MINIMIZING DISTORTION FOR MULTI-PATH VIDEO STREAMING OVER AD HOC NETWORKS

Eric Setton, Xiaoqing Zhu and Bernd Girod

Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA94305-9510, USA
{esetton, zhuxq, bgirod}@stanford.edu

Invited Paper

ABSTRACT

We present a video distortion model analyzing the performance of multi-path routing for low latency video streaming, in congestion-limited ad hoc networks. In such environments, a single node transmitting multimedia data may have an impact on the overall network conditions and may need to limit its rate to achieve the highest sustainable video quality. For this purpose, optimal routing which seeks to minimize congestion and distributes traffic over multiple paths is attractive. To predict the end-to-end rate-distortion tradeoff, we develop a model which captures the impact of quantization and of packet loss on the overall video quality. Network simulations are performed to confirm the validity of the model in different streaming scenarios over different numbers of paths.

1. INTRODUCTION

In ad hoc networks, nodes self-organize to create a mesh, in which each node can act as a source, a destination or a relay for traffic. The flexibility offered in such networks may be leveraged in a variety of contexts. For example, in disaster areas or in search-and-rescue operations, it is very appealing to be able to rapidly deploy a wireless ad hoc network without the need of a fixed infrastructure. However, because of adverse channel conditions and potential node mobility, traditional networking tasks such as routing can be challenging even when the number of nodes is limited.

Routing algorithms generally share the goal of minimizing the number of hops between source and destination, and several distributed protocols have been successfully tailored to wireless ad hoc networks [1]. An alternate routing strategy is to minimize the overall network congestion. The solution to this classic problem, analyzed in particular in [2], relies on multiple paths to efficiently make use of the network resources. This property is important for streaming time-critical multimedia data over a bandwidth-constrained network since the high data rate might overwise saturate an individual link.

Prior work on distortion models for streamed packetized video has focused on the rate-distortion tradeoff on error-prone channels [3]. Recently, work has also been done to analyze the impact of different error patterns on the reconstructed video quality [4]. In both cases, however, the bandwidth of the network is over-provisioned and the influence of the client on the network is not considered. This is not the case in a wireless ad hoc network, where the action of a single user may affect the overall conditions. The effect of network congestion on streamed video quality has been discussed in our previous work [5], where a congestion-minimized multi-path routing scheme is proposed to make efficient use of the bandwidth. In this paper, we emphasize the video distortion model, and present additional experimental results which confirm its validity.

The rest of the paper is organized as follows. In the next section, we present a video distortion model which captures the influence of both the encoder performance and network congestion. In Section 3, we briefly describe the congestion-optimized routing scheme, used in our experiments, for video streaming over ad hoc networks. Simulation results for two video sequences are presented in Section 4, where we discuss the impact of different parameters on the reconstructed video quality.

2. VIDEO DISTORTION MODEL

For live video streaming applications, video packets are transmitted over the network and need to meet a playout deadline. Decoded video quality at the receiver is therefore affected by two factors: encoder compression performance and distortion due to packet loss or late arrivals. Assum-

This work was supported, in part, by National Science Foundation Grant CCR-0325639.

ing an additive relation of these two independent factors, a video distortion model can be derived based on [3].

With the *Mean Squared Error (MSE)* criterion, distortion of the decoded video is:

$$D_{dec} = D_{enc} + D_{loss}. \tag{1}$$

The encoder distortion may be modelled by :

$$D_{enc} = D_0 + \theta/(R - R_0), \tag{2}$$

where R is the rate of the video stream, and the parameters D_0, θ and R_0 are estimated from empirical rate-distortion curves via regression techniques.

In addition, we suppose, as in [3], that D_{loss} is linearly related to the packet loss rate :

$$D_{loss} = \kappa P_{loss}, \tag{3}$$

where the scaling factor κ depends on the encoding structure, e.g., the ratio of intra-coded blocks, and P_{loss} reflects the combined rate of random losses and late arrivals.

