Adaptive forward error correction for interactive streaming over the Internet

Fernando Silveira Filho   Edson H. Watanabe   Edmundo de Souza e Silva
Systems Engineering and Computer Science Department
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
{fernando, edson, edmundo}@land.ufrj.br

***Abstract*— Despite the emergence of voice over IP and video-conferencing services, recent studies have shown that the Internet still may not provide acceptable quality for interactive streaming applications. Among the major parameters that affect quality of service are packet loss rates and loss burst sizes. To mitigate the negative impact of losses, forward error correction (FEC) techniques can be employed. These work by adding extra information to the data stream. So, while FEC can potentially reduce the perceived effects of packet losses, it increases bandwidth requirements which in turn may increase the loss rate. Therefore, it is important to choose, in real time, the proper FEC scheme to provide the best performance to the application. We evaluate the performance of parity-based FEC schemes using an analytical loss model. Then, we develop an adaptive mechanism for FEC selection using a predictive model. We present the results of simulation experiments using real Internet measurements.**

I. Introduction

In the last few years, interactive multimedia services such as Voice over IP (VoIP) and video-conferencing, have changed from promising new applications to reality. The increasing demand for audio and video services in the Internet has spawned a number of commercial applications plus some very popular free tools such as Skype™, Google Talk™and Windows Live™Messenger. Nonetheless, some studies have shown that the current Internet infrastructure is not ready to provide acceptable quality to these applications [1], [2]. One-way delay, jitter and packet losses are the most consequential impairments to quality of service (QoS) in interactive streaming applications.

While jitter is usually mitigated through playout scheduling mechanisms [3], [4], there is a number of alternatives for dealing with the effects of packet losses [5]. Techniques for recovering from errors in a data stream are based in either *automatic repeat request* (ARQ) or *forward error correction* (FEC). Retransmission schemes based on ARQ introduce end-to-end delays that are generally not suited for interactive communications. Forward error correction is a more attractive alternative when delay constraints are stringent.

Forward error correction can be either media-specific or media-independent. The former involves replicating media units with a possibly lower quality codec, while the latter uses error correcting codes in order to produce additional bits in the data stream that can be used to recover lost packets.

Among media-independent FEC techniques, parity coding performs an exclusive-OR operation over a block of packets in order to produce an additional payload that can be used in case a single packet is lost within the protected block. In [6], the authors proposed the use of interleaved parity codes to recover lost packets in voice transmission over the Internet.

Since FEC adds overhead to the original data stream, it also has the potential to increase network load even when no correction is needed [7]. Therefore, there must be some sort of adaptive control mechanism in order to regulate the amount of redundancy used in each situation. The work in [8] developed an algorithm for FEC control based on estimates of the loss rate after reconstruction for each scheme in a set of available FEC settings.

The work in this paper is inspired by that of [8], albeit with some essential differences. First, that reference uses past packet loss measurements to infer a Gilbert model for the channel. In our approach, we use a hierarchical hidden Markov model that switches the network state between different Gilbert models and attempts to predict future channel statistics. Second, the mechanism described in [8] is used to control the amount of redundancy in a media-specific FEC scheme proposed by the IETF for packet audio [9]. In this paper, we use a media-independent parity based FEC scheme [6]. We also chose to focus our experiments on voice transmission, and we compare the performance of our approach to that of the technique in [8]. Despite this limitation of scope, our methodology is readily applicable to other sorts of traffic such as that of video-conferencing.

The rest of this paper is organized as follows. Section II reviews the parity based FEC technique proposed in [6] and develops an exact formula for evaluating the performance of these codes under the assumption of a Gilbert loss model. In section III our methodology for adaptive FEC control is presented and then, in section IV, we report experimental results on the performance of this approach. Finally, section V summarizes our contributions and outlines our future research directions.

II. Forward Error Correction

The work in [6] describes how interleaved parity FEC can be used to recover from packet losses. In that approach, a single parity block is generated from a chain of n packets. If k similar chains are interleaved, then k parity blocks are generated to protect a window of $w = kn$ packets. Following the terminology from that paper, we refer to this setting as a k:w scheme.

Figure 1 was taken from [6] and it shows an example of this algorithm. In the figure there are two interleaved chains in every window of six packets and thus it represents a 2:6 scheme. Each chain consists of three packets and it is protected by a single parity block, which is added to the payload of a packet in the next window.

