

원형 ad hoc 방식의 음성 컨퍼런스 시스템

김철한
(주) 휴림 인터랙티브

Voice Conference System Based On Circular ad hoc

Chul-Han Kim
Hurim Interactive

Abstract - 최근 컴퓨터 및 통신 기술의 발전으로 인터넷 전화 서비스가 증가하고 있다. 지금까지의 인터넷 전화 서비스는 1:1 방식의 서비스가 주류를 이루었으나, 점차 다자간 음성 채팅, 원격 강의 등과 같은 N:N 방식의 음성 컨퍼런스 서비스에 대한 수요가 급증하고 있다. 그러나 기존의 음성 컨퍼런스 시스템은 별도의 컨퍼런스 서버 구축, 부하 집중, 사용자 수의 증가에 따른 음성 트래픽 증가 등의 문제로 일정한 통화 품질을 유지하기 어렵다는 단점이 있다. 따라서, 본 논문에서는 이와 같은 문제를 개선하기 위한 ad hoc 방식의 음성 컨퍼런스 시스템을 제안하였다. 그리고 이 시스템의 성능 분석을 통하여 최소 4명의 참여자까지 일정한 음성 품질이 보장됨을 확인함으로써 본 논문에서 제안하는 시스템의 타당성을 검증하였다.

1. 서 론

최근 컴퓨터 및 통신 기술의 발전으로 다양한 멀티미디어 서비스가 등장하고 있다. 그 중 대표적인 예가 인터넷 전화 서비스이다. 인터넷 전화 서비스는 회선 교환망을 사용하는 기존의 전화 서비스와는 달리 패킷 교환망을 사용하기 때문에, 패킷의 지연·지터·손실, 안정적인 대역폭 보장의 어려움으로 기존 전화망 수준의 음질을 제공하지는 못하지만, 저렴하고 다양한 서비스를 제공하는 등 많은 장점을 가지고 있다. 뿐만 아니라, 인터넷 인프라의 발전, 데이터 압축 기술, 단말의 성능 향상 등의 이유로 인터넷 망이 가진 기술적인 한계를 극복하여 높은 수준의 QoS를 기반으로 일대일 및 컨퍼런스 형태의 음성 통신 제공이 가능하게 되었다.

지금까지의 인터넷 전화는 1:1 방식의 서비스가 주류를 이루었으나, 컴퓨터와 통신기술, VoIP 기술의 발전으로 점차 다자간 음성 채팅, 인터넷 방송, 원격교육 및 회의 등과 같은 N:N 방식의 음성 통화 서비스에 대한 수요가 급증하고 있다. 이러한 서비스에 맞추어, N:N 방식의 음성 컨퍼런스 서비스를 가능하게 할 수 있도록 하는 연구가 필요하다.

기존의 음성 컨퍼런스 시스템은 집중형과 분산형의 두 가지 유형으로 분류된다. [1] 집중형은 음성 컨퍼런스를 위해 MCU(Multipoint Control Unit)라는 별도의 서버를 두는 방식으로, MCU는 그룹별로 음성 데이터를 합성하고 분배하는 역할을 한다. 이러한 방식은 고가의 MCU 서버를 별도로 구축해야 하는 부담과, MCU로 모든 음성 데이터들이 집중됨으로 인해, MCU의 부하를 급

