

分组交换网络语音回放控制及性能分析^{*}

谭献海 温碧丽 金炜东

(西南交通大学信息科学与技术学院 成都 610031)

摘要 语音通信中的回放控制是保证语音平滑性的关键。本文深入研究分组交换网中的语音回放控制及其性能分析问题。在总结语音通信系统的特点和端到端时延的组成及其计算方法的基础上,提出了一个简单可行的语音业务回放控制模型,采用双令牌桶对语音流量进行整形。基于最新的网络演算理论,推导出了给定端到端时延、语音到达曲线和网络服务曲线条件下的语音回放时延、需要分配的速率和缓冲区长度的计算公式。最后通过应用实例分析验证了本文的分析结论。

关键词 分组交换网络,语音业务,回放,性能分析,网络演算

Playback Control and its Performance Analysis of Voice Traffic over Packet-Switch Network

TAN Xian-Hai WEN Bi-Li JIN Wei-Dong

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031)

Abstract Playback control is the key guarantee scheme of voice smoothing. In this paper, the problems of voice playback control and its performance analysis are studied. Based on the summary of the characteristic of voice communication and its end-to-end components, a simple and effective voice playback control model is proposed, which exploiting double token buckets as voice traffic shaper. Based on the novel network calculus theory, the formulations of playback delay, desired rate and buffer size are derived, given the end-to-end delay, the voice arrival curve and the network service curve. Finally, the validity and effectiveness of the analytic results of the paper is verified with an application example.

Keywords Packet-switch network, Voice traffic, Playback, Performance analysis, Network calculus

1 引言

随着网络的发展,Internet 正逐渐演变为一个通用的网络平台,各种各样的应用均可通过 Internet 来传输,语音通信也将逐渐迁移到 Internet 上。Internet 的分组交换方式与传统 PSTN 的电路交换方式有很大的区别,分组交换网中的语音通信控制问题是一个值得深入研究的技术问题。目前国内外的研究者在这方面的研究大都是定性分析,很少涉及到具体如何实现分组交换网络中的语音回放控制,以获取希望的服务质量^[5,6](如时延、时延抖动和丢包率等)的问题。这正是本文的研究动机。

网络演算^[1~3]是最近 10 多年在国外发展起来的一种网络分析方法,利用它可以方便地计算确定服务条件下多个网络节点的端到端时延和队列长度等。本文采用网络演算理论研究分组交换网中的语音回放控制及其性能分析问题。在简单介绍网络演算基本概念及结论的基础上,总结了语音通信系统的特点和端到端时延的组成及其计算方法,以确定语音通信端到端时延、排队时延和回放时延之间的关系。提出了一个简单可行的语音业务回放控制模型,采用双令牌桶对语音流量进行整形,以便考虑语音通信的突发性。基于最新的网络演算理论,推导出了给定端到端时延、语音到达曲线和网络服务曲线条件下的语音回放时延、需要分配的速率和需要的缓冲区长度的计算公式。最后通过应用实例分析验证了本文的分析结论。

2 网络演算基本概念

作为确定服务排队分析数学基础的网络演算^[1]最早由 Cruz 提出,许多研究者对它进行了精炼和深化^[2,3],它采用包络方法将网络的到达过程和服务过程分别描述为到达曲线和服务曲线,到达曲线描述给定时间间隔内数据流产生的最大流量,服务曲线表示系统提供给流的最小服务。借助最小加代数运算,就可以方便地计算多个网络节点的端到端时延和队列长度等。下面简要介绍本文用到的网络演算的基本概念及结论。

定义 1(最小加卷积) 函数 f 和 g 的最小加卷积运算为:

$$(f \otimes g)(t) = \begin{cases} 0, & t < 0 \\ \inf_{s \in [0, t]} [f(t-s) + g(s)], & t \geq 0 \end{cases}$$

定义 2(最小加反卷积) 函数 f 和 g 的最小加反卷积运算为:

$$(f \oslash g)(t) = \sup_{s \in [0, \infty]} [f(t+s) - g(s)]$$

定义 3(达到曲线) 给定一个广义增函数 $\alpha(t)$, 若对于所有的 $t \geq 0, s \leq t$, 输入函数 x 满足 $x(t) - x(s) \leq \alpha(t-s)$, 则称 α 是 x 的到达曲线, $x(t) \leq \inf_{0 \leq s \leq t} \{\alpha(t-s) + x(s)\} = (\alpha \otimes x)(t)$ 。

