Decreasing Distortion Using Low Delay Codes For Bursty Packet Loss Channels

Emin Martinian, Carl-Erik W. Sundberg

Email: emin@allegro.mit.edu, sundberg@ibiquity.com

Abstract

The strict delay constraints of real-time communication applications in packet networks limit the use of ARQ (retransmission systems) and error correction codes with extensive interleaving and decoding over long intervals. Since packet losses can introduce significant impairments, we study the effectiveness of low delay channel coding techniques to increase transmission quality across links with bursty losses.

Specifically, we consider the benefits of the newly discovered class of low delay convolutional codes known as maximally short codes. By analyzing a Gaussian source transmitted over a Gilbert-Elliott Channel, we demonstrate that these codes can achieve significant gains in comparison to uncoded transmission schemes or traditional coded schemes employing Reed-Solomon block codes. To complement and validate the theoretical analysis we also present results from informal listening tests with a Voice over IP application.

Keywords

VoIP, low delay channel coding, erasure channels, packet loss, multidescriptive coding, convolutional codes.

I. Introduction

Recently there has been a great deal of interest in real-time communication applications for packet switched networks such as the Internet. In contrast to traditional data services such as email and File-Transfer-Protocol (FTP), the strict delay requirements for applications such as Voice over IP

(VoIP), video-conferencing, tele-medicine, etc. limit the use of ARQ (retransmission systems) and forward error correction (FEC) codes with extensive interleaving and decoding over long intervals. In addition, while most FEC codes are designed for memoryless channels, recent work suggests that packet losses in the Internet tend to occur in bursts [6]–[13].

Since packet losses due to congestion, errors, or other transmission problems can introduce significant distortions, we study the effectiveness of channel coding techniques to increase the quality of transmission across links with bursty losses. Long pauses in interactive communication are unacceptable, therefore a key requirement of these systems is low delay. A new class of short, low delay, convolutional codes are presented in [18], [19]. These codes have the property that for a burst of packet losses of a certain length, the required decoding delay and guard interval in terms of received packets are the shortest possible.

A great deal of previous work considers FEC based techniques for the transmission of data or multimedia over networks especially in multicast applications (e.g. see [14]–[17] and references therein for some recent examples). However, most of this work except for [1]–[6] does not consider decoding delay as an important but limited resource.

Prototype systems such as [1] and [2] provide a way to empirically investigate a host of practical issues. Using one of these prototypes, Bolot et al. [6] consider a rate allocation problem which allocates bits between an IETF standardized FEC scheme and source coding. Although the IETF FEC scheme works by repeating a coarse version of packet i in later packets as opposed to compressing the data stream and adding true erasure correction capability, in [6] the authors report that adding FEC provides significant benefits.

Altman et al. [4], [3] study the benefits of the IETF scheme for packets sent over an $M/M/1/K$ queue where performance is measured according to the total volume of data received. In contrast to [6], Altman et al. conclude that adding redundancy with this type of FEC always reduces overall quality.

Our contribution is to show that the average distortion in transmitting a Gaussian source over a Gilbert-Elliott Channel can be significantly reduced using maximally short codes. In Section II we introduce some notation and briefly review the new FEC code constructions in [18], [19]. In Section III we study the mean square error distortion reductions possible when a sequence of independent, identically distributed (i.i.d.) Gaussian random variables are transmitted over a Gilbert-Elliott Channel. Specifically, we show that maximally short codes outperform both uncoded systems and systems using conventional Reed-Solomon block codes. In trading off the level of source coding versus FEC coding, we consider a rate allocation problem analogous to [6] and therefore do not increase the size or number of packets transmitted in contrast to [4], [3]. To complement and validate the theoretical predictions, we present results from informal listening tests in Section IV. We close with some concluding remarks in Section V.

II. Code Constructions

A rate $R = k/n$ packet code is a mapping from a sequence of k-unit symbols, $\vec{x}[i] = (x_0[i], x_1[i], ..., x_{k-1}[i])$, to n-unit symbols, $\vec{y}[i] = (y_0[i], y_1[i], ..., y_{n-1}[i])$. A burst of length B starting at time i is defined as the erasure (*i.e.* loss) of one or more symbols from the set $\{\vec{y}[i], \vec{y}[i+1], ..., \vec{y}[i+B-1]\}$. If the earliest that $\vec{x}[i]$ can be recovered due to such a burst is once $\vec{y}[i+B]$, $\vec{y}[i+B+1]$, ..., $\vec{y}[i+T]$ are received then the decoding delay for recovering $\vec{x}[i]$ is T.

Such a code can be used to transmit a stream of n-unit data symbols (*i.e.* source packets), $\vec{u}[i] = (u_0[i], u_1[i], ..., u_{n-1}[i])$, over a packet switched network as follows. The original data is compressed via a lossy source code into k-unit symbols, $\vec{x}[i] = (x_0[i], x_1[i], ..., x_{k-1}[i])$, creating room for $n-k$ redundant units per packet. The k-unit compressed data symbols are then encoded to form a stream of n-unit coded symbols which are transmitted with 1 symbol per packet. Note that neither the packet size or the number of packets is changed in this coding scheme thus facilitating a fair comparison between the performance of coded and uncoded systems.

