

多媒体流间同步参数计算

The Calculation of Parameters for Multimedia Inter-Stream Synchronization

李宇辉 李吉桂

(华南师范大学计算机科学系 广州 510631)

Abstract Proper configuration of boundary buffer for multimedia objects(MOs) is a vital means to satisfy the requirement of QOP in real-time multimedia communication. This paper analyzes the time sequence of MOs transmission and the mathematical relationship of QOP parameters in master/slave media inter-stream synchronization, gives the minimum and maximum number of buffers for pre-fetching MOs. The calculation result has been verified in a practical multimedia communication system.

Keywords Multimedia synchronization, Quality of presentation, Boundary buffer

1 引言

在分布式多媒体同步控制中,基于有界缓冲区(Boundary Buffer)的同步方式是其中一种主要的强制同步方式^[1]。为同步源设置有界缓冲区,将有助于减轻或消除因网络抖动或其他影响所带来的不连续现象,这是保证流内各个媒体单元的实时和连续的重要手段。对于流间同步,设置有界缓冲区的方式也是十分必要的,如快进或重播某些媒体单元等同步操作,都需要缓冲区的支持。

人们在为同步设置缓冲单元时,往往只是凭经验来定出合适的媒体缓冲单元数,效果不甚理想^[2],难以满足一定的媒体表现质量(QOP: Quality Of Presentation)要求。为此,如何分析分布式多媒体通信中相关参数(如媒体缓冲区大小、最大抖动、传输速率等)对媒体 QOP 的影响,并据此得出这些参数之间的数学关系及量化结果,一些文献对此作了探讨^[2-4]。文[3]运用受限分块法计算出在 ATM 网络中的视频流流内同步所需要的视频缓冲单元数,同时用分块法计算出相同条件下保持音频流流内同步的音频缓冲单元数。文[3]中的计算并不采纳“预取”(Pre-Fetching)方法(即接收端在缓冲了一定数量的媒体单元后才开始播放),因此在不同流中出现不同延迟时,将难以控制流间同步。事实上采取“预取”的方法有利于平滑刚切入接收时的突发媒体流,并保持开始点的同步。文[2]采取预取的方法,计算了在网络延迟有界并且分组按序到达(如虚连接)的情况下,为了保证多媒体表现的连续性,目标结点应至少和最多设置的媒体缓冲单元数。文[2]中计算的是为保证流内媒体同步所需设置的媒体缓冲区大小。文[2]的计算结果给实时多媒体通信系统中的同步

控制设计提供了可供参考的依据。

本文将在文[2]的基础上继续探讨在理想的网络状态下(只考虑网络延迟,且网络延迟有界,并保证包的按序到达,和包出错或丢失概率很小),为保证流间同步要求,各媒体流所应设置的媒体缓冲区的大小。

2 流内同步的有界缓冲单元参数计算

下面给出文[2]的计算结果(定理 1 和定理 2),为了便于说明问题,本文沿用文[2]中的术语或计算参数约定:

同步单位时间(T):决定多媒体间同步的最小单位,根据参考点同步模型的思想,同步单位时间的大小应以某一基媒体的采集或播放频率为参考点。如:若视频每秒 25 帧,则 T 为 1/25 秒。

多媒体对象(MO):又称同步粒度,是一个同步单位时间 T 内所包含的各种媒体信息及其相互约束关系的集合。多媒体对象可以理解为媒体单元和其他相关信息的集合,如在一帧视频的播放时间 T 内,一个多媒体对象也就是与同步有关的一帧的视频、音频、文本等信息。

Dmin:在网上传输一个多媒体对象的最小延迟时间;

Dmax:在网上传输一个多媒体对象的最大延迟时间;

0:多媒体对象的表现周期,即一个同步单位时间;

Bp:在多媒体表现开始前应至少预置的多媒体对象数;

Bc:需要开设的最大缓冲单元数。

[Dmin, Dmax]:定义了网络传输的随机延迟。

定理 1 为了消除因多媒体对象传输与回放速度

不匹配时可能造成的表现不连续,目标结点必须至少预取 $Bp = \left\lceil \frac{D_{\max} - D_{\min}}{\theta} \right\rceil$ 个多媒体对象。

