

# 一种新的 VoIP 自适应缓冲算法

陈伟,黎忠文

(厦门大学信息科学与技术学院,福建 厦门 361005)

**摘要:**网络延迟与缓冲的矛盾是 VoIP 应用中的一个重要问题。介绍了 VoIP 应用中几种当前主要的缓冲算法,分别分析了它们的优缺点,提出了新的自适应缓冲算法,称为 FISD 算法,对现有的代表性算法以及 FISD 算法分别进行了仿真实验。结果表明,在网络延迟抖动较大时,新算法可以有效地提高语音质量。

**关键词:**VoIP;网络延迟;抖动;自适应缓冲

**中图分类号:**TP301.6

**文献标识码:**A

**文章编号:**1673-629X(2009)01-0005-04

## A New Algorithm for VoIP Adaptive Buffer

CHEN Wei, LI Zhong-wen

(School of Info. Sci. & Tech., Xiamen University, Xiamen 361005, China)

**Abstract:** The contradiction between network delay and the buffer is an important issue of the application of VoIP. Introduces several major current buffer algorithms in the application of VoIP, and analyses their advantages and disadvantages respectively. Then a new adaptive buffer algorithm is proposed, which is named FISD algorithm. The representation of the existing algorithms and FISD algorithm simulation experiments were carried out. The results show that when the network delay jitter is intense, the new algorithm can effectively improve voice quality.

**Key words:** VoIP; network delay; jitter; adaptive buffer

## 0 引言

VoIP (Voice over Internet Protocol) 是利用 IP 网络传输语音的技术,目前得到广泛的应用。在 VoIP 应用中, QoS (Quality of Service) 是至关重要的。延迟与延迟抖动是 QoS 的两项重要参数。数据包在源端发出,在目的端接收,期间这段时间称为数据包的延迟。由于 IP 网络采用的是无连接尽力服务 (Best Effort) 的工作方式,所以存在分组丢失、失序到达、延迟抖动等情况。传统数据传输业务可以通过等待数据包和重传数据包来解决这些问题,但是语音传输业务的流畅播放要求与实时性要求使得这个方法不能用于 VoIP 应用。

在 VoIP 应用中,解决这些问题的一个重要方法是在接收端设置缓冲区,从而达到吸收网络抖动,降低

丢包率的效果。但是缓冲区大小的设置对 QoS 的影响甚大:缓冲区设置过大会增大语音延迟,影响会话的交互性;缓冲区设置过小则无法吸收延迟的抖动,丢包率上升。

为了更合理地设置缓冲区大小,从而解决缓冲与延迟的矛盾,出现了自适应缓冲算法,所谓自适应缓冲,即控制缓冲区的大小随着网络环境的变化动态地作出变化。国内外的 VoIP 研究者提出了许多不同的自适应缓冲算法。这些算法在不同程度上缓解了缓冲与延迟的矛盾,但是在网络延迟抖动较大时,这些算法往往无法及时适应恶劣的网络环境,使得在这种网络环境下 VoIP 应用难以得到令人满意的效果,对 VoIP 应用造成了很大的局限性。

为了解决在网络延迟抖动剧烈情况下现有自适应算法无法有效发挥作用的问题,文中提出了 FISD 算法,这是一种借鉴了 AIMD 算法思想的新自适应缓冲算法,利用该算法可以使缓冲区对波动较大的网络环境迅速做出反应并做出调整,从而保持 VoIP 通话业务的流畅性能。

文中对现有的代表性缓冲算法以及文中提出的 FISD 算法分别进行了仿真实验。实验结果表明,在网

收稿日期:2008-04-03

基金项目:福建省 2004 年自然科学基金资助项目(A0410004);厦门大学院士基金资助项目(0630-E23011);厦门大学新世纪优秀人才支持基金(0000-X07116)

作者简介:陈伟(1985-),男,福建莆田人,硕士研究生,研究方向为 VoIP 语音质量;黎忠文,博士,教授,CCF 会员,研究方向为实时系统高安全和高可靠技术。

络延迟抖动较大时,新算法的缓冲延时与丢包率综合性能有显著提高,可以有效地提高语音质量。

## 1 当前现有的缓冲算法

目前缓冲算法主要有以下几类(如图 1 所示):

