



基于被动测量的流媒体业务性能分析

张杰^{1,2} 程光^{1,2}

(¹东南大学计算机科学与工程学院, 南京 21189)

(²教育部计算机网络和信息集成重点实验室(东南大学), 南京 211189)

摘要: 随着流媒体业务的广泛使用, 用户对流媒体业务的性能要求不断提高。现有的网络测量技术无法满足现实需求。服务质量(QoS)是衡量网络流媒体业务的重要指标。本文在分析流媒体业务 QoS 测量原理的基础上, 设计并实现一个流媒体业务性能测量系统, 采用基于端到端的流媒体 QoS 被动测量方法, 对流媒体业务网络层和应用层的性能指标进行了研究和分析。研究结果能为流媒体业务系统的性能优化提供参考。

关键词: QoS; 性能测量; 流媒体业务

Streaming media services performance measurement based on passive testing

Zhang Jie^{1,2} Cheng Guang^{1,2}

(¹School of Computer Science and Engineering, Southeast University, Nanjing 211189, China)

(²Key Laboratory of Computer Network and Information Integration, Ministry of Education, Southeast University, Nanjing 211189)

Abstract: With the development of streaming media business, users' demands on network service performance are improving. Network measurement can not meet real demands. Quality of service (QoS) is an important indicators to measure streaming media services. Based on the analysis of streaming media service QoS, we design and implement a streaming media business performance measurement system, propose an end-to-end passive measurement method of streaming media QoS, study and analysis streaming media services' network layer and application layer performance index. The result can provide a reference for optimizing the streaming media business system.

Key words: QoS; Performance Measurement; Streaming Media

1 引言

随着 Internet 技术的快速发展, 各种新型网络应用不断涌现。其中, 流媒体业务更是发展迅猛。根据中国互联网络信息中心最新发布的中国互联网络发展状况统计报告显示, 截至 2013 年 12 月, 中国网络视频用户规模达 4.28 亿, 较上年底增加 5637 万人, 增长率为 15.2%。网络视频使用率为 69.3%, 与上年底相比增长 3.4 个百分点, 是受众最多的互联网业务之一^[1]。用户对流媒体业务提出服务质量要求, 服务提供商为了留住用户, 必须确

保良好的用户体验质量。服务提供商希望得到准确可靠的测量数据, 能够对提供的网络业务质量进行评估, 从而发现问题, 检测存在的或潜在的网络故障。这是一个亟待解决的问题。

由于各种原因, 在以 IP 协议为核心的互联网上提供流媒体业务时, 流媒体业务的性能经常随着网络和时间而变化, 不能满足用户的需求。因此, 运营商需要通过流媒体业务性能测量, 对流媒体业务的特征、可靠性、可用性等一系列性能参数进行获取与分析, 从而实时掌握业务的运行情况, 着力改善业务质量。流媒体业务性能测量能够提供一个中立、客观的流媒体业务评估结果, 可以作为用户提供流媒体业务选择的依据。对于网络运营商来说, 通过流媒体业务性能测量提供的端到端的性能数据, 可以了解网络当前的运行状态, 并辅助进行承载网的故障诊断, 提高用户访问流媒体业务的

作者简介: 张杰, (1988-), 男, 硕士研究生, E-mail: jiez.csu@foxmail.com; 程光, (1973-)男, 教授, 博导, E-mail: gcheng@njnet.edu.cn.



数量和质量。因此，流媒体业务性能测量不论是对内容提供商还是运营商都具有重要意义^[2]。

流媒体业务测量从测量内容上看，可以按照网络协议层次分为 3 类：网络层性能测量、传输层性能测量和应用层性能测量。网络层性能测量指标包括连通性、带宽利用率、吞吐率等；传输层性能测量主要针对某个特定的流媒体传输会话，测量流媒体传输的丢失率、延迟和延迟抖动等服务质量参数；应用层性能指标往往针对特定应用，旨在反映网络对应用提供的服务能力。例如测量流媒体播放器的缓冲时间、视频帧丢失情况等^[3]。

