

# 实时流媒体传输技术研究综述<sup>\*</sup>

辣鸡!

熊永华<sup>1</sup>, 吴 敏<sup>1</sup>, 贾维嘉<sup>2</sup>

(1. 中南大学 信息科学与工程学院, 长沙 410083; 2. 香港城市大学 计算机科学系, 香港 999077)

**摘 要:** 为推动实时流媒体传输技术的进步, 提高 Internet 流媒体服务质量, 将实时流媒体传输流程分为用户层、编码层、流处理层、传输控制层和网络层, 分层总结和综述了实时流媒体传输过程中的主要研究问题和研究进展, 阐述了一些目前主流的基于 UDP 和 TCP 的实时流媒体传输技术, 并且比较分析了它们的优缺点。最后指出低比特率、高丢包率与无线、强干扰环境下保证流媒体服务质量是未来的研究方向, 移动实时流媒体业务则是实时流媒体传输技术应用的必然趋势。

**关键词:** 实时流媒体; 传输; 用户数据报协议; 传输控制协议

**中图分类号:** TP393

**文献标志码:** A

**文章编号:** 1001-3695(2009)10-3615-06

doi:10.3969/j.issn.1001-3695.2009.10.004

## Survey on real-time multimedia streaming transmission technology

XIONG Yong-hua<sup>1</sup>, WU Min<sup>1</sup>, JIA Wei-jia<sup>2</sup>

(1. School of Information Science & Engineering, Central South University, Changsha 410083, China; 2. Dept. of Computer Science, City University of Hong Kong, Hong Kong 999077, China)

**Abstract:** In order to promote the development of real-time multimedia streaming transmission technology and improve quality of service of the Internet streaming, this paper divided the process of transmitting real-time multimedia streaming into several layers, named as user layer, code layer, streaming construct layer, transport control layer, and network layer. Summarized and surveyed the key issues and major progresses with respect to all layers. Reviewed the principal technologies of real-time multimedia streaming based on UDP and TCP with some comparisons between them. How to ensure the quality of service of multimedia streaming under the circumstance of wireless network with low bite rate, high packet loss ratio, and strong disturbance, was viewed as research direction of real-time multimedia streaming transmission technology, so that it's undoubted that the mobile real-time multimedia streaming service should be the application trend of the technology.

**Key words:** real-time multimedia streaming; transmission; UDP; TCP

## 0 引言

实时流媒体是随着 Internet 与多媒体技术的发展而逐渐兴起的一种多媒体网络传输技术, 其典型传输方式为: 连续采样的多媒体信息经编码压缩后, 从源端连续、实时地发送到网络, 接收端从网络收到部分数据后就可以开始解码播放, 而后续接收的数据则持续不断地存入本地缓存, 形成可连续播放的媒体流。与传统 FTP、e-mail 等业务相比, 实时流媒体对于实时性、网络带宽、容错性、媒体同步、播放平滑性等方面均有严格要求。例如, 当语音延迟超过 250 ms, 视频延迟超过 1 s 时, 就能明显感觉到停顿<sup>[1]</sup>; 当网络可用带宽小于 28 kbps 时, 采用高效编码压缩技术后的流媒体仍易导致严重的网络拥塞; 在通过 Internet 传输时, 要求能在一定程度上允许丢包的发生, 并保持视频、音频等信息同步; 当网络状况出现波动时, 要求能通过设置播放缓冲机制等消除抖动, 保证播放的平稳性等。

当前的 Internet 是以 best-effort 方式工作的异构网络, 其终

端接入方式、可用带宽、延迟抖动和丢包等因素都是动态变化和不可预知的, 难以满足实时流媒体传输的要求。为使实时流媒体能够适应 Internet 的特点, 提高流媒体服务质量, 经过研究人员十多年的努力, 在实时流媒体传输技术方面已形成了一些日趋实用的研究成果与标准, 如自适应传输<sup>[2,3]</sup>、自适应编码<sup>[4,5]</sup>、流媒体重传机制<sup>[6,7]</sup>、应用层组播<sup>[8,9]</sup>、RTP<sup>[10]</sup>、RTCP<sup>[11]</sup>、TRFC<sup>[12]</sup>等, 并在 Skype、PPLive、eyeBeam 等国内外一些相关软硬件产品中得到应用。

尽管如此, 实时流媒体的服务质量仍有相当大的改进空间, 而在异构、低比特率、高丢包率、强干扰等网络环境中, 如 21 世纪新兴的 3G、WiMAX、Ad hoc、工业无线网络等环境下, 实时流媒体传输技术更需要进一步的研究与完善。只有深入分析实时流媒体传输的整个过程, 发现传输过程所需要解决的主要问题, 总结现有的相关研究成果, 才能更好地推动实时流媒体传输技术的进步。为此, 本文结合典型实时流媒体传输的流程, 对基于 UDP 和 TCP 两个主流传输层协议的实时流媒体传输技术进行了综述。

