

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

VANDERSILVIO DA SILVA

**Proposta de Metodologia para Avaliação de
Redes de Voz sobre IP**

Dissertação apresentada como requisito parcial
para a obtenção do grau de Mestre em Ciência
da Computação

Prof. Dr. João Netto
Orientador

Porto Alegre, junho de 2006.

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Silva, Vandersilvio da

Proposta de Metodologia para Avaliação de Redes de Voz sobre IP / Vandersilvio da Silva – Porto Alegre: Programa de Pós-Graduação em Computação, 2006.

92 f.:il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre, BR – RS, 2006. Orientador: João Netto;

1.Redes. 2.VoIP. 3.Qualidade. I. Netto, João. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. José Carlos Ferraz Hennemann

Vice-reitor: Prof. Pedro Cezar Dutra Fonseca

Pró-Reitora de Pós-Graduação: Profa. Valquiria Linck Bassani

Diretor do Instituto de Informática: Prof. Philippe Olivier Alexandre Navaux

Coordenador do PPGC: Prof. Flávio Rech Wagner

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

AGRADECIMENTOS

À minha esposa Marilaine , pelo carinho, incentivo e compreensão.

Aos amigos e colegas de trabalho da Equipe de Redes da Digitel, pela ajuda,
amizade e incentivo.

Ao Guilherme Marshall pela parceria, disponibilidade e troca de informações.

Á Digitel por permitir as ausências no horário de trabalho.

Aos professores do programa de Pós-graduação do II, pelo ensino oferecido.

Ao meu orientador Dr. João César Netto, pelos ensinamentos, oportunidade e
confiança depositada.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	7
LISTA DE FIGURAS	9
LISTA DE TABELAS	10
RESUMO	11
ABSTRACT	12
1 INTRODUÇÃO	13
1.1 Contextualização.....	13
1.2 Objetivo da Dissertação	13
1.3 Organização do Trabalho	13
2 PROTOCOLOS DE VOZ SOBRE IP	15
2.1 Real-time Transport Protocol – RTP.....	16
2.2 RTP Control Protocol – RTCP.....	18
2.3 Session Initiation Protocol – SIP	18
2.3.1 Componentes de um ambiente de rede SIP	19
2.3.2 Acesso ao ambiente de rede SIP	20
2.3.3 Requisições e respostas SIP	20
2.3.4 Processo de Execução de Conexões SIP.....	21
2.4 O Padrão H.323.....	22
2.4.1 Componentes de um Sistema H.323	23
2.4.2 Protocolos H.323	23
2.4.3 Processo de Execução de Conexões H.323	25
2.5 O Padrão MGCP.....	27
2.6 O Padrão H.248/MEGACO	28
2.7 Conclusões do capítulo 2	30
3 FATORES DE QUALIDADE DE VOZ	31
3.1 Perda de Pacotes	31
3.2 Perda em Rajadas	32
3.3 Atraso.....	33
3.4 Variação de Atraso ou <i>Jitter</i>	35
3.5 Eco.....	35
3.6 Momento da falha	36

3.7	Peculiaridades dos codecs	37
4	MÉTODOS DE AVALIAÇÃO DA QUALIDADE DE VOZ	38
4.1	Mensuração Subjetiva da Qualidade da Fala	39
4.1.1	ACR – MOS.....	39
4.1.2	DCR – DMOS.....	40
4.2	Mensuração Objetiva Intrusiva da Qualidade da Fala.....	41
4.2.1	PSQM e PSQM+.....	41
4.2.2	MNB	42
4.2.3	PAMS.....	42
4.2.4	PESQ.....	42
4.3	Mensuração Objetiva Não-Intrusiva da Qualidade da Fala.....	44
4.3.1	Modelo E.....	44
4.3.2	Extended E-Model	47
4.3.3	Modelo de Rede Neural Artificial	49
4.3.4	P.SEAM	50
4.4	Métodos de Avaliação em fase de Padronização.....	52
4.4.1	P.AAM.....	52
4.4.2	P.VTQ.....	52
4.5	Análise sobre os Métodos de Avaliação de Qualidade da Voz.....	54
4.5.1	Comparação entre Medições Ativas e Medições Passivas	54
4.5.2	Considerações sobre Extração de Parâmetros	56
5	DIMENSÕES DA AVALIAÇÃO DA QUALIDADE DE VOZ SOBRE IP	57
5.1	Definição de Qualidade da Chamada	57
5.1.1	Qualidade da Escuta.....	57
5.1.2	Qualidade da Conversação.....	57
5.1.3	Qualidade da Transmissão	58
5.2	Métricas Primárias de Qualidade de Voz sobre IP	58
5.2.1	Clareza	59
5.2.2	Atraso e <i>Jitter</i>	60
5.2.3	Eco	60
5.3	Qualidade de Serviço Percebida (QoE)	60
5.4	Análise sobre as Dimensões da Qualidade da Voz sobre IP	62
5.4.1	A diferença entre os Métodos de Medida de QoS e QoE	62
5.4.2	Considerações sobre os Métodos de medida de Escuta, de Conversação e de Estrutura	62
6	TRABALHOS CORRELATOS	63
6.1	Emissão de CDRs por <i>Gateways</i> de Voz	63
6.2	Monitoração Ativa	64
6.3	Monitoração Passiva em Pontos Intermediários da Rede.....	66
6.4	Monitoração nos Pontos Iniciais e Finais da Chamada	67
7	METODOLOGIA PROPOSTA PARA AVALIAÇÃO DE REDES DE VOZ SOBRE IP	71
7.1	O Procedimento em 7 Blocos	71
7.1.1	Bloco OBJETIVO.....	71
7.1.2	Bloco AMBIENTE	72
7.1.3	Bloco FERRAMENTAS.....	73

