Proceedings Volume 3

IEEE INFOCOM'96

The Conference on Computer Communications

Fifteenth Annual Joint Conference of the IEEE Computer and Communications Societies

Networking the Next Generation

March 24-28, 1996
San Francisco, California

Sponsored by
IEEE Computer Society
IEEE Communications Society

IEEE Computer Society Press
Los Alamitos, California

Washington • Brussels • Tokyo

IEEE Computer Society Press
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1264

IEEE Computer Society Press Order Number PR07292 (set of 3 volumes)
IEEE Order Plan Catalog Number 96CB35887
ISBN 0-8186-7292-7 (paper)
IEEE Order Plan ISBN 0-8186-7293-5
Microfiche ISBN 0-8186-7294-3
ISSN 0743-166X

Additional copies may be ordered from:

IEEE Computer Society Press
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1264
Tel: +1-714-821-8380
Fax: +1-714-821-4641
Email: cs.books@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: +1-908-981-1393
Fax: +1-908-981-9667
misc.custserv@computer.org

IEEE Computer Society
13, Avenue de l'Aquilon
B-1200 Brussels
BELGIUM
Tel: +32-2-770-2198
Fax: +32-2-770-8505
euro.ofc@computr.org

IEEE Computer Society
Ooshima Building
2-19-1 Minami-Aoyama
Minato-ku, Tokyo 107
JAPAN
Tel: +81-3-3408-3118
Fax: +81-3-3408-3553
tokyo.ofc@computer.org

Editorial production by Penny Storms
Cover by Joseph Dagle/Studio Productions
Printed in the United States of America by KNI, Inc.

The Institute of Electrical and Electronics Engineers, Inc.

Control Mechanisms for Packet Audio in the Internet

Jean-Chrysostome Bolot Andrés Vega-García
INRIA
B. P. 93
06902 Sophia-Antipolis Cedex
France
{bolot, avega}@sophia.inria.fr

Abstract

The current Internet provides a single class best effort service. From an application's point of view, this service amounts in practice to providing channels with time-varying characteristics such as delay and loss distributions. One way to support real time applications such as interactive audio given this service is to use control mechanisms that adapt the audio coding and decoding processes based on the characteristics of the channels, the goal begin to maximize the quality of the audio delivered to the destinations. In this paper, we describe and analyze a set of such control mechanisms. They include a jitter control mechanism and a combined error and rate control mechanism.

These mechanisms have been implemented and evaluated over the Internet and the MBone. Experiments indicate that they make it possible to establish and maintain reasonable quality audioconferences even across fairly congested connections.

1 Introduction

The transmission of voice over packet switched networks was an active research area in the late 70's and the early 80's [29]. Much of the work then focused on using packet switching for both voice and data in a single network. Packet voice, and more generally packet audio applications, have recently become again of interest. This interest has been fueled by the availability of supporting hardware (microphones now come standard with most workstations), of increased bandwidth throughout the Internet, and by the development of the MBone [7]. A variety of audio tools such as *vat* [17] or *Nevot* [24] have been available for a few years, and they have been used to audiocast conferences. Recently, several more tools have been announced, which claim to provide toll-quality workstation or PC audio over the Internet for a fraction of the cost of a telephone call (see [5] for pointers to these tools and other information related to packet audio).

However, the Internet provides a simple single class best effort service. From a connection's point of view, the best effort service amounts in practice to offering a channel with time-varying characteristics such as delay and loss distributions [2, 21]. These characteristics are not known in advance since they depend on the (*apriori* unknown) behavior of other connections throughout the network. This makes it essentially impossible to provide performance guarantees such as minimum loss rate or maximum delay. Thus, it is not clear how well applications with minimum guaranteed requirements such as audio applications can work over the Internet. Experimental evidence suggests that, although the quality of the audio delivered by Internet tools has improved, audio quality is still mediocre in many audio conferences. This is clearly a concern since audio quality has been found to be more important than video quality or audio/video synchronization to successfully carry out collaborative work [15].

It should be pointed out that bad audio quality is often caused by problems having little to do with either the network service or the audio tools themselves. The experience accumulated with the audiocasting of MICE [20] and IETF meetings suggests that badly tuned or set up microphones and speakers are responsible for many such problems. However, all these can be addressed by users at their own sites. Furthermore, their impact is expected to decrease as users become familiar with the tools and the tools themselves become more user friendly. In any case, the most persistent problems with audio quality are caused by the network, or rather by the impact of traffic in the network on the stream of audio packets. Two approaches have emerged to tackle this problem.

One approach is to extend current protocols and switch scheduling disciplines to provide the desired requirements. This approach requires that admission control, reservation, and/or sophisticated scheduling mechanisms be implemented in the network. These mechanisms are not yet implemented in the Internet, and their design, analysis, and evaluation is still an active research area [26]. Thus, we have not pursued this approach so far.

