On Retransmission-Based Error Control for Continuous Media Traffic in Packet-Switching Networks

Bert J. Dempsey *Jörg Liebeherr* *Alfred C. Weaver*

Computer Science Department
University of Virginia
Charlottesville, VA 22903
U. S. A.

Abstract

Distribution of continuous media traffic such as digital audio and video over packet-switching networks has become increasingly feasible due to a number of technology trends leading to powerful desktop computers and high-speed integrated services networks. Protocols supporting the transmission of continuous media are already available. In these protocols, transmission errors due to packet loss are generally not recovered. Instead existing protocol designs focus on preventive error control techniques that reduce the impact of losses by adding redundancy, e.g., forward error correction, or by preventing loss of important data, e.g., channel coding. The goal of this study is to show that retransmission of continuous media data often is, contrary to conventional wisdom, a viable option in most packet-switching networks. If timely retransmission can be performed with a high probability of success, a retransmission-based approach to error control is attractive because it imposes little overhead on network resources and can be used in conjunction with preventive error control schemes. Since interactive voice has the most stringent delay and error requirements, the study focuses on retransmission in packet voice protocols. An end-to-end model of packet voice transmission is presented and used to investigate the feasibility of retransmission for a wide range of network scenarios. The analytical findings are compared with measurements of packet voice transmission over a campus backbone network.

Key Words: Error Control, ARQ, Continuous Media, Packet-Switching, Retransmission, Packet Voice.

1 Introduction

Continuous media services such as voice and video were traditionally carried over circuit-switching networks. Using packet-switching networks for these services has become increasingly attractive due to technology trends that have enabled high-speed multiservice networks. Since voice and video are inherently variable rate sources [4, 12], statistical multiplexing gains in packet-switching networks result in a more efficient use of network resources than circuit switching. However, the distribution of continuous media is fundamentally different from traditional reliable data transfer, such as file transfer and remote login, since continuous media is sensitive to end-to-end delays and variations of the delays.

The distribution of continuous media across a packet-switching network requires consideration of encoding schemes, end-to-end network delays, network delay variations, and packet loss, all of which significantly affect the playback quality at the receiving site:

- In recent years, considerable progress has been made in the design of efficient digital encoding techniques for analog audiovisual data [11]. The selection of an encoding scheme represents a trade-off between consumption of bandwidth in the network and playback quality at the receiving site since low-bit-rate encoding schemes result in a less precise reconstruction of the original analog signal.

- In an interactive continuous media session, human perception factors produce a requirement for bounded roundtrip delays. If roundtrip delays are too long, the interactive nature of the session is degraded.

- Statistical multiplexing introduces variations in the network delay experienced by individual packets. These variations are referred to as *delay jitter*. Delay jitter can lead to interruptions in the continuous playback of the continuous media stream at the receiver.

- Unlike data transmission, most continuous media data does not require reliable delivery, though its tolerance for packet loss is low. Techniques for robust signal processing in the presence of packet loss can significantly improve loss tolerances, but even the loss of a single packet may noticeably degrade playback quality at the receiver.

In this paper we study error control for voice transmission in packet-switching networks since interactive packet voice has very stringent delay and error requirements. While our investigation focuses on voice transmission, most of our concepts apply to other forms of continuous media traffic, e.g., low-bandwidth video. We examine the feasibility of retransmission-based error recovery for continuous media traffic. In order to be effective, retransmissions of lost packets must be completed within the delay constraints of the packet stream. However, if timely retransmission can be achieved with a high probability of success, a retransmission-based approach to error control is attractive because it imposes little overhead on network resources. Note that retransmission-based error recovery can be used in conjunction with extant preventive error control schemes such as forward error correction or channel coding.

We employ analytical modeling techniques to investigate the effectiveness of retransmission for different network scenarios. To explore the relationship between our theoretical results and the dynamic behavior of voice transmission over existing networks, we present measurements of the delays experienced by voice transmission running over a contemporary campus backbone network. Our results indicate that retransmission-based error recovery can be effective for many end-to-end transmission scenarios in current networks.

The remainder of this paper is structured as follows. In Section 2 we review issues that must be addressed by protocols for voice distribution in packet-switching networks. In Section 3 we develop an analytic model for the end-to-end transmission of packet voice and derive a performance metric for timely retransmission in the presence of errors. We present examples where we apply the performance metric under variations of network parameters. In Section 4 we provide measurements of voice packets on a multiple-segment local area network and compare the empirically obtained data with our theoretical findings. In Section 5 we present the conclusions of the paper.

2 Protocol Issues

Continuous media protocols must have mechanisms to address all factors that may degrade the quality of remote playback. In this section we briefly discuss important issues for maintaining high quality voice transmission, and discuss how these issues are resolved in extant packet voice protocols. An important consideration in the design of packet voice protocols is that speech is actually an alternating series of activity periods, or *talkspurts*, followed by silence periods with the activity periods constituting only around 40% of the total time [4].

2.1 Encoding and Packetization

The packet voice source continuously collects and buffers digitized voice samples. After a fixed period of time, the so-called *packetization interval*, voice samples collected by the audio hardware are placed into a network packet, and the packet is submitted to the network. Typical packetization intervals range from 10 – 50 ms [19].

Given a fixed packetization interval, the encoding scheme determines the actual number of bits per packet. The ubiquitous pulse code modulation (PCM) encoding scheme for voice [6] samples every 125 μs with 8 bits per sample to yield a 64 Kbit/s channel. Bandwidth reduction can be achieved through the use of fewer bits per sample, less frequent sampling, suppression of transmission during silence periods, and compression of the digitized data. Adaptive differential pulse code modulation (ADPCM) [7], for example, encodes only the difference between consecutive samples, reducing the number of bits per sample to $2-5$ bits. Coding techniques with even lower bit rates, e.g., Linear Predictive Coding (LPC), exist, though speech fidelity is frequently poor [11].

2.2 Roundtrip Delay

Behavioral studies [5, 13] have shown that roundtrip delays above a certain threshold degrade the interactive nature of the conversation. Quantifying this factor is difficult since individual human users have different tolerances for delay and these tolerances vary with the application. High-quality voice applications require less than 200 ms roundtrip delays, but delays of up to 600 ms have been shown to be acceptable [13].

Since current packet-switching networks do not provide a bounded delay service, voice protocols must provide mechanisms that can cope with highly variable end-to-end delays. Adjustments of the packetization interval and buffering of voice packets at the receiver are widely used to compensate for unpredictable delays.

2.3 Delay Jitter

If the network delay of voice packets is not constant, e.g., due to statistical multiplexing, discontinuity of the voice playback at the receiver can occur. We refer to these discontinuities as *gaps*. Gaps are commonly addressed through buffering at the receiving site. The first packet in a talkspurt is artificially delayed at the receiver for a period of time known as the *control time*. The control time builds up a buffer of arriving packets sufficient to provide continuous playback in the presence of delay jitter. Note however, that the control time cannot be arbitrarily large due to constraints on the roundtrip delay.

The use of a control time to compensate for delay jitter requires mechanisms to identify the beginning of talkspurts and to determine the control time. The latter is difficult since it requires knowledge of the network delay distribution. Numerous methods have been proposed for estimating the control time of a talkspurt, based on network delay measurements [15], on stochastic assumptions about the network delay [1, 2], or both [16].

2.4 Error Control

The impact of packet loss on voice quality varies since interpolation can mitigate the effects of lost samples and not all samples contain equally important information. In any case, the tolerance to packet loss is low and even the loss of even a single packet may be perceptible during playback.

Packet-level error control for packet voice streams must be designed so as to provide the best possible quality for the stream. Conventional error control techniques are unacceptable since they do not consider the delay sensitivity of voice data. Hence, researchers have dismissed a retransmission-based approach to error control, focusing instead on open-loop techniques that recover or limit the effects of losses.

