

# 모바일 VoIP 시스템

이경희\* · 설동명\* · 광지영\* · 김연희\* · 김두현\*

## 1. 서 론

통신망의 고속화, 컴퓨터 처리 속도의 향상, 데이터의 압축 및 복원 알고리즘의 개발 등으로 인해 다양한 멀티미디어 서비스들이 개발되고 있다. 이들 서비스 중의 하나로서 VoIP(Voice Over Internet Protocol) 서비스가 있다. 이는 인터넷 프로토콜(IP)을 이용하여 음성신호를 실시간으로 전달하는 기술을 의미하며, 모바일 VoIP 시스템은 통신 환경을 모바일 환경으로 발전시킨 것을 의미한다.

이들 시스템간의 호환성을 위해 사용되어 지는 프로토콜은 기존의 멀티미디어 단말 규격 혹은 영상 회의 시스템의 규격을 대부분 이용하고 있다. 이들 영상 회의와 관련된 표준안 중 대표적인 것이 ITU-T(International Telecommunication Union-Telecommunication Sector)의 H.323[7]과 IETF(Internet Engineering Task Force)의 SIP(Session Initiation Protocol)[11]이 있다.

본 고는 모바일 환경의 VoIP 시스템에 대한 것으로서, VoIP 시스템을 위한 대표적인 국제 표준 규약의 구조 및 구성에 대한 것과 모바일 환경하에서 VoIP 시스템을 구성하는 것에 대해 기술하였다.

## 2. VoIP 관련 표준 프로토콜

통신 프로토콜의 국제 표준을 권고하는 기관들 중 영상 회의와 관련된 권고안을 제안한 대표적인 기구로서 ITU-T와 IETF를 들 수 있다. 이들 기관이 제정한 국제 표준 프로토콜 중 VoIP 시스템에서 많이 사용되어 지는 것이 ITU-T의 H.323 프로토콜과 IETF의 SIP이다. 본 장에서는 ITU-T와 IETF에서 권고한 VoIP 관련 프로토콜에 대해 기술한다.

### 2.1 ITU-T의 H.323

#### 2.1.1 ITU-T의 권고안

ITU-T의 권고안 중 H.320[6], H.324[8], H.323[7]은 영상 회의와 관련된 표준안이다. 이중 H.320은 ISDN(Integrated Services Digital Network)상의 영상 회의 표준 권고안이며, H.323은 LAN(Local Area Network)을 위한 표준 권고안이고, H.324 권고안은 GSTN(Generalized Switched Telephone Network)을 위한 영상 회의 표준 권고안이다. 각 시스템들이 사용하는 오디오, 비디오 압축/복원 알고리즘은 유사하나 전송 프로토콜 및 콘트롤 프로토콜은 전달 매체의 특성에 따라 다른 프로토콜 스택을 사용한다.

이 중 H.324 시스템은 기존의 전화망을 이용하

\*한국전자통신연구원

표 1. ITU-T H.32x 영상 회의 규격 및 관련 권고안

권고(승인)	권고 제목	비디오	오디오	스트림 전송	제어	다지점	일반 인터페이스
H.320 (1990)	Narrow-band ISDN visual telephone systems and terminals equipment	H.26x	G.7xx	H.221	H.230, H.242	H.231, H.243	I.400
H.323 V1/V2 (1998)	Packet-based multimedia communications systems	H.26x	G.7xx	H.225.0	H.245	H.323	TCP/IP
H.324 (1996)	Terminal for low bitrate multimedia communication	H.26x	G.7xx	H.223	H.245		V.34 modem

여 오디오, 비디오 등 멀티미디어 통신을 하는 시스템의 규약이다. 이는 Internet Protocol을 이용하지 않고 전화망에서 모뎀을 이용하여 오디오, 비디오, 콘트롤 데이터 등의 디지털 데이터를 멀티플렉싱한 후 모뎀을 통해 상대방에게 전송한다. 기본적인 H.324 시스템을 확장하여 현재의 이동통신망과 같은 환경에 적합하도록 콘트롤 데이터 및 멀티미디어 데이터의 전송시 기존 전화망을 이용하는 것보다 전송 에러의 발생 확률이 높은 환경에도 이용할 수 있도록 표준안을 보강한 것으로 H.324M이라는 프로토콜이 있다. 그러나 H.324M도 H.324와 마찬가지로 IP상에서 동작하는 프로토콜이 아니고 모뎀을 이용하여 오디오, 비디오, 콘트롤 데이터 등을 멀티플렉싱하여 송수신한다.

