Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Communications 28 (2005) 125–135
computer communications
www.elsevier.com/locate/comcom

Adaptive coding schemes for support of multicast audio to wireless devices

Leann Christianson*, Kevin Brown

Department of Math and Computer Science, California State University, Hayward, CA 94542, USA

Received 29 January 2004; revised 2 July 2004; accepted 27 July 2004
Available online 7 September 2004

Abstract

The mobile wireless environment, characterized by dynamically varying bandwidth and signal quality, provides a harsh environment for audio and video applications. Insufficient bandwidth leads to losses and delays which significantly reduce the quality of playback at the receiver. Such an environment demands application support for bandwidth adaptation. Audio applications in particular have traditionally operated at fixed rates and are prime candidates for enhancement. We propose a DCT-based audio decomposition (DAD) scheme for real-time general audio streams in the multicast environment which allows audio quality to be traded off against reductions in bandwidth usage. We compare the performance of DAD, using objective and subjective measures, against two other schemes which also support adaptation to restricted bandwidth.

Keywords: Coding; Adaptive; Wireless; Multicast

1. Introduction

The mobile wireless environment, characterized by dynamically varying bandwidth and signal quality, provides a harsh environment for real-time applications. The wireless networks in use today typically do not offer guarantees on minimum bandwidth or maximum delay. As such, the bandwidth provided to a connection will vary over time as traffic levels rise and fall or as users encounter obstacles or signal interference.

Real-time applications such as audio and video conferencing and shared application control, however, have stringent requirements regarding maximum delay and minimum bandwidth. A reduction in available bandwidth will result in dropouts in audio streams, loss of video frames, possible loss of synchronization between streams, and difficulties in shared control of applications as timing requirements may be exceeded.

In this paper, we concentrate on the particular problem of supporting audio applications operating in a wireless multicast environment. Applications common to this environment include CD-quality digital radio broadcasts, multicasts of live concerts, and delivery of audio tracks to accompany video-on-demand services. Multicast may be used in these situations to efficiently deliver the same audio stream to a number of receivers which can effectively scale from a few receivers to many thousands.

Audio applications pose a specific challenge as they typically produce data streams at *fixed* data rates (e.g. 1.4 Mbps for uncompressed stereo CD-quality audio or 192 Kbps for stereo MP3 streams). When available bandwidth drops below the level needed to transport the stream, packet losses, bit errors, and delays that degrade the quality of the received signal may occur. Lost or delayed packets and bit errors cause dropouts and static which lead to an audio stream which sounds choppy and may be unintelligible to the user. The human ear has been shown to be very sensitive to discontinuities in speech and other types of general audio signals.

* Corresponding author.
E-mail addresses: leann@csuhayward.edu (L. Christianson), kbrown@csuhayward.edu (K. Brown).

0140-3664/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2004.07.012

A possible solution to this lack of bandwidth is to reduce the quality of the audio stream so that continuous playback is possible within the restrictions of the available bandwidth. A lower quality audio stream will require less bandwidth for transmission. It is not reasonable, however, to simply deliver the lowest quality (and hence lowest bandwidth) audio stream at all times in an attempt to reduce packet losses and delays. Application users should expect to receive high quality audio when bandwidth is available and lower quality audio when bandwidth is restricted. Hence, the application must *adapt* its output to the available bandwidth with *continuity* of the transmission being the primary goal. When considering the multicast environment, however, a second issue arises, that of defining 'available bandwidth'. Some receivers may be connected via high-speed wireless LANs (e.g. 11 Mbps) and others via low-speed CDPD (19 Kbps). Hence, a high data rate would overwhelm the CDPD users, but a low data rate would not take advantage of the bandwidth available to the wireless LAN users. As such, adaptation must be done on a *per-receiver* basis.