Forward error correction (FEC) [3, 17] provides robustness in the presence of packet loss by adding redundant information to the original samples. If only a small number of packets is lost, the added redundancy enables a reconstruction of the original voice data at the receiver. The ability to recover lost information strongly depends on the degree of redundancy. In addition to considerable processing overhead, FEC-based error control results in increased network bandwidth consumption. Thus, FEC contributes to network congestion, and, since losses in the network are most often due to congestion, may even be the cause of packet loss.

Channel coding refers to a class of approaches that separate the voice signal into multiple data streams with different priorities. The priorities are used by network switches to selectively discard low priority packets which carry information that is less crucial in reconstructing the voice signal. Channel coding techniques have been shown to provide a graceful degradation of playback quality in a variety of loss scenarios [10, 20, 23]. For PCM-encoded voice, packet loss rates of over 5% on the channel carrying the least significant information have been reported as tolerable when using small (32-byte) packets [21]. A drawback to channel coding, however, is that the network is required to support selective discarding of packets during periods of congestion.

3 An Analytical Model for Retransmission of Packet Voice Data

In this section we present an analytical retransmission model for error control of a voice packet stream in a packet-switching network. Through analysis of the model we are able to quantify the gain in quality from the retransmission of packet losses in the network. Given an encoding scheme, voice quality depends primarily on maintaining the continuity of the playback of each talkspurt at the receiver. The loss of voice quality results from discontinuities due to delay jitter or packet loss. Since they both cause gaps, delay jitter and packet loss cannot be considered separately. Timely retransmission is of little value if discontinuities due to delay jitter degrade the quality of the playback. For this reason we define the performance metric for measuring transmission quality as the probability of continuous playback, i.e., a playback without gaps, of an entire talkspurt. Under a given packet loss scenario, this metric accounts for the quality degradation due to delay jitter *and* to untimely retransmission. The quality degradation due to delay jitter alone is found by computing the metric under the assumption of error-free transmission.

Our model considers all protocol issues reviewed in the previous section. Since packets in different talkspurts rarely interfere with each other, We model the transmission of a single talkspurt within a packet voice stream. The sender introduces packets into the network with a deterministic spacing as given by the packetization interval. The network delay of each packet is assumed to follow an arbitrary but fixed distribution. Since packets are spaced relatively far apart in time, each packet in the voice stream will experience a network delay independent of the other packets in the stream. Note, however, that due to the requirement for in-order playback of the packets in a talkspurt, the model enforces in-sequence delivery of packets at the voice receiver. The in-sequence delivery requirement is independent of the underlying network service. For a connectionless network, correct sequencing is provided by a higher layer protocol, while for a connection-based network such as an ATM network, sequential delivery is provided by the network itself.

In Section 3.1 we give a detailed description of an end-to-end retransmission model. In Section 3.2 we develop analytic expressions for the effectiveness of retransmission of lost voice data. In Section 3.3 we present numerical examples of our results.

3.1 End-to-End Model for Packet Voice Transmission

Starting at time $t = 0$, the sender generates voice packets after packetization intervals of fixed length $\overline{x}$. A talkspurt is assumed to consist of a fixed number of N packets. The transmission times of packets are assumed to be negligible compared to the packetization interval. Thus, $\overline{x}$ represents the distance between packets at the entrance of the network.

The network delay experienced by a packet is described by a distribution F_D, and D_j is used to denote the network delay of the jth packet in a talkspurt. Network delays are assumed to be independent. However, since the model considers in-sequence delivery of packets to the voice receiver, which is enforced by either the network or a higher layer protocol, the dependence between packet delays due to resequencing is considered.

When the first packet of a talkspurt arrives to the receiver, playback of the talkspurt is delayed for the duration of the so-called control time, denoted by V ($V \geq 0$). Thus, playback of the talkspurt is started at $t = D_1 + V$. The playback duration of each packet is identical to the packetization interval $\overline{x}$.

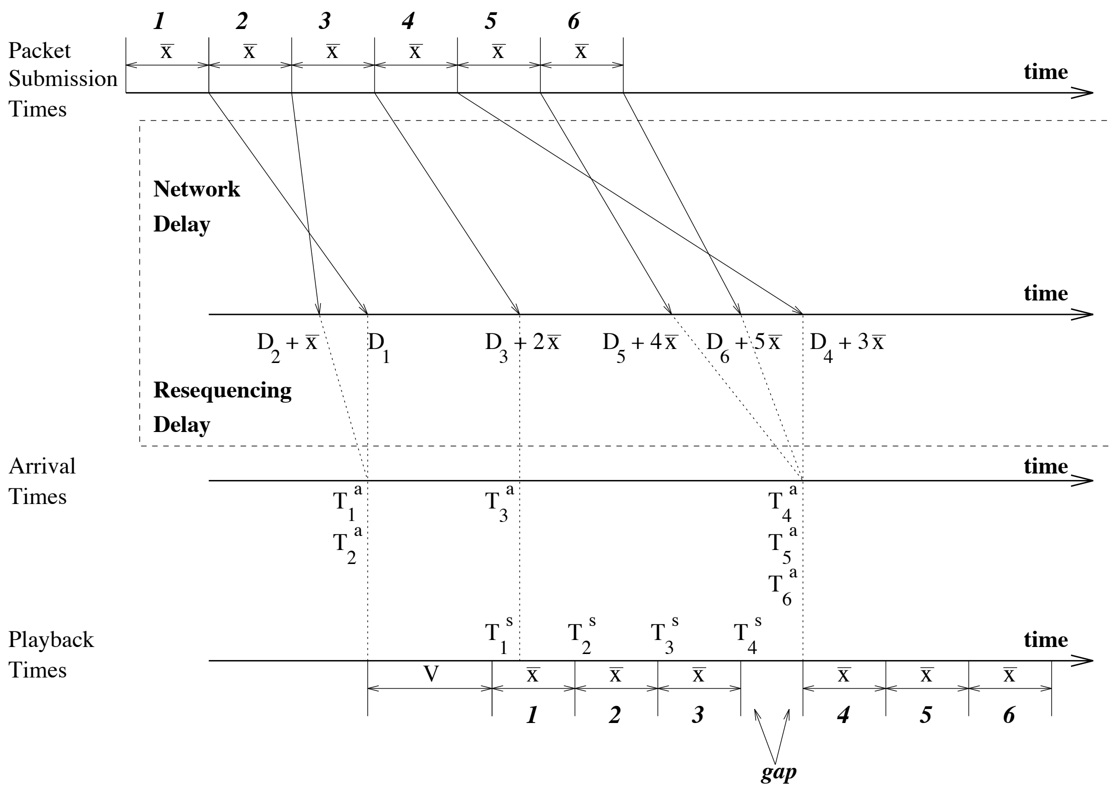

Figure 1: Transmission Model of a Talkspurt.

The end-to-end transmission model is summarized in Figure 1 for a talkspurt consisting of $N = 6$ packets. In the top of the Figure we show the transmission of packets with a distance of $\overline{x}$ time units. Each packet experiences an independent network delay with D_j denoting the network delay of the jth packet in the talkspurt. Due to the additional delay required for proper sequencing, the arrival time of the jth packet at the receiver, denoted by T_j^a, is given by

$$T_j^a = \max_{i=1,\ldots,j} \{D_i + (i-1)\overline{x}\} \tag{1}$$

Assuming that no playback discontinuities have occurred before the arrival of the jth packet, the scheduled playback time for the jth packet, denoted by T_j^s, is fully determined once the first packet has arrived at the receiver. As shown on the bottom timeline in Figure 1, T_j^s is given by

$$T_j^s = D_1 + V + (j-1)\overline{x} \tag{2}$$

In Figure 1, the fourth packet arrives after its playback time, that is, $T_4^s < T_4^a$, causing a discontinuity, or *gap*, in the playback of the talkspurt. Once a gap occurs in the playback of a talkspurt, playback synchronization is lost, and all subsequent packets in the talkspurt, e.g., the fifth and sixth packets in Figure 1, begin playback after their scheduled playback times.