In the M/M/1 model, the delay distribution of packets on a single link is exponential :

$$\mathbf{Prob}\{Delay > T\} = e^{-\lambda T}, \tag{4}$$

where λ is determined by the average delay:

$$\mathbf{E}\{Delay\} = 1/\lambda = L/(C - R). \tag{5}$$

In (5), C is the capacity of the link, R is the rate, and L is the packet size. The case that we consider differs from this scenario: video packets are sent at regular intervals over multiple network paths, each composed of multiple hops. Nevertheless, we suggest modelling the distribution by an exponential :

$$\mathbf{Prob}\{Delay > T\} = e^{-(C'-R)T/L'}, \tag{6}$$

in which case C' and L' are determined empirically. Together with a random packet loss rate P_r due to transmission errors, the total packet loss rate is:

$$P_{loss} = P_r + (1 - P_r)e^{-(C'-R)T/L'}, \tag{7}$$

where T denotes the maximum delay allowed for each packet, due to its playout deadline. Combining (1)-(7), the received video distortion can be expressed as:

$$\begin{aligned} D_{dec} &= D_0 + \theta/(R - R_0) + \\ &\quad \kappa(P_r + (1 - P_r)e^{-(C'-R)T/L'}). \end{aligned} \tag{8}$$

The proposed formula models the impact of the rate on the video distortion. At lower rates, reconstructed video quality is limited by coarse quantization, whereas at high rates, the video traffic leads to higher network congestion, hence more packet drops due to delays and reduced reconstruction quality. For live video steaming in a bandwidth-limited environment, we therefore expect to achieve maximum decoded quality for some intermediate rate.

3. CONGESTION-MINIMIZED MULTI-PATH ROUTING

The problem of congestion-optimized routing on a generic network has been studied extensively, see, e.g. [2]. In [5], we analyze the benefits of such schemes for low latency video streaming and focus in particular on optimal traffic partitioning over a set of predetermined paths. This problem has a low complexity convex formulation and can be solved efficiently to the global optimum.

This type of routing has two main advantages. The use of multiple paths in a bandwidth-limited network allows the video to be transmitted at a higher rate, without overwhelming an individual link. When the set of routes is not independent, congestion-optimized traffic partitioning still achieves efficient bandwidth utilization and outperforms heuristics such as load balancing. The benefit of multi-path routing observed in [5] is different from the increase in robustness provided by combining path diversity and multiple description coding, which initially motivated multi-path video streaming [6].

In the experimental results presented in the next section, we will use the congestion-optimized traffic partitioning scheme. As an example, Fig. 1 illustrates the solution for streaming 100 kbps of data from Node 1 to Node 5 in a network of 15 nodes randomly placed in a 100m-by-100m square. Note however that our video distortion model may be applied to other single-path or multi-path routing algorithm as shown in [5].

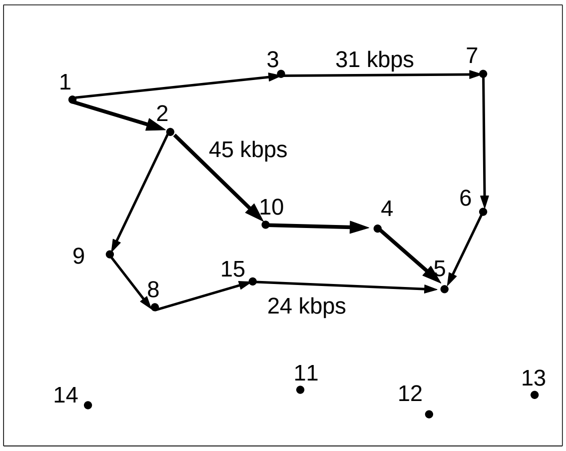

Fig. 1. Congestion-optimized traffic partitioning of 100 kbps over 3 paths

4. SIMULATION RESULTS

Simulations are performed over the 15-node network described in Section 3 using the network simulator ns-2 [7].

Video traffic is streamed from Node 1 to Node 5. We use the Constant Bit Rate (CBR) model in ns-2 for video traffic and assume each frame fits in one packet. Cross traffic is specified as a random process, occupying up to 50% of the capacity on each link, with exponentially distributed packet sizes and arrivals. We generate and decode the QCIF video sequences *Foreman* and *Mother and Daughter* using an H.264 codec, at 30 frames per second at various quantization levels and for different *Group of Pictures (GOP)* lengths. Packets are dropped if they do not arrive at the receiver by their playout deadline. In this case previous frame concealment is used. Decoded video quality is measured in terms of *Peak-Signal-to-Noise-Ratio (PSNR)* of the luminance component.

Figure 2 illustrates how the model parameters are determined. The fit of the encoder distortion is shown for a GOP length of 15. The alignment of empirical points in the second graph of Fig. 2 justifies the exponential model chosen for the delay distribution. These results are obtained for streaming the video packets over 3 paths as described in Section 3; the playout deadline is 350 ms. Similar observations are made for different numbers of paths, GOP lengths and playout deadlines. The encoder distortion parameters D_0, R_0 and θ, are sequence dependent and can be estimated offline separately from C' and L' which only depend on the network condition. This is also true for the sensitivity to packet loss, κ.