A known strategy to solve the per-receiver adaptation problem is to encode and transmit signals at a selection of data rates and allow the receiver to choose the rate which 'best' fits her available bandwidth. Choosing a stream which is transmitted at a lower rate would imply a lower quality stream, where the loss of quality is dependent upon the method used to reduce the data rate. A stream whose data rate matches the currently available bandwidth, however, would be much less likely to suffer dropped packets and delays. The choice here is between a reduction in quality which can be specified and controlled by the receiver and a reduction in quality which is randomly effected by the network as it drops and delays packets which cannot be delivered due to bandwidth restrictions. Clearly, controlled reduction is preferable as the lowest priority information may be eliminated rather than a random selection of packets.

Fig. 1 shows the difference between these mechanisms for a mono CD-quality audio stream. The top signal is the original audio signal, which was sampled at 44,100 Hz using 16 bit samples. It requires 705.6 Kbps of allocated bandwidth for transmission without loss. The middle and bottom signals both require only half that amount (352 Kbps), but the reduction in bandwidth is achieved in two different ways. The bottom signal is the result of the original signal suffering a 50% loss of data packets during transmission. The lost packets cause large gaps, or silence periods, in the signal. The result is poor quality music that is choppy and often annoying or unintelligible to the user. The middle signal is result of recoding the original signal and lowering the signal resolution by discarding the least important 50% of signal information. Although the middle and bottom signals require the same amount of bandwidth, the effect of the restricted bandwidth on the two signals is very different. We will show in Section 5 that listeners prefer the lower resolution sample to the higher resolution sample with loss. The key, again, is that the reduction in bandwidth is controllable rather than random and playback remains continuous.

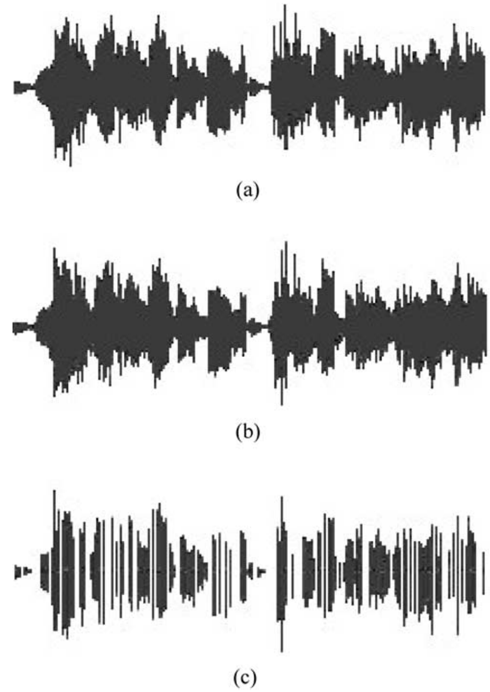

Fig. 1. (a) Original audio signal. (b) Audio signal transmitted with 50% bandwidth reduction via DAD mechanism. (c) Audio signal transmitted with 50% bandwidth reduction due to random packet loss.

The encoding scheme that we propose in this paper is called DCT-based audio decomposition (DAD). It provides support for bandwidth adaptation in the multicast environment, but can easily be modified to work in the unicast environment as well. Here, we focus on the multicast approach. Audio samples are transformed using the discrete cosine transform (DCT), and then separated into groups based on their importance to the signal. A base group provides minimal audio resolution, and each additional group increases the resolution seen at the receiver (at the cost of increased bandwidth usage). Groups of DCT coefficients are transmitted as separate multicast streams, and the receiver chooses to subscribe to as many groups as she can feasibly receive. As available bandwidth may change over time, the receiver may periodically decide whether to subscribe to an additional group (to increase quality) or cancel a subscription to an existing group (to decrease bandwidth usage).

The rest of the paper is organized as follows. Section 2 contains background and related work, and Section 3 presents our design. Section 4 presents objective results and Section 5 presents qualitative results from listening tests. Section 6 describes the implementation of the DAD multicast tool and Section 7 concludes the paper.