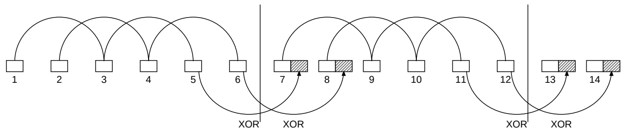

Fig. 1. Two interleaved chains of parity FEC with a window of six packets.

Clearly, the overhead of the k:w scheme is k parity blocks for every w packets, or $k/w = 1/n$. If $k = w$, then this scheme is equivalent to simply sending copies of the original data after k packet transmissions, with a unit overhead of 100%. It is important to notice that, in this particular case, it is often more convenient to compress the redundant data with a media-specific codec. If this can be done, then the overhead measured in bytes will be actually smaller than 100%, depending on the specific codec. However, this reduced overhead is achieved through a degradation in media quality.

Another parameter of importance to choose the appropriate scheme is the added delay that results from recovering a lost packet. For a k:w scheme, if a packet is lost, the receiver must wait, in the worst case, for w subsequent transmissions until the parity data arrives and the recovery can be performed.

A. *Performance Analysis of FEC Schemes*

In the following development, we assume that the packet loss process seen by an end receiver can be described as a 2-state Markov chain, also known in the literature as the Gilbert model. Several studies have been proposed in attempt to characterize the loss processes experienced by traffic flows in communication networks [10], [11], [12], and it has been argued that the Gilbert model is not able to emulate the behavior of loss traces with long-term correlations. Despite these shortcomings, the Gilbert model is commonly applied in performance evaluation studies [8], [13], due to its analytical simplicity and the good results it provides in practical applications.

Let X_t denote the t-th packet outcome, with $X_t = 1$ representing a packet loss and $X_t = 0$ a successful transmission. Under the assumption of a Gilbert model, we let $p = P(X_t = 1|X_{t-1} = 0)$ and $q = P(X_t = 0|X_{t-1} = 1)$, as illustrated in Figure 2. The steady state probabilities $\pi_0 = P(X_t = 0)$ and $\pi_1 = P(X_t = 1)$ are given, respectively, by $q/(p + q)$ and $p/(p + q)$.

Consider a single chain of packets protected by a parity payload in a k:w FEC scheme. Such a chain contains $n = w/k$ packets. In the long run, the loss fraction seen by a packet flow after the FEC reconstruction can be regarded as the expected

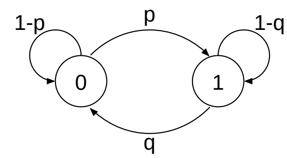

Fig. 2. Packet loss model used in this analysis.

number of packets in a given chain that can not be recovered, divided by n. We show how to evaluate this measure, by conditioning on whether the FEC payload is correctly received.

As illustrated in Figure 1, in the k:w scheme, the parity data is included k packets *after* the last protected packet in the chain. Since our packet loss model, the Gilbert model, is a time reversible stochastic process, we can assume, with no loss in generality, that the FEC payload is included k packets *before* the first protected packet in the chain. This simplification is used merely for the sake of analytical simplicity as, in the implementation, FEC is indeed transmitted after the original data. However, under the assumption of time reversibility, the results of both analyses must be the same.

Furthermore, since consecutive packets in a FEC chain are separated by $k - 1$ other packets, the analysis of losses in the chain makes use of the k-step transition probabilities of the Gilbert model. We let P_{ij} be transition probability from i to j in this k-step Gilbert model.

For conciseness of notation, let X_0 be the outcome of the packet that contains the FEC data and $X_k, X_{2k} \ldots, X_{nk}$ represent the outcomes of the n remaining packets in the chain, which are protected by the redundancy in X_0. Also, let L be the random variable representing the total number of losses in this chain of packets, i.e., $L = \sum_{i=1}^{n} X_{ik}$. Finally, we denote as L' the number of losses in the same chain that can not be recovered using the FEC payload. The expected value $E[L']$ can be computed by conditioning on the result of X_0:

$$E[L'] = E[L'|X_0 = 0]\pi_0 + E[L'|X_0 = 1]\pi_1. \qquad (1)$$

Whenever $X_0 = 1$, i.e., the FEC packet is lost, L' is simply the original number of losses in the chain, L. If, otherwise, FEC can be employed, then L' will be zero if the total number of losses in the chain is at most one. In case more than one packet is lost within a same chain, the FEC payload will be useless for recovery. With these considerations, we can write:

$$\begin{aligned} E[L'|X_0 = 1] &= E[L|X_0 = 1] \\ &= \textstyle\sum_{i=1}^{n} P_{11}^{(i)}, \qquad (2) \\ E[L'|X_0 = 0] &= \textstyle\sum_{j=2}^{n} jP(L = j|X_0 = 0) \\ &= E[L|X_0 = 0] - P(L = 1|X_0 = 0) \\ &= \textstyle\sum_{i=1}^{n} P_{01}^{(i)} - P(L = 1|X_0 = 0), \qquad (3) \end{aligned}$$

where $P_{01}^{(i)}$ and $P_{11}^{(i)}$ are transition probabilities in i steps for the k-step Gilbert model.

B. *FEC Scheme Diversity*

Figure 3 plots the efficiency of three different k:w schemes. Namely, we compare the schemes 1:1, 3:6, and 6:6. For each

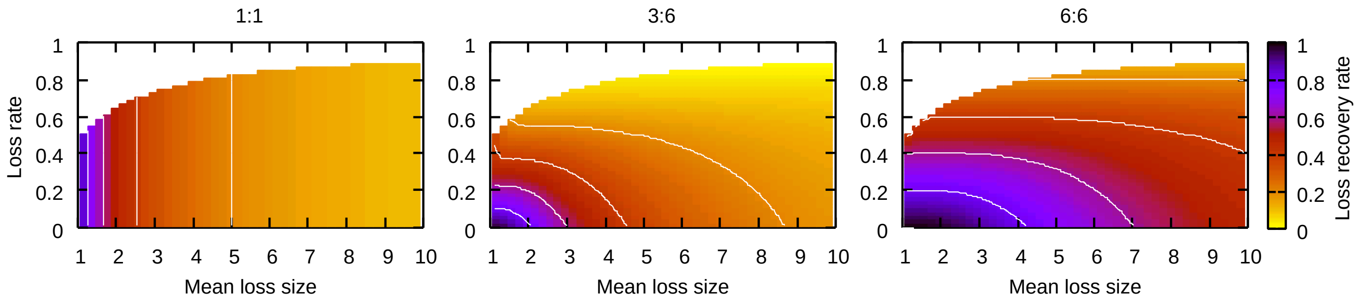

Fig. 3. Loss recovery efficiency for three different schemes.

scheme we consider a packet loss model as in Figure 2 and change its loss rate, r, and the mean loss burst size, b, across a sampled space. The values of r and b alone are sufficient to determine the parameters of a Gilbert model since $q = 1/b$ and $p = qr/(1-r)$. Given these parameters, we then evaluate the expected fraction of lost packets recovered by the scheme. Following the notation from the previous section this can be evaluated as $1 - E[L']/E[L]$. Each graphic also shows a contour plot of levels 0.2, 0.4, 0.6 and 0.8.

The left plot in Figure 3 shows that the 1:1 scheme is expected to recover more than 80% of the losses if the mean loss size does not exceed one packet. The 6:6 scheme, in the rightmost plot, is seen to provide a much better resistance to losses of larger sizes, even for considerable loss rates. Although both these schemes have the same overhead (100%), the delay incurred in recovering a packet with the 6:6 scheme is always much larger than that of the simpler 1:1 scheme. In the middle plot of Figure 3, the 3:6 scheme is capable of recovering moderate losses with a 50% overhead. The recovery delay in the 3:6 scheme is at most 6 transmissions.

One can learn from Figure 3 and the previous discussion that different FEC schemes may perform best at specific situations. For instance, there is no need to use the 6:6 scheme, unless one expects both high loss rates and large loss burst sizes. Similarly, if some additional delay may be tolerated, then the 3:6 scheme should provide a much better cost-benefit ratio than the 1:1 setting in the most common scenarios of packet loss.

III. Proposed Methodology

In this section we describe the two components of our adaptive FEC control mechanism. First, a hierarchical packet loss model is described that enables us to predict the parameters of a Gilbert model in the short-term future. Second, the adaptive FEC selection mechanism is proposed.

A. Predictive Packet Loss Model

We use a hidden Markov model (HMM) that contains a separate Gilbert model in each of its hidden states. The work in [11] proposed the use of HMMs to model packet loss events in communication networks. Each hidden state in a HMM represents a specific network condition, i.e., congestion level. In [11], each state is characterized by a single parameter: the loss fraction at that state. In our approach, each hidden state defines a Gilbert model, allowing for different loss rates and mean loss burst sizes.

