Autonomic QoS Optimization of Real-time Internet Audio using Loss Prediction and Stochastic Control

Lopamudra Roychoudhuri*
Department of Computer Science
Carroll University,
Voorhees 317, 100 N. East Ave,
Waukesha, WI 53186
Phone: (262) 524-7363, Fax: (262) 524-7397
Email: lroychou@carrollu.edu

Ehab S. Al-Shaer
Department of Software & Information Systems,
College of Computing and Informatics,
University of North Carolina,
9201 University City Blvd,
Charlotte, NC 28223
Phone: (704) 687-8663, Fax: (704) 687-6065
Email: ealshaer@unc.edu

Abstract

Quality of Internet audio is highly sensitive to packet loss caused by congestion in the links. Packet loss for audio is normally rectified by adding redundancy using Forward Error Correction (FEC). Alternatively, path diversity mechanisms are used to improve reliability and thus session quality. To achieve optimized receiver audio quality for transmissions using single or multiple paths, we propose a self-adaptive joint Error and Rate Control framework based on packet loss prediction and on-line quality assessment. The Error Control chooses appropriate FEC proactively to preserve quality with optimal bandwidth, using a Markov Decision Process (MDP) and a stochastic inventory control, a novel approach for multimedia error recovery. The Rate Control uses a quality optimization model to determine the optimal dispersion over single or multiple paths. We present results using simulation and Internet experiments to show the superiority of our mechanism over other similar techniques.

Keywords

VoIP, QoS, Proactive Mechanism.

INTRODUCTION

Deployment of Next Generation Networks and service convergence is likely to make IP technology the main vehicle for carrying interactive voice and video. Quality of Internet multimedia is highly sensitive to packet loss (Cole, R.G. & Rosenbluth, J.H., 2001; Markopoulou, A. P., Tobagi, F. A. & Karam, M. J., 2002) caused by congestion in the links. Packet loss for real-time UDP audio and video is normally rectified using FEC (Lin, S., Costello, D.J., 1983), where a number of redundant packets are sent with the original packets. FEC is shown to be the most common technique for maintaining acceptable quality in presence of loss (Jiang, W. & Schulzrinne, H., 2002). But the challenge with FEC is its bandwidth overhead, as FEC must be sufficient but not excessive, and timely, in order to be effective. Thus FEC degree and duration should be chosen adaptively in response to the network conditions of
bandwidth degradation and packet loss. Alternatively, path diversity mechanisms, where session packets are dispersed over multiple paths, are used to improve reliability and thus multimedia session quality in overlay/p2p systems (Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001; Apostolopoulos, J., 2001; Fei, T., Tao, S., Gao, L. & Guerin, R., 2006), or multihoming networks (Akella, A., Pang, J., Maggs, B., Seshan, S. & Shaikh, A., 2004; Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001). Packet dispersion can be a viable option assuming that at least one path will provide good performance to maintain session quality (Savage, S., Collins, A., Hofmann, E., Snell, J. & Anderson, T., 1999). In particular, packet dispersion is significantly beneficial to reduce the effect of high degree of burst loss (Zlatokrilov, H. & Levy, H., 2004). But compared to FEC, path diversity is a relatively costly solution, since it can incur more sudden changes in one-way delay (Tao, S., Xu, K., Estepa, A. et. al., 2005), disrupting a smooth playback. Thus the challenge is to provide an error control solution that adaptively combines the benefits of both FEC and dispersion techniques under different network conditions in order to provide optimal quality. It is also imperative that such error control is combined with an efficient rate control in order to provide bandwidth-friendly transmission with an effective degree of loss recovery.

In this paper we present a self-adaptive joint error and rate control mechanism that ensures an optimal receiver quality at real-time by taking proactive control actions, based on packet loss prediction (Roychoudhuri, L. & Al-Shaer, E., 2005) and on-line quality assessment (Roychoudhuri, L. & Al-Shaer, E., 2005; Roychoudhuri, L., Al-Shaer, E. & Settimi, R., 2006). The Loss Predictor is an end-to-end monitor that tracks the one-way delay and inter-packet gap of in-line stream packets, as well as short- and long-term trends, and indicates the current degree and severity of congestion, hence the likelihood of packet loss in the next window of packets. The Audio Quality Assessor is a passive monitor that assesses the receiver audio quality objectively for a real-time audio stream in terms of Mean Opinion Score (MOS), the ITU standard of voice quality assessment (ITU-T Recommendation P.800, 1996). The Error Control recovers individual path loss and maintains receiver quality in the short term using optimal FEC. The FEC degree and duration are chosen dynamically using an MDP and stochastic inventory control, a novel approach in the area of multimedia error control. The Rate Control detects changing bandwidth and uses a rate quality optimization model to proactively diversify optimal codec/bitrate combination over single or multiple paths. This is superior to the reactive feedback used in current sender based single-path (Bolot, J. & Vega-Garcia, A., 1996; Mohamed, S., Cervantes-Perez & Afifi, H., 2001) and multi-path (Nguyen, T. & Zakh, A., 2003) rate and error control mechanisms that depend on RTCP feedback of packet loss. Second, our mechanism is user quality centric, as opposed to ad-hoc reaction to network packet loss using static FEC (Jiang, W. & Schulzrinne, H., 2002). The sole purpose of the error control and rate adaptation actions is to optimize receiver quality at real time using objective audio quality assessment. The control decisions are made to improve or retain the end quality above a specific threshold, and are not simply based on network condition targets, such as bandwidth or packet loss levels. That may mean that certain degree of loss may be tolerable in terms of quality and actions will be taken only when required, ensuring the efficiency of the mechanism. The framework is currently designed for audio, but is easily extensible to video applications.

RELATED WORK

As an initial work in this area, Bolot and Garcia presented a combined Error and Rate control mechanism for audio that adapts to the loss feedback from the receiver using RTCP (Bolot, J. & Vega-Garcia, A., 1996). Padhye, Christensen, Moreno (Padhye, C., Christensen, K. & Moreno, W., 2000) improved on Bolot’s algorithm by considering the history of packet loss before changing the amount of redundancy. Bolot and Fosse-Parisis presented an FEC technique that adapts to varying loss conditions in the network with RTCP feedback (Bolot, J., Fosse-Parisis, S., & Towsley, D., 1999). Mohamed et. al. (Mohamed, S., Cervantes-Perez & Afifi, H., 2001) designed a system that dynamically adjusts codec parameters based on RTCP feedback and MOS speech quality scores generated by neural networks. All these mechanisms react to packet loss using RTCP feedback from the receiver. In contrast, we predict loss and take error and rate control actions accordingly. Jiang and Schulzrinne showed the effectiveness of static FEC techniques to repair bursty loss (Jiang, W. & Schulzrinne, H., 2002). In contrast, we generate dynamic FEC based on current loss probability using an MDP and a stochastic inventory control model that is novel in the context of multimedia error recovery.