We assume that during the transmission of a talkspurt there is an arbitrary period, the so-called *error period*, during which zero or more packets are dropped by the network. We assume only one error period per talkspurt, but allow multiple consecutive packets to be dropped in an error period. Since packets of a talkspurt arrive at the receiver in the order in which they are transmitted, the receiver detects losses as soon as a packet arrives out-of-sequence. After detecting a loss, a retransmission procedure is initiated. The time to recover lost packets via retransmission is fully determined by a roundtrip network delay. Processing times for retransmissions at the receiver or the sender are assumed to be small and not considered in our model. Also, we assume that the sequence of lost packets in an error period can be retransmitted in a single packet. This assumption is realistic for network or transport layer protocols since the maximum packet size is much larger than the typical voice packet. Thus, denoting the retransmission time by R, the distribution function of R is given by $F_R = F_D \otimes F_D$ where $\otimes$ is the convolution operator.

3.2 Analysis of End-to-End Model

In this subsection we develop an exact analytic expression for the probability of continuous playback of a talkspurt. First, we derive the desired probability assuming an error-free scenario. Then, we extend our expression to consider error periods.

3.2.1 Probability of Continuous Playback Without Errors

We are concerned with the occurrence of a gap in the playback of a talkspurt. We thus define random variables G_i $(1 \le i \le N)$ that indicate the presence of a discontinuity in the playback. By setting

$$G_i := \begin{cases} V & \text{if } i = 1 \\ 0 & \text{if } G_{i-1} = 0,\ i \neq 1 \\ \max\{0, T_i^s - T_i^a\} & \text{otherwise} \end{cases} \tag{3}$$

we obtain $G_i = 0$ if a packet with index i or less arrives after its playback time. Since the arrival of the first packet sets the playback schedule and cannot cause a gap, that is, $T_1^s - T_1^a = V$, we set $G_1 = V$. For a talkspurt with N packets, $G_N > 0$ indicates that no discontinuity has occurred during playback of the entire talkspurt.

Note that in equation (3), $G_i = 0$ for the ith packet is feasible in two scenarios. Either no gap has occurred before packet i and packet i arrives after its playback time, or a gap has occurred before the arrival of packet i. Therefore, by denoting $\mathrm{P}\{G_i = 0\}$ as the probability of $G_i = 0$, we obtain for $2 \le i \le N$:

$$\mathrm{P}\{G_i = 0\} \quad = \quad \mathrm{P}\{G_{i-1} = 0\} + \mathrm{P}\{G_{i-1} > 0 \ \mathit{and} \ T_i^s < T_i^a\} \tag{4}$$

The second term on the right side of equation (4) is the probability that the ith packet causes the first gap in the playback of the talkspurt. In this case, all packets with index less than i have arrived before their respective playback times. It follows that packet i could not have arrived earlier than any packet with a smaller index, that is,

$$T_i^a = \max_{j=1,\ldots,i} \{D_j + (j-1)\overline{x}\} = D_i + (i-1)\overline{x} \tag{5}$$

Substituting equation (5) into equation (4) yields

$$\mathrm{P}\{G_i = 0\} \quad = \quad \mathrm{P}\{G_{i-1} = 0\} + \mathrm{P}\{G_{i-1} > 0 \ \textit{and} \ V + D_1 < D_i\} \tag{6}$$

Here, we can recursively compute the probability for continuous playback, i.e., $G_N > 0$, by

$$\mathrm{P}\{G_N > 0\} \quad = \quad 1 - \int_0^\infty \mathrm{P}\{G_N = 0 \mid D_1 = t\} dF_D(t) \tag{7}$$

3.2.2 Probability of Continuous Playback in the Presence of Errors

In the presence of errors, gaps in the playback of a talkspurt may result from delay jitter or from a failure of the retransmission procedure to recover lost packets before their playback times. In this subsection we calculate the probability of continuous playback of a talkspurt, given that the network loses k consecutive packets, say packets with index $n-k, n-k+1, \ldots, n-1$. We exclude the loss of the first packets in a talkspurt, i.e., $n > k+1$. Note that in a talkspurt containing an error period, the arrival time of the jth packet is obtained by calculating the latest packet arrival with index less than j that is not lost in the network. Thus we obtain for T_j^a that

$$T_j^a = \begin{cases} \max\limits_{l=1,\ldots,j} \{D_l + (l-1)\overline{x}\} & \text{if } j \leq n-k-1 \\ \max\limits_{l=1,\ldots,n-k-1,n,n+1,\ldots,j} \{D_l + (l-1)\overline{x}\} & \text{otherwise} \end{cases} \tag{8}$$

For calculating the probability of continuous playback in the presence of errors, we must consider two cases which can result in gaps. First, a gap may be due to an untimely retransmission of the lost packets. We refer to this case as an *error gap* or *E-gap*. Second, a gap may result from excessive delay variations, independent of the loss of packets. This case is referred to as a *jitter gap* or *J-gap*. Thus, $\mathrm{P}\{gap\}$, the probability of a gap in the playback of a talkspurt, is given by

$$\mathrm{P}\{gap\} \quad = \quad \mathrm{P}\{E\text{-}gap\} + \mathrm{P}\{no\ E\text{-}gap\}\mathrm{P}\{J\text{-}gap \mid no\ E\text{-}gap\} \tag{9}$$

The calculation of $\mathrm{P}\{gap\}$ is performed in three steps. We first calculate the probability of an *E-gap*. Next, using an approach similar to that of the previous subsection, we define random variables G_i that indicate the occurrence of a *J-gap* at or before the arrival of packet i. We then calculate

P{*J-gap* | *no E-gap*}, which is the probability of a *J-gap* in the talkspurt under the condition that the lost packets are retransmitted in a timely fashion. Note that a jitter gap can occur before or after the lost packets are due for playback. In the latter case, the discontinuity occurs independent of the retransmission procedure.

First we consider the probability of gaps due to untimely retransmission. Recall that the time necessary for retransmission is denoted by R with $F_R = F_D \otimes F_D$ and the arrival of the nth packet, that is, the first packet after the sequence of lost packets, invokes the retransmission procedure. For retransmission to be untimely, the difference between the playback time of the $(n-k-1)$st packet, i.e., the first lost packet, and the arrival time of the nth packet must be less than R, the time necessary for retransmission. Thus, the probability of an *E-gap* is

$$\begin{aligned} \mathrm{P}\{E\text{-}gap\} &= \mathrm{P}\{T^s_{n-k} - T^a_n < R\} \\ &= \mathrm{P}\left\{ \max_{j=1,\ldots,n-k-1,n} \{D_j + (j-1)\overline{x}\} > V + D_1 + (n-k-1)\overline{x} - R \right\} \end{aligned} \tag{10}$$

Using fixed values for the retransmission time, $R = r$, and the delay of the first packet, $D_1 = t$, we obtain from equation (10) that

$$\begin{aligned} &\mathrm{P}\left\{ \max_{j=2,..,n-k-1,n} \{t, D_j + (j-1)\overline{x}\} > (V + t + (n-k-1)\overline{x} - r) \mid D_1 = t,\ R = r \right\} = \\ &\quad 1 - \left(\mathrm{P}\left\{V + (n-k-1)\overline{x} > r\right\} F_D(V + t - k\overline{x} - r) \prod_{j=2}^{n-k-1} F_D(V + t + (n-k-j)\overline{x} - r) \right) \end{aligned} \tag{11}$$