Figure 3 shows the tradeoff between decoded video quality and transmission rate for both sequences. Video packets are streamed over 6 paths and the playout deadline is set to 350 ms or 500 ms. In all the curves, the low rate region follows closely the encoder rate-distortion performance as few packets experience excessive delays. When the rate exceeds a certain threshold the decoded video quality quickly degrades, as predicted by the model. For a shorter playout deadline, the system is more susceptible to congestion and this degradation occurs at lower rates. The sharpness of the degradation depends on the quality of the

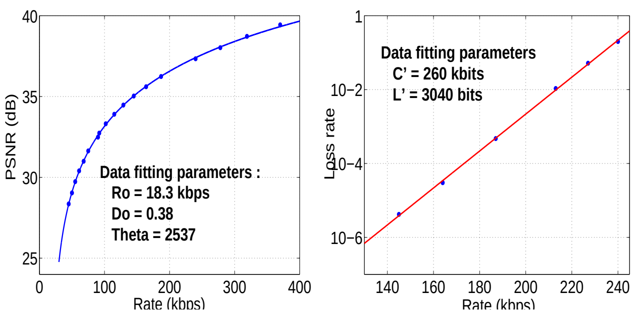

Fig. 2. Fitting the encoder distortion is done via nonlinear regression; the exponential fit of the delay distribution is a simple linear regression of its logarithm.

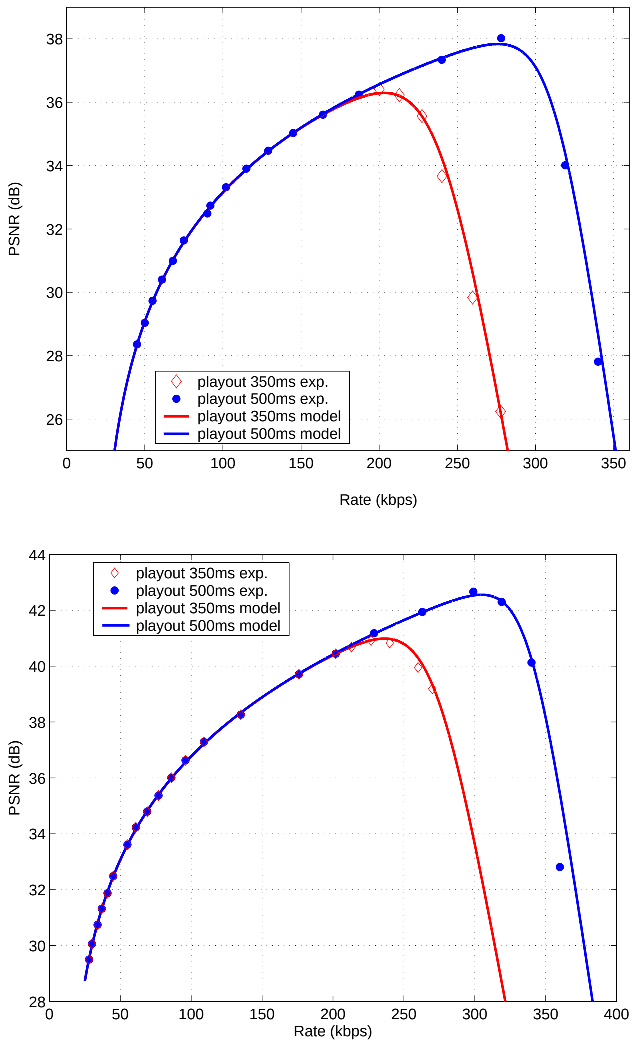

Fig. 3. Rate-PSNR performance for live video streaming of *Foreman* (top) and *Mother and Daughter* (bottom) using 6-path routing with congestion-minimized traffic partition.

error concealment performed in the case of packet loss. Due to the motion in the *Foreman* sequence, errors are more difficult to conceal than in the *Mother and Daughter* sequence. Therefore, the quality drop beyond the optimal operation point is sharper.

Figures 4 and 5 compare the influence of the GOP length for 3-path routing, for the *Foreman* sequence. Figure 4 shows the results for transmission over a network where packets are lost only due to congestion (and not due to channel errors). Not surprisingly, streams with longer GOPs have better compression efficiency, but are less resilient to network congestion at high rates. Note that for any given quality, the minimum rate is always achieved by encoding the video with longer GOPs.