Another approach is to adapt applications to the

0743-166X/96 $5.00 © 1996 IEEE

service provided by the network. This amounts in practice to adapting applications to the time-varying characteristics of the connection over which the application data packets are sent, the goal being to maximize the quality of the data delivered to the destinations. Experimental evidence suggests that the quality of the audio depends essentially on the number of lost packets and on the delay variations between successive packets. Thus, the most important network characteristics for audio applications are the delay variance (or jitter), and the loss distributions. Furthermore, for live audio applications such as audioconferences, the average end-to-end delay must be small to allow interactions between participants.

The goal then in this approach is to develop mechanisms that attempt to eliminate or at least minimize the impact of packet loss and delay jitter on the quality of the audio delivered to the destinations. We have developed a set of such mechanisms. One mechanism adjusts the playout time of audio packets at the destination, the objective being to minimize the impact of delay jitter. A second mechanism adds redundancy information in the audio packets sent by the source, the objective being to minimize the impact of packet loss. A third mechanism controls the rate at which packets are sent over a connection, the objective being to match the send rate to the capacity of the connection and hence to minimize packet loss. The second and third mechanisms both attempt to minimize the impact of packet loss, and they really are two sides of a joint error/rate control mechanism.

These mechanisms have been implemented in a new audio tool developed at INRIA. For lack of space (and as suggested by reviewers) we do not describe in the paper the jitter control mechanism. We focus instead on the rate and error control mechanisms. In Section 2, we describe the structure of the audio tool. In Section 3, we characterize the loss process of audio packets, and describe and evaluate a packet loss recovery scheme. In Section 4, we describe and evaluate a joint error and rate control scheme. Section 5 concludes the paper.

2 The audio tool

The structure of the audio tool is shown in Figure 1 below. It is being developed within the MICE project in collaboration with a group at University College London (UCL). Work at UCL has focused on device-independent audio input, efficient mechanisms for silence detection, automatic gain control, and echo cancellation, and on the evaluation of the auditory quality of the signal delivered to the destinations. Work at INRIA has focused on coding schemes, and on jitter, rate, and error control mechanisms.

The coding schemes available at this time use 8-kHz sampled speech with bit rates varying from a few kb/s to 64 kb/s. Specifically, they include a 64-kb/s μ-law PCM, various adaptive delta modulation (ADM) coders with rates varying from 16 kb/s (for ADM2) to 56 kb/s (for ADM6), a 13 kb/s GSM coder, and a 4.8 kb/s LPC low bit rate coder. Work is underway to include wideband speech coders. The PCM, ADM6, ADM5, and GSM coders deliver high quality audio with MOS scores above 3.5. The ADM2, ADM3, and LPC coders delivers audio with a somewhat lower quality. However, even a mediocre low bit rate coder turns out to be useful for error control purposes (refer to Section 3). The boxes in the figure which involve one of the control mechanisms of interest in the paper have been highlighted. They include the redundancy box (which involves the error control mechanism), the congestion information and feedback information boxes (which involve the error/rate control mechanism), and the playout buffer box (which involves the jitter control mechanism).

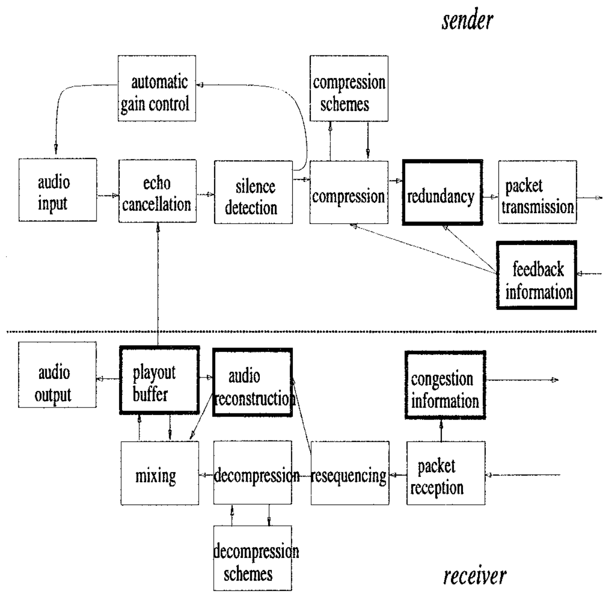

Figure 1: Structure of the audio tool

The audio packets are sent from the source to the destination(s) using IP (or its multicast extension), UDP, and RTP. To each audio packet is associated a timestamp and a sequence number. The timestamp is used to measure end-to-end delays, and the sequence number is used to detect packet losses.