2.1.2 H.323의 구성 요소

H.323 시스템은 IP망을 기본 통신망으로 이용하는 영상 회의 규약이다. H.323 영상 회의 시스템을 구성하는 구성 요소로는 단말기, MC(Multipoint Controller), MP(Multipoint Processor), Gateway, Gatekeeper, MCU(Multipoint Control Unit) 등이 있다. Terminal은 오디오를 필수적으로 입출력할 수 있어야 하며 다른 단말기 혹은 Gateway, Gatekeeper 등과 연결되는 종단(endpoint)이다. MC는 다자간 회의를 콘트롤하는 부분으로 참석하는 단말기의 오디오, 비디오 입출력 능력(capability)

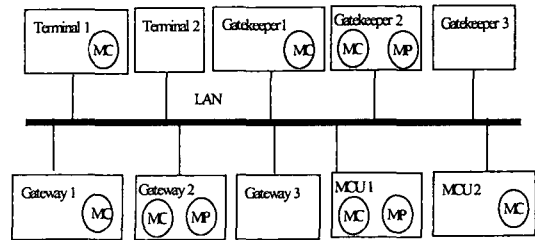


그림 1. H.323 영상 회의 시스템의 구성 요소

정보를 분석하여 회의에서 공통적으로 사용할 수 있는 통신 모드를 결정하고 멀티캐스트 주소 등을 관리한다. MC는 MCU에 필수적인 컴포넌트로서 포함되어지며 ad hoc 다자 회의(ad hoc multipoint conference)에 사용되어진다. MP는 다자간 영상 회의에서 오디오 스트림들 혹은 비디오 스트림들을 믹싱하는 등 멀티미디어 데이터를 가공, 처리하는 부분이다. MCU는 MC와 MP를 가지고 있는 요소로서 다자간 영상 회의를 지원하는 장치이다. Gateway는 PSTN(Public Switched Telephone Network) 등 다른 통신망상에 존재하는 단말기와의 통신을 위한 요소이다. Gatekeeper는 단말기, MCU, Gateway 등을 위한 주소 변환, 접근 제어 등을 행하며 H.225.0에 정해진 RAS(Registration, Admission and Status) 메시지를 이용하여 정보를 교환한다.

2.1.3 H.323 단말의 구조

H.323은 ITU-T에서 권고한 것으로서 LAN을

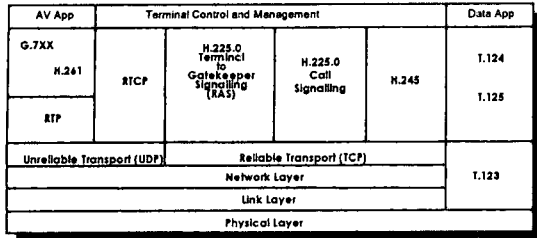


그림 2. H.323 단말의 구조

기반으로 하는 영상 회의 시스템의 표준 규약이다. 현재 출시되고 있는 대부분의 H.323 제품들은 TCP/IP를 통신 프로토콜로 이용하고 있으며 H.320, H.324 시스템과 다르게 오디오, 비디오, 콘트롤 신호를 다중화하여 전송하지 않고 독립적인 채널을 이용하여 송수신한다.

H.323을 구성하는 프로토콜 스택은 크게 H.225.0[3], H.245[5], 오디오 코덱, 비디오 코덱 등이다.