In equation (11), we have used the independence of network delays. Now we uncondition the expression in equation (11) by integrating over the retransmission delay and the delay of the first packet in the talkspurt. That is,

$$\begin{aligned} \mathrm{P}\{E\text{-}gap\} &= \int_t \int_r 1 - \ \mathrm{P}\ \{V + (n-k-1)\overline{x} > r\} F_D(V + t - k\overline{x} - r) \cdot \\ &\qquad \left(\prod_{j=2}^{n-k-1} F_D(V + t + (n-k-j)\overline{x} - r) \right) dF_R(r) dF_D(t) \\ &= 1 - \int_t \int_{r=0}^{V+(n-k-1)\overline{x}} F_D(V + t - k\overline{x} - r) \prod_{j=2}^{n-k-1} F_D(V + t + (n-k-j)\overline{x} - r) dF_R(r) dF_D(t) \end{aligned} \tag{12}$$

In the following we assume that retransmission is timely, that is, an *E-gap* has not occurred. We compute P{*J-gap* | *no E-gap*}, the probability that a jitter gap occurs in the playback of a talkspurt under the assumption of timely retransmission. Similarly to equation (3), we define random variables G_i, such that $G_i = 0$ if a packet with index i or less arrives after its playback time. With our assumption that packets $n-k, n-k+1, \ldots, n-1$ are lost and do not arrive to the receiver we obtain for $1 \le i \le n-k-1$ or $n \le i \le N$

$$G_i := \begin{cases} V & \text{if } i = 1 \\ G_{n-k-1} & \text{if } i = n \\ 0 & \text{if } G_{i-1} = 0, i \ne 1, i \ne n \\ \max\{0, T_i^s - T_i^a\} & \text{otherwise} \end{cases} \tag{13}$$

With this definition, $\mathrm{P}\{G_N = 0\}$ is the desired probability that, assuming timely retransmission, a jitter gap occurs in the playback of a talkspurt. Since the definition of G_i is recursive, we must compute $\mathrm{P}\{G_i = 0\}$ for all values of i. From the definition in equation (13), $\mathrm{P}\{G_1 = 0\} = 0$. That is, the first packet sets the playback schedule and cannot cause a jitter gap. Also, the nth packet cannot cause a jitter gap since its arrival invokes the retransmission procedure, which could not be timely if the nth packet arrived after its playback time. Hence $\mathrm{P}\{G_n = 0\}$ is defined to be the probability that a jitter gap occurs before the nth packet, i.e., $\mathrm{P}\{G_n = 0\} = \mathrm{P}\{G_{n-k-1} = 0\}$. Two cases remain to be considered: $2 \le i \le n-k-1$ and $n \le i \le N$.

We first calculate the probability of a *J-gap* at or before the ith packet where $2 \le i \le n-k-1$. In this case the ith packet arrives early enough to ensure successful retransmission of the lost packets, but the ith packet can cause a jitter gap by arriving after its playback time. That is, the conditions for the ith packet resulting in a jitter gap are

$$T_i^a > T_i^s \qquad \text{and} \qquad T_i^a \le T_n^a < T_{n-k}^s - R \qquad \text{for } 2 \le i \le n-k-1 \tag{14}$$

Hence, for $2 \le i \le n-k-1$,

$$\mathrm{P}\{G_i = 0\} \;=\; \mathrm{P}\{G_{i-1} = 0\} + \mathrm{P}\{G_{i-1} > 0 \;\; and \;\; T_i^s \le T_i^a \le T_{n-k}^s - R\} \tag{15}$$

The last term in equation (15) is the probability that the first jitter gap in the playback occurs upon the arrival of the ith packet. By using a similar argument as for the derivation of equation (5), the arrival time of the ith packet is $D_i + (i-1)\overline{x}$, and equation (15) can be rewritten as

$$\begin{aligned} \mathrm{P}\{G_i = 0\} = {} & \mathrm{P}\{G_{i-1} = 0\} \\ & + \mathrm{P}\{G_{i-1} > 0 \;\; and \;\; V + D_1 \le D_i \le V + D_1 + (n-k-i)\overline{x} - R\} \end{aligned} \tag{16}$$

Equation (16) is equivalent to

$$\begin{aligned} \mathrm{P}\{G_i = 0\} = {} & \int_t \int_r \mathrm{P}\{G_{i-1} = 0 \mid D_1 = t, R = r\} dF_R(r) dF_D(t) \\ & + \int_t \int_r \mathrm{P}\{G_{i-1} > 0 \mid D_1 = t, R = r\} (F_D(V + t + (n-k-i)\overline{x} - r) - F_D(V + t)) dF_R(r) dF_D(t) \end{aligned} \tag{17}$$

We have shown how to compute $\mathrm{P}\{G_i = 0\}$ for $i \leq n-k-1$. Next we consider $\mathrm{P}\{G_i = 0\}$ for $i > n$. Since packets with index greater than n cannot affect the retransmission procedure, we obtain with the definition in equation (13):

$$\mathrm{P}\{G_i = 0\} \quad = \quad \mathrm{P}\{G_{i-1} = 0\} + \mathrm{P}\{G_{i-1} > 0 \quad \mathit{and} \quad T_i^s \leq T_i^a\} \qquad \text{for } n+1 \leq i \leq N \qquad (18)$$

Note that the second term on the right side of equation (18) is the probability that the ith packet causes the first gap in the playback of the talkspurt. In this case, all packets with index less than i have arrived before their respective playback times. It follows that packet i could not have arrived earlier than any packet with a smaller index, that is,

$$T_i^a = \max_{j=1,\ldots,n-k-1,n,n+1,\ldots,i} \{D_j + (j-1)\overline{x}\} = D_i + (i-1)\overline{x} \qquad \text{for } n+1 \leq i \leq N \qquad (19)$$

Then, equation (18) yields, for $n+1 \leq i \leq N$,

$$\mathrm{P}\{G_i = 0\} \quad = \quad \int_t \mathrm{P}\{G_{i-1} = 0 \mid D_1 = t\} + \mathrm{P}\{G_{i-1} > 0 \mid D_1 = t\}\mathrm{P}\{V + t \leq D_i\} dF_D(t) \quad (20)$$

At this point we can compute $\mathrm{P}\{gap\}$, the probability of a gap in the playback of a talkspurt, from equation (9) since we have $\mathrm{P}\{E\text{-}gap\}$ in equation (12), and $\mathrm{P}\{J\text{-}gap \mid no\ E\text{-}gap\}$ by recursively evaluating equations (17) and (20).

3.3 Numerical Examples

We present numerical examples that apply our analysis for determining the effectiveness of retransmission without disrupting the continuous playback of voice packets for various network and protocol parameters. We present three examples, where in each example the effectiveness of retransmission is expressed in terms of the probability of maintaining playback continuity during a talkspurt, as derived in the previous subsection.

- Example 1 shows the sensitivity of the probability of continuous playback in the presence of errors to different control time values.
- Example 2 shows the degree to which the network delay distribution influences retransmission-based error recovery.
- Example 3 shows the effects of the average network delay on retransmission-based error recovery.