In the area of packet dispersion, a path diversity transmission approach sending different subsets of video packets over different paths ensured lower burst loss than over a single path (Apostolopoulos, J., 2001). The authors evaluated two packet dispersion scheduling policies (Zlatokrilov, H. & Levy, H., 2004), and concluded that the use of packet dispersion over multiple paths reduced Noticeable Loss Rate of VoIP and improves session quality. How-
ever, none of the approaches proposed any adaptive multi-path error or rate control solution. The authors proposed an adaptive FEC scheme for multi-path video streaming using multi-senders and single-sender (Nguyen, T. & Zakhor, A., 2003) architectures to one receiver. In this work, the receiver calculated the loss rate after the fact and derived an optimal sending rate for each sender. In contrast, our approach provides proactive error and rate control based on packet loss prediction that considers the path characteristics and target quality. The authors proposed redundant voice streams encoded using multiple descriptions coding over multiple independent paths (Liang, Y., Steinbach, G. & Girod, B., 2001) in order to reduce latency and loss rates. Their main contribution was to provide a playout scheduling of multiple streams based on a Lagrange cost function, which took packet delay probability into consideration. This is in contrast to our packet loss prediction mechanism, where we identify variation of delays relative to degree of congestion, and delay trends to estimate the likelihood of loss in next packet train. In (Tao, S., Xu, K., Estepa, A. et. al., 2005) the authors used an adaptive diversity only path switching algorithm by measuring actual quality of the paths and switching to an alternate path whenever the alternative path quality exceeded the original path quality by a constant value. In contrast, we measure projected quality based on Loss Predictor, and combine FEC and path diversity in a single mechanism. We also use a probabilistic loss predictor for multiple bottleneck paths, as opposed to our previous work (Roychoudhuri, L. & Al-Shaer, E., 2005), where we presented a loss predictor for a single bottleneck path.

JOINT RATE/ERROR CONTROL SYSTEM OVERVIEW

Our mechanism assumes the existence of multiple paths of uncorrelated characteristics between the sender and the receiver. It is assumed that such disjoint paths do exist in the underlying topology such as overlay and multihoming networks. In multihoming networks alternate paths are provided by different network providers, as described in (Akella, A., Pang, J., Maggs, B., Seshan, S. & Shaikh, A., 2004; Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001). The congested paths can be bypassed using alternate paths (Savage, S., Collins, A., Hoffman, E., Snell, J. & Anderson, T., 1999), as the performance bottleneck is often not at the edge link, but in the backbone or inter-AS links (Akella, A., Pang, J., Maggs, B., Seshan, S. & Shaikh, A., 2004). Identification of disjoint alternate paths based on path quality monitoring using network performance is a well-researched topic in overlay and p2p systems (Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001; Apostopoulos, J., 2001; Fei,T., Tao,S., Gao,L. & Guerin,R., 2006), but not in the focus of this paper.

Figure 1 describes the components of our framework. The Rate control initially calculates an optimal mix of bitrates of a multi-bitrate codec, distributed over single or multiple paths. It solves a rate-quality optimization problem to ensure the highest possible audio quality and transmission rate under the current network conditions of available bandwidth, end-to-end delay and packet loss. It uses feedbacks from the Loss Predictor (see Loss Predictor subsection), the Audio Genome (see On-Line Audio Quality Assessment subsection), and the Bandwidth and Delay Monitors for this purpose. The Loss Predictor uses active probing over multiple paths in order to estimate the individual path loss probability that is used in the path selection decision by the Rate Control. The feedback from Audio Genome is used to assure that the optimal quality is maintained at the receiver. As the session proceeds, the Error Control takes adaptive actions in the short term to minimize packet loss by dynamically changing degree and duration of FEC on each path, maintaining optimal quality and bandwidth overhead. The Error Control is activated if likelihood of loss increases on a certain path, as indicated by the Loss Predictor. The Error Control is designed as an MDP that determines the optimal type, degree and frequency of FEC by optimizing a reward to improve quality and minimize bandwidth and delay overhead.

Under significant bandwidth degradation causing loss bursts that are not optimally repairable by FEC, the error control becomes infeasible over one or more paths. In this case the long-term Rate Control is reactivated to calculate optimal rate and dispersion over alternate paths (assuming such paths exist [10] (Savage, S., Collins, A., Hoffman, E., Snell, J. & Anderson, T., 1999) under the changed network conditions, in order to maintain the audio quality at the receiver. The following subsections and sections provide details on the individual components of our system.
Packet Loss Predictor

The Loss Predictor expresses the probability of loss in the next packet train of a UDP transmission by passively analyzing (1) changes in the available bandwidth, manifested as inter-packet gap and end-to-end delay variations of in-line stream packets as “evidences”, and (2) near-past history of congestion in terms of observed loss patterns and trends of gap and delay variations. In our approach, we identify the baseline delay and inter-packet gap as the delay and gap under no congestion. In contrast, various degrees of increase or decrease of the delay and gap occur before and after loss occurrences due to cross traffic at the bottleneck links. We identify the delay and gap expansion at the capacity saturation point of a path as loss thresholds, (Figure 2 - vertical lines denoting packet loss), after which packet loss is more likely. We track the short-term and long-term trends as indications of congestion build-up and release, and accordingly derive the likelihood of packet loss by detecting loss thresholds.

In our new approach [31] (Roychoudhuri, L. & Al-Shaer, E., 2007) over our previous work [30] (Roychoudhuri, L. & Al-Shaer, E., 2005), the Loss Predictor is formalized as a Bayesian probability measure of packet loss in the next projection window based on the evidences of inter-packet gap expansion and delay variation. We first estimate the projected delay $\hat{D}$ and expansion $\hat{Y}$ in the next projection window, based on the current delay $D$, expansion $Y$ and the short-term and long-term trends. For the projected delay and gap expansion, we calculate the Delay Distance and Expansion Distance metrics that are measures of the delay increase and the gap expansion in relation to the baseline and loss thresholds. The probability measures of loss are classified into two categories - (1) $L_2$, derived from the evidence of Delay Distance, and (2) $L_e$, derived from the evidence of Gap Expansion Distance.