流媒体业务测量的分类标准有多种^[4]。根据测量的方式，可分为主动测量和被动测量；根据测量点的多少，可分为单点测量与多点测量；根据测量的内容，可分为拓扑测量与性能测量；根据测量点的位置，可分为内部测量、中间点测量和端点测量。

本文的研究思路是从流媒体业务最终用户的角度测量和评估流媒体服务的性能。采用端到端的被动测量方式，通过在客户端采集视频流数据，获取流媒体视频流在网络层、传输层和应用层的性能指标，全面反映流媒体服务性能，综合评估流媒体业务服务质量。

本文的组织结构如下：第二节介绍流媒体业务性能测量研究现状；第三节分析了影响流媒体视频质量的因素；第四节介绍流媒体测量指标及系统设计；第五节分析实验结果；第六节总结并阐述下一步的工作。

2 现有视频质量评估方法的研究

视频质量评估从方法上可以分为主观评价方法和客观评价方法，前者凭感知者主观感受评价对象的质量；后者依据模型给出的量化指标衡量视频图像质量。

2.1 视频业务质量的主观评价方法

国际电信联盟 (ITU) 在 ITU-R BT.500-11 中推荐使用平均主观得分法 (MOS) 来评价接收者感知的媒体质量^[5]。它是指在受控的测试环境下，将待评价的图像序列播放给评价者观看，并记录他们的打分，然后对所有评价者的打分进行统计，得出平均分 (MOS) 作为评价结果。接收者采用 5 分制评价，其中 5 分为很好 (Excellent)，4 分 (Good)，3 分为中 (Fair)，2 分为差 (Poor)，1 分为不可

接受 (Unsatisfactory)。如下图 1 所示：

得分	视频失真分级	视频质量分级
5	不能觉察 (Imperceptible)	优秀 (Excellent)
4	能觉察，但不明显 (Perceptible, but not annoying)	良好 (Good)
3	能明显觉察，还可以容忍 (Slightly annoying)	一般 (Fair)
2	令人讨厌 (Annoying)	差 (Poor)
1	很讨厌 (Very annoying)	很差 (Bad)

图 1 视频主观评价得分标准

MOS 为接收媒体质量提供了定量指标，但是对流媒体而言，这种评价方法存在着弊端。MOS 的特性使得必须有接收者参与，这就花费大量的人力和时间，并且这种方法不能避免主观因素的影响。常用的主观评价方法有：双刺激连续质量分级法 (DSCQS)，单刺激连续质量评价法 (SSCQE)。

2.2 视频业务质量的客观评价方法

国际电信联盟所属的视频质量专家组提出了根据对原始参考视频需要的程度，视频客观质量评估方法可以分成三大类：全参考视频客观质量评估，部分参考视频客观质量评估和无参考视频客观质量评估^[6]。

(1) 全参考视频客观质量评估方法

全参考视频客观质量评估需要完整的原始参考视频，通过比较失真视频和原始参考视频，得到失真视频的质量，例如目前广泛使用的峰值信噪比 (PSNR) 等。

(2) 部分参考视频客观质量评估方法

部分参考方法是指在原始参考视频和失真视频上分别施加某种运算，各自得到少量的统计数据，然后对统计数据进行分析，从而评估视频质量。常用方法有：基于数字水印的方法。

(3) 无参考视频客观质量评估方法

无参考方法不需要任何原始参考视频的信息。现有的大多数无参考方法是通过视频的处理和分析提取视频序列中出现的某些失真特征 (如模糊、块效应等)，然后根据各类失真特征来判定视频序列的质量。常用方法有：基于神经网络的视频质量评估。



3 流媒体视频影响因素分析

当前有限的网络带宽决定了视频要在网络上实现高速传播,必须对视频进行压缩处理,从而缩减视频文件的大小以适应网络带宽。这些处理必然会带来视频质量的下降。同时,不可靠的网络传输同样会对视频质量产生不可预期的损伤^[7]。

