TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 15/16 pp320-331
Volume 16, Number 3, June 2011

M3FEC: Joint Multiple Description Coding and Forward Error Correction for Interactive Multimedia in Multiple Path Transmission*

ZHANG Hui (张 辉)[1,2], JIANG Wei (江 伟)[1], ZHOU Jin (周 晋)[2],
CHEN Zhen (陈 震)[2,3], LI Jun (李 军)[2,3,**]

1. Department of Automation, Tsinghua University, Beijing 100084, China;
2. Research Institute of Information Technology, Tsinghua University, Beijing 100084, China;
3. Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China

Abstract: The best-effort internet has inherent limitations on the end-to-end performance for interactive multimedia communications. This paper presents a multiple description coding (MDC) and forward error correction (FEC) based multiple path transmission schemes for interactive multimedia (M3FEC), which improves the end users' experience by maximizing a rate-distortion (R-D) based optimization problem. The proposed model considers both the network diversity and the application's stringent requirements, and combines the individual merits of the three promising technologies of multiple path overlay routing, MDC and FEC. Extensive numerical analysis and PlanetLab experiments demonstrate that M3FEC successfully combats packet losses, error propagation, and unpredictable network dynamics. This method also significantly increases distortion for interactive multimedia by over 10 dB than traditional IP-layer single path transmission in poor network environments, and outperforms the performance achieved by using MDC or FEC alone.

Key words: interactive multimedia; multiple path overlay routing; multiple description coding (MDC); forward error correction (FEC); rate-distortion (R-D)

Introduction

With recent advances in multimedia and network technologies, various interactive multimedia applications are becoming more popular among internet users, including video calling, instant messaging, massive multiplayer online games, distance learning and telemedicine. Interactive multimedia applications demand uninterrupted bandwidth with rigid bounds on packet loss and end-to-end delays to achieve a minimally acceptable quality. However, current predominant internet protocols and mechanisms do not favor delay sensitive applications but are developed for data transmission and only provide best-effort services. Thus new methods are needed for timely and reliable delivery of interactive multimedia.

This paper presents a reliable interactive multimedia transmission scheme over lossy packet networks such as the internet by combining the following three popular technologies: (1) network-level technology of multiple overlay path routing; (2) source coding technology of multiple description coding (MDC); (3) channel coding technology of forward error correction (FEC).

The emergence of overlay techniques[1] triggers research on sending packets simultaneously over multiple paths by routing traffic through intelligent relay nodes at strategic locations in the internet without the physical network support. The benefits of multiple paths with fine diversity include improved fault tolerance and link recovery[2,3], reduced delay variance[4],

Received: 2009-06-29; revised: 2011-03-28
* Supported by the National Natural Science Foundation of China (No. 90718040) and NEC Laboratories China (No. LC-2008-055)
** To whom correspondence should be addressed.
E-mail: junl@tsinghua.edu.cn; Tel: 86-10-62796400

the conversion of burst packet losses to isolated packet losses[4], the provision of larger aggregate bandwidth[5,6] and load balancing[7]. These improvements provide some interesting benefits for interactive multimedia to temper the unpredictability and time-varying nature of the best-effort internet. Many approaches have also been proposed to find appropriate relay nodes to increase path diversity using topology heuristics, such as earliest divergence rule (EDR)[8], disjoint multi-path QoS routing (DMQR)[9], and widest disjoint paths (WDP)[10]. Their basic idea is to establish overlay paths which take different underlying physical routes from the default IP-layer path between the sender and receiver.

MDC is a source coding approach which generates several independent descriptions each with its own prediction process and state. These descriptions deliver a basic quality when they are individually decoded and each additional description further refines the quality. MDC works particularly well with multiple path transmission, in which the different descriptions are explicitly sent over different routes. Hence, as long as losses do not occur simultaneously in every path, the receiver is guaranteed an acceptable media quality. MDC has been shown to be effective in coping with burst packet losses and subsequent error propagation among video frames[2,11]. Consequently, MDC is envisioned as a promising solution for interactive multimedia applications.

FEC is a channel-level error correction coding method, which enables the receiver to detect and correct errors without interactions by adding redundant information at the sender side. Compared with another widely used error correction method retransmission, FEC reduces the time needed to recover the missing packets at the expense of bandwidth consumption by the redundant packets. Therefore, this technique is quite attractive for interactive multimedia applications.

Various studies have applied MDC or FEC to multimedia applications. Apostolopoulos[2] used path diversity along with MDC to improve the reliability of streamed media, and showed how an erasure in a description could be recovered by other descriptions as long as the errors did not occur simultaneously. CoopNet[12] introduced MDC into the P2P stream by using MD-FEC to code the source into several substreams and building multiple multicast trees from the sources to the receivers, with each tree disseminating a separate description of the media content. Apostolopoulos et al.[11] improved the performance of streaming by exploiting the path diversity provided by existing CDN infrastructure and MDC. Their results showed that MDC requires about 50% fewer CDN servers than conventional streaming techniques to achieve the same distortion at the clients. Liu et al.[13] proposed a multistream coding and transmission scheme, redundancy free multiple description (RFMD) coding, specifically designed for P2P VoD systems. Begen et al.[14] developed an MDC over multiple paths model for interactive multimedia, simulation results with MPEG-2 and NS-2 show that PSNR improvements ranging from 0.73 to 6.07 dB can be achieved. Yu et al.[15] proposed and analyzed a joint MDC and FEC coding approach for interactive multimedia applications, which was quite instructive for our research. However, they assumed that the source was generated from Gaussian variables and gave only some numerical simulations.

Inspired by the above work, this paper introduces MDC and FEC into interactive multimedia communications in a path diversity environment to make the coding scheme more adaptive to the network environment and to efficiently exploit existing network resources, thus providing a more pleasant viewing experience for the end users. The major contributions of this paper are twofold. First is an explicit rate-distortion (R-D) based framework M3FEC which provides joint MDC and FEC transmissions with optimized coding and transmission parameters for the given network conditions. Second is an analysis of the effectiveness and reliability of M3FEC for interactive multimedia by comparisons with several other transmission schemes such as MDC, FEC, and single path.

1 Preliminaries

1.1 Average R-D in MDC

R-D theory was created by Shannon in his foundational work on information theory[16]. In this theory, the rate R, denotes the number of bits per data sample to be stored or transmitted after source encoding, while the distortion D, is defined as the variance of the difference between the input and output signals (i.e., the mean squared error, MSE). Let $\{X_k\}_{k=1}^{M}$ denote a sequence of source symbols input to the source encoder and $\{\hat{X}_k\}_{k=1}^{M}$ denote the reconstructed signal after source decoding. Then the average distortion (MSE) will be

$$D = \mathrm{E}\left\{\frac{1}{N}\sum_{k=1}^{N} d(X_k, \hat{X}_k)^2\right\} \tag{1}$$

where $d(,)$ is the squared Euclidean norm between X_k and $\hat{X}_k$.