This does not hold when a random packet loss rate of

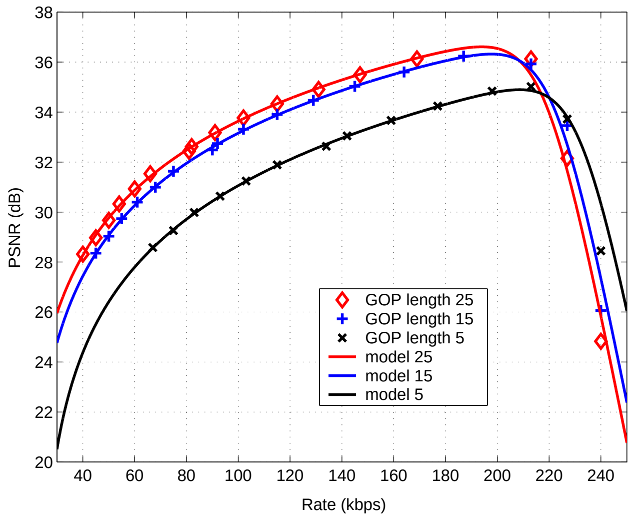

Fig. 4. Rate-distortion performance for live video streaming using 3-path optimal routing and different coding structures.

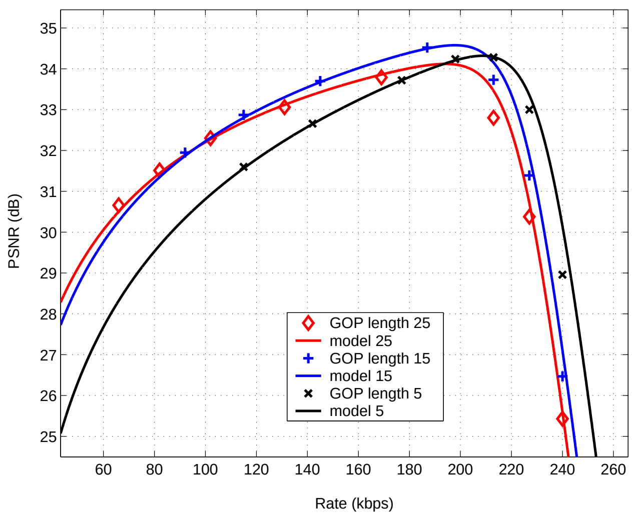

Fig. 5. Rate-distortion performance for live video streaming using 3-path optimal routing and different coding structures, with 1% of random packet loss

1% is introduced on each path as shown in Figure 5. At very low bit rates, there is no loss due to congestion and the best GOP length is 25, as this stream is encoded with the highest efficiency. As the rate increases, however, the video stream with the longest GOP suffers most severely from packet drops, whereas the intermediate GOP length of 15 achieves the best tradeoff between coding efficiency and robustness against network congestion. For the two sets of experiments, the model parameters are the same and can predict the performance of the different coding structures.

5. CONCLUSION

We propose a video distortion model for live video streaming in wireless ad hoc networks. The model incorporates contributions from both encoder distortion and packet loss due to congestion. In this context the optimal operating bit rate is the one for which the compressed video stream achieves highest quality without creating significant congestion on the network. Experimental results for two sequences are presented for different number of paths, loss rates, coding structures and playout deadlines. The model captures the influence of these different parameters and can be used to predict the end-to-end rate-distortion performance.

6. ACKNOWLEDGMENTS

The authors would like to thank Professors Andrea Goldsmith and Fouad Tobagi at Stanford University for many insightful discussions.

7. REFERENCES

[1] Elizabeth M. Royer and Chai-Keong Toh, "A review of current routing protocols for ad hoc mobile wireless networks," *IEEE Personal Communications*, pp. 46–55, April 1999.

[2] Leonard Kleinrock, "Queuing Systems, Volume II: Computer Applications," *Wiley Interscience, NY*, 1976.

[3] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–32, June 2000.

[4] Yi J. Liang, John G. Apostolopoulos, and Bernd Girod, "Analysis of packet loss for compressed video: does burst length matter?," *IEEE International Conference on Acoustics, Speech and Signal Processing, Hong Kong*, April 2003.

[5] Eric Setton, Xiaoqing Zhu, and Bernd Girod, "Congestion-optimized multipath streaming of video over ad hoc wireless networks," *International Conference on Multimedia and Expo 2004, submitted*.

[6] John G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *Proc. VCIP, San Jose, USA*, pp. 392–409, Jan. 2001.

[7] "The Network Simulator - ns-2," *www.isi.edu/nsnam/ns/*.