We combine the probability measures into a unified Loss Predictor by normalizing and assigning positive weights to each information source as following:

$$f(L_2(D,Y),L_e(D,Y)) = w_2 \cdot L_2(D,Y) + w_e \cdot L_e(D,Y),$$

and $w_2 + w_e = 1$.

We use an exponentially weighted moving average $\hat{L}_{1 \rightarrow 2}$ of the predictor data over the past long term window, where $\hat{L}_{1 \rightarrow 2}$ is the new average based on the current observed $L_2$ and average $\hat{L}_{1 \rightarrow 2}$ and $\alpha$ is chosen to be a high value of 0.9 to efficiently capture the changes in the predictor, yet smooth out the measurement noise.

Threshold Accuracy for multiple bottleneck paths: We conducted a large set of Internet experiments over a period of one year using a number of sites on the RON/NetBed from emulab.net, a wide-area network testbed consisting of sites located around the Internet. We gathered one way delay and packet loss data on both the forward and reverse paths in order to get accurate picture of the correlation of delay variation and packet loss at the receiver end. Unlike the consistent baseline delay, a wide range of delay loss thresholds was noticed in our measurements (Roychoudhuri, L., Al-Shaer, E. & Brewster, G.B., 2005). This was possibly due to large number of hops and multiple bottlenecks over Internet paths, and variable degrees of cross-traffic causing unpredictable queuing delay variations and packet loss at the intermediate links. For some sites the loss threshold varied considerably inside one session, whereas we saw a consistent unimodal loss threshold for others inside and across sessions. We can characterize the behavior of the loss thresholds in the following categories (Figure 3):

- A consistently strong threshold - A prominent threshold is possible if the congestion occurs mostly at a single bottleneck link significantly narrower than other links. In Figure 3, 57% of the observations show a prominent grouping around a strong mode (Column A).
- Scattered (many thresholds) - The second most observed case is that of scattered thresholds with no clear mode value (Column B), due to variable degrees of transient cross traffic through different parts of the path.
- Two or more distinct thresholds - Multiple distinguishable thresholds can occur when there are non-homogeneous buffer sizes in the intermediate routers, and no prominent bottleneck link (Figure 3, Columns C and D). Each of these needs to be considered as individual thresholds.

In the case of multiple thresholds we consider the accuracy of each threshold, i.e. the probability of loss at each threshold, in the calculation of our loss prediction. The probability measures $L_2$ and $L_e$ are expressed as weighted
sums of loss probabilities at mutually exclusive loss thresholds. For example, assuming that a threshold is defined for every observed loss, let \( \Phi_k \) be the delay loss thresholds in increasing order of delay values. Then \( L_d \) is derived as the following:

\[
D^k \quad \text{is the measured one way delay for packet } k, \quad w^k_d, \quad \text{the accuracy of the delay loss threshold } \Phi_k, \quad l
\]

\[
\text{probability of loss at delay } \lambda, \quad M_e(\lambda) \quad \text{is the Delay Distance w.r.t. threshold } \Phi, \quad L_e \quad \text{is derived similarly based on Expansion Distance and gap loss thresholds. The details can be found in [31] (Roychoudhuri, L. & Al-Shaer, E., 2007).}
\]

Predictor simulations show average 15% false negatives and 10% false positives under reasonable conditions. Internet experiments show promising results: less than 30% false negatives in 83% cases for loss rates >1%, and less than 10% false positives in 90% cases (Roychoudhuri, L. & Al-Shaer, E., 2007).

**Predictor active probing over multiple paths**: Active probing to evaluate alternate path conditions is a commonly used technique in path diversity (Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001; Tao, S., Xu, K., Estepa, A. et. al., 2005). In a multi-path transport, we need a predictor for each path to estimate loss probability on each path. A probe with a small packet size (20 bytes) is sent on the alternate paths that matches the frequency interval (average of 20ms) of audio packets on the original path. The resulting probe size of average 8kbps does not consume significant bandwidth and is small compared to the path capacities.

<<< Fig4_two_state_Markov.tif >>>

Figure 4: 2-State Markov Process.

**Two-state Markov Chain Model for Packet Loss Prediction**: A two-state discrete time Markov process has been used as an approximation of Internet packet loss due to congestion (Yajnik, M., Moon, S., Kurose, J. & Towsley, D., 1999). The loss prediction, if accurate, should also be represented using a two-state Markov process, dividing the range \([0,1]\) of possible Predictor values into two states, namely good \((0 \leq P_{\text{pred}} \leq \lambda)\) and bad \((\lambda < P_{\text{pred}} \leq 1)\), where \(\lambda \geq 0.5\). The process transits from an uncongested good(0) state, with small probability of packet loss, to a congested bad(1) state with high loss probability, and vice versa (Figure 4). \(P\) and \(Q\) are the transition probabilities from good to bad state and vice versa: \[ P = P[\xi_0 = 1|\xi_0 = 0, q = P[\xi_1 = 0|\xi_1 = 1]. \]

The maximum likelihood estimators [4] (Billingsley, P., 1961) of \(P\) and \(Q\) are \[ \hat{P} = \frac{n_{11}}{n_{10}}, \hat{Q} = \frac{n_{01}}{n_{00}}, \] where \(n_{ij}\) = number of times in the observed time-series \(f\) follows \(i\), and \(n_i\) = number of \(i\) 's in the trace \(\xi = \{0, 1\}\).

The data can be represented as alternating series of 0 s, or good run lengths, and series of 1 s, or bad run lengths. The respective good and bad run length distributions for this model are \(f_{\text{good}} = (1-p)^{f-1}\) and \(f_{\text{bad}} = (1-q)^{f-1}\), where \(f = 1, 2, \ldots, \infty\).

<<< Fig5_bad_run_dist.tif >>>

Figure 5: Cumulative Distribution of 'Bad' Runs of Loss Prediction.

To validate this model, we conducted a set of simulation experiments using ns-2 (Network Simulator (NS2)), where a CBR stream of 64kpbs flowed from a source to the sink through more than 10 intermediate hops, and a number of Pareto cross traffics of various packet sizes and rates flowed through intermediate hops at dispersed
points in time, causing different degrees of congestion resulting in stream packet loss at the intermediate links. We represented the predictor data as good and bad run lengths and evaluated the models to see if the observed data can indeed be represented by a 2-state Markov process. We also tested with various values of $\bar{\alpha}$, the 'goodness' threshold, to determine a best fit. In Figure 5, we find that the cumulative distribution of the bad run lengths for the total experiment set closely match the cumulative distributions predicted by the 2-state Markov process model for $\bar{\alpha} = 0.5$. Since the loss prediction can thus be approximated as a two state Markov process, we use a Markov Decision Process for modeling an error control based on the Predictor feedback, presented in Error Control section.