The reconstructed media quality, expected as the average distortion, is based on a recent analytic model[17], which fits scalable video codecs such as MPEG-4 FGS quite well. The expression of this cited model is

$$D(R) = 2^{aR+b\sqrt{R}+c} \tag{2}$$

where a, b, and c are parameters. For offline-encoded video sequences, these parameters can be pre-computed. For real-time encoding, initial values are assumed according to the characteristics of the communication scenarios, with these values are refined as more frames are encoded.

Consider the illustrative R-D curves in Fig. 1. The actual R-D curves were produced by encoding the CIF format test sequences Foreman and Coastguard by a standard MPEG-4 encoder FFmpeg[18] with the default parameters. These empirical results demonstrate that the R-D model of Eq. (1) fits the actual R-D curves quite well and its accuracy is sufficient for our purposes.

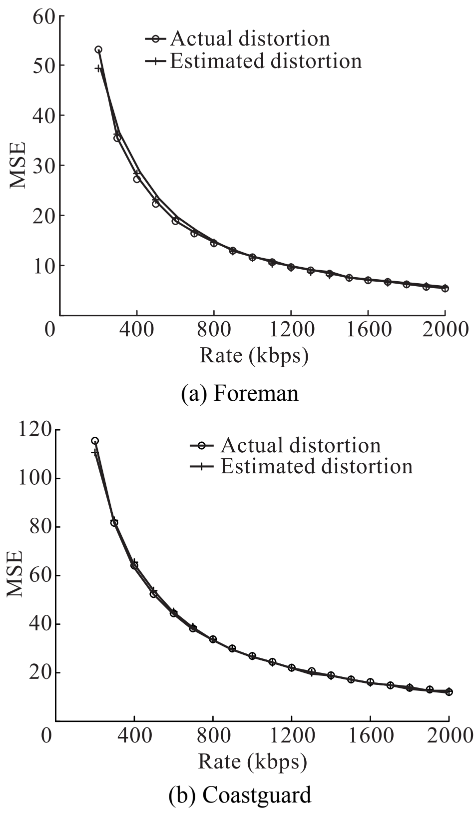


Fig. 1 Actual and estimated R-D curves

MDC was invented by Bell Laboratories in the 1970s for speech communication over the telephone. There are now various techniques for generating multiple descriptions. This paper uses the straightforward time-domain partitioning method with two descriptions, which separates even and odd numbered frames of a sequence into two groups and encodes them individually.

The choice of two descriptions is due to the following considerations. The delay variance measured in terms of the standard deviation decreases with $1/\sqrt{N}$ [2], where N is the number of independent paths. Therefore, there are diminishing improvements with increasing number of paths. In a real internet environment, typically a small number of paths should provide a good balance between diverse path finding complexity and performance achievements. As an example, the choice of two paths yields a relatively simple solution and also straightforwardly couples to two MDC descriptions, where each independently decodable stream is sent over a separate path.

Two description scenarios have four possible cases at the receiver side determined by which description arrives and whether it is on-time. Let D_{00} (D_{11}) denote the distortion when both descriptions arrive intact and on time (are lost or delayed). Similarly, let D_{01} (D_{10}) denote the distortion when the first (second) description arrives intact and on time, but the other is lost or delayed. In this case, the missing description is concealed with the help of the received description. Thus, the expected distortion will be

$$D = p_{00}D_{00} + p_{01}D_{01} + p_{10}D_{10} + p_{11}D_{11} \tag{3}$$

where p_{00}-p_{11} denote the corresponding probabilities of achieving the distortion for D_{00}-D_{11} which will be derived in Section 1.2.

If both descriptions arrive intact and on time, the distortion will be the average of the individual description, then

$$D_{00} = \frac{1}{2}\times 2^{aR_1+b\sqrt{R_1}+c} + \frac{1}{2}\times 2^{aR_2+b\sqrt{R_2}+c} \tag{4}$$

where a, b, and c are the constants in Eq. (2) and R_i denotes the rate in terms of the number of bits per source sample for the corresponding description.

The expressions for D_{01} (D_{10}) in this paper are referred to in previous work[14]. When only one description is successfully received, the missing frames can be reconstructed using the received frames. However, their qualities will be lowered because of imperfections in the reconstruction. Begen et al.[14] used a scaling factor, γ, to decide the distortion increases and express D_{01} as

$$D_{01}=\frac{1}{2}\times 2^{aR_1+b\sqrt{R_1}+c}+\frac{1}{2}\times\gamma\times 2^{aR_1+b\sqrt{R_1}+c} \tag{5}$$

Their experiments indicated that this expression provides a reasonable estimation of the actual distortion. For interactive applications (such as video telephony), a default value of 1.5 reasonably reflects most application characteristics.

The final distortion term in Eq. (3), D_{11}, is calculated by assuming an average value for each pixel and then computing the expected error.

1.2 FEC packet loss characteristics

The reed-solomon (RS) code is the most widely used FEC coding method for data storage on CDs and DVDs and for data transmission such as DSL and WiMAX. $\mathrm{RS}(n,k)$ represents k packets of source data encoded at the sender to produce $n-k$ redundant packets in such a way that any subset of k packets in the total n packets suffices to reconstruct the source data. This code allows the receiver to recover from up to $n-k$ losses in a group of n encoded packets. The irrecoverable loss probability with $\mathrm{RS}(n,k)$ coding for Bernoulli packet loss where packet losses are independent and identically follows formula (6)[19] is

$$p_F=\frac{1}{k}E[Y]=p-\sum_{i=1}^{n-k}\binom{n-1}{n-k-i}(1-p)^{k+i-1}p^{n-k-i+1} \tag{6}$$

where p is the average packet loss ratio along the path.

However, packet losses are known to be correlated and often occur in bursts[20] and the packet loss burst has significant impact on the FEC performance. Several models have been developed to capture the stochastic characteristics of the underlying packet loss process, including the Gilbert model[20], the general k order Markov Chain[21], and the hidden Markov model[22]. Compared with the later two, the Gilbert model gives an acceptable level of accuracy and yet its computational complexity is relatively lower.