证明略。

定理 2 为了保证多媒体表现的连续性,目标结点最多只需设置 $Bc = \left\lceil \frac{2(D_{\max} - D_{\min})}{\theta} \right\rceil$ 个多媒体对象的缓冲单元。

证明略。

3 流间同步的有界缓冲单元参数计算

文[4]只给出了流内同步时的最大和最小的媒体对象的缓冲单元数。本文将在文[2]的基础上继续求解在相同假设条件下流间同步时各流所需设置的最大和最小媒体对象的缓冲单元数。

Y. Ishibashi 和 S. Tasaka 将媒体流划分成主、从媒体流,通过从媒体流始终和主媒体流保持同步,实现媒体流的同步^[1]。为方便起见,本文的流间同步计算采取主、从媒体流划分,同时为了便于求解,只考虑双流的情况,即只有一个主媒体流和一个从媒体流(只考虑双流的情况已能满足大部分应用的需要,如只需要视频和音频同步的情况)。另外,假设只在目标结点处进行媒体流间同步,源结点处不作同步控制,因此,源结点只要一产生媒体流,就将其发送,而目标结点则需要设置一定的缓冲区对媒体流进行预取。

问题描述:由定理 1 和 2,已知主媒体流 MS 为了维持连续的媒体表现,需要预设至少 MBp 个和最多 MBc 个多媒体对象的缓冲单元,求为了保证从媒体流 SS 媒体表现的连续性和保证从媒体流 SS 与主媒体流 MS 在 MS 开始播放时的同步,从媒体流 SS 需要至少预设多少个和最多预设多少个多媒体对象的缓冲单元?

主媒体流在整个同步过程中是独立的,它只需要维持其自身的连续性,而从媒体流则需要在同步过程中进行调整,以维持与主媒体流的相对时序关系,因此,主媒体流所需要设置的多媒体对象的缓冲单元数,可由定理 1 和定理 2 确定。对于从媒体流,一方面,为了保证它本身播放的连续性,它也必须预设相应的缓冲区(由定理 1 和定理 2 确定),另一方面,由于主媒体流预先填充满了若干个缓冲单元后,才开始播放,因此对于从媒体流来说,它应该有足够的缓冲来保存在主媒体流播放之前到达的多媒体对象,否则将丢失部分媒体单元,造成一开始就不能同步。

为了定量地讨论这个问题,我们先定义以下参数:

θ_m :主媒体流 MS 多媒体对象的表现周期,即一个同步单位时间;

θ_s :从媒体流 SS 多媒体对象的表现周期,即一个

同步单位时间;

根据以上所设,我们将从媒体流的设置规则归纳为下述定理 3 和定理 4。

定理 3 若目标结点中的主媒体流 MS 的预取多媒体对象的单元数为 p,则为了保证从媒体流 SS 媒体表现的连续性和保证从媒体流 SS 与主媒体流 MS 在 MS 开始播放时的同步,从媒体流 SS 需要至少预设

$$B_{\min} = \text{MAX}$$

$$\left\{ \left\lceil \frac{D_{\max} - D_{\min}}{\theta_s} \right\rceil, \left\lceil \frac{(p-1)\theta_m}{\theta_s} + 1 \right\rceil \right\}$$

个多媒体对象的缓冲单元。

证明:先求解为了保证从媒体流 SS 与主媒体流 MS 的开始播放时的同步,从媒体流 SS 至少应预设的多媒体对象的缓冲单元数 $B_{\min}$ 。

由于主媒体流 MS 在第 p 个多媒体对象到达目标结点后,MS 开始播放预存的第 1 个多媒体对象,设 t_p 为 MS 的第 p 个多媒体对象在源结点处的发送时刻,则在 t_p 之前,源结点处已至少产生 q 个 SS 多媒体对象,并已发送或正要发送。q 由下式计算得出:

$$q = \left\lfloor \frac{t_p}{\theta_s} \right\rfloor + 1 = \left\lfloor \frac{(p-1)\theta_m}{\theta_s} \right\rfloor + 1 \quad (1)$$

