

# H.323 기반 VoIP 어플리케이션에서의 대역폭 향상을 위한 방법

김기훈\*, 박동선, 이승상, 박종빈  
전북대학교 공과대학 정보통신학과

## Bandwidth enhancement scheme for VoIP application based on H.323

Kihun Kim\*, Dongsun Park, Seungsang Lee, Jongbin Park  
Information and Communication Engineering Department  
Chonbuk National University

E-mail : \*khkim@multilab.chonbuk.ac.kr, dspark@moak.chonbuk.ac.kr ,  
twobitmail.chonbuk.ac.kr , nicedugy@chollian.net

### Abstract

In this paper, we propose a scheme that applies to the VoIP application based on H.323 protocol to enhance the bandwidth efficiency. We multiplex the audio and video stream. In this scheme, audio frame is carried with video stream. And we apply not only multiplexing but also RTP header compressing to the real audio/video stream to increase the bandwidth efficiency. With the multiplexing and RTP header compressing, we gain the bandwidth efficiency. In the finite network environment, we can assign bandwidth to other users who want to use other service, and other VoIP users. If we can apply the real time network situation to the our VoIP application, we can get more efficient performance.

### I. Introduction

인터넷 서비스는 우리 생활에 있어 필수적인 요소로 자리 잡아가고 있으므로, 앞으로 그에 따른 더 많은 요구사항들을 충족해 주어야 할 것이다. VoIP 서비스 역시 그러한 요구사항들에 대해 민감하게 대처하지 않을 수 없다.

실시간 데이터를 처리하는 VoIP 서비스가 현재 네트워크에서 고려해야 할 요구사항들로는 보안과 QoS 그리고 제한된 네트워크 자원을 어떻게 효율적으로 사

용해야 할 것인가로 나누어 볼 수 있다. 보안 문제는 현재 네트워크에서 다른 어플리케이션들 역시 가지고 있는 문제로써 개인의 사적 자유와 비밀성을 보장해 주어야 한다. QoS 보장 문제와 자원의 효율성은 따로 생각할 수 없는 문제이기 때문에 본 논문에서도 이를 고려하고 있다. 품질이 우수한 서비스를 제공해 주기 위해서는 네트워크 자원을 많이 사용하여야 하고, 제한된 네트워크 자원의 효율성을 극대화하기 위해서는 많은 사용자 수용하여야 하는데, 이는 실시간 데이터를 처리하는 VoIP 와 같은 어플리케이션에서는 통화의 품질 저하를 발생 시킬 수 있다. 그래서 QoS 를 보장해주는 범위 내에서 자원의 효율성을 향상시키기 위한 방법들이 연구 되어지고 있다[8].

본 논문의 구성은 II장에서 이해를 돕기 위한 H.323 프로토콜 스택과 헤더 압축 기술 그리고 사용되어질 음성 코덱에 대해 간단히 기술하고 있으며, III장에서는 제안하고자 하는 방법을 기술하고, IV, V장에서는 제안된 방법을 적용하였을 때 얻을 수 있는 효과를 예상해 볼 것이며, 끝으로 결론을 맺기로 하겠다.

### II. Background

이 장에서는 본 논문과 관련된 H.323 표준들과 RTP

헤더 압축 기법, 그리고 실제 사용되어 지고 있는 영상과 음성 코덱에 대해 간단히 살펴보겠다.

## 2.1 H.323 프로토콜 표준

H.323 기반에서의 VoIP 를 구현하기 위한 관련 프로토콜로는 H.323[1] 단말기 사이의 콜 셋업을 설정하는데 사용되는 H.225[6] 콜 시그널링 프로토콜이 있으며, H.323 단말기나 게이트웨이가 게이트키퍼에 등록하게 되는 H.225 RAS(Registration, Admission, Status)프로토콜이 있고, 두 단말기 사이에 주고받는 컨트롤 메시지와 단말기의 동작에 필요한 미디어 전송 채널에 대한 정보 전송, 플로우 컨트롤 메시지 같은 정보를 주고 받는 H.245[7] 컨트롤 시그널링 프로토콜이 있다. 그리고 실제 영상과 음성을 전송하는 프로토콜로는 RTP[9]가 있으며 RTP 를 통해 전송되는 미디어 전송에 대한 리포팅 서비스를 제공해 주는 RTCP 로 구성되어 있다. 기타 다른 프로토콜에 대해서는 본 논문과의 관련성이 적기 때문에 기술하지 않을 것이며, 기본적인 프로토콜 구조는 그림 1에서 살펴보기로 하겠다.

