Adaptive Playout Scheduling Using Time-scale Modification

Yi Liang, Nikolaus Färber
Bernd Girod, Balaji Prabhakar
Outline

- QoS concerns and tradeoffs
- *Jitter adaptation* as a playout scheduling scheme
- Packet scaling using improved time-scale modification technique
- Loss concealment in compatible with adaptive playout
- Performance comparison and audio demos
QoS Concerns at the Receiver

Over best-effort network ... 

Delay Jitter

- Obstructs proper reconstruction of voice packets at the receiver
- Impairs interactivity of conversations

Delay

Packet Loss

- Impairs speech quality
Playout Algorithm (1) - Fixed Deadline

- Use buffer to absorb delay variations and playout voice packets at fixed deadline – *jitter absorption*
- Voice packets received after deadline are discarded

![Diagram](image)
Buffer Delay vs. Late Loss

Fixed playout deadline and *jitter absorption*:
- The playout rate is constant
- The tradeoff is between buffering delay and late loss
Playout Algorithms (2) - Adaptive Playout

- Monitor delay variation and adapt playout time accordingly - *jitter adaptation*
- Slow down playout when delay increases to avoid loss; speed up playout when delay decreases to reduce delay
Adaptive Playout and Jitter Adaptation

- Scaling of voiced packets in highly dynamic way
- Playout schedule set according to past delays recorded
- Improved tradeoff between buffering delay and late loss
- Playout rate is not constant
Packet Scaling (1)

- Based on WSOLA [Verhelst 93]
- Improved to scale short individual voice packets
- “In-and-out black box operation”, no algorithmic delay, smooth transitions
- Preserves pitch
Packet Scaling (2)

- STD network delay = 20.9 ms
  Max. jitter = 112.0 ms
- STD total delay = 10.5 ms
- Packets scaled: 18.4%
- Scaling ratio: 50% - 200%

- DMOS: 4.5

**DMOS scaling**: degradation is
5 - inaudible
4 - audible but not annoying
3 - slightly annoying
2 - annoying
1 - very annoying
Loss Concealment

- Based on [Stenger 96]
- Using information from both sides, delay minimized to one packet time
- Integrates nicely into system when adaptive playout is used
- 20% random packet loss:
  Original: ♫
  Loss: ♫  Concealed: ♫
Comparison of Different Algorithms

1. Method which uses fixed playout time throughout the whole session;

2. Method which estimates delay dynamically but only adjusts playout time during silence periods [Ramjee 94, Moon 98];

3. Method which dynamically estimates and adjusts playout time, and scales packets within talkspurts using time-scale modification.
Performance Comparison

Traces measured between a host at Stanford and hosts in:

1) Chicago
2) Germany
3) MIT
4) China
Overall Performance

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Buffer Delay</th>
<th>Loss Rate</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alg. 2</td>
<td>55 ms</td>
<td>10%</td>
<td>2.6</td>
</tr>
<tr>
<td>Alg. 3</td>
<td></td>
<td>4%</td>
<td>3.7</td>
</tr>
<tr>
<td>Original</td>
<td></td>
<td></td>
<td>4.4</td>
</tr>
</tbody>
</table>

Quality Score
- Excellent: 5
- Good: 4
- Fair: 3
- Poor: 2
- Bad: 1
Conclusions

- Improved time-scale technique to work on individual packets with minimum delay
- WSOLA based loss concealment integrates nicely into system
- Adaptive playout and jitter adaptation significantly reduce buffering delay and late loss, which results in improved overall performance

- Small playout rate variation can be traded for lower delay and lower loss rate
- Playout scaling depends on audio scaling; scaling of individual packets is almost inaudible