This paper adopts the Gilbert model, which can be represented by a discrete-time two-state Markov chain where the current state of the stochastic process depends only on the previous value. The state diagram of the Gilbert model is shown in Fig. 2. If the process is in the Good state G, the receiver receives all packets, while if the process is in the Bad state B, all packets are lost. The process can be characterized by the two probabilities of the chain changes from Good to Bad (p_{GB}) and vice versa (p_{BG}), or by the two parameters of average loss ratio p and the packet correlation ρ, where

$$\begin{aligned}p_{\mathrm{GB}}&=p(1-\rho),\\ p_{\mathrm{BG}}&=(1-p)(1-\rho)\end{aligned} \tag{7}$$

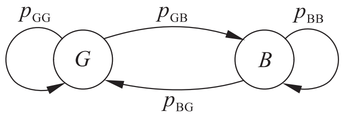


Fig. 2 State diagram of the Gilbert model

The steady state probabilities of the related Gilbert model can be expressed as $\pi(G)=\dfrac{p_{\mathrm{BG}}}{p_{\mathrm{GB}}+p_{\mathrm{BG}}}=1-p$ and $\pi(B)=\dfrac{p_{\mathrm{GB}}}{p_{\mathrm{GB}}+p_{\mathrm{BG}}}=p$ respectively. The packet correlation ρ provides an average measure of how the states of two consecutive packets are correlated to each other. In particular, if $\rho=0$, the loss process is memory less (Bernoulli process), and when the value of ρ increases, the states of two consecutive packets become more and more correlated.

To analyze the FEC effects on the Gilbert packet loss, this paper borrows the framework given by Wu and Radha[23], which presents a rather elegant approach for evaluating the probability of receiving i packets among n packets transmitted over Gilbert erasure channels. They constructed another Markov process by extending the two-state Gilbert model using the number of correctly received packets as the indexes for the states of the extended Markov process. For example, if the receiver correctly receives i packets and is in a good state, then the process is in state G_i, while if the receiver correctly receives i packets and is in a bad state, then the process is in state B_i as shown in Fig. 3.

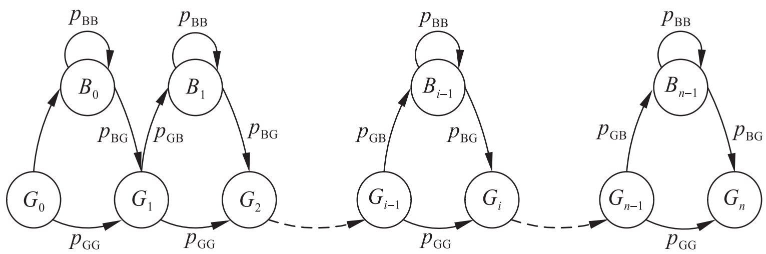


Fig. 3 State diagram of the extended Gilbert model

It should be pointed out that the extended Markov process can start from G_0 or B_0 and end in any state. The initial state probabilities G_0 and B_0 are the steady state probabilities of the original Gilbert model $\pi(G)$ and $\pi(B)$, respectively. Then, the probability that the sender transmits n packets and the receiver correctly receives i packets is equivalent to the probability that the process starts at G_0 or B_0 and ends in state G_i and B_i after n stages. Let $\phi_{G_0G_i}(n)$ represent the multistep transition probability from G_0 to G_i, then

$$\phi_{G_0G_i}(n) = P\{s(n) = G_i \mid s(0) = G_0\} \tag{8}$$

where $s(j)$ is the state of the extended Markov process at time index j. Wu and Radha[23] prove that:

$$\phi_{G_0G_i}(n) = \begin{cases} 0, & i=0; \\ \sum_{m=1}^{\min(i,n-i)} \binom{i}{m}\binom{n-i-1}{m-1} p_{GB}^{m} p_{BG}^{m} p_{GG}^{i-m} p_{BB}^{n-i-m}, & 0<i<n; \\ p_{GG}^{n}, & i=n \end{cases} \tag{9}$$

$$\phi_{G_0B_i}(n) = \begin{cases} \sum_{m=0}^{\min(i,n-i-1)} \binom{i}{m}\binom{n-i-1}{m} p_{GB}^{m+1} p_{BG}^{m} p_{GG}^{i-m} p_{BB}^{n-i-m-1}, & 0\leqslant i<n; \\ 0, & i=n \end{cases} \tag{10}$$

$$\phi_{B_0G_i}(n) = \begin{cases} \sum_{m=0}^{\min(i-1,n-i)} \binom{i-1}{m}\binom{n-i}{m} p_{GB}^{m} p_{BG}^{m+1} p_{GG}^{i-m-1} p_{BB}^{n-i-m}, & 0<i\leqslant n; \\ 0 & i=0 \end{cases} \tag{11}$$

$$\phi_{B_0B_i}(n) = \begin{cases} P_{BB}^{n}, & i=0; \\ \sum_{m=0}^{\min(i-1,n-i-1)} \binom{i-1}{m}\binom{n-i}{m+1} p_{GB}^{m+1} p_{BG}^{m+1} p_{GG}^{i-m-1} p_{BB}^{n-i-m-1}, & 0<i<n; \\ 0, & i=n \end{cases} \tag{12}$$

Let $\phi(n,i)$ represent the probability that the sender transmits n packets and the receiver correctly receives i packets, then

$$\phi(n,i) = \pi(G)(\phi_{G_0G_i}(n) + \phi_{G_0B_i}(n)) + \pi(B)(\phi_{B_0G_i}(n) + \phi_{B_0B_i}(n)) \tag{13}$$

Consequently, the desired probability $\phi(n,i)$ can be correctly estimated by any two parameters that could characterize the underlying Gilbert process. Based on the above theory frameworks, this paper defines a random variable $Y (0 \leqslant Y \leqslant k)$ to be the number of data packets that could not be recovered in each FEC block. Then,

$$P[Y=y] = \begin{cases} \sum_{i=k}^{n} \phi(n,i), & y=0; \\ \sum_{i=0}^{\min(y-1,n-k)} \{\pi(G)[\phi_{G_0G_{k-y}}(k)(\phi_{G_0G_i}(n-k) + \phi_{G_0B_i}(n-k)) + \phi_{G_0B_{k-y}}(k)(\phi_{B_0G_i}(n-k) + \phi_{B_0B_i}(n-k))] + \\ \quad \pi(B)[\phi_{B_0G_{k-y}}(k)(\phi_{G_0G_i}(n-k) + \phi_{G_0B_i}(n-k)) + \phi_{B_0B_{k-y}}(k)(\phi_{B_0G_i}(n-k) + \phi_{B_0B_i}(n-k))]\}, & 0<y\leqslant k \end{cases} \tag{14}$$

This expression is the summation of the probability that the receiver successfully receives $k-y$ original packets and less than $y-1$ redundant packets. Therefore, the irrecoverable packet error probability with $\mathrm{RS}(n,k)$ will be

$$p_F = \frac{1}{k}E[Y] = \frac{1}{k}\sum_{y=0}^{k} yP[Y=y] \tag{15}$$

2 Assumptions and Modeling

2.1 Assumptions and notations

This section begins the expatiation of M3FEC with a set of assumptions.

(1) The original real-time streams for the same pair of nodes are generated and encoded at a uniform speed across time.

(2) The encoded source is packetized into the same size packets with packet along the same path sent in time slots of fixed duration.