이중 H.225.0은 크게 RAS, Q.931, 그리고 RTP[9], RTCP[10]로 구성되어 있다. RAS는 게이트키퍼와 시그널링을 할 때 사용되어지는 메시지와 절차를 정의하고 있다. Q.931은 상대방에게 호출 메시지를 송수신하는 규약으로 TCP와 같은 신뢰성있는 통신 프로토콜을 이용하여 ASN.1(Abstract Syntax Notation 1)으로 인코딩된 메시지를 송수신한다. RTP는 UDP(User Datagram Protocol)을 기반으로 하여 유니캐스트 및 멀티캐스트를 이용하여 오디오 및 비디오 데이터를 전송한다. RTCP는 RTP를 이용하여 전송되는 데이터의 지연(delay), 지터(jitter) 및 동기를 콘트롤하는 정보를 송수신하는데 사용되어지는 프로토콜이다.

H.245는 영상 회의 진행을 위한 콘트롤 및 상태 정보(indication)를 전송하는 프로토콜로서 각 메시지들은 ASN.1 표기로 기술되어져 있으며 메시지 송수신시에 메시지를 ASN.1 규약에 따라 인코딩하여 전송한다. H.245에는 여러 개의 독립적인

state machine을 가진 엔티티(entity)들이 포함되어 있어 상대방 단말기와 request, response 메시지를 송수신하면서 엔티티에서 정해진 절차를 수행하거나 command 혹은 indication 메시지를 이용하여 회의 콘트롤 절차를 수행한다.

## 2.2 IETF의 SIP

IETF의 SIP는 임의의 세션을 생성할 때 필요한 메시지를 전달하는 규약이다. 콘트롤 메시지 전달에서 SIP는 텍스트 기반의 문자열 메시지를 송수신하지만 H.323은 ASN.1(Abstract Syntax Notation 1) 방식으로 메시지를 인코딩하여 송수신한다. 따라서 상대방과의 메시지 송수신시 SIP는 H.323에 비해 메시지를 인코딩하고 디코딩하는데 H.323에 비해 상대적으로 간결할 뿐만 아니라 호출을 시작하고 데이터 통화가 최초로 시작될 때까지 걸리는 메시지 송수신의 회수도 적은 장점 등이 있다.

SIP는 네트워크 세션을 만들고 수정하며 해제하는 프로토콜이다. 현재 SIP는 멀티미디어 세션 제어와 인터넷 전화 서비스에서 사용되는 프로토콜로 많이 개발되고 있으며 그 대표적인 특성은 간결성(lightweight), 전달 망에 독립성(transport independent), 텍스트 기반(text-based)의 프로토콜이라고 할 수 있다. SIP의 특성들을 살펴보면 첫째 간결성, 단지 6가지의 방법(method)으로만 구성되어 있어 간결하고 이들 방법들은 서로 조합하여 멀티미디어 세션을 완벽하게 제어할 수 있다. 둘째 전달 망에 독립성, 다양한 데이터 전달 프로토콜(UDP, TCP, ATM.)들을 사용할 수 있다. 셋째 텍스트 기반, 쉽게 구문을 분석할 수 있고 확장성이 좋다[11].

SIP의 기본 개념은 이 메일 주소나 전화번호 등으로 구별할 수 있는 사용자나 호스트간의 세션

을 설정하는 것이다. 사용자들은 접속 위치가 바뀌거나 다른 접속 장치를 사용하여도 같은 식별자를 유지할 수 있다. 식별자는 네트워크 제공자나 전화 서비스 제공자가 부여할 수 있다.

SIP는 호출을 위한 메시지와 절차를 정의하고 있으며, 통화에 관련된 부가 정보의 송수신을 위해 SDP(Session Description Protocol)를 이용하고 있다. 통화관리를 위한 서버로 Registrar, Redirect Server, Proxy Server를 정의하여 이용하며, 단말 측에서 호출을 개시하는 User Agent Client, 호출을 수신하는 User Agent Server를 규정하고 있다.

### 2.2.1 SIP의 구성 요소

SIP 영상 회의 시스템의 구성 요소 중 서버로 분류되는 것들에는 Redirect Server, Proxy Server, Registrar 등이 있으며 단말측에서 운영되는 요소로 User Agent Server, User Agent Client로 불리는 User Agent 등이 있다.

User Agent Client는 상대방에게 호출 메시지를 생성하여 생성된 호출 메시지를 Redirect Server, Proxy Server 혹은 User Agent Server 등으로 전달하는 역할을 수행한다. User Agent Server는 서버들 혹은 User Agent로부터 전달되는 호출 메시지를 수신하여 이를 처리하는 역할을 수행한다. 일반적으로 User Agent Server와 User Agent Client는 같은 단말내 혹은 하나의 모듈로 구현되어 진다.