设 MS 的第 p 个多媒体对象的最早到达时间(p 经历了 $D_{\min}$ 之后到达的时间)为 a'_p ;设 SS 的第 q 个多媒体对象的最迟到达时间(q 经历了 $D_{\max}$ 之后到达的时间)为 a'_q 。其中:

$$a'_p = t_p + D_{\min} = (p-1)\theta_m + D_{\min} \quad (2)$$

则 $B_{\min}$ 将在 $a'_p \leq a'_q$ 的条件下取得。其时间关系由图 1 所示。

在 $a'_p \leq a'_q$ 的情况下,从媒体流 SS 中所有在 a'_p 时刻之前到达目标结点的多媒体对象都必须被缓存,若存在 SS 的多媒体对象 x ($x < q$),它经过最小延迟 $D_{\min}$ 到达目标结点,且其到达目标结点的时间比 a'_p 早并且它的下一个多媒体对象到达目标结点的时间比 a'_p 迟,则所求 x 的值就是 $B_{\min}$ 。

令 a'_x 为 x 的最早到达时间,则有:

$$a'_x = (x-1)\theta_s + D_{\min} \quad (3)$$

由上述讨论,有

$$a'_x \leq a'_p$$

将式(2)、(3)代入上式,整理后有:

$$x \leq \frac{(p-1)\theta_m}{\theta_s} + 1$$

由于 x 为正整数,故 x 的最大值

$$B_{\min} = x_{\max} = \left\lfloor \frac{(p-1)\theta_m}{\theta_s} + 1 \right\rfloor \quad (4)$$

综合定理 1 和式(4),为了保证从媒体流 SS 媒体表现的连续性和保证从媒体流 SS 与主媒体流 MS 在开始播放时的同步,从媒体流 SS 需要至少预设的多媒体对象的缓冲单元数为:

$$BS_{\min} = \text{MAX} \left\{ \left\lceil \frac{D_{\max} - D_{\min}}{\theta_s} \right\rceil, \left\lfloor \frac{(p-1)\theta_m}{\theta_s} + 1 \right\rfloor \right\}$$

证毕。

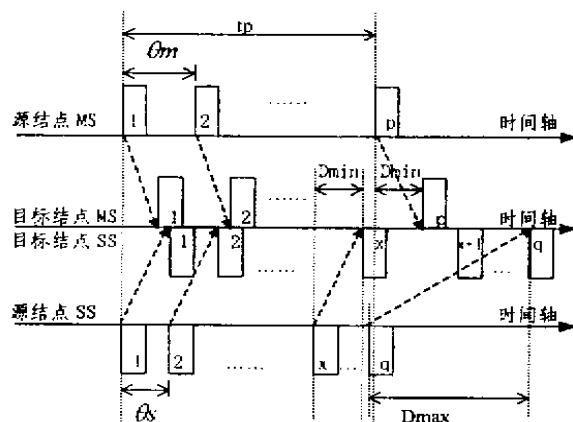


图1 SS获取最小缓冲单元数的情况

定理4 若目标结点中的主媒体流MS的预取多媒体对象的单元数为 p ,则为了保证从媒体流SS媒体表现的连续性和保证从媒体流SS与主媒体流MS在MS开始播放时的同步,从媒体流SS最多只需设置

$$BS_{\max} = \text{MAX} \left\{ \left\lceil \frac{2(D_{\max} - D_{\min})}{\theta_s} \right\rceil, \left\lfloor \frac{(p-1)\theta_m + D_{\max} - D_{\min}}{\theta_s} + 1 \right\rfloor \right\}$$

个多媒体对象的缓冲单元。

证明:先求解为了保证从媒体流SS与主媒体流MS在开始播放时的同步,从媒体流SS最多只需设置的多媒体对象的缓冲单元数 $B_{\max}$ 。

根据定理3的分析,在MS的多媒体对象 p 发送的时刻 t_p 之前,源结点处已至少产生 q 个SS多媒体对象,并已发送或正要发送。 q 由(1)式给出。

令 a'_p 为多媒体对象 p 的最迟到达目标结点的时间,且有

$$a'_p = t_p + D_{\max} = (p-1)\theta_m + D_{\max} \quad (5)$$

令 a'_q 为多媒体对象 q 的最早到达目标结点的时间,且有

$$a'_q = t_q + D_{\min} = (q-1)\theta_s + D_{\min} \quad (6)$$

其中 t_q 为 q 在源结点处发送的时刻。

当 p 以最迟到达时间到达目标结点, q 以最早到达时间到达目标结点,且满足 $a'_q \leq a'_p$ 时, $B_{\max}$ 将可在此条件下取得,其时间关系由图2所示。

