

网络自适应传输控制策略研究

高东日, 魏海平, 姜 东

(辽宁石油化工大学信息工程学院, 抚顺 113001)

摘要: 针对流媒体传输中的拥塞问题, 对网络自适应传输控制进行了研究, 结合 RTP 协议的传输特性, 提出了一种基于带宽的网络自适应传输策略并阐述了这种策略的控制机理。该策略能够根据网络资源状况动态调节传输速率、匹配网络带宽, 实现了网络自适应传输。

关键词: 实时传输协议; 实时传输控制协议; 网络状态估计; 自适应传输控制策略

Research on Network Adaptive Transmission Control Strategy

GAO Dongri, WEI Haiping, JIANG Dong

(College of Information Engineering, Liaoning University of Petroleum & Chemical Technology, Fushun 113001)

【Abstract】 After studying the stream media congestion question, the network adaptive transmission control and the transmission characteristic of RTP protocol, this paper proposes a kind of network adaptive transmission control strategy based on the bandwidth and expatiates this transmission strategy control mechanism. According to network resource condition, this strategy adjusts dynamically transmission speed to match the network bandwidth change, and achieves network adaptive transmission.

【Key words】 Real-time transport protocol(RTP); Real-time transport control protocol(RTCP); Network state estimate; Adaptive transmission control strategy

多媒体流的实时传输对带宽和延迟比较敏感, 要求有一定的服务质量保证。目前的 Internet 只提供尽力而为的服务, 它的带宽、延迟是随时间变化的, 在网络上实时传输音频、视频流, 必须采取防止网络拥塞的措施, 仅仅依靠传统的一些互联网协议是无法满足实际需求的。为了解决目前面临的问题, 互联网工程任务组 IETF 陆续提出了一些更高效的协议, 如 Real-time Transport Protocol(RTP)/Real-time Transport Control Protocol(RTCP)、Resource Reservation Protocol(RSVP)、Real-Time Streaming Protocol(RT-SP)等。这些协议共同完成了一个多媒体流的传输协议系统。目前通常采取的网络传输策略有 2 种: (1)预留网络资源, 来保证足够的网络带宽, 这需要整个网络通路都必须支持资源预留协议 RSVP; (2)在终端系统中采取措施, 使其发送到网络上的数据量能够与变化的网络带宽相匹配。由于目前网络环境中的许多路由器等网络设备还不支持资源预留, 因此在复杂的网络环境中无法实现预留网络资源策略, 只能通过调整发送端的传输速率来匹配网络带宽的变化, 达到多媒体数据的网络自适应传输。

1 实时传输协议 RTP

RTP 协议是多媒体数据流的实时传输协议。RTP 协议充分体现了应用层分帧这一通信协议的设计思想, 允许用户了解、调整甚至制定连续媒体的打包方案。由于该协议能够确保具体应用对于每一个 RTP 包所携带的数据内容有足够的了解, 因此使得应用程序可以根据 RTP 包头中的序号、时间戳等内容及连续媒体的编码方式的特点, 有效地判断各种差错以及网络流量的变化, 分析实时数据的传输质量的影响程度, 并在考虑具体应用特点、网络环境等因素的基础上, 选用最为恰当的方法灵活地完成拥塞控制、差错控制、同步控制等任务, 以更好地满足实时应用的要求。整个 RTP 协议包

括: 负责数据传输的 RTP 和负责控制部分的 RTCP。

1.1 RTP 数据包

RTP 包由头部和数据部分组成。RTP 包包含一个固定长度的 RTP 报头、可能为空的有效源列表和有效载荷数据。媒体数据直接跟在 RTP 报头之后。

1.2 实时传输控制协议 RTCP

RTCP 是 RTP 协议中的控制功能部分, 和 RTP 一起提供流量和拥塞控制服务。RTCP 的主要功能如下:

(1)QoS 监视和拥塞控制: RTCP 包的报告块中提供了 QoS 监视的必要信息。如 RR 中含有所接收的最高顺序号、丢失率、累计丢失的包数、包间抖动等。

(2)媒体间同步: RTCP 的 SR 包中包含实际时间 NTP 时间戳和相应的 RTP 时间戳, 可用于不同媒体间的同步。

(3)识别信息: RTP 数据包只能通过 SSRC 来识别源, 而 RTCP 的 SEDS 包中有足够的文本信息, 如 CNAME、用户名等。

(4)会话大小估计和控制信息量的调节: 参与会话的每个成员周期性地发送 RTCP 包, 各站点可据此估计和计算出参与会话的人数, 调节实时控制信息量, 使控制信息量和媒体业务量达到平衡。

