Distortion-Optimal Receiver Grouping for MD-FEC coded Video Streaming

Hang Zhang, Adarsh K. Ramasubramonian, Koushik Kar, and John W. Woods

Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

Email: zhangh10@rpi.edu, adarshkrishnan@gmail.com, {koushik, woods}@ecse.rpi.edu

August 21, 2012

Distortion-Optimal Receiver Grouping for MD-FEC coded Video Streaming

Hang Zhang, Adarsh K. Ramasubramonian, Koushik Kar and John W. Woods

Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Email: zhangh10@rpi.edu, adarshkrishnan@gmail.com, {koushik, woods}@ecse.rpi.edu

Abstract

Multiple Description with Forward Error Correction (MD-FEC) coding provides the flexibility, easy adaptivity and distortion-rate optimality that are desirable for delivering streaming video in a network environment with time-varying bandwidth fluctuations and random packet losses. In this work, we consider the issue of how diverse receivers of a video stream should be grouped – where each group receives a MD-FEC coded bitstream optimized for that group – so that the average video distortion is minimized across all receivers. We show that a *sequential* grouping solution is optimal for linear distortion-rate functions. For non-linear distortion-rate functions, while the optimal grouping structure may not be sequential in general, we observe that the approximation factor attained by the best sequential solution can be characterized in terms of the "degree of convexity" of the distortion-rate function. Numerical experiments with realistic distortion-rate functions reveal that the difference between the globally optimal grouping solution and the best sequential solution, is typically small. We provide a dynamic programming based polynomial-time algorithm to compute the best sequential solution.

Index Terms

video streaming, MD-FEC, grouping, distortion, PSNR, dynamic programming

I. Introduction

Video (and multimedia) networking is arguably one of the most important emerging trends in communication networking today [1]. Streaming multicast video (live or pre-recorded), which is envisioned by many to become the "killer" application in the coming decade, will decide whether the Internet and the current wireless technologies can live up to the explosive growth of video applications, and the multimedia demands of the next-generation network users [2]. As networks become more diverse and dynamic in nature, they pose significant challenges to effective delivery of streaming video. Firstly, the receivers that request and receive the streaming video can have heterogeneous access link capacities. Secondly, the available bandwidths on the source-to-receiver paths or the receiver's access links can vary dynamically with time. Finally, the streaming video delivery solution must scale to a large number of interested subscribers (receivers) who can dynamically join and leave the network.

Multiple Description coding with *Forward Error Correction* (MD-FEC), introduced in [3], provides a promising technology for effective video streaming over typical network environments. MD-FEC allows the necessary flexibility, easy adaptivity and distortion-rate optimality that are necessary or desirable requirements for delivering streaming video (as described earlier). With MD-FEC coding, video is coded as multiple *descriptions*, and different parts of the video are protected from channel losses through differentiated redundancy (FEC) provisioning. The overall distortion that can be attained through MD-FEC coded video streaming would however depend on the number of MD-FEC streams that can be used to serve a given set of receivers. On one extreme, if each receiver is served a different MD-FEC coded bitstream (simulcast), the overall distortion will be the minimum. However, this approach does not scale with number of receivers. On the other extreme, all receivers of a video can be grouped together and served a single bitstream. Such an approach is scalable, but would lead to larger average distortion as compared to a simulcast solution. The problem that we address is how a given set of M (typically large) receivers each with different path bandwidths (resulting in different packet-loss rates across the receivers), should be divided into Q (typically small) groups, where each group is served a separate MD-FEC coded bitstream; the overall objective in this grouping strategy is to minimize the average video distortion across all M receivers.

Grouping of receivers into a given number of groups also becomes necessary when we take into account a typical large-scale video distribution system (network) where a set of relay servers is used to distribute video from a single source server to a (possibly large) number of receivers. In such a scenario, each source forwards the video content to the relay nodes (relay servers), and each relay node in turn encodes the video and sends it to a subset of all receivers. Each receiver group in this context corresponds to the set of receivers assigned to receive the video stream from a single relay node.

For this optimal receiver grouping problem, the key results that we present in this work are as follows. We first show that for linear distortion-rate functions, there exist optimal solutions that have a *sequential* structure, which
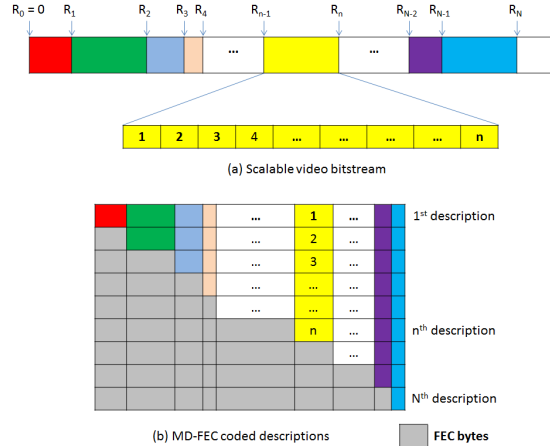
Fig. 1. MD-FEC coding basics: given a Group-Of-Pictures (GOP) of scalable-coded video bitstream organized from Most Significant Bit (R_0) to less significant bits (R_N), suppose we want to encode this GOP into N descriptions, we first run an optimal bit allocation scheme and divide the bitstream into N sections, marked with source-rate break points $R_0, R_1, R_2, ..., R_N$, where $R_0 \leq R_1 \leq R_2 \leq ... \leq R_N$ and $R_0 = 0$. Section n ($n \in [1, N]$), contained between rate points R_{n-1} and R_n is further split into n equal size subsections. These subsections are encoded by a Reed-Solomon (N, n) code vertically at block level to generate parity blocks. Each row in (b) corresponds to a description.