设SS有 k 个多媒体对象在 a'_q 与 a'_p 之间到达,若存在第 $(q+k)$ 个多媒体对象是以最早时间到达目标结点,则有 $B_{\max} = q+k$ 。

令 a'_x 为SS中第 x ($x=q+k$)个多媒体对象的最早到达目标结点的时间,则有

$$a'_x = t_x + D_{\min} = (x-1)\theta_s + D_{\min} \quad (7)$$

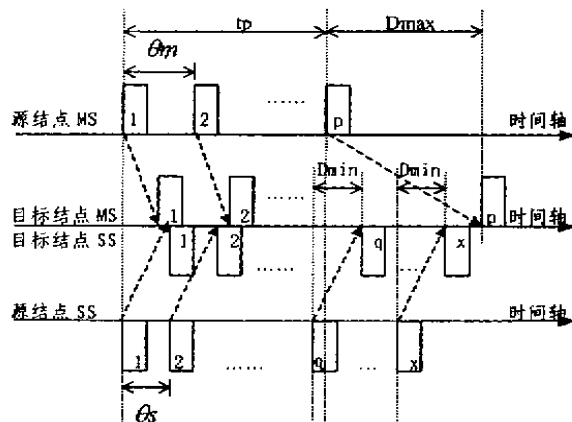


图2 SS获取最大缓冲单元数的情况

其中 t_x 为 x 在源结点处发送的时刻。此时有

$$a'_x \leq a'_p$$

将式(5)、(7)代入上式,整理后有:

$$x \leq \frac{(p-1)\theta_m + D_{\max} - D_{\min}}{\theta_s} + 1$$

由于 x 为正整数,故 x 的最大值

$$B_{\max} = x_{\max} = \left\lfloor \frac{(p-1)\theta_m + D_{\max} - D_{\min}}{\theta_s} + 1 \right\rfloor \quad (8)$$

综合定理2和式(8),为了保证从媒体流SS媒体表现的连续性和保证从媒体流SS与主媒体流MS在开始播放时的同步,从媒体流SS最多需要预设的多媒体对象的缓冲单元数为:

$$BS_{\max} = \text{MAX}$$

$$\left\{ \left\lceil \frac{2(D_{\max} - D_{\min})}{\theta_s} \right\rceil, \left\lfloor \frac{(p-1)\theta_m + D_{\max} - D_{\min}}{\theta_s} + 1 \right\rfloor \right\}$$

证毕。

4 应用实例

为了进一步研究上述定理3和定理4的有关结果,我们设计了以下应用实例。

在多媒体通信系统里,声音往往具有特殊的地位,它所承载的信息一般地比视频较为重要。因此,我们在一个基于IP组播的实时多媒体通信系统^[5]的设计与实现中,把客户端的音频作为主媒体流,而视频作为从媒体流。为了保证媒体播放的连续性和视频与音频的同步,依据以下的已知参数,根据上述定理1、2、3和4,分别计算音频和视频所需要预取的最大和最小缓冲单元数。

对于多媒体会议系统等视频应用来说,ITU标准中建议的最大端到端的总延迟不要超过150ms^[1],所

(下转第68页)

```

BookNo=BookNo+1
Form1.Text1.Text=Form1.Text1.Text & "Book" &
    BookNo & "：" & targetNode.Text & vbCrLf
End If

```

Next nodesize

2.3 SAX 和 DOM 的优缺点对比

	SAX	DOM
代	能轻松处理拥有上百万个节点的大文档,且内存的消耗不会随文档增大而增多,能显著节约内存和提升处理效率	能执行强大的 XSLT 变换,用多种不同类型的风格样式单为同一个文档创建出多个视图
	能随时根据找到信息的情况终止文档解析,从而避免了无谓的遍历文档	DOM 模型的树形结构将自动维护包含文档上下文信息的复杂数据结构,能执行复杂的 Xpath 过滤功能
点	SAX 可用高层的对象,如股票代码,新闻等从几个大文档中非常简单快速地创建出新文档,采用 DOM 则只能利用底层的元素,属性和处理指令来构造新文档;	在内存中建立文档的树形结构,能随机访问数据,并能随意修改文档内容并保存修改
	SAX 是属于底层的 API,给程序员更多的控制权	DOM 属于高层的 API,编程方便,得到更广泛的支持
缺	不能保存文档的上下文信息;不能随机访问文档结构;不能实现复杂的查询;目前缺乏浏览器的支持	不能适应大文档的处理;占用系统资源多;处理速度不如 SAX 快;不能终止解析过程

