Rate-Constrained Scalable Video Transmission over the Internet*

Bo Hong, Aria Nosratinia
Department of Electrical Engineering
University of Texas at Dallas, Richardson, TX 75080
E-mail: {bhong,aria}@utdallas.edu

Abstract

Forward error correction (FEC) are used widely to address the packet loss problem for Internet video. In this paper, we show that packet size plays a very important role in rate allocation of FEC-based methods. We develop a probabilistic framework for the solution of rate allocation problem in the presence of overhead. We apply this analysis to an unequal error protection scheme combined with Fine Granularity Scalability (FGS) in the MPEG-4 video standard. Given a total available bandwidth, optimal assignment of FEC and packet size is achieved simultaneously by minimizing end-to-end distortion.

1 Introduction

Streaming video applications are becoming increasingly popular over networks such as the Internet. Internet, a best-effort delivery network, suffers packet losses arising from network congestion. A video packet is also useless when it arrives late. Error control techniques such as FEC (Forward Error Correction) and ARQ (Automatic Repeat reQuest) are used to maintain video quality. Of these two, FEC is more suitable for streaming applications, since retransmissions may exceed timing requirements of streaming video, and also degrade network performance.

Given the total available bandwidth, finding the optimal bit allocation is very important in FEC-based video, because the FEC bit rate is deducted from the total rate. We want to give proper protection to the source, but we also should prevent unwanted FEC rate expansion. The rate of packet headers is often ignored in allocating bit rate. Actually, this packetization overhead has significant influence on system performance in many cases, as we show in this paper. Reducing the packet size will increase the rate of packet headers, thus reducing the available rate for the source and its FEC codes. On the other hand, smaller packet size will allow a larger number of packets, in which case it can be shown the efficiency of FEC codes improves.

---

*This work was supported in part by the National Science Foundation under Grant CCR-9985171.

A natural, and so far unanswered, question then is: how important is the tradeoff introduced by header overheads and packet size? And what is the optimal point in this tradeoff? This paper offers answers to both these questions. Our experiments show that the impact of packet header (overhead) cannot be ignored without significant loss of performance. Furthermore, we develop a probabilistic framework for the solution of rate allocation problem in the presence of overhead.

We solve the overhead problem in the context of one of the important video coding modalities for networks: scalable video [1, 2]. Past work in FEC-protected scalable video for networks has ignored the effect of packet overhead [3, 4], or at most has minimizing overhead [5, 6] without considering the effect on efficiency of FEC. We show that significant gains are possible, without much pain, by a judicious allocation of rate while keeping the overhead in mind. We implement our solution on the MPEG-4 Fine Granularity Scalability (FGS) mode [7]. FGS, a very simple and flexible layered coding method, has a two-layer structure where the enhancement layer is progressively coded. To show the flexibility of our technique, we use an unequal error protection scheme with FGS, and present an overall solution for packet size, code rate for the base layer, and code rate for the enhancement layer.

The paper is organized as follows: Section 2 gives some backgrounds and present our scheme in detail. Section 3 shows simulation results. Finally, concluding remarks are given in Section 4.

2 Rate Constrained Video Transmission Scheme

2.1 FGS coding with Unequal Error Protection

In FGS, a video stream is coded into two layers: a base layer and an enhancement layer. A fine-granular (bit-plane coding of DCT coefficients) scheme is used in coding the enhancement layer. FGS is suitable for Internet streaming video, because a server may transmit the video stream at a wide range of bit rates. This property of FGS coding provides a solution to the heterogeneous problem of the Internet in video transmission. The details of FGS coding can be found in [8].

On the error-control side, Reed-Solomon (R-S) codes are used in this paper because they are convenient to analyze and are flexible. A (n, k) R-S code has k information packets within n packets. The decoder can correct up to any $n - k$ packet erasures when the position of lost packets are known. FEC codes are applied across packets, with unequal error protection to different layers according to their importance.