(3) Various multiple path overlay routing algorithms[9,23] are used by a pair of communication nodes to select two paths with excellent diversity and delay quality. This analysis then only investigates the coding and transmission scheme between them.

(4) The loss and delay varying behaviors of the two selected paths are approximate and uncorrelated. Note that this hypothesis is not always true but highly depends on the degree of physical link disjointness.

Table 1 summarizes the key notation for this analysis. Subscript i represents the parameters for the i-th path, channel or description. The path parameters have no subscript when using single path transmission only or the parameters for all paths are the same.

Table 1 Notation and meaning

Notation	Meaning	Default value or unit
p_i	Original packet loss ratio along the i-th path	n/a
ρ_i	Correlation between two consecutive packets along the i-th path	n/a
p_{GB_i}, p_{BG_i}	Probability of good/bad state transformation to bad/good state for the i-th path	n/a
$\pi_i(G), \pi_i(B)$	Steady state probabilities for the Gilbert process along path i	n/a
$\phi_{G_0G_j}(n)$	Multistep transition probability from G_0 to G_j after n stages	n/a
$\mathrm{RS}(n_i, k_i)$	Reed-Solomon code parameter of the i-th path (k_i original data packets and $n_i - k_i$ redundant packets)	n/a
p_{Fi}	Packet loss ratio for the i-th path with FEC, which is a function of p_i, ρ_i and the FEC protection level	n/a
R_i	Source coding rate in terms of the number of bits per source symbol for the i-th path	bits/pixel
D	Average distortion	n/a
a, b, c	Parameters in R-D model, refer to Eq. (2)	n/a
p_{uv}	$u/v = 0/1$, where u and v represent the 1st and the 2nd paths. 0 denotes packets successfully received and 1 denotes packets lost or delayed, refer to Eq. (3)	n/a
γ	Scaling factor in MDC decoding to decide the distortion increases	n/a
B_{a}	Maximum allowable bandwidth	kbps
B_i	Actual bitrates allocated on each path	kbps
$W \times H$	Video frame resolution (in pixels)	352×288 pixel/frame
F	Frame rate	25 frames/s
C	Known constant that depends on the chroma sub-sampling format $B_i = R_i \times (W \times H \times F \times C)$	1.5 (for 4:2:0 YUV video)
t_i	One way end-to-end delay for the i-th path	100 ms
t_{r}	One way interactive application delay requirement	300 ms
N	Number of communication paths	2 or 1

2.2 Constraints

This section presents several constraints that make the model better satisfy the actual interactive multimedia requirements and tolerate the Internet's time-varying properties.

First, interactive multimedia applications have stringent end-to-end delay requirements, with packets arriving after their decoding or play-out deadline being useless as shown in Fig. 4. In FEC coding, the $n-k$ redundant packets can only be computed after the k original packets have been generated, but this delay cannot be ignored in real-time applications. Let t_f denote the time when the first packet in an FEC block is being generated by the sender and t_l denote the time when the last packet in the same FEC block is received. To guarantee that each packet in the same FEC block is received on time, it requires $t_\mathrm{l}-t_\mathrm{f} \leqslant t_\mathrm{r}$. According to assumptions (1) and (2) in Section 2.1, this inequality is equivalent to

$$\frac{k_i}{B_i/l_i}\left(1+\frac{n_i-k_i}{n_i}\right)+t_i \leqslant t_\mathrm{r} \tag{16}$$

where $B_i = R_i \times (W \times H \times F \times C)$ is the bandwidth consumption on the i-th path after source coding and before channel coding, l_i is the length of the payload in each packet, and t_i is the one way end-to-end delay for the i-th path. Since t_i varies in time, the one way delay requirement t_r is set to 300 ms instead of 400 ms declared in the ITU standard G.114[24] to handle the internet delay dynamics.

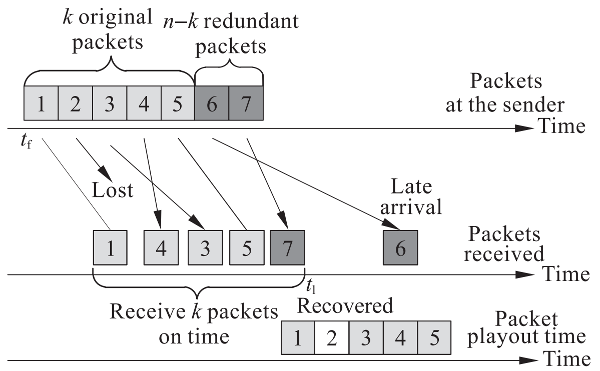


Fig. 4 Timing constraints and packet loss with FEC

Secondly, it is desired that different transmission schemes consume similar bandwidth resources for fairness performance comparisons. Thus the proposed model requires the following inequality:

$$\sum_{i=1}^{N}(h+l_i)(B_i/l_i)\frac{n_i}{k_i} \leqslant B_\mathrm{a} \tag{17}$$

where h is the header length in each packet, including the RTP header, UDP header, and IP header, and B_a is the maximum bandwidth that can be consumed on all the paths.

Thirdly, considering the MTU limitation and some routers assign a low priority for large packets, the packet size is limited to

$$h+l_i \leqslant l_{\max} \tag{18}$$

2.3 M3FEC model

The M3FEC transmission scheme described here is illustrated in Fig. 5. M3FEC encodes the original streams into two independent descriptions at average rates of R_1 and R_2 bits per sample. The output of each encoder is then sent to the corresponding FEC encoder. The FEC codec uses the widely used Reed-Solomon code[25]. To avoid additional delays of the encoded packets, the packets produced by the source encoder are assumed to be immediately available for transmission over the network, with a copy of each packet sent to the FEC encoder to generate the redundant packets. From the analysis in Sections 1 and 2, the M3FEC transmission can be concluded as the following optimization problem:

$$\begin{aligned}\min\Bigg\{D = {} & p_{00}\left(\frac{1}{2}\times 2^{aR_1+b\sqrt{R}_1+c}+\frac{1}{2}\times 2^{aR_2+b\sqrt{R}_2+c}\right)+p_{01}\left(\frac{1}{2}\times 2^{aR_1+b\sqrt{R}_1+c}+\frac{1}{2}\times\gamma\times 2^{aR_1+b\sqrt{R}_1+c}\right)+\\ & p_{10}\left(\frac{1}{2}\times 2^{aR_2+b\sqrt{R}_2+c}+\frac{1}{2}\times\gamma\times 2^{aR_2+b\sqrt{R}_2+c}\right)+p_{11}D_{11}\Bigg\},\end{aligned}$$