**On-Line Audio Quality Assessment**

Figure 6: Comparison of codecs: single loss.

Figure 7: Regression models and Observed MOS - G.729.

Audio codecs have a diverse range of compression degrees and underlying technologies. Audio quality of any speech processing system is generally described in terms of MOS (Mean Opinion Score) (ITU-T Recommendation P.800, 1996), the formal subjective measure of received speech quality, which is a real number between 1 and 5, where 1 is 'Bad' and 5 is 'Excellent'. The main factors that significantly influence the audio quality in IP telephony thus include codec type, loss rate, loss burst, inter-loss gap, delay, and recency (Bolot, J. & Vega-Garcia, A., 1996; Clark, A.D., 2001; Cox, R. & Perkins, R., 1999). In our simulation (Roychoudhuri, L. & Al-Shaer, E., 2005) and Internet experiments (Roychoudhuri, L., Al-Shaer, E. & Brewster, G.B., 2005) we observed that different codecs react differently and non-linearly under similar loss distribution scenarios. As an example, in Figure 6, GSM, a codec of medium bitrate 13kbps, performs as the best codec in high loss range, better than codecs of higher and lower bitrates. G.729, a codec of low bitrate 8kbps, performs well in high loss conditions (2.45 at gap 5, i.e. 20% loss), but worse than others under low loss. However, the challenge here to establish a framework that derives audio quality on-line considering these factors. ITU specified E-model (ITU-T Recommendation G.107, 1998; ITU-T Recommendation G.108, 1999) provides a computational model to derive relative impairments to voice quality and to estimate subjective MOS. But ITU provides no analytic methods that can directly measure the impairment due to random loss conditions of bursts and inter-loss gaps.

Audio Genome (Roychoudhuri, L. & Al-Shaer, E., 2005; Roychoudhuri, L., Al-Shaer, E. & Settimi, R., 2006) is a statistical framework that quantifies the effects of packet loss on various codecs and bitrates (of a multi-rate codec G.722.2 (ITU-T Recommendation G.722.2, 2002)) by considering a wide range of loss rates, inter-loss gaps, and loss bursts of degree 1 to 4, the most occurring burst sizes as observed in the Internet (Bolot, J. & Vega-Garcia, A., 1996). The Audio Genome approach can be described as follows.

- **Generation of audio clips with packet loss scenarios:** We drop packets from audio clips using a periodic drop framework for a set of chosen codecs. It is worth noting that the framework is not limited to these codecs, as other codecs can easily be added by following these steps.

- **MOS evaluation and observations:** Using PESQ (Perceptual Evaluation of Speech Quality, an ITU objective speech quality testing scheme (ITU-T Recommendation P.862, 2001)), we compare 'pure' and 'poisoned' audio clips to deduce MOS scores under loss. We also observe the characteristics and behavior of each codec under various packet loss conditions.

- **Codec Quality Function Derivation:** We deduce codec quality functions for the collected data under loss using multiple polynomial regression analysis techniques.

For example, the following is the estimated regression equation for the data for G.729, where $y$ is MOS, $x$ is the inter-loss gap length, and $\bar{\alpha} = 1 + \alpha$.

$$y = -2.41603 + 1.35213 \times \ln(x) - 0.05901 \times (\ln(x))^2 + 3.4438 \times D_1 - 0.34901 \times \ln(x) \times D_1 + 2.13615 \times D_2 - 0.3143 \times \ln(x) \times D_2$$
The regression equations for loss bursts of 1, 2 and 3 are derived by setting $U_1, U_2, U_3$ to 1 respectively, with the rest of $D$’s equal to 0. The equation for burst $\geq 4$ is derived by setting all $D$’s to 0. In Figure 7 the regression equations match the observed MOS for G.729 well.

- **Online prediction of audio quality:** We utilize the resulting repository of quality regression functions to deduce the MOS by analyzing a series of (inter-loss gap, loss burst) values collected from the ongoing session.

For all codecs, under all loss scenarios, Audio Genome shows high accuracy of 96%-98% in average with low standard deviation of 0.07-0.12 and minimum accuracy of 91%. Though it is meant for audio, the design of this component can be extended to a Video Genome by following similar steps with video codecs.

**ERROR CONTROL**

Figure 8: Single-Product Stochastic Inventory Control Model for Audio Quality.

Figure 9: Quality model under a $[r, s]$ rule.

The objective of the short term self-adaptive Error Control is to maximize receiver quality under current and predicted loss conditions by adding FEC that consumes optimal bandwidth. This is achieved by an inventory control MDP that optimizes the reward based on audio quality improvement less bandwidth overhead at each epoch. The degree and duration of FEC is dynamically determined based on the Loss Predictor and the loss patterns observed so far by the Loss Monitor.

**Markov Decision Process Model for Error Control**

We use a single-product stochastic inventory control model [28] (Puterman, M., 1994), where a decision to use additional error is made at the beginning of an epoch $t$. Thus the state of the system is the projected session quality at the start of an epoch, the action is to apply an amount of error control over next window of packets, and the objective is to maximize a reward (quality improvement less bandwidth and delay costs) over the decision making horizon (Figure 8).

Without loss of generality, we assume that a single-bitrate stream is divided among $k$ alternate paths from the sender to the receiver with a ratio $w_i, i = 1, \ldots, k$, where $\sum_i w_i = 1$, an initial distribution of which is determined by the Rate Control. We use a decision rule [28] (Puterman, M., 1994) to assign error control at each epoch given the system quality projected by the path Predictors, stated as: Order sufficient error control to raise the audio quality to be $\geq \Sigma$ whenever the projected quality at the beginning of an epoch, as determined by the Predictor states for all paths, is $\leq \Sigma$. Do not order error control if the projected quality is $\geq \Sigma$. Figure 9 illustrates this decision process. The following is the MDP model.

**Decision Epochs.** Event-driven, based on the Predictor feedback frequency for each path from the receiver to the sender.