7.1.4	Bloco ABORDAGEM	75
7.1.5	Bloco APLICAÇÃO	77
7.1.6	Bloco ANÁLISE.....	77
7.1.7	Bloco REPLANEJAMENTO	77
7.1.8	O Conjunto Final com todos os Blocos	77
7.2	Validação do Método.....	78
7.2.1	Bloco OBJETIVO.....	78
7.2.2	Bloco AMBIENTE	78
7.2.3	Bloco FERRAMENTAS.....	80
7.2.4	Bloco ABORDAGEM	80
7.2.5	Bloco APLICAÇÃO	80
7.2.6	Bloco ANÁLISE.....	80
7.2.7	Bloco REPLANEJAMENTO	80
7.2.8	Bloco (nova) ABORDAGEM.....	81
7.2.9	Bloco (nova) APLICAÇÃO.....	81
7.2.10	Bloco (nova) ANÁLISE.....	81
8	CONCLUSÃO.....	82
	REFERÊNCIAS.....	84
	GLOSSÁRIO.....	92

LISTA DE ABREVIATURAS E SIGLAS

AAM	Acoustic Assessment Model
CDR	Call Detail Record
CODEC	COder / DECoder
ETSI	European Telecommunications Standards Institute
IETF	Internet Engineering Task Force
IMT	InterMachine Trunk
ISDN	Integrated Service Digital Network
ITU-T	International Telecommunication Union Telecom Standardization
MCU	Multipoint Controller Units
MEGACO	Media Gateway Control
MGC	Media Gateway Controler
MIPS	Million Intruction Per Second
MOS	Mean Opinion Score
PAMS	Perceptual Analysis Measurement System
PESQ	Perceptual Evaluation of Speech Quality
PRI	Primary Rate Interface
PSQM	Perceptual Speech Quality Measurement
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RNP	Rede Nacional de Ensino e Pesquisa
RSVP	Resource ReSerVation Protocol
RTCP	Real-time Transport Control Protocol
RTCP-XR	Real-Time Control Protocol eXtended Reports
RTP	Real-time Transport Protocol
RTP	Real-time Transport Protocol
RTPC	Rede Telefônica Pública Comutada
SCN	Switched Circuit Network

SG	Signaling Gateway
SIP	Session Initiation Protocol
SNMP	Simple Network Management Protocol
SS7	Signaling System 7
STFC	Sistema de Telefonia Fixa Comutada
TCP	Transmission Control Protocol
VAD	Voice Activity Detection