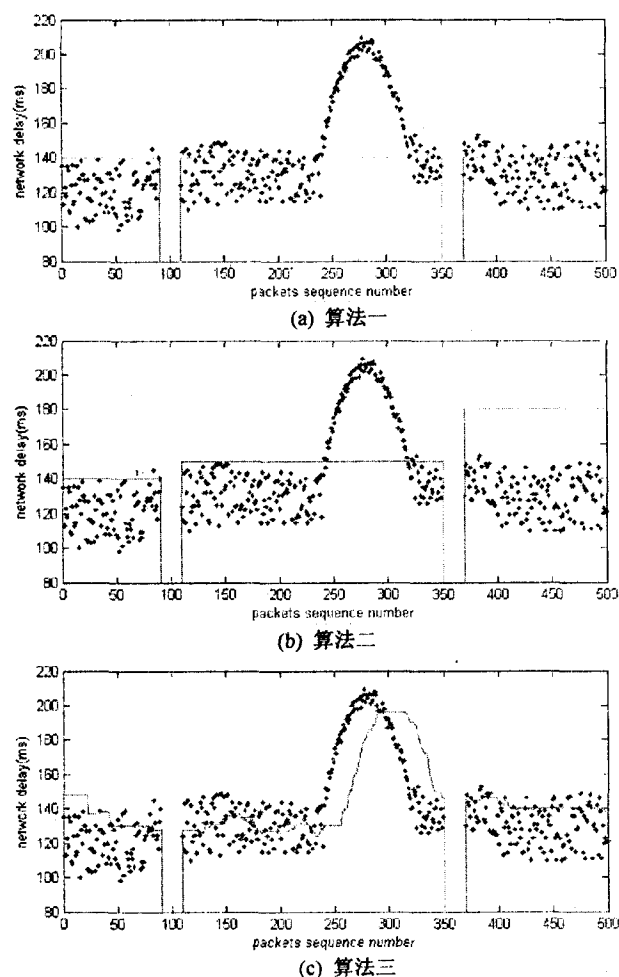


图 1 现有算法

(1)始终采用固定的缓冲,无法根据网络环境进行自适应变化,会话期间所有语音包采用固定的延迟时间,如图 1(a)所示。这种方法实现简单,但是效果很差。

(2)一般来说,语音业务有静默与发言交替出现的特点,故存在通话段(talkspurt)<sup>[1]</sup>。通话段可通过语音活动检测(Voice Activity Detection, VAD)(也称为静音抑制)技术来分隔。第二类缓冲算法<sup>[1,2]</sup>以通话段为单位来调整缓冲大小,即同一个通话段中的所有语音包采用相同的缓冲,而不同通话段之间可调。如图 1(b)所示。这类算法较第一类算法有所改进,但是当通话段较长或网络延迟变化较大时,这类算法的弊端就显现出来。

(3)第三类缓冲算法<sup>[3,4]</sup>是利用语音包播放时间修正技术(如 WSOLA<sup>[5]</sup>)或者改变语音播放速度,在

通话段内部也进行缓冲时间的自适应修改,这种方法更为灵活。如图 1(c)所示。不过目前这类算法多是根据大量历史数据推测后续语音包的延迟情况,当出现数据包延迟突增时,无法及时改变缓冲大小,无法充分发挥这类算法的优势。

## 2 FISD 算法

加性增乘性减(Additive Increase Multiplicative Decrease, AIMD)算法<sup>[6]</sup>是目前 IP 网络传输层广泛应用的拥塞控制算法。该算法的大致思想是:当网络畅通时,加性增大拥塞窗口,从而充分利用网络带宽;当网络阻塞时,乘性减小拥塞窗口,从而迅速让出过分占用的带宽,防止出现网络拥塞。

文中提出的 FISD(Fast Increase Slowly Decrease)算法借鉴了 AIMD 的思想,并结合所解决的问题做了一些修改。

当缓冲区过小时,以较大比例增加缓冲区大小,即以原缓冲区大小乘以一个大于 1 的常数,得到新的缓冲区大小。使之可以尽快达到足够的缓冲区水平。该行为类似于 AIMD 算法中“乘性减小”拥塞窗口的行为,不过此处是以较大比例乘性放大缓冲区。

当缓冲区过大时,以较小比例减小缓冲区大小,在用户无法察觉音频的细微变化的前提下实现对缓冲区大小的调整。该行为与 AIMD 的“加性增加”不完全相同。AIMD 的“加性增加”是以固定步长增加拥塞窗口大小,而本算法是将原缓冲区大小乘以一个略小于 1 的常数,得到新的缓冲区大小。作这个修改的原因是当缓冲区大小被放大过多时可以更快地回到合理水平,而当缓冲区大小放大水平较低时,该方法的效果与“加性”减少缓冲区效果相当。