**States.** Two states depending on the value of projected session audio quality $\bar{S}_t$ at the receiver at beginning of interval $t$, $S_t = \{S_1, S_2\}$, where $S_1 = \{\bar{S}_t < \bar{S}\}$ is the good state, and $S_2 = \{\bar{S}_t \geq \bar{S}\}$ is the bad state. The projected session quality $\bar{S}_t$ is the weighted sum of the individual path projected qualities $\bar{S}_t^f, f = 1, \ldots, k$ due to expected packet loss in path Predictor states $P_{tf}$, $f = 1, \ldots, k$, have known stationary probability distributions $\bar{P}_{tf}^f, P_{tf}^f = \bar{P}_{tf}^f$. We assume that Predictor states $P_{tf}$, $f = 1, \ldots, k$ have known stationary probability distributions $\bar{P}_{tf}^f, P_{tf}^f = \bar{P}_{tf}^f$. We assume that Predictor states $P_{tf}$, $f = 1, \ldots, k$ have known stationary probability distributions $\bar{P}_{tf}^f, P_{tf}^f = \bar{P}_{tf}^f$.

**Actions and Decision Rules.** The actions are $A_{S_1} = \{0\}, A_{S_2} = \{\bar{S}_t - \bar{S}_t\}$, where $\bar{S}_t$ denotes no error control, and $\bar{S}_t$ denotes error control that improves the quality from $\bar{S}_t$ back to at least $\bar{S}_t$. The decision rule is thus:
Rewards. Let $F_k$ be the amount of error control used at beginning of interval $k$, and $[L(F_k)]_i$ and $[H(F_k)]_i$ be the relative quality improvement and cost (bandwidth and delay overhead). The reward is $\sum_{i=1}^{n} [L(F_k)]_i - [H(F_k)]_i \cdot$ We present actual formulas for reward measurement in the Evaluation section.

Transition Probabilities. The probabilities are weighted averages of individual path predictor Markov model transition probabilities under no error control and error control actions respectively. Ideally, when error control is applied, there should be no transition from good state to bad state, and a sure transition from bad to good.

Error Control Enumeration

The above model poses two problems. (i) We need to determine projected quality $F_k$ due to predicted packet loss in the next window on each of the $K$ individual paths in Predictor states $F_k$, $k = 1, \ldots, K$. (ii) If $F_k \leq F$, we need to enumerate the error control $E$ that will improve the session quality by $E(F - \xi)$. 

Projected Quality for Path Predictor States: The Loss Monitor at the receiver keeps track of loss patterns observed in good and bad Predictor states for each path $l$ in terms of average loss burst lengths, inter-loss gap lengths, and loss event lengths (a series of losses occurring 'sufficiently close' of one another [19] (Koodli, R. & Ravikanth, R., 2002), with respective averages $\overline{\text{loss burst}}, \overline{\text{inter-loss gap}}$, and $\overline{\text{loss event}}$. We assign loss bursts, inter-loss gaps and loss events to the Predictor states as follows: (i) Loss bursts and loss events belong to the predictor state immediately preceding it. (ii) For an inter-loss gap, if the first loss is in state $\alpha$, we need to add a loss burst of degree 2 to Audio Genome and derive quality score $\xi_{\alpha}$. 

Determination of Dynamic FEC Degree and Duration: In a conservative approach, we identify all the paths for which $F_k \leq F$ and apply the following steps on each path such that each $F_k$ is raised over $\xi$.

1. Determine optimal loss patterns for projected quality $E$ : We first determine a loss distribution $\{(\text{Gap}_{\alpha}, \text{Burst}_{\alpha})\}$ that is tolerable for each path in terms of quality. For each loss burst value $\text{Burst}_{\alpha}$ stored in Genome for the codec in use, we choose $\text{minGap}_{\alpha}$ to be the minimum gap such that the value of the regression function $\left[[(\text{Gap}_{\alpha}, \text{Burst}_{\alpha})] \right]$ exceeds $\xi$. We match the current $\{(\text{Gap}_{\alpha}, \text{Burst}_{\alpha})\}$ to a closest acceptable pattern using Algorithm 1. If $\text{Burst}_{\alpha} = 1$, that is, majority of the losses are single losses, the acceptable pattern $\{(\text{Gap}_{\alpha}, \text{Burst}_{\alpha})\}$ is $\{(\text{minGap}_{\alpha}, 1)\}$. If the average loss burst is of degree 2 or higher, we select the optimal loss burst to be the burst length with the $\text{minGap}_{\alpha}$ that is the closest to the current average gap, i.e. $\text{minGap}_{\alpha} = \min \{\text{minGap}_{\alpha} \mid \text{Burst}_{\alpha} \geq 2\}$. In this case $\{(\text{Gap}_{\alpha}, \text{Burst}_{\alpha})\}$ is $\{(\text{minGap}_{\alpha}, \text{Burst}_{\alpha})\}$.

Algorithm 1 OptimalPattern

1: for all $j \notin \text{Genome Loss Burst table do}$
2: $\text{minGap}_j \leftarrow \arg \min \{f(\text{Gap}_j, f) > \xi\}$ \{f = Genome regression function.\}
3: end for
4: if $\text{Burst}_{\alpha} = 1$ then
5: $\text{Burst}_{\alpha} \leftarrow 1; \text{Gap}_{\alpha} \leftarrow \text{minGap}_{\alpha}$
2. Perform appropriate Error Control: FEC Degree. We use Reed-Solomon (RS) [21] (Lin, S., Costello, D.J., 1983) encoding, where \( k \) original packets are used to produce total \( n \) packets to recover a burst of maximum \((n - k)\) packets with an added overcode \( \frac{n - k}{k} \). The FEC degree \((n - k)\) is determined by \( \left(\frac{\text{avgBurst}}{\text{burst}_{\text{OPT}}} + 1\right) \), the minimum loss burst length that needs to be rectified in order to achieve acceptable quality. \( n \) is determined by the delay budget, the number of packets we can afford to wait at the receiver before reconstruction.

FEC Duration. Obviously, the optimum quality inter-loss gap is greater than the current average gap, i.e. \( \frac{n - k}{k} > 1 \). We need to apply FEC at least for the duration of \( \left(\frac{\text{max(lossEvent)} - \text{avgBurst}_{\text{OPT}}}{\text{avgBurst}_{\text{OPT}} - \text{avgBurst}_{\text{OPT}} + 1}\right) \) packets in order to recover the intolerable loss. We also need to apply FEC for at least the average loss event length in the current state. The FEC duration is thus chosen to be \( \left(\frac{\text{max(lossEvent)} - \text{avgBurst}_{\text{OPT}}}{\text{avgBurst}_{\text{OPT}} - \text{avgBurst}_{\text{OPT}} + 1}\right) \).