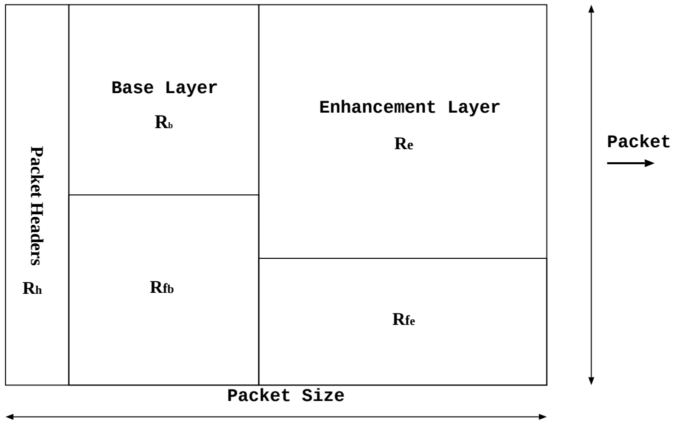

Figure 1: Structure of a Packet Block

Typically, bit streams of all layers are interleaved into one block of packets (BOP). The structure of a BOP is shown in Figure 1, where R_b is the base layer bit rate, R_e is the enhancement layer bit rate, R_{fb} is the FEC bit rate for the base layer, R_{fe} is the FEC bit rate for the enhancement layer, and R_h is the bit rate for packet headers. More FEC is applied to the base layer than the enhancement layer, because the enhancement layer is useless without the base layer.

In this paper, we focus on pre-encoded video. The base layer rate R_b is therefore fixed. The rate for the enhancement layer is variable, because the bitstream of the enhancement layer is progressive. The other rates, as well as packet lengths, are all variable subject to an overall bit budget.

2.2 Channel Model

Because of its simplicity and effectiveness, a two state Markov model is used to simulate packet loss patterns over the Internet. The two states of the model are G and B (see Figure 2). In state G, a packet is received correctly, while in state B, a packet is lost. This model is described by average packet loss probability P_b and average burst packet loss length L_b. The transition probability P_{BG} and P_{GB} can be easily computed using P_b and L_b:

$$P_{BG} = \frac{1}{L_b} \tag{1}$$

$$P_{GB} = P_{BG}\frac{P_b}{1-P_b} \tag{2}$$

Following the development of [9], let $g(v)$ denote the probability that an error-free interval length is $v-1$, i.e., $g(v) = Pr(0^{v-1}1|1)$, where $'1'$ denotes a lost packet and 0^{v-1}

denotes $v-1$ consecutive successfully received packets. Similarly, let $G(v)$ denote the probability that an error-free interval length is greater that $v-1$, i.e., $G(v) = Pr(0^{v-1}|1)$. We have:

$$
\begin{aligned}
g(v) &= \begin{cases} 1-P_{BG}, & v=1 \\ P_{BG}(1-P_{GB})^{v-2}P_{GB}, & v>1 \end{cases} \qquad (3) \\
G(v) &= \begin{cases} 1, & v=1 \\ P_{BG}(1-P_{GB})^{v-2}, & v>1 \end{cases} \qquad (4)
\end{aligned}
$$

Let $R(m,n)$ be the probability that $m-1$ out of $n-1$ packets are lost following a lost packet. The probability of m packet losses within a n-packet block is:

$$P(m,n) = \sum_{v=1}^{n-m+1} P_b\, G(v)\, R(m, n-v+1), 1 \leq m \leq n. \qquad (5)$$

In this paper, we assume that the average packet loss probability P_b is independent of packet size. This assumption is valid for Internet video because the congested router is the dominant reason for packet loss over the Internet, therefore, packet size has no significant effect on the packet loss probability. This assumption has been experimentally verified by Wenger [10] and used by Gallant [6].

For wireless links, the dominant error mechanism is bit error. The longer the packet size, the higher the probability of unrecoverable bit errors, and the higher the probability of packet loss. To adapt our algorithm for wireless links, further analysis is needed. In this work, we only focus on Internet video.

2.3 Packet size optimization

Universal Datagram Protocol (UDP) and Real-time Transport Protocol (RTP) [11] are typically used as transfer layer protocols in video transmission over the Internet. The minimum packet header length is 40 bytes. To maintain a reasonable level of overhead, packet size should be substantially more than 40 bytes. At the same time, 1500 bytes is the

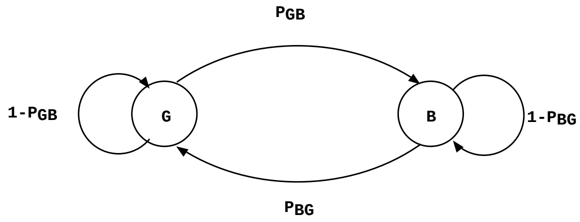

Figure 2: Channel Model

upper bound for packet size due to the Maximum Transfer Unit (MTU) of the Ethernet. Currently, many routers will split a packet longer than 1500 bytes into at least 2 packets. Loss of any packet fragment will make the whole original packet useless.