In our model, transitions between hidden states may occur only at embedded points every S packet outcomes. We assume that, while the local packet statistics may be well-represented by a Gilbert model, the parameters of this model may change over time, at a slower time scale, governed by a hidden Markov chain. We refer to our model as the *hierarchical Gilbert hidden Markov model*.

Let $\{Y_t\}$ denote the underlying n-state Markov chain. The initial state distribution is given by the n-dimensional vector π, with $\pi_i = P(Y_1 = i)$. The state transition probabilities are controlled by the $n \times n$ matrix $\mathbf{A} = \{a_{ij}\}$, where $a_{ij} = P(Y_t = j | Y_{t-1} = i)$.

Once the model enters a state at time t, a group of S packet outcomes, denoted $x_t = [x_{t,1}, \ldots, x_{t,S}]$ is generated according to the parameters of the Gilbert model determined by Y_t. Such parameters are:

$$r_i = P(X_{t,1} = 1 | Y_t = i), \tag{4a}$$
$$p_i = P(X_{t,s} = 1 | X_{t,s-1} = 0,\, Y_t = i), \quad 1 < s \le S, \tag{4b}$$
$$q_i = P(X_{t,s} = 0 | X_{t,s-1} = 1,\, Y_t = i), \quad 1 < s \le S. \tag{4c}$$

We refer to the entire model as the tuple $\lambda = (\pi, \mathbf{A}, \mathbf{r}, \mathbf{p}, \mathbf{q})$, where $\mathbf{r}$, $\mathbf{p}$, $\mathbf{q}$ are vectors containing the respective parameters r_i, p_i, q_i, for each state i.

Given a sample of T batches of packet loss measurements $\{x_1, \ldots, x_T\}$, we are interested in the set of maximum likelihood estimators for λ. First, we define the following measures, using the notation from [14]:

$$\alpha_t(i) = P(x_1, \ldots, x_t, Y_t = i | \lambda), \tag{5a}$$
$$\beta_t(i) = P(x_{t+1}, \ldots, x_T | Y_t = i, \lambda), \tag{5b}$$
$$\gamma_t(i) = P(Y_t = i | x_1, \ldots, x_T, \lambda), \tag{5c}$$
$$\xi_t(i,j) = P(Y_t = i, Y_{t+1} = j | x_1, \ldots, x_T, \lambda). \tag{5d}$$

These can be efficiently evaluated through the forward-backward recursions (see [14] for details). Given the measures in Equation (5), it can be shown through the application of the EM method [14], that the model parameters may be iteratively estimated using the following formulae:

$$\pi_i = \gamma_1(i), \tag{6a}$$

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \tag{6b}$$

$$r_i = \frac{\sum_{t=1}^{T} \mathbb{1}\{x_{t,1} = 1\}\gamma_t(i)}{\sum_{t=1}^{T} \gamma_t(i)}, \tag{6c}$$

$$p_i = \frac{\sum_{t=1}^{T} S_{x_t}^{01}\gamma_t(i)}{\sum_{t=1}^{T}(S_{x_t}^{01} + S_{x_t}^{00})\gamma_t(i)}, \tag{6d}$$

$$q_i = \frac{\sum_{t=1}^{T} S_{x_t}^{10}\gamma_t(i)}{\sum_{t=1}^{T}(S_{x_t}^{10} + S_{x_t}^{11})\gamma_t(i)}, \tag{6e}$$

where $S_{x_t}^{ij}$ is the number of transitions from i to j in x_t, with both i and j being either 0 or 1. The development of these expressions can be found in [15].

B. Adaptive FEC Control

The goal of our mechanism is to keep the perceived loss rate below some pre-determined threshold, θ. To achieve this, we use our model from section III-A in order to predict the Gilbert model that characterizes the packet loss process in the near future. Then, we use this prediction, together with the analytical development from section II-A, to choose an efficient FEC scheme that will satisfy our loss rate constraint.

In our experiments, we will model the packet loss process with a hierarchical Gilbert HMM with the parameter S chosen so that it corresponds to 1 second of packet transmissions. The hidden Markov chain presented in our results has only 3 states. We also tested our methodology with models of more states, but the results we obtained with only 3 states are often much better.