implies that finding the best (optimal) sequential grouping also provides us a solution that is globally optimal. We observe that for non-linear distortion-rate functions, optimal sequential solutions need not be globally optimal; however, the difference in the overall distortion attained by the two solutions is bounded by the "degree of convexity" of the distortion-rate curve. We show that the optimal sequential grouping solution (for both linear and non-linear distortion-rate functions) can be cast as a dynamic programming problem, and is computable in polynomial-time. Through numerical experiments with realistic distortion-rate functions, we compare the performance of the optimal sequential grouping solution (computed by solving the dynamic program), the globally distortion-optimal solution (not necessarily sequential, and computed through enumeration) and a baseline equal partitioning approach. We also demonstrate the benefits of a per-group multirate (MD-FEC) solution over a per-group unirate video streaming approach. We observe that the best sequential grouping solution typically attains a distortion that is very close (sometimes equal) to the minimum attainable distortion.

II. PROBLEM FORMULATION

We first provide a brief overview of MD-FEC. MD coding [4] involves splitting the source data into two or more descriptions in such a way that even if a subset of descriptions is received, the receiver would still be able to decode the video, albeit at a lower quality. Priority encoded transmission (PET) [5] was introduced to improve the transmission of priority ordered data, e.g. the I, P, and B frames of MPEG2, on lossy packet networks, by generating MD codes with the help of parity bytes. The MD-FEC algorithm [3] was developed to generate descriptions that are distortion-optimal for a video source over a single lossy link between a source and a receiver. For this purpose, one can use Reed-Solomon codes (which satisfy the Maximal Distance Separable or MDS property) of type (N, n), for $n = 1, \ldots, N$ where N is the number of descriptions that we generate per GOP. As illustrated in Figure 1, the RS encoding for each section is done vertically and the FEC bytes are arranged below the corresponding input source symbols. If the receiver obtains n descriptions, then it will be able to decode all the source data up to rate R_n (the first n sections). MD-FEC video coding nicely adapts itself to changes in the available capacities and the packet loss rates. The optimization algorithm returns the rate break-points $\{R_n\}_{n=1}^N$ that would minimize the distortion seen by the receiver, when the loss statistics of the link connecting the source and the receiver is known.

We consider MD-FEC in a network model that includes one source (the origin server or server for the video content), and M receivers (subscribers of the video). The M receivers are then grouped into Q groups, where each group is served by a separate (independent) MD-FEC coded bitstream via multicast. In a video distribution network as described in the previous section, each group may correspond to the set of receivers assigned to (served by) a single relay server. The focus of this work is on finding a receiver grouping strategy that is optimal or near-optimal, and yet computationally efficient.

Towards this end, we study the class of *sequential* grouping solutions, which – as we show later in this work – attains the above desirable properties. A sequential grouping solution is of *multi-threshold* type, i.e., there

exists a set of $Q + 1$ bandwidth thresholds, $\theta_0, \theta_1, \cdots \theta_Q$, (where $\theta_0 = 0$, and θ_Q is the maximum possible receiver path bandwidth) such that any receiver k with path bandwidth in $(\theta_{q-1}, \theta_q]$ is a part of group q, for $q = 1, \cdots, Q$.

We consider a scalable video bitstream that is coded into N descriptions, where each description is of rate Δ. We assume discrete bandwidth levels, where the minimum bandwidth granularity is Δ and all receivers have path bandwidths in multiples of Δ. Let ρ_n (given) be the number of receivers whose bandwidth is $n\Delta$, for $n = 1, \cdots, N$. We can see that $\sum_{i=1}^{N} \rho_i = M$. Note that there are $O(Q^M)$ ways to put M receivers into Q groups, which is a very large number. Our optimization objective is to minimize the average distortion:

$$\min_{\{\rho_{n,q}, R_{n,q}\}} \frac{1}{M} \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q} D(R_{n,q}), \tag{1}$$

$$\text{subject to } \sum_{q=1}^{Q} \rho_{n,q} = \rho_n, \ n = 1, 2, ..., N, \tag{2}$$

and for every $q \in \{1, \ldots, Q\}$,

$$\sum_{n=1}^{N} \alpha_n R_{n,q} \le N\Delta, \ R_{1,q} \le R_{2,q} \le \ldots R_{N,q}, \tag{3}$$

where $D(R)$ is the distortion-rate function, assumed to be convex decreasing, $\rho_{n,q}$ is the number of receivers with path bandwidth $n\Delta$ that are assigned to group q, $R_{n,q}$ is the source rate decodable by receivers at bandwidth level $n\Delta$ in group q, and $\alpha_n = \frac{N}{n(n+1)}$ for $1 \le n \le N - 1$, and $\alpha_N = 1$. We minimize the objective over the variables $\rho_{n,q}, R_{n,q}$ for $n = 1, \ldots, N, q = 1, \ldots, Q$.