定义 4(服务曲线) 考虑一个系统 S , 通过该系统的数流具有输入函数 x 和输出函数 y , 若对于所有的 $t \geq 0$, 存在

^{*}项目简介:西南交通大学科学研究基金项目(2005A03),国家自然科学基金项目(No. 60572143)。谭献海 副教授,博士生,主要研究方向为计算机网络。温碧丽 硕士研究生,主要研究方向为计算机网络。金炜东 教授,博导,主要研究方向为职能信息处理、系统仿真。

$0 \leq t_0 \leq t$, 满足 $y(t) - x(t_0) \geq \beta(t - t_0)$, 则称系统 S 为数据流提供服务曲线 $\beta(t)$, $y(t) \leq \inf_{0 \leq s \leq t} \{\beta(t-s) + x(s)\} = (\beta \otimes x)(t)$ 。

定理 1(节点串联) H 个系统串联后提供的总服务曲线 β 为这些系统提供的服务曲线 $\beta_i, i=1, 2, \dots, H$ 的最小加卷积, 即 $\beta = \beta_1 \otimes \beta_2 \otimes \dots \otimes \beta_H$ 。对于两个延迟速率(latency-rate)服务曲线, 串联后的等效服务曲线为:

$$\beta = \beta_{R1, T1} \otimes \beta_{R2, T2} = \beta_{R1 \wedge R2, T1 + T2}$$

定理 2(积压界) 假定一个数据流受限于到达曲线 α , 通过一个提供服务曲线 β 的系统传输, 则系统的积压(backlog)上界为到达曲线与服务曲线之间的垂直偏差的最大值, 即:

$$B(t) = v(\alpha, \beta) = \sup_{u \geq 0} \{\inf \{\alpha(u) - \beta(u)\}$$

定理 3(时延界限) 假定一个数据流受限于到达曲线 α , 通过一个提供服务曲线 β 的系统传输, 则系统的时延上界为到达曲线与服务曲线之间的水平偏差的最大值, 即:

$$d(t) \leq h(\alpha, \beta) = \sup_{s > 0} [\inf \{T; T \geq 0 \text{ and } \alpha(s) \leq \beta(s+T)\}]$$

3 语音业务回放控制模型

3.1 语音业务的特点

语音业务具有许多与视频、数据等业务不同的特点^[10]。其中第一个特点是低数据率。传统的语音编码器是 G. 711 及其变种^[7], 采用 PCM 编码, 每 125ms 产生 8 比特采样数据, 对应的数据速率为 64kbps。最近 10 年, 提出了几个采用码激励线性预测(CELP)技术的新的语音编码标准, 这些编码标准以增加编码时延为代价来降低码速, 其中 G. 729A^[8] 的速率为 8kbps, G. 723^[9] 的速率为 5.3kbps 或 6.4kbps, 如表 1 所示。

表 1 G. 711, G. 729A 和 G. 723. 1 的帧大小 f , 预取时延 l 和处理时延 p

Encoding scheme	G. 711	G. 729A	G. 723. 1
(Nominal bit rate r)	(64kbit/s)	(8kbit/s)	(5, 3/6, 4kbit/s)
Frame size f	125 μ s	10ms	30ms
Look-ahead l	0	5ms	7.5ms
Processing delay p	Negligible	Less than 10ms	Less than 30ms

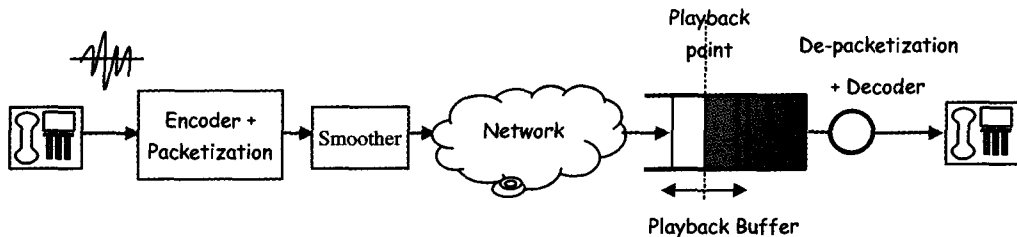


图 1 语音回放控制方案

3.3 端到端时延分析

语音通信的端到端时延指的是从发送端语音产生的时刻开始, 到接收端收听到语音为止的之间差, 主要包括: 模拟信号采用的编码时延 D_{enc} , 是下列三种时延之和: 帧尺寸 f , 预处理(look-ahead)时延 l 和处理时间 p , 具体取值如表 1 所示。一般地, 编码比特流的速率 r 越小, 帧尺寸、预处理时延和处理时间就越大。分组化时延 D_{enc} , D_{enc} 为一个包中包含的帧数 k 的函数, $D_{pack} = (k-1)f$ 。语音报文中网络中的传输时延, 传输路径上的每一跳均包括传输时间 T_h , 排队时间 Q_h 和传播时延 P_h , 其中传播时延由链路的长度决定, 局域网的传播

