# CONGESTION-OPTIMIZED ROUTING AND SCHEDULING OF VIDEO OVER WIRELESS AD HOC NETWORKS

*Bernd Girod, Eric Setton and Xiaoqing Zhu*

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

*Invited Paper*

## ABSTRACT

We propose a dynamic scheduling and traffic partitioning scheme for a video source communicating over a mobile ad hoc wireless network. This cross-layer design approach operates at both the network layer and the application layer. It indicates along which route(s) the sender should operate in order to obtain the highest allowable rate while limiting self-congestion. In addition, the congestion-distortion optimized scheduling takes fully advantage of this rate allocation to select and transmit the packets of the video stream which will maximize video quality. Experiments performed over a simulated network, with H.264 encoded video, show performance gains for the proposed technique.

## 1. INTRODUCTION

Recently, there has been a growing interest in ad hoc wireless networks, where nodes communicate with each other without the support of a fixed infrastructure. In such a network, each node can act as a source, a destination, or as a relay for the traffic of other nodes. While this provides appealing features of rapid deployment and flexible configuration for many applications, technical challenges also arise when the network is required to support real-time media applications such as video streaming. The demanding rate and delay constraints are difficult to accommodate in a network where bandwidth is scarce and shared among users, node mobility may require frequent route updates and energy is limited at each node. The combination of these challenges calls for novel design paradigms.

Unlike in the conventional network structure with protocols independently designed for each layer, cross-layer design allows information sharing across the different layers, and considers adaptive power control, media access control, routing, scheduling and source coding jointly for efficient utilization of network resources. This idea has been explored both in its theoretical aspects, with convex optimization formulations [1][2], and for more practical systems concerns [3][4].

We propose to exploit the advantages of cross-layer information exchange for designing both the network layer and the application layer of a video streaming system operating in a mobile wireless ad hoc network using the 802.11 protocol. In this situation, instead of adjusting the rate of a user independently at the transport layer, as the congestion control in TCP [5] or TCP-Friendly Rate Control (TFRC) [6], one can do so according to the importance of the transmitted content. For video streaming specifically, this importance may be reflected in the rate-distortion (RD) characteristics of the encoded video.

The aim of the paper is to allocate rate optimally to the source and take fully advantage of this rate via smart scheduling. The network layer traffic partitioning algorithm will consider the tradeoff between low rates which do not allow high quality video streams and high rates which create too much self-congestion to allow for a timely delivery of the media packets. At the application layer, the scheduler will decide which video packets to transmit to yield a high video quality while abiding to the rate constraint. One of the questions we will try to answer is whether transmitting all the packets of a lower quality video stream is better than streaming a pruned representation encoded at a higher quality.

The rest of the paper is organized as follows. In the next section, we describe a rate-distortion model of video which captures both the encoder performance and the impact of self-congestion on the video streaming. The algorithm for congestion-distortion optimized

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

scheduling over a multi-hop wireless path is explained in Section 3. In Section 4, we apply the proposed scheme and compare its performance with other approaches.

## 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. Assuming an additive relation of these two independent factors, a video distortion model can be derived based on [7].

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 $D_0$, $\theta$ and $R_0$ are estimated from empirical rate-distortion curves via regression techniques.

As explained in [8], we propose to model the distortion due to packet loss by:

$$D_{loss} = \kappa(P_r + (1 - P_r)e^{-(C-R)T/L}), \tag{3}$$

where the scaling factor $\kappa$ depends on the encoding structure, e.g., the ratio of intra-coded blocks, and the other factor reflects the combined rate of random losses and late arrivals. In (3), $P_r$ and $T$ respectively reflect the random packet loss rate and the time within which each packet should reach the receiver (typically a few hundreds of milliseconds). The parameters $C$ and $L$ are related to the maximum allowable rate and the average packet size, however they need to be determined dynamically for a given route.

Combining the different equations, the received video distortion can be expressed as:

$$\begin{aligned} D_{dec} &= D_0 + \theta/(R - R_0) + \\ &\quad \kappa(P_r + (1 - P_r)e^{-(C-R)T/L}). \end{aligned} \tag{4}$$

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-DISTORTION OPTIMIZED SCHEDULING

In this section, we describe how to determine an optimal transmission schedule for the packets of a video stream. This schedule indicates when the packets of the stream will be sent from the server to the client, if at all, assuming a discrete set of transmission times. To limit the exponential number of possible schedules, the time horizon covered by the schedule is limited. Furthermore, rather than optimizing jointly the schedule for all the packets of the stream, only a small number of packets are selected and the optimization is performed iteratively for each packet.

The aim of CoDiO is to determine a schedule minimizing the expected Lagrangian cost $D + \lambda\Delta$, where $D$ is the distortion of the received video stream and $\Delta$ is the end-to-end delay which serves as the congestion metric. To minimize this objective function, CoDiO selects the most important packets in terms of video quality, and transmits them in an order which minimizes the average backlog of the bottleneck queue. For example, the I frames of a video stream will be transmitted in priority whereas B frames might be dropped. In addition, CoDiO will avoid transmitting packets in large bursts as this has the worse effect on the queuing delay. In the following, we briefly describe how to evaluate the expected end-to-end delay and distortion corresponding to a given transmission schedule of a set of packets. This elementary step is repeated several times to evaluate the performance of different schedules and choose the schedule which performs best.

