II.4. Feedback of Rate and Loss Information for Networked Video

Prof. Chia-Wen Lin
Department of CS
National Chung Cheng University
886-5-272-0411 ext. 33120
cwlin@cs.ccu.edu.tw

Outline

- Motivation for feedback and adaptation
- Transport of video over the best-effort internet
- Source adaptation using rate-based feedback control
- Rate characterization and smoothing
- The SAVE algorithm
- Evaluation criteria
- Experiments and simulation results
- Conclusion

Motivation for Feedback and Adaptation

- Three drawbacks of an open-loop transport of video :
  - The requirement for conservative admission control
  - Large buffers (impacting delay)
  - The potential for packet loss (impacting quality)

Motivation for Feedback and Adaptation (Cont.)

Transport of Video over the Best-Effort Internet

- Two types of transporting video over the best-effort internet :
  - Unicast (Point to Point)
  - Multicast (Point to Multipoint)

The AIMD Rate Control Algorithm

If (Congested)

λ = max{$\lambda \times \alpha$, minimum_rate};

If (UnCongested)

λ = max{$\lambda + \alpha$, minimum_rate};

λ: source rate;
α: reduction factor ($\leq$ 1);
β: increase factor

Transport of Video over the Best-Effort Internet

- RAP : Involves a rate-adaptive source at a server and a corresponding rate-adaptive receiver at every client, with the receiver acknowledging every packet
- The interpacket gap (IPG) is adjusted in RAP in response to congestion.The adjusting strategy is the additive increase, multiplicative decrease
- Real-time streams that do not use TCP (e.g., they may use UDP) may gain an unfair advantage over TCP-transported data
  - Possible Solution: “TCP-friendly” protocols

Supporting Video in A Multicast Environment

- McCanne proposed to move the rate adaptation to the receivers by means of receiver-driven layered multicast (RLM).
  - Combines layered video compression with a layered transmission scheme
  - RLM receivers use join experiments to add a new layer at well chosen times
  - RLM groups use shared learning to improve join experiments.

Source Adaptation Using Rate-Based Control

- Constraints on rate required for transporting real-time video
- Rate renegotiation mechanisms and adaptation

Rate Constraints for Transporting Real-time Video

- Buffering and delay constraints: (Buffer underflow or overflow)
  - End-to-end delay
- Source adaptation constraints
  - Received quality, delays, …
- Lookahead constraints (e.g., for the stored video)
- Implicit versus explicit feedback
- Signaling frequency and latency constraints
- Rate prediction error
  - Uncertainties on the network and source rates

Rate Renegotiation Mechanism and Adaptation

- RCBR uses renegotiation of the traffic parameters of a constant bit rate connection by means of ATM signaling messages.
- The control mechanism proposed by Kanakia et al. (3b) is based on predicting the evolution of the system over time and using that prediction to compute a target sending rate for each frame of video data.

Rate Renegotiation Mechanism and Adaptation

If $X_{n-k} = 0$

$\lambda_n = \lambda_{n-1} + \delta$

else

$\lambda_n = \mu_n + (X^* - X_n)/(\text{gain} * F)$

X^* : The target value of the buffer occupancy at the bottleneck for this video flow.

$1/F$: The frame rate for this video

μ_n : The service rate

X_n : The queue occupancy

Rate Characterization and Smoothing

$S(t) = \sum s_i$ is the cumulative amount of data sent up to time t

l_i denote the sizes of frame in a stream of N video frames

To avoid of underflow requires that $L(t) \leqq S(t)$

To avoid of overflow requires that

$S(t) \leqq U(t) := \min\{L(t-1) + B, L(N)\}$

The SAVE Algorithm

- SAVE (Smoothed Adaptive Video over Explicit rate networks) is a source algorithm used for transporting compressed video in conjunction with explicit rate-based feedback control in the network.
- This algorithm comprises two parts
  – The rate request algorithm : specifies how the source requests bandwidth from the system
  – The frame quantization algorithm : specifies how the frame size are controlled to avoid excessive delay

Heuristics for the Requested Rate

- Because of the GOP structure, one determinant of the required rate will be the smoothed rate :

  $r_{sm} = f_{sm}/\tau$

  τ is the interframe time

  f_{sm} is the ideal frame size that required by the encoder to encode the frame at ideal perceptual lossless quality

Rate Request & Frame Quantization Algorithms

$$r_{sm}(n) = (\tau\, w_{sm})^{-1} \sum_{i=0}^{w_{sm}-1} f(n-i)$$

$$r_{max}(n) = (\tau\, w_{max})^{-1} \max f(n-I)$$

$$i = 0, 1, \ldots, w_{max}-1$$

- The requested rate at frame n is :