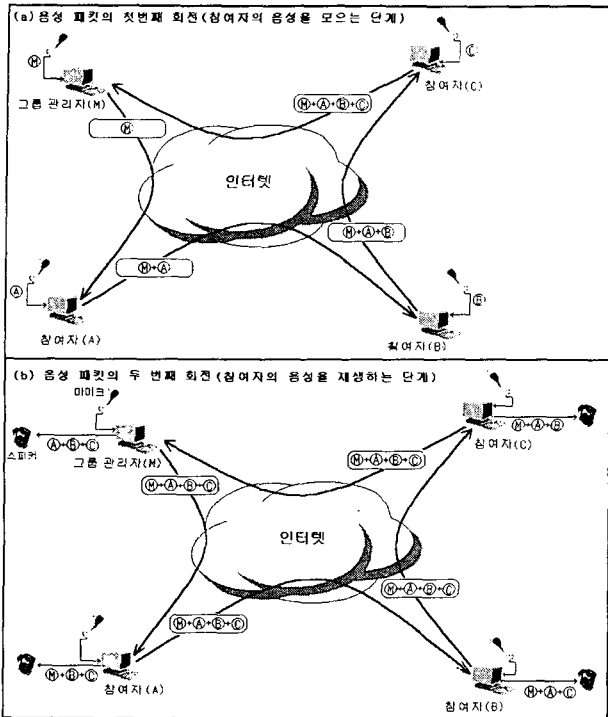
증시킴으로써 음성 품질을 균일하게 보장하지 못하는 단점이 있다. 분산형은 별도의 MCU를 필요로 하지 않으며, 각 참여자들 간에 채널을 설정하여 음성 데이터의 합성을 각 참여자가 담당하는 방식으로, 참여자의 수가 증가함에 따라 네트워크 트래픽과 각 참여자가 처리해야 할 음성 데이터의 양이 급격히 증가하여 음성 품질을 보장하기 어렵다는 단점이 있다.

따라서, 본 논문에서는 기존 멀티캐스팅 시스템이 안고 있는 문제점인 별도의 멀티캐스팅 서버 구축, 부하 집중, 네트워크 트래픽 증가 문제를 개선하기 위한 원형 ad hoc 방식의 음성 컨퍼런스 시스템을 제안하고 실제 구현한 결과의 성능 분석을 통하여 이 모델의 타당성을 검증하였다.

2. 본 론

2.1 원형 ad hoc 방식의 컨퍼런스 시스템.

그림 2.1은 원형 ad hoc 방식의 음성 컨퍼런스 시스템을 나타낸 그림이다. 이는 음성 컨퍼런스 소프트웨어가 탑재된 일반 단말과 인터넷만으로 구성된 즉, MCU와 같은 별도의 서버 없이 각 그룹의 참여자들끼리 음성을 주고 받는 분산형 시스템이다.

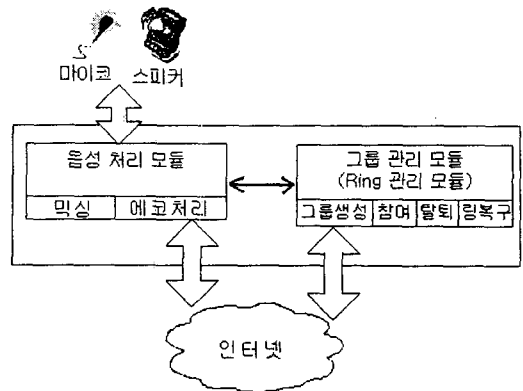


[그림 2.1 원형 ad hoc 방식의 음성 컨퍼런스 시스템]

그림 2.2는 각 단말에 설치된 음성 통신용 소프트웨어의 기능별 구성을 나타내는데, 이는 크게 음성 처리 모듈과 그룹 관리 모듈로 나누어진다.