Ideally the code rate, R, should be near unity so that the original data does not have to be excessively compressed. Intuitively, though, lower rate codes have more redundancy and therefore provide more robustness to packet losses, lower delay in recovering from lost packets, or both. Using simple extensions of guard space bounds in [20], it is possible to show that if a rate R packet code can correct all erasure bursts of length B with decoding delay at most T, it must satisfy [19]

$$T/B \geq \max[1, R/(1-R)]. \tag{1}$$

To construct codes achieving this bound, let $P\{u_1, u_2, ..., u_{ms+s}\}$ be the s parity check symbols for a systematic, $(n, k, d) = (2s + ms, ms + s, s + 1)$ Reed-Solomon block code [21] with input $(u_1, u_2, ..., u_{ms+s})$. Then $\mathcal{C}_{m,s}$ is the rate $R = (ms + 1)/(ms + 1 + s)$ code defined by the mapping

$$\begin{aligned}\vec{y}[i] = (x_0^{ms}[i], P\{&x_0[i-1], x_0[i-2], ..., x_0[i-s],\\ &x_1^{s}[i-s-1], x_{s+1}^{2s}[i-2s-1], ..., x_{(m-1)s+1}^{ms}[i-ms-1]\})\end{aligned} \tag{2}$$

where $x_a^b[j]$ denotes $(x_a[j], x_{a+1}[j], ..., x_b[j])$. Such codes inherit linearity from the constituent Reed-Solomon codes and are therefore time-invariant convolutional codes [22]. In [19] we show that the decoding delay required by this family of codes to correct a burst of length s is exactly $T = ms + 1$ thus meeting the decoding delay bound with equality. Furthermore, via periodic interleaving of degree λ, the code $\mathcal{C}_{m,s}$ can be transformed into a code capable of correcting a burst of length λs with delay λT (and therefore continuing to meet the delay bound with equality).

A. Example Codes

To illustrate how the encoding rule in (2) yields a maximally short code we consider the rate 3/5 code $\mathcal{C}_{1,2}$ constructed in Table I. Imagine that $\vec{y}[i]$ and $\vec{y}[i+1]$ are lost in a burst of length 2. When $\vec{y}[i+2]$ is received, the decoder has $x_1^2[i-1]$ (since it was not lost in the burst) and $P\{x_0[i+1], x_0[i+0], x_1^2[i-1]\}$ (the 2 parity check symbols for the $(6,4,3)$ Reed-Solomon block code with input $x_0[i+1]$, $x_0[i]$, $x_1^2[i-1]$). Therefore the decoder recovers $x_0[i]$ and $x_0[i+1]$ at time $i+2$. When $\vec{y}[i+3]$ is received, the decoder has received $x_0[i+2]$ (since it was not lost in the burst), $x_0[i+1]$ (which was recovered in the previous decoding step), and $P\{x_0[i+2], x_0[i+1], x_1^2[i]\}$ (the 2 parity check symbols for the $(6,4,3)$ code with input $x_0[i+2]$, $x_0[i+1]$, $x_1^2[i]$). Therefore the decoder recovers $x_1^2[i]$ at time $i+3$. A similar argument shows that the decoder recovers $x_1^2[i+1]$ at time $i+4$. Hence the code corrects any burst of length 2 with decoding delay 3 meeting the bound in (1) with equality.

TABLE I
ENCODING EXAMPLE FOR THE CODE $\mathcal{C}_{1,2}$.

$$\begin{aligned}
\vec{y}[i-1] &= \left(x_0^2[i-1], P\{x_0[i-2], x_0[i-3], x_1^2[i-4]\}\right)\\
\vec{y}[i+0] &= \left(x_0^2[i+0], P\{x_0[i-1], x_0[i-2], x_1^2[i-3]\}\right)\\
\vec{y}[i+1] &= \left(x_0^2[i+1], P\{x_0[i+0], x_0[i-1], x_1^2[i-2]\}\right)\\
\vec{y}[i+2] &= \left(x_0^2[i+2], P\{x_0[i+1], x_0[i+0], x_1^2[i-1]\}\right)\\
\vec{y}[i+3] &= \left(x_0^2[i+3], P\{x_0[i+2], x_0[i+1], x_1^2[i+0]\}\right)\\
\vec{y}[i+4] &= \left(x_0^2[i+4], P\{x_0[i+3], x_0[i+2], x_1^2[i+1]\}\right)
\end{aligned}$$

For rates of the form $R = k/(k+1)$ (*i.e.* $m = k-1$ and $s = 1$), the maximally short codes developed in [19] take on the particularly simple structure of single parity check convolutional codes [18]. In this special case, the encoding rule in (2) used with periodic interleaving of degree λ can be reduced to

$$\vec{y}[i] = \left(x_0^{ms}[i], \sum_{j=0}^{k-1} x_j[i - \lambda(j+1)] \right) \tag{3}$$

where additions are carried out in the appropriate finite field.

For example, Table II illustrates the rate 2/3 single parity check convolutional code $\mathcal{C}_{1,1}$ with periodic interleaving of degree 2. If $\vec{y}[i]$ and $\vec{y}[i+1]$ are lost, successful decoding is accomplished via

$$x_0[i] = y_2[i+2] \ominus x_1[i-2] \quad (4)$$

$$x_0[i+1] = y_2[i+3] \ominus x_1[i-1] \quad (5)$$

$$x_1[i] = y_2[i+4] \ominus x_0[i+2] \quad (6)$$

$$x_1[i+1] = y_2[i+5] \ominus x_0[i+3], \quad (7)$$

thus demonstrating that this code can recover from a burst of 2 lost packets with decoding delay 4 meeting (1) with equality.