  $r_{req}(n) = \beta \max \{r_{sm}(n), r_{max}(n), r_{ar}(n)\}$

- The encoded size of frame n is :

  $f_{enc}(n) = \min \{f(n), \max\{f_{avail}(n-1), r \times f(n)\}\}$

Evaluation Criteria

- Source delay
- Quality and Adaptation
- Robustness to Network Feedback Delay
- Channel Capacity and Multiplexing Gain
- Sensitivity to Algorithm Parameters

Evaluation Criteria (Cont.)

- Source delay :

  The primary concern for the work is the aggregate delay introduced in the source buffer and the network.

- Quality and Adaptation :

  Cropping : The reduction from the ideal rate to the encoded rate.

  Evaluate the pattern of cropping over the entire sequence of frames.

Evaluation Criteria (Cont.)

- Robustness to Network Feedback Delay :

  The rate allocation mechanism of the explicit rate network is not expected to instantaneously allocate a rate in response to requests.
- Channel Capacity and Multiplexing Gain :
- Sensitivity to Algorithm Parameters

Experiments and Simulation Results

- Modeling network characteristics with frame-level simulations

  if $n > \delta$

  $r_{all}(n) = r_{req}(n - \delta)$

  else

  $r_{all}(n) = r_0$

Experiments and Simulation Results

- Modeling network characteristics with frame-level simulations

$p(n) = \max\{1, C/R(n)\}$

if $n > \delta$

$r_{all}(\text{n}) = \text{P}(n - \delta)\text{r}_{req}(n - \delta)$

else

$r_{all}(\text{n}) = r_0$

Experiments and Simulation Results

- Modeling network characteristics with frame-level simulations

The buffer evolution is :

$b(n) = f_{enc}(n) + \max\{0, b(n-1) - r_{all}(n-1) \times \tau\}$

Network Configuration of Cell-Level Simulation

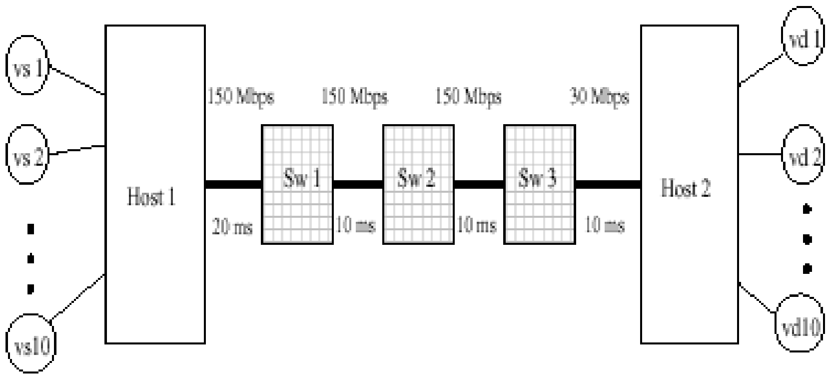
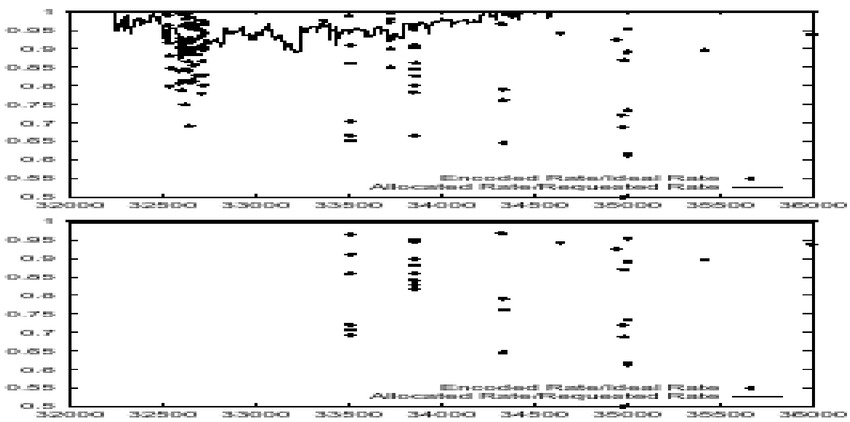
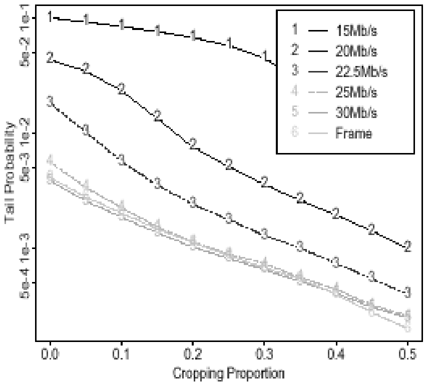