$$\text{where}\begin{cases}\dfrac{k_i}{B_i/l_i}(1+\dfrac{n_i-k_i}{n_i})+t_i \leqslant t_r, \quad \displaystyle\sum_{i=1}^{N}(h+l_i)(B_i/l_i)\frac{n_i}{k_i} \leqslant B_\mathrm{a}, \quad B_i = R_i\times(W\times H\times F\times C), \quad h+l_i \leqslant l_{\max},\\ P_i[Y=y] = \displaystyle\sum_{j=0}^{\min(y-1,n_i-k_i)}\{\pi_i(G)[\phi_{iG_0G_{k-y}}(k_i)(\phi_{iG_0G_j}(n_i-k_i)+\phi_{iG_0B_j}(n_i-k_i))+\phi_{iG_0B_{k-y}}(k_i)(\phi_{iB_0G_j}(n_i-k_i)+\phi_{iB_0B_j}(n_i-k_i))]+\\ \qquad\qquad \pi_i(B)[\phi_{iB_0G_{k-y}}(k_i)(\phi_{iG_0G_j}(n_i-k_i)+\phi_{iG_0B_j}(n_i-k_i))+\phi_{iB_0B_{k-y}}(k_i)(\phi_{iB_0G_j}(n_i-k_i)+\phi_{iB_0B_j}(n_i-k_i))]\},\\ p_{Fi} = \dfrac{1}{k_i}\displaystyle\sum_{y=1}^{k_i} yP_i[Y=y], \quad p_{11}=p_{F1}\cdot p_{F2}, \quad p_{01}=(1-p_{F1})\cdot p_{F2}, \quad p_{10}=p_{F1}(1-p_{F2}), \quad p_{00}=1-p_{11}-p_{01}-p_{10},\\ i=1,2\end{cases} \tag{19}$$

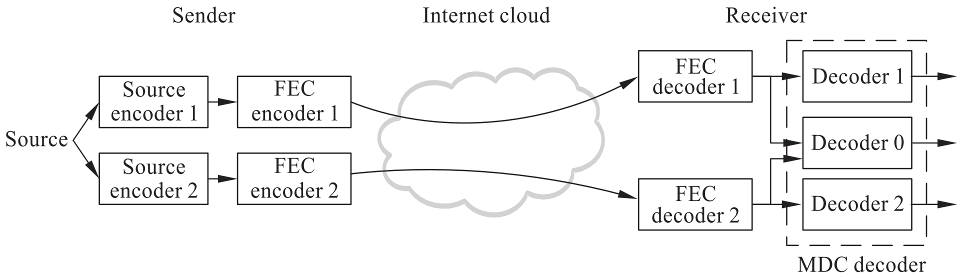


Fig. 5 M3FEC transmission scheme

where ϕ related variable could be calculated using Eq. (9) to (12). $t_i, p_{\mathrm{GB}_i}, p_{\mathrm{BG}_i}$ and B_{a} are input network parameters for the optimization. The FEC redundancy n_i and k_i, source the encoded bitrates B_i and the payload length l_i are decision variables by the transmission scheme. The expected distortion D is the optimization objective.

If $n = k = 1$, single MDC transmission scheme without FEC can be specialized as

$$\min\left\{D = p_{00}\left(\frac{1}{2}\times 2^{aR_1+b\sqrt{R_1}+c} + \frac{1}{2}\times 2^{aR_2+b\sqrt{R_2}+c}\right) + p_{01}\left(\frac{1}{2}\times 2^{aR_1+b\sqrt{R_1}+c} + \frac{1}{2}\times\gamma\times 2^{aR_1+b\sqrt{R_1}+c}\right) + p_{10}\left(\frac{1}{2}\times 2^{aR_2+b\sqrt{R_2}+c} + \frac{1}{2}\times\gamma\times 2^{aR_2+b\sqrt{R_2}+c}\right) + p_{11}D_{11}\right\} \tag{20}$$

$$\text{where}\ \begin{cases} \dfrac{1}{B_i/l_i} + t_i \leqslant t_r, \quad \displaystyle\sum_{i=1}^{N}(h+l_i)(B_i/l_i)\frac{n_i}{k_i} \leqslant B_{\mathrm{a}}, \quad B_i = R_i\times(W\times H\times F\times C), \quad h+l_i \leqslant l_{\max}, \\ p_{11} = p_1 p_2, \quad p_{01} = (1-p_1)p_2, \quad p_{10} = p_1(1-p_2), \quad p_{00} = 1-p_{11}-p_{01}-p_{10}, \quad i = 1,2. \end{cases}$$

For comparison, Fig. 6 illustrates the corresponding FEC merely with no MDC scheme over two paths. This scheme is equivalent to the previous predominant M2FEC scheme[19] and can be formulated as

$$\min\{D = (1-p_F)2^{aR+b\sqrt{R}+c} + p_F D_{11}\}$$

$$\text{where}\ \begin{cases} \dfrac{k}{\sum_{i=1}^{N} B_i/l}\left(1+\dfrac{n-k}{n}\right) + \max(t_i) \leqslant t_r, \quad (h+l)\left(\displaystyle\sum_{i=1}^{N} B_i/l\right)\frac{n}{k} \leqslant B_{\mathrm{a}}, \quad B = R\times(W\times H\times F\times C) = \displaystyle\sum_{i=1}^{N} B_i, \\ h+l \leqslant l_{\max}, \quad p_F = \dfrac{1}{k}E[Y] = \dfrac{1}{k}\displaystyle\sum_{y=1}^{k} yP[Y=y], \\ P[Y=y] = \displaystyle\sum_{i=\max(0,y-k_2)}^{\min(y,k_1)} \sum_{j=0}^{\min(y-1,(n_1-k_1)+(n_2-k_2))} \sum_{m=\max(0,j-(n_2-k_2))}^{\min(j,n_1-k_1)} \{\pi_1(G)[\phi_{1G_0G_{k_1-i}}(k_1)\bullet \\ \quad (\phi_{1G_0B_m}(n_1-k_1) + \phi_{1G_0G_m}(n_1-k_1)) + \phi_{1G_0B_{k_1-i}}(k_1)(\phi_{1B_0B_m}(n_1-k_1) + \phi_{1B_0G_m}(n_1-k_1))] + \\ \quad \pi_1(B)[\phi_{1B_0G_{k_1-i}}(k_1)(\phi_{1G_0B_m}(n_1-k_1) + \phi_{1G_0G_m}(n_1-k_1)) + \phi_{1B_0B_{k_1-i}}(k_1)(\phi_{1B_0B_m}(n_1-k_1) + \phi_{1B_0G_m}(n_1-k_1))]\}\times \\ \quad \{\pi_2(G)[\phi_{2G_0G_{k_2+i-y}}(k_2)(\phi_{2G_0B_{j-m}}(n_2-k_2) + \phi_{2G_0G_{j-m}}(n_2-k_2)) + \phi_{2G_0B_{k2+i-y}}(k_2)(\phi_{2B_0B_{j-m}}(n_2-k_2) + \\ \quad \phi_{2B_0G_{j-m}}(n_2-k_2))] + \pi_2(B)[\phi_{2B_0G_{k_2+i-y}}(k_2)(\phi_{2G_0B_{j-m}}(n_2-k_2) + \phi_{2G_0G_{j-m}}(n_2-k_2)) + \\ \quad \phi_{2B_0B_{k_2+i-y,}}(k_2)(\phi_{2B_0B_{j-m}}(n_2-k_2) + \phi_{2B_0G_{j-m}}(n_2-k_2))]\} \end{cases} \tag{21}$$