**收稿日期:** 2009-02-19; **修回日期:** 2009-04-11 **基金项目:** 国家杰出青年科学基金资助项目(60425310); 香港城市大学战略发展计划资助项目(7002102, 7002214)

**作者简介:** 熊永华(1979-), 男, 湖北洪湖人, 讲师, 博士研究生, 主要研究方向为多媒体通信、无线网络(yhxiang@mail.csu.edu.cn); 吴敏(1963-), 男, 广东化州人, 教授, 博士, 主要研究方向为过程控制、鲁棒控制和智能系统; 贾维嘉(1957-), 男, 贵州贵阳人, 教授, 博士, 主要研究方向为多媒体通信、移动计算和分布式系统。

1 流媒体传输流程及主要研究问题

一个典型的基于 Internet 的端到端文字、语音、视频实时流媒体会话流程如图 1 所示。

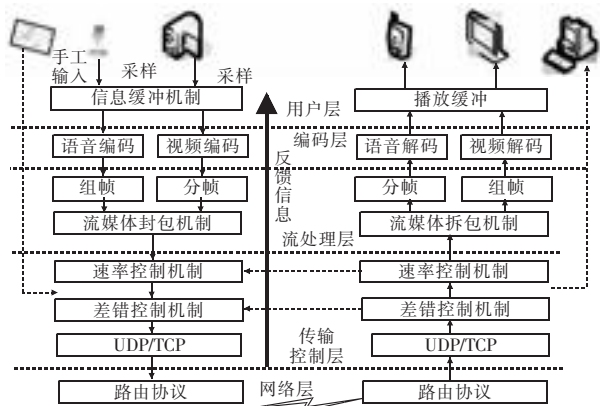


图 1 实时流媒体传输流程示意图

本文将这一流程分为用户层、编码层、流处理层、传输控制层和网络层。文字等人工输入的信息频率低、信息量小,可直接将其封装成 UDP 或 TCP 的载荷后传输,仅需基本的差错控制;语音、视频等需设备连续采样的数据,频率高、信息量大,还需经过编码压缩、分帧、组帧、流处理、缓冲播放等操作,而且采样数据实时性、连续性、有序性的特点也对传输的流量控制与差错控制机制等提出了更高要求。本章将描述各层主要功能,分析传输过程的主要研究问题和主流传输技术,总结相关研究进展。

1.1 用户层

用户层主要实现数据的采样输入、媒体信息成帧和缓冲播放等功能。

1) 采样输入

语音与视频信息的采样频率由麦克风、摄像头等硬件设备的性能决定。一般来说,一旦语音与视频设备开始启动,采样频率就固定了,如果需要变更采样频率,则需要停止会话并重新设置设备的属性。因此通常认为设备的固有采样频率是恒定的。

2) 媒体信息成帧

采样的多媒体数据通过用户层信息缓冲机制进行成帧处理,以便于编码和传输。

对于语音信息,每次生成的语音帧应包含多个连续的采样点,如对于 8 000 次/s 的语音采样设备,若设置语音信息缓冲机制的周期为 20 ms(称为语音帧长),即形成 50 个/s 语音帧,则每个语音帧应包含 8 000/50 = 160 个连续采样点。视频的缓冲机制有所不同,如对于采样频率 35 fps 的摄像头,若设置缓冲机制成帧频率为 10 fps(称为视频帧率),则缓冲机制每秒从 35 帧图像中抽取 10 帧送入视频编码器,同时丢弃其他视频帧。

成帧后的多媒体数据(称为语音帧或视频帧)送入编码层进行编码压缩处理,并通过网络传输到接收方,经过相应的解码后,还不能立即播放,而是需要首先送入接收方播放缓冲区(playout buffer)。

3) 播放缓冲

接收方播放缓冲是一种延迟缓冲机制,作用在于消除 IP 包的延迟抖动,如发送方每隔 100 ms 发送一个视频 IP 包(假设一个视频帧只被封装在一个 IP 包中),但这些 IP 包的到达间隔时间往往不等于发送间隔 100 ms,因此接收方在 IP 包到达以后不是立即播放,而是将其放入播放缓冲区中等待时间段  $D_{\text{playout}}$  以后再进行播放。

如果  $D_{\text{playout}}$  设置过小,当包到达较迟时,播放缓冲区将溢出,媒体包将错过其播放时间,等同于包丢失。另一方面,增加  $D_{\text{playout}}$  的大小将降低缓冲区溢出的可能性,但同时又会增加端到端的延迟。

图 2 描述了  $D_{\text{playout}}$ 、播放速率、丢包等之间的关系。可以看出,包到达间隔上下波动, $D_{\text{playout}}$  的设置应维持缓冲区溢出与端到端延迟的平衡性。目前对  $D_{\text{playout}}$  的设置静态法和动态法两种。静态法通常设置缓冲区为两个帧的大小,即  $D_{\text{playout}}$  为两倍采样间隔<sup>[13]</sup>;动态法则需要根据缓冲区的填充水平、网络延迟状况等进行动态调整<sup>[14~16]</sup>。