음성 처리 모듈은 단말기에 장착된 마이크변조와 같은 입력 장치로부터 사람이 말하는 음성 신호를 디지털 신호로 변환하여 저장하고 이를 수신한 음성과 믹싱하여 원형의 경로를 따라 믹싱된 음성 데이터를 인터넷으로 전송하며, 수신한 음성 데이터에서 자신의 음성을 제거하여 스피커와 같은 출력장치로 재생해 주는 역할을 담당한다. 그룹 관리자가 녹음한 음성 패킷은 원형의 경로를 따라 두 번 회전하는데, 음성 패킷의 첫 번째 회전에서는 각 참여자들이 말한 음성을 모으고, 두 번째 회전에서는 첫 번째 회전에서 모았던 음성을 재생해 준다. 음성 패킷의 첫 번째 회전에서, 참여자들이 말한 음성 데이터는 원형의 경로를 따라 돌면서 다른 참여자들의 음성과 믹싱되어 전달되는데, 둘 이상의 음성을 믹싱한 음성 데이터의 비트수는 원래의 음성 데이터와 크기가 같아 참여자 수가 증가하더라도 네트워크 트래픽이 증가하지 않는다. 음성 패킷의 두 번째 회전에서 각 참여자가 수신 받은 음성 데이터를 재생할 때 첫 번째 회전에서 믹싱한 자신의 음성을 제거하는 에코 처리 작업을 수행한다. 또한, 묵음을 검사[2]하여 묵음인 경우 음성 처리 과정을 생략하여 음성 처리 시간을 최소화하는 방법을 사용한다.

그룹 관리 모듈은 그룹 관리와 관련된 그룹 생성, 참여, 탈퇴, 복구 기능을 가지고 있으며, 그룹 관리자가 전체적인 그룹 관리 기능을 수행한다. 그리고, 그룹 참여자들 사이에 원형의 연결을 설정 및 갱신하며, 비정상적인 상황으로 인한 원형의 회선수에 이를 복구하는 기능을 수행한다. 한편, 그룹 관리를 책임지는 그룹 관리자와 일반 참여자는 그룹 생성시에 결정된다.

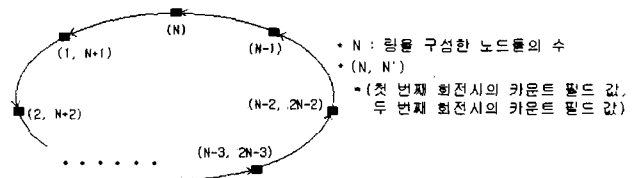


[그림 2.2 단말 소프트웨어의 기능별 구성]

2.1.1 음성 처리

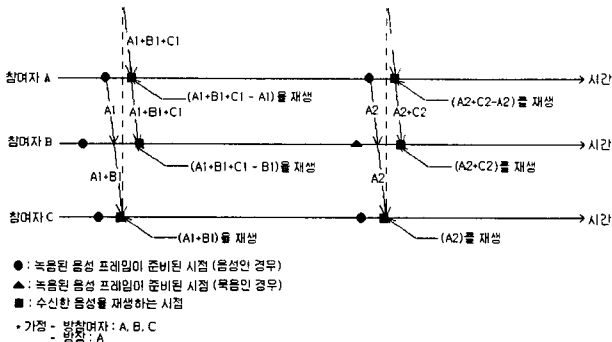
그림 2.3은 음성처리 모듈의 동작을 보여주고 있다. 그룹 참여자를 N명으로 가정하면, 음성 처리 모듈은 수신 모듈을 통해 다른 참여자의 음성 데이터를 전달 받아, 해당 음성 데이터의 카운트 필드 값을 조사한다. 카운트 필드 값은 그룹 관리자에서 생성된 음성으로부터 시작하여 그 값이 $2N-2$ 가 될 때까지 원형을 돌면서 참여자를 지날 때마다 1씩 증가한다. 음성 처리 모듈은 패킷의 카운트 필드 값에 따라 수신한 음성 패킷을 다르게 처리한다. 첫 번째 회전에서는(즉, 카운트 필드 값이 $1 \sim N-2$ 일 경우에는), 수신된 음성 프레임에 자신의 녹음된 음성을 믹싱하여 다음 참여자에게 전달하고, 마지막($N-1$)일 경우에는 수신된 음성 프레임에 자신의 음성을 믹싱하여 다음 참여자에게 전달한다. 그리고, 동시에 자신의 음성이 추가되지 않은 수신된 음성 프레임을 재생한다. 두 번째 회전에서는(즉, 카운트 필드 값이 $N \sim 2N-3$ 일 경우에는), 수신된 음성 프레임을 다음 참여자에게 전달하면서, 수신 프레임에서 자신의 음성 프레임을 제거한 후 재생한다. 두 번째 회전의 마지막($2N-2$)일 경우에는 음성 프레임의 전달 없이 수신한 음성 프레임에서 자신의 음성만을 제거하고 재생한다. 이 방법은 N명의 참여자가 있는 음성 통신 그룹에서 모든 참여자들의 믹싱된 음성 프레임이 N번째가 아닌 $N-1$ 번째에서 형성된다는 점을 인식하여 고안된 것으로, 음성 패킷의 회전 시간을 줄이는 효과가 있다.