For the network delay, we consider delay distributions with different variance, namely Erlang distributions with k exponential phases, denoted by Erlang-k or E_k, for $k \geq 1$. Networks with large

Example	Packetization Interval ($\overline{x}$)	Packets in Talkspurt (N)	Avg. Network Delay	Network Delay Distribution (F_D)
1	20 ms	20	15 ms	E_2
2	20 ms	20	15 ms	E_1, E_2, E_6
3	20 ms	20	2,10,20,30,40 ms	E_2

Table 1: Network Parameters.

delay variations are modeled by E_1, that is, an exponential distribution; for moderate and low delay variations we use, respectively, E_2 and E_6. In accordance with our empirical delay measurements (see Section 4) and studies on wide-area connections (e.g., [18]), we select E_2 as the default network delay distribution. The selection reflects that delay variations over short periods of times, such as the duration of a talkspurt, are generally modest.

The parameters for our examples are presented in Table 1. The default average network delay is set to 15 ms. The default packetization interval is fixed at $\overline{x} = 20$ ms, a value commonly used in extant voice protocols [19], and the default number of packets in a talkspurt is $N = 20$. Thus, each talkspurt has a duration of 400 ms, a value motivated by our empirical measurements of packet voice traffic in Section 4.

3.3.1 Example 1: Effects of the Control Time on Retransmission

Recall that in our end-to-end model we specify a single error period during the transmission of a talkspurt, but multiple consecutive packets can be lost during an error period. Here, we consider error periods in which zero, one, two, or three packets are lost. An error period in which i packets are lost is referred to as an *i-error* scenario.

Figure 2 shows the probability of continuous playback of the talkspurt under variation of the control time. The figure depicts four curves representing the respective error scenarios. The *0-error* scenario is included in order to consider the effects of delay jitter on the playback continuity of talkspurts whose end-to-end transmission is error-free. From the *0-error* curve we see that a control time of roughly $V = 60$ ms is required to compensate for the delay jitter in the network. The *1-error* curve gives the probability that playback is continuous for a talkspurt during whose transmission exactly one packet is lost in the network and subsequently retransmitted. With $V = 60$ ms, the *1-error* curve shows that approximately 70% of single-packet losses are successfully recovered through retransmission. As the control time is lengthened, successful retransmission is more likely, and at $V = 100$ ms, successful retransmission in both the *1-error* and *2-error* scenarios occurs in over 90%

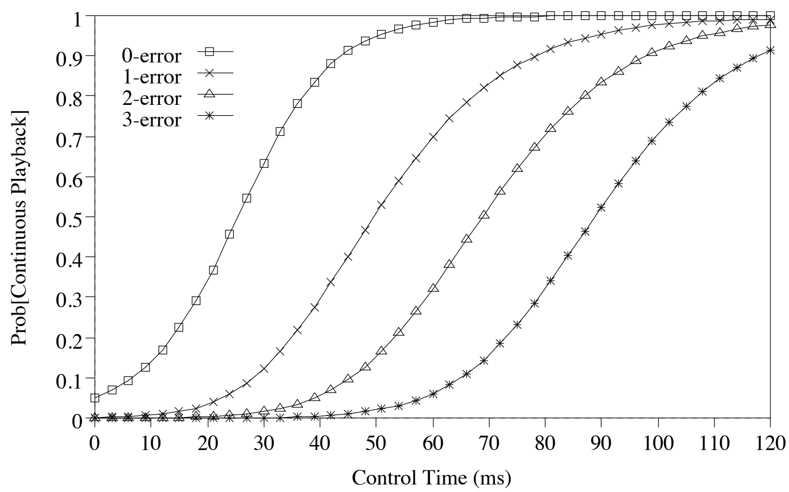

Figure 2: Retransmission Effectiveness for an E_2 (Erlang-2) Network Delay Distribution.

of the cases.

Recall that the feasible range of control time values is determined by the end-to-end delay restriction. In our example, the sum of the packetization interval and the network delay on the average account for only 35 ms of the total end-to-end delay. Hence control times on the order of $V = 100$ ms are feasible for all but the most stringent delay requirements.

The packetization interval plays an important role in the retransmission algorithm. In a k-error scenario, the average amount of time that elapses between the occurrence of packet loss and its discovery at the receiver is $k\overline{x}$ since the receiver discovers packet loss when the first out-of-sequence packet arrives. Hence the probability of successful retransmission in a k-error scenario will be low when $V < k\overline{x}$. This can be observed in Figure 2 where the recovery rate for control times of less than $k\overline{x}$ is roughly 5%, e.g., for the *3-error* scenario at a control time of $V = 60$ ms, 6% of retransmission attempts are successful. The influence of the packetization interval is graphically evident in Figure 2: to achieve a fixed probability of successful retransmission, the control times required for the *1-error* and *2-error* scenarios differ by approximately the size of the packetization interval. The same relationship is observed between the control times for the *2-error* and *3-error* scenarios.

3.3.2 Example 2: Effects of Network Delay Variation on Retransmission

Here we study the effects of different network delay distributions on the probability of successful retransmission. We select E_1 to represent networks with high delay variations, E_2 for moderate delay variations, and the E_6 for low delay variations. Before examining retransmission behavior,

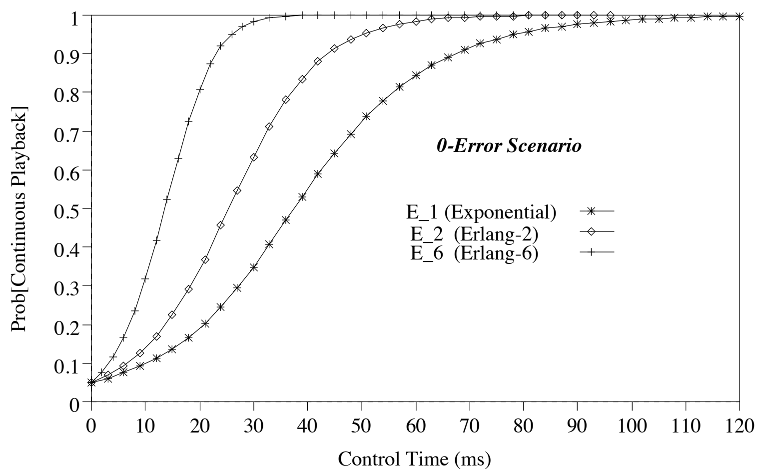

Figure 3: Delay Jitter for Different Network Delay Distributions.

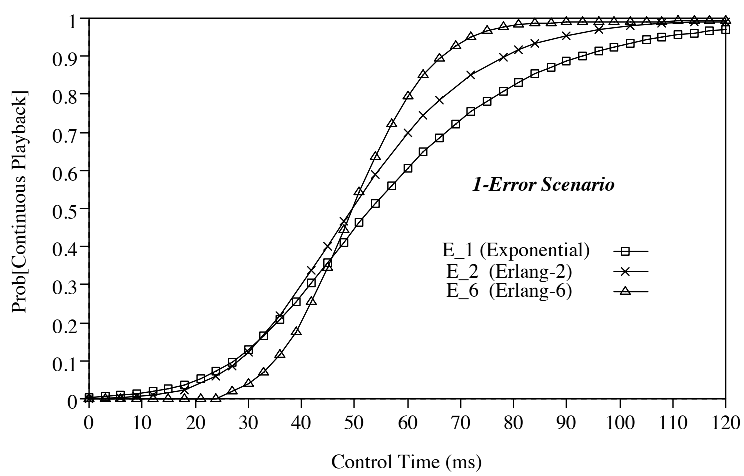

Figure 4: Retransmission Effectiveness for Different Network Delay Distributions.

we first consider the differences in the amount of delay jitter in the network under these delay distributions.

Figure 3 shows the probability of continuous playback of a talkspurt for the respective network delay distributions, given that the transmission of the talkspurt is error-free. The size of the control times necessary to compensate fully for the delay jitter in the network varies widely: $V = 100$ ms for E_1, but only $V = 60$ ms for E_2, and $V = 30$ ms for E_6.