### 3.1. Determining the end-to-end delay

Unlike rate in the RaDiO framework [9], end-to-end delay is not additive and the contribution of each packet cannot be derived separately. For a given transmission schedule, the rate output by the server may be used to derive the average delay on the network path, given the capacity of this path. This in turn leads to the average value of the end-to-end delay over the time horizon considered.

In the scenario considered, the network layer, based on the video distortion model derived in the previous section, determines an optimal operating rate $R^*$. The CoDiO scheduler considers this rate as an average rate constraint. For this purpose the route chosen by the network layer is modelled as a succession of high bandwidth links followed by a bottleneck link of capacity $R^*$. Congestion-distortion optimized scheduling is then performed for this hypothetic path. More details on the algorithm can be found in [10].

To derive the queuing delay on this hypothetic bottleneck queue we assume the delay for a packet is constant until it reaches the bottleneck. Knowing the capacity of the virtual bottleneck link and the transmitted rate, the size of the queue at this virtual link may easily be computed. A typical illustration of the size of the queue as a function of time is shown in Fig. 1. For practical purposes, as the traffic conditions on the network path may vary and to account for the inexactitude of the model, the bandwidth of the link is corrected dynamically based on round-trip time estimates computed from acknowledgement packets.

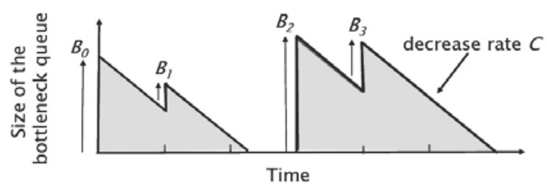


**Fig. 1**. Backlog at the bottleneck queue

### 3.2. Determining the video distortion

The expected value of the distortion for the video stream decoded by the client is computed as in [11]. Namely, if copy error concealment is used, an undecodable frame is replaced with the nearest correctly decoded frame for display. Hence, to capture the effect of packet loss on the video quality, only a limited number of display outcomes need to be identified and associated with different distortions. Let $D(s, f)$ denote the distortion resulting from substituting frame $s$ to frame $f$, the expected distortion when displaying frame $f$ is:

$$D(f) = \sum_{s} D(s, f) Pr\{s\} \quad (5)$$

In Eq. (5), $Pr\{s\}$ represents the probability that frame $s$ is displayed instead of $f$. This probability may be computed, as described in [11], by combining the probabilities that different packets do not reach the client by their playout deadline.

## 4. SIMULATION RESULTS

We use the network simulator NS-2 [12] to evaluate the performance of the proposed cross-layer joint optimization scheme. The simulated network consists of 10 mobile nodes within a 100m-by-100m square. Each node follows the random walk mobility model with average speed 3 m/s. Routing is updated once every 1.0 s in both schemes. The entire simulation duration is 100 s.

A video streaming session is set up from Node 3 to Node 5. The *Foreman* CIF video sequence is encoded by the H.264 codec at 30 frames per second at different qualities, using a typical IBBP... coding structure with GOP length 16. The playout deadline is 150 ms, typical for live streaming scenarios. In the experiments, packet losses are caused by link failures or overflow of transmission queues due to congestion. Packets arriving at the receiver after their deadlines are discarded. Error concealment is performed by replacing an undecodable frame with the nearest correctly decoded frame.

## 5. CONCLUSION

In this paper, we propose a cross-layer design for the network layer and the application layer of a video streaming system. This design optimizes both the routing and the scheduling of a video stream transmitted over a wireless ad hoc network. In this environment, the impact of self-congestion needs to be taken into account to derive an optimal operating rate. Subsequently, a scheduler determines which packets to transmit, and when, to optimize the video quality while abiding to this loose rate constraint. Network experiments show the benefits of such an approach.

## 6. REFERENCES

[1] Y. Wu, P. Chou, Q. Zhang, W. Zhu K. Jain, and S-Y. Kung, "Network planning in wireless ad hoc networks: a cross-layer approach," *IEEE Journal on Selected Areas on Communications, submitted October 2003*.

[2] R. L. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multi-hop wireless networks," *Proc. INFOCOM, San Francisco, USA*, pp. 702–711, Mar. 2003.

[3] S. Mao, S. Lin, S. S. Panar, Y. Wang, and E. Celebi, "Video transport over ad hoc networks: Multistream coding with multipath transport," vol. 21, pp. 1721 – 1737, December 2003.

[4] T. Yoo, E. Setton, X. Zhu, A. Goldsmith, and B. Girod, "Cross-layer design for video streaming over wireless ad hoc networks," in *Multimedia Signal Processing Workshop (MMSP), Sienna, Italy*, October 2004.

[5] M. Allman, V. Paxson, and W. R. Stevens, "Tcp congestion control," *RFC 2581*, Apr. 1999.

[6] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "Tcp friendly rate control (tfrc): Protocol specification," *RFC 3448*, Jan. 2003.

[7] 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.

[8] Eric Setton, Xiaoqing Zhu, and Bernd Girod, "Minimizing distortion for multipath video streaming over ad hoc networks," *International Conference on Image Processing 2004, submitted*.

[9] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *Microsoft Research Technical Report MSR-TR-2001-35*, Feb. 2001.

[10] E. Setton and B. Girod, "Congestion-Distortion Optimized Scheduling of Video," *MMSP 2004, to be submitted*.

[11] M. Kalman, P. Ramanathan, and B. Girod, "Rate-distortion optimized streaming with multiple deadlines," *Proc. International Conference on Image Processing, Barcelona, Spain*, Sept. 2003.

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