서버들 중 Registrar는 사용자의 등록을 받아 관리하는 역할을 수행한다. Registrar는 독립적으로 구현될 수도 있지만 Redirect Server 혹은 Proxy Server 등에 포함되어져 구현된다. Redirect Server는 User Agent Client로부터 전달되는 호출 메시지를 분석하여 호출을 요구한 User Agent Server의 주소를 알려주어 User Agent Client가

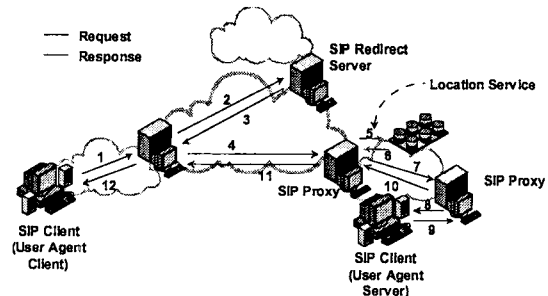


그림 3. SIP 서비스 시스템 구성도

원하는 User Agent Server를 호출하여 통화가 되도록 콘트롤하는 역할을 수행한다. Proxy Server는 User Agent Client가 전달하는 메시지를 수신하여 해당하는 User Agent Server를 호출하도록 하며 User Agent Client가 송신하는 메시지들을 User Agent Server로 전달함과 동시에 User Agent Server가 송신하는 메시지를 User Agent Client로 전달하는 기능을 수행한다.

### 2.2.2 SIP 메시지의 구성

SIP 프로토콜은 텍스트 기반으로 되어 있으므로 많은 부분이 HTTP 프로토콜의 메시지와 구성에 있어서 매우 유사하다. 즉, SIP 메시지는 헤더(header) 필드와 바디(body)로 구성되며 텍스트 문자로 작성된다. 헤더 필드에는 SIP 제어 정보 및 경로 정보 등이 포함되며, 바디부분에는 호 설정 시 단말 시스템 양측의 오디오 및 비디오 코덱과 같은 성능에 대해 서로의 정보를 알리고 협상하기 위한 내용이 기술되는데, 이는 SDP(Session Description Protocol)에 따라 작성된다. 또한, 바디부분은 다양한 멀티미디어 정보를 포함할 수 있으며, 이러한 바디 정보는 SIP 헤더필드인 content-type헤더필드에서 MIME(Multipurpose Internet Mail Extension) 타입으로 표시된다.[11]

SIP 메시지의 종류는 6개의 기본 메소드(Method)와 그에 대한 응답으로 구성되는데, 요청(Request) 메시지와 응답(Response) 메시지의 형태

이다. 6개의 기본 메소드에는 호 설정 요청을 위한 INVITE, 호 설정을 확인하는 ACK, 호 설정을 종료하는 BYE, 호 설정 도중에 호를 취소하는 CANCEL, 위치 등록을 위한 REGISTER, 그리고 호 설정 없이 상대방의 성능을 파악하는 OPTION 등이 있으며, 이 밖에 필요에 따라 extension 메소드를 정하여 이용할 수 있다.

### 3. 모바일 VoIP 시스템 구성의 예

#### 3.1 모빌리티(Mobility)

개인용 PC(Personal Computer)를 이용하는 경우에 비해서 PDA(Personal Digital Assistant) 혹은 이동 전화를 이용하는 경우에는 단말을 휴대하고 위치를 쉽게 이동할 수 있다. 이와 같은 이동 환경에서 사용자들에게 제공되어지는 모빌리티(mobility)는 크게 세가지가 있다. 이 중 첫째는 사용하는 단말이 이동하여도 계속하여 서비스를 받을 수 있는 단말 모빌리티(terminal mobility), 둘째는 사용자가 이동하여 이동한 곳의 새로운 단말을 이용하여도 계속하여 서비스를 받을 수 있는 사용자 모빌리티(user mobility), 그리고 셋째로 어느 단말 혹은 어디에서도 원하는 서비스를 지속적으로 받을 수 있는 서비스 모빌리티(service mobility) 등이 있다.