Long packets reduce the overhead, but shorter packets make for more efficient FEC codes. We demonstrate the latter effect by an example (for the moment, the overhead is disregarded). Assume we want to send 40,000 bytes either in 40 packets of 1000 bytes or in 200 packets of 200 bytes each. We transmit packets over a Markov channel with $P_b = 0.097$ and $L_b = 9.97$. Assuming a code rate of 3/4, the first case uses a (40,30) R-S code which, using Equation 5, will give a probability of block error of 0.16. On the other hand, the smaller packets can be sent using a (200,150) R-S code (same rate) which results in a block error probability of 0.06.

2.4 Optimization Problem Formulation

For the FGS coding scheme, the progressive enhancement layer can be truncated anywhere. Given the total bandwidth R and the fixed base layer rate R_b, we need to allocate available bandwidth (which is $R - R_b$) to R_e, R_{fb}, R_{fe} and R_h. Therefore, this problem can be formulated as minimizing expected distortion D subject to the rate constraint.

$$\begin{aligned} &\min_{R_{fb},R_e,R_{fe},R_h} D, \\ &\text{subject to } R_{fb} + R_e + R_{fe} + R_h \leq R - R_b \end{aligned} \tag{6}$$

where D is the end-to-end expected distortion and R is the total available rate for the packet block. Naturally, we assume that $R > R_b$.

Suppose we use a (n, m_1) R-S code for the base layer and a (n, m_2) R-S code for the enhancement layer where n is the packet number in the block. We have:

$$\begin{aligned} D &= (1 - P_1 - P_2)D_0 + P_1 D_1 + P_2 D_2 \\ &= \sum_{i=0}^{n-m_2} P(i,n)D_0 + \sum_{i=n-m_2+1}^{n-m_1} P(i,n)D_1 + \sum_{i=n-m_1+1}^{n} P(i,n)D_2 \end{aligned} \tag{7}$$

where P_2 is the probability that more than $n - m_1$ packets are lost, which means both layers are lost. D_2, the corresponding distortion, is a fixed value calculated based on conditional replenishment (repeating the frame). In this paper, we do not consider error concealment, although our methods can be generalized to include it. D_0 is the distortion when both layers can be reconstructed using R-S decoding. P_1 is the probability that the number of lost packets is between $n - m_2 + 1$ and $n - m_1$, which means the base layer is received correctly and part of the enhancement layer is lost. In this situation, although we cannot get the whole enhancement layer, the part before the first packet error will be useful because data is progressive. D_1 is the corresponding expected distortion.

$$D_1 = \sum_{j=0}^{n-i} p_j d_j \tag{8}$$

where p_j is the probability that the first packet error is in packet $j+1$ given i out of n packets have been lost, and d_j is the corresponding distortion. Let A be the event that the first error is in packet $j+1$ and B be the event that i out of n packets are lost. p_j can be calculated by:

$$\begin{aligned} p_j &= p(A|B) = \frac{p(A,B)}{P(i,n)} \\ &= \frac{p(A)R(i,n-j)}{P(i,n)} \end{aligned} \tag{9}$$

where

$$p(A) = \begin{cases} (1-P_b)P_{GB} + P_b(1-P_{BG}), & j = 0 \\ (1-P_b)(1-P_{GB})^j P_{GB} + P_b P_{BG}(1-P_{GB})^{j-1} P_{GB}, & 1 \leq j \leq n-i \end{cases} \tag{10}$$

The complexity of the above algorithm may be too high for some real-time applications. In this situation, equal error protection (EEP) can be used as a suboptimal but simple alternative.

3 Experimental Results

We use a MPEG-4 MoMuSys codec with FGS coding to demonstrate the performance of our scheme. Simulation results are obtained based on two CIF format sequences: Foreman and Coastguard. The frame rate is 10 frames/s. In our experiments, a BOP includes 10 frames corresponding to 1 second of video. The first frame is an I frame, and following frames are P frames. We use one BOP in each sequence, base layer rates are about 160 *kbps* (Foreman) and 125 *kbps* (Coastguard) respectively.

In Figure 3, the proposed method is compared with several fixed packet size schemes. The total available bandwidth is $480kbps$ and the average burst length $L_b = 9.97$. When the channel condition is good, long packets should be used, because error protection is a small part of the rate and the rate consumed by overhead is more important than the effectiveness of FEC. Compared to the shortest packet size (200 bytes), the optimal scheme achieves up to 1 dB improvement.