In our model and MD-FEC solution, we implicitly assume that a receiver with path bandwidth $n\Delta$ suffers an apparent (packet) loss of $\frac{n\Delta}{N\Delta} = \frac{n}{N}$. We relax this assumption later in our simulation study in Section IV-B, where we also consider random losses.

In general, it can be shown that there exists an optimal grouping solution where all receivers at the same bandwidth level are assigned to the same group. In other words, there is an optimal solution in which for each n, there exists a $\tilde{q}(n)$ such that $\rho_{n,\tilde{q}(n)} = \rho_n$. The solutions that we find in this work satisfy this property.

III. ANALYTICAL RESULTS

In this section, we analyze the properties of the optimal sequential (multi-threshold) grouping solution, and analyze theoretically how it compares with the "globally optimal" solution. Note that in the globally optimal solution, no constraint is placed on the policy type, so this solution may or may not be sequential. We have proved that the optimal sequential solution can be computed in polynomial-time using dynamic programming.

A. Global optimality of sequential grouping for linear distortion-rate functions

Distortion-rate functions are decreasing, theoretically convex functions of the rate. To obtain important insights to the grouping problem, however, we first analyze the special case of *linear* distortion-rate functions. If the range of rate variations is "small," then the convex distortion-rate functions are well approximated by a linear function; otherwise the linear distortion-rate functions that we analyze can be viewed as the outer (upper) approximation of the actual convex distortion-rate function.

Proposition 1: For linear distortion-rate functions, there exists a sequential MD-FEC grouping solution that is globally optimal, i.e., attains the minimum value in (1).

The proof of this result can be found in Section VI Appendix.

The above result states that if distortion-rate functions are linear, the optimal sequential MD-FEC grouping solution, i.e., a sequential grouping solution with MD-FEC coding, where the grouping thresholds are chosen "optimally", performs as good as any other possible solution. We next analyze the case of general convex distortion-rate functions.

B. Optimality properties of sequential grouping for general convex distortion-rate functions

For general convex distortion-rate functions, it can be shown that the optimal sequential grouping is not necessarily globally optimal. As an example, consider 3 bandwidth levels (1, 2, 3), and 2 groups, i.e. $N = 3, Q = 2$, and Gaussian distortion-rate function $D(R) = \sigma_g^2 2^{-2R}$, for $R \ge 0$. The user population across the bandwidth levels is: $\rho_1 = 1$, $\rho_2 = 10$, $\rho_3 = 15$, which can be represented with the vector $\overrightarrow{\rho} = (1, 10, 15)$ to represent the population. Set $\sigma_g = 1$ and $\Delta = 1$ for computational simplicity.

In the globally optimal solution, the receivers at bandwidth levels 1 and 3 (1 and 15 receivers, respectively) are grouped together; the rate values ($R_{n,1}$) obtained are as follows: $R_{1,1} = 0.1822 = R_{2,1}, R_{3,1} = 2.6356$. Receivers at bandwidth level 2 (10 receivers) form a group by themselves, and the rate values obtained are $R_{1,2} = 0, R_{2,2} = 2.0 = R_{3,2}$. The average distortion is 0.0689.

In the optimal sequential solution, however, the user at bandwidth level 1 forms a group by itself, with the corresponding rate value being 1.0, i.e., $R_{1,1} = 1.0 = R_{2,1} = R_{3,1}$. Receivers at bandwidth levels 2 and 3 (10 and 15 receivers, respectively) constitute the second group, and the rate value for all these receivers is 2.0, i.e., $R_{1,2} = 0, R_{2,2} = 2.0 = R_{3,2}$. The average distortion is 0.0697. So we observe that the optimal sequential distortion value is about 1% more than the globally minimal distortion value. Next, we upper-bound this difference in performance in terms of the convexity of the distortion-rate function.

Given a decreasing, convex distortion-rate function D, let D_L represent its linear (outer) relaxation, constructed as a line joining $(0, D(0))$ and $(N\Delta, D(N\Delta))$. We have $D_L \geq D$ for $R \in [0, N\Delta]$. Assume the maximal ratio between $D_L(R)$ and $D(R)$ in this range is r, i.e. $r = \max_{0 \leq R \leq N\Delta} \frac{D_L(R)}{D(R)}$. Then we have the following result, showing that the distortion of the optimal sequential solution differs from the globally optimal distortion by at most the multiplicative factor r.