사용자의 이동성을 제공하는 기본적인 예로는 가정용 전화를 휴대전화로 포워딩시키는 것을 예로 들 수 있다. 이를 구현하기 위해서는 통화를 관리하는 서버에게 사용자의 인식자(ID 혹은 전화번호)와 최종 이동 단말의 위치(전화번호 혹은 IP 주소 등)를 알려 줌으로써 호출이 자동적으로 포워딩되게 하여 원하는 곳에서 호출을 받도록 하면 된다. H.323 시스템의 경우에는 게이트키퍼에게 이동한 곳의 위치(IP 주소)를 알려 주면 되고 SIP의 경우에도 Registrar에게 이동한 곳의 위치

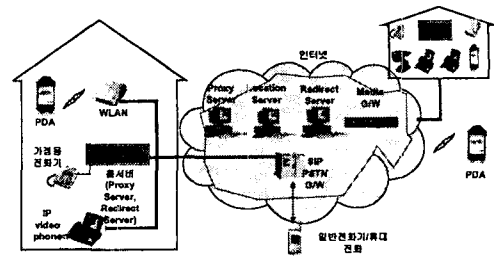


그림 4. SIP 기반 모바일 VoIP 시스템의 구성

를 알려 주면 된다.

단말의 이동성은 크게 두가지로 나누어 생각할 수 있다. 하나는 통화 대기중에 이동하는 것이고 다른 하나는 통화 중에 이동을 하는 것이다. 이때 통화 대기 중에 다른 지역으로 이동하여도 계속 서비스를 받게 하는 것을 로밍(roaming)이라고 하며 통화 중에 인접한 다른 지역으로 이동하여도 계속 통화를 할 수 있게 하는 것을 핸드오프(hand-off)라고 한다. 따라서 사용자 인터페이스 이하의 프로토콜 계층에서 로밍과 핸드오프의 기능을 수행하게 하여 사용자에게 사용의 편의성을 제공하도록 하여야 한다.

#### 3.2 시스템 구성의 예

본 절에서는 SIP 기반하여 모바일 VoIP 시스템을 구성한 예를 보인다. 그림 4에 구성을 나타낸 그림이 있다. 그림에서 VoIP 프로토콜을 이용하는 이동 단말은 PDA 등이 해당되고 이는 WLAN(Wireless Local Area Network) 등을 이용하여 무선통신망에 접속되도록 한다. 그리고 서버들은 Location Server, Proxy Server, Redirect Server 등 SIP에서 구성 요소로 이용되어 지는 서버들은 태내 혹은 인터넷 상에 배치하였다.

모바일 단말들간의 통신 뿐만이 아니라 기존 PSTN 혹은 휴대 전화들과도 연동이 되게 하려면 SIP-PSTN G/W 등에서 시그널링 게이트웨이와 미디어 게이트웨이 등의 기능 수행이 요구되어

진다.

사용자 혹은 단말의 이동시에 Location Server 등의 서버들에게 사용자의 위치 변경 등록을 하고 통화 중인 경우 RTP 및 이하의 계층에서 통신을 계속 진행하도록 하면 사용자 모빌리티와 단말 모빌리티를 제공할 수 있을 것이다.

#### 4. 결론

VoIP는 인터넷 프로토콜을 이용하여 음성 신호를 전달하는 기술이며 VoIP 시스템들을 위한 표준 규약으로는 대표적인 것이 ITU-T의 H.323과 IETF의 SIP가 있다.

모바일 환경에서는 기본적으로 사용자 모빌리티, 단말 모빌리티, 서비스 모빌리티를 고려하여야 하며, 이를 제공하기 위해서는 물리적인 통신망과 사용자 등록 서버 등을 이용하여야 할 것이다.