TABLE II
ENCODING EXAMPLE FOR THE CODE $\mathcal{C}_{1,1}$ WITH INTERLEAVING OF DEGREE 2.

$$\vec{y}[i-2] = \left(x_0^1[i-2], x_0[i-4] \oplus x_1[i-6]\right)$$
$$\vec{y}[i-1] = \left(x_0^1[i-1], x_0[i-3] \oplus x_1[i-5]\right)$$
$$\vec{y}[i+0] = \left(x_0^1[i+0], x_0[i-2] \oplus x_1[i-4]\right)$$
$$\vec{y}[i+1] = \left(x_0^1[i+1], x_0[i-1] \oplus x_1[i-3]\right)$$
$$\vec{y}[i+2] = \left(x_0^1[i+2], x_0[i+0] \oplus x_1[i-2]\right)$$
$$\vec{y}[i+3] = \left(x_0^1[i+3], x_0[i+1] \oplus x_1[i-1]\right)$$
$$\vec{y}[i+4] = \left(x_0^1[i+4], x_0[i+2] \oplus x_1[i+0]\right)$$
$$\vec{y}[i+5] = \left(x_0^1[i+5], x_0[i+3] \oplus x_1[i+1]\right)$$

B. Comparison to MDS Block Codes

We can gain some insight into the advantages of maximally short codes obtained via (2) by comparing them to maximum distance separable (MDS) block codes (*e.g.* Reed-Solomon codes)

[21] and near-MDS block codes (*e.g.* Low Density Parity Check (LDPC) codes [14]). While the latter can be decoded quickly in the sense that they require fewer computations than traditional MDS codes, for our purposes, LDPC codes are essentially equivalent to Reed-Solomon codes. This is because an (n, k) LDPC block code requires that at least roughly k of the n transmitted symbols must be received before successful decoding can occur. Thus our comments regarding the decoding delay (in terms of packets received) for Reed-Solomon block codes apply equally well to LDPC codes.

An RS block code can be used to transmit a stream of n-unit data symbols over a packet switched network as follows. Original data blocks of n symbols are compressed into blocks of k symbols (without changing the symbol size) via a lossy source code. The ith group of k symbols, $x[i \cdot k]$, $x[i \cdot k + 1]$, ..., $x[i \cdot k + k - 1]$, is then encoded into the ith group of n symbols, $y[i \cdot n]$, $y[i \cdot n + 1]$, ..., $y[i \cdot n + n - 1]$, using an (n, k) RS code (again without changing the symbol size). Each coded symbol is then transmitted in a separate packet. As before neither the packet size nor the number of packets is changed relative to an uncoded system.

For example, in a rate 3/5 RS block code system, the ith group of 3 packets $x[3i]$, $x[3i+1]$, and $x[3i + 2]$ is encoded into $y[3i]$, $y[3i + 1]$, ..., $y[3i + 4]$ via a $(5, 3)$ RS block code. If the packets corresponding to $y[3i]$ and $y[3i+1]$ are lost in a burst of length 2, then the decoder must wait until $y[3i + 2]$, $y[3i + 3]$, and $y[3i + 4]$ are received before the lost packets can be recovered resulting in a decoding delay of 4 to recover $y[3i]$. By contrast the rate 3/5 maximally short code $\mathcal{C}_{1,2}$ in Table I can recover from any burst of length 2 with decoding delay 3. In general, a rate $R = k/n$ RS block code system designed to correct bursts of length $B = n - k$ can require decoding delay $n - 1$ while a maximally short code based system can be constructed to correct all bursts of length $B = n - k$ with decoding delay k provided R is of the form $(ms + 1)/(ms + s + 1)$. Thus the worst case decoding delay for a Reed-Solomon code is $n - k - 1$ packets longer than for the corresponding maximally short code.

III. Overall Distortion And Coding Gains

For a fixed capacity link, adding diversity via the codes discussed in Section II, decreases the bits available for source coding. For a channel with many erasures, however, we expect that the decrease in source coding rate is compensated by the increased reliability introduced by the channel code. On the other hand, if the channel were perfect and no erasures occurred then little or no channel coding should be used. Thus the appropriate division of the link capacity between a channel code and a source code depends upon the severity of the channel as illustrated in Fig. 1 for some example channel coding rates. To more precisely analyze this hypothesis we investigate the average distortion obtained for various coded and uncoded transmission schemes for some simple source and channel models.

A. Source And Channel Models

Although various authors disagree on the nature and severity of packet losses in the Internet, a number of researchers have found empirical evidence that packet losses tend to occur in bursts [6]–[13]. Many researchers use a simple two state Markov chain (sometimes referred to as a Gilbert-Elliott Channel) to model bursty packet loss both for its analytical tractability and its reasonable agreement with experimental data. Consequently, in our analysis we consider packet losses to occur according to the Gilbert-Elliott Erasure Channel (GEEC) shown in Fig. 2. We conjecture qualitatively similar results would be obtained on more complicated models (e.g. an $M/M/1/K$ queue as in [4], [3]).

The GEEC model consists of two states: good and bad. A packet is never lost in the good state and always lost in the bad state. The probability of entering the bad state given that the last packet was not lost is α and the probability of entering the good state given that the last packet was lost is β. The quantity $\rho = (1-\beta)/\alpha$ captures the temporal correlation. A value of $\rho > 1$ corresponds to a bursty channel since losses are ρ times more likely to occur given that the last packet was lost instead of correctly received. A value of $\rho = 1$ corresponds to a Bernoulli channel since conditioning on whether the last packet was lost does not change the probability of the next packet being lost. In keeping with the simple channel model, we model the source as a zero mean, i.i.d., discrete-time, Gaussian process with unit variance.

B. Average Distortion

To investigate the advantages of coding, we study the average mean square error distortion reduction for a Gaussian source transmitted over a Gilbert-Elliott channel using various channel codes. As with our channel model, we conjecture that qualitatively similar results hold for a large class of reasonable source and distortion models. Intuitively, this is because the only feature essential to our analysis is the diminishing marginal returns usually present in quantization and compression. Specifically, according to [23], the minimum mean square error distortion in quantizing a unit variance Gaussian with R bits/sample is 2^{-2R}. Therefore in going from $R = 1$ to $R = 2$ bits/sample we reduce the distortion by $1/2 - 1/4 = 1/4$ while going from $R = 2$ to $R = 3$ bits/sample only reduces the distortion by $1/4 - 1/64 = 3/64$. Thus at some point allocating bits to FEC is more effective than allocating them to source coding.