Proposition 2: For convex distortion-rate functions, the distortion attained by the optimal sequential MD-FEC grouping solution is within a factor of r of the globally optimal distortion value.

The proof of this result can be found in Section VI Appendix.

More precisely, let

$$H(\overrightarrow{\rho}, \overrightarrow{R}) = \frac{1}{M} \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q} D(R_{n,q})$$

denote the distortion-rate function that we want to minimize, as a function of the decision variables (vectors) $\overrightarrow{\rho}, \overrightarrow{R}$. Let H^* be the globally optimal distortion (1), and $\overrightarrow{\rho}^*_S, \overrightarrow{R}^*_S$ be the optimal sequential MD-FEC grouping solution. Then, the above result states the following:

$$H(\overrightarrow{\rho}^*_S, \overrightarrow{R}^*_S) \leq rH^*.$$

Our numerical/experimental studies with $N \leq 8$ and $Q \leq 3$ reveal that in general, the performance of the optimal sequential solution is much closer to the globally optimal distortion than that predicted by Proposition 2; in many cases, they are exactly the same. Numerical studies for larger values of N, Q could not be conducted due to the high computational complexity of running the globally optimal solution (which requires enumeration over a number of points that is exponential in the problem parameters).

C. Computation of optimal sequential grouping

We next present a polynomial-time algorithm based on dynamic programming that finds the optimal sequential grouping solution for general convex distortion-rate functions. Let $K(i, i+1, ..., j)$ be the minimal total distortion of receivers with bandwidth level indices $i, i + 1, ..., j$ when they are put into one group. Note that once the set of receivers in a group is given, the minimal total distortion value can be computed by solving a convex optimization problem, and the rate allocations ($R_{n,q}$) correspond to an MD-FEC solution. This implies that finding $K(i, i+1, ..., j)$ in general requires solving a convex optimization problem, which can be solved exactly, or be approximated to any desired approximation factor, in polynomial-time [3]. Now we introduce the algorithm:
Algorithm **SEQOPT-MDFEC:**
(1) Initialization: Initialize an $(N + 1) \times (Q + 1)$ matrix $J(\cdot, \cdot)$ as,
$\quad J(0, 0) = 0, J(0, 1) = 0, ..., J(0, Q) = 0$;
$\quad J(1, 0) = \infty, ..., J(N, 0) = \infty$.
(2) Iterative update: For $1 \leq n \leq N$, $1 \leq q \leq Q$,

$$J(n, q) = \min \begin{cases} J(n, q - 1), \\ J(n - 1, q - 1) + K(n), \\ J(n - 2, q - 1) + K(n - 1, n), \\ ...... \\ J(0, q - 1) + K(1, 2, ..., n). \end{cases}$$

(3) Output minimal average distortion $\frac{J(N,Q)}{M}$.

Proposition 3: On termination of **SEQOPT-MDFEC**, $J(N, Q)/M$ corresponds to the minimum average distortion that can be attained by any sequential grouping solution with MD-FEC coding.

The proof of this result can be found in Section VI Appendix.

We now consider the computation time of the algorithm. The computation time is dominated by step 2 (the iterative update procedure), which requires $O(N^2Q)$ computations of a convex programming problem (MD-FEC computation for a single group/stream), for finding $K(i, i+1, ..., j)$. As mentioned earlier, $K(i, i+1, ..., j)$ can be computed or approximated closely in polynomial-time.

IV. NUMERICAL RESULTS

In this section we evaluate SEQOPT-MDFEC, or optimal sequential grouping with MD-FEC coding, by analyzing simulation results based on real video sequences. We show how it compares with the globally optimal solution (which may not be sequential) on some representative examples. We then compare SEQOPT-MDFEC and three other sequential receiver grouping and video streaming methods:

(i) *Equal partitioning with MD-FEC:* In this solution, the bandwidth levels are divided into the groups evenly (and in a sequential manner), and MD-FEC is applied within each group. This is a special case of sequential grouping (one in which the thresholds $\theta_0, \theta_1, \cdots, \theta_Q$ are evenly spaced), but is not necessarily the optimal.

(ii) *Equal partitioning with unirate:* The grouping in this solution is the same as in (i), but unirate video coding is used instead of MD-FEC. In unirate video coding, the video is only sent at a single rate per group. Note that the unirate solution, where there is a single rate point per group, is a special case of the MD-FEC solution. The rate assigned to a group is computed by minimizing the distortion for that group.

(iii) *Optimal unirate grouping:* In this case, unirate video streaming used, but the grouping solution is optimized (taking unirate transmission into account). A dynamic programming procedure similar to the SEQOPT-MDFEC algorithm can be used to obtain the optimal grouping in this case; this has been observed before in [6] which addresses the optimal grouping question for unirate video streaming.