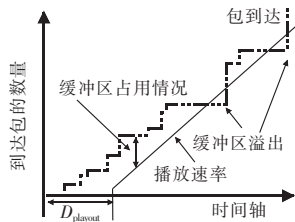


图 2  $D_{\text{playout}}$  与丢包及端到端延迟的关系

1.2 编码层

编码层主要对语音帧和视频帧进行编码压缩与解码处理,为流处理层提供符合要求的语音和视频编码帧。

1) 语音编/解码

语音编码器均会带来一定的处理延时,但是经过高度压缩后的语音又具有更高的传输效率。广泛使用的语音编码压缩算法有 G. 723.1、G. 729、AMR (adaptive multi-rate) 等。G. 723.1 压缩算法延迟较大但压缩率较高,适合低比特率可视电话;G. 729 延迟小、压缩率适中,适合 Internet 语音传输;AMR 延迟适中但压缩率可根据网络可用带宽进行自适应调整,生成多个不同档次的码率,一般用于 GSM、3G 网络等移动通信领域。

当发送方的语音压缩算法与接收方的解码算法不一致时,需要设置编码转换机制使得双方的编/解码标准完全一致,才能保证语音的正确播放。因此设计具有强通用性的语音网关,确保发送方与接收方能够相互转码,是一个实用性很强的研究课题<sup>[17,18]</sup>。

2) 视频编/解码

常用的视频编码压缩标准有国际电信联盟 ITU-T 所制定的 H. 26x 系列(H. 261、H. 263 和 H. 264)标准以及国际标准化组织 ISO/IEC 所制定的 MPEG-x 系列(MPEG-1、MPEG-2 和 MPEG-4)标准。

H. 26x 主要应用于实时视频通信。其中 H. 263 是最早用于 64 kbps 以下低码率视频编码的 ITU-T 标准,也是目前使用最为广泛的一种编码方式,WCDMA 3G 视频通信就是采用的 H. 263 标准;H. 264 是 ITU 与 IEC 共同制定的最新视频编码标

准,引入了帧内和帧间预测新技术,具有较高的压缩比,但同时也大大增加了算法复杂度,对编码设备要求较高,目前尚未广泛使用。

MPEG-x 标准主要应用于视频存储、DVD、广播电视等。其中 MPEG-4 是 2003 年发布的最新视频压缩标准,扩展了旧的 MPEG-1、MPEG-2 等标准以支持低比特率编码、交互式存储媒体和数字版权管理等。

1.3 流处理层

1.3.1 组帧与分帧

一个编码后的语音帧往往只有几个到几十个字节,如 AMR 编码语音帧约为 31 Byte,而一个编码后的视频帧则依据场景的复杂程度,从几千字节到数百千字节不等,如 H.263 编码视频帧通常为数十千字节。因此,若 MTU (maximum transmission unit) 为 1 500 Byte 的 IP 包仅携带一个几十字节语音帧,则会大大降低 IP 包的有效载荷;相反,一个视频帧的大小往往又大于 MTU。因此有必要将多个语音帧组合成一个大的语音帧(组帧),而将一个视频帧分成多个小的视频帧(分帧),如图 3 所示。

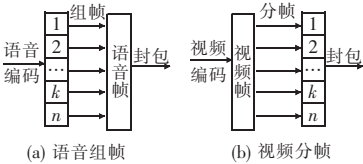


图 3 语音组帧和视频分帧过程

在图 3(a) 中,设音频编码频率为  $n$  fps,每帧大小为  $i$  Byte,包头大小为  $p$  Byte,若每  $k(k \leq n)$  个语音帧组成一个大的语音帧,则有  $k \times i + p < \text{MTU}$ ,且有效载荷为  $k \times i / \text{MTU}$ ,发送周期为  $k / n$  s,因此第  $k$  个语音帧所具有的组帧延迟为  $(k - 1) / ns$ 。显然缓存的语音帧越多,即  $k$  值越大,语音延迟越大;反之,若  $k$  值减少,则语音延迟缩小,但有效载荷降低,不利于提高路由效率,易引起语音包被路由丢弃。因此语音组帧过程中  $k$  值的设定,应该综合考虑应用程序对于语音质量、丢包率等的要求。

在图 3(b) 中,一个大小为  $i$  Byte 的视频帧可被分成约  $n$  个大小为  $\text{MTU} - p$  的子帧,由于各子帧经过网络传输后并不能同时到达接收方,而接收方只能等到所有子帧都到达后才能将各子帧按分帧顺序进行重新组帧,若有子帧在网络中丢失,将造成接收方无法组帧或者组帧错误。由此可见,一个视频帧所具有的传输延迟应该等于各个视频子帧的最大传输延迟,视频被分成的子帧越多,则它的传输延迟就有可能越大,而子帧丢失的概率也会增加,因此发送方的视频分帧应该依照 MTU 的大小进行,尽可能保证视频包的满载荷。

1.3.2 流媒体封包机制

语音与视频在发送方进行组帧与分帧后,需要使用封包机制将帧封装成具有特殊包头的 IP 包,接收方则进行相应的拆包操作。流媒体封包与拆包机制实际上是双方在传输多媒体时所达成的一个关于媒体类型、大小、包序号、有效时间、时间戳、同步类型等方面的传输协议。