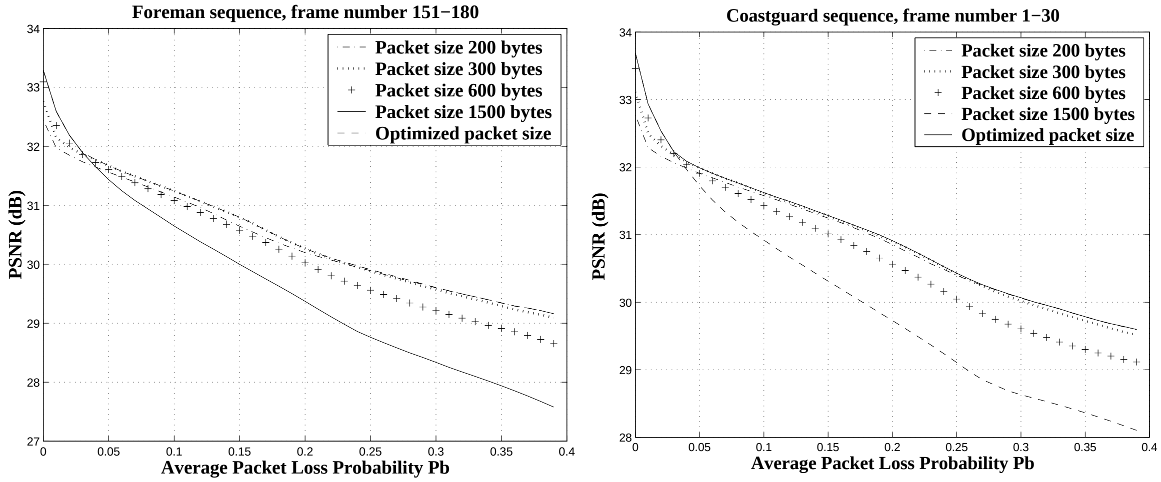

Figure 3: Performance using different packet sizes, Foreman (left), Coastguard (right)

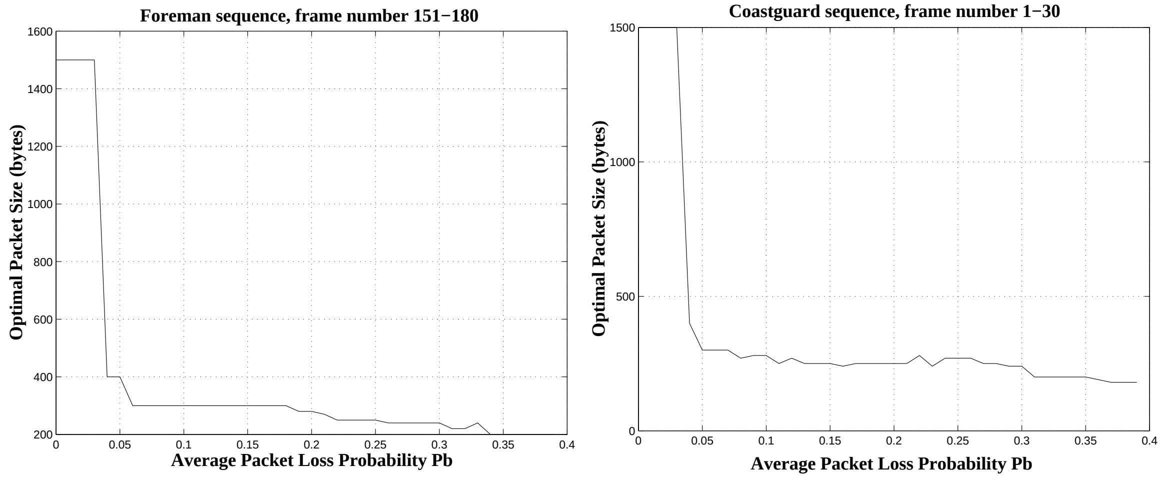

Figure 4: Optimal packet size for sequences Foreman (left) and Coastguard (right)

When the channel condition is bad, the efficiency of Reed-Solomon codes becomes more important, thus one must reduce the packet size (and suffer the overhead) to get better error protection. Compared to the longest packet size (1500 bytes), 0.5 dB improvement is achieved when $P_b = 0.1$ and 1.0 dB gain is achieved when $P_b = 0.2$. This tradeoff of packet size, driven by the role of overhead and efficiency of FEC, is demonstrated in Figure 4.