Note that performance comparison of our proposed solution (SEQOPT-MDFEC) with (i) helps us identify the benefits of optimal grouping alone, for MD-FEC streaming. Performance comparison of the proposed solution and (iii) allows us to evaluate the benefits of using MD-FEC alone, when receiver grouping is done optimally for the two cases. Performance comparison of the proposed solution and (ii) would show the joint benefits of optimal sequential grouping and MD-FEC coding over a simple baseline grouping solution and unirate coding.

In the following, performance is measured in terms of average PSNR, which is popularly used to quantify video quality. The average PSNR measure is equivalent to the average distortion (D) measured in terms of MSE, and the two are related as follows: $PSNR = 10 \log_{10}(255^2/D)$. We assume that all packet losses follow a binomial distribution. The results shown below are for GOP 1 of the CIF@30fps *Foreman* video; we have processed all 18 GOPs of the Foreman video and verified that the results are similar in nature to those for the first GOP presented here.

In this case, it also turns out that the globally optimal solution is sequential in nature, when the number of bandwidth levels is restricted to 8, and the receiver population across bandwidth levels follows a uniform distribution. Thus SEQOPT-MDFEC results in the globally optimal distortion value for these settings.

A. Results comparison

For the results discussed below, the number of bandwidth levels (also equal to N) is 32. The number of groups (Q) is varied from 1 to 6, and the total bandwidth $N\Delta$ considered is 1 Mbps. We consider two cases based on the number of receivers at each bandwidth level.

(1) Uniform distribution: $\overrightarrow{\rho} = (0, 0, 0, 0, 1, ..., 1)$, i.e. it includes 32 bandwidth levels, where the lowest 4 bandwidth levels have no receivers, and the rest have one (same number of) receiver(s) per bandwidth level.

(2) Gaussian distribution: $\rho_n = Ae^{-\frac{(n-\mu)^2}{2\sigma^2}}$, where A=1000, μ=(1+32)/2=16.5, σ=25 for $n = 5, 6, \ldots, 32$. In addition, $\rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$.

The reason we set the number of receivers at very low bandwidth levels to be zero is that the video cannot be decoded at those rates.

The average PSNRs of SEQOPT-MDFEC, the proposed solution (labeled 'Optimal Sequential; MDFEC' in the figures), and solutions (i), (ii) and (iii) as described above, are shown in Figures 2 and 3.

From the figures, we observe that average PSNR increases as the number of groups increases, as expected, but with diminishing returns. Our solution (SEQOPT-MDFEC) provides the best performance in all cases, but the difference with the optimal sequential unirate solution (iii) goes away for more than two groups. The results also show that MD-FEC provides significant benefits, but only for small number of groups; for larger number of groups, the grouping strategy makes a greater difference, and the performance difference between the optimal grouping and equal partitioning cases is significant even when the number of groups is five or more.
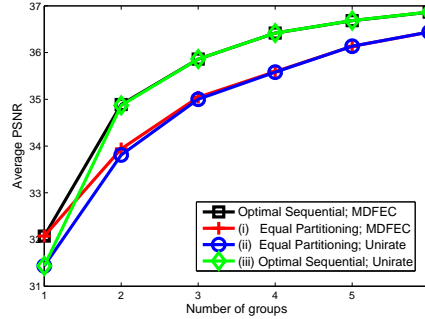
Fig. 2. Average PSNR versus number of groups for the four grouping/coding strategies with *uniform* receiver distribution across bandwidth levels.
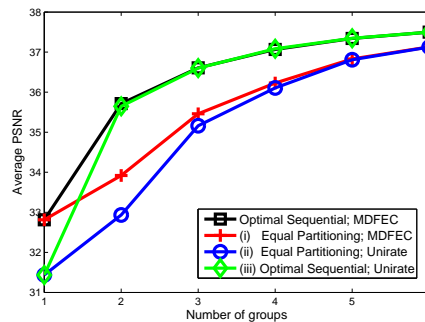

Fig. 3. Average PSNR versus number of groups for the four grouping/coding strategies with *Gaussian* receiver distribution across bandwidth levels.

B. Results with random path losses

So far in our formulation, analysis and simulation, we have assumed only "apparent losses" of packets in the network, i.e. deterministic losses that happen only due to bandwidth limitations of the paths from the source to the receivers. In addition to these inevitable losses, there can be random losses due to faulty/noisy links (as in wireless networks), buffer overflows, etc. Next, we consider such additional losses in our simulation model; since such loss rates are typically small, this additional loss rate is set between 0 and 0.2 in our simulations. The rest of the parameters remain the same as in Section IV-A. While we have studied both cases of uniform and Gaussian receiver distributions (across the bandwidth levels) as before, we only show the results for the Gaussian case below.

