Basic principles of Voice over IP

Dr. Peter Počta

{pocta@fel.uniza.sk}

Department of Telecommunications and Multimedia
Faculty of Electrical Engineering
University of Žilina, Slovakia
Outline

- VoIP Transmission Chain
- Speech Codecs
- Transport Protocols and Data Units
- Jitter Buffer principles
- Packet Loss Recovery methods
VoIP Transmission chain

Fig. 1: Schematic representation of the speech processing steps involved in a VoIP user interface and subsequent transmission (adopted by Raake)
Speech codecs

- To reduce the bandwidth costs of voice transmission

**The reduction of speech data:**
- By exploiting properties of speech production (e.g. estimating vocal tract parameters)
- To a lesser extent of auditory perception (e.g. spectral masking)

**Classification of coding algorithms:**
- Waveform codecs (quantization of actual waveform)
- Parametric codecs (based on speech production model)
- Hybrid codecs (as a combination of thereof)
**Speech codecs**

**Narrowband codecs**

- Narrowband: 300-3400 Hz (classical telephony)
- Operate on different frame sizes, from 10 to 40 ms

**Narrowband codecs:**

- G.711 (PCM, 64 kbit/s)
- G.723.1 (dual rate codec operating at 5.3 or 6.3 kbit/s)
- G.729 (CS-ACELP, 8kbit/s)
- Internet Low Bit-rate Codec (iLBC, 15.2 and 13.33 kbit/s)
- AMR-NB (eight rates from 4.75 to 12.2 kbit/s)
Speech codecs
Wideband codecs

- Wideband: 50-7000 Hz (classical telephony)
- Operate on different frame sizes, from 10 to 40 ms

Wideband codecs:
- G.722 (PCM, 64 kbit/s)
- G.722.1 (Modulated Lapped Transform, 32 kbit/s)
- AMR-WB (Bit rates from 6.6 to 23.85 kbit/s)
- G.729.1 (WB extension of G.729, 24 kbit/s)
VoIP Transport Protocols
Basic overview

- Packet-based networks (no permanent physical connection)

The buildup of link:
- Communication setup (agreement between the VoIP end-points (port and codecs negotiations, connection oriented - TCP, SIP, H.323 – signaling protocols)
- Connection established (based on RTP and UDP connectionless)
- Packetization of codec speech data
- Sending speech packet from source to destination station
## VoIP Transport Protocols

<table>
<thead>
<tr>
<th>Protocol/Technology</th>
<th>Main Header Information and Task</th>
<th>OSI Model Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>Timestamp: For example, interpacket delay can be calculated as the difference between the respective timestamps. Sequence number: Indicates whether packets are missing or arrived out of order. Payload type: Signalizes the type of payload, for example, speech coded according to G.729A (see text).</td>
<td>‘[…] Transport protocol implemented in the application layer’ (Tanenbaum, 2003)</td>
</tr>
<tr>
<td>UDP</td>
<td>Destination and source ports: Indication of the sending and receiving applications.</td>
<td>Transport layer</td>
</tr>
<tr>
<td>IP</td>
<td>Source and destination addresses: Packet routing.</td>
<td>Network layer</td>
</tr>
<tr>
<td>Ethernet, WLAN, and so on</td>
<td>Media access.</td>
<td>Sublayer of data link layer</td>
</tr>
</tbody>
</table>

Table 1: Protocols and media access technologies involved in VoIP packetization (adopted by Raake)
Fig. 2: VoIP: Speech payload nested in a packet, with headers added by different protocols (example for VoIP in an Ethernet-based LAN) (adopted by Raake)
VoIP data units

Table 2: Header sizes of different protocols involved in VoIP (adopted by Raake)

<table>
<thead>
<tr>
<th>Protocol</th>
<th>RTP</th>
<th>UDP</th>
<th>TCP</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet header (minimum, [Bytes])</td>
<td>12</td>
<td>8</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>
Real VoIP Transmission

- Different conditions in case of subsequent packets transmission (congestion or traffic load changing → resulting in delay variation (jitter))
- Limited values of delay in case of efficient speech communication (values in next part of presentation)
Jitter Buffer principles

- To compensate for jitter a typical VoIP application buffers incoming packets in jitter buffer before playing them.
- Allows slower packets to arrive on time in order to be playout.
- Threshold for jitter buffer size (not too long or too short) \(\rightarrow\) loss/delay trade-off
  - too short \(\rightarrow\) higher losses (voice quality suffers)
  - too long \(\rightarrow\) higher buffering delay (disrupts interactive speech communication)
Jitter Buffer principles

- Currently the management of the playout buffer is not specified by any standard and is vendor specific.
- Not possible to get info about implementation of jitter buffers in commercial applications.
- The playout buffer module has a strategic value from the vendor’s perspective and is usually kept confidential.

Types of jitter buffer schemes:
- Fixed
- Adaptive (mainly deployed)
Jitter Buffer principles

**Jitter buffer monitors:**
- The timestamps $t_i$ (RTP)
- The reception time $r_i$ of the $i$-th packet

**Jitter buffer adjusts:**
- The playout deadline $p_i$
- The playout delay or deadline is adjusted by compressing or expanding silent periods between consecutive talkspurts.

Fig.3: Adaptive playout buffer control mechanism at a receiver (adopted by Narbutt)
Jitter Buffer principles

Mechanisms for playout time adjusting:

- **Per talkspurt** (The playout time is calculated only for the first packet of the incoming talkspurt. Any subsequent packets of that talkspurt are played out with the rate equal to the generation rate at the sender. (not sufficient for long talkspurts and high delay variation))

- **Per packet** (Proper reconstruction of continuous output speech is achieved by scaling individual voice packets using the “time-scale modification technique”. (more sufficient for long talkspurts and high delay variation))
Jitter buffer principles

Four groups of playout buffer algorithms:

- **Reactive algorithms** that perform continuous estimation of network delays and jitter to calculate playout deadlines.

- **Histogram-based algorithms** that maintain a histogram of packet delays and choose the optimal playout delay from that histogram.

- **Algorithms that monitor packet loss ratio or buffer occupancy** and adjust the playout delay accordingly.

- **Algorithms that aim in maximizing user satisfaction.**
Packet Loss Recovery methods

- To conceal the losses from user perspective

**Methods:**
- Receiver-based methods
- Sender-based methods

**Mainly used approaches:**
- *Packet Loss Concealment (PLC) - Receiver-based method*
  - No additional data are sent from sender to receiver
  - Lost packets are compensated at receiver side
  - Approaches: Silence insertion, Noise substitution, Packet repetition
Packet Loss Recovery methods

- **Forward Error Correction (FEC) - Sender-based method**
  - Duplicate version of coded data is transmitted by sender
  - Used at the receiver to restore the lost info
  - High quality also in case of higher losses and burstiness

- **Low-Bitrate Redundancy (LBR) – Sender-based method**
  - Redundant version of speech signal is coded at lower bitrate and sent in addition to receiver
  - Used at the receiver side for the replacement of lost packets
  - Less effective than FEC and more than PLC
References

Thank you for your attention!

Questions?