수신 패킷의 카운트 필드 값	각 단말 노드에서의 액션
믹싱 loop	$1 \sim (N-2)$: 수신 데이터와 녹음 데이터를 믹싱하여 전달
	$N-1$: 수신 데이터와 녹음 데이터를 믹싱하여 전달
재생 loop	$N \sim (2N-3)$: 수신 데이터를 그대로 전달
	$2N-2$: 수신 데이터에서 자신의 음성을 제거하고 재생



[그림 2.3 음성 처리 동작 메커니즘]

그림 2.4는 위의 방식을 적용하였을 때 각 참여자들이 음성을 처리하는 녹음시점과 재생시점을 시간축 상에 표현한 것으로 이는 각 참여자들마다 일정 시간 간격으로 음성 프레임을 재생하고 있음을 보여준다.



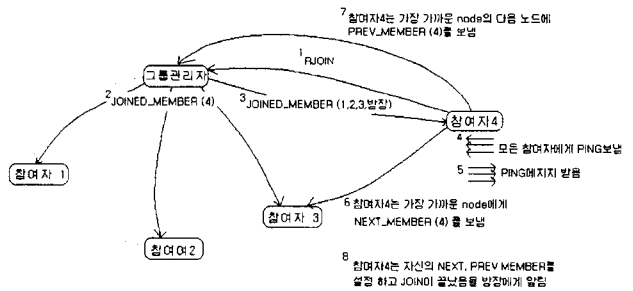
[그림 2.4 음성 프레임의 녹음 및 재생 시점]

2.1.2 그룹 관리

그룹 관리는 기본적으로 그룹 생성, 참여, 탈퇴, 복구의 기능을 가지고 있으며, 그룹의 생성, 참여, 탈퇴시 각 참여자들을 가상적인 원형으로 구성하는 역할을 한다.

그룹 생성은 특정한 그룹을 생성하는 기능을 수행한다. 이때 그룹 생성자는 그룹 관리자가 되어 다른 사용자들의 그룹 참여를 기다리고, 참여자들의 정보나 그룹 정보 전체를 관리한다.

그룹 참여는 일반 사용자가 특정 그룹에 참여하는데 필요한 작업을 수행한다. 그림 2.5는 그룹 참여과정을 보여주고 있다. 그룹에 참여하려는 새로운 참여자는 기존의 참여자들에게 핑(ping) 메시지를 전송하여 가장 가까운 거리의 참여자를 검출한 후, 가장 가까운 참여자와 그 다음 참여자 사이에 들어가 그룹 참여 과정을 수행한다. 이 방법은 그룹 참여시 최단 거리의 원형을 구성하여 음성 패킷의 지연을 최소화 한다.



[그림 2.5 그룹 참여 과정]

그룹 탈퇴는 한 참여자가 특정 그룹에서 탈퇴하는데 필요한 작업을 수행한다. 그룹에서 탈퇴하고자 하는 참여자가 그룹 관리자에게 그룹 탈퇴 메시지를 보내면, 그룹 관리자는 나머지 참여자에게 탈퇴할 참여자의 정보를 알려주고, 탈퇴할 참여자의 이전 참여자와 다음 참여자를 연결시켜 주어 그룹 탈퇴 작업을 수행한다.