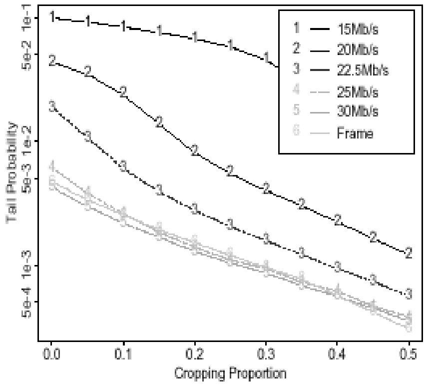
Figure 3: NETWORK CONFIGURATION FOR CELL-LEVEL SIMULATION

Impact of Network Congestion & Feedback Delay on Quality



Behavior of Frame Delay





Behavior of Frame Delay (Cont.)

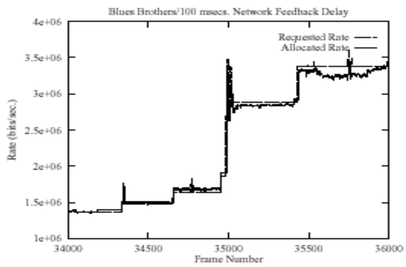


Figure 6: NETWORK SIMULATION RATES. Trace A, 1 trace within a multiplexed set. For short periods, allocated rate is lower than requested due to network contention.

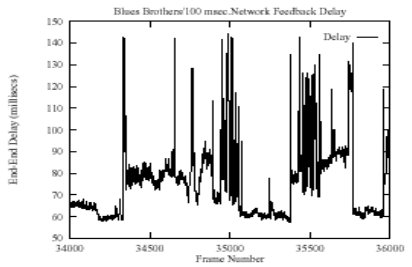


Figure 7: END-TO-END NETWORK DELAY. Trace A, 1 trace with a multiplexed set.

Multiplexing Gain & Impact on Quality

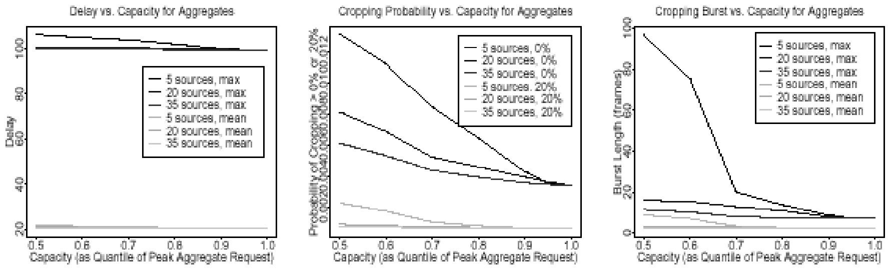
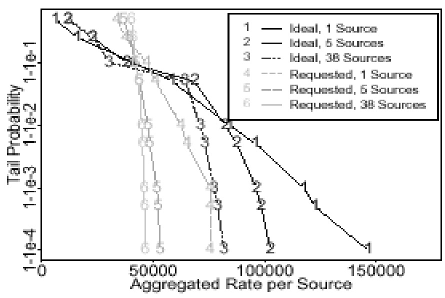


Figure 8: VARIATION WITH AGGREGATION SIZE OF IMPACT OF RATE-REDUCTION THROUGH REDUCED CAPACITY. For aggregations of 5, 20, 35 sources, averaged impact on single traces within aggregate; capacity expressed as quantile of peak aggregate requested rate. TOP: mean and max. delay. MIDDLE: proportion of cropping > 0% and > 20%. BOTTOM: mean and max. burst length of cropping > 20%.

Multiplexing Gain & Impact on Quality (Cont.)



igure 9: RATE SMOOTHING AND AGGREGATION Tail distribution, for ideal and requested rates per ource, for aggregations of 1,5 and 38 sources of Set E. High quantiles of requested rate are lower for equested rate.

Multiplexing Gain & Impact on Quality (Cont.)