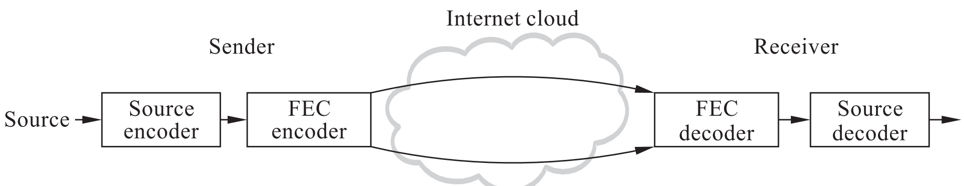


Fig. 6 FEC alone transmission scheme

If the number of paths N is equal to one and $n=k=1$, traditional single path transmission without any FEC or MDC scheme can be specialized as

$$\min\{D=(1-p)2^{aR+b\sqrt{R}+c}+pD_1\}$$

$$\text{where}\quad \begin{cases}\dfrac{1}{B/l}+t\leqslant t_{\mathrm{r}};\\(h+l)(B/l)\leqslant B_{\mathrm{a}};\\B=R\times(W\times H\times F\times C);\\h+l\leqslant l_{\max}\end{cases}\tag{22}$$

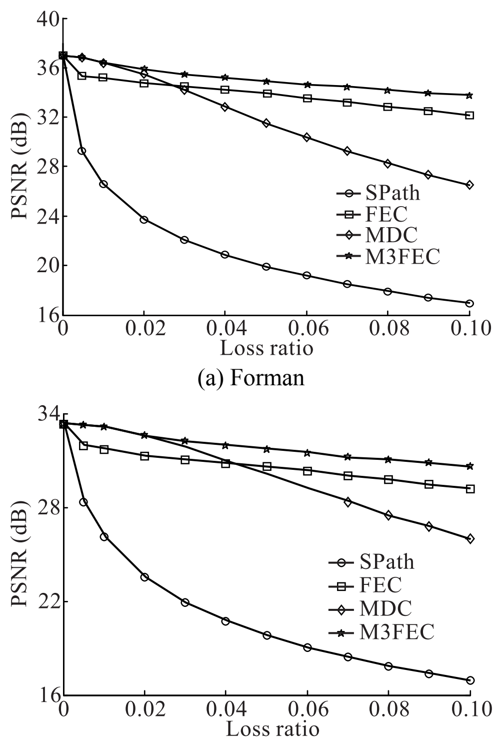
3 Evaluation

Section 2.3 models the transmission schemes of single path (SPath), FEC, MDC and M3FEC as optimization problems using Eq. (19) to (22) respectively. The ultimate purpose of this paper is to propose a joint MDC and FEC approach for interactive multimedia and to show that the joint approach M3FEC outperforms the transmission scheme using MDC or FEC alone. Therefore, this section evaluates the efficiency of M3FEC by providing quantitative answers to the following three questions. (1) Compared with the other three schemes, whether and how much M3FEC improves media quality under various network environments? (2) How well M3FEC adapts to internet dynamics? (3) How well M3FEC performs for interactive multimedia under real internet communications?

3.1 Distortion improvement

This numerical analysis set several typical network scenarios, and compares the expected media quality by solving the optimization problem proposed in Section 2.3. The distortion, D, is converted to the PSNR by $\mathrm{PSNR}=10\log_{10}\left(\dfrac{255^2}{\mathrm{MSE}}\right)$ to make the comparison results intuitively meaningful. The parameters in the R-D model (2) were acquired by fitting the two standard test sequences Forman and Coastguard. The packet loss ratio of each path ranged from 0 to 0.1 synchronously.

As shown in Fig. 7, M3FEC performs best; followed by FEC, MDC, and SPath. On the average, M3FEC improves the PSNR significantly by over 10 dB than SPath (see Table 2). As the packet loss increases, both MDC and FEC help slow the media quality degradation. When few packet losses occur (less than or equal to 2%), the MDC performance is similar to M3FEC. However, when network conditions become worse, the MDC decoder alone cannot recuperate all the packet losses or error propagation, thus its performance decreases rapidly. In contrast, FEC performs much better in poor network conditions, but is always 1 dB worse than M3FEC when packet losses occur. Intuitively, without the help of MDC, FEC tends to increase its coding redundancy to combat packet loss, resulting in a lower coding rate, R, for the same bandwidth occupation as M3FEC.



(a) Forman

(b) Coastguard

Fig. 7 Simulation curves with B_{a} =1000 kbps, t_i =100 ms and ρ_i =0.6

Table 2 Average PSNR in simulation

Sequence	B_{a} (kbps)	PSNR (dB)				
		ρ_i	SPath	FEC	MDC	M3FEC
Foreman	500	0.2	22.1	32.5	30.6	33.0
		0.6	22.1	31.4	30.6	32.1
	1000	0.2	22.5	35.6	32.2	36.1
		0.6	22.5	34.2	32.2	35.3
	1500	0.2	22.7	37.5	33.0	38.0
		0.6	22.7	36.0	33.0	37.3
Coastguard	500	0.2	21.5	29.0	28.4	29.6
		0.6	21.5	28.1	28.4	29.0
	1000	0.2	22.0	32.0	30.3	32.6
		0.6	22.0	31.0	30.3	32.0
	1500	0.2	22.2	34.0	31.3	34.5
		0.6	22.2	32.8	31.3	34.0

In fact, we conducted a large number of experiments and obtained similar results to the curves shown in Fig. 7. Due to space limitations, only parts of the results are summarized in Table 2.

3.2 Dynamic adaptability

The internet behavior varies significantly over time. This section encodes Foreman and Coastguard with the MPEG-4 codec FFmpeg, and introduces some perturbations during the transmission simulation to investigate the dynamic adaptability of different schemes. The results in Fig. 8 show the effect of delay fluctuations with a Gamma distribution with the packet loss related parameters having perturbation of 20% (84% of the end-to-end delay processes were shifted with a Gamma shape distribution with a long tail[26]). The results show that the standard deviation is reduced 90% by M3FEC compared with SPath, 55% compared with MDC, and 80% compared with FEC.