인터넷을 이용하는 사용자가 증가함에 따라 VoIP의 수요도 증가할 것이며 이동 환경으로 통신 환경이 급변함에 따라 모바일 VoIP 시스템도 각광을 받을 것으로 예상된다. 이들 VoIP 서비스 사용자들을 위해서는 저렴한 가격과 우수한 음성 혹은 영상 품질을 제공해야 할 것이다. 따라서 기존 아날로그 혹은 이동 전화에 비교하여 지연 혹은 지터, 패킷 손실 등을 축소하거나 이러한 품질 저하 요인에도 음질의 손실을 최소로 줄이는 노력이 지속적으로 이루어져야 하며, 사용자 혹은 단말의 이동시의 모빌리티 제공을 위한 방법연구 그리고 이와 아울러 VoIP를 이용한 다양한 서비스의 개발도 병행되어야 할 것이다.

#### 참고 문헌

- [ 1 ] ITU-T, "H.221: Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices," ITU-T, July 1997
- [ 2 ] ITU-T, "H.223: Multiplexing protocol for low bit rate multimedia communication," ITU-T, March 1996
- [ 3 ] ITU-T, "H.225.0 Media Stream Packetization and Synchronization on Non-Guaranteed Quality of Service LANs," ITU-T, May 1996
- [ 4 ] ITU-T, "H.242: System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s," ITU-T, July 1997
- [ 5 ] ITU-T, "H.245: Control protocol for multimedia communication," ITU-T, Feb. 1998
- [ 6 ] ITU-T, "H.320: Narrow-band visual telephone systems and terminal equipment," ITU-T, July 1997
- [ 7 ] ITU-T, "H.323: Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service," ITU-T, Feb. 1997
- [ 8 ] ITU-T, "H.324: Terminal for low bit-rate multimedia communication," ITU-T, Feb. 1998
- [ 9 ] Schulzrine, Casner, Frederick, and Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 1889, Internet Engineering Task Force, Feb. 1996
- [10] Schulzrine, "RTP Profile for Audio and Video Conferences with Minimal Control," RFC 1890, Internet Engineering Task Force, May 1996
- [11] Handley, Schulzrinne, Schooler, and Rosenberg, SIP: Session Initiation Protocol, RFC 2543, Internet Engineering Task Force, November 2000



이 경 희

- 1990년 경북대학교 컴퓨터공학과 졸업(공학사)
- 1992년 경북대학교 대학원 컴퓨터공학과 졸업(공학 석사)
- 1999년 전자계산조직응용기술사 취득
- 1992년~현재 한국전자통신연구원 선임연구원
- 관심분야: 실시간 스트리밍, 영상 회의 시스템, 그룹웨어
- e-mail: kyunghee@etri.re.kr



곽 지 영

- 1999년 전남대학교 컴퓨터공학과 졸업(학사)
- 2001년 전남대학교 일반대학원 컴퓨터공학과 졸업(공학 석사)
- 현재: 한국전자통신연구원 컴퓨터소프트웨어기술연구소 정보가전연구부 모바일협동작업연구팀 재직 중
- 관심분야: Traffic Control, Scheduling, Performance Evaluation, VoIP, SIP
- e-mail: jiyoung@etri.re.kr



설 동 명

- 1990~1996: 충남대학교 컴퓨터공학과(공학사)
- 1997~1998: 충남대학교 컴퓨터공학과(공학석사)
- 1999~현재: 한국전자통신연구원 연구원
- 관심분야: Multimedia Collaboration, VoIP(Voice over IP), 인터넷 정보가전 멀티미디어 서비스
- e-mail: dmsul@etri.re.kr



김 연 희

- 2001년 서울대학교 컴퓨터공학부(공학사)
- 2001년~현재 한국전자통신연구원 모바일협동작업연구팀, 연구원
- 관심분야: 모바일 네트워킹, 공동작업프로토콜



김 두 현

- 1981~1985: 서울대학교 컴퓨터공학과(공학사)
  - 1985~1987: 한국과학기술원 전산학과(이학석사)
  - 1991~1993: 미 스탠포드연구소 객원연구원
  - 1993. 12.: 정보처리기술사 취득
  - 1987~현재: 한국전자통신연구원 책임연구원  
멀티미디어그룹웨어연구팀장
  - 관심분야: Multimedia Collaboration, VoIP(Voice over IP),  
인터넷 정보가전 멀티미디어 서비스
  - e-mail: doohyun@etri.re.kr
-