FEC Infeasibility. FEC on a path can be infeasible due to two reasons: (i) The average loss burst is so large that \((n - k)\) may exceed delay budget. (ii) Even if \((n - k)\) does not exceed the delay budget, FEC overcode may be too high in ratio to the quality gain, i.e. \( \frac{n - k}{k} > 1 \). In such cases, the Rate Control is invoked to re-calculate the path dispersion.

RATE CONTROL

The proactive Rate Control provides optimum rate dispersion of codecs/bitrates over single or multiple paths while maintaining an optimized receiver audio quality, based on feedback from Loss Predictor and Audio Genome. Combining multiple bitrates is a viable and flexible option for rate control. For a multi-bitrate codec G.722.2 that allows switching between 9 bitrates, we observed almost identical quality degradation patterns for all bitrates under all loss scenarios [33] (Roychoudhuri, L. & Al-Shaer, E., 2005). The average of all observations for G.722.2 as a single codec significantly reduced the data that needed to be stored. In addition, although different bitrates of G.722.2 have different quality MOS, subjective quality testing combining various bitrates in different ratios showed acceptable audio quality, and was observed to be a linear combination of individual bitrate qualities.

Rate-Quality Optimization Model

In our earlier work [32] (Roychoudhuri, L. & Al-Shaer, E., 2004) on a single path, given a set of \( B \) codec bitrates, we derived the optimal codec ratio that would maximize the receiver audio quality under the current constraints of predicted available bandwidth, end-to-end delay and packet loss. We used the predictor \( P \) as the congestion detector, 0 indicating no congestion, and 1 indicating capacity saturation leading to packet loss. We extend the single path multi-bitrate rate control problem over multi-path transmission by considering the predictor values as congestion indicators observed for each path, and reducing the available bandwidth of each path by a respective congestion detection factor.

Problem Definition: Given \( B \) bitrates and \( k \) independent paths (different bandwidth, one-way-delays and loss rate), derive the optimal bitrate combination for the audio stream distributed over these paths that maximizes the audio quality.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mode</th>
<th>Bitrate (kbps)</th>
<th>Delay for 1 byte (ms)</th>
<th>Packet size (bytes)</th>
<th>MOS Under no loss</th>
<th>MOS for 1% single loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>Mode 7</td>
<td>23.05</td>
<td>1.64</td>
<td>59</td>
<td>3.92</td>
<td>3.4</td>
</tr>
</tbody>
</table>
Problem Formulation: Let $p_{ij}$, $u_{ij}$, and $d_{ij}$ be the predictor values, one-way-delays and available bandwidths for $k$ paths. $f_i$ is the normalized path congestion detection factor, where $f_i = \frac{p_i}{\sum_{t=1}^{k} p_t}, i=1,...,k$. Let $b_j$ and $d_j$ be the bit-rates and delays for $n$ modes of a multi-bitrate codec. Let $r_{ij}$, $q_{ij}$ be the the bitrate ratios and MOS under current loss conditions for $n$ codec modes over $k$ paths. The optimization problem is thus:

$$
\max \quad e = \sum_{i=1}^{k} \sum_{j=1}^{n} q_{ij} r_{ij}
$$

subject to

Available bandwidth constraints:

$$
\sum_{j=1}^{n} b_j r_{ij} \leq u_{ij} = (1 - f_i), i=1,...,k
$$

Delay constraints:

$$
\sum_{j=1}^{n} d_j r_{ij} + u_{ij} d_{ij} \leq 400, i=1,...,k
$$

Quality constraints:

$$
\sum_{i=1}^{k} \sum_{j=1}^{n} q_{ij} r_{ij} = 1
$$

Transmission rate constraint:

The objective function is the audio quality to be maximized, and is the sum of the product of codec bitrate ratios and quality scores under current loss condition, as determined by Audio Genome. The first constraint allows distributing the transmission over the $k$ paths proportional to the individual predicted available bandwidths. The total codec delay, expressed as the sum of products of encode/decode delay and percentage of each codec bitrate plus the one-way-delay of each path should not exceed 400ms. The quality sum should be bounded by 4.3 (highest possible score) and $\phi$, a target quality.

Figure 10: G.722.2 Bitrate ratio on Paths 1 and 2 with avail-bw 12.5k – no loss.

Figure 11: G.722.2 Bitrate ratio on Paths 1 and 2 with avail-bw 12.5k – under 1% loss.

Evaluation of Rate-Quality Optimization Problem

As a simple case of multi-path, we used a two-path topology with one `primary' (Path 1). The transmission was diverted through Path 2 only if congestion was detected on Path 1 (similar to Andersen, D.G., Balakrishnan, H., Kaashoek, M.F. & Morris, R., 2001), indicated by the Path 1 predictor. We tested the feasibility of our Rate-Quality optimization model for a two path topology under no-loss and a moderate 1% single loss scenarios. We chose 3 bitrates of G.722.2 (Table I) to test the feasibility of our Rate-Quality optimization model in the case of a multi-bitrate codec. We used GLPK (Gnu Linear Programming Kit, GNU) to solve the Linear Programming problems. The problems consisted of 6 variables (a reasonable number for codec combinations) that ran fast to provide feasible solutions. The MOS scores were derived from Audio Genome. The available bandwidth values were varied from moderate 100kto near congestion 12.5kand OWD was kept at 40ms for both paths. In higher bandwidth scenarios the codec percentage on both paths consisted of Mode 7, the solutions tending towards Path 2 from Path 1 as the Path 1 predictor increased, and completely switching to Path 2 when the predictor reached 1. The solutions leaned towards Path 2 earlier in case of lower available bandwidths. Figures 10 and 11 show the near-congestion cases of available bandwidth 12.5k for both paths. The transmission used combinations of bitrates on both paths, especially up to 77% Mode 4 (low bitrate but relatively high quality) under no loss (Figure 10). The target quality was chosen to be 3.5 (`Acceptable') in no loss, but was reduced to 3 (`Fair') in case of loss. In the case of loss (Figure 11), Mode 1 was
chosen in a greater ratio (up to 48%) due to its fair quality maintenance under loss. The optimal quality was maintained above the target qualities in all cases. Thus the solutions matched what was expected intuitively, verifying the validity of the model.