语音业务的第二个特点是高的突发。一个语音通信周期包括话音区(talk-spurt)和静音(silence)区两部分, 话音区和静音区交替出现, 可用一个两状态的 Markov 链来建模。典型地, 话音区和静音的长度分别为 1.65 和 1.35 秒^[5]。使用静音抑制技术, 可以消除静音部分。

语音业务的第三个特点是对时延、时延抖动和丢包率有严格的要求。根据 ITU-T 的 G. 114 建议^[11], 可容忍的单向端到端时延小于 400ms, 时延抖动最大不超过 80ms。由于 Internet 具有长相关特性, 其报文丢失是相关的, 如果出现报文, 后面的报文也很可能发生丢失, 因而连续的报文丢失周期很容易突破 60ms, 文[12]的主观测试结果表明连续 60ms 以上的丢包周期会影响接收语音的可理解性。为此, 一般要求语音通信具有较低的丢包率(如 10^{-5})。

3.2 语音回放控制模型

根据语音的特点, 需要对语音业务进行有效的回放控制, 以控制其时延、时延抖动和丢包率符合规定的要求。本文采用的回放控制方案如图 1 所示。语音源产生的模拟语音信号首先经过采样、量化和编码, 变成离散信号, 然后进行分组化处理。在分组报文进行网络传输之前, 为了保证语音信号能获得确保的服务质量, 必须对语音信号进行监管, 一般可以采用监控(policing)或整形(shaping)来实现, 符合合约规定的流量提供快速转发服务, 而对超出合约的流量进行丢弃、标记或提供尽力而为的服务。经过整形后的语音流量经过网络传送到接收端。在接收端, 为了保证语音播放的连续性, 减少或消除语音传输的时延抖动, 首先将接收到的语音数据放入回放缓冲区进行缓存, 等待规定的时间(回放时间)后在从缓冲区中取出语音数据, 进行解包和解码等处理, 最后交给语音播放器播放。

本语音回放控制方案的关键技术问题是: (1) 系统时延的组成及端到端时延的确定; (2) 语音整形及语音到达特性的确定; (3) 根据语音的端到端时延要求和语音到达特性, 确定网络需要给该语音流量预留的有效带宽, 以及接收端的回放时延及回放缓冲区的大小。

时延可以忽略不计, 大陆内(如哈尔滨到海南三亚)的传播时延一般小于 30ms, 大陆之间(如北京到美国的圣地亚哥)的传播时延一般在 50ms~100ms 之间。网络平均传输时延可通过 RFC3611 的 RTP 控制协议扩展报告(RTCP XR)^[13] 来获得。

在接收端产生的时延主要包括语音回放时延 D_{play} 、解包和解码等处理时延 D_{dec} 。由于目前的处理器的能力都很强, 可以忽略发送端及接收端的处理时延(p 和 D_{dec}), 因此, 语音通信端到端的时延为:

$$D_{e2e} = T_f + l + \sum_{hc \text{ path}} (T_h + Q_h + P_h) + D_{play} \quad (1)$$

其中 $T_f = kf$, k 为帧数, f 为帧大小。

式(1)中的变量为排队时延和回放时延。其中排队时延可以有会话协商过程给定,如 RSVP 的 ADSpec^[6] 中包含有路径的时延特性信息,包括排队时延 C_{tot} , D_{tot} , 以及路径的固定时延(包括网络节点的处理时延和链路的传播时延)。端到端的排队时延为:

$$V_{tot} = \frac{C_{tot}}{R} + D_{tot}$$

其中 R 为网络为语音通信预留的速率。可见,给定语音通信系统的端到端时延,就可以确定对应的回放时延要求。

3.3 语音流量整形

本文采用的语音流量整形方案如图 2 所示。为了考虑语音流量的突发性,采用双令牌桶。其中参数为峰值速率 p 和最大分组长度 M 的第一个桶用于控制突发,而语音通信的长期行为由参数为平均速率 r 和桶大小 b 的第二个令牌桶来控制。经整形器整形后的语音流量的到达曲线为:

$$a^{TSpec}(t) = \begin{cases} M + pt, & t < T \\ +rt, & t \geq T \end{cases}, T = \frac{b-M}{p-r} \quad (2)$$