(1) 视频压缩对视频质量的损伤

原始视频数据体积非常大,要在互联网上传输视频数据就必须对其进行压缩。视频压缩算法对视频造成损伤的一个关键点就是对 DCT 系数矩阵进行量化。量化过程造成了视频图像的细节丢失。视频压缩导致视频出现失真,不同的编码方法,将产生不同程度的各种视觉失真效应。例如,基于块的编码算法产生的方块效应较为严重,而基于小波变换的压缩编码,最易引起的视觉失真是模糊和震荡效应。

(2) 网络传输对视频质量的损伤

网络传输失真包括丢包、延时和抖动等因素。流媒体视频对网络失真比较敏感,特别是丢包造成的失真。丢包是媒体包在传输过程中由于网络拥塞、故障等情况造成的丢失。丢包对视频质量的影响取决于丢失的信息。如语法信息的丢失和语义信息的丢失对视频质量的影响就不相同。一般情况下,语法信息(如包头信息)的丢失对视频质量的影响大于语义信息(如运动矢量 MV)对视频质量的影响。因此网络丢包可能造成视频的局部或者全部的丢失。MPEG-2 编码是分层次的,在每个层次都有相应的数据头部和数据载荷。如果将数据头部丢失,将导致解码器无法解码,造成视频的全部丢失。如果只是一个帧的部分片(slice)信息丢失,则会导致视频的部分丢失。根据包的大小,丢包所影响的可能是个视频帧内的局部区域,也有可能整个视频帧的丢失。

现有视频标准采用基于运动补偿的预测、正交变换和熵编码器技术的混合编码框架,编码效率的提高在很大程度上仍依赖于视频帧内及帧间的相关预测和补偿。因此,由于网络传输错误而导致的误差积累会严重影响到解码视频的质量。例如,空间预测编码技术使得在一个视频帧内沿空间方向积累错误。帧间预测编码中通常使用已解码的视频帧中某些宏块进行运动补偿,因此将会导致空-时方向的错误积累。丢包对视频图像所造成的损伤可分为空

域损伤和时域损伤。空域损伤导致图像帧中空间结构信息被破坏。空域损伤的表现效果有:拉伤、不规则块现象和扭曲等。时域损伤是指视频内容的连续性在时间轴上受到的破坏。时域损伤可导致帧冻结或帧重复。

4 流媒体测量方法

本文针对基于 RTSP/RTP 协议流媒体业务,采用端到端的流媒体 QoS 被动测量方法,通过仿真实验环境,在客户端被动监测流媒体业务性能指标,从而了解流媒体业务的性能状况。

4.1 流媒体测量过程

图 2 中描述的是基于 RTSP^[8]协议的流媒体视频点播系统客户端与服务器端的交互过程。从图中可以看出,客户端与服务器端的交互过程可以分为四个步骤:(1)用户在浏览器中向 Web 服务器发送请求,Web 服务器把相关的流媒体服务器地址、资源路径及编码类型等信息提供给客户端。客户端启动流媒体播放器,与流媒体服务器进行连接。(2)客户端通过 RTSP 协议发送控制播放命令。服务器端收到 RTSP 协议中的播放命令后便开始向客户端发送流媒体数据。(3)客户端接收服务器端传输的视频和音频数据,并向服务器的发送 RTCP 接收报告,(4)数据传输完毕后,客户端通过 RTSP 协议发送释放命令结束与服务器的连接^[9]。

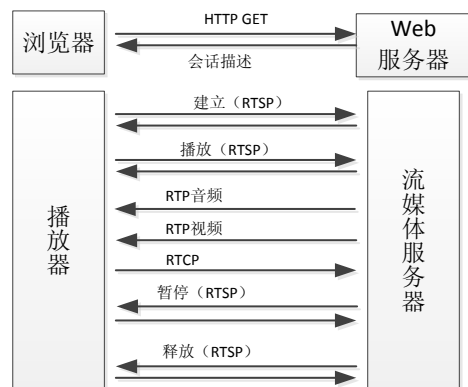


图 2 流媒体业务交互过程

根据流媒体业务的交互过程,在客户端抓取数据报文,通过获取利用 RTP 包的顺序号、时间戳以及 RTCP 携带的控制信息,可以提取、计算出网络层和应用层的性能指标。测量的具体过程如图 3 所