3 A loss recovery mechanism

Anecdotal evidence suggests that audio quality is still mediocre in many audio connections because of packet losses. This makes it important to implement an efficient loss recovery mechanism for audio applications. We address this problem in this section. Our main result is that open loop error control mechanisms based on forward error correction are adequate to reconstruct most lost audio packets. We describe one such mechanism and report on improvements of audio quality obtained with it.

Analysis of the loss process

Many different quantities can be used to characterize the loss process of audio packets. The obvious measure is the average loss, or unconditional loss probability. Let l_n denote a boolean variable which is set to 1 if packet n is lost, and 0 otherwise. The average loss is thus equal to the expected value of l_n. We denote $ulp = E[l_n]$. However, ulp does not characterize the burstiness of the loss process, or equivalently the correlation between successive packet losses. One way to capture such correlation is to consider the conditional probability that a packet is lost given that the previous packet was lost. We denote $clp = P[l_{n+1} = 1 | l_n = 1]$.

We have analyzed clp and ulp in the Internet using measurements and analysis. The measurements have all be done using the PCM coder with 320-byte packets (or 40 ms of speech) between INRIA Sophia Antipolis and University College London (UCL) in the UK. Figure 2 shows the evolutions of the number of consecutively lost packets as a function of n measured at 3:00 pm. The average loss $ulp = 0.21$ is quite high

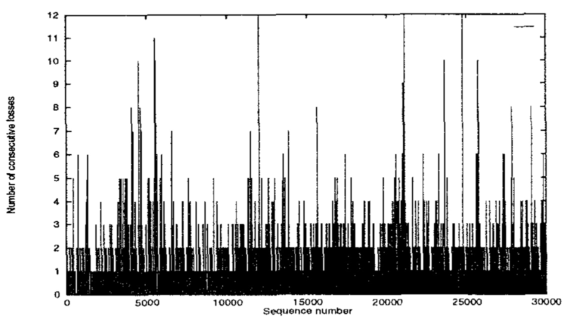

Figure 2: Evolutions of the number of consecutively lost packets

because the INRIA-UCL connection is heavily loaded during daytime. However, it appears that most loss periods involve one or two packets. This observation is confirmed by looking at the frequency distribution in Figure 3, which shows the frequency distribution of the number of consecutive losses (i.e. the number of occurrences of n consecutive losses for different n) corresponding to the trace in Figure 2. The slope of the distribution decreases linearly near the origin.

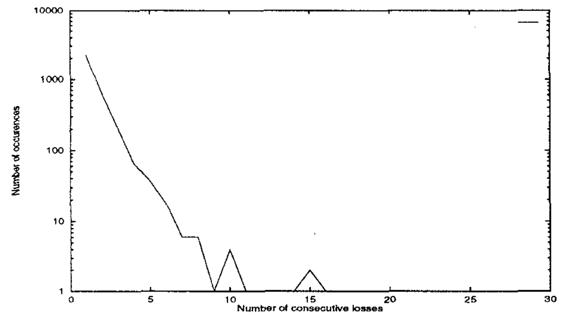

Figure 3: Frequency distribution of the number of consecutively lost packets

Since the figure is drawn on a log scale, this indicates that the probability distribution decreases geometrically fast away from the origin.

We have examined the loss process of audio packets over unicast connections other than the INRIA-UCL connection, and over multicast connections as well. In all cases, we have found that the frequency distribution of the number of consecutively lost packets is similar to that described above [4].

Packet loss recovery schemes

A loss recovery scheme is required if the number of lost audio packets is higher than that tolerated by the listener at the destination. Loss recovery is typically achieved in one of two ways. With closed loop mechanisms such as Automatic Repeat Request (ARQ) mechanisms, packets not received at the destination are retransmitted. With open loop mechanisms such as Forward Error Correction (FEC) mechanisms, redundant information is transmitted along with the original information so that (at least some of) the lost original data can be recovered from the redundant information.

ARQ mechanisms are not generally acceptable for live audio applications because they increase end to end latency. Furthermore, they do not scale well to large multicast environments. FEC is an attractive alternative to ARQ for providing reliability without increasing latency [1]. However, the potential of FEC mechanisms to recover from losses depends crucially on the characteristics of the packet loss process in the network. FEC mechanisms are more effective when lost packets are dispersed throughout the stream of packets sent from a source to a destination. Thus, our measurements above indicate that FEC is particularly well suited for audio applications over the Internet.

Many FEC mechanisms proposed in the literature involve exclusive-OR operations, the idea being to send every nth packet a redundant packet obtained by exclusive-ORing the other n packets [25]. This mech-

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | sales@docketalarm.com | 1-866-77-FASTCASE

fastcase® Smarter legal research.