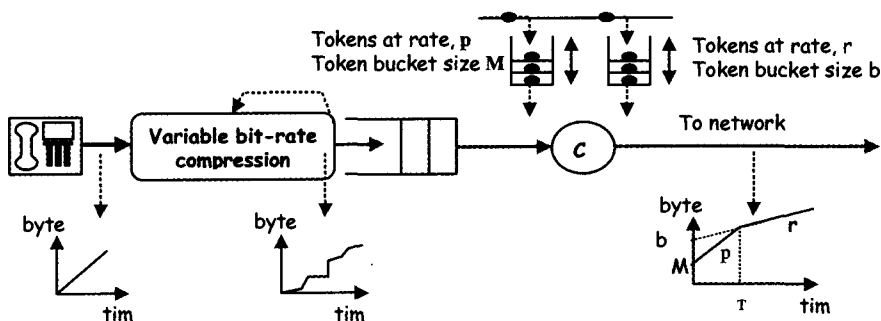


图 2 整形器和到达曲线

4 语音回放控制参数的确定

4.1 单级网络

下面根据网络演算的基本概念及其结论,推导语音回放控制系统的有效带宽、回放时延和需要的缓冲区长度大小。语音回放控制系统的到达曲线如式(2)。服务曲线采用基于延迟速率(latency-rate)调度器^[4]的确定服务线性函数:

$$\beta_{R,V_{tot}}(t) = R[t - V_{tot}]^+ = \begin{cases} 0, & t \leq V_{tot} \\ R(t - V_{tot}), & t > V_{tot} \end{cases} \quad (3)$$

式中 $V_{tot} = \frac{C_{tot}}{R} + D_{tot}$ 。 R 为排队服务策略分配给流的速率, $R \geq r$ 。 C_{tot} 和 D_{tot} 分别为基于分组的调度器与理想流模型^[4](GPS)之间的速率相关和速率无关时延偏差,对于许多调度器,特别是常用的 WFQ(Weighted Fair Queueing), $C_{tot} = M$, $D_{tot} = M/c$, 其中 c 为链路速率。

根据网络演算基本结论,语音回放控制系统的有效带宽、回放时延和需要的缓冲区长度大小之间的关系如图 3 所示。可以求得给定上述到达曲线和服务曲线条件下系统的最小回放时延为:

$$D_{play} = \begin{cases} \frac{T(p-R)+M}{R} + L_{tot}, & p \geq R \geq r \\ \frac{M}{R} + L_{tot}, & R \geq p \geq r \end{cases} \quad (4)$$

需要分配给语音业务的速率为:

$$R = \begin{cases} \frac{pT+M+C_{tot}}{D_{play}+T-D_{tot}}, & D_{play} \geq \bar{d} \\ \frac{M+C_{tot}}{D_{play}-D_{tot}}, & D_{play} < \bar{d} \end{cases} \quad (5)$$

其中 $\bar{d} = \frac{M+C_{tot}}{p} + D_{tot}$

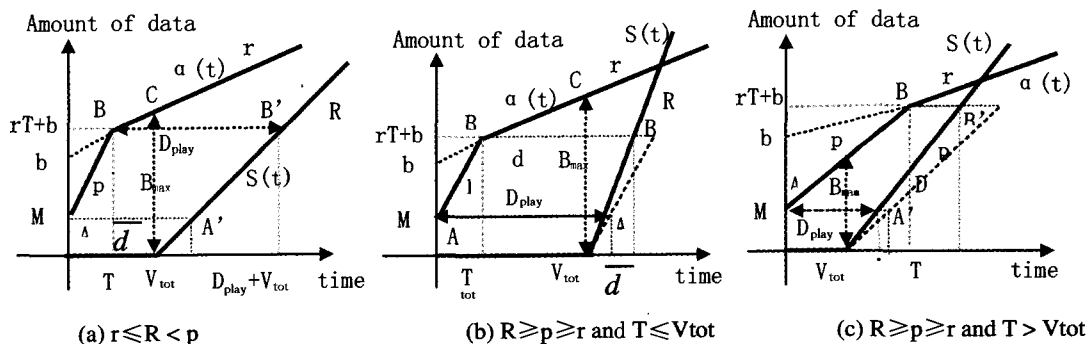


图 3 语音回放控制系统参数确定方法

需要的缓冲区容量为:

$$B_{max} = \begin{cases} b + \frac{(r-R)(b-M)}{p-r} + RV_{tot}, & p \geq R \geq r, V_{tot} \leq T \\ M + pV_{tot}, & R \geq p \geq r, V_{tot} \leq T \\ b + rV_{tot}, & V_{tot} > T \end{cases} \quad (6)$$

4.2 多级网络

对于多级网络,如发送端和接收端分别通过接入网连接到骨干网这样的三级网络系统,其语音回放控制系统的有效带宽、回放时延和需要的缓冲区长度大小的计算方法与上述的分析过程类似,只需要用整个网络系统的服务曲线代替上面分析中的服务曲线即可。

(下转第 119 页)

表 1 比较次数的统计表

数据集组比较	比较范围 δ	比较次数
(10,10)比较:	1	(3171 * 3170)/2 次
(11,10)比较:	1	2748 * 3171 次
(11,11)比较:		(2748 * 2747)/2 次
(12,11)比较:	1	2445 * 2748 次
(12,12)比较:		(2445 * 2444)/2 次
...
(150,136)比较:	16	9 * 12
(150,137)比较:		9 * 9
(150,138)比较:		9 * 16
(150,139)比较:		9 * 22
(150,140)比较:		9 * 12
(150,141)比较:		9 * 9
(150,142)比较:		9 * 21
(150,143)比较:		9 * 10
(150,144)比较:		9 * 9
(150,145)比较:		9 * 10
(150,146)比较:		9 * 11
(150,147)比较:		9 * 17
(150,148)比较:		9 * 15
(150,149)比较:		9 * 10
(150,150)比较:		(9 * 8)/2
总比较次数		84592370

说明:表中(i,j)表示出度为i的子集和出度为j的子集的比较。

参 考 文 献

- Gibson D, Kleinberg J, Raghavan P. Inferring Web Communities from Link Topology. In: Proc. of the 9th ACM Conf. on Hypertext and Hypermedia. Pittsburgh, PA, USA, 1998. 225~234
- Reddy P K, Kitsuregawa M. Inferring Web Communities through relaxed-cocitation and power-law. KITSUREGAWA Lab:[Annual Report]. 2001. 27~46
- Broder A, Galssman S, et al. Sytactic clustering of the Web. In: Proc. of the sixth international WWW Conf. April 1997. 391~404
- Shivakumar N, et al. Finding near-replicas of documents on the Web. WebDB, 1998. 204~212
- Wang Y, Kitsuregawa M. Clustering of Web Search Results with Link Analysis. KITSUREGAWA Lab:[Annual Report]. 2001. 18~26
- Broder A Z, Kumar R, Maghoul F, et al. Graph structure in the Web. Computer Networks, 2000. 309~320
- Dean J, Henzinger M R. Finding Related Pages in the World Wide Web. In: Proc. of the eighth international WWW Conf. Toronto, Canada: 1999. 389~401

(上接第 115 页)

三级网络构成一个串联的网络,串联后的网络提供的总服务曲线 β 为三个子系统提供的服务曲线 $\beta_i, i=1,2,3$ 的最小加卷积,即 $\beta = \beta_1 \otimes \beta_2 \otimes \beta_3$ 。如果三个网络均采用延迟一速率服务曲线,则串联后总的服务曲线为: $\beta = \beta_{R1, V1} \otimes \beta_{R2, V2} \otimes \beta_{R3, V3} = \beta_{R1 \wedge R2 \wedge \beta_3, V1 + V2 + V3}$ 。

5 应用实例分析

设语音系统采用 G. 729A 编码,语音帧的格式化时间为 $T_f = 30$ ms,预处理时延为 $l = 5$ ms,总的端到端传输时延、排队时延和传播时延分别为 5.80 和 30 ms。设语音允许的最大单向端到端延迟为 $d_{max} = 250$ ms。

