A Cross-Layer Perspective on Real-time Video Transmission

Robert W. Heath, Jr.
www.profheath.org

Video-Aware Wireless Networks (VAWN)
Final Meeting

January 23, 2014
Outline

- Real-time Video Transmission
  - Challenges and Opportunities

- Lessons Learned for Real-time Video
  - Mitigating Losses in Scalable Video
  - Mitigating Losses in Non-scalable Video
  - Mitigating Delays in Real-time Video
# Real-Time: Live & Interactive Video

<table>
<thead>
<tr>
<th>Buffering</th>
<th>Content encoded on the fly ➔ Minimal buffering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport</td>
<td>Tight delay constraints ➔ Unreliable (UDP) ➔ APP losses</td>
</tr>
</tbody>
</table>
| Impairments | **Source:** Compression distortions  
| | **Network:** Losses, excessive delay  
| | **Channel:** Transmission errors |
Video Codec Considerations

- Inter-frame (and Intra-frame) prediction:
  - Represent signal using a reference and a residual

- Rate-adaptive and/or scalable representation
  - Enable several rate-quality operating points

- Error concealment
  - Conceal losses using frame copy or motion copy

**Consequences:**

- Temporal/spatial structural dependencies
- Different error propagation patterns
- Unequal response to losses & distortions
Cross-layer Design for Real-time Video

- Enables exploiting video structure

**Challenges:**
- Multi-timescale adaptation
- Designing for video quality at lower layers
- Minimal additional cross-layer overhead
<table>
<thead>
<tr>
<th>Lessons learned on</th>
<th>Scope</th>
<th>Codec</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitigating Losses in Scalable Video</td>
<td>Single user</td>
<td>Scalable</td>
<td>Compression + Losses</td>
</tr>
<tr>
<td>Mitigating Losses in Non-scalable Video</td>
<td>Single user</td>
<td>Non-scalable</td>
<td>Compression + Losses</td>
</tr>
<tr>
<td>Mitigating Delay in Real-time Video</td>
<td>Multiuser</td>
<td>Any</td>
<td>Compression + Delay</td>
</tr>
</tbody>
</table>
Mitigating Losses in Scalable Video


Summary of Lessons Learned

1. Online learning is a key tool for adaptive real-time transmission
   - Regularity of natural scene statistics enables quick convergence
   - Enables adapting to video temporal dynamics, scene changes
   - Natural approach since content is not pre-encoded

2. UEP provides reasonable gains for UDP-based transmission
   - Lower reliability at APP layer enables higher gains
   - Generally, UEP is most beneficial at low SNR
   - Lower gains with sophisticated retx, e.g. HARQ

3. Adaptation of scalable video layer transmission
   - Avoid buffer starvation due to rapid adapt. in fast-varying channels
   - Quality adaptation preferable over frame rate adaptation
   - Frame rate adaptation performance highly dependent on EC method
1. Select video layers to transmit

**Minimize source distortions**
- **Constraint:** Playback buffer starvation
- **Timescale:** GoP level ➔ Use CDI

2. Unequally protect video layers

**Minimize channel distortions**
- **Constraint:** Per-layer target PER
- **Timescale:** Packet level ➔ Use CSI
**Objective:** Provide a perceptual quality guarantee

**Challenge:** Analytical modeling of loss effect on quality difficult

**Proposed Approach:** Learn quality-loss mapping online

1. Applies to real-time / live video streaming
2. Adapts to video temporal dynamics, scene changes

Apply local linear regression at target video quality using \( w \) observations (GoPs)

\[
\log_{10}(\rho_{k,t}[n]) = (b(\bar{q})^T(B^T W(\bar{q}) B)^{-1} B^T W(\bar{q})) \cdot \log_{10}(\rho_{k,t}[n - w : n - 1])
\]

- **Non-parametric:** Fit a linear model locally to minimize weighted least squares
- **Locality:** Captured through distance-based kernel weights
- Target video quality translates to lower bound on quality with combined losses
**Demo: Cross-layer Adaptation for Scalable Video**

**Baseline 1:** Equal error protection
- **Drawback:** Base layer losses, Low rate

**Baseline 2:** Link adaptation only
- **Drawback:** Frequent frame freezes

**Proposed:** APP + PHY adaptation

**Baseline 2:**
1) APP: SVC Rate Adaptation
2) PHY: Equal Error Protection

**Baseline 1:**
1) APP: SVC Rate Adaptation
2) PHY: Equal Error Protection
Mitigating Losses in Non-Scalable Video


1. **Loss visibility is useful / inexpensive for packet prioritization**
   - Side information of 3-4 bits sufficient to converge to full gain
   - Gains most pronounced in low mobility scenarios
   - Gains highly dependent on GoP structure

2. **MIMO channels provide a convenient way to achieve prioritization gain**
   - Unequal modulation enables both high quality/lower latency for important packets
   - Adapting MIMO mode based on packet importance variability can provide gains
   - With limited feedback, quality drops (obviously) but gain in quality improves

3. **Packet discarding vs. rate adaptation**
   - If rate adaptation not possible, “smart” packet discarding is useful
   - Rate adaptation enables good balance of source-channel distortions
   - Loss visibility is more useful for prioritization (at BS) than discarding (at router)
Relevance of Loss Visibility

• Loss visibility estimate [LinKanZhi10]
  \[ \log \left( \frac{v_p}{1 - v_p} \right) = \beta_0 + \sum_{i=1}^{F} \beta_i \cdot x_{pi} \]

