Chapter 10
Multimedia Network Communications and Applications

Quality of Multimedia Data Transmission
Multimedia over IP
Transport of MPEG-4
Media-on-Demand (MOD)
Multimedia over Wireless Networks
Characteristics of Multimedia Data

• Multimedia network communication is same as computer network communication because both are deal with data communication. Multimedia data has the following characteristics:
  
  • **Voluminous** — they demand very high data rates, possibly dozens or hundreds of Mbps.
  
  • **Real-time and interactive** — they demand low delay and synchronization between audio and video for “lip sync”. In addition, applications such as video conferencing and interactive multimedia also require two-way traffic.
  
  • **Sometimes bursty** — data rates fluctuate drastically, e.g., no traffic most of the time but burst to high volume in video-on-demand.
Quality of Multimedia Data Transmission: Quality of Service (QoS)

- **Quality of Service (QoS)** depends on many parameters:
  - **Data rate**: a measure of transmission speed, often kilobits per second (kbps) or megabit per second (Mbps).
  - **Latency (maximum frame/packet delay)**: maximum time needed from transmission to reception, often measured in milliseconds (msec). EX. In voice communication, when the round-trip delay exceeds 50 msec, echo becomes a noticeable problem; when one-way delay is longer than 250 msec, talker overlap will occur since each caller will talk without knowing the other is also talking.
  - **Packet loss or error**: a measure (in percentage) of error rate of the packetized data transmission. Packets get lost or garbled or delivered late over the internet. When it approaches 10%, it becomes intolerable. Sometimes simple-error-recovery method used for real-time multimedia to replay the last packet.
Quality of Multimedia Data Transmission: Quality of Service (QoS)

- **Jitter (or delay jitter):** a measure of smoothness of the audio/video playback, related to the variance of frame/packet delays. A large buffer (jitter buffer) can be hold enough frames to allow the frame with longest delay to arrive, to reduce playback jitter. This increases the latency and may not be desirable in real-time and interactive applications.

- **Sync skew:** a measure of multimedia data synchronization, often measured in milliseconds (msec). For a good lip synchronization, the limit of sync skew is +/- 80 msec between audio and video. For video with speaker and voice the limit of sync skew is +/- 120 msec if video precedes voice and 20 msec if voice precedes video.
Quality of Multimedia Data Transmission:
Multimedia Service Classes

- Based on previous measures, multimedia services classified as following:
  - **Real-Time** (also *Conversational*): two-way traffic, low latency and jitter, possibly with prioritized delivery, e.g., voice telephony and video telephony.
  - **Priority Data**: two-way traffic, low loss and low latency, with prioritized delivery, e.g., E-commerce applications.
  - **Silver**: moderate latency and jitter, strict ordering and sync. One-way traffic, e.g., streaming video, or two-way traffic (also *Interactive*), e.g., web surfing, Internet games.
  - **Best Effort** (also *Background*): no real-time requirement, e.g., downloading or transferring large files (movies).
  - **Bronze**: no guarantees for transmission.
### Table 16.1: Requirement on Network Bandwidth / Bit-rate

<table>
<thead>
<tr>
<th>Application</th>
<th>Speed Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telephone</td>
<td>16 kbps</td>
</tr>
<tr>
<td>Audio-conferencing</td>
<td>32 kbps</td>
</tr>
<tr>
<td>CD-quality audio</td>
<td>128–192 kbps</td>
</tr>
<tr>
<td>Digital music (QoS)</td>
<td>64–640 kbps</td>
</tr>
<tr>
<td>H. 261</td>
<td>64 kbps–2 Mbps</td>
</tr>
<tr>
<td>H. 263</td>
<td>&lt; 64 kbps</td>
</tr>
<tr>
<td>DVI video</td>
<td>1.2–1.5 Mbps</td>
</tr>
<tr>
<td>MPEG-1 video</td>
<td>1.2–1.5 Mbps</td>
</tr>
<tr>
<td>MPEG-2 video</td>
<td>4–60 Mbps</td>
</tr>
<tr>
<td>HDTV (compressed)</td>
<td>&gt; 20 Mbps</td>
</tr>
<tr>
<td>HDTV (uncompressed)</td>
<td>&gt; 1 Gbps</td>
</tr>
<tr>
<td>MPEG-4 video-on-demand (QoS)</td>
<td>250–750 kbps</td>
</tr>
<tr>
<td>Videoconferencing (QoS)</td>
<td>384 kbps–2 Mbps</td>
</tr>
</tbody>
</table>
Table 16.2: Tolerance of Latency and Jitter in Digital Audio and Video

<table>
<thead>
<tr>
<th>Application</th>
<th>Avg Latency Tolerance (msec)</th>
<th>Avg Jitter Tolerance (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-end videoconf. (64 kbps)</td>
<td>300</td>
<td>130</td>
</tr>
<tr>
<td>Compressed voice (16 kbps)</td>
<td>30</td>
<td>130</td>
</tr>
<tr>
<td>MPEG NTSC video (1.5 Mbps)</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>MPEG audio (256 kbps)</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>HDTV video (20 Mbps)</td>
<td>0.8</td>
<td>1</td>
</tr>
</tbody>
</table>
Quality of Multimedia Data Transmission: Perceived QoS

- Although QoS is commonly measured by the above technical parameters, QoS itself is a “collective effect of service performances that determine the degree of satisfaction of the user of that service”.

