of Seoul

### 요 약

인터넷 사용자가 급격히 증가함에 따라 인터넷을 통한 멀티미디어 서비스에 관심이 집중되고 있다. 인터넷을 통한 화상회의나 인터넷 폰, 음성 메일 등이 그 대표적인 예이다. 그 중에서도 대부분의 사람들은 전자 메일을 통해 정보를 공유한다. 현재 대부분의 메일 Agent 프로그램들은 음성 메일을 보낼 수 없으며, 또한 웹을 기반으로 구축된 메일 시스템도 그러하다. 본 논문에서는 Web을 기반으로 하여 음성 메일을 전송하는 시스템을 설계하였다. 현재 웹 응용 시스템을 구축하는 핵심 기술로서 부상하고 있는 서버릿을 사용하여 사용자를 대신해 메일을 전송하는 Mail Agent를 구현하였다. 실시간의 음성처리가 필요하므로 이를 위해 RTP를 사용하였으며, 사용자의 음성을 Capture하고 RTP를 통해 Mail Agent로 보내기 위해 애플릿을 사용하였다. 본 시스템의 모든 컴포넌트들을 자바로 구현함으로써 최대한의 이식성을 얻을 수 있었다.

## 1. 서론

최근 컴퓨터 및 통신 기술의 발전으로 통신 사용자가 늘어남에 따라 정보의 획득과 다양한 서비스를 제공받기 위한 매체로서 인터넷이 각광받고 있다. 그 중에서도 화상 회의나 인터넷 폰, 음성 메일 등의 멀티미디어 서비스에 관심이 집중되고 있다.

대부분의 기존 시스템에서는 메일 전송 표준 프로토콜인 SMTP(Simple Mail Transfer Protocol)를 사용하고 있으며 메시지 형식으로는 RFC 822의 규격에 따르고 있다. 그러나 이를 기반으로 한 전자 메일은 7bit의 아스키 텍스트만을 처리할 수 있다는 치명적인 약점을 가지고 있다.

따라서 아스키 텍스트를 포함한 다양한 문자 집합과 다양한 형태의 데이터 전송을 위해 MIME(Multipurpose Internet Mail Extension)이 제안되었다.

실시간으로 사용자가 음성 메일을 보내기 위해서는 실시간 스트리밍을 지원하는 새로운 프로토콜을 필요로 하게 되는데, 이러한 목적으로 IETF에서 제안한 것이 바로 RFC 1889 RTP(Realtime Transfer Protocol)이다.

웹을 통한 멀티미디어 메일 서비스에 대한 연구는 이루어져 왔으나 실시간의 스트리밍을 요구하는 시스템에 대한 고려는 없었다[1].

사용자는 이 시스템을 통해 별도의 소프트웨어 도움 없이 이 서비스를 제공하는 웹사이트에 접속만 함으로써 언제든지 음성 메일을 보낼 수 있다.

## 2. 관련연구

### 2.1 CODEC

음성 메일을 보내기 위해서는 먼저 사용자의 음성 신호를 디지털 신호로 바꾸어야 한다. 이는 보통 PCM(Pulse Code Modulation) 방식을 통해 이루어지며, 결과 디지털 신호의 크기는 다음과 같은 공식에 의해 계산될 수 있다.

$$\text{Storage Space} = \text{Capture Size} \times \text{Sampling Resolution} \times \text{Channel}$$

만약 CD 음질의 데이터를 위의 식에 대입해 보면(44KHz × 16bits per Sample × 2 Channels), 1초당 약 166K를 필요로 한다는 것을 알 수 있다. 따라서 소리 데이터의 저장공간 요구량을 줄이기 위해 CODEC을 사용한다[5]. 사람의 음성 데이터에는 주로  $\mu$ -law, a-law, G.723.1 등의 방식을 사용한다. 인터넷을 통해 음성 데이터를 주고 받을 경우에는 그 음성 데이터의 크기를 최대한 줄이는 것이 요구되는데, 위의 CODEC을 사용하여 압축할 경우 합리적인 효과를 거둘 수 있다.

### 2.2 POP3와 IMAP4

인터넷 상의 조그만 노드에서는 SMTP 서버와 같은 MTS(Message Transport System)를 유지하는 것이 비실용적이다[2]. 그럼에도 불구하고 이러한 노드에서는 UA(User Agent)의 도움으로 메시지를 다룰 필요성이 있는데 이를 위한 프로토콜이 POP3와 IMAP4이다. 이는 각각 RFC 1939와, RFC 1730에 기술되어 있다.