Our mechanism is composed of two kinds of events: (a) the model parameters are periodically re-estimated in order to reflect the long-term changes in the network conditions; (b) at a lower time scale, the current model parameters are used together with recent measurements to predict a Gilbert model that best characterizes packet losses in the short-term future.

1) Parameter Estimation: Model parameters are estimated once every minute, by applying the equations in (6) for a number of iterations. The sample used for this training includes the last 3 minutes of packet loss measurements. Once the parameters are determined, we assign a specific FEC policy to each of the states in the HMM. Namely, if the Gilbert model in state i provides a loss rate which is already smaller than our threshold θ, then the policy for state i is not using any redundancy. Otherwise, we attempt to find a k:w setting, within a library of available schemes, for which the loss rate after reconstruction is below θ. If more than one scheme satisfy this condition, then we choose the one with the smallest overhead and reconstruction delay. On the other hand, if there are no FEC schemes that can satisfy the loss constraint, we choose the one which provides the closest loss rate to θ.

In the experiments we report on section IV, our loss rate constraint θ was chosen to be 3%. In addition, the schemes that we consider for determining the policy of a particular state are restricted to all k:w settings such that the reconstruction delay is at most 6 packet intervals.

2) Network State Prediction: Once every 5 seconds, we evaluate the distribution of the hidden state in the HMM given the outcomes of packet loss measurements in the latest 5 seconds. This can be easily obtained through the forward recursion, [14]. Using this information, we evaluate the distribution of the hidden state in each of the 5 seconds until the next prediction. Namely, if ψ is the distribution in the last second before prediction, then $\psi^{(t)} = \psi\mathbf{A}^t$, for $t = 1, \ldots, 5$, is the distribution of the state in each of the next 5 seconds.

After these distributions are obtained, we apply an heuristic rule to determine which among the 3 states in the model is more characteristic of the future network conditions in each second. For a given second t, in the prediction window, this rule works as follow:

- Let the 3 model states be sorted such that, $\psi_1^{(t)} \geq \psi_2^{(t)} \geq \psi_3^{(t)}$
- If $\psi_1^{(t)} \geq \Omega$, where Ω is a pre-defined parameter of the heuristic, then the policy chosen for the t-th second is that of state 1.
- Otherwise, the policy chosen for the t-th second is a composition of the policies in states 1 and 2.

The intuition behind this mechanism, is that, if we are certain that we will be in state 1, i.e., $\psi_1^{(t)} \geq \Omega$, then we need only to use the FEC policy previously associated with that state. If there is not enough confidence to make that decision, we then apply the policies of both states 1 and 2 in parallel. If the policies in states 1 and 2 are the same, then we only use it once. In our experiments, we chose $\Omega = 0.50$, specifying that we only need a majority of 50% in order to trust what is the current network state.

IV. Experimental Results

We carried out an experiment to assess the goodness of our methodology for adaptive FEC control. These experiments are based on a set of real packet loss measurements for VoIP traffic. We compare the performance of our method with that of another proposal for dynamic FEC selection from the literature, which is described in [8].

A. Packet Loss Measurements

In our experiments, we had at our disposal an extensive set of end-to-end measurements performed between four academic sites, two of these located in Brazil and the other two in the United States. These measurements display a large variety of network conditions, ranging from hours with no packet loss to complete link outages. In all measurements performed, CBR traffic was generated using tools available in the public domain, [16].

Each traffic generation session lasted for an hour and a total of 998 sessions were performed at different periods in the years of 2001, 2002 and 2004. In any given day of experiments, the sessions were conducted at three different periods, usually centered around the peak of usage in many of the links transversed, taking into account the time differences between the end points. The traffic pattern was chosen to emulate the behavior of a simplified Voice over IP (VoIP) tool, sending 50 packets with 324 bytes each per second. From our packet traces we produce a binary sequence $\{x_i\}_{i=1}^{T}$, where x_i is 1 in order to indicate a loss or 0 otherwise.

Many of the 998 collected traces exhibit packet loss processes that are not interesting for our packet loss recovery experiment, such as measurements of extremely low average loss rates. We selected 194 traces whose loss rates are between 1% and 30%, and containing consecutive loss periods no longer than 30 seconds.

B. Discussion

In the experiments, we compare our own methodology for FEC selection with another approach in the literature. Namely, we consider the mechanism proposed in [8] for voice over IP. Using the packet traces from section IV-A, we simulate the behavior of both control mechanisms and compute a set of measures over the loss recovery process.

