

文章编号: 0529-6579 (2000) 02-0016-03

提高桌面会议系统中音频质量的方法研究*

汤 庸¹, 彭重嘉², 印 鉴³, 李显济², 李 松²

(1. 中国科技大学研究生院计算机系, 北京 100039;

2. 广东工业大学计算机系; 3. 中山大学计算机系)

摘 要: 讨论影响视频会议系统音频质量的主要原因, 提出一种动态设置缓冲区改善音质的方法, 并建立 Petri 网络模型分析其可行性.

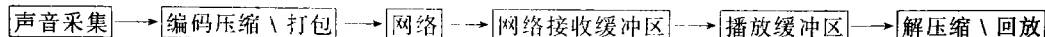
关键词: 音频; 服务质量; 桌面会议系统; 计算机支持的协同工作

中图分类号: TP393 **文献标识码:** A

视频会议是多媒体技术在网络上的一个重要应用. 在视频会议中, 对于实时的交互, 人们较为容忍视频信号的不连续, 对于控制信号, 250 ms 的延迟都可接受. 但对于声音, 几毫秒的延迟抖动都会让人感到不满意^[1,2]. 在因特网上造成延时抖动的主要原因包括: ① 不同的数据包选择不同的路由, 使存储-转发交换延迟以及查找接点所需要的时间不同, 从发送端到接收端的时延不同, 数据包到达接收端时的间隔时间也就不同; ② 由于拥塞, 分组被丢弃而请求重发, 或者因为不同路由器上的队列处理时间不同而产生的时延不同, 也会造成数据包到达接收端的时间间隔不同; ③ 网络传输材料也会引起的微弱的物理抖动, 比如金属导体中传输延迟会随温度的变化而变化, 这些抖动产生的影响很小.

1 增加接收端的缓冲区来改善音质

视频会议通常要求某一时刻只有一个会议参加者有发言权, 而且发言后, 并不希望得到实时的回应, 即声音是单向的. 所以, 增加语音的额外延迟而不影响语音的质量, 在接收端增加一个接收缓冲区, 等接收到的语音包达到一定数量后才开始回放. 即



将缓冲区设定为动态的, 以提高资源的利用率. 演讲者在拥有发言权后, 开始声音采集、发送语音包, 经过网络传输后, 语音包到达接收端的网络接收缓冲区 (根据使用不同的协议, 如 TCP/IP 或 UDP, 会有确认信号返回发送端或无需确认), 再放到播放缓冲区. 当播放缓冲区里的语音包到达一定数量后, 系统从缓冲区中取走语音数据, 送到声卡设备中去解压缩, 并按照数字化语音时的速率和格式回放. 由于网络拥塞的原因, 接收端接收

* 基金项目: 广东省自然科学基金资助项目 (960090)

收稿日期: 1999-07-08 作者简介: 汤 庸 (1964~), 男, 广东工业大学计算机系教授

数据的速度要小于回放时的速度,因此,随着播放的继续,播放缓冲区中的语音包会逐渐减少到零,这时就停止回放,直到缓冲区中的语音包重新达到一定的数量后,再开始回放。这时会增加一个新的播放延时。

另外,为了两端时间匹配,需要采用静音技术。即当发送端演讲者停顿时,不传输语音信号,这段时期叫静默期^[2],这样语音流就不是连续的,而是由静默期分开的一些突发期。这样可使接收端和发送端的时钟失配在静默期得到消除。接收端在回放语音信息时,由于网络传输造成的时延抖动,有可能会使播放缓冲区中的数据不够而产生播放间隙。如果播放缓冲区的参数设置得恰当,可以使发送端的静默期与接收端的播放间隙匹配,减少播放间隙对音质的影响(图1)。在 t_1 时刻,发送端开始发送语音包,经过网络传输,在 t_2 时刻,该包到达网络接收缓冲区, $t_2 - t_1$ 表示第1个语音包的传输延迟。并且该语音包进入播放缓冲区存储起来,在 t_3 时刻,开始从播放缓冲区中播放第1个语音包。在图1中“静默期”所示的位置为演讲者有一个停顿。在时刻 t_4 时,由于播放缓冲区中已经无语音包,因此,要经过一段时间,才又开始播放。从图1中还可看到,到达网络接收缓冲区的语音包的时间间隔不同,表示网络的时延抖动。但是,可以看出在有语音包时,回放的速度是同语音采集的速度相一致的。图中所示是一种理想的情况,即播放缓冲区中无数据时,产生的播放间隙刚好与发送端的静默期相匹配。

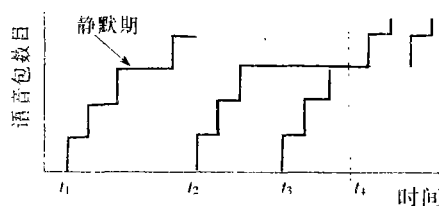


图1 语音包与时间

Fig.1 Audio and time

2 模型的 Petri 网描述与分析

本文用有色时间 Petri 网来描述该系统模型并分析它的运行情况,图2是该系统采用 UDP 协议的 Petri 网模型。在该模型中,在发言期间,只有语音数据,并且是单向流动的。