LISTA DE FIGURAS

Figura 2.1: Principais Protocolos de voz sobre IP.....	15
Figura 2.2: Formato de um pacote RTP.....	16
Figura 2.3: Processo de execução de uma conexão SIP	21
Figura 2.4: Pilha de Protocolos H.323.....	24
Figura 2.5: Estabelecimento da chamada	25
Figura 2.6: Sinalização de Controle.....	26
Figura 2.7: Término da Chamada	27
Figura 2.8: Controle de ações do MGCP.....	28
Figura 2.9: MGCP ou H.248/MEGACO no controle de Gateways	29
Figura 3.1: Perda por Rajada versus perda Aleatoriamente Distribuída.....	32
Figura 3.2: Valores PESQ para comparação entre <i>codecs</i>	33
Figura 4.1: Classificação de Métodos de Medição	38
Figura 4.2-Métodos baseados em:(a) Parâmetros, (b) Sinais, (c) Comparação.....	39
Figura 4.3: Filosofia básica do PESQ.....	43
Figura 4.4 : Métodos Não-Intrusivos.....	44
Figura 4.5 : Conexão referência do Modelo E	46
Figura 4.6 : Comparação entre o Extended E-Model (VQmon), o MOS (Subjective) e o Modelo E	48
Figura 4.7 : Modelo computacional do Extended E-Model	49
Figura 4.8 : Diagrama conceitual do modelo ANN.....	50
Figura 4.9 : Diagrama em blocos do P.563	51
Figura 4.10 : Arquitetura da PROBE PsyVoIP	53
Figura 4.11 : Comparação de processamento e acuracidade entre métodos... ..	55
Figura 4.12 : Comparação entre P.563 e ACRMOS.....	55
Figura 5.1 : Métricas Primárias de Qualidade de Voz	59
Figura 5.2 : Dimensão da Qualidade Percebida.....	61
Figura 5.3 : Inter-relacionamento de fatores de QoS.....	61
Figura 6.1 : Arquitetura do Sistema ECAS	64
Figura 6.2 : Monitoração Ativa da Qualidade de Voz	65
Figura 6.3 : Monitoração Passiva da Qualidade de Voz com várias <i>probes</i>	66
Figura 6.4 : Arquitetura da Biblioteca VQuality	68
Figura 6.5 : Uso de VQmon em pontos finais de chamada.....	68
Figura 6.6 : Estabelecimento de chamada com EMBAC-DP	69
Figura 7.1 : Tomada de informações para definir a ABORDAGEM.....	75
Figura 7.2 : Procedimento completo com os sete blocos	78
Figura 7.3 : Cenário para a Validação do Método.....	79

LISTA DE TABELAS

Tabela 4.1: Relação entre fator R e MOS	45
Tabela 4.2 : Descrição dos parâmetros da figura 4.2.....	46

RESUMO

A redução de custo com telefonia através do uso de voz sobre IP tem disparado a busca de soluções que transformem redes IP originalmente dedicadas a transporte de dados em redes para transporte de voz.

Esta dissertação tem por objetivo apresentar uma metodologia para sistematizar a avaliação de redes para o tráfego de voz sobre IP de acordo com as possibilidades disponíveis no cenário a ser avaliado.

Inicialmente é dada uma visão geral de voz sobre IP, apresentando os protocolos utilizados, os fatores que influenciam na qualidade da voz e os métodos de avaliação de qualidade da voz.

Na seqüência são apresentados trabalhos correlatos a avaliação de qualidade de aplicações de voz sobre IP.

E por fim descreve-se a proposta de uma metodologia para sistematizar a avaliação de redes com VoIP.

Palavras-Chave: VoIP, QoS, Avaliação de Qualidade, Monitoração.

Proposal of Methodology for Evaluation of Voice over IP Networks

ABSTRACT

The use of voice over IP telephony was started with solutions to adapt existent data networks to carrier voice streams. The use of monitoring techniques, QoS and signaling protocols can be combined on a such design.

Our goal is to present a methodology to evaluate and choose the probing points and the voice quality evaluation techniques to be used in network redesign.

An overview about VoIP protocols and parameters that change the voice quality are presented as well as some related works on evaluating voice quality based on network parameters.

A proposed methodology is presented, with a case study to show how one can choose the right combination of probing points with some voice quality measurement technique.

Keywords: VoIP, QoS, Quality Measurement, Quality Assessment, Quality Monitoring.

1 INTRODUÇÃO

Esta introdução apresenta uma contextualização sobre redes com voz sobre IP e suas respectivas avaliações de qualidade, os objetivos desta dissertação e sua organização.

1.1 Contextualização

O apelo de redução de custo com telefonia através do uso de voz sobre IP tem disparado a busca de soluções que transformem redes IP originalmente dedicadas a transporte de dados em redes para transporte de voz.

Uma má avaliação da capacidade da rede para este tipo de tráfego pode gerar a frustração das expectativas.

Para evitar esta frustração é necessário medir objetivamente a capacidade da rede e levantar os fatores que influenciam a qualidade da voz .

Para realizar uma avaliação consistente é necessário conhecer os tipos de avaliação, e utilizar um método sistemático de medição.

1.2 Objetivo da Dissertação

Esta dissertação tem por objetivo apresentar uma metodologia para sistematizar a avaliação de redes para o tráfego de voz sobre IP de acordo com as possibilidades disponíveis no cenário a ser avaliado.

1.3 Organização do Trabalho

Este trabalho está estruturado em oito capítulos. O primeiro capítulo apresenta uma contextualização sobre redes com voz sobre IP e suas respectivas avaliações de qualidade, os objetivos desta dissertação e sua organização.

O segundo capítulo apresenta a descrição dos principais protocolos utilizados em voz sobre IP.

O terceiro capítulo descreve os fatores que influenciam na qualidade da voz.

O quarto capítulo detalha os métodos de avaliação de qualidade da voz.

O quinto capítulo apresenta diversas formas de como a qualidade da voz pode ser entendida, mostrando abordagens diferentes e específicas para o termo generalizado “