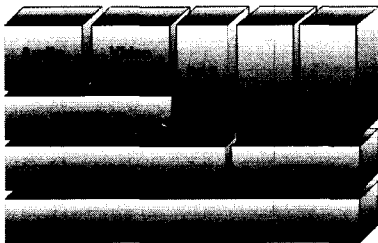


그림 1 H.323 프로토콜 스택

## 2.2 RTP 헤더 압축

본 논문에서 응용할 RTP 헤더 압축은 Casner 와 Jacobson 이 제안한 압축 방법[10]중 IP 와 UDP 를 제외한 RTP 부분의 압축만을 활용할 것이다.

## 2.3 오디오 및 비디오 코덱

단말기가 지원해야 할 오디오 코덱은 강제적으로 G.711 이 있으며 나머지 코덱은 각 단말이 지원해 줄

수 있는 미디어 정보를 교환하는 과정에서 선택할 수 있다. 비디오의 경우 만약 단말이 비디오를 지원해 준다면 H.261 코덱을 지원해 줄 수 있어야 하며 나머지 코덱은 음성과 같이 미디어 정보를 교환하는 과정에서 선택할 수 있다.

본 논문에서 적용해 볼 두 오디오 코덱에 대한 특성을 표 1에서 살펴 보기로 하겠다.

Codec	G.723.1	G.729A
Bitrate(Kb/s)	5.3 / 6.3	8
Frame size(bytes)	20 / 24	10
Frame period(ms)	30	10
Lookahead delay(ms)	7.5	5
Total encoding delay(ms)	37.5	15
Decoding delay(ms)	18.75	7.5
Quality	Fair	Good

표 1 코덱 특성표

## III. Proposed Scheme

본 논문에서 제안하는 방법은 압축률이 높은 음성 코덱을 사용 할 경우, 발생되는 데이터의 양이 매우 적기 때문에 발생된 데이터를 하나의 패킷에 넣어 전송할 경우, 실제 데이터의 전송률보다 데이터를 보내기 위한 헤더의 크기가 더 커지게 된다. 그리고 여러 프레임의 하나의 패킷에 전송하려 할 경우 프레임의 기다림으로 발생하는 지연 때문에 수신측의 음성 품질이 저하될 수 있다. 그래서 제안하는 방법은 영상을 전송하는 패킷에 상대적으로 크기가 작은 음성 데이터를 올려 전송함으로써 IP/UDP 그리고 압축된 RTP 헤더 크기만큼의 대역폭을 절약할 수 있을 것이라 기대를 할 수 있다.

다음은 본 논문에서 사용할 RTP 헤더의 형식과 각각의 비트가 의미하는 것을 표 2에서 살펴 보기로 하겠다.

0	0	R	M	Sequence Number
Full RTP Header(only video)				

(a) 영상만 전송하는 경우

0	1	R	M	Sequence Number
Full RTP Header(audio)				

(b) 헤더가 압축되지 않은 음성을 전송하는 경우

1	0	R	M	Sequence Number
Audio data				

(c) 헤더가 압축된 음성을 함께 전송하는 경우

1	1	R	M	Sequence Number
Video data				

(d) 영상 헤더만 압축해서 전송하는 경우

표 2 사용할 영상 RTP 헤더 형식

네 가지 형식의 헤더를 정의하고 있으며 (a)의 경우 영상 데이터에 음성이 전송되지 않으며 영상의 RTP 헤더 역시 압축이 되지 않은 경우를 의미한다. 그리고 M 비트는 실제 RTP 헤더에서 마커비트를 의미하며 동일하게 사용되어 진다. (a)의 경우에서 의미는 부여 되지 않으며 정의된 헤더 부분만을 떼어 버리고 영상을 처리하는 루틴으로 데이터를 넘겨주게 된다. (b)의 경우 영상 데이터에 음성이 전송되며, 음성의 RTP 헤더가 압축되지 않은 경우이다. 실제 어플리케이션에서는 Full RTP 헤더 부분과 그 이후에 있는 일정한 오디오 데이터의 크기만큼 만을 오디오 처리 루틴으로 넘겨주게 된다. (c)의 경우 음성의 압축된 RTP 헤더와 음성 데이터를 포함 하고 있으므로 정의한 헤더 부분과 코덱에 따라 일정한 오디오 데이터만을 음성 처리 루틴으로 넘겨주게 되면 된다. (d)의 경우 음성이 전송되지 않고, 영상의 RTP 헤더가 압축된 형태를 나타낸다. 그리고 본 논문에서 가정하고 있는 것은 Sequence Number 를 4 비트로 가정하였기 때문에 16 개의 패킷을 전송한 이후에는 송신측과 수신측의 동기를 맞춰주기 위해 압축되지 않은 헤더를 가지고 있는 패킷을 전송하여야 한다. 그리고 헤더가 압축된 경우의 Timestamp 는 Sequence Number 를 통해 얻을 수 있다. 이때 영상과 음성의 Timestamp 의 증가는 상수 값으로 일정하다.