示。

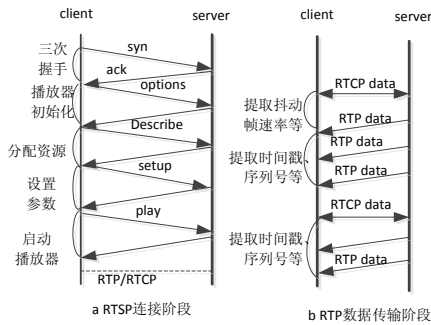


图3 具体测量过程

4.2 流媒体测量指标的获取

对流媒体视频的客观评价指标包括传输层指标和应用层指标两部分。

网络层测量的性能指标及获取方法如下：

(1) 丢包率

丢包率是指一段时间内到达接收方的包的数量与发送方发送的总数的比值。丢包率测量的前提是 RTSP 连接已经建立，客户端发出 PLAY 方法的请求后，服务器开始传送数据。测量方法：利用 RTP 包的顺序号(sequence number)。假设客户端接收到的数据包的最小序列号为 min_seq，接收到的数据包中最大的序列号为 max_seq，收到的数据包个数为 N(统计得到)，则传输过程中报文丢包率为： $1 - N / (max_seq - min_seq)$ 。

(2) 往返时延

往返时延是指 TCP 连接建立后，客户端通过 RTSP 协议发送请求报文，服务器端收到请求后发送回相应的 RTSP 响应报文的时延。

往返时延的计算方法：客户端发送给服务器端的数据包为 C，服务器给客户端的数据包为 S，双向时延 D 的计算如下：

$$\text{If } (C(j).seq = S(i).ack)$$

$$\text{then } D(i) = \text{Time}(S(i)) - \text{Time}(C(j))$$

其中，seq 是数据包的 TCP 包头中的 Seq 字段，代表本数据包的序列号。ACK 是数据包 TCP 头中的 ACK 字段，代表数据包的确认数据包字段，或者说是希望收到对方下一个数据包的序列号，Time 是数据包被截取时的本地时间戳。RTSP 协议中，响应报文的确认序列号等于请求报文的序列号。

(3)时延抖动

时延抖动是指相邻两个时延之间的差值。通过将一段时间内的时延抖动取平均值，可以得到一段时间内的平均时延抖动。包 i 和包 j 之间增加的延迟差计算公式如下：设 R_j 代表第 j 个包的到达时刻， S_j 代表第 j 个包的 RTP 时间戳值，则第 i 个 RTP 报文与第 j 个 RTP 报文间的抖动为 $D(i, j)$ 。

$$D(I, j) = (R_j - R_i) - (S_j - S_i) = (R_j - S_j) - (R_i - S_i)$$

(4) 乱序率

如果本次收到的数据包的编号比上一次收到的编号小，则称为发生一次乱序。乱序率等于发生乱序的数据包与所有收到数据包的比值。

通过解析网络层抓取的 RTP 报文，可以得到视频帧的相关信息。应用层测量指标及获取方法如下：

(1) 帧持续时间

视频帧的传输往往需要多个数据报文，对每一个视频帧，必要要所有的报文都达到播放器缓冲区后才能解码播放。因此，帧在缓冲区中有段缓存时间。帧持续时间就是视频帧中第一个到达缓冲区的 RTP 报文时间和最后一个到达的 RTP 报文时间的差值。

(2) 帧丢失比率

一个视频帧中出现连续丢包或者包含视频帧头部的 RTP 报文丢失时，可认为整个视频帧丢失。计算丢帧比率就是计算所有丢失的帧的数目与整个视频帧数目的比值。同时，丢帧比率还可以根据帧的类型进一步细化为 I-帧丢失比率、P-帧丢失比率和 B-帧丢失比率。