Figure 4 shows the corresponding performance results, when all receivers have the same additional loss rate of 0.1. In this case, it can be shown that SEQOPT-MDFEC can again compute the optimal sequential grouping solution for MD-FEC coding (Proposition 3). For Figure 5, the additional loss rates for different receivers are different, and follows a uniform distribution between 0 and 0.2. Note that in the latter case, the notion of a "sequential solution" is not well-defined. In this case, we order the receivers in terms of their *effective path bandwidths*, i.e., path bandwidth $\times$ (1 - loss rate), where the loss rates vary between 0 and 0.2, as mentioned above. The sequential solution is now defined in terms of these effective path bandwidths (as opposed to the raw path bandwidths), and the algorithm SEQOPT-MDFEC is utilized to compute the solution that is labeled "Optimal Sequential; MDFEC" in Figure 5. Numerical studies on some small instances of the problem revealed that this grouping strategy, along with MD-FEC, attains close-to-minimal distortion in most cases (the latter being computed through enumeration over all possible groups).

C. Discussion

Comparing the results in Figures 4 and 5 with those in Figures 3, we observe that while the general trends are similar, use of MD-FEC makes a significant difference in the lossy case (as compared to unirate), both with optimal grouping and equal partitioning, and even when the number of groups is five or more. We also observe that better performance of our solution (SEQOPT-MDFEC) as compared to the baseline solution (ii) comes in nearly equal measure due to optimal sequential grouping the use of MD-FEC. We also observed that
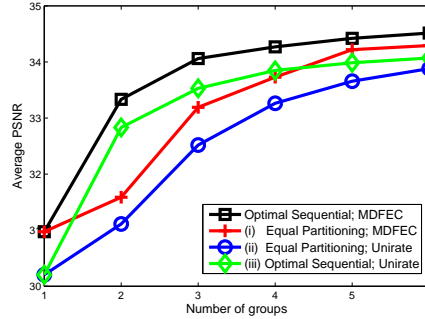
Fig. 4. Average PSNR versus number of groups in four grouping/coding strategies, for *same loss rate* across receivers, and *Gaussian* receiver distribution across bandwidth levels.
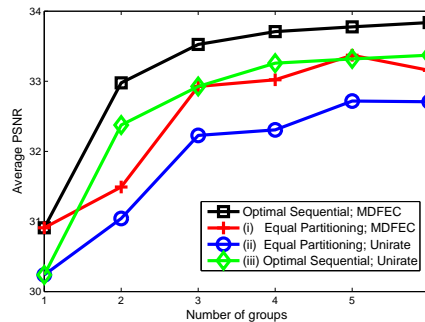

Fig. 5. Average PSNR versus number of groups in four grouping strategies, for *different loss rate* across receivers, and *Gaussian* receiver distribution across bandwidth levels.

the performance difference (between SEQOPT-MDFEC and other solutions that do not use optimal grouping or MD-FEC coding or both) is more significant when (i) receivers are distributed unevenly across bandwidth levels, (ii) losses due to path bandwidth limitations and bandwidth-independent losses are both present, and (iii) packet loss rates vary across receivers. Thus in realistic network scenarios involving video streaming, optimal sequential grouping and MD-FEC coding are likely to complement each other, and attain good PSNR (distortion) performance across a wide range of network characteristics.

V. CONCLUSION AND FUTURE WORK

For the distortion-optimal receiver grouping problem for MD-FEC video streaming, we showed that the best sequential grouping solution is globally optimal when the distortion-rate function is linear. In general cases where the distortion-rate function is non-linear, we further showed that the same solution is approximately optimal – by a factor that depends on the degree of convexity of the distortion-rate curve. The best sequential grouping solution can be obtained by a dynamic programming algorithm in polynomial time, and it performs significantly better than the equal partitioning approach in numerical experiments. We also observed that in terms of performance benefits obtained, optimal sequential grouping and MD-FEC coding seem to nicely complement each other.

In future work, we plan to consider additional constraints on the bandwidths used for serving each group of receivers, which may arise due to access capacity limitations of the relay servers. We also plan to evaluate our sequential grouping solution on other test video sequences.

REFERENCES

[1] "Future looks bright for IPTV," *Satellite Today*, May 2005.
[2] "Web could collapse as video demand soars," *Daily Telegraph*, April 2008.
[3] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," *Asilomar Conf. on Signals, Systems, and Comp.*, vol. 1, pp. 342–346, Oct. 1999.
[4] V.K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, Sep. 2001.
[5] J. Edmonds M.Luby A. Albanese, J. Blomer and M. Sudan, "Priority encoded transmission," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1737–1744, Sep. 2006.
[6] M.S. Kim Y.R. Yang and S.S. Lam, "Optimal partitioning of multicast receivers," *Proc. ICNP*, pp. 129–140, 2000.