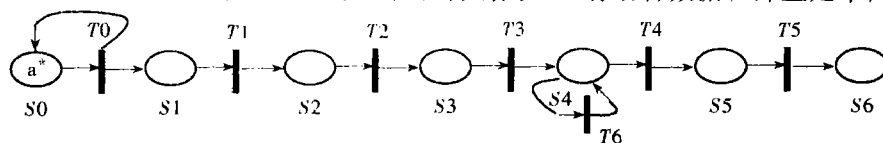


图2 Petri 网模型

Fig.2 Petri net model

模型中有3种不同属性的托肯:一种是小圆点 a ,表示语音包,它的变迁可解释为语音包在网络中的传输和回放。另2种是元素 f 和 t ,用于控制变迁 T_4 的发射。

在图2中, $T_0 \sim T_5$ 表示变迁,只对小圆点托肯有效,权函数为1。 T_6 是对有色元素 $\{f, t\}$ 的变迁,变迁函数为 $g(f) = t$,该变迁有效的条件是库所 S_4 中的元素大于值 M_0 ; $g(t) = f$,该变迁有效的条件是库所 S_4 中的元素等于0。各个库所含义为: S_0 表示声音采集,每个元素表示一个语音包,容量为1。 S_1 表示对声音编码及打包,容量为1,在该库所的停留时间为语音包处理时间。 S_2 表示网络,在该库所的停留时间 $time$ 为近似于正态分布的随机变量,容量可设为无限。 S_3 表示网络接收缓冲区,容量为1,在该库所的停留时间为语音包在该缓冲区的处理时间。 S_4 表示播放缓冲区,容量为 M ,变迁有效的条件是库所中的有色托肯 t 存

在时, S_5 表示回放处理, 容量为 1, 元素在该库所停留的时间为 t_0 , 在 t_0 的时间内回放一个语音包. S_6 表示回放, 容量不限.

该模型的初始标识为: 库所 S_0 中有一个托肯 a , 库所 S_4 中有一个托肯 f . 模型的运行情况如下: 当演讲者得到发言权后, 该模型建立. 在初始状态下, 变迁 T_1 每隔一个随机时间 t_i 有效, 因为对于以太网或令牌环网来说, 一个站点得到发送数据包的权力的时间是随机的. 变迁 T_1 发射时, 一个语音包进入库所 S_1 , 同时又回到库所 S_0 , 相当于产生了一个新的语音包. 托肯 a 经过 T_0, T_1, T_2, T_3 一系列的与时间相关的变迁后到达库所 S_4 . 初始时, 库所 S_4 中有一个有色托肯 f . 当库所 S_4 中的托肯 a 的数目达到 M_0 时, 变迁 T_6 发生, 有色托肯 f 变为 t , 这时变迁 T_4 有效. 托肯 a 进入 S_5 , 在其中停留时间 t_0 后进入库所 S_6 , 可解释为回放了一个语音包. 根据 Petri 网的变迁有效的基本原则, 一个变迁有效还必须满足条件: 输出库所的容量必须足够容纳输入的托肯. 因此, 在上述模型中, 变迁 T_4 发生的一个必要条件是 S_5 中的容量为 0. 另外, 如果 S_4 中的托肯数减少为 0, 则变迁 T_6 发生, 有色托肯 t 变为 f . 相当于回到初始状态下的情形.

通过分析可知, 无论各个库所的停留时间为多少, 系统都可以运行, 所以该系统不会进入死锁状态. 但是, 模型还会有这样的问题, 在库所 S_4 中的有色托肯为 f , 且库所 S_0 产生的托肯数目不足 M_0 时, 这些托肯将在库所 S_4 里不再发生变迁. 它的实际意义就是这些小于 M_0 的语音包不能被回放. 为解决这个问题, 可以给变迁 T_4 增加一个外部事件控制. 即设置一个定时器, 当该定时器在足够长的时间内检测到没有托肯 a 进入 S_4 时, 该外部事件激发变迁 T_4 发生.

本文用 VC++ 6.0 实现了上述模型实验系统, 对在接受端设置一个额外的缓冲区来平滑声音的方法进行了模拟实验. 结果表明该方法在提高声音的平滑度上有很强的效果, 并且得出了缓冲区的大小对声音平滑影响的一般性结论^[3]. 缓冲区设置算法可参见文[3].

参考文献:

- [1] 汤庸, 魏际洲, 李显济, 等. 音频视频协同工作机制及其实现研究[J]. 软件学报, 1998, 9(增刊): 75 ~ 79.
- [2] 何宝宏, 郭玉钗, 林宗楷, 等. 因特网语音工具中音质技术的研究[A]. 史美林, 李红臣. 第一次全国 CSCW 学术会议论文集[C]. 北京: 电子工业出版社, 1998. 199 ~ 203.
- [3] 彭重嘉. 因特网上提高声音质量的研究[D]. 广州: 广东工业大学, 1999.

Research on Quality of Audio in Desktop Conferencing Systems

TANG Yong^{*}, PENG Chong-jia, YIN Jian, LI Xian-ji, LI Song

Abstract: This paper discusses the quality of audio in desktop conferencing systems, proposes a method of dynamic buffers of audio back-play to improve the quality of audio. The method is verified by a Petri net modeling.

Keywords: audio; QoS; desktop conferencing systems; CSCW

^{*} Department of Computer, Guangdong University of Technology, Guangzhou 510090, China