Recall from section III-B that, in our methodology, we select the most appropriate FEC scheme from a restricted set of choices. More specifically, we consider only the k:w settings such that $w \leq 6$.

Our evaluation of FEC strategy quality is based on two primary metrics: the *loss recovery rate* and the *FEC overhead*. The first of these measures is the fraction of packet losses that were recovered by the FEC reconstruction process. The latter is the amount of redundancy units relative to the number of data units. In our analysis, a data unit refers to the original media payload sent over the network. A redundancy unit, on the other hand, is a parity block over a chain of packets in a k:w algorithm described in section II.

Another metric we compute is the ratio between the number of recovered packets and the total amount of overhead sent. We consider this as a measure of *FEC efficiency*. Intuitively, this measure represents the fraction of the redundancy units that were actually useful in recovering a packet loss.

We refer to the approach in [8] as the *reference method*. It is important to note that, in the reference method, the FEC schemes considered are restricted to the set 1:1, 2:2 and 3:3, so that the redundant copy of data may be compressed with a lower bitrate codec. Such compression will decrease the FEC overhead, and thus increase efficiency. However, our intention is to measure only the gain in efficiency obtained by the adaptive FEC control. In order to provide a better comparison to the reference method, we also evaluate a restriction of our proposal which uses only the schemes 1:1, 2:2 and 3:3. This restricted method can also be used with media-specific compression techniques, and therefore, it can achieve any reduction in overhead that can be obtained with the reference method.

Table I displays the three aforementioned metrics averaged over the 194 packet loss traces, for the reference method of [8], our own proposed method for media-independent FEC, and the restricted version which allows for media-specific correction. It can be verified that both our proposal and its restricted version, not only recover more packet losses but also use considerably less overhead than the reference method.

TABLE I
COMPARISON OF DIFFERENT STRATEGIES FOR FEC SELECTION AVERAGED OVER THE 194 TRACES.

FEC selection	Recovery	Overhead	Efficiency
Reference method	23.2%	76.6%	1.4%
Proposed method	25.6%	14.9%	8.4%
Restricted method	28.7%	30.8%	7.8%

Figure 4 shows the efficiency metric for the three control mechanisms evaluated separately for each of the 194 traces. The traces were sorted so that the curve corresponding to the proposed method for media-specific FEC is non-decreasing. It is clear that at the same time as the reference method falls below the curve for our media-specific proposal, our media-independent method is more efficient than both of them if the redundancy units have the same size as the original data units.

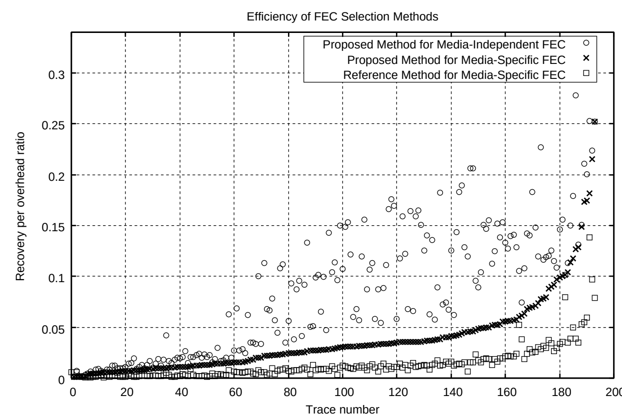

Fig. 4. Ratio between recovery and overhead for each trace.

As a qualitative example of our proposed method for media-independent FEC, Figure 5 shows the difference for a sample trace between the original packet loss process and the perceived loss rates after the reconstruction process has been applied.

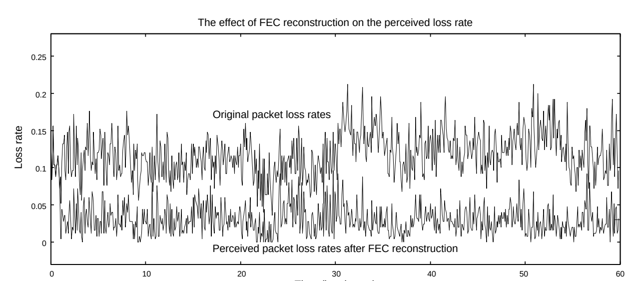

Fig. 5. Normalized loss rates after FEC reconstruction in 5 second intervals for a sample trace.