(3) 帧损伤比率

网络状态不稳时会发生丢包，大多数时候丢包情况是离散的。当一个视频帧中丢包数很少时，会对视频帧造成损伤，而不是视频帧丢失。计算帧损伤比率就是计算所有损伤的帧数目与整个视频帧数目的比值。

4.3 系统设计

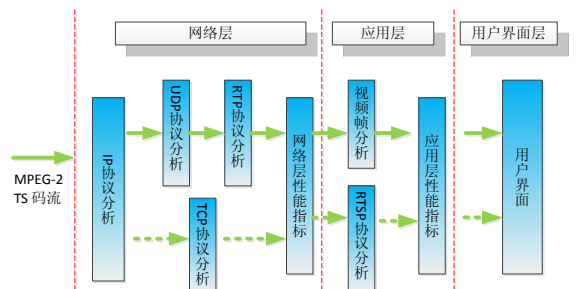


图4 系统整体框架



如上图 4 所示，本系统整体架构可分为三个层次：

(1)网络层

网络层是整个系统的基础。网络层接收数据包，根据 TCP 报文和 UDP 报文进行分类，对 TCP 报文统计 TCP 连接建立时延，进而解析在其之上的 RTSP 协议；对 UDP 报文根据 RTP 等协议分析数据包，提取数据包头的关键字段，计算出当前网路状况指标（如丢包、时延、抖动等），同时根据 RTP 包头的序号字段对乱序的视频数据进行重新排序。

(2)应用层

应用层是系统的关键。在该层对 TS 码流进行分析，进而得到整个视频流中所有的帧信息。在此基础上统计整个视频流中帧的数目、帧的类型，帧丢失数目、不同类型帧丢失数目等信息；对 RTSP 协议进行解析，计算请求响应时延，请求成功率等反映用户与服务器交互情况的指标参数。这些参数对评估网络性能和视频质量有重要作用。

(3)用户界面层

在用户界面层可以直观的查看所有反映网络性能和视频质量的参数。系统显示出整个视频码流的统计信息，包括网络层的时延、丢包率、乱序率；应用层的帧序号、帧类型、丢帧率、丢帧类型等。通过给出系统对整个视频质量的评价，并检测出这些指标参数对视频质量造成的影响。

5 实验分析

5.1 流媒体测量方案

本文的测量方案通过搭建仿真实验环境来模拟整个流媒体的传输过程。在实验中，使用开源软件 live555 作为流媒体服务器，在客户端使用播放器 VLC 来接收流媒体视频。在安装有 fedora 系统的服务器上使用 Zebra 来实现路由器功能，用 Netem 对路由器端口进行丢包和时延控制，从而模拟不同的网络环境。如下图 5 所示。

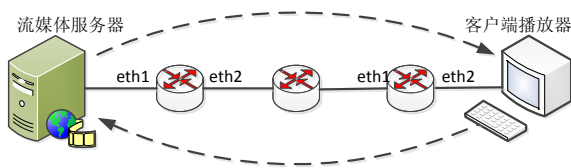


图 5 测量方案示意图

在进行实验时，将视频源文件采用 MPEG-2 标准进行编码后流媒体服务器中，在客户端使用播放器 VLC 作为接收流媒体数据的播放器。在流媒体系统中，采用 RTSP 协议控制客户端与服务器进行交互，采用 RTP/RTCP 协议传输流媒体数据。在客户端采集流媒体数据报文，报文的采集是通过基于 WinPcap 编程实现的。通过分析采集到的数据报文，获取流媒体业务的性能指标。测试视频分为标清视频和高清视频两类。服务器上保存有 720P 的标清视频和 1080P 的高清视频，标清的视频像素为 1024×768，高清视频的视频像素为 1920×1080。测试中我们通过使用 Netem 中的命令，限制路由器接口的带宽以模拟现实网络中接入网带宽不足以及损伤等现象。