- In other words, it has everything to do with how the user perceives it. For example, in real-time multimedia:
  - Regularity is more important than latency (i.e., jitter and quality fluctuation are more annoying than slightly longer waiting).
  - Temporal correctness is more important than the sound and picture quality (i.e., ordering and synchronization of audio and video are of primary importance).
  - Humans tend to focus on one subject at a time. User focus is usually at the center of the screen, and it takes time to refocus especially after a scene change.
Prioritized Delivery

- Used to alleviate the perceived deterioration (high packet loss or error rate) in network congestion.
  - **Prioritization for types of media**: Transmission algorithms can provide prioritized delivery to different media.
  - **Prioritization for uncompressed audio**: PCM audio bitstreams can be broken into groups of every nth sample.
  - **Prioritization for JPEG image**: The different scans in Progressive JPEG and different resolutions of the image in Hierarchical JPEG can be given different priorities.
  - **Prioritization for compressed video**: Set priorities to minimize playback delay and jitter by giving highest priority to I-frames for their reception, and lowest priority to B-frames.
Multimedia over IP

- A *broadcast* message is sent to all nodes in the domain, a *unicast* message is sent to only one node, and a *multicast* message is sent to a set of specified nodes.

- **IP-Multicast:**
  - Anonymous membership: the source host multicasts to one of the IP-multicast addresses — doesn’t know who will receive.
  - Potential problem: too many packets will be traveling and alive in the network — use time-to-live (TTL) in each IP packet.
  - Vital for applications such as mailing lists, group file transfer, audio/video conferencing.
Multimedia over IP (cont.)

- **RTP (Real-time Transport Protocol)**
  - Designed for the transport of real-time data such as audio and video streams:
    - Primarily intended for multicast.
    - Used in *nv* (*network video*) for MBone, Netscape LiveMedia, Microsoft Netmeeting, and Intel Videophone.
  - Usually runs on top of UDP which provides efficient (but less reliable) connectionless datagram service:
    - RTP must create its own *timestamping* and *sequencing* mechanisms to ensure the ordering.
Multimedia over IP (cont.)

- **RTCP (Real Time Control Protocol)**
  - A companion protocol of RTP:
    - Monitors QoS in providing feedback to the server (sender) and conveys information about the participants of a multi-party conference.
    - Provides the necessary information for audio and video synchronization.
  - RTP and RTCP packets are sent to the same IP address (multicast or unicast) but on different ports.
Multimedia over IP (cont.)

- **RSVP (Resource ReSerVation Protocol)**
  - Developed to guarantee desirable QoS, mostly for multicast although also applicable to unicast.
    - A general communication model supported by RSVP consists of $m$ senders and $n$ receivers, possibly in various multicast groups (e.g. Fig. 16.4(a)).
  - The most important messages of RSVP:
    - A **Path** message is initiated by the sender, and contains information about the sender and the path (e.g., the previous RSVP hop).
    - A **Resv** message is sent by a receiver that wishes to make a reservation.
  - Main Challenges:
    - There can be a large number of senders and receivers competing for the limited network bandwidth.
    - The receivers can be heterogeneous in demanding different contents with different QoS.
    - They can be dynamic by joining or quitting multicast groups at any time.
• Fig. 16.4: A scenario of network resource reservation with RSVP.
About the Example in Fig. 16.4

- Fig. 16.4 depicts a simple network with 2 senders (S1, S2), three receivers (R1, R2, and R3) and 4 routers (A, B, C, D):
  - In (a), **Path** messages are sent by both S1 and S2 along their paths to R1, R2, and R3.
  - In (b) and (c), R1 and R2 send out **Resv** messages to S1 and S2 respectively to make reservations for S1 and S2 resources. Note that from C to A, two separate channels must be reserved since R1 and R2 requested different data streams.
  - In (d), R2 and R3 send out their **Resv** messages to S1 to make additional requests. R3’s request was merged with R1’s previous request at A and R2’s was merged with R1’s at C.
Multimedia over IP (cont.)