Figure 4 shows the probability of continuous playback of a talkspurt for the respective network delay distributions, given that a single-packet loss occurs during the transmission of the talkspurt. For small control times, the higher variation of E_1-distributed delays results in a greater probability of successful retransmission than the other distributions. However, E_1 results in a lower probability

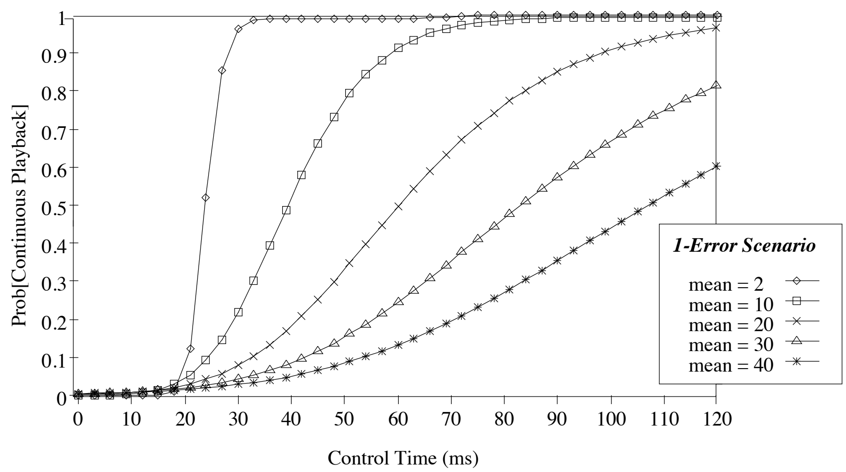

Figure 5: Retransmission Effectiveness for Different Average Network Delays.

of successful retransmission than E_2 and E_6 for larger control time values. Hence the delay distributions with low variation will inherently recover more packets successfully. For instance, when the control time is $V = 70$ ms, the probability of successful retransmission is approximately 0.75 for E_1, 0.85 for E_2, and 0.95 for E_6.

3.4 Example 3: Effects of Average Network Delay

In this example we examine the effects of varying the mean network delay for a fixed network delay distribution, i.e., E_2. All other parameters are the same as in Example 1. The *1-error* curves for average network delays of 2, 10, 20, 30, and 40 ms are shown in Figure 5.

Figure 5 shows the effect of increasing the average network delay on retransmission effectiveness. Retransmissions are rarely successful whenever $V \leq 20$ ms since, as discussed in Example 1, the packetization interval represents the average time required for the discovery of a single-packet loss at the receiver. The curve for a mean network delay of 2 ms graphically illustrates this lower bound on the control time required for a high probability of successful retransmission.

The curves reveal a strong interaction between retransmission effectiveness and the one-way network delay. As the average network delay increases, so does the absolute size of delay jitter in the network. Retransmission behavior is also linked to the roundtrip network delay, amplifying the effects of increases in one-way network delays. For example, consider a control time of $V = 80$ ms in Figure 5. If the average network delay is 10 ms, this control time will allow for the recovery of almost all single-packet losses. If the average network delay is 20 ms, successful recovery falls to 80%, for 30 ms to approximately 50%, and for 40 ms to approximately 25%.

For the network parameters in Example 3, we conclude that our retransmission-based approach to error control will not be advantageous if the mean network delay is at or above roughly 30 ms. For average network delays below 30 ms, the results suggest that there is a high probability for successful retransmission of small burst losses.

4 Empirical Measurements of Voice Traffic

In this section we turn to empirical network measurements of voice transmission and compare our analytical results with the dynamic behavior of a contemporary network, the backbone network of the University of Virginia. This network is representative of large enterprise networks consisting of Ethernet segments and FDDI rings that are connected by high-performance routers. We present delay and loss measurements for voice transmission across a multiple-segment path in this network. Then we extrapolate these measurements to explore the question of whether retransmission-based error control for real-time traffic is feasible in this network environment. In Section 4.1 we outline our data collection method. In Section 4.2 we present representative samples of the empirical data collected, and in Section 4.3 we provide an interpretation of the measurements in the context of retransmission-based error recovery.

4.1 Experimental Approach

To obtain a network traffic profile for packet voice, we used the packet voice capability of the INRIA Videoconferencing System (IVS), a public domain software videoconferencing tool [22], running on a Sun SPARCstation. We selected 32 Kbit/s ADPCM encoding with silence detection and the default packetization interval of 20 ms.

In the measurement experiment, the workstation that generates the IVS-based packet voice traffic sends a stream of datagrams to a process on a remote machine using the UDP/IP protocol stack. The size and interpacket times for the datagram stream are taken from approximately 5 minutes of male monologue speech processed with the IVS package. When the remote machine receives a datagram, it immediately reflects the datagram back to the traffic-generating workstation.

Packet delays were measured by a software-based network monitoring tool [14] running on an Intel 486 workstation, which was attached to the same local-area network as the traffic-generating workstation. This network monitor node captured each packet on the local network segment and timestamped the packet with a clock accuracy of ± 1 ms. To calculate roundtrip delays, packets were timestamped as they left and as they returned to the local network segment. The roundtrip time of a packet thus represents endsystem processing at the remote machine as well as the time to cross the network twice. Examination of sequence numbers in packet headers verified that the network monitor did not drop packets.

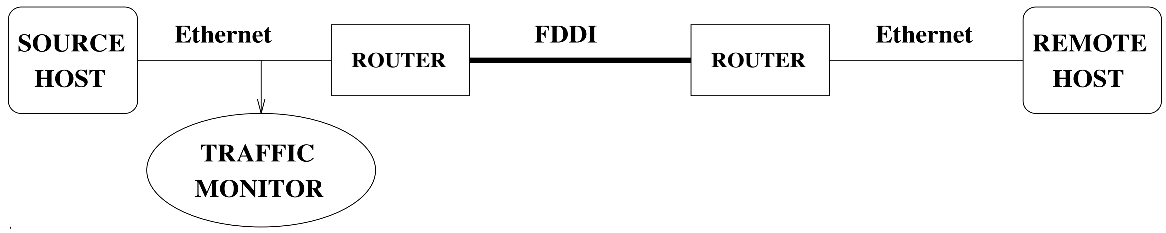

Figure 6: Network Path for Voice Transmission.

We measured the characteristics of a path through the network illustrated in Figure 6. This path consists of three network segments: an Ethernet network on which the source machine is located, an FDDI backbone, and an Ethernet network on which the remote machine is located. The traffic-generating machine and the remote machine are desktop workstations that were lightly loaded while measurements were being taken.

The data presented here was taken in at 11:00 AM on a weekday. [1] There was a background load on the local Ethernet segment of approximately 5 Mbits/s on average during each experiment.

4.2 Empirical Packet Delays and Losses

The measured roundtrip delays for our experiment are shown in Figure 7. Each data point represents the roundtrip delay of one packet. The negative spikes indicate dropped packets. In Figure 7 we observe that for the first 90 seconds the network delay is low, i.e., most of the roundtrip times are in the range of 2 – 8 ms. However, starting at 90 seconds into the experiment, the network experiences large fluctuations in roundtrip times with peaks up to nearly 200 ms. This behavior continues for approximately 60 seconds, at which point the network roundtrip times return to a lower range with a few delays in the 30 – 50 ms range.

For most of the experiment, the empirical data indicates a very fast path across the network and through the kernel-level processing at the remote machine. Figure 8 shows the roundtrip times during the time period labeled as the 'high delay period' in Figure 7. In Figure 8, the roundtrip delays of individual packets display a pattern whereby the end-to-end delays for consecutive packets associated with a spike are approximately 20 ms apart. Since the packets were transmitted from the source 20 ms apart, this spacing indicates that the packets were queued at some transient bottleneck and then released in a burst.