为方便不同厂家的客户端之间也能相互通信,IETF RFC 1889 定义了一种媒体的封包与组包机制,即 RTP (real-time transport protocol)。RTP 的包头提供了时间标签、序列号以及其他具有实时性特征的结构用于控制适时数据的流放,因此适

用于交互式视频/音频数据的实时传输。一个使用 RTP 的 IP 包格式如图 4 所示。

IP 包头	20 或 24 Byte					
UDP/TCP 包头	UDP 包头=8 Byte,TCP 包头≥20 Byte					
RTP 包头	V	P	X	CC	M	PC sequence number
时间戳 time stamp						
同步源标志 SSIS(synchronization source identifiers)						
贡献源标志 CSRS(contributing source identifiers)						
编码器包头	payload					

图 4 使用 RTP 的 IP 包的构成

编码器包头与语音和视频数据被封装在 RTP 信息包中,作为 RTP 的载荷(payload),每个 RTP 信息包又被封装在 UDP 或 TCP 的消息段中,作为 UDP 或 TCP 的 payload,整个 UDP/TCP 的包头与其载荷则一起作为 IP 数据包的 payload。

RTP 包头中,V、P、X、CC、M、PT 与 sequence number 分别表示协议的版本号、填充标志、扩展标志、贡献源标志计数、重要标记、载荷类型与系列号。由于同一时刻的采样数据有可能通过多个 RTP 包进行传输,只有确切知道了包的序列号、数据戳以及同步源标志,接收方才能正确地拆包并构建出完整的媒体帧。

应用程序也可不遵循标准 RTP 而采用自定义协议,但是使用标准 RTP 的应用程序更容易与其他厂家的网络应用程序配合运行,如两个不同公司所开发的 IP 语音电话软件,若都遵循相同标准的 RTP,则使用这两个不同电话软件的用户之间有可能相互通信。

1.4 传输控制层

1.4.1 速率控制机制

速率控制机制通过检测网络与丢包状况,再结合接收端的播放质量、缓冲区大小等反馈信息,对流处理层的分帧机制、封包组包机制、编码层的编码参数,以及用户层的采样速率、信息缓冲机制等作出相应调整,使得发送端的发送速率能够有效匹配网络的可用带宽。当前的速率控制机制可以分为两类:

a)探测法。在传输多媒体数据时周期性地采用加法增加、乘法减少发送速率的方法进行网络可用带宽检测与发送速率的调整。设  $p$  为周期性检测得到的实时丢包率, $P_{th}$  为满足最低媒体播放要求时的丢包率, $P_{th}$  的大小由编/解码器的容错性能与接收端对媒体的播放质量所决定, $r$  为发送方的发送速率,AIR 为加法增加因子, $a$  为乘法减少因子,maxR 与 minR 分别为发送端能达到的最大与最小速率,则发送速率的基本调整算法为<sup>[5]</sup>

if( $p \leq P_{th}$ )

then  $r = \min \{ (r + \text{AIR}), \text{maxR} \}$

else  $r = \max \{ (a * r), \text{minR} \}$

其中:丢包率  $p$  由接收方负责计算并反馈给发送方,发送方根据所探测到的丢包率来估计网络可用带宽,从而周期性地对发送端的发送速率进行调整。

b)模型法。通过数学模型直接估算当前的可用网络带宽,再周期性地调整发送速率,较著名的带宽估算模型是 Floyd 等人<sup>[19]</sup>提出的 TCP 友好速率模型。

$$T = \frac{\text{MTU}}{\text{RTT}} \cdot \frac{C}{\sqrt{p}}$$

其中: $T$  是 TCP 连接的吞吐量;MTU 是链路最大传输单元的大小;RTT(round trip time)是网络往返时间; $p$  是丢包率。MTU



和  $C$  根据接入网的不同取常量;  $RTT$  和  $p$  由接收端计算并将结果周期性地反馈给发送端。发送端以  $T$  为限制条件, 周期性地调整发送速率, 确保其小于  $T$ 。目前有许多对该模型的推论与扩展<sup>[20~22]</sup>。

发送端利用探测法或模型法对发送速率进行调整的方式主要有:

a) 流处理层调整。通过流处理层进行分帧、组帧及封包机制调整, 如 IETF 在 RFC 3550 定义了 RTCP。RTCP 与 RTP 配合使用, 发送端和接收端周期性地分别发送发送者报告和接收者报告, 传送已发送的数据包的数量、丢失的数据包数量等统计资料, 从而对所传输的数据流服务质量进行监控, 动态改变传输速率。RTCP 信息可在一定程度上反映当前的网络状况, 但由于 RTCP 报告的发送周期通常长达 5 s, 通过 RTP 与 RTCP 速率调整难以满足很多实时性较强的应用, 研究人员提出了一些基于 RTP/RTCP 的改进方案<sup>[23,24]</sup>。

b) 速率自适应编码调整。通过调整编码器的编码速率使之适应可用网络带宽。传统编码机制通过改变编码器量化参数  $QP$  (quantization parameter) 调整编码速率, 帧率的微小调整也能导致播放质量的显著变化, 因此通常要求用户层具有恒定的帧率, 从而导致编码速率难以达到低比特率的水平。基于此, 适用于用户层可变帧率的低比特率自适应编码机制受到更多的关注<sup>[4,5]</sup>。

c) 用户层帧率调整。在设备固定采样频率不变时, 通过调整缓冲机制的成帧频率进行帧率调整, 使得用户层的信息率经过编码压缩后能基本匹配目标发送速率。用户层可以通过预测或计算媒体帧被丢弃的概率, 有选择性地丢弃一些不重要的, 或者在网络传输时丢失概率较大的媒体帧, 这也是一种应用较多的用户层调整策略<sup>[2,3]</sup>。