2 网络状态估计策略

由于在实际应用中 RTP/RTCP 是运载在 UDP 上, 而 UDP 是不可靠的协议, 它没有 TCP 的流量控制等能力, 因此常用的算法是基于已丢失报文的评估算法。

2.1 数据包丢失率的计算

发送端根据接收者报告 RR 计算包丢失率、数据包到达间隔抖动和回路时间等, 确定网络负载的状况, 判别网络信道状况。在这几个指标中通常只把平滑后数据包丢失率作为网络带宽调整的依据, 它基本上能反映网络的传输状况。

作者简介: 高东日(1980 -), 男, 硕士生, 主研方向: 计算机网络安全与多媒体; 魏海平, 教授; 姜 东, 硕士生

收稿日期: 2006-02-02 **E-mail:** gaodongri@163.com

(1)计算该发送时隙丢失的数据包数 $L(n)$:根据发送端一个发送时隙前后接收的 2 个 RR 包中包丢失的累计数之差,就是该发送时隙丢失的数据包数,即

$$L(n)=C(n)-C(n-1) \quad (1)$$

其中, $C(n)$ 是第 n 个发送时隙后共丢失的数据包数。

(2)计算该时隙中应收到的数据包的数 $R(n)$ 。2 个接收数据包中的最高序列号扩展的差就是接收者在该时隙中应收到的数据包数。

$$R(n)=H(n)-H(n-1) \quad (2)$$

其中, $H(n)$ 是第 n 个发送时隙后收到的数据包的最高序列号。

(3)计算该时隙中数据包丢失率 $F(n)$, $L(n)$ 与 $R(n)$ 之比即在该时隙中数据包丢失率。

$$F(n)=L(n)/R(n) \quad (3)$$

数据包丢失率是监控器估计网络信道状况的指标,但不能直接利用丢失率判断网络信道状况并据此调整视频的发送,因为会使发送数据变动得过于频繁,也使接收端的图像质量不稳定。在利用数据包丢失率估计网络信道状况之前,先对其作平滑处理。可设 $t(n)$ 是平滑后的丢失率,其平滑关系为

$$T(n)=(1-a)*T(n-1)+a*F(n), 0 \leq a \leq 1 \quad (4)$$

当 a 值增加时,当前数据包丢失率 $F(n)$ 对平滑的结果 $T(n)$ 影响增大;当 a 值取低时,以前的平滑结果 $T(n-1)$ 对当前产生的平滑结果 $T(n)$ 影响增大。根据实验的统计结果来确定 a 值,会使平滑处理的结果削弱网络信道服务质量 QoS 振荡作用,用平滑后的数据包丢失率进行网络信道状况的判断有利于合理调整视频的发送。

2.2 网络状态估计算法

根据平滑后的数据包丢失率与事先确定的阈值的关系,可以判断网络上的负载情况,依此将网络状态分为 3 类:轻载,满载和阻塞。

网络状态的分类算法:设定 2 个阈值 k_1 、 k_2 (测试给定值),平滑后 RTP 数据包的丢失率为 $t(n)$ 。

- (1)当 $t(n) < k_1$ 时,视为轻载;
- (2)当 $t(n) > k_1$ 且 $t(n) < k_2$ 时,视为满载;
- (3)当 $t(n) > k_2$ 时,视为阻塞。

下阈值 k_1 的值应该设置的尽量小,以消除服务质量 QoS 振荡的影响;上阈值 k_2 选择的值虽然使丢包的程度很高,并已经影响了数据的传输,但解码后的视频效果在可以接受的范围之内。阈值 k_1 和阈值 k_2 的选择具有任意性,但必须得到实验验证。对于视频数据来源来说, k_1 的值为 3%, k_2 的值为 5%较为合适。网络状态分类如图 1 所示。

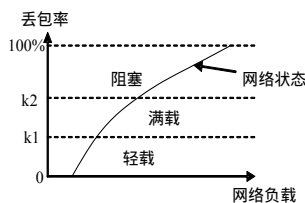


图 1 网络状态分类

3 自适应网络传输控制算法

基于传输层和应用层自适应网络技术有多种,在主动自适应方法中,应用系统的接收端与发送端连接的传输结构有单播和组播。

3.1 单播方式下的调整算法

在单播的传输结构中,即点对点的情况,决策机制可直

接按照网络的 3 种状态映射为“提高”、“保持”或“降低”视频帧率的发送。单播方式下的帧率调整算法如图 2 所示。

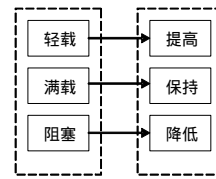


图 2 单播方式下的帧率调整算法

3.2 组播方式下的调整算法

在组播连接环境下,即点对多点的情况,许多接收端同时接收实时数据,并送回接收报告,这种情况下不能因为某端点连接造成很高的数据包丢失率,就将视频数据的发送带宽降低。为避免上述问题,可采用如下算法:

(1)当某接收端连接有很高的数据包丢失率时,决策机制直接按照网络的 3 种状态去调整发送带宽。该算法的优点是可以避免任何阻塞,接收端都能接收视频服务,缺点是带宽较好的客户端接收的视频服务质量较差。

(2)当某接收端连接有很高的数据包丢失率时,将其与发送端的连接改为单播的传输结构,然后再根据接收端的类型状态去调整带宽。该算法的优点是可以避免任何阻塞,接收端都能接收视频服务,缺点是增加了发送端的工作负担。

(3)计算由接收端观察到的网络状态的数目,根据轻载、满载和阻塞的数目比例作出调整发送帧率决策。假设处于轻载、满载和阻塞的数目分别为 N_{mild} 、 N_{full} 、 N_{block} ,则总数目 $N_{sum}=N_{mild}+N_{full}+N_{block}$ 。

处于不同状态的接收端数目和总数的比例决策如下:

```

If (Nblock/Nsum > Kdecrease)
    Decision = Decrease;
Else if (Nfull/Nsum > Khold)
    Decision = Hold;
Else
    Decision = Increase;

```

其中, $K_{decrease}$ 和 K_{hold} 为降低发送帧率和保持原状态的 2 个阈值,通常情况分别取值为 10%较为适宜。该算法中带宽降低决策优先级最高,仅当处于轻载状态的数目比例大于 80%时才增加发送帧率。

3.3 发送码流的调整算法

当多媒体源接收到接收方发来的 RTCP 报文时,可以根据报文中包含的发送方的报文丢失率调节 RTP 流的发送速率。当报文的丢失率超过了一定的阈值,可以降低报文的发送速率。如果丢失率没有超过阈值,可以增加报文的发送率。调整策略一般采取缓升快降的方法:如果收到一个降低消息时,使用乘法因子 a 迅速降低帧的发送率;收到提高的消息时,使用加法因子 b 增大帧的发送率。

缓升快降算法的表达式为

$$X(n+1)=a*X(n) \quad t(n) > K_2 \text{ 或 } X(n)+b \quad t(n) \leq K_1 \quad (0 \leq a < 1, b \geq 0) \quad (5)$$

其中,参数 a 是降低因子,参数 b 是提高常量, $x(n+1)$ 和 $x(n)$ 分别是 $n+1$ 和 n 时刻的 RTP 流的速度, $t(n)$ 是单位时间内报文的丢失率,阈值 k_1 和阈值 k_2 是报文丢失率的门限,反映用户对报文丢失率的容忍程度。

采用上述算法可以快速地调整发送速率以适应网络变化,但对于利用 RTCP 作为反馈的 RTP,需要使用变量增长的方法。即式(5)中的参数 a 不再是常量,而是一个变量。

设 X 是上一次发生拥塞时的 RTP 流的速率, R 是常增量, K 是判决恢复过是否足够接近上一次发生拥塞的速度 X 的阈值, 是一个相对量。当网络发生拥塞时, 无论 RTP 流的当前的速度如何, 都会将其速度按式(5)的算法减小。在网络中没有拥塞发生时, 参数 a 按如下逻辑变化:

如果网络中未曾发生过拥塞, 即 RTP 流第一次进入网络时, 则

$$a=R \quad (6)$$

如果网络中曾发生过拥塞, 则有:

$$a=R \quad (7)$$

(1) 如果 $X(n)+R < X$ 且网络中没有拥塞, 则

$$a=(X-X(n))/2 \quad (8)$$

(2) 如果 $X(n)+R \geq X$, $X-X(n) > K$, 且网络中没有拥塞, 则

$$a=R \quad (9)$$

本算法的基本思想是利用网络发生拥塞时的信息来控制增大发送速率的幅度。当网络发生时, 该算法记录此时的 RTP 流的速度, 作为上一次拥塞速度 X 。同时 RTP 流的速度减少按式(5)进行。此时的 RTP 流的速度如果满足式(8), 即 RTP 流还未足够接近上次发生拥塞时的速率; 若 RTP 流的速度满足式(9), 表明 RTP 流充分接近上次发生拥塞时的速率, 而网络仍能正常工作, 说明上次网络拥塞是临时的。