### 2.3 MTA(Message Transfer Agent)와 UA(User Agent)

전자 메일 시스템은 크게 MTA와 UA 두 가지로 구성된다. 실제 전자메일을 전송하는 부분을 MTA라고 하며, 사용자가

\* 본 연구는 첨단정보기술 연구센터를 통하여 과학재단의 지원을 받았다

응용 프로그램으로 전자 메일을 작성하고 MTA에 접근하는 부분을 UA라 한다.

### 2.4 RTP(Realtime Transfer Protocol)

HTTP 또는 FTP와 같은 일반 목적의 프로토콜과 달리 RTP는 엄격한 시간 제한을 가진 미디어 스트림을 전송하기 위해 제안되었다. 화상 회의 시스템이나 음성 채팅, VOD 등의 시스템에 걸쳐 그 응용 범위가 넓으며, 또한 화상 회의 프로토콜인 H.323에서도 RTP를 이용한다.

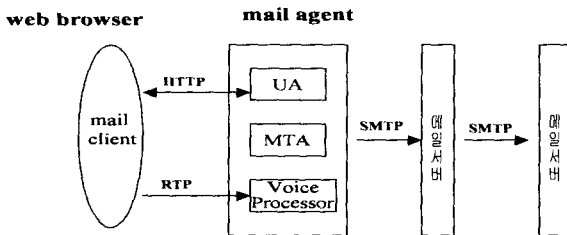
### 2.5 Servlet, JMF, JavaMail

Servlet은 효율적인 Web 응용 시스템을 구축할 수 있게 해주며[3, 4], JMF(Java Media Framework)을 통해 자바 애플릿이나 응용 프로그램에서 쉽게 멀티미디어 데이터를 조작할 수 있다[8]. JavaMail은 자바를 통해 메일을 전송하거나 수신할 수 있게 하는 프레임워크를 제공한다[7].

## 3. 음성 메일 전송 시스템

### 3.1 시스템 구성도

본 시스템의 송수신 과정은 그림 1과 같다.



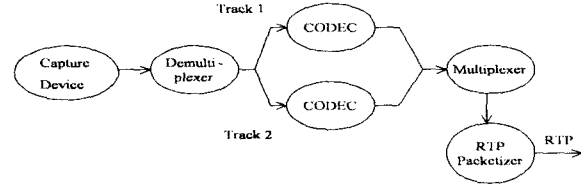
<그림 1> 음성 메일 전송 시스템

음성 메일 전송 시스템의 핵심 기능은 mail agent이다. mail agent는 서블릿으로 구현되며, 그 구성 요소로서 UA, MTA, Voice Processor를 가진다. mail client는 단지 메일 전송에 필요한 정보만을 사용자로부터 입력받아 mail agent에 전달하는 역할을 수행한다. Voice Processor는 RTP를 통해 들어오는 사용자의 음성 신호를 처리한다. UA는 HTML Form문에 의해 일은 전자메일의 구성요소 From, Date, To, Send 등에 기반하여 메시지를 만드는 역할을 하며, 만든 메시지를 MTA에 전달한다. MTA는 SMTP를 통해 MIME 형태의 메일을 메일 서버로 전송하는 역할을 한다.

### 3.2 구성 모듈

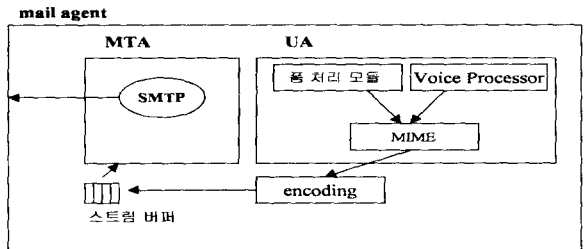
메일 클라이언트는 자바 애플릿과 form-based HTML로 구성된다. 전자 메일 주소 등 메일 전송시 필요한 정보는 form 문을 통해 mail agent로 전송된다. 자바 애플릿은 음성을 capture 장치로부터 PCM 방식으로 capturing하는 음성 capture 모듈과 PCM 음성 신호를 압축하는  $\mu$ -law CODEC, 마지막으로 압축된 음성 신호를 RTP 프로토콜을 통해 mail agent로 전송하는 RTP Packetizer로 구성된다. 애플릿의 실행은 브라우저 내의 보안 관리자에 의해 철저히 관리되므로 애플릿을 통해 지역 시스템에서 직접 음성을 capture하는 것은 허

용되지 않는다. 이를 해결하기 위해 signed applet을 사용하거나 또는 각 브라우저의 특정 보안 정책을 사용할 수 있다[6]. 애플릿의 음성 처리는 그림 2와 같다



<그림 2> 음성 신호 처리

mail agent는 사용자를 대신해 메일을 전송해 주는 서블릿으로 만들어진 핵심 모듈이다. CGI 방식이 사용자로부터 요청이 있을 때마다 새로운 프로세스를 사용하여 처리하는 반면 서블릿은 쓰레드로 처리되기 때문에 서버의 성능 향상에 크게 도움이 된다[4]. mail agent의 구성요소는 그림 3에 있다.



<그림 3> mail agent의 구성 요소

구성 모듈별로 기능을 살펴보면 다음과 같다.

- 폼 처리 모듈은 사용자가 HTML 폼에 입력한 데이터를 꺼내 오는 역할을 한다.
- Voice Processor는 RTP를 통해 사용자의 음성 신호를 서버에 임시적으로 wave format으로 저장하는 역할을 한다.
- MIME 처리 모듈과 encoding 모듈은 폼 처리 모듈 Voice Processor로부터 얻은 정보를 MIME type으로 메시지를 생성하고 Base64 방식으로 인코딩하는 역할을 한다.
- 마지막으로 SMTP 모듈은 SMTP 프로토콜을 사용하여 MIME 메시지를 메일 서버로 전송하는 역할을 한다.

java에서는 JMF와 JavaMail을 통해 메일 전송과 수신, 멀티미디어 데이터 처리에 관련된 프레임워크를 제공하므로 이를 분히 활용하여 구현하였다.

### 3.3 MIME Message 생성

JavaMail API는 메일 시스템을 구성하는 객체를 정의하고 추상 클래스들을 제공한다. 예를 들어, API는 Message, Set Transport와 같은 클래스들을 정의한다. API는 필요할 때 새로운 기능을 추가하거나 새로운 프로토콜을 만들고자 할 때 사용되어서 사용될 수 있다. API는 또한 추상 클래스의 하위 클래스로서 RFC822와 RFC2045에 순응하는 MimeMessage, MimeBodyPart 등도 제공한다. 본 논문에서는 MIME 메시지를 만드는 데 있어 이러한 클래스들을 전적으로 활용하였다.

본 논문에서는 음성 메일에 대한 MIME 메시지에 사용

음성 메시지와 글 메시지 둘 다 포함시켜야 하므로 multipart를 사용하였다. multipart는 한 메시지 안에 여러 가지 콘텐츠를 포함시킬 수 있는 구조로서 각 part는 boundary에 의해 구별된다. 그림 4는 multipart 메시지를 보여준다.

메일 전송 프로토콜인 SMTP는 7bit 데이터만 전송 가능하므로 8bit 데이터로 구성된 멀티미디어 데이터는 7 bit 데이터 형식으로 변형되어야 한다. Base64기법을 이용하여 멀티 미디어 데이터 7bit로 변형하였다.

```
Return-Path: <siba@spring.rit.co.kr>
Received: from autumn.rit.co.kr (IDENT:root@autumn.rit.co.kr
[210.92.44.35])
    by spring.rit.co.kr (8.9.3/8.9.3) with SMTP id BAA05191
    for <siba@spring.rit.co.kr>; Sat, 4 Mar 2000 01:23:08 +0900
Message-ID: <649260196.952074240918.JavaMail.root@autumn.rit.co.kr>
Date: Fri, 3 Mar 2000 18:04:00 +0900 (GMT+09:00)
From: siba@spring.rit.co.kr
To: siba@spring.rit.co.kr
Subject: =?EUC-KR?B?xde9usau?= // test란 문자였음
Mime-Version: 1.0
Content-Type: multipart/mixed;
boundary="=-1559040857.952074240740.JavaMail.root@autumn.rit.co.kr"
X-UIDL: aac45d052881b4fa588fc966a2be3a8

---1559040857.952074240740.JavaMail.root@autumn.rit.co.kr
Content-Type: text/plain; charset=EUC-KR
Content-Transfer-Encoding: base64

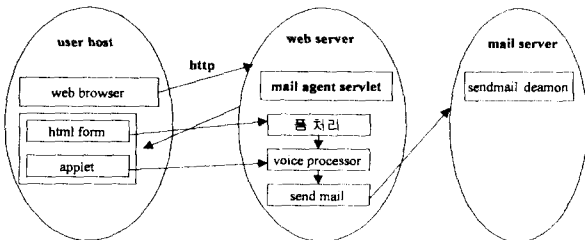
xde9usau // test란 문자였음
---1559040857.952074240740.JavaMail.root@autumn.rit.co.kr
Content-Type: audio/basic; name=ulaw.wav
Content-Transfer-Encoding: base64
Content-Disposition: attachment; filename=ulaw.wav

// 실제 ulaw.wav에 실린 byte stream이 뒤 따른다.
---1559040857.952074240740.JavaMail.root@autumn.rit.co.kr---
```