#### 1.4.2 差错控制机制

丢包是 Internet 中不可避免的现象, 文本、数字信息可以容忍较长的延迟但是不能允许丢包, 而诸如语音、视频等多媒体信息, 对于传输延迟的要求较高, 但可以通过差错隐藏、前向纠错等方式来容忍少量丢包。

##### 1) 差错隐藏

尽管每一个流媒体数据包的丢失都会对播放产生影响, 但具有好的纠错与容错机制的编/解码器可在一定程度上隐藏丢失的信息, 减少丢包对播放质量的影响。空间重构和时间重构是两种使用较多的错误隐藏的方法<sup>[25]</sup>, 空间重构在一个视频帧的某个数据包丢失时, 采用空间相近的像素信息来重构丢失的数据包, 适用于帧内信息重构, 而时间重构则在整帧丢失时, 采用上一帧的信息来重构丢失的帧, 适用于帧间信息重构。

当丢包较多以至于编/解码器无法容错, 将造成播放错误或者播放时出现严重抖动、语音无法分辨、马赛克等现象, 通常采用两种方式进行处理: 对数据包进行前向纠错编码和重传丢失的数据包。

##### 2) 前向纠错 (forward error correction, FEC)

FEC 是在媒体数据包中加入一些冗余信息, 从而在数据包丢失时能够通过冗余信息重构丢失的数据包, 常用的视频前向纠错编码机制有通道纠错、视频信号源纠错和联合通道与视频源纠错三种。

通道前向纠错编码在每个数据包中加入额外的块编码区

信息, 因此当有部分数据包丢失时, 可以根据这些额外的块编码进行丢失包的重构。通道前向纠错编码的缺点在于, 增加了传输的数据量, 加大了视频帧的延迟, 当丢失的数据包数量大于额外传输的数据包数量时, 纠错效果较差。

视频源信号前向纠错编码是在前向纠错编码基础上专为 Internet 视频传输作的改进。通道前向纠错的块编码区与视频信息无关, 也适用于其他数据, 但在视频源纠错编码中, 块编码区的冗余信息则主要是视频编码特征, 从而有效减少了冗余的信息量和延迟, 但因为对数据包进行了分组特征提取, 在纠错恢复时降低了视频质量。通道与视频信号源纠错编码符合香农的信号分离理论, 在很多场合得到了广泛应用<sup>[26]</sup>。联合通道与视频源纠错编码方法是对两种 FEC 方法的综合<sup>[27]</sup>。

##### 3) 重传丢失的数据包

重传丢失的数据包是在延迟允许的情况下对丢包最有效的处理方式, 重传机制一般遵循基本的请求重传协议 (automatic repeat-request, ARQ)。

重传一个丢失的数据包需要花费额外的  $RTT$ , 因此很多场合不选择重传的方式来处理媒体丢包。但是如果媒体从采样到播放所允许的最大延迟大于 1.5 倍  $RTT$  时, 接收端将感觉不到重传的发生, 即重传对于接收端是透明的<sup>[28]</sup>。由此可见, 在  $RTT$  很小的局域网中以及在 Internet 传输时接收方播放延迟较大的场合, 重传机制都是可行的。目前主要有基于接收端<sup>[7]</sup>、基于发送端<sup>[29]</sup>和基于混合模式<sup>[6]</sup>三类重传机制。

#### 1.4.3 基于 UDP 和 TCP 的实时流媒体传输技术比较

UDP 本身不具有流量控制、拥塞控制和差错控制机制, 使用 UDP 进行流媒体传输时通常结合 RTP/RTCP 与 TRFC 一起使用进行流量控制, 如文献<sup>[5,23,24,30]</sup>等, 但仍需要自行设计差错控制、重传、流量控制等策略, 并保证传输的 TCP 友好性; 同时还要求编/解码器具有一定的纠错、容错、错误隐藏等功能, 实现起来较为复杂。但这种方式的优点在于: 平均延迟较小, 实时性更强; 包的有效载荷较高, 传输效率高; 自定流量控制策略, 在网络可用带宽预测精确时, 流量波动较小; 可扩展性更强, 对不同流媒体类型和不同质量要求可制定不同的流量、拥塞及重传控制策略。