If no coding is used the average distortion is

$$\mathcal{E}_{\mathrm{U}} = \Pr[\text{packet lost}] \cdot D(0) + (1 - \Pr[\text{packet lost}]) \cdot D(C) \tag{8}$$

where $D(C)$ is the distortion–rate function for a Gaussian source over a link of capacity C bits/sample. The average probability of packet loss is simply the steady state probability of being in the bad state of the channel, $\alpha/(\alpha+\beta)$. Combining this with the well-known fact that the distortion–rate function for a unit variance Gaussian is 2^{-2C} [23], we obtain

$$\mathcal{E}_{\mathrm{U}} = \frac{\alpha}{\alpha+\beta} + \left(1 - \frac{\alpha}{\alpha+\beta}\right) \cdot 2^{-2C}. \tag{9}$$

If a rate $R = (ms+1)/(ms+s+1)$ maximally short code is used, then only $R \cdot C$ bits/sample will be available for source coding. In this case the average distortion will be

$$\mathcal{E}_{\mathrm{MS}}(m,s,\lambda) = \mathrm{PLP}(m,s,\lambda) \cdot D(0) + [1 - \mathrm{PLP}(m,s,\lambda)] \cdot D\left(C \cdot \frac{ms+1}{ms+1+s}\right). \tag{10}$$

where $\mathrm{PLP}(m,s,\lambda)$ is the packet loss probability for the maximally short code $\mathcal{C}_{m,s}$ with periodic interleaving of degree λ derived in [19]. By plotting the average signal to distortion ratio (SDR) for an uncoded system $(1/\mathcal{E}_{\mathrm{U}})$ as well as for systems using maximally short codes $(1/\mathcal{E}_{\mathrm{MS}})$ in Fig. 3, we see that our analytical results confirm our intuitive expectations from Fig. 1.

Similarly, if an (n,k) systematic Reed-Solomon code is used as described in Section II-B, then the average distortion is

$$\mathcal{E}_{\mathrm{RS}}(n,k) = \mathrm{PLP}'(n,k) \cdot D(0) + [1 - \mathrm{PLP}'(n,k)] \cdot D\left(C \cdot \frac{k}{n}\right). \tag{11}$$

where $\mathrm{PLP}'(n,k)$, the average packet loss probability for the Reed-Solomon code, can be obtained through a straight-forward (though tedious) computation.

We can find the distortion for the best maximally short code, $\mathcal{E}^*_{\mathrm{MS}}$, by minimizing $\mathcal{E}_{\mathrm{MS}}(m,s,\lambda)$ subject to the the decoding delay constraint $\lambda(ms+1) \leq T$. Similarly, we can find the distortion for the best Reed-Solomon code, $\mathcal{E}^*_{\mathrm{RS}}$, by minimizing $\mathcal{E}_{\mathrm{RS}}(n,k)$ subject to the decoding delay constraint $n-1 \leq T$.

To investigate the advantages of maximally short codes and Reed-Solomon codes relative to uncoded transmission, we plot $\mathcal{E}_{\mathrm{U}}/\mathcal{E}^*_{\mathrm{MS}}$ and $\mathcal{E}_{\mathrm{U}}/\mathcal{E}^*_{\mathrm{RS}}$: the gain in signal to distortion ratio (SDR) for such codes relative to an uncoded system for loss rates and levels of burstiness roughly corresponding to those observed in the Internet (Fig. 4) [12] and in wireless networks (Fig. 5) [24]. As the plots show, the robustness added by channel coding can significantly reduce the distortion for the coded stream over a range of parameters.

For example, in Fig. 4, we see that significant gains in SDR are achievable via coding even at relatively short decoding delays. Furthermore we see that maximally short codes outperform Reed-Solomon codes when the decoding delay is less than 9. As the decoding delay increases, the maximally short codes still possess superior burst correcting capability, but the larger minimum distance of Reed-Solomon codes allows them to correct occasional non-bursty patterns of losses which defeat the maximally short codes.

Figure 5 shows that in channels with longer bursts the performance gains of maximally short codes are even more dramatic. When long enough delays are allowed, however, Reed-Solomon codes will again eventually perform better due to their larger minimum distance and consequent advantage in correcting non-bursty losses. In practice, this suggests that a hybrid of maximally short codes and codes with good minimum distance (*e.g.* Reed-Solomon codes) might achieve even better performance.

IV. Informal Mean Opinion Score Listening Tests

A. Purpose

In the previous sections we described constructions of various burst erasure correcting codes and theoretically analyzed their advantages in reducing distortion for an i.i.d. Gaussian source. While these simple models allow us to obtain precise analytical results, they do not necessarily demonstrate the effectiveness of such codes in practice. For example, issues such as non-rate-distortion optimal compression, error concealment, and the complicated nature of human perception of distortions need to be considered as well.

Therefore, in order to supplement the analytical results and demonstrate the effectiveness of these codes for increasing the quality of Voice over IP, we conducted informal listening tests for speech signals with some representative channel codes and source coders. We studied channels of varying burstiness with uncoded packet loss rates ranging from 0 to 17% with a delay constraint of 60 ms and discovered the benefits of low delay channel codes are noticeable even at uncoded loss rates as low as 4%. As expected, the benefits of these codes increase as the loss rate increases.