As seen in Figure 7, packet loss did occur in the experiment. Over the entire data set, the loss rate is small, but, during the high delay period in Figure 7, packet loss affects over 4% of the talkspurts. However, over the entire duration of the experiment, losses are irregularly spaced in

[1] Other data sets corroborate the essential aspects of the presented measurements [9].

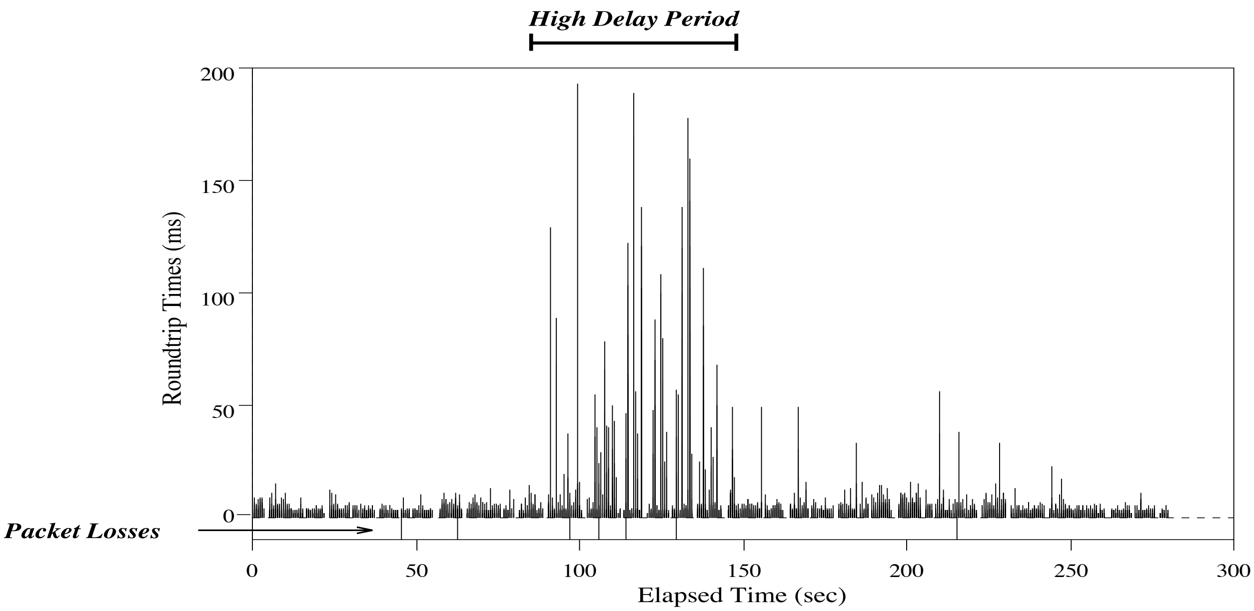

Figure 7: Measured Roundtrip Times of Voice Packets.

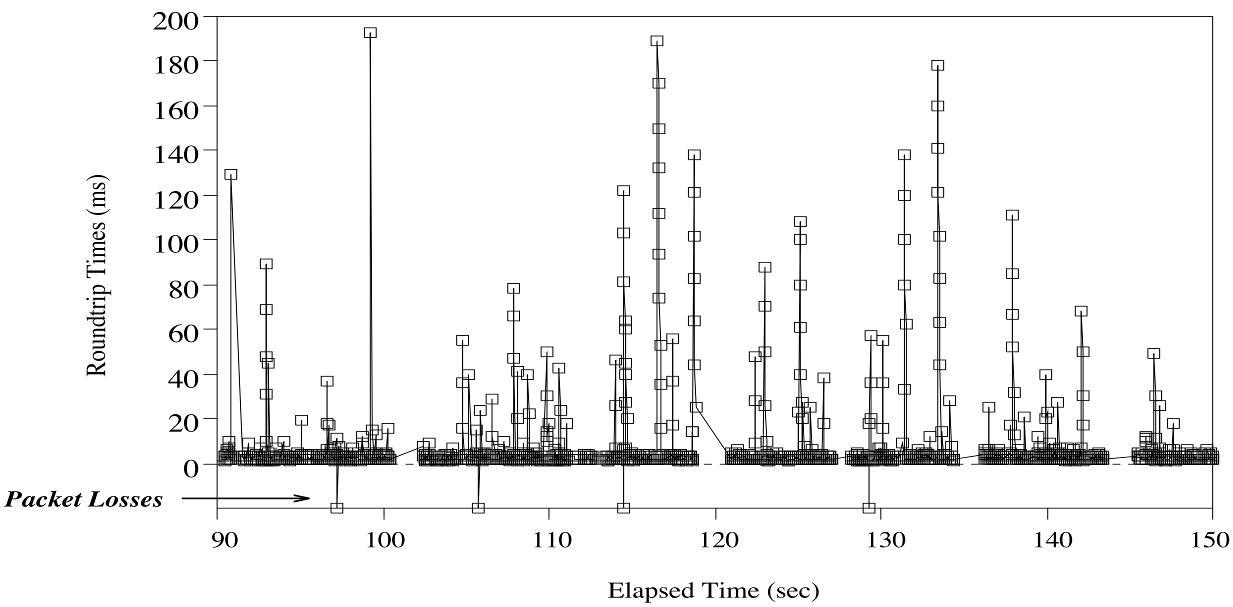

Figure 8: Roundtrip Times of Voice Packets in the High Delay Period of Figure 7.

time and only loosely correlate with periods of high network delays.

4.3 Retransmission Probabilities

To investigate the effectiveness of retransmission-based error recovery for the data sample presented in the previous subsection, we perform the following construction. We first extrapolate the talkspurt boundaries in our measurements. For each talkspurt, using the measured network delay of its packets and a given control time value, we determine the packet arrival times and the playback schedule at the voice receiver. Then, as with the *k-error* scenario in Section 3.1, we introduce an error period in which k consecutive packets of the talkspurt are dropped in the network and subsequently retransmitted. We estimate the retransmission time for the dropped packets and determine whether the retransmission is timely. This calculation is repeated for all possible positions for a *k-error* period in all talkspurts. In this way we extrapolate the probability of playback continuity of the talkspurts at the receiver under *k-error* scenarios, and the construction allows us to compare the empirical measurements with our analytical results from Section 3.

Recall that we measure roundtrip delays of packets in the network rather than their unidirectional latencies. A common methodology in measurement studies is to estimate the one-way network delay by dividing in half the measured roundtrip delay. However, as in the wide-area study presented in [8], our measurements suggest that a roundtrip delay measurement is unlikely to be the sum of two unidirectional delays of approximately the same magnitude. Thus we use the measured roundtrip delays as estimates of the unidirectional network delay from the source machine to the remote machine.

Since the IVS package does not explicitly mark talkspurt boundaries, we define a new talkspurt to begin whenever there are no packet transmissions from the traffic-generating workstation for a period of greater than 80 ms duration. (The resulting talkspurts in the experimental data have a mean duration of 400 ms.) We estimate the retransmission time as twice the average network delay of the four packets arriving after the last packet that is dropped.

Figure 9 depicts the *k-error* curves resulting from our construction. The *0-error* curve in Figure 9 converges slowly towards continuous playback for all talkspurts. A control time of approximately 100 ms will achieve continuous playback at the receiver for over 95% of the talkspurts in the data, while a 120 ms control time is required for continuous playback of over 98% of the talkspurts.