JOINT RATE-ERROR CONTROL EVALUATION RESULTS

We evaluated the Joint Rate>Error Control for single and two-path scenarios by using both simulation and Internet experiments. The purpose was to evaluate its efficiency in terms of optimal quality and bandwidth maintenance under a wide range of packet loss scenarios, and to compare our mechanism with other similar techniques. The factors to consider in measuring the efficiency of the mechanism are the quality improvement, minus the bandwidth and delay costs incurred due to the application of error control and path diversity. We define the rewards of Error Control as follows. For single path, the cost is primarily determined by the bandwidth cost as the most significant detrimental factor. Hence we define the single-path Reward metric as a combination of two factors having disparate measurement units. The quality improvement ratio is a non-negative fraction, with values 0 when both qualities are equal and close to 1 when quality is significantly improved with FEC. The FEC overcode is kept less or equal 1 by choosing amount of redundancy that is never greater than the number of original packets (i.e. \( \frac{n - k}{k} \leq 1 \)). For two-path, on the other hand, the cost includes the extra delay incurred by path switching. The two-path Reward metric is thus

\[
\beta = \frac{Q_f - Q_n}{Q_f} - \delta
\]

where \( Q_f = \) Quality after applying error control, \( Q_n = \) Quality with no error control, and \( \delta = \) FEC overcode. We choose relative amounts in order to express \( \beta \) as a ratio of two factors having disparate measurement units. The quality improvement ratio is a non-negative fraction, with values 0 when both qualities are equal and close to 1 when quality is significantly improved with FEC. The FEC overcode is kept less or equal 1 by choosing amount of redundancy that is never greater than the number of original packets (i.e. \( \frac{n - k}{k} \leq 1 \)).

For single path, RS(10,k) was chosen as the Dynamic FEC scheme simulating a short delay budget. Dynamic FEC was compared with two static schemes of degrees RS(10,9) and RS(10,8) (with low overcodes 11% and 25% respectively) that provided acceptable quality, and were more bandwidth friendly than used in (Jiang, W. & Schulzrinne, H., 2002) (RS(3,2), RS(2,1) and RS(4,3) with 50%, 100% and 33% overcodes) and (Nguyen, T. & Zakhor, A., 2003) (RS(30,23) with 30% overcode and longer delay budget). For two-path, we compared the following: (i) when only Rate Control was used, but no Error Control on either path (Diversity-only), (ii) when Error Control was used on the first path, but no Rate Control (Dynamic-FEC-only), (iii) a combination of both Error and Rate Control when dynamic FEC was infeasible on first path (Diversity, Dynamic-FEC), and lastly (iv) when neither was used. We also show the effectiveness of our mechanism compared to the adaptive path-switching (Diversity-only) algorithm proposed by (Tao, S., Xu, K., Estepa, A. et. al., 2005).

For simulation, we created a n-hop single path topology using ns-2, where a CBR stream flowed from sender to receiver through intermediate Droptail links. The capacity of the links was varied between 760kbps and 10Mbps. Pareto (shape = 1.9 for infinite variance) and TCP cross traffics of various packet sizes (100 bytes to 1000 bytes) and rates (200k up to 9M) were randomly injected through the intermediate links from 10 sources per intermediate link, causing variable degrees of congestion resulting in different degrees of stream packet loss at the intermediate links. The end-to-end one-way-delay was varied between high values of 250ms and 500ms in order to have observable variations. In two-path topology, each path was an n-hop single path. We collected a wide dataset exhibiting loss rates from 1% to 20% and variable loss bursts degrees. We fed this data to the Rate/Error Control (developed using C and GLPK) that generated the proactive FEC degrees and durations for single path and FEC-dispersion combination for two paths. For simplicity, a single bitrate G.722.2 Mode 7 was chosen. The decision rule parameters \( z \) and \( w \) were chosen to be 3 (‘Fair’ MOS) and 2 (‘Poor’ MOS) respectively.
For Internet experiment, we conducted two sets of experiments on PlanetLab (http://www.planet-lab.org/), a wide-area network testbed consisting of sites located around the Internet. In Experiment Set I, a "Master" program from a US site sent a 64k stream of 3 minute speech segment to 4 US and 6 international sites. The stream rate in Set I was chosen to be considerably low in order to generate low degrees of loss (1% to 3%). In Set II the send rate was varied from 204kto 1M, in order to create greater congestion and higher loss burst scenarios (upto 20%). We selected alternate paths to every destination by selecting intermediate hops from PlanetLab nodes, and ensured that the one-way-delay was comparable to the original path, and sent same stream data on original and alternate paths simultaneously. We ran Set I every 4 hours for 12 days, and Set II for 12 days, with a total of 864 datasets. Loss occurred in 131 datasets for Set I and 180 primary datasets for Set II, and were selected for our analysis.

**Rate-Error Control on Single Path**

Figure 12: Single Path Simulation: Example Dynamic FEC.

Figure 13: Single Path Simulation: Quality Maintenance by Dynamic FEC when expected quality falls under $\sigma$.

Figure 14: Single Path Simulation: Comparison of Dynamic FEC with other Schemes: Quality comparison of average MOS.

Figure 15: Single Path Internet Experiments: Average and Minimum MOS of Dynamic FEC compared to No FEC.

Figure 16: Single Path Dynamic FEC over static FEC Schemes: Bandwidth Improvement in Simulation.

Figure 17: Single Path Dynamic FEC over static FEC Schemes: Bandwidth Improvement in Internet Experiments.

Figure 18: Single Path Dynamic FEC over static FEC Schemes: Reward Comparison in Simulation.

**Simulation**: Figure 12 shows the one-way-delay, predictor value, and the dynamic FEC in an example simulation scenario with a low loss degree of 3%. The Loss Predictor is consistently in bad state before and around packet loss occurrences. Based on Predictor and Loss Monitor feedbacks, the Error Control selected FEC degrees to be 1 or 2 (recovering single losses and loss burst of 2 respectively), and FEC durations. The dynamic FEC improved the quality in the range of 264% to 368% over no FEC (Figure 13). It should be noticed that the quality under dynamic FEC was consistently maintained around 3.7 (above 'Acceptable' MOS 3.5). For the whole experiment set, we noticed that the average MOS for no FEC significantly degraded as loss rate increased (Figure 14). In contrast, the average MOS with Dynamic FEC maintained optimal values above 3.5. The Dynamic FEC quality was consistently better than RS(10,9), and minimally less than RS(10,8). The bandwidth consumption of Dynamic FEC showed significant improvement over bandwidth-friendly RS(10,8) and RS(10,9) in the complementary cumulative frequency distribution of the bandwidth improvement in Figure 16. Dynamic FEC showed more than 70% improvement over RS(10,8) in 100% cases, and more than 35% improvement over RS(10,9) in 80% cases. Average bandwidth overcode was maintained well below RS(10,9) (Figure 14). As a combined metric of quality and bandwidth cost, the Reward showed superiority of Dynamic FEC over static FEC schemes (Figure 18), where 92% cases maintained a reward of 0.75 or higher.
**Internet Experiments:** We used low loss (1% to 3%) datasets from Experiment Set I. Figure 15 shows the average and minimum MOS values of Dynamic FEC and no FEC for 9 sets of data from Internet Experiment set I. The quality for Dynamic FEC shows low variation, where the average MOS and the minimum MOS are close to or above ‘Acceptable’ MOS 3.5 and ‘Fair’ MOS (3.0 or $\Xi$) respectively in most cases. In contrast, the quality for no FEC varies significantly, the average being mostly below ‘Acceptable’. Similar to simulation, Dynamic FEC showed bandwidth improvement over both RS(10,8) and RS(10,9) (Figure 17). This is also reflected in the Reward comparison in Figure 19. Interestingly, RS(10,9) performed comparable to Dynamic FEC in both quality and bandwidth consumption, and thus $\Xi$ of Dynamic FEC is only marginally better than RS(10,9). Under low loss situations of Experiment Set I, most of the losses are single bursts, triggering the choice of degree by dynamic FEC to be RS(10,9) in most cases.