VI. APPENDIX

A. Proof for Proposition 1

Proof: Assume $D(x) = -ax + b$, $a > 0, b > 0$. As mentioned above, our aim is to minimize:

$$\min_{\{\rho_{n,q}, R_{n,q}\}_{n=1,\ldots,N, q=1,\ldots,Q}} \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q} D(R_{n,q}).$$

Since we have two vector variables $\rho_{n,q}$ and $R_{n,q}$, it is better to consider them separately. Assume $\rho_{n,q}$ is fixed, then consider $R_{n,q}$ for given $\rho_{n,q}$. Let D_q be the distortion of group q, $1 \le q \le Q$.

$$D_q = \sum_{n=1}^{N} \rho_{n,q} D(R_{n,q})$$

Then the total distortion D_{tot} would be:

$$D_{tot} = \sum_{q=1}^{Q} D_q$$

From the expression of D_q, $R_{n,q}$ and $R_{n,p}$ are independent if $q \ne p$. Hence we can minimize them respectively. Then

$$
\begin{aligned}
D_q &= \sum_{n=1}^{N} \rho_{n,q}(-a \cdot R_{n,q} + b) \\
&= b \sum_{n=1}^{N} \rho_{n,q} - a \sum_{n=1}^{N} \rho_{n,q} R_{n,q}
\end{aligned}
$$

For $\forall q$, $R_{n,q}$ need to satisfy the constraints:

$$\sum_{n=1}^{N} \alpha_n R_{n,q} \le N\Delta, \; R_{1,q} \le R_{2,q} \le \ldots R_{N,q},$$

where $\alpha_n = \frac{N}{n(n+1)}$ for $1 \le n \le N-1$, and $\alpha_N = 1$.

Equivalently,

$$N[R_{1,q} + \sum_{n=2}^{N} \frac{R_{n,q} - R_{n-1,q}}{n}] \le N\Delta.$$

That is,

$$\frac{R_{1,q}}{\Delta} + \sum_{n=2}^{N} \frac{R_{n,q} - R_{n-1,q}}{n\Delta} \le 1.$$

So let $\beta_{1,q} = \frac{R_1}{\Delta}$, and $\beta_{n,q} = \frac{R_{n,q} - R_{n-1,q}}{n\Delta}$ for $2 \le n \le N$, $1 \le q \le Q$.

Then we have $R_{n,q} = \sum_{i=1}^{n} \beta_{i,q} \cdot i\Delta$, where $\beta_{i,q}$ satisfies $\sum_{i=1}^{N} \beta_{i,q} \le 1$ and $\beta_{i,q} \ge 0$.

Then the expression of D_q in terms of $\beta_{i,q}$ is:

$$
\begin{aligned}
D_q &= b \sum_{n=1}^{N} \rho_{n,q} - a \sum_{n=1}^{N} \rho_{n,q} R_{n,q} \\
&= b \sum_{n=1}^{N} \rho_{n,q} - a \sum_{n=1}^{N} \rho_{n,q} \sum_{i=1}^{n} \beta_{i,q} \cdot i\Delta
\end{aligned}
$$

This is a linear objective function with linear constraints. So the optimal solution would be obtained at some vertex(vertices) of the polyhedral feasible region. Thus in the optimal solution, $\{\beta_{1,q}, \ldots, \beta_{N,q}\} = (0, \ldots, 0, 1, 0, \ldots, 0)$, i.e. there is $\widehat{n}$ such that $\beta_{n,q} = 1$ if $n = \widehat{n}$ and $\beta_{n,q} = 0$ if $n \ne \widehat{n}$. This can be interpreted as: if $n < \widehat{n}$, $R_{n,q} = 0$; if $n > \widehat{n}$, $R_{n,q} = R_{\widehat{n},q}$. Each group has an $\widehat{n}$ that can be considered to be the index of the dominating bandwidth level of the group, as long as there are some users in that group.

Assume the globally optimal solution is not sequential, then we can construct a sequential solution at least as good as the globally optimal solution. Suppose the dominating bandwidth level indices for Q groups are n_1, $n_2, \ldots, n_Q$ respectively. Without loss of generality, we assume $n_1 < n_2 < \ldots < n_Q$. Then we construct the

following sequential solution: we put users with bandwidth level indices $1, 2, ..., n_2 - 1$ in group 1, put users with bandwidth level indices $n_i, n_i + 1, ..., n_{i+1} - 1$ in group i (for i from 2 to $Q-1$), put users with bandwidth level indices $n_Q, n_Q + 1, ..., N$ in group Q. This sequential solution would not be worse than the global optimal solution, since the distortion of receivers at any bandwidth level would not get worse after the construction. ∎

B. Proof for Proposition 2

Proof: In general, the total distortion is:

$$H(\overrightarrow{\rho}, \overrightarrow{R}) = \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q} D(R_{n,q})$$

and its linear relaxation is:

$$H_L(\overrightarrow{\rho}, \overrightarrow{R}) = \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q} D_L(R_{n,q})$$

D_L, which is defined in section 3.2, is the outer linear relaxation of D. defined in section 3.2. Let $H(\rho^*, \overrightarrow{R}^*)$ be the global minimum of H, $H_L(\rho_L^*, \overrightarrow{R}_L^*)$ be the global minimum of H_L, $H(\rho_S^*, \overrightarrow{R}_S^*)$ be the optimal sequential solution to H, and $H_L(\rho_{LS}^*, \overrightarrow{R}_{LS}^*)$ be the optimal sequential solution to H_L. Besides, as defined in section 3.2, r is the max ratio of $D_L(R)$ over $D(R)$, i.e. $r = \max\{\frac{D_L(R)}{D(R)}\}$.