• Set of features
  o Calculated at the slice level
  o Source-related
    • Motion (Magnitude, variance, camera)
    • Residual energy
    • Before / after scene cut, Etc…
  o Codec-related
    • Temporal error propagation
    • SSIM, maximum MSE per MB
    • Far conceal, other scene conceal, Etc…
How Packet Prioritization Works in MIMO

**Proposed:** LV-based Precoding: Symbols/packet mapped to one stream

**Baseline:** Conventional Precoding: Symbols/packet mapped to all spatial streams

LV = loss visibility = 💙
Objective: Maximize throughput weighted by per-packet loss visibility

Interpretation: Maximize cumulative utility of packets per unit time

Optimal packet-stream mapping:
- Decompose MIMO channel, order spatial streams and threshold loss visibility values

Per-stream Link Adaptation: Select MCS to maximize per-stream throughput

Stream Ordering: Order streams by corresponding packet error probability

Thresholding: Fraction of packets through stream proportional to throughput

\[
\text{Fraction of packets through stream } i = \frac{\text{Stream } i \text{ throughput}}{\text{Sum throughput}}
\]
Cross-layer View of LV-based Transmission

Example: Foreman video sequence, \( S = 3 \) packet classes

Estimated LV Distribution

Corresponding Thresholding Policy

Rate / PER per stream

- Loss visibility distribution
- Per-stream Modulation order
- Retransmission overhead

APP layer

PHY layer

MAC layer

Fraction of packets through stream \( i \) = \( \frac{\text{Rate}(1 - \text{Post-retx PER})}{\text{Sum throughput}} \)

High priority packets sent at higher rate AND with higher reliability
Example MIMO Packet Prioritization Result

LV-based Prioritization

No Prioritization

4x4 MIMO, 3 streams, 5 dB average SNR, up to 4 retransmissions per packet
Mitigating Delays in Real-time Video


1. Diverse QoS requirements have high impact on resource allocation
   - Different use cases require vastly different resources to achieve same quality
   - Max SINR scheduling is highly suboptimal for diverse QoS users
   - Statistical QoS-based user admission is key for fast fading channels

2. Insights on quality-maximizing resource allocation
   - Stringent QoS users with bad channels consume most resources and get lower quality
   - Don’t opt for full fairness in quality – too expensive, especially in macro cells
   - Variability in R-D characteristics across users is useful for quality optimization
Compute max source spectral efficiency

Compute BW allocation

Channel statistics

Target QoS

R-D

User 1

User K

Compute BW-QoS exponent product

Compute max source spectral efficiency

Compute BW allocation

Compute BW-QoS exponent product

Compute max source spectral efficiency

Compute BW allocation

\[ \ln \left( \mathbb{E}_y \left\{ (1 + \gamma_k)^{ \frac{d_k T}{\ln(2)}} \right\} \right) = \frac{\ln(P^\text{th})}{D^\text{th}} \]

\[ \frac{R^*_k(B_k)}{B_k} = \frac{\ln \left( 1/P^\text{th}_k \right)}{D^\text{th}_k T d_k} \]

\[ \frac{\partial Q_k(R^*_k(B^*_k))}{\partial R^*_k(B^*_k)} \times \frac{R^*_k(B^*_k)}{B^*_k} = \rho \]

\[ \theta^*_k = \frac{\ln \left( 1/P^\text{th}_k \right)}{T R^*_k D^\text{th}_k} \]

\[ B^*_k = \frac{d_k}{\theta^*_k} \]

\[ \text{monotonically decreasing in } B \]

\[ \overrightarrow{\text{monotonically decreasing in } B} \]

Guaranteed convergence to optimal resource allocation
Different QoS Requirements are Important

- **User 1:** Typical live streaming application \((D_{th} = 2 \text{ sec}, P_{th} = 0.1, R_{min} = 185 \text{ Kbps})\)
- **User 2:** Typical video conferencing session \((D_{th} = 0.3 \text{ sec}, P_{th} = 0.1, R_{min} = 185 \text{ Kbps})\)

Diverse QoS users require vastly different resources to achieve same quality

Under diverse QoS requirements, Max SINR scheduling highly suboptimal
Benefits of Quality Maximizing Resource Allocation

Simulation parameters: $D_{th} = 0.3$ sec w.p. $\frac{1}{2}$ or 2 sec w.p. $\frac{1}{2}$, $P_{th} = 0.1$, Rayleigh channels, $B = 20$ MHz

- **Scheduling gain:** Drop some BW-intensive users to max total users
- **Resource allocation gain:** Better partitioning based on QoS / channel

**Capacity & coverage improvement while meeting delay bound per user**
Directions of Future Work
Possible Extensions

**Frequency Selective Channels**
- Relaxed QoS user
- Tolerate low SIR
- Higher quality & QoS met

**Interference Management**
- Stringent QoS user

**Multicell Extensions / Spatial Models**
- Analyze Network-level QoS Metrics

**Non reference (NR) based Adaptation**
- NR Quality Assessment
- Client-driven Adaptation
Questions?

![Diagram showing different thrusts and optimizations]

<table>
<thead>
<tr>
<th>Thrusts 1 and 2: User-level Perceptual Optimization</th>
<th>Thrust 3: Network-level Perceptual Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-scalable Video</td>
<td>Scalable Video</td>
</tr>
<tr>
<td>Video Feature Extraction</td>
<td>Loss Visibility Estimation</td>
</tr>
<tr>
<td>Prioritized Retransmission</td>
<td>Packet loss visibility</td>
</tr>
<tr>
<td>MIMO Mode Selection</td>
<td>Unequal Error Protection</td>
</tr>
<tr>
<td>Prioritized Precoder</td>
<td>CSI</td>
</tr>
</tbody>
</table>

APP (Server), Transport, Network, MAC, PHY (BS), Client (MS)