실제 어플리케이션에 본 논문에서 제안한 방법을 구현하였을 때, 실제 통신에서 적용할 것인지의 여부는 각 단말기의 H.245 능력 교환 과정에서 결정하며 네트워크 상황에 따라 그 사용 여부를 능동적으로 결정할 수 있다. 예를 들어 단말기가 RoundTripDelay 나 단말기가 구성 되어진 네트워크에서 능동적으로 네트워크 상황을 판단하여 단말에 전송하여 준다면, 이를 통해 위의 헤더 구조에 있는 R 비트를 이용하여 더 이상 제안

된 방법을 사용할 것인지의 여부를 상대 단말에 알릴 수 있다.

또한 다자회의의 상황에서, MCU(Multipoint Control Unit : 기본적으로 MC 를 가지고 있으며 MP 를 포함할 수 있음)가 포함되었을 경우, MCU 는 MP(Multipoint Process : 다자 회의의 데이터 가공 모듈로 음성 합성 및 분배 등의 역할 수행)를 가지고 있으면, 단말에서 MCU 로의 영상과 음성 데이터의 전송은 본 논문에서 제안한 방법을 적용할 있고, MCU 에서 각 단말로의 데이터 전송은 표준 방식을 따른다.

#### IV. Performance Estimation and Related Works

이 장에서는 본 논문에서 제안한 방법을 실제 어플리케이션에 적용하였을 경우 얻어지는 효율성과 그렇지 않을 경우를 나눠 표현하기로 하겠다.

제안된 방법을 사용하지 않고 영상과 음성을 각각의 채널을 통해 전송하였을 경우 전체 크기는

$$T_{size} = T_{vsize} + T_{asize} \text{-----} \textcircled{1}$$

①과 같으며, T\_vsize 와 T\_asize 는 각각 ②, ③과 같이 표현할 수 있다.

$$T_{vsize} = vsize + Num_v * (IP(20) + UDP(8) + RTP(12)) \text{---} \textcircled{2}$$

$$T_{asize} = asize + Num_a * (IP(20) + UDP(8) + RTP(12)) \text{----} \textcircled{3}$$

vsize : 단위 시간당 코덱에서 발생하는 영상 데이터

asize : 단위 시간당 코덱에서 발생하는 음성 데이터

T\_size : 영상과 음성을 합한 전체 크기

T\_vsize : 헤더를 포함한 전체 영상 데이터 크기

T\_asize : 헤더를 포함한 전체 음성 데이터 크기

Num\_v : 영상이 전송된 패킷의 수

Num\_a : 음성이 전송된 패킷의 수

제안된 방법을 적용하였을 경우는 T\_size 는 다음과 같이 표현할 수 있다.

$$T_{size} = T_{vhsz} + T_{ahsz} \text{-----} \textcircled{4}$$

$$T\_vhsz = vsz + Num\_vo * (IP + UDP + RTP(or cRTP) + 1) + Num\_vh * (IP + UDP + cRTP(2)+1)-----⑤$$

$$T\_ahsz = asz + (Num\_a - Num\_vh) * (IP + UDP + RTP) + Num\_vh * (2)-----⑥$$

T\_vhsz : 제안된 방법을 사용하였을 경우 영상 크기

T\_ahsz : 제안된 방법을 사용하였을 경우 음성 크기

Num\_vo : 영상만 전송되는 경우

Num\_vh : 영상과 음성이 함께 전송된 경우

Num\_v : Num\_vo + Num\_vh

cRTP : 압축된 RTP 헤더 (2 byte)