**Rate-Error Control on Two Paths**

Figure 20: Two-path: Bursty Loss, Dynamic FEC and Diversity Indicator on a Primary Path in Simulation.

Figure 21: Example Loss Patterns observed on the Internet.

Figure 22: Two-path - Quality Comparison of Four Schemes: Simulation.

Figure 23: Two-path - Quality Comparison of Four Schemes: Internet Experiments.

**Simulation:** We introduced heavy cross-traffic in the first path causing loss burst of high degrees, and used a second path with moderate loss rate of 3% as the diversity path when FEC on the first path was not feasible. We created loss degrees from moderate 5% to high 20% on the primary path, with moderate (2%) to high (15%) degree of burstiness, measured as the percentage of loss bursts greater than 4 packets. In an example scenario (Figure 20), the loss burst lengths exceeded 10 packets a number of times. The Error Control chose RS(10,5) as the maximum feasible FEC (5 redundant for 5 original, FEC overcode 1) to cater to the burstiness of the observed loss on this path. Whenever FEC was infeasible on the first path, Rate Control diverted the packets to the second path. It should be noted that such high burstiness, though not common, is observed in the Internet (Markopoulou, A. P., Tobagi, F. A. & Karam, M. J., 2002). An example dataset from our experiment Set II showed both high (burst lengths greater than 20, oval 2) and moderate loss bursts (burst lengths between 10 and 20, ovals 1 and 3) during the same session (Figure 21).

In Figure 22 the average combined quality using (Diversity, Dynamic FEC) was considerably better than all other mechanisms, with 100% of results showing average quality greater than or equal to 3.7. (Diversity, Dynamic-FEC) produced better quality than Diversity-only by rectifying feasible loss bursts on both paths using Dynamic FEC. It is interesting note that the Diversity-only mechanism showed better quality than Dynamic-FEC-only. This is due to the fact that high burst degrees, irreparable by the FEC-only scheme, occurred more frequently than low burst degrees. In Figure 26 FEC-only performed better than Diversity-only in the case of loss with low to medium (upto 5%) burstiness, where 100% cases for FEC-only had quality greater than or equal to 3.5. Clearly, FEC can rectify losses and improve quality if the burstiness is at a lower degree, otherwise diversity is a better choice.
Next we compared our system with the adaptive path switching algorithm proposed in (Tao, S., Xu, K., Estepa, A. et al., 2005), which we refer to as APS. APS measures actual quality of the paths based on the network loss and delay conditions observed so far, and switches to an alternative path whenever the alternative path quality exceeds the original path quality by a constant value 0.1. In Figure 24 our mechanism showed superiority in Reward (\( \bar{R} \)). The reasons are two-fold: first, we measure *projected* quality based on Loss Predictor state over next window and take proactive FEC action, which results in much higher degree of quality improvement, even compensating for the FEC bandwidth cost that APS does not have. Second, since APS is solely based on path switching, the number of path switches is higher than ours (Figure 25) that incurs more delay overhead and reduces the Reward for APS.

**Internet Experiments:** From Set II we selected data that exhibited comparatively higher loss rates and degrees of burstiness. We selected the corresponding alternate path with low degree of loss. In Fig 23, the combined (Diversity, Dynamic FEC) showed better average quality than Dynamic-FEC-only, and was comparable to Diversity-only. Under high burst degrees, the path diversity was chosen more often, as FEC was considered to be infeasible by Error Control. The burst losses observed in our Internet experiments were mostly of high degree (\( \mathcal{E} \leq 5 \)), as opposed to the simulation data that exhibited a fair mix of bursts of both higher and lower degrees. In Figure 27 Rate-Error control had less number of path switches than APS, even though it provided mostly diversity comparable to APS. Our mechanism takes path diversity action only when FEC is infeasible on primary path, thus avoiding some ‘unnecessary’ path switches made by APS.

**CONCLUSION AND FUTURE WORK**

We present a joint Proactive Rate and Error Control framework based on packet loss prediction and audio quality assessment. The Error Control chooses dynamic FEC using an MDP and a stochastic inventory control based on loss prediction, a novel approach for multimedia error recovery. The Rate Control maximizes receiver audio quality by determining the optimal rate dispersion over single or multiple paths using a rate-quality optimization model. Our self-adaptive mechanism combines the benefits of both FEC and packet dispersion to ensure a bandwidth-friendly transmission with effective degree of loss recovery. The proactive and dynamic nature of various components makes the framework superior to reactive mechanisms that use ad-hoc FEC and static feedbacks.

The joint error and rate control, utilizing the Loss Predictor and Audio Genome, has shown promising results in simulation and using Internet data. Dynamic FEC was superior to other similar techniques in single paths in both target quality maintenance and bandwidth consumption. It maintained a reward of 0.75 or higher in 92% cases, and showed 70% improvement over similar techniques in bandwidth consumption in 100% cases. In multiple paths, (Diversity, Dynamic-FEC) maintained average quality greater than or equal to 3.7 in 100% cases and performed superior to Dynamic-FEC-only and Diversity-only schemes. It is in our future work to prove its viability for the real Internet. We also will extend the framework to video.

**REFERENCES**


Cox, R. & Perkins, R. (1999). Results of a Subjective Listening Test for G.711 with Frame Erasure Concealment, Committee contribution T1A1.7/99-016


GNU Linear Programming Kit: http://www.gnu.org/software/glpk/


ITU-T Recommendation G.722.2. (2002). Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)


ITU-T Recommendation P.862. (2001). Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs