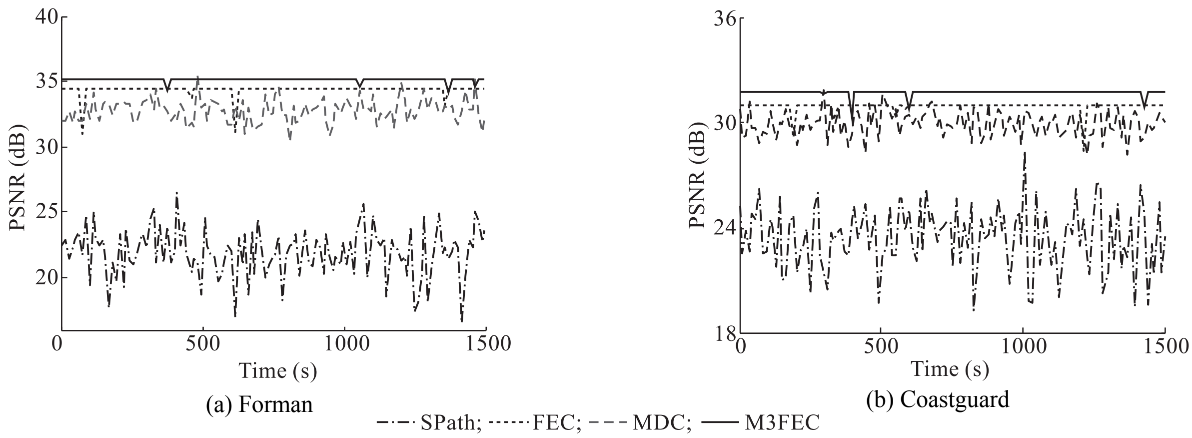


Fig. 8 PSNR fluctuations with an end-to-end delay having a Gamma like distribution and packet loss fluctuations of 20% with B_a =1000 kbps, t_i =100 ms, p_i =0.03, and ρ_i =0.6

3.3 PlanetLab experiment

To evaluate the effectiveness of M3FEC in a real world environment, an on-line transmission module was implemented on Planetlab as shown in Fig. 9. There were three types of nodes: sender, forwarder, and receiver in the measurement. The senders collected round-trip delay and Gilbert packet loss information for each path with the cooperation of the forwarders and receivers. After calculating the transmission scheme by addressing proposed optimization, the standard CIF format source sequence was encoded using corresponding source encoding and channel encoding parameters. Then, in accordance with the senders' scheduling, RTP packets with a timestamp and sequence number were sent out directly to the receiver along the default IP-layer path or to the forwarder which relayed the packets to the receiver. At the receiver side, original stream was reconstructed by reordering and decoding the received packets. Since interactive multimedia has stringent delay requirements, packets arriving beyond 400 ms one-way end-to-end delay was considered to exceed the receiver's play-out time and dropped in the tests.

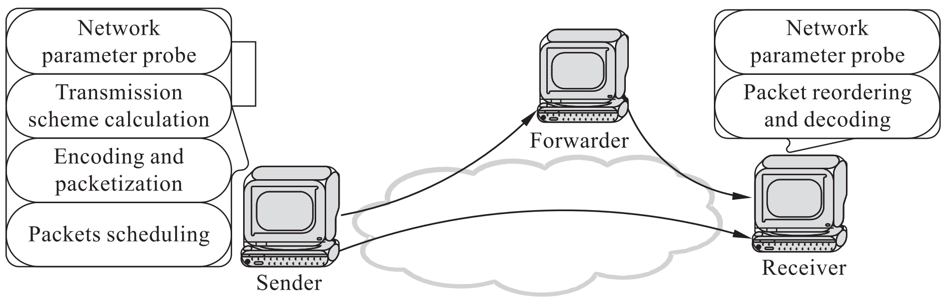


Fig. 9 Structure of M3FEC on-line module

Each group of sender, forwarder, and receiver was defined as a triple. The tests have 100 triples covering 127 distinct nodes, scattered over 28 countries and 3 continents. The node distribution is given in Fig. 10

and Table 3. Although this was only a median scale test bed, the tests contained several typical long-distance communication scenarios, such as Asia to North American, Europe to North American, Asia to Europe, and Education network to Commercial network. The overlay paths were selected using a previously proposed algorithm[27], which provides excellent delay and path diversity at the same time.

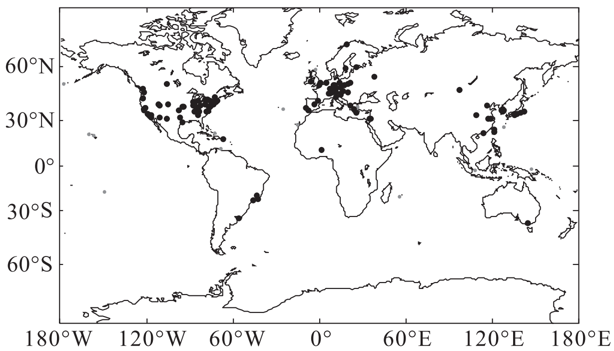


Fig. 10 Geographical distribution of the selected nodes

Table 3 Intercontinental distribution statistics of the selected nodes

North American	Europe	Asia	South America	Oceania	com/org/net
49	39	17	4	1	17

In the measurements, each type of transmission lasted 1 min and was conducted consecutively on one triple. The results are summarized in Fig. 11 and Table 4. In Fig. 11, the network environment deteriorates along the *x*-axes. With relatively small packet losses, the three multiple path transmission schemes, FEC, MDC and M3FEE have similar performance characteristics; while with more lossy conditions, M3FEC exhibits more effectively in avoiding distortion degradation, with nearly 17 dB PSNR improvement over SPath in the worst case.

Table 4 Average PSNR in PlanetLab experiment

Sequence	B_a (kbps)	PSNR (dB)			
		SPath	FEC	MDC	M3FEC
Foreman	500	26.4	31.9	30.2	33.4
	1000	27.6	34.0	30.9	36.2
	1500	28.1	35.9	32.0	37.4
Coastguard	500	24.5	29.2	28.1	30.0
	1000	25.1	31.5	30.0	32.6
	1500	25.4	33.4	30.8	33.9

