The Flow Control of Audio Data Using Distributed Terminal Mixing in Multi-point Communication

Young-Mi Kim and Dae-Joon Hwang

School of Information and Communication Engineering, Sungkyunkwan University, 300 Chunchun-Dong, Jangan-Ku, Suwon-city, South-Korea, 440-746
ultraym@skku.edu, djhwang@skku.ac.kr

Abstract. This paper describes an efficient audio flow control method in the point of quantitative performance using audio-mixing, compared to existing P2P(Peer To Peer) method. In comparison with existing P2P method, using central mixing and distributed terminal mixing method, we achieved advance at the point of global network usage and each terminal's CPU load, and additionally we expect more session, more terminal can be served by same amount of network bandwidth and computers. By using P2P method in audio communication, speaker and listener must connect to each other. So it has the critical defect that as the participants grows more and more, the network bandwidth usage, each terminal's CPU load will grows rapidly. So the number of participants in same session will be extremely restricted. In comparison with P2P method, the central mixing method has the great advantage at the points of network usage and terminals CPU load. Regardless of the number of speakers and listeners, all the participants can speak and listen with all other participants by using just one stream's amount of data size and CPU load. But all the network usages and CPU loads of "Audio decompression->Buffering->Mixing->Audio Compression" are concentrated on central server. So the number of sessions and terminals can be participated in one server will be highly restricted. This study solves the problems of server's CPU load and network load by using the distributed terminal mixing method.

1 Introduction

As Internet technologies have dramatically improved and Internet services have been widely expanded and adopted in recent years, the demand for multimedia data services is much higher than ever before in our daily lives. This is remarkably true with real-time audio/video communication services, which have been deployed very much in quantity and require ensuring accurate and real-time transfer based on data loss restoration, however, considerable data volume led to too much load on networks.

In addition, the rapid advancement of computer and communication network technologies contributed to make the speed of processor as well as network much faster. However, in the other hand, there appeared more Internet-users, protocols and programs requiring considerable throughput. So there has always been and will continue to be some kind of effort to minimize required computer processing capacity and network usage in order to save costs.

H. Legind Larsen et al. (Eds.): FQAS 2006, LNAI 4027, pp. 671 – 680, 2006.

As we do not have any specific organization responsible for controlling network resources, each network program has its own way of leveraging allocated bandwidth and servicing users. Even if some two programs' functionalities are identical, their performance and/or network bandwidth consumption may be different from each other, depending on transfer method. Especially in multi-point multimedia network programs, there are a lot of contributing factors to this kind of difference. These factors can be grouped into multimedia data compression and multimedia data transfer. Multimedia data compression has been evolved consistently. Standard organizations continue to provide CODECs with better video output and compression rate, which are being deployed to a wide variety of areas, including multimedia-enabled programs, consumer electronics, etc. The development of multimedia transfer technology which helps save network resources doesn't seem so fast as that of CODEC, however. The reason is in most cases, P2P-based approach has been deployed to multi-point area, which doesn't make full use of the opportunities for saving bandwidth. This aspect would lead to a condition where there are only several multi-point communication sessions occupying the entire network.

There is a protocol for real-time multi-point stream, called IP-multicasting.[1][6] With that protocol, we can deliver the stream to desired targets using just one transmission, but because there are not enough routers supporting IP-multicasting, in practice, we must use also another protocol with it to complete all transmission. And IP-multicasting has no gain on CPU load. And one of important defect is that it can't guarantee the delivery of data we sent. It's very difficult work to restore the lost packets to get a continuous stream in real-time in especially multi-point communication.[2]

In addition, CODEC evolvement required more CPU load. In a multi-point multimedia communication, you should transfer data to and/or receive & process data from more than one person. To process more than one stream in a single computer, it should be assumed that the computer has enough processing capacity. We cannot always make sure any user's computer would meet the requirement, however.[3][4]

Although it's not the main topic of this study, we should consider the sequence from the capturing audio signal to playing audio "analog audio signal->conversion to digital signal -> compression -> sending network packets -> receiving network packets -> decompression -> conversion from digital to analog signal -> playing with audio output device". And the compression and decompression occupy the most of load on CPU.

In this paper, we're going to focus on improvement on these issues below.

- Global network usage
- Server / terminal network usage
- Server / terminal CPU usage
- The number of sessions and terminals that can be accommodated on the same number of server and network bandwidth.

We suggested the central mixing method for reduction of global network usage and terminal's CPU load. This is the method that all the audio data must be collected in central mixing server first, and the mixing server mixes the data and sends the mixed data to each participant terminal.