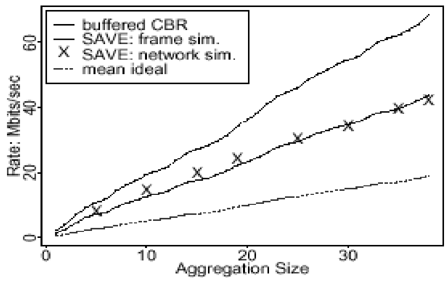


Figure 10: UTILITY OF EXPLICIT RATE FOR MULTIPLEXING GAIN. Bandwidth requirements, normalized by mean ideal rate, for 1- to 38-fold aggregates of Set E. LOWER CURVE: SAVE, 90^{th} percentile of aggregate requested rate. UPPER CURVE, CBR rate for 100ms source buffer. Also shown: mean ideal rate, and capacity estimate based on full network simulation described below.

Multiplexing Gain & Impact on Quality (Cont.)

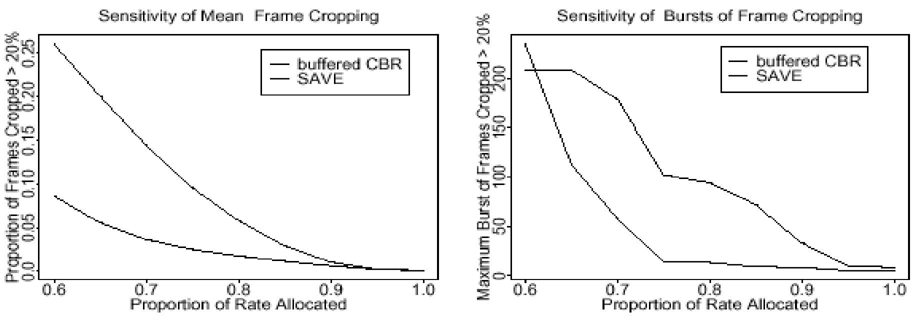


Figure 11: COMPARATIVE SENSITIVITY OF FRAME CROPPING IN SAVE AND BUFFERED CBR TO SYSTEMATIC RATE-REDUCTION OR UNDERALLOCATION OF BANDWIDTH. LEFT: average cropping more sensitive for SAVE. RIGHT: Burst cropping is more sensitive for buffered CBR.

Sensitivity to Smoothing Parameters

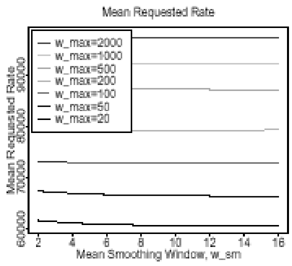


Figure 12: MEAN REQUESTED RATE VS. w_{sm}, trace set A, for varying w_{max}

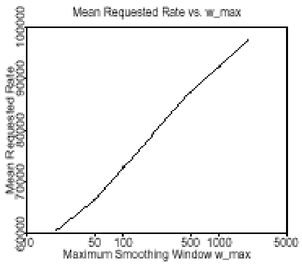


Figure 13: MEAN REQUESTED RATE VS. w_{max}, averaged over w_{sm}

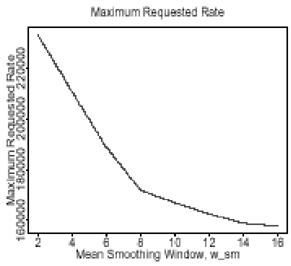


Figure 14: MAXIMUM REQUESTED RATE VS. w_{sm}, trace set A. Curves coincide for w_{max} from 20 to 2000.

Sensitivity to Smoothing Parameters (Cont.)

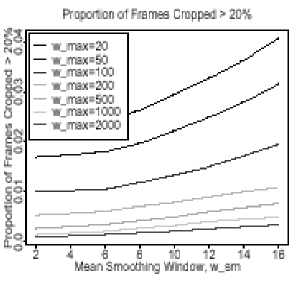


Figure 15: PROPORTION OF FRAMES CROPPED > 20% VS. w_{sm}, trace set A for varying w_{max}

Figure 16: PROPORTION OF FRAMES CROPPED > 20% VS. w_{max}, trace set A, averaged over w_{sm}.

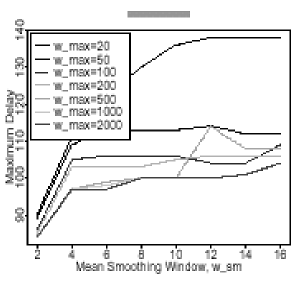


Figure 17: MAXIMUM DELAY VS. w_{sm}, trace set A for varying w_{max}.

Summary

- Source adaptation using rate-based feedback control
- Rate smoothing
- The SAVE algorithm
- Evaluation criteria