Since H_L is the linear relaxation, the optimal sequential solution is globally optimal. Therefore $H_L(\rho_L^*, \overrightarrow{R}_L^*) = H_L(\rho_{LS}^*, \overrightarrow{R}_{LS}^*)$. Then we have:

$$
\begin{aligned}
H(\overrightarrow{\rho}^*, \overrightarrow{R}^*) &= \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q}^* D(R_{n,q}^*) \\
&\geq \sum_{q=1}^{Q} \sum_{n=1}^{N} \rho_{n,q}^* \frac{1}{r} D_L(R_{n,q}^*) \text{ (Because } r = max\{\frac{D_L(R)}{D(R)}\} \text{)} \\
&= \frac{1}{r} H_L(\overrightarrow{\rho}^*, \overrightarrow{R}^*) \\
&\geq \frac{1}{r} H_L(\overrightarrow{\rho}_L^*, \overrightarrow{R}_L^*) \text{ (Because } \overrightarrow{\rho}_L^*, \overrightarrow{R}_L^* \text{ is assumed optimal for } H_L \text{)} \\
&= \frac{1}{r} H_L(\overrightarrow{\rho}_{LS}^*, \overrightarrow{R}_{LS}^*) \text{ (Because for } H_L, \text{ optimal sequential solution is globally optimal.)} \\
&\geq \frac{1}{r} H(\overrightarrow{\rho}_{LS}^*, \overrightarrow{R}_{LS}^*) \text{ (Because } H_L \text{ is a relaxation of } H \text{)} \\
&\geq \frac{1}{r} H(\overrightarrow{\rho}_S^*, \overrightarrow{R}_S^*) \text{ (Because } (\overrightarrow{\rho}_{LS}^*, \overrightarrow{R}_{LS}^*) \text{ is sequential and is feasible, and } (\overrightarrow{\rho}_S^*, \overrightarrow{R}_S^*) \\
&\qquad \text{ is the optimal sequential solution.)}
\end{aligned}
$$

Therefore,

$$H(\overrightarrow{\rho}_S^*, \overrightarrow{R}_S^*) \leq r \times H(\overrightarrow{\rho}^*, \overrightarrow{R}^*)$$

∎

C. Proof for Proposition 3

Proof:

To show the correctness of the dynamic programming solution, we need equivalently to show $J(n, q)$ solves the subproblem for $0 \leq n \leq N, 0 \leq q \leq Q$. Let $J^*(n, q)$ be the minimal total distortion attained with receivers at the first n bandwidth levels and q groups. We want to show $J(n, q) = J^*(n, q)$.

Firstly we consider $q = 0$, i.e. there is no group. Then it is obvious that $J^*(n, q) = 0$ if $n = 0$ and $J^*(n, q) = \infty$ if $n \geq 1$. In the initialization of the dynamic programming algorithm in section 3.3, we set $J(0, 0) = 0$ and $J(n, 0) = \infty$ when $n \geq 1$. Hence $J(n, q) = J^*(n, q)$ holds when $q = 0$.

Next we want to show $J(n, q) = J^*(n, q)$ holds given that $J(k, q - 1) = J^*(k, q - 1)$ for $0 \leq k \leq n$. Assume $J(n, q) \neq J^*(n, q)$, then $J(n, q) > J^*(n, q)$ since $J^*(n, q)$ is the minimum. Let $(1, 2, ..., i_{q-1})$ be the bandwidth level indices of users in the first $(q - 1)$ groups in $J^*(n, q)$. As defined in the algorithm,

$$J(n,q) = \min \begin{cases} J(n, q-1), \\ J(n-1, q-1) + K(n), \\ J(n-2, q-1) + K(n-1, n), \\ ...... \\ J(0, q-1) + K(1, 2, ..., n). \end{cases}$$

Then we have:

$$\begin{aligned} J(n,q) &\leq J(i_{q-1}, q-1) + K(i_{q-1}+1, ..., N) \\ &= J^*(i_{q-1}, q-1) + K(i_{q-1}+1, ..., N) \\ &= J^*(n,q) \end{aligned}$$

(4)

Contradiction. Therefore $J(n,q) = J^*(n,q)$ holds. Hence $J(n,q)$ solves the problem that put the receivers at first n bandwidth levels into q groups. Let $n = N$ and $q = Q$, $J(N,Q)$ solves the problem of N bandwidth levels and Q groups. In other words, $J(N,Q)$ gives the minimal total distortion, and then $\frac{J(N,Q)}{M}$ gives the minimal average distortion.

■