V. Conclusion

In this paper, we presented an algorithm for adaptive FEC selection that can be used in interactive streaming applications. Our method uses a predictive packet loss model in order to determine a Gilbert model that characterizes the network state in the future. Then, our method evaluates the expected loss rate after the reconstruction process for each setting in a group of media-independent FEC schemes, and chooses one of them using an intuitive heuristic rule. Both the analytical development used to determine the expected loss rate after FEC, and the decision mechanism for choosing a FEC scheme are original contributions of this article.

Using real traces we collected over the Internet, we compared the performance of our approach to that of a media-specific FEC control mechanism previously proposed in the literature. Our method not only recovers more packets but it does so more efficiently than the reference method, when we restrict the FEC schemes available to our decision mechanism to those used in [8].

It is important to notice that all the computations required for the entire control mechanism are sufficiently fast to be executed in a real-time application. We have plans to incorporate this mechanism into an existing interactive streaming application.

VI. Acknowledgements

We would like to thank Flavio P. Duarte for providing us with an implementation of the hidden Markov model training algorithms. David S. Boechat and Hugo H. C. Sato have also been very helpful in the implementation of other tools for the experiments that have led to this paper.

References

[1] A. P. Markopoulou, F. A. Tobagi, and M. J. Karam, "Assessing the quality of voice communications over Internet backbones," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 747–759, October 2003.

[2] D. Loguinov and H. Radha, "End-to-end Internet video traffic dynamics: statistical study and analysis," in *Proceedings of IEEE INFOCOM*, June 2002, pp. 723–732.

[3] R. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne, "Adaptive playout mechanisms for packetized audio applications in wide-area networks," in *Proceedings of the IEEE INFOCOM*, June 1994, pp. 680–688.

[4] S. B. Moon, J. Kurose, and D. Towsley, "Packet audio playout delay adjustment: performance bounds and algorithms," *ACM/Springer Multimedia Systems*, vol. 5, no. 1, pp. 17–28, January 1998.

[5] C. Perkins, O. Hodson, and V. Hardman, "A survey of packet loss recovery techniques for streaming audio," *IEEE Network*, vol. 12, no. 5, pp. 40–48, September 1998.

[6] D. R. Figueiredo and E. de Souza e Silva, "Efficient mechanisms for recovering voice packets in the Internet," in *Proceedings of the IEEE GLOBECOM*, vol. 3, December 1999, pp. 1830–1837.

[7] E. Altman, C. Barakat, and V. M. Ramos R., "Queueing analysis of simple FEC schemes for IP telephony," in *Proceedings of the IEEE INFOCOM*, 2001, pp. 796–804.

[8] J. C. Bolot, S. F. Parisis, and D. F. Towsley, "Adaptive FEC-based error control for Internet telephony," in *Proceedings of the IEEE INFOCOM*, March 1999, pp. 1453–1460.

[9] C. Perkins, *et al.*, "RTP payload for redundant audio data," RFC 2198, IETF, September 1997.

[10] M. Yajnik, S. B. Moon, J. F. Kurose, and D. F. Towsley, "Measurement and modeling of the temporal dependence in packet loss," in *Proceedings of the IEEE INFOCOM*, March 1999, pp. 345–352.

[11] K. Salamatian and S. Vaton, "Hidden Markov modeling for network communication channels," in *Proceedings of the ACM SIGMETRICS*, June 2001, pp. 92–101.

[12] P. Ji, B. Liu, D. F. Towsley, Z. Ge, and J. F. Kurose, "Modeling frame-level errors in GSM wireless channels," *Performance Evaluation*, vol. 55, no. 1-2, pp. 165–181, January 2004.

[13] X. Yu, J. W. Modestino, and X. Tian, "The accuracy of Gilbert models in predicting packet-loss statistics for a single-multiplexer network model," in *Proceedings of IEEE INFOCOM*, March 2005.

[14] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–285, February 1989.

[15] F. Silveira Filho and E. de Souza e Silva, "Modeling the short-term dynamics of packet losses," *To appear in ACM Performance Evaluation Review*.

[16] E. de Souza e Silva and R. M. M. Leão, "The TANGRAM-II Environment," in *TOOLS '00: Proceedings of the 11th International Conference on Computer Performance Evaluation: Modelling Techniques and Tools*. London, UK: Springer-Verlag, March 2000, pp. 366–369.