为了明确地描述 FISD 算法,现定义以下变量:

$S(i)$ :第  $i$  个语音包从源端发送的时间;

$R(i)$ :第  $i$  个语音包到达接收端的时间;

$P(i)$ :第  $i$  个语音包在接收端的播放时间;

$L(i)$ :第  $i$  个语音包的播放时间长度。 $L(i) = P(i+1) - P(i)$ ;

$Dn(i)$ :第  $i$  个语音包由网络传输导致的延迟, $Dn(i) = S(i) - R(i)$ ;

$Db(i)$ :第  $i$  个语音包由接收端缓冲导致的延迟, $Db(i) = R(i) - P(i)$ ;

$Dt(i)$ :第  $i$  个语音包从发送到播放的总延迟, $Dt(i) = Dn(i) + Db(i) = S(i) - P(i)$ ;

$DtPr$ :算法预测的语音包总延迟;

$Lpkt$ :一个语音包包含的语音长度,通常为 20ms;

$Nprp$ :用于初始化的数据包个数,是一个常数;

DnAverage:最近  $N_{prp}$  个数据包的平均延迟;  
 VAverage:最近  $N_{prp}$  个数据包的近似延迟方差;  
 $C\_extend$ :扩大系数。是一个大于1的常数;  
 $C\_shorten$ :缩小系数。是一个小于1的常数。

其中,  $C\_extend * C\_shorten > 1$ 。

FISD算法描述:

1)初始化。

在通话前先收集  $N_{prp}$  个数据包,计算它们的传输延迟总和  $\Sigma Dn$ ,并将  $DnAverage = \Sigma Dn / N_{prp}$  作为  $DtPr$  的初始值。计算  $VAverage$ 。 $i$  赋值为0。

2)接收数据包。

(1)接收第  $i$  个数据包。

(2)根据数据包的发送时间  $S(i)$  推测下一个数据包的发送时间  $S(i+1) = S(i) + L_{pkt}$ ;推测下一个数据包的播放时间  $P(i+1) = S(i+1) + DtPr$ 。

(3) $i = i + 1$ 。

3)循环。

(1)在  $P(i)$  以前等待接收第  $i$  个数据包。若超时仍未接收到,则视为丢包,  $DtPr = DtPr * C\_extend$ ;否则

(a)将数据包作播放长度修正后进入播放队列,并于  $P(i)$  时间播放。

(b)更新  $DnAverage$ 、 $VAverage$  近似值。

(c)若  $DtPr > DnAverage + VAverage$ ,  
 则  $DtPr = \text{Min}(DtPr * C\_shorten, DtPr - 1\text{ms})$ ;

否则  $DtPr$  保持原值。

(2)推测下一个数据包的发送时间  $S(i+1) = S(i) + L_{pkt}$ 。推测下一个数据包的播放时间  $P(i+1) = S(i+1) + DtPr$ 。

(3) $i = i + 1$ 。

4)算法结束。

从 FISD 算法的描述可以看出,算法通过较大幅度修正少量语音包的播放长度,牺牲这部分语音包的语音质量,使得缓冲时间长度能在短时间内增大到足够大的程度,从而对突发的延迟高峰迅速地作出反应。当对延迟高峰出现误判时,可通过对大量语音包的播放长度做轻微缩短使缓冲时间恢复到合理水平。所以该算法在网络数据包延迟波动较小时,与其他算法相比并无特别出色之处;当网络数据包延迟波动大时,可以有效地吸收抖动并保持较佳的语音质量。