We observed an interesting phenomenon when the chosen rate for the base layer consumes most of the bandwidth. Ordinarily, a bad channel would require a small packet size, as indicated in experiments of Figure 3. However, when the base layer bitrate dominates the bandwidth, the algorithm dispenses with the enhancement layer and also increases packet length (decreases overhead) to make room for the base layer and whatever protection for it that can be afforded. This interesting phenomenon is illustrated in Figure 5. The Foreman sequence is used in this experiment. The total available rate is $240kbps$. The base layer rate is about $160 kbps$ which means $80 kbps$ bandwidth is available for R_{fb}, R_e, R_{fe}, and R_h. When $P_b \geq 0.05$, this $80 kbps$ cannot give the base layer enough protection, therefore, we have to reduce the rate of packet headers and allocate more rate to FEC codes of the base layer. This leads to longer packets.

Figure 6 gives the rate used by the enhancement layer for the Foreman sequence under different channel conditions. It shows that when the channel is good, we have enough rate to send and protect the base layer, so more rate will be allocated to the enhancement layer. When the packet loss probability increases, we have to increase the error protection for both layers, thus reduce the rate allocated to the enhancement layer.

Both the base layer and the enhancement layer are optimally protected in our algorithm. However, some schemes [5, 6] deem the enhancement layer unimportant and leave it unprotected. To demonstrate the importance of protecting the enhancement layer, we changed our algorithm by setting the rate of FEC codes for the enhancement layer to zero. Figure 7 shows the results. In this experiment, the total channel rate is $480 kbps$. We see that when the channel is almost perfect, i.e., $P_b \leq 0.03$, the quality of video is virtually no different with or without protecting the enhancement layer. However, in the majority of cases, it is best to protect the enhancement layer and take the corresponding FEC rate into account in the bit allocation problem. The results show that up to 1.5 dB improvement is achieved in this way. For Foreman sequence, when $P_B \geq 0.35$, two curves merge because no rate is allocated to the enhancement layer.

4 Conclusion

In this paper, the impact of packet size is studied in video transmission over a noisy channel. We show that packet size should be optimized to balance the effect of packet headers and the efficiency of FEC codes. We develop a probabilistic framework for the solution of rate allocation problem in the presence of packet overhead. Experimental results show that packet size has significant effect on the reconstructed video quality and network resources are utilized efficiently with our approach.

References

[1] R. Aravind, M. R. Civanlar, and A. Reibman, "Packet Loss Resilience of MPEG-2 Scalable Video Coding Algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 426–435, October 1996.

[2] S. R. McCanne, *Scalable Video Coding and Transmission for Internet Multicast Video.* PhD thesis, University of California, Berkeley, 1996.

[3] U. Horn, K. Stuhlmuller, and B. Girod, "Robust Internet Video Transmission Based on Scalable Coding and Unequal Error Protection," *Signal Processing: Image Communications*, vol. 15, pp. 77–94, September 1999.

[4] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Resource Allocation for Multimedia Streaming Over the Internet," *IEEE Transactions on Multimedia*, vol. 3, pp. 339–355, September 2001.

[5] J. M. Boyce, "Packet Loss Resilient Transmission of MPEG Video over the Internet," *Signal Processing: Image Communications*, vol. 15, pp. 7–24, September 1999.

[6] M. Gallant and F. Kossentini, "Rate-Distortion Optimized Layered Coding with Unequal Error Protection for Robust Internet Video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 357–372, March 2001.

[7] ISO/IEC 14496-2, "Coding of Audio-Visual Objects, Part2: Visual," July 2000.

[8] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 301–317, March 2001.

[9] E. O. Elliott, "A Model of the Switched Telephone Network for Data Communications," *The Bell System Technical Journal*, vol. 44, pp. 89–109, January 1965.

[10] S. Wenger, "Proposed error patterns for intenet experiments," *ITU-T Study Group 16 H.263+ Video Experts Group*, vol. Q15109, October 1999.

[11] H.Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-time Applications," In RFC 1889, Internet Engineering Task Force, January 1996.