5.2 实验分析

(1) 标清视频测试

通过对标清视频进行测试，得到标清视频码率速率约为 10Mbps。因此对标清视频，我们将线路速率从 11Mbps 开始，依次为 11M, 10.5M, 10M, 9.5M, 9M, 8.5M, 8M, 7.5M 和 7M。在客户端获得这 10 组数据并进行分析。在视频测试的基础上同步进行了基于双刺激损伤评价 DSIS 的主观视频评估对比测试，邀请了 10 位观察者对这段高清视频在不同带宽下进行重复观看，记录用户体验的主观描述数据，并按视觉效果的好坏进行打分，最后求出 MOS 平均值。

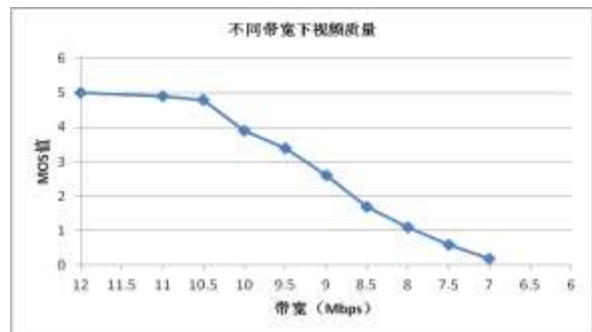


图 6 标清视频质量随带宽变化

从图 6 中可以看出，视频质量随着带宽下降而下降，当带宽小于 10.5Mbps 时，视频质量下降迅速。通过对获取的视频码率报文数据进行分析，可以获得视频码流在不同带宽条件下性能参数。下表 1 显示了通过被动测量方式获取的各种网络性能参数。



表 1 标清视频网络性能参数

带 宽 (Mbp s)	MO S 值	丢包 率 (%)	时延 (ms)	抖 动 (ms)	帧丢 失率 (%)	帧损 伤率 (%)
10.5	4.8	2.4	19	35	4	63.1
10	3.9	7.7	20	55	9.2	85.9
9.5	3.4	13.2	21	80	10.6	87.2
9	2.6	19.4	25	101	14.9	83.2

从上表中可以看出,随着带宽的下降,网络的时延、抖动、丢包都相应的增加。媒体丢包率小于2.4%时, MOS 值大于4,视频质量优秀;当媒体丢包率大于2.4%小于19.4%时, MOS 值变化较快,持续递减,视频质量逐渐下降;当媒体丢包率大于17.5%时, MOS 值小于2,视频基本无法观看。同时, RTP 报文的丢失导致应用层视频帧的丢失。通过上表可以看出,视频帧丢失率与丢包率有着对应关系,当帧丢失率小于4%时, MOS 值大于4,视频能正常观看;当帧丢失率大于4%小于14.9%时, MOS 值持续递减,能看出视频失真但是还能够忍受;当帧丢失率大于14.9%时, MOS 值小于2,视频基本无法观看。通过对比帧损伤率与带宽和 MOS 值的关系发现,无论带宽是大是小,帧损伤率都保持在很大的比值,这说明报文丢包大部分都是离散丢包的,连续丢包的情况很少。因此,丢包导致的视频帧损伤影响范围很大。

(2) 高清视频测试

通过对高清视频进行测试,得到高清视频的码率15.5Mbps。因此对标清视频,我们将线路速率从16.5Mbps开始,依次为16.5M,16M,15.5M,15M,14.5M,14M,13.5M,13M,12.5M和12M。在客户端获取这10组数据并进行分析。在视频测试的基础上同步进行了基于双刺激损伤评价 DSIS 的主观视频评估对比测试,邀请了10位观察者对这段高清视频在不同带宽下进行重复观看,记录用户体验的主观描述数据,并按视觉效果的好坏进行打分,最后求出 MOS 平均值。