B. Hypothesis

Motivated by the theoretical results in Section III, we expect that for a reasonable set of parameters, decreasing the source coding rate in order to add redundancy for FEC coding will increase the overall quality relative to an uncoded system as illustrated in Fig. 1.

C. Procedure

We conducted informal listening tests using various levels of channel coding for an adaptive differential pulse code modulation source coder [25]. In each test, the speech coder used an erasure concealment algorithm to mitigate the effect of lost packets. We simulated two channels: the Gilbert-Elliott Erasure Channel (GEEC) discussed previously and a block loss channel.

The block loss channel model is as follows. Initially the channel is in the good state and the initial packet is not lost. Each time a packet is not lost there is a probability α that the next β packets will be lost. After the loss of β packets, the channel returns to the good state as illustrated in Fig. 6.

In each test, the total rate of the transmitted data is 64 kb/s. The original speech samples were quantized with 8 bit μ-law PCM, sampled at 8 kHz, and converted into 10 ms packets. The listening tests were conducted in 2 sessions over 2 days. For each channel condition and encoding format, the 4 listeners were presented with 4 stimuli, each lasting 7-9 seconds. The 4 stimuli consisted of 2 male and 2 female speakers each reciting 2 sentences in a normal voice without noise. In each session, the first 20% of the stimuli consisted of an unscored training phase. After the training phase, the listeners were asked to score the remaining samples according to the following range: (1) bad, (2) poor, (3) fair, (4) good, (5) excellent.

We used four different encoding formats: no channel coding, a rate 3/4 maximally short (MS) code, a rate 1/2 MS code, and a rate 1/2 MS code combined with a multidescriptive source coder. All of the coding formats, except for no channel coding, required a decoding delay of 6 packets which corresponds to 60 ms. Further details about the encoding formats are below:

1. **No Channel Coding:** In this format the entire 64 kb/s were allocated to the source coder. Unfortunately, the highest encoding rate supported by our encoder was 40 kb/s. Informal listening tests indicate that our 40 kb/s encoder performs almost as well as 8 bit μ-law PCM. Hence we used our 40 kb/s encoder as a proxy for a full 64 kb/s encoder. Despite this flaw we believe our results are still meaningful in understanding the qualitative behavior of channel coding in this scenario. Of course, to increase the credibility of these results we would redo our tests using a full 64 kb/s encoder for this format.

2. **Rate 3/4 MS Code With $\lambda = 2$ Interleaving:** In this format we allocated 48 kb/s to the source coder and used a rate 3/4 MS code with interleaving degree 2 as the channel code. As in the previous format, we used our 40 kb/s encoder as a proxy for the proposed 48 kb/s encoder.

3. **Rate 1/2 MS Code With $\lambda = 6$ Interleaving:** In this format we allocated 32 kb/s to the source coder and used a rate 1/2 MS code (*i.e.* repetition coding) with interleaving degree 6 as the channel code.

4. **Rate 1/2 Multiple Descriptions Delay & Duplicate Code With $\lambda = 6$ Interleaving:** In this format we described the source with a pair of 32 kb/s descriptions [25], transmitted in packets i and $i+6$. If either the description in packet i or packet $i+6$ is lost then a coarse reconstruction can be recovered from the remaining description. If neither packet is lost, then the two descriptions can be combined for a high quality description, [25].

D. Results

Tables III-VIII show the results of the listening tests for various channel conditions from mildest to most severe. Tables IV, V, VI, and VIII were listened to in the first session and Tables III and VII were listened to in the second session. The first column in each table indicates the encoding format, the second column indicates the packet loss rate, the third column is the mean opinion score (MOS) and the fourth column is the standard deviation for the MOS.

These results confirm the behavior we predicted in Fig. 1. In the case where no packet loss occurs, Table IV shows that no channel coding produces the highest quality signal followed by the rate 3/4 code and rate 1/2 code. The multiple descriptions code also performs quite well since when no packet loss occurs, both descriptions can be combined for a high quality reproduction.

As the packet loss rate is increased in Table IV, the performance of the uncoded stream drops precipitously while the rate 3/4 and rate 1/2 codes are mostly unaffected. The performance of the multidescriptive code also drops below the channel coded streams but not as far as the uncoded stream since the multidescriptive code still contains some redundancy to resist packet loss. This trend continues in Table V.

In Table VI, the packet loss rate increases enough that the rate 1/2 coded stream has the highest quality. This trend continues in Tables VII and VIII. In addition, as the channel becomes more sever in Table VIII, the multidescriptive code does better than the rate 3/4 code.

TABLE III

Results for listening test with no erasures.

Encoding Format:	Loss Rate	MOS	σ
No Coding	0	3.75	0.75
Rate 3/4 MS Code	0	3.623	0.78
Rate 1/2 MS Code	0	3.19	0.53
Multiple Descriptions	0	3.69	0.46

TABLE IV

Results for listening test on GEEC with $\alpha = 0.05, \beta = 0.8$

Encoding Format:	Loss Rate	MOS	σ
No Coding	0.039	2.38	0.60
Rate 3/4 MS Code	0.004	3.56	0.70
Rate 1/2 MS Code	0.001	3.25	0.66
Multiple Descriptions	N/A	2.69	0.46

TABLE V

Results for listening test on GEEC with $\alpha = 0.075, \beta = 0.85$

Encoding Format:	Loss Rate	MOS	σ
No Coding	0.065	2.06	0.56
Rate 3/4 MS Code	0.016	3.50	0.71
Rate 1/2 MS Code	0.005	3.19	0.73
Multiple Descriptions	N/A	2.56	0.70

TABLE VI

Results for listening test on Block Loss Channel with $\alpha = 0.05, \beta = 2$.