본 논문에서 제안한 방법을 실제 어플리케이션에 적용시켰을 경우 Num\_v 는 코덱의 압축률과 영상의 변화에 따라 변할 수 있는 요소이나 ②, ③식에서는 동일한 값을 갖게 된다. Num\_a 의 경우 코덱에 따라 일정한 값을 갖게 되며, 실제 전체 성능에 영향을 미치는 변수는 Num\_vh 가 된다. 여기서 Sequence Number 가 4 비트인 것을 고려하였을 때 Num\_vh 와 Num\_v 의 관계는

$$Num\_vh / Num\_v = 1/16-----⑦$$

이 되고, 제안된 방법에서 발생하는 1 바이트의 오버헤드를 고려하였을 때 역시 적용하지 않은 것에 비해 좋은 효율성을 예상할 수 있다. 그리고 음성의 경우 실제 어플리케이션에 구현되어지는 방법에 따라 달라질 수 있으나, 압축률이 높은 G.723.1[4]이나 G.729A[5]를 사용할 경우 영상쪽에서의 효율증대를 수치 상으로 기대할 수 있다. 예를 들어 음성 코덱이 G.729A 라고 했을 때, 영상과 함께 전송되어지는 음성의 비율이 전체의 1/4, 2/4, 3/4 일 경우 대략 사용하지 않았을 경우의 81%, 62%, 43%의 대역폭을 사용하게 됨으로써 나머지 부분을 다른 사용자나 혹은 추가적인 서비스 요구에 사용될 수 있다. 단말에서 음성 코덱을 G.711 과 같은 데이터의 양이 큰 코덱을 사용하였을 경우 크기가 작은 코덱을 사용하였을 때의 효과를 기대할 수는 없을 것이다. 또한 이런 경우 추가적인 복잡도가 발생할 수 있으므로 적용하지 않는 것이 바람직할 것이다. 본 논문에서 제안하는 요지는 어플리케이션에서 어느 정도 복잡도는 발생할 수 있으나 그것에 비해 제한된 네트워크의 자원을 사용하는데 있어 효율성을 재고해 볼 수 있다는 것이다.

## V. Conclusion and future works

기존의 방법과 비교해 보았을 때 대역폭의 효율을 향상시킬 수 있으므로 실제 어플리케이션으로의 적용을 생각해 볼 수 있을 것이다. 하지만 사용하는 음성 코덱이 G.711 과 같이 데이터율이 큰 경우 약간의 대역폭 향상은 기대할 수 있으나 이에 따르는 실제 어플리케이션 상에서의 복잡도 또한 고려하여야 할 것이다.

앞으로 진행 방향은 실제 어플리케이션에 적용하였을 경우 대역폭 뿐만 아니라 앞에서 언급한 QoS(Delay 에 의해 발생하는 음성이나 영상의 품질 저하)와 관련된 부분도 포함하여야 진행해야 할 것이다[3].

## 참고문헌

- [1] ITU-T Recommendation H.323, "Packet-based multimedia communication system," ITU, Nov. 2000.
- [2] Jonathan Davidson, "Voice over IP Fundamental," Cisco Press, Mar.2000
- [3] ITU-T Recommendation G.114, "One way transmission time," ITU, Feb.1996.
- [4] ITU-T Recommendation G.723.1, "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.4 kbit/s," ITU, Mar.1996.
- [5] ITU-T Recommendation G.729 Annex A, "C source code and test vectors for implementation verification of the G.729 reduced complexity 8kbit/s CS-ACELP speech coder," ITU, Nov.1996.
- [6] ITU-T Recommendation H.225, "Call signalling protocols and media stream packetization for packet-based multimedia communication systems," ITU, Nov.2000.
- [7] ITU-T Recommendation H.245, "Control Protocol for multimedia communication," ITU, Nov.2001.
- [8] H.P.Sze, Soung C.Liew, Senior Member IEEE, Jack Y.B, Danny C.S.Yip, "A Multiplexing Scheme for H.323 Voice-Over-IP Applications," in IEEE JOURNAL, Sep.2002
- [9] H.Schulzrinne, S.Casner, R.Frederick, and V.jacobson, "RTP: A transport protocol for real-time applications," in Proc. IETF RFC 1889, Jan.1996.
- [10] S.Casner and V.Jacobson, "Compressing IP/UDP/RTP headers for low speed serial links," in Proc. IETF RFC 2508, Feb.1999.