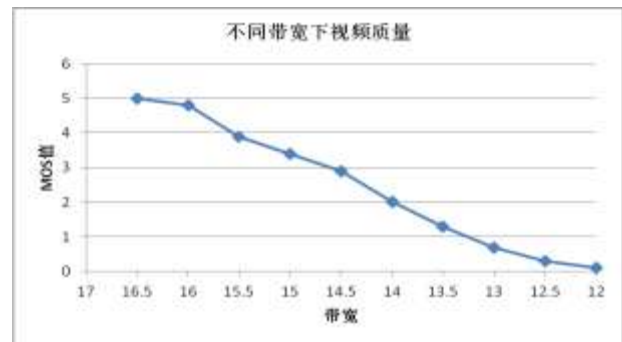


图 7 高清视频质量随带宽变化

从图7可以看出,当带宽小于16Mbps时,视频质量下降迅速。通过对获取的视频码率报文数据进行分析,可以获得视频码流在不同带宽条件下性能参数。表2中显示了通过被动测量方式获取的各种网络性能参数。

表 2 高清视频网络性能参数

带 宽 (Mbps)	MOS 值	丢包 率 (%)	时延 (ms)	抖 动 (ms)	帧丢 失率 (%)	帧损 伤率 (%)
16	4.8	0.8	15	50	0.7	36.6
15.5	3.8	3.8	18	60	0.3	77.3
15	3.4	7.3	19	73	0.6	84.4
14.5	2.7	11.2	20	64	0.8	86.1

从上表中可以看出,随着带宽的下降,网络的时延、抖动、丢包都相应的增加。媒体丢包率小于0.8%时, MOS 值大于4,视频质量优秀;当媒体丢包率大于0.8%小于11.2%时, MOS 值变化较快,持续递减,视频质量逐渐下降;当媒体丢包率大于11.2%时, MOS 值小于2,视频基本无法观看。同时,视频帧丢失率也随之带宽的下降而增大。视频帧损伤率依然保持很高的值。

通过对比标清视频和高清视频我们发现,高清视频比标清视频对丢包的敏感性更高。这是因为高清视频图像更大,因此一个完整的视频帧中任意一个报文丢失的概率也更大。同时我们也发现,当网络带宽大于视频码流速率时,带宽对视频质量影响不大;当网络带宽小于视频码流速率时,带宽改变对视频质量有很大的影响。



6 总结

本文在分析流媒体业务 QoS 测量原理的基础上,通过基于端到端的流媒体 QoS 被动测量方法,设计并实现了一个流媒体业务的性能测量系统;通过仿真实验,对流媒体业务网络层和应用层的性能指标进行了研究,深入分析了带宽、丢包、时延抖动以及帧丢失、帧损伤对流媒体业务性能及流媒体视频质量的影响。上述工作能及时向流媒体服务器反馈信息,实现动态的质量控制,并且为流媒体业务系统的优化,流媒体视频质量的评估提供依据。

目前只是研究了网络层和应用层的性能指标对流媒体业务性能的影响,对于如何评估用户感知到的流媒体视频质量,目前还没有一个具体的方法。下一步的工作需要研究能够评估视频质量的模型。

参考文献

- [1] 中国互联网络信息中心.中国互联网络发展状况统计报告[R] 2014.01.
- [2] 张玫,曾彬,黎文伟,张大方.面向端用户的流媒体业务性能监测与分析[J].计算机科学与工程,2010,32(12):7~11.
- [3] ITU-T Rec E. 800-1994, Terms and definitions related to quality of service and network performance including dependability[J].1994
- [4] 张宏莉,方滨兴,胡铭曾,姜誉,张树峰. Internet 测量与分析综述[J].软件学报, 2003, 4(1):11~16.
- [5] ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television pictures [S]. 2002.
- [6] Z. Wang, AC Bovik. A universal image quality index. IEEE Signal Processing Letters. 2002, 9(3):81-84.
- [7] 姚继先, 张远, 朱雨涵. 基于特征提取的视频通信质量客观评价模型[A]. 图像图形技术研究与应用 2009—第四届图像图形技术与应用学术会议论文集[C].北京, 2009
- [8] 李慧彬. 基于 RTP/RTCP 协议的 QoS 研究[J]. 机械管理开发. 2008, 23(1). 79-80
- [9] 解冲锋, 杨明川, 王琳. 基于 RTP/RTCP 协议的流媒体业务性能测量[J]. 网络信息, 2003. 7