Encoding Format:	Loss Rate	MOS	σ
No Coding	0.065	1.56	0.61
Rate 3/4 MS Code	0.011	2.88	0.78
Rate 1/2 MS Code	0	3.06	0.43
Multiple Descriptions	N/A	2.88	0.60

TABLE VII

Results for listening test on GEEC with $\alpha = 0.125, \beta = 0.75$

Encoding Format:	Loss Rate	MOS	σ
No Coding	0.105	1.44	0.61
Rate 3/4 MS Code	0.042	2.44	0.61
Rate 1/2 MS Code	0.011	2.88	0.70
Multiple Descriptions	N/A	2.19	0.53

TABLE VIII

Results for listening test on Block Loss Channel with $\alpha = 0.05, \beta = 6$

Encoding Format:	Loss Rate	MOS	σ
No Coding	0.167	1.06	0.24
Rate 3/4 MS Code	0.085	1.56	0.70
Rate 1/2 MS Code	0.007	2.75	0.66
Multiple Descriptions	N/A	2.38	0.70

E. Conclusion

The informal listening tests confirm our hypothesis that adding channel coding increases the signal quality on bursty channels. Furthermore our results show that more severe channels require more powerful channel codes. This suggests two possible strategies. When the channel is known, the appropriate level of channel coding can be chosen a priori and fixed. When the channel is time-varying, an adaptive multi-rate scheme similar to those in [26], [27] can be used to continuously choose the appropriate code.

Another interesting result of the experiment is the behavior of the multidescriptive code. Ideally such a code would perform almost as well as the uncoded stream when no packet loss. An ideal multidescriptive code would also perform well when packet loss occurs since each description can serve as a coarse representation of the source. The former goal is satisfied as shown in Table III. However as shown in the following tables, when packet loss occurs, the reconstruction from a single descriptions is significantly degraded. Apparently the multidescriptive code does not perform as well as a higher rate channel code. Effectively the multidescriptive code corresponds to more channel coding and less source coding than the rate 3/4 code but less channel coding and slightly more source coding than the rate 1/2 code.

V. Conclusions

In this paper we studied the benefits of various low delay burst erasure correcting codes. After reviewing the maximally short codes in [18], [19], we described how these codes could be used to decrease the average distortion in transmitting multimedia across links with bursty losses. We presented theoretical calculations of expected distortion for the simple model of a Gaussian source transmitted over a Gilbert-Elliott Channel. The calculations predicted that maximally short codes could achieve significant gains in the signal to distortion ratio relative to both uncoded schemes and systems employing Reed-Solomon block codes. Finally, the theoretical predictions regarding the value of coding were verified by performing informal listening tests using various channel codes with an adaptive differential pulse code speech coder employing error concealment.

Since multimedia communication over packet networks is still an emerging field, there exists a wide array of possibilities for future work. Some important issues include developing and analyzing more accurate models for packet losses, investigating alternative code constructions, and characterizing the interaction of data compression and error concealment schemes with burst correcting codes. One particularly interesting area of investigation includes multi-mode decoders for hybrid codes which can correct dense bursts of erasures with low delay and potentially correct random (*i.e.* non-bursty patterns) of erasures or even errors when longer delays are allowed.

Acknowledgments

The authors would like to thank the members of the Speech and Audio Communications group at Bell Labs, Lucent Technologies for their help, encouragement and advice. In particular we would like to thank Peter Kroon, Herve Taddei, Sean Ramprashad, Frank Baumgarte, and Yair Ghitza for their aid in designing and conducting listening tests. We would also like to thank Dr. Nikolaos Sidiropoulos for directing us to [1]–[4] and providing other helpful comments. Finally, we are indebted to the anonymous reviewers for their careful reading of the initial manuscript and many valuable suggestions.

References

[1] “Rat: Robust audio tool,” http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/, Multimedia Integrated Conferencing for Eurpopean Researches, University College London.

[2] Andres Vega Garcia and Sacha Fosse-Parisis, “Freephone audio tool,” http://www-sop.inria.fr/rodeo/fphone/, High-Speed Networking Group, INRIA Sophia Antipolis.

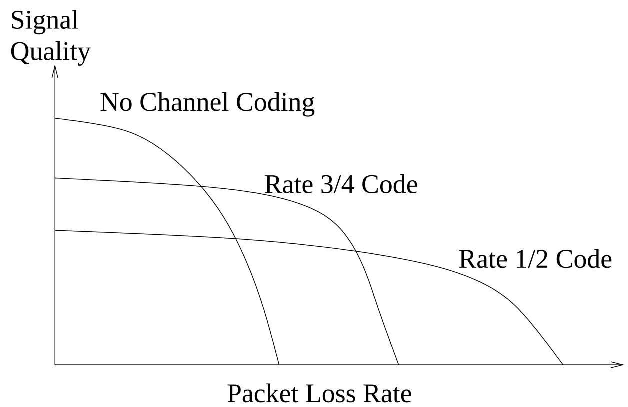

Fig. 1. Conceptual plot of the signal quality we expect for various levels of channel coding as a function of packet loss rate. We expect that for mild channels little or no coding will produce the highest quality signal at the receiver. As the packet losses become more pronounced, however, we expect that the added robustness of coded streams will produce better quality signals at the receiver.

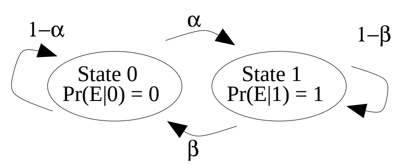

Fig. 2. Diagram of the Gilbert-Elliott Erasure Channel (GEEC). Erasures always occur in the bad state (state 1), and never occur in the good state, (state 0).