假定系统采用的编码规则为 G. 729A,每 10ms 产生一个数据帧,每个数据帧长 10 字节。通常一个 IP 包包括 3 个数据帧,则单方向产生数据包的速率为 33 包/s,每个数据包包含 30 字节的语音编码数据。在以网上传输时一个 IP 电话数据包的大小为 84 字节。取峰值速率为平均速率的 4 倍,则整形系统的参数为 $(M, p, b, r) = (84 \text{ 字节}, 11088 \text{ 字节/秒}, 400 \text{ 字节}, 2772 \text{ 字节/秒})$ 。采用 IP 网络,假定语音数据流的传输路径长度为 5 跳,各跳的 MTU 均为 1518 字节,链路速率为 $c = 10$ Mbps,网络节点采用 WFQ 调度策略。

根据上述给定条件,系统允许的最大回放时延为:

$$D_{play} = 250 - (30 + 5 + 5 + 80 + 30) = 100\text{ms}$$

$$C_{tot} = 5 * M = 420 \text{ 字节}, D_{tot} = 5 * MTU / c = 6.072\text{ms}.$$

$$T = (b - M) / (p - r) = 40\text{ms}$$

$$\bar{d} = \frac{M + C_{tot}}{p} + D_{tot} = 51.5\text{ms}, \bar{d} < D_{play}.$$

代入上面的式(5),计算得:

$$R = \frac{pT + M + C_{tot}}{D_{play} + T - D_{tot}} = 7071 \text{ 字节/秒} \approx 2.5r.$$

$$V_{tot} = \frac{C_{tot}}{R} + D_{tot} = 65.4\text{ms}$$

代入式(6)可以求得需要的回放缓冲区长度为: $B_{max} = b + rV_{tot} = 581$ 字节。

结束语 本文深入研究了语音通信系统的特点和端到端

时延的组成及其计算方法,提出一个简单可行的语音业务回放控制模型,采用双令牌桶对语音流量进行整形,充分考虑了语音通信的突发性特点。基于最新的网络演算理论,推导出了给定端到端时延、语音到达曲线和网络服务曲线条件下的语音回放点(回放时延)、需要分配的速率和需要的缓冲区长度的计算公式,并简单论述了由接入网和骨干网组成的多网络环境下上述公式的使用方法。最后通过应用实例分析验证了本文的分析结论。分析结果表明,对于 250ms 的单项端到端时延,100ms 左右的回放时延要求,需要分配的速率和缓冲区比较合理,带宽利用率达 40% 左右。

参 考 文 献

- Cruz R L. Quality of Service Guarantees in Virtual Circuit Switched Networks. IEEE Journal of Selected Area in Communication, 1995, 13(6): 1048~1056
- Boudec J Y L, Thiran P. Network Calculus: A Theory of Deterministic Queuing System for the Internet. Heidelberg: Springer-Verlag, 2004
- Firoiu V, Le Boudec J-Y, Towsley D, Zhang Zhi-Li. Theories and models for Internet quality of service. Proceedings of the IEEE, 2002, 90(9): 1565~1591
- Yuming J. Relationship between guaranteed rate server and latency rate server. Computer Networks, 2003, 43(3): 307~315
- Shenker S, Partridge C, Guérin R. Specification of guaranteed quality of service. RFC 2212, 1997
- Wroclawski J. The use of RSVP with IETF integrated services. RFC2210, September 1997
- Recommendation G. 711, Pulse Code Modulation (PCM) of Voice Frequencies. ITU, Nov. 1988
- Annex A to Recommendation G. 729, Coding of Speech at 8kbit/s using Conjugate Structure Algebraic-Code-Excited Linear-Prediction (CSACELP), Annex A: "Reduced Complexity 8 kbit/s CS-ACELP Speech Codec". ITU, Nov. 1996
- Recommendation G. 723. 1, Speech Coders; Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s. ITU, March 1996
- Karam M, Tobagi F. Analysis of the delay and jitter of voice traffic over the Internet. In: Proc. of Infocom 2001. http://citeseer.ist.psu.edu/karam01analysis.html
- Recommendation G. 114. Transmission systems and media general characteristics of Internet telephone connections and Internet telephone circuits one-way transmission time. http://www.pesq.org, 1996
- Gruber J, Strawczynski L. Subjective Effects of Variable Delay in Speech Clipping in Dynamically Managed Voice Systems. IEEE Transactions on Communications, 1985, COM-33(8)
- Friedman T, Caceres R, Clark A. RTP control protocol extended reports (RTCP XR). RFC 3611. http://www.ietf.org, 2003