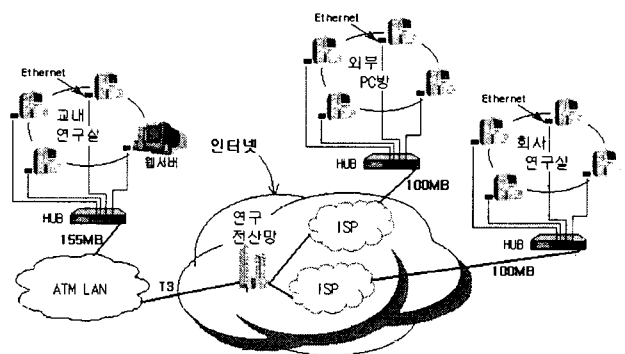
그룹 복구는 한 참여자의 단말이 비정상 종료되어 원형을 회손시키는 경우 종료된 참여자를 제외한 나머지 참여자들로 구성된 그룹으로 원형을 복구하는 기능을 수행한다. 원형 복구는 비정상 종료된 참여자 검출, 비정상 종료된 참여자 탈퇴의 2단계로 이루어지는데, 각 참여자가 매 시간 마다 음성 패킷을 제대로 수신하였는지 검사하여 일정시간 동안 음성 패킷을 수신하지 못하면 자신의 이전 참여자가 비정상 종료되었는지 알아보기 위해 이전 참여자에게 IS_ALIVE라는 메시지를 보낸다. 이 메시지에 대한 응답인 IM_ALIVE라는 메시지를

time-out이 될 때까지 수신하지 못하면, 자신의 이전 참여자가 비정상 종료되었다고 판단하여 그 정보를 그룹 관리자에게 알려 비정상 종료된 참여자를 검출한다. 다음으로, 그룹 관리자는 비정상 종료된 참여자를 탈퇴시키는 작업을 수행하는데 그 과정은 그룹 관리자가 한 참여자를 탈퇴시키는 과정과 동일하다.

2.2 성능 분석

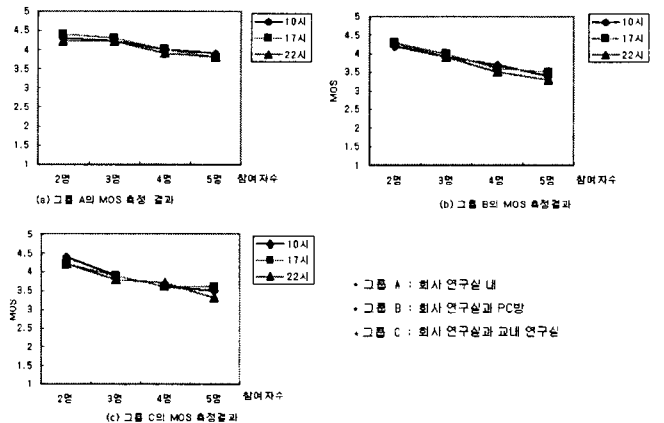
본 논문에서는 원형 ad hoc 방식의 음성 컨퍼런스 시스템의 성능을 측정하기 위해 시험 망을 구축하고 동작 시험 및 성능 분석을 하였다.

그림 2.6은 본 논문에서 구현한 시스템의 시험 망을 보여주고 있다. 시험 망은 특정 ISP 업체로부터 전용선을 임대 받아 사용하는 회사 연구실과 PC방과 ATM을 백본으로 하는 교내 연구실로 구성하였다.



[그림 2.6 시험 망 구성도]

본 논문에서 구현한 원형 ad hoc 방식의 음성 컨퍼런스 시스템의 성능을 측정하기 위해 테스트 그룹을 셋으로 나누어 그룹 참여자와 통화 품질에 대한 MOS(Mean Opinion Score)를 측정하였다. 그림 2.7은 MOS 측정결과를 그래프로 보여주고 있다.