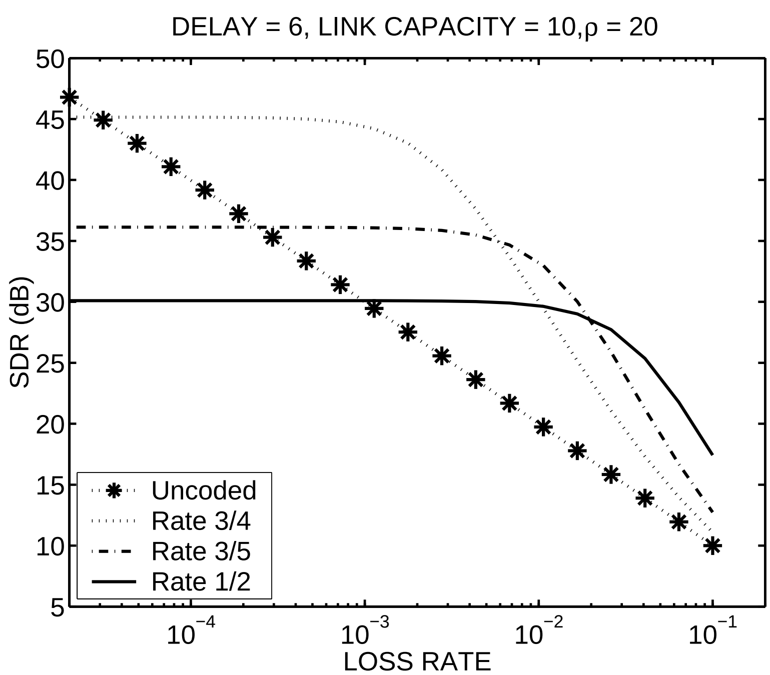

Fig. 3. The figure shows the average signal to distortion ratio for an uncoded system and systems using MS codes of various rates on a Gilbert-Elliott channel where packet losses are 20 times more likely given that the previous packet was lost. The total bit rate is 10 bits/sample and a decoding delay of 6 packets is allowed for the various coded systems.

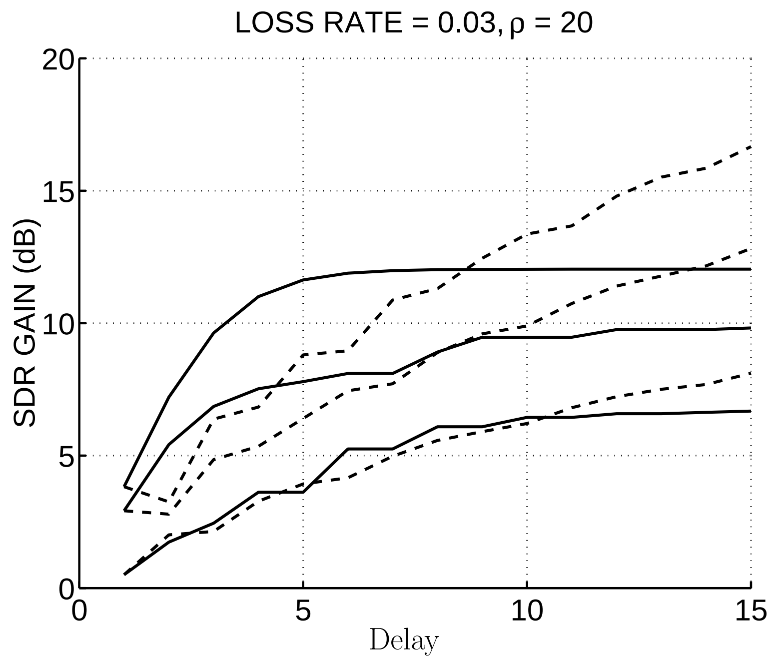

Fig. 4. This figure shows the increase in signal to distortion ratio (SDR) relative to an uncoded system when transmitting a Gaussian source over a Gilbert-Elliott Channel. The solid lines represent the performance for maximally short codes at link capacities of 10, 8, and 6 bits/sample from top to bottom, while the broken lines represent the performance for Reed-Solomon block codes at link capacities of 10, 8, and 6 bits/sample from top to bottom. The uncoded packet loss probability is .03 and a packet loss is 20 times more likely given that the last packet was lost.

[3] E. Altman, C. Barakat, and V.M. Ramos, "Queueing analysis of simple fec schemes for ip telephony," in *INFOCOM 2001*, 2001, vol. 2, pp. 796–804.

[4] Eitan Altman and Alain Jean-Marie, "Loss probabilities for messages with redundant packets feeding a finite buffer," *Journal of Selected Areas in Communications*, vol. 16, no. 5, pp. 778–787, June 1998.

[5] G. M. Schuster et al, "Forward error correction system for packet based real time media," US Patent 5,870,412.

[6] J. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive fec-based error control for internet telephony," in *Proc. of Infocom'99*, March 1999.

[7] J. Bolot, "Characterizing end-to-end packet delay and loss in the internet," *Journal of High Speed Networks*, pp. 289–298, September 1993.

[8] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and modelling of the temporal dependence in packet loss," in *Proceedings of The Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '99)*, 1999, pp. 345–352.