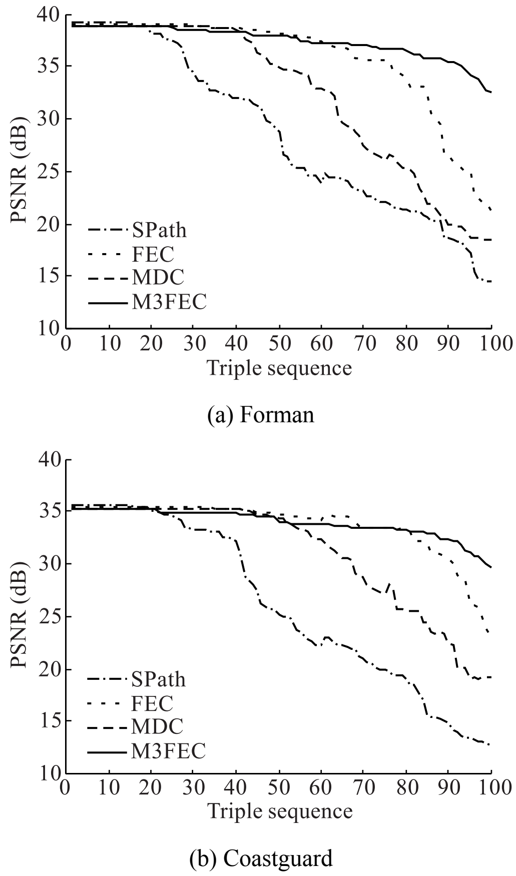


Fig. 11 PSNR of different schemes for each PlanetLab triple (B_a =1500 kbps)

4 Conclusions and Future Work

This paper describes M3FEC—a joint MDC and FEC transmission scheme for interactive multimedia communications. The high error resilience of this scheme is demonstrated theoretically and empirically through extensive numerical analysis and actual Internet measurements. M3FEC yields the maximal video quality at the receiver by taking both the network conditions and application requirements into consideration. The series of experiments concentrate on video because it places the greatest load on the network in terms of bandwidth, as well as delay and packet loss. The method can easily be extended to voice transmission.

As indicated in Section 2, in order to simplify the analysis, this paper employs abstract and simplified network models and architectures, which may cause some inaccuracies with the M3FEC scheme. It will be our valuable but challenging future work to employ more general network models into M3FEC.

Reference

[1] Andersen D, Balakrishnan H, Kaashoek F, et al. Resilient overlay networks. *Operating Systems Review* (*ACM*), 2001, **35**(5): 131-145.

[2] Apostolopoulos J G. Reliable video communication over lossy packet networks using multiple state encoding and path diversity. In: Proceedings of Visual Communications and Image Processing (VCIP). San Jose, CA, USA, 2001: 392-409.

[3] Gummadi K P, Madhyastha H V, Gribble S D, et al. Improving the reliability of Internet paths with one-hop source routing. In: Proceedings of 6th USENIX Symposium on Operating Systems Design and Implementation (OSDI). San Jose, CA, USA, 2004: 183-197.

[4] Zhang Hui, Tang Li, Chen Zhen, et al. Empirical study on multi-path benefits on interactive video. In: Proceedings of the 2007 IEEE International Conference on Multimedia & Expo (ICME). Beijing, China, 2007: 2154-2157.

[5] Lee S J, Banerjee S, Sharma P, et al. Bandwidth-aware routing in overlay networks. In: Proceedings of IEEE INFOCOM. Phoenix, AZ, USA, 2008: 2405-2413.

[6] Jain M, Dovrolis C. Path selection using available bandwidth estimation in overlay-based video streaming. *Computer Networks*, 2008, **52**(12): 2411-2418.

[7] Apostolopoulos J G, Trott M D. Path diversity for enhanced media streaming. *IEEE Communications Magazine*, 2004, **42**(8): 80-87.

[8] Fei T, Tao S, Gao L, et al. How to select a good alternate path in large peer-to-peer systems? In: Proceedings of IEEE INFOCOM. Barcelona, Spain, 2006: 1106-1118.

[9] Zhuang X, Zhu S. Disjoint multipath QoS routing. In: Proceedings of 4th Annual IEEE Consumer Communications and Networking Conference (CCNC). Las Vegas, NV, USA, 2007: 556-559.

[10] Nelakuditi S, Zhang Z L, Du D H C. On selection of candidate paths for proportional routing. *Computer Networks*, 2004, **44**(1): 79-102.

[11] Apostolopoulos J G, Wong T, Tan W T, et al. On multiple description streaming with content delivery networks. In: Proceedings of IEEE INFOCOM. New York, NY, USA, 2002, (3): 1736-1745.

[12] Padmanabhan V N, Wang H J, Chou P A. Resilient peer-to-peer streaming. In: Proceedings of IEEE International Conference on Network Protocols (ICNP). Atlanta, GA, USA, 2003: 16-17.

[13] Liu Zhengye, Shen Yanming, Panwar S, et al. Efficient substream encoding and transmission for P2P video on demand. In: Proceedings of International Packet Video Workshop. Lausanne, Switzerland, 2007.

[14] Begen A C, Altunbasak Y, Ergun O, et al. Multi-path selection for multiple description video streaming over overlay networks. *Signal Processing*: *Image Communication*, 2005, **20**(1): 39-60.

[15] Yu Xunqi, Modestino J W, Fan D. Joint multiple description coding and FEC for delay-constrained applications in congested networks. In: Proceedings of IEEE Information Theory Workshop (ITW). Lake Tahoe, CA, USA, 2007: 547-552.

[16] Shannon C E. A mathematical theory of communication. *Journal of Bell Syst. Tech.*, 1949, **27**: 379-423.

[17] Dai Min, Loguinov D, Radha H M. Rate-distortion analysis and quality control in scalable Internet streaming. *IEEE Transactions on Multimedia*, 2006, **8**(6): 1135-1146.

[18] FFmpeg. http://ffmpeg.org/. 2009.

[19] Zhang Hui, Zhou Jin, Li Jun. M2FEC: An effective FEC based multi-path transmission scheme for interactive multimedia communication. *Journal of Visual Communication and Image Representation* (*JVCI*), 2009, **21**(2): 120-128.

[20] Jiang Wenyu. Qos measurement and management for internet real-time multimedia services [Dissertation]. Department of Computer Science, Columbia University, USA, 2003.

[21] Yajnik M, Moon S, Kurose J, et al. Measurement and modelling of the temporal dependence in packet loss. In: Proceedings of IEEE INFOCOM. New York, NY, USA, 1999, (1): 345-352.

[22] Salamatian K, Vaton S. Hidden Markov modeling for network communication channels. *Performance Evaluation Review*, 2001, **29**(1): 92-101.

[23] Wu Mingquan, Radha H. Network embedded FEC for overlay and P2P multicast over channels with memory. In: Proceedings of Conference on Information Sciences and Systems (CISS). Baltimore, MD, USA, 2005.

[24] One way transmission time, ITU-T recommendation G.114, 2003.

[25] Reed-Solomon code. https://systems.cs.colorado.edu/projects/softmac/browser/trunk/fec/. 2008.

[26] Bovy C J, Mertodimedjo H, Hooghiemstra G, et al. Analysis of end-to-end delay measurements in Internet. In: Proceedings of the Passive and Active Measurement Workshop (PAM). Fort Collins, CO, USA, 2002.

[27] Tang Li, Huai Yin, Zhou Jin, et al. A measurement study on the benefits of open routers for overlay routing. *Journal of Communications* (*JCM*), 2009, **4**(9): 714-723.