实际网络采集的延迟数据,如图2所示,其中,横轴为数据包序号,纵轴为延迟时间。图例如图3所示。实验结果如图4所示,其中,横轴为平均的缓冲延时,纵轴为丢包率。

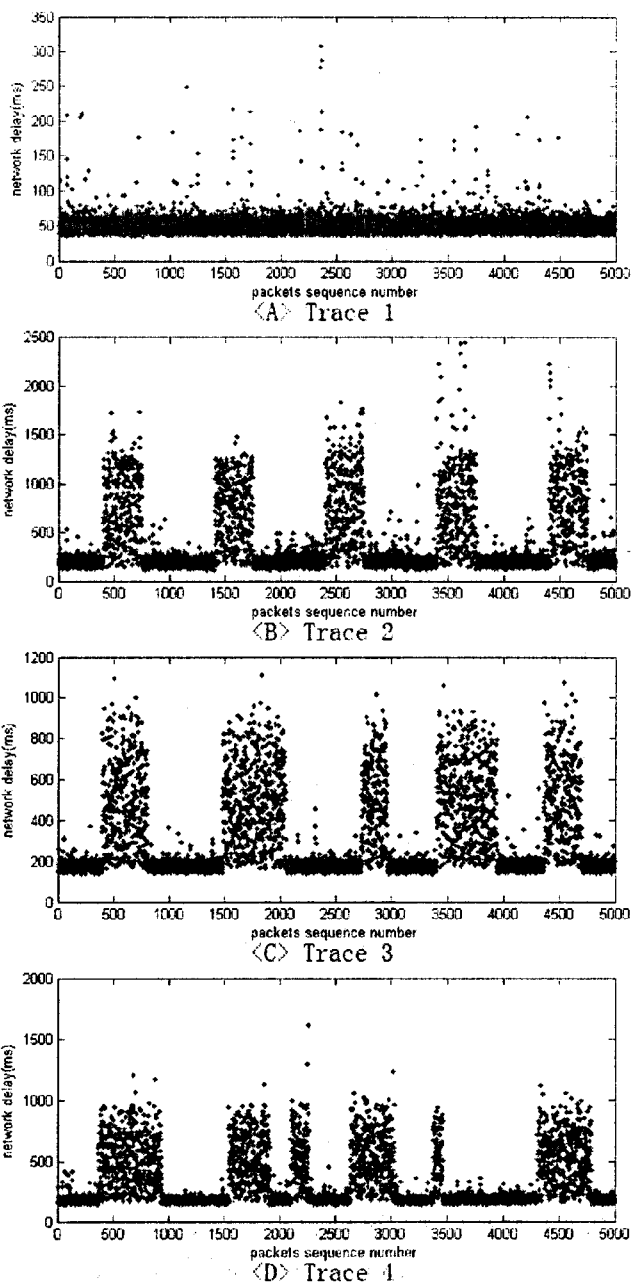


图2 网络延迟数据

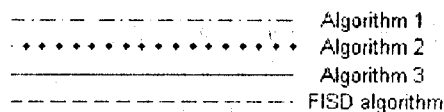


图3 图例

### 3 仿真实验

为了评估这些算法的性能,笔者分别用当前的3种算法和文中提出的算法进行仿真。实验采用4组从

### 4 实验结果分析

从实验结果可以看出,在网络延迟抖动稳定(trace1)时,FISD算法与其他3个算法效果无明显差

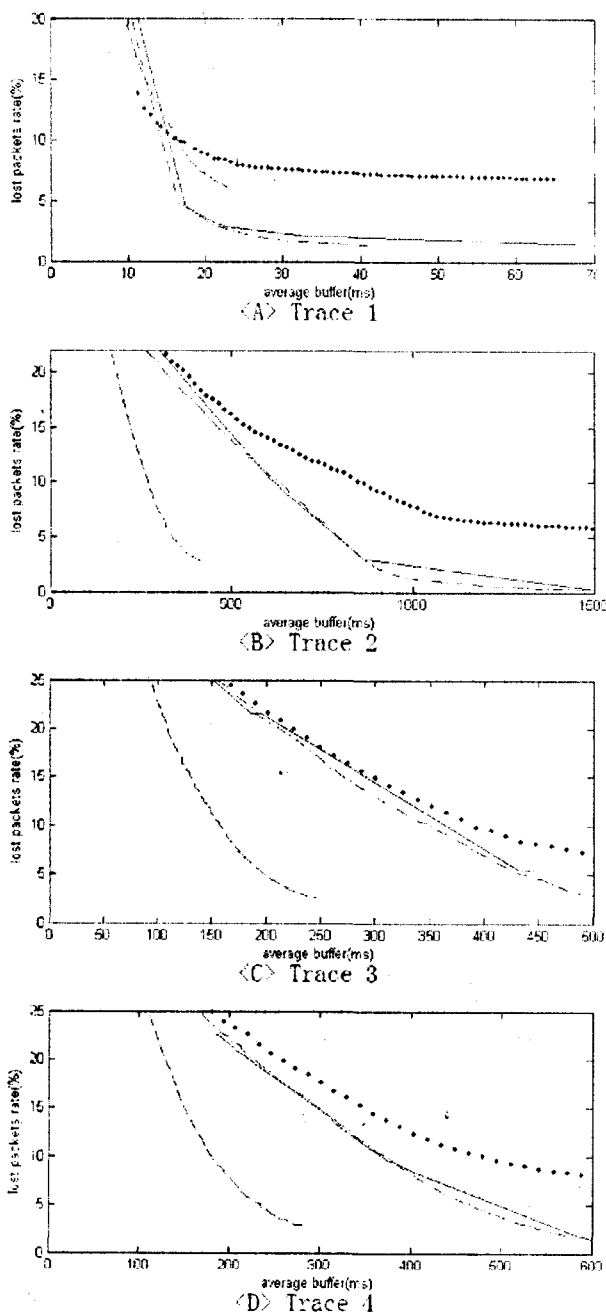


图 4 实验结果

别,在网络延迟抖动较大时(trace2,3,4),与其他3个算法相比,FISD算法在缓冲延时和丢包率的综合性能方面有着明显的优势。以Trace4的结果为例。如图5所示,丢包率为10%时,算法1、2、3的平均缓冲延时分别约为360ms、480ms、360ms,而算法4的平均缓冲延时约为170ms,大大低于其他3个算法。当延时为200ms时,算法4的丢包率为7%至8%,算法1、2、3的丢包率都在20%以上。

## 5 结束语

延迟与缓冲的矛盾是VoIP应用中的一个重要问

题。缓冲算法在一定程度上缓解了二者的矛盾,但是现有的缓冲算法在网络延迟抖动剧烈情况下无法有效发挥作用。文中分析了VoIP应用中当前主要的缓冲算法的优缺点,提出了FISD算法,并对这些算法进行了仿真实验。结果表明,在网络延迟抖动较大时,新算法可以有效地提高语音质量。下一步将研究如何根据网络实际情况得到更合适的扩大系数和缩小系数,以取得更高的服务质量。

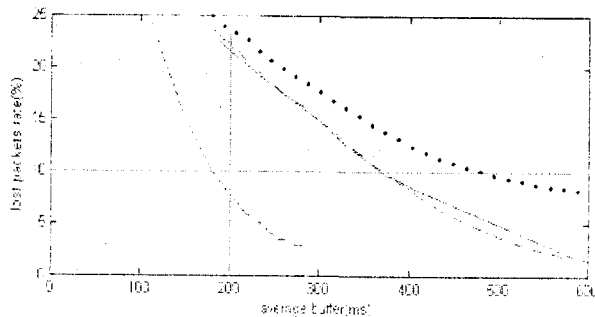


图 5 Trace4 实验结果分析

## 参考文献:

- [1] Pinto J, Christensen K J. An Algorithm for Playout of Packet Voice based on Adaptive Adjustment of Talkspurt Silence Periods[C]//In Proceedings of the 24th Annual IEEE Conference on Local Computer Networks. Lowell, Massachusetts, USA: IEEE Computer Society, 1999: 224 - 231.
- [2] Ramjee R, Kurose J, Towsley D, et al. Adaptive Playout Mechanisms for Packetized Audio Applications in Wide - Area Networks[C]//In Proceedings of IEEE INFOCOM '94. Toronto, Ontario, Canada: IEEE Computer Society, 1994: 680 - 688.
- [3] Liang Ji J, Farber N, Girod B. Adaptive Playout Scheduling Using Time - Scale Modification In Packet Voice Communications[C]//In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Salt Lake City, Utah, USA: IEEE Computer Society, 2001: 1445 - 1448.
- [4] 陈明义, 钱敏, 谭勇. 提高IP电话QoS的综合去抖动同步策略研究[J]. 计算机工程与设计, 2006, 27(23): 4491 - 4493.
- [5] Verhelst W, Roelands M. An Overlap - add Technique Based on Waveform Similarity (WSOLA) For High Quality Time - scale Modification of Speech[C]//In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 93. Minneapolis, Minnesota, USA: IEEE Computer Society, 1993: 554 - 557.
- [6] Kurose J E, Rose K W. Computer Networking: A Top - Down Approach Featuring the Internet (影印版)[M]. 3rd Edition. 北京: 高等教育出版社, 2005: 175 - 178.