<그림 4> multipart message

### 3.4 웹 브라우저에서의 실행 과정

웹 브라우저를 통해 해당 시스템에 HTTP Protocol를 통해 접속하게 되면, HTML문서와 함께 음성을 처리하는 애플릿을 다운로드 받는다. 애플릿은 사용자의 음성을 capture하여 RTP를 통해 서버의 Voice Processor로 보내게 되며, Voice Processor는 그 음성 신호를 wave 포맷의 파일로 임시 저장한다. 사용자가 최종적으로 메일 전송에 필요한 모든 정보를 form문에 기입하고 mail agent에 보내게 되면 mail agent는 Voice Processor에 의해 생성된 wave파일과 form문을 통해 얻은 정보에 근거하여 multipart로 구성된 MIME형식의 메시지를 생성하여 메일 서버로 SMTP를 통해 전송하게 된다.



<그림 5> 음성 메일의 전송 동작 구조

## 4. 구현

본 논문에서 제안한 음성 메일 전송 시스템은 Java 1.1 VM 이상을 지원하는 브라우저에서 작동된다. 메일 클라이언트는 음성 capture 및 전송을 위해 애플릿을 사용한다. 클라이언트를 대신하여 메일을 전송해 주는 mail agent는 자바 서블릿 기술을 사용하여 구현하였다.

applet의 녹음시작 버튼을 통해 음성 녹음이 시작되며, 녹음 끝 버튼을 통해 음성 녹음을 종료한다. 마지막으로 HTML form문에 적정 정보를 입력한 후 '메일전송' 버튼을 통해 메일을 보낼 수 있다.

## 5. 결론

자신의 데스크탑 PC가 없어도 언제든지 메일을 송·수신할 수 있다는 이점 때문에 Web을 통한 메일 서비스는 큰 매력에 갖는다. Web을 통해 다양한 콘텐츠를 가진 메일을 송수신할 수 있는 시스템이 요구되고 있는 시점이다. 따라서 웹을 통해 음성 메일을 보낼 수 있는 시스템을 개발하였다. 음성 메일을 보내고자 하는 사용자는 웹 서버에 접속함으로써 언제든지 그 서비스를 이용할 수 있다. 웹 기반의 시스템이므로 서버에 부하가 집중될 수 밖에 없으므로, mail agent는 사용자의 요구를 thread로 처리하는 서블릿을 사용하여 구현하였다. 또한 모든 컴포넌트들이 자바를 이용해 개발되었으므로 어느 플랫폼에서도 수정 없이 사용 가능하다.

Web을 통해 음성 메일을 받아 보는 것도 이 시스템을 약간 확장하여 쉽게 구현할 수가 있다. 기존 PSTN망을 통해 인터넷 계정에 음성 메일을 보낼 수 있는 시스템을 만드는 것이 다음 과제라 하겠다.

## [참고 문헌]

- [1] 김성후, "WWW상에서 운용되는 전자메일 시스템의 설계 및 구현", 한국정보과학회 논문지(B), 제 23권 제 2호, 1996, pp. 1213-1216.
- [2] H. Schulzrinne, "RTP: A Transport Protocol for Real-Time Applications," RFC1939, 1996.
- [3] D. Ayers, et al., Professional Java Server Programming, Wrox Press, 1999.
- [4] D. R. Callaway, Inside Servlets, Addison-Wesley, 1999.
- [5] L. deCarmo, Core Java Media Framework, Prentice Hall, 1999.
- [6] MageLang Institute, "Fundamentals of Java Security," <URL:http://developer.java.sun.com/developer/onlineTrain/ing/Security/Fundamentals/index.html>
- [7] Sun, JavaMail API Design Specification, Sun, 1998.
- [8] Sun, Java Media Framework API Guide, Sun, 1999.