3 DOM 和 SAX 的结合使用的设想

我们可以利用给 SAX 体系结构添加索引信息来绕过 DOM 模型的尺寸限制,这对于那些内容并不经常更新的大 XML 文档进行只读操作来说是很好的选择。具体做法在 SAX 解析器顺序读 XML 文档时利用一个索引程序定位每一个元素的开始和结束位置来建立一个有关元素位置的索引表。当执行查询时,标记名在 XML 文档中的位置是通过文件定位指令(如 Seek)在标记位置索引表中检索来得到的,由于 XML 文档的结构完整性,所以一旦给定的元素及其内容被检索

出来,那么在查询中的子元素也肯定在检索出来的信息的子集中,通过选择性的抽取元素,就能够最终检索到任何需要节点,而且还可以将检索到的节点位置信息添加到节点位置索引表中以便于后继再次查询时提高效率。

参考文献

- 1 孙一中,等. XML 理论和应用基础. 北京:北京邮电大学出版社,2000
- 2 Goldfarb C F,等. XML 实用技术. 北京:清华大学出版社,1999
- 3 Available at: <http://www.microsoft.com/xml/>
- 4 Available at: <http://www.w3.org/TR/REC-xml>
- 5 Available at: <http://www.meggison.com/SAX>

(上接第52页)

以 $D_{max}=150ms$, 在速率为 $10Mbit/s$ 的以太网上,传输一个多媒体单元(一帧视频图像或一个音频片段,本例中其最大长度设为 $4kB$)的最小延迟时间 D_{min} 为:

D_{min} = 一个多媒体单元的大小/网络的最大传输速率

$$D_{min}=4kB/10Mbit/s=3.2ms$$

在本例中,音频片段的采集间隔 $\theta_m=1000ms$,视频帧速为 15 帧/秒,因此 $\theta_s=66.7ms$ 。

音频需要预设的最小和最大的缓冲单元个数分别为:

$$AB_{min}=\lceil (150ms-3.2ms)/1000ms \rceil=1$$

$$AB_{max}=\lceil 2 \times (150ms-3.2ms)/1000ms \rceil=1$$

视频需要预设的最小和最大的缓冲单元个数分别为:

$$VB_{min}=\text{MAX}\{\lceil (150ms-3.2ms)/66.7ms \rceil, \lceil (1-1) \times 1000ms/66.7ms + 1 \rceil\}=3$$

$$VB_{max}=\text{MAX}\{\lceil 2 \times (150ms-3.2ms)/66.7ms \rceil, \lceil (1-1) \times 1000ms + 150ms - 3.2ms / 66.7ms + 1 \rceil\}=5$$

根据计算结果,我们将客户端音频的接收缓冲单

元数设为 1 ,而视频的接收缓冲单元数设为 5 。实验结果表明,系统能够很好地实现多媒体的实时连续性通信并且音视频达到很好的同步,满足 QCIF 视频的表现质量要求。

小结 在分布式多媒体通信系统里,适当地为同步源设置多媒体对象缓冲区,是保证多媒体通信的实时性和连续性,以及实施流间媒体同步控制的必要手段之一,本文对同步时所需要的有界缓冲区大小的计算结果,尽管是在较为理想的通信条件下得出,但它仍然为实际的多媒体通信系统设计提供了可供参考的同步参数之间的量化依据。实际应用表明,运用本文的参数计算结果,能够很好地利用系统资源,并有效地保证了多媒体通信中的媒体表现质量要求。

参考文献

- 1 胡晓峰,李国辉. 多媒体系统. 人民邮电出版社,1997
- 2 郑庆华,等. 分布式多媒体同步中表现质量的参数计算. 通信学报,1999,10(20)
- 3 张明德,王永东. 视频会议系统原理与应用. 北京:希望电子出版社,1999
- 4 魏铁军,陈俊亮. 基于漏窗机制实时自适应多媒体同步. 通信学报,1999,5(20)
- 5 李宇辉. 基于 IP 组播的实时媒体传输研究与实现. [学位论文]. 广州:华南师范大学,2000