- **RTSP (Real Time Streaming Protocol): Streaming Audio and Video:**
  - Audio and video data that are transmitted from a stored media server to the client in a data stream that is almost instantly decoded.
  - **RTSP Protocol:** for communication between a client and a stored media server (Fig. 16.5).
    - **Requesting presentation description:** the client issues a DESCRIBE request to the Stored Media Server to obtain the presentation description — media types, frame rate, resolution, codec, etc.
    - **Session setup:** the client issues a SETUP to inform the server of the destination IP address, port number, protocols, TTL (for multicast).
    - **Requesting and receiving media:** after receiving a PLAY, the server started to transmit streaming audio/video data using RTP.
    - **Session closure:** TEARDOWN closes the session.
• Fig. 16.5: A possible scenario of RTSP operations.
Multimedia over IP (cont.)

- **Internet Telephony**: Main advantages of Internet telephony over *POTS (Plain Old Telephone Service)*:
  - Uses packet-switching — network usage is much more efficient (voice communication is bursty and VBR encoded).
  - With the technologies of multicast or multipoint communication, multi-party calls are not much more difficult than two-party calls.
  - With advanced multimedia data compression techniques, various degrees of QoS can be supported and dynamically adjusted according to the network traffic.
  - Good graphics user interfaces can be developed to show available features and services, monitor call status and progress, etc.
Transport of MPEG-4

- **Delivery Multimedia Integration Framework (DMIF) in MPEG-4:** An interface between multimedia applications and their transport. It supports:
  - Remote interactive network access (IP, ATM, PSTN, ISDN, mobile).
  - Broadcast media (cable or satellite).
  - Local media on disks.
- A single application can run on different transport layers as long as the right DMIF is instantiated.
- Fig. 16.9 shows the integration of delivery through three types of communication mediums.
- **MPEG-4 over IP:** MPEG-4 sessions can be carried over IP-based protocols such as RTP, RTSP, and HTTP.
• Fig. 16.9: DMIF — the multimedia content delivery integration framework.
Media-on-Demand (MOD)

- Interactive TV (ITV) and Set-top Box (STB): ITV supports activities such as:
  - TV (basic, subscription, pay-per-view).
  - Video-on-demand (VOD).
  - Information services (news, weather, magazines, sports events, etc.).
  - Interactive entertainment (Internet games, etc.).
  - E-commerce (on-line shopping, stock trading).
  - Access to digital libraries and educational materials.

- The fundamental difference between ITV and conventional TV is that ITV invites user interactions; hence need two way traffic (downstream and upstream). Also, ITV is rich in information and multimedia content.
Multimedia over Wireless Networks

• Mainly concerned with sending video robustly over wireless channels, e.g., for a video conferencing.
• Since wireless data transmissions incur the most data loss and distortion, error resilience and error correction become primary concerns.
• Characteristics of wireless handheld devices worth keeping in mind when designing multimedia transmission:
  • Both the handheld size and battery life limit the processing power and memory of the device – low complexity in encoding and decoding.
  • Due to memory constraints and other reasons, real-time communication will likely be required.
  • Wireless channels have much more interference than wired channels
  • error resilient coding is very important.
3GPP QoS Requirements for Multimedia Transmission

- **Synchronization** — video and audio should be synchronized to within 20 msec.
- **Throughput** — the minimum video bit-rate to be supported is 32 kbps. Video rates of 128 kbps, 384 kbps and above should be supported as well.
- **Delay** — the maximum end-to-end transmission delay is defined to be 400 msec.
- **Jitter** — the maximum delay jitter (maximum difference between the average delay and the 95th percentile of the delay distribution) is 200 msec.
- **Error Rate** — the video conferencing system should be able to tolerate a frame error rate of $10^{-2}$ or bit error rate of $10^{-3}$ for circuit switched transmission.
Synchronization Loss

- **Loss of decoder synchronization**: For digital video coding, when there is damage to a packet containing variable bit-length data, that error, if unconstrained, will propagate all the way throughout the stream.
- Other than synchronization loss, errors in prediction reference frames cause much more damage to the signal quality than errors in frames not used for prediction.
- Similarly, if the video is scalable, an error at the base layer will deteriorate the quality of a video stream more than in enhancement layers.
Synchronization Loss (Cont’d)

• MPEG-4 defines additional error-resilient tools that are useful for coding under noisy and wireless channel conditions:
  • A data partitioning scheme will group and separate header information, motion vectors, and DCT coefficients into different packets, and put synchronization markers between them.
  • An adaptive Intra frame refresh mode is allowed where each MB can be coded independently of the frame as an Inter or Intra block according to its motion, to assist with error concealment.
Synchronization Loss (Cont’d)

- **Sender-Receiver Feedback** techniques can be used if a back channel is available to the encoder:
  - According to the bandwidth available at any moment, the receiver can ask the sender to lower or increase the video bit-rate (transmission rate control), which combats packet loss due to congestion.
  - If the stream is scalable, it can also ask for enhancement layers.
  - Receiver can notice damage in a reference frame, and request that the encoder use a different reference frame for prediction – a reference frame that the decoder has reconstructed correctly.