[9] M. S. Borella, D. Swider, S. Uludag, and G. B. Brewster, "Internet packet loss: measurement and implications

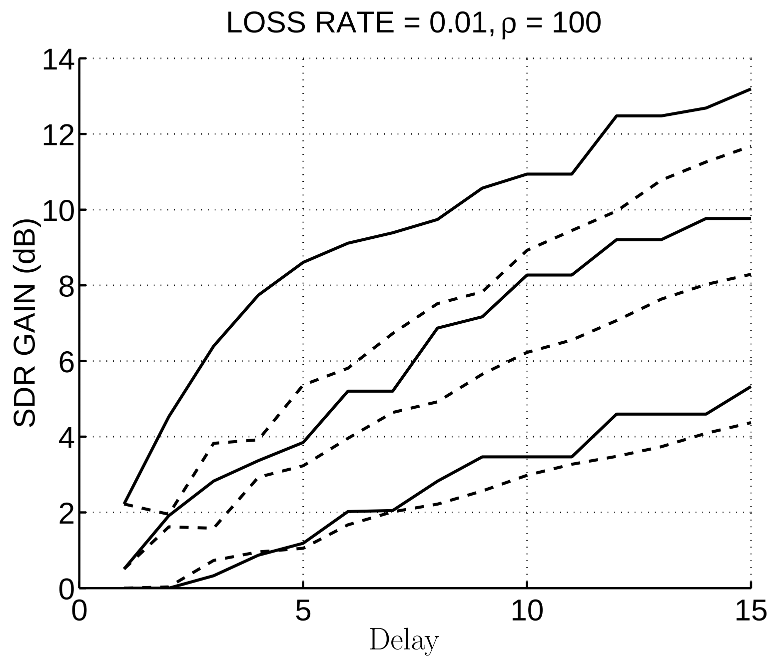

Fig. 5. This figure shows the increase in signal to distortion ratio (SDR) relative to an uncoded system when transmitting a Gaussian source over a Gilbert-Elliott Channel. The solid lines represent the performance for maximally short codes at link capacities of 10, 8, and 6 bits/sample from top to bottom, while the broken lines represent the performance for Reed-Solomon block codes at link capacities of 10, 8, and 6 bits/sample from top to bottom. The uncoded packet loss probability is .01 and a packet loss is 100 times more likely given that the last packet was lost.

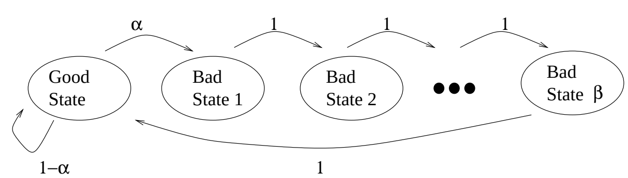

Fig. 6. Diagram of the block loss Channel. Erasures never occur in the good state and always occur in the bad states. Thus erasures always occur in bursts of β packets.

for end-to-end qos," in *Proceedings of the 1998 ICPP Workshops on Architectural and OS Support for Multimedia Applications/Flexible Communications Systems/Wireless Networks and Mobile Computing*, 1998, pp. 3–12.

[10] Wenyu Jiang and Henning Schulzrinne, "Qos measurement of internet real-time multimedia services," Tech. Rep., Columbia University, Decemeber 1999.

[11] M. Arai, A. Chiba, and K. Iwasaki, "Measrurement and modeling of burst packet losses in internet end-to-end communications," in *Proceedings of the 1999 Pacific Rim International Symposium on Dependable Computing*, 1999, pp. 260–267.

[12] Vern Paxson, "End-to-end internet packet dynamics," *IEEE Transactions On Networking*, vol. 7, no. 3, pp.

277–292, June 1999.

[13] H. Sanneck, G. Carle, and R. Koodli, "A framework model for packet loss metrics based on loss runlengths," in *Proc. of SPIE/ACM SIGMM Multimedia Computing and Networking Conference*, January 2000.

[14] J. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," in *Proc. ACM SIGCOMM 1998*, Vancouver Canada, September 1998, pp. 56–67.

[15] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1737–1744, Nov 1996.

[16] P.A. Chou, A.E. Mohr, A. Wang, and S. Mehrotra, "Error control for receiver-driven layered multicast of audio and video," *IEEE Transactions on Multimedia*, pp. 108–122, 2001.

[17] Wai-Tian Tan and A. Zakhor, "Video multicast using layered fec and scalable compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 373–386, March 2001.

[18] Emin Martinian and Carl-Erik W. Sundberg, "Low delay burst erasure correction codes," to appear in the *Proc. International Conference on Communications* 2002, 2002.

[19] Emin Martinian and Carl-Erik W. Sundberg, "Channel codes with low delay for burst erasure channels," submitted to *IEEE Transactions on Information Theory*.

[20] G. D. Forney Jr., "Burst-correcting codes for the classic bursty channel," *IEEE Transactions on Communications Technology*, vol. COM-19, no. 5, pp. 772–781, October 1971.

[21] S. Lin and D. J. Costello Jr., *Error Control Coding: Fundamentals and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1983.

[22] R. Johannesson and K. Sh. Zigangirov, *Fundamentals of Convolutional Coding*, IEEE Press, New York, NY, 1999.

[23] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley and Sons, Inc., 1991.

[24] Almudena Konrad, Ben Y. Zhao, Anthony D. Joseph, and Reiner Ludwig, "A markov-based channel model algorithm for wireless networks," in *Proceedings of Fourth ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2001)*, 2001.

[25] Cheng-Chieh Lee, "Diversity control among multiple coders: A simple approach to multiple descriptions," in *2000 IEEE Workshop On Speech Coding,* Proceedings, Delevan, Wisconsin, September 2000, pp. 69–71.

[26] T. Hindelang, M. Kaindl, J. Hagenauer, M. Schmatuz, and W. Xu, "Improved channel coding and estimation for adaptive multi rate (amr) speech transmission," in *IEEE 51st Vehicular Technology Conference*, Tokyo, Japan,

May 2000, vol. 2, pp. 1210–1214.

[27] D. J. Goodman and C-E. W. Sundberg, "Combined source and channel coding for matching the speech transmission rate to the quality of the channel," *Bell Syst. Tech. J.*, vol. 62, no. 7, pp. 2017–2036, September 1983.