[그림 2.7 MOS 측정 결과 그래프]

MOS 측정은 인터넷 사용의 Busy Hour인 오전 10시와 오후 5시 그리고 오후 10시로 나누어 실시하였다. 그림 2.7(a)를 살펴보면, 그룹 A의 MOS 측정 결과를 보여주고 있다. 그룹 A의 경우 동일 LAN 내에서 그룹을 구성함으로써, 전달 지연이 거의 발생하지 않아 참여자의 증가에 따른 MOS 값이 민감하게 반응하지 않음을 알 수 있다. 그림 2.7(b)의 그룹 B나 그림 2.7(c)의 그룹 C에서는 서로 다른 네트워크 간에 그룹을 구성함

으로써, 전달 지연이 거의 발생하지 않아 참여자의 증가에 따른 MOS 값이 민감하게 반응하지 않음을 알 수 있다. 그림 2.7(b)의 그룹 B나 그림 2.7(C)의 그룹 C에 서는 서로 다른 네트워크 간에 그룹을 구성하므로, 참여자의 증가로 인한 전달 지연이 상대적으로 커져 MOS 값이 그룹 A에 비해서는 민감하게 변하였으나, 측정 결과 대부분의 MOS 값이 3.5 이상으로 우수한 통화 품질을 나타내었다.

3. 결 론

지금까지의 인터넷 전화는 1:1 방식의 서비스가 주류를 이었으나, 컴퓨터와 통신기술, VOIP 기술의 발전으로 점차 다자간 음성 채팅, 인터넷 방송, 원격교육, 원격 회의 등과 같은 N:N 방식의 음성 컨퍼런스 서비스에 대한 수요가 급증하고 있다.

음성 멀티캐스팅 서비스를 제공하기 위해서는 별도로 고가의 MCU를 구축해야 하는데, 이 경우 서버로 부하가 집중되어 장비의 성능 저하가 발생하며, 트래픽의 폭주로 인해 일정한 통화 품질을 유지하기가 어려운 문제점이 있다.

따라서, 본 논문에서는 원형 ad hoc 방식을 이용한 음성 컨퍼런스 시스템을 제안하였다. 이는 음성 컨퍼런스를 위한 별도의 서버를 사용하지 않고, 각 그룹별로 단말들을 가상의 원형으로 구성하여 원형의 경로를 따라 음성 데이터를 전달하는 시스템이다.

이 시스템은 음성 데이터를 회전시킬 때 참여자들의 음성을 하여 참여자 수가 증가하여도 각 참여자가 처리해야 할 데이터의 양을 일정하게 유지하고 네트워크 부하를 최소화하여 일정한 수준의 음성 품질을 보장할 수 있는 장점이 있다.

아울러 본 논문에서는 원형 ad hoc 방식을 이용한 음성 컨퍼런스 시스템을 구현하여, 다양한 네트워크 상에서 동작시험 및 성능 분석을 하였다. 시험결과 시간대에 관계없이 최소 4명까지는 3.5 이상의 MOS 값을 나타내었고, non Busy Hour일 때는 5명 까지도 일정한 음성 품질이 보장됨을 확인하였다. 이 시스템은 별도의 전용 하드웨어 시스템 없이 각 단말에 소프트웨어만 탑재하면 양방향 음성 컨퍼런스 서비스를 제공할 수 있기 때문에 향후 다양한 형태의 음성 컨퍼런스에 활용이 가능하다.

(참 고 문 헌)

- [1] ITU-T Recommendation H.323, "Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service", Nov. 1996
- [2] 황인환, "인터넷 전화에서 통화 품질 향상을 위한 목음 처리 기법", 정보과학회 2000 가을학술 발표 논문집(III), pp222-224, V.27, Oct. 2000