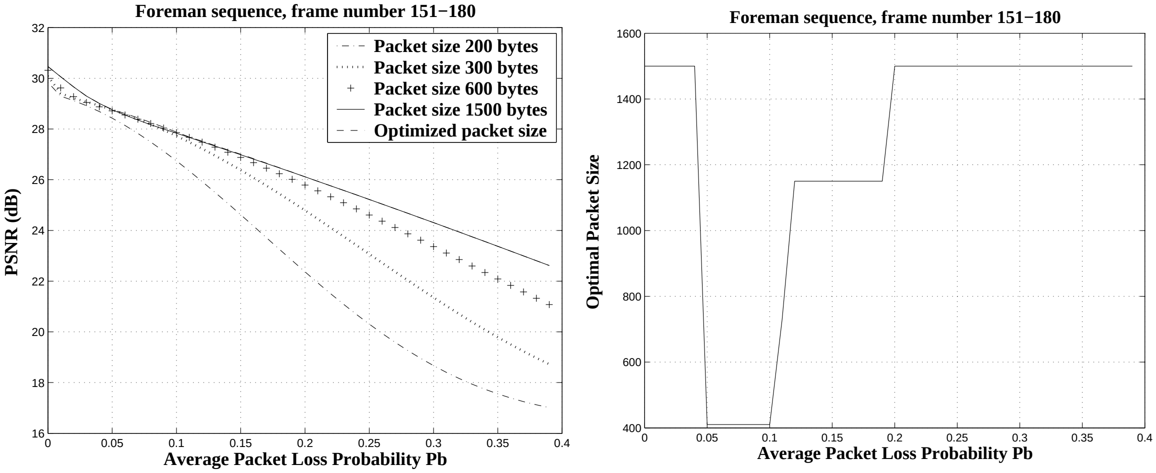

Figure 5: Performance (left) and optimal packet size (right), $R = 240kbps$

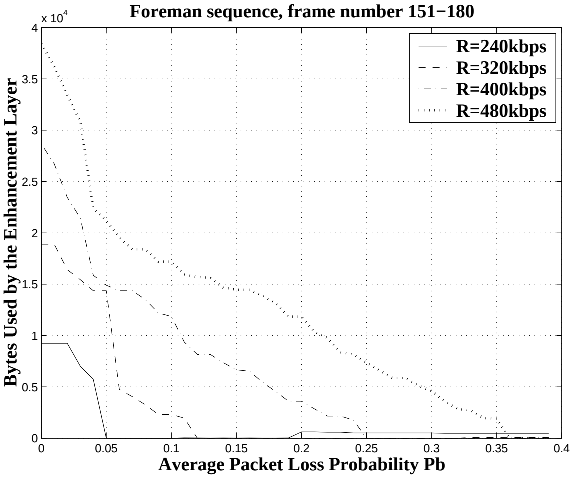

Figure 6: Bandwidth used by the enhancement layer

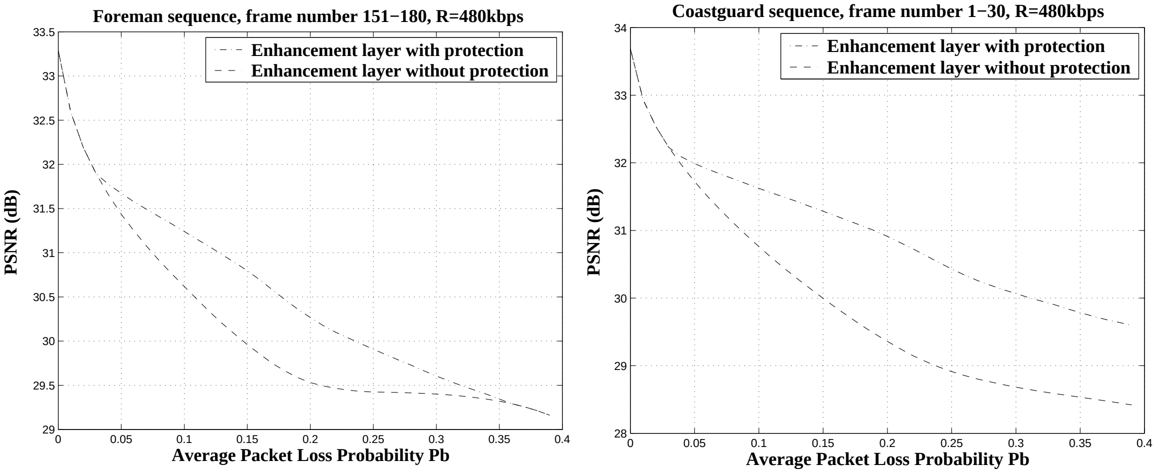

Figure 7: Performance (left) and optimal packet size (right), $R = 240kbps$