Note that the curves in Figure 9 have the same characteristic shape of the analytical curves in Section 3. The "knee" of the curve for the probability of successful recovery of k lost packets is located at the control time value of k times the packetization interval. Beyond this critical value, the curves rise very rapidly and then converge slowly. The rapid rise reflects the large amount of traffic

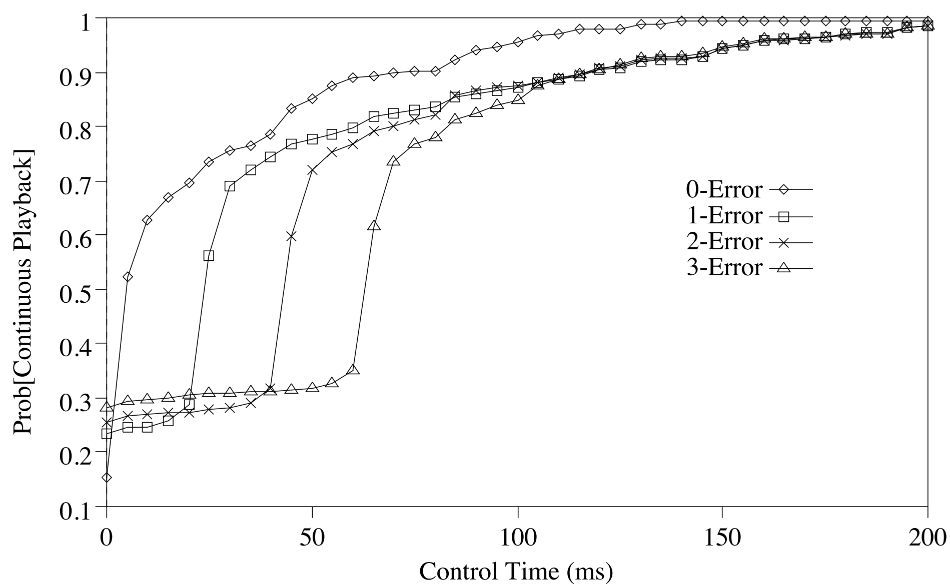

Figure 9: Extrapolated Retransmission Probabilities for Variable Control Times.

that experiences low delays in our measurements. The slow convergence is due to the difficulty of retransmission-based loss recovery during periods of high delays or high delay variations, e.g., the data shown in Figure 8.

The curves depicted in Figure 9 clearly show that, under the assumption that all packets are equally likely to be lost, a retransmission-based approach provides a high probability of timely error recovery. For example, a control time of 100 ms allows for recovery of over 80% of burst losses consisting of three or fewer consecutive packets.

5 Conclusions

In this paper we have investigated the feasibility of performing retransmission-based error recovery for continuous media traffic in packet-switching networks. We developed an analytical end-to-end model of voice transmission and provided an exact transient analysis to quantify the effectiveness of retransmission of voice packets. We presented an empirical study of voice transmission over a contemporary campus-wide network. The empirical measurements showed that retransmission is a viable option in many packet-switching networks.

Our results conclusively demonstrate the applicability of retransmission-based error recovery to packet voice streams in packet-switching networks. The control times required at the receiver to achieve successful retransmission with a high probability have been shown both analytically and empirically to be within the range of feasible control times for interactive voice streams (< 100 ms).

The control time necessary in order to compensate fully for the delay jitter in the network generally provides for a high probability of successful retransmission for single-packet losses. However, our studies indicate that increasing the control time by only a small amount can significantly increase the probability of successful retransmission.

Our measurements indicate that it is sufficient for packet voice protocols to use a fixed control time at the receiver, rather than adapting control times to fluctuations in the network delay. In our backbone network an adaptive algorithm for estimating the control time is of limited utility since the control times required for successful retransmission during periods of high network delay often exceed the end-to-end delay bound for interactive voice. Since the majority of packets experiences low delay, a conservative fixed control time yields a high probability of successful retransmission, without introducing the overhead associated with adaptive protocols.

Acknowledgment

The authors thank Matthew Lucas for his help in collecting the empirical voice measurements.

References

[1] G. Barberis. Buffer Sizing of a Packet-Voice Receiver. *IEEE Transactions on Communications*, COM-29(2):152–156, February 1981.

[2] G. Barberis and D. Pazzaglia. Analysis and Design of a Packet-Voice Receiver. *IEEE Transactions on Communications*, COM-28(2):217–227, February 1980.

[3] E. Biersack. Performance Evaluation of Forward Error Correction in ATM Networks. *ACM SIGCOMM '92*, 22(4):248–258, August 1992.

[4] P. T. Brady. A Technique for Investigating On-Off Patterns of Speech. *Bell System Technical Journal*, 44(1):1–22, January 1965.

[5] P. T. Brady. Effects of Transmission Delay on Conversational Behavior on Echo-Free Telephone Circuits. *Bell System Technical Journal*, 50(1):115–134, January 1971.

[6] CCITT. Recommendation G.711 - Pulse Code Modulation of Frequencies, October 1984.

[7] CCITT. Recommendation G.727 - 5-, 4-, 3-, 2-Bits/Sample Embedded ADPCM, July 1990.

[8] K. Claffy, G. Polyzos, and H. Braun. Measurement Considerations for Assessing Unidirectional Latencies. *Internetworking: Research and Experience*, 4:121–132, 1993.

[9] B. Dempsey. *Retransmission-Based Error Control for Continuous Media Traffic in Packet-Switched Networks.* PhD thesis, Department of Computer Science, University of Virginia, May 1994.

[10] M.W. Garrett and M. Vetterli. Joint Source/Channel Coding of Statistically Multiplexed Real-Time Services on Packet Networks. *IEEE/ACM Transactions on Networking*, 1(1):71–81, February 1993.

[11] N. Jayant. High Quality Networking of Audio-Visual Information. *IEEE Communications*, 31(9):84–95, September 1993.

[12] G. Karlsson and M. Vetterli. Packet Video and Its Integration into the Network Architecture. *IEEE Journal on Selected Areas in Communications*, 7(3):739–751, June 1989.

[13] E. Klemmer. Subjective Evaluation of Transmission Delay in Telephone Conversations. *Bell System Technical Journal*, 46:1141–1147, July 1967.

[14] J. McNabb. WIRETAP User's Manual, 1993. Internal Report, Department of Computer Science, University of Virginia.

[15] W.A. Montgomery. Techniques for Packet Voice Synchronization. *IEEE Journal on Selected Areas in Communications*, SAC-1(6):1022–1028, December 1983.

[16] W.E. Naylor and L. Kleinrock. Stream Traffic Communication in Packet-Switched Networks: Destination Buffering Considerations. *IEEE Transactions on Communications*, COM-30(12):2527–2534, December 1982.

[17] H. Ohta and T. Kitami. A Cell Loss Recovery Method using FEC in ATM Networks. *IEEE Journal on Selected Areas in Communications*, 9(9):1471–1483, December 1991.

[18] A. Sanghi, A. Agrawala, and B. Jain. Experimental Assessment of End-to-End Behavior on the Internet. *IEEE INFOCOM '93*, pages 867–874, March 1993.

[19] H. Schulzrinne. Voice Communication Across the Internet: A Network Voice Terminal. Technical report, University of Massachusetts, July 1992.

[20] K. Sriram and W. Whitt. Traffic Smoothing Effects of Bit Dropping in a Packet Multiplexer. *IEEE Transactions on Communications*, 37(7):703–712, July 1989.

[21] J. Suzuki and M. Taka. Missing Packet Recovery Techniques for Low-Bit-Rate Coded Speech. *IEEE Journal on Selected Areas in Communications*, 7(5):707–717, June 1989.

[22] T. Turletti. H.261 Software Codec for Videoconferencing over the Internet. Technical Report 1834, Institut National de Recherche en Informatique et en Automatique (INRIA), January 1993.

[23] N. Yin. Congestion Control for Packet Voice by Selective Packet Discarding. *IEEE Transactions on Communications*, 38(5):674–683, May 1990.