4 基于 RTP 协议的自适应的网络传输模型

宽带视频的实时性传输需要一定的网络带宽。尤其当有多个用户同时访问同一视频发布服务器的时, 容易产生并发拥塞和性能瓶颈。这时输出码流必须动态调整以适应网络状况的变化。图 3 为基于带宽自适应的网络传输模型, 在发送端和接收端之间构造多个闭环的负反馈系统, 发送端通过实时的探测网络传输状况, 根据相关的控制策略实时的调整输出码率, 从而实现了自适应传输。

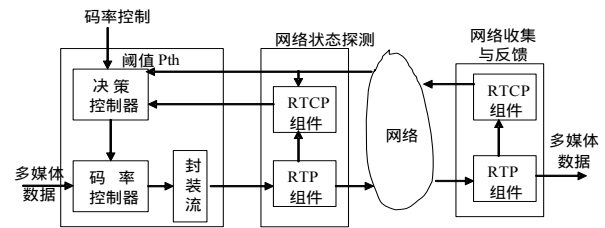


图 3 基于带宽的网络自适应传输模型

码率控制是基于网络可用带宽确定视频通信发送码率的技术。在系统自适应网络传输过程中, 接收端不断检测码流传输质量因素, 并最终反馈给压缩编码端 RTCP 组件; RTCP 包中含有已发送的数据包的数量、丢失的数据包的数量等统计资料, 经过处理后将信息交到决策控制器; 决策控制器根据相应策略控制码率控制部分各组件, 完成该发送码流的动态整合, 并交发送端 RTP 组件完成多媒体流的传输; 接收端接收多媒体数据并交后续的网络播放器进行解码播放。

5 结论

本文的多媒体流自适应网络传输控制模型是基于远程视频监控系统的应而提出的。当网络带宽和信道质量发生变化时, 这种策略能够自动适应网络信道的状况, 保证视频传输的服务质量, 一定程度上解决了网络负载能力受限和拥塞突变的状态。

参考文献

- 1 丁贵广, 郭宝龙. 多媒体数据压缩标准化的现状与发展[J]. 计算机工程与应用, 2002, 38(1): 104-107.
- 2 黎洪松. 数字视频技术及其应用[M]. 北京: 清华大学出版社, 1997.
- 3 罗明宇, 卢锡城, 韩亚欣. Internet 多媒体实时传输技术[J]. 计算机工程与应用, 2000, 36(9): 119-120.

(上接第 90 页)

一个综合的服务结果, 把天气情况、交通情况、餐饮住宿费用情况等信息向用户提供, 帮助用户制订旅游计划。整个旅游服务模型将异构旅游服务资源进行了高效整合, 最终向用户提供旅游综合服务, 从而验证了多主体系统服务模型的可行性和有效性。

4 结论

多主体系统已经在多个领域得到研究和应用, 面向服务的软件工程思想也被人们逐步所重视。本文结合多主体系统理论和面向服务的软件设计思想, 设计出了一个面向服务的多主体系统, 在多主体系统 MAGE 上进行实现, 并在城市应急联动与社会综合服务系统中得到了应用。下一步的主要工作是: (1) 加强服务描述模块, 继续完善服务描述语言 SDLSIN, 使其描述功能更加强大; (2) 研究主体服务协商问题和协作问题, 让面向服务的多 Agent 系统更加智能化; (3) 考虑面向服务的多主体环境的安全性问题。

参考文献

- 1 Wooldridge M. 多 Agent 系统引论[M]. 石纯一译. 北京: 电子工业出版社, 2003.
- 2 蒋运承. 基于主体的智能 Web 中的服务研究[D]. 北京: 中国科学院研究生院, 2004.
- 3 Shi Zhongzhi, Dong Mingkai, Zhang Haijun, et al. Agent-based Grid Computing[C]. Proc. of International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Wuxi, 2002: 164-169.
- 4 史忠植. 智能主体及其应用[M]. 北京: 科学出版社, 2000-12.
- 5 The OWL Services Coalition. OWL-S: Semantic Markup for Web Services, OWL-S 1.0 Release[EB/OL]. 2003. (<http://www.daml.org/services/owl-s/1.0/>).
- 6 Jiang Yuncheng, Shi Zhongzhi, Zhang Haijun, et al. Dynamic Service Matchmaking in Intelligent Web[J]. Journal of Web Engineering, 2004, 2(3): 131-147.