TCP 采用 AIMD 机制进行拥塞控制, 采用滑动窗口进行流量控制, 采用请求重发机制进行差错控制<sup>[31]</sup>。TCP 的发送速率乘法减小的特性使得传输速率陡降, 这会大大影响流媒体传输的平滑性, 而包重传机制也会增加数据包的延迟抖动和乱序, 降低流媒体的播放质量, 因而有研究者认为 TCP 的特性并不适合对于有时延要求并可容忍一定丢包率的流媒体传输<sup>[32]</sup>。但是 TCP 仍具有很多优势: TCP 是基于连接的可靠传输, 无须再另外设计差错控制与重传策略, 减少了对编/解码器的要求, 而且这更加适合于能够容忍一定程度的抖动但对丢包十分敏感的流媒体; TCP 对于网络的拥塞状况能迅速作出反应, 便于应用程序能及时地作出速率调整; 当前 Internet 上 85% ~ 95% 的应用是基于 TCP 的, 很多防火墙乐于接受 TCP 流但是却阻止非 TCP 的流<sup>[28]</sup>; 不需要考虑 TCP 友好的问题, 因为动态的 TCP 友好协议实现起来非常困难; 基于 TCP 的流媒体系统更易于实现, 运行也更加稳定, 很多商业的流媒体软件, 如 RealPlayer、Skype、QQ 等, 为了便于实现与维护, 都是直接使用 TCP 作为其默认的传输协议<sup>[33]</sup>。

由此可见,TCP 与基于 UDP 的方式各有优缺点。对基于 UDP 的 TFRC 流媒体传输的研究,主要在于基于 UDP/RTP/RTCP 的流量控制、拥塞控制、差错控制、自适应编码、端主机缓冲区管理、新的流媒体传输协议的探讨、流媒体的性能评估等方面,如文献[23,24,30,34]。对 TCP 流媒体传输的研究,主要集中在端主机缓冲区管理与流量自适应机制<sup>[35]</sup>、修改 TCP 以适应流媒体<sup>[28,36]</sup>、TCP 流媒体传输理论模型<sup>[33,37]</sup>、异构网络 TCP 流媒体传输<sup>[38]</sup>以及 TCP 流媒体传输的性能评估<sup>[39]</sup>等方面。

### 1.5 网络层

流媒体网络路由方面的研究主要集中在资源预留与路径交换方面,如文献[40,41]等。由于当前的 Internet 网络层并没有对流媒体提供 QoS 保障,绝大部分网络路由器并不能区分流媒体数据包与一般数据包,研究工作基本停留在探讨与小规模模拟实验阶段,随着路由器软硬件设备性能的不断改善,提供流媒体 QoS 保障的网络路由器的大规模推广应用将为期不远。

## 2 实时流媒体传输技术的应用

根据媒体源实时性的高低,可将流媒体传输技术的应用归纳为以下几类:

a) 源端媒体数据完全实时采样,实时发送的应用,如 VoIP 系统、视频监控、可视电话、远程医疗以及最新的 3G 实时通信系统等<sup>[42]</sup>,特点是对实时性的要求高于对质量的要求。

b) 源端媒体数据来自于媒体文件的应用,如视频点播、远程教育、网络广告等,特点是媒体以文件形式存于源端,接收端根据需求内容、播放质量等要求进行点播,对质量要求较高。

c) 源端既有采样数据又有存诸文件的应用,此时源端对分段时间的实时采样进行存储,形成分段实时数据文件后再进行发送,如 IPTV 交互式网络电视、P2P 直播等。

P2P 直播(应用层组播)是近年来实时流媒体传输技术应用的一个热点。其核心是将所有用户终端组织成一棵以流媒体服务器为根节点的组播树,通过用户终端上的数据中继实现实时传输<sup>[8]</sup>。由于流媒体服务器需要存储短期的实时采样数据,并将其分成若干传输单元以后再进行传输,基于 P2P 的网络直播与现场直播相比也有一定的延迟。最早的 P2P 直播商业软件是 2004 年的 CoolStreaming 系统<sup>[9]</sup>,目前已形成 PPLive、PPStream、UUSee、Rox、MySee、MSN Video 等诸多商业化网络直播或点播软件<sup>[43]</sup>。随着 3G 网络基础设施的完善和移动终端功能的增强,融合语音、文字、图像等多种形态的移动实时流媒体业务,如手机 P2P 直播<sup>[44,45]</sup>、移动网络会议等<sup>[46,47]</sup>,将成为实时流媒体传输技术应用的必然趋势。

## 3 结束语

本文认为在以下几个方面还需要更进一步的研究:

a) 低比特率、高丢包率环境下的服务质量问题。当前的实时流媒体传输技术在丢包率小于 3%,带宽大于 28 kbps 时才能满足基本的 QoS 要求。因此 Internet 流媒体的服务质量还存在很大的改善空间,实时流媒体传输的速率控制、拥塞控制及差错控制仍需不断改进完善。

b) 无线网络和强干扰网络环境下的速率自适应传输问题。Wi-Fi、WiMAX、3G 等无线网络与有线网络相比,涉及到包调度、带宽分配、网络切换、移动性支持等问题,太空通信、工业无线网等强干扰网络,则涉及长肥管道、高误码率等问题。这些均对实时流媒体的速率自适应传输技术提出了新的要求。

c) 使用流媒体专用路由与路径交换机制环境下,实时流媒体传输流程及策略的研究。一旦具有流媒体 QoS 支持的路由器在 Internet 上普及,Internet 将能够为实时流媒体提供一定程度的 QoS 保障,此时如何制定新的流媒体传输流程及相关策略也是未来的研究方向。

### 参考文献:

- [1] WU Da-peng, HOU Yi-wei, ZHU Wen-wu, *et al.* Streaming video over the Internet: approaches and directions[J]. *IEEE Trans on Circuits Systems for Video Technology*, 2001, 11(3): 282-300.
- [2] ZHANG Qian, GUO Qu-ji, NI Qiang, *et al.* Sender-adaptive and receiver-driven layered multicast for scalable video over the Internet [J]. *IEEE Trans on Circuits Systems for Video Technology*, 2005, 15(4): 482-495.
- [3] JAMMEH E A, FLEURY M, GHANBARI M. Rate-adaptive video streaming through packet dispersion feedback [J]. *IEEE Trans on Communications*, 2009, 3(1): 25-37.
- [4] CONKLIN G, GREENBAUM G, LILLEVOLD K, *et al.* Video coding for streaming media delivery on the Internet [J]. *IEEE Trans on Circuits Systems for Video Technology*, 2001, 11(3): 269-281.
- [5] SHOAIB M, WAHEED M. Streaming video in cellular networks using scalable video coding extension of H. 264-AVC [C]//Proc of IEEE International Conference on Wireless Communications, Networking and Mobile Computing. 2008: 1-4.
- [6] BONG J O, YU S B, KYEONG D M, *et al.* Efficient retransmission architecture of digital broadcast services over IPTV networks [J]. *IEEE Trans on Consumer Electronics*, 2008, 54(1): 65-70.
- [7] HUSZAK A, IMRE S. Source controlled and delay sensitive selective retransmission scheme for multimedia streaming [C]//Proc of IEEE Symposium on Computers and Communications. 2007: 191-196.
- [8] ZHANG Xin-yan, LIU Jiang-chuan, LI Bo, *et al.* CoolStreaming/DO-Net: a data-driven overlay network for peer-to-peer live media streaming [C]//Proc of the 24th Annual Joint Conference on IEEE Computer and Communications Societies. 2005: 2102-2111.
- [9] XU Chang-qiao, MUNTEAN G M, FALLON E, *et al.* A balanced tree-based strategy for unstructured media distribution in P2P networks [C]//Proc of IEEE International Conference on Communications. 2008: 1797-1801.
- [10] SCHULZRINNE H, CASNER S, FREDERICK R, *et al.* RFC 1889, RTP: a transport protocol for real-time applications [S]. 1996.
- [11] SCHULZRINNE H, CASNER S, FREDERICK R, *et al.* RFC 3550, RTP: a transport protocol for real-time applications [S]. 2003.
- [12] HANDLEY M, FLOYD S, PADHYE J, *et al.* RFC 3448, TCP friendly rate control (TFRC): protocol specification [S]. 2003.
- [13] TAO Shu, XU Kua. Improving VoIP quality through path switching [C]//Proc of the 24th Annual Joint Conference on IEEE Computer and Communications Societies. 2005: 2268-2278.
- [14] HIRANNAIA H R M, JASTI A, PENDSE R. Influence of codecs on adaptive jitter buffer algorithm [C]//Proc of the 66th IEEE Vehicular

- Technology Conference. 2007:2015-2019.
- [15] McNEILL K M, LIU Ming-kuan, RODRIGUE Z. Adaptive jitter buffer play-out scheme to improve VoIP quality in wireless networks [C]//Proc of IEEE MILCOM. Washington DC: IEEE Computer Society, 2006:1-5.
- [16] OKLANDER B, SIDI M. Jitter buffer analysis [C]//Proc of the 17th IEEE International Conference on Computer Communications and Networks. 2008:1-6.
- [17] HUANG Yong-feng, FENG Dan, ZHANG Jiang-ling. The study of the voice gateway architecture [C]//Proc of the 14th International Conference on High Performance Computing in the Asia-Pacific Region. Washington DC: IEEE Computer Society, 2000:704-705.
- [18] KAITRUNGRIT D, DAILEY M N, WUTIWIWATCHAI C. Thai voice application gateway [C]//Proc of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. 2008:101-104.
- [19] FLOYD S, FALL R K. Promoting the use of end-to-end congestion control in the Internet [J]. *ACM Trans on Networking*, 1999, 7(4):458-472.
- [20] GHAZALI O, HASSAN S. TCP-friendly layered multicast protocol for multimedia streaming [C]//Proc of IEEE International Conference on Electronic Design. 2008:1-6.
- [21] 陈志刚, 邓晓衡, 张连明, 等. 无线网络中 TCP 友好流媒体传输改进机制 [J]. *通信学报*, 2007, 8(5):22-29.
- [22] YAN Jin-yao, KOSTAS K, MARTIN M, *et al.* Media-and TCP-friendly congestion control for scalable video streams [J]. *IEEE Trans on Multimedia*, 2006, 8(2):196-206.
- [23] JIN Hai-peng, RAYMOND H, WANG Jun. Performance comparison of header compression schemes for RTP [C]//Proc of IEEE Wireless Communications and Networking Conference. Piscataway: IEEE Press, 2004:1691-1696.
- [24] GHARAVI H, BAN K, CAMBIOTIS J. RTCP-based frame-synchronized feedback control for IP-video communications over multipath fading channels [C]//Proc of IEEE International Conference on Communication. 2004:1512-1516.
- [25] 马宇峰, 魏维, 杨科利. 视频通信中的错误隐藏技术 [M]. 北京: 国防工业出版社, 2007.
- [26] ZHANG Sheng-li, ZHU Yu, LIEW S C, *et al.* Joint design of network coding and channel decoding for wireless networks [C]//Proc of IEEE Conference on Wireless Communications and Networking. 2007:779-784.
- [27] RADIC J, BEGUSIC D, ROZIC N. Joint source-channel, space time turbo channel decoding [C]//Proc of the 9th International Conference on Telecommunications. 2007:173-178.
- [28] YANG Xiao-gang, SHAN Bao-song, YANG Ya-ping. A rate control protocol for high speed streaming video [C]//Proc of IEEE International Conference on Information Systems. Washington DC: IEEE Computer Society, 2008:7-12.
- [29] LOGUINOV D, RADHA H. On retransmission schemes for real-time streaming in the Internet [C]//Proc of IEEE INFOCOM. 2001:1310-1319.
- [30] TOUFIK A, AHMED M, RAOUF B. Adaptive packet video streaming over IP networks: a cross-layer approach [J]. *IEEE Journal on Selected Areas in Communications*, 2005, 23(2):385-401.
- [31] ALLMAN M, PAXSON V, STEVENS W. RFC 2581, TCP congestion control [S]. 1999.
- [32] REJAIE R, HANDLEY M, ESTRIN D. RAP: an end-to-end rate based congestion control mechanism for real-time streams in the Internet [C]//Proc of IEEE INFOCOM. 1999:1337-1345.
- [33] WANG Bing, KUROSE J, SHENOY P. A model for TCP-based video streaming [R]. Amherst: University of Massachusetts, 2007.
- [34] FU Zheng-hua, MENG Xiao-qiao, LU Song-wu. A transport protocol for supporting multimedia streaming in mobile Ad hoc networks [J]. *IEEE Journal on Selected Areas in Communications*, 2003, 21(10):1615-1626.
- [35] XIONG Yong-hua, WU Min, JIA Wei-jia. Rate adaptive real-time video transmission scheme over TCP using multi-buffer scheduling [C]//Proc of the 9th IEEE International Conference for Young Computer Scientists. 2008:354-361.
- [36] EVENSEN K, PETLUND A, GRIWODZ C, *et al.* Redundant bundling in TCP to reduce perceived latency for time-dependent thin streams [J]. *IEEE Communications Letters*, 2008, 12(4):324-326.
- [37] KIM T, AVADHANAM N, SUBRAMANIAN S. Dimensioning receiver buffer requirement for unidirectional VBR video streaming over TCP [C]//Proc of IEEE International Conference on Image Processing. 2006:3061-3064.
- [38] HALEPOVIC E, WU Qian, WILLIAMSON C, *et al.* TCP over WiMAX: a measurement study [C]//Proc of IEEE International Symposium on Modeling, Analysis and Simulation of Computers and Telecommunication Systems. 2008:1-10.
- [39] SHAKKOTTAI, KUMAR S. TCP performance over end-to-end rate control and stochastic available capacity [J]. *ACM Trans on Networking*, 2001, 9(2):377-391.
- [40] YOO H, YOON J, LEE S, *et al.* A mobile reservation protocol for video streaming [C]//Proc of IEEE International Conference on Advanced Communication Technology. 2008:1842-1845.
- [41] TAO Shu, XU Kuai, XU Ying, *et al.* Exploring the performance benefits of end-to-end path switching [J]. *ACM SIGMETRICS Performance Evaluation Review*, 2004, 32(1):418-419.
- [42] XIONG Yong-hua, WU min, JIA Wei-jia, *et al.* Design and implementation of instant message system based on 3G [C]//Proc of the 26th Chinese Control Conference. 2007:656-660.
- [43] HUANG Cheng, LI Jin, KEITH W. Can Internet video-on-demand be profitable [J]. *ACM SIGCOMM Computer Communication Review*, 2007, 37(4):133-144.
- [44] SUMINO H, ISHIKAWA N, KATO T. Design and implementation of P2P protocol for mobile phones [C]//Proc of the 4th IEEE Workshop on Pervasive Computing and Communications. 2006:162-166.
- [45] MUELLER U, YOUNG M, GEFFLAUT A. Running the Windows P2P infrastructure on mobile phones [C]//Proc of the 7th IEEE International Conference on P2P Computing. Washington DC: IEEE Computer Society, 2007:245-246.
- [46] VENOT S, LU Yan. On-demand mobile peer-to-peer streaming over the JXTA overlay [C]//Proc of IEEE International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies. 2007:131-136.
- [47] WANG Lei. EPMD: a 3G faced P2P e-commerce platform for mobile devices [C]//Proc of International Conference on Computing, Communication, Control, and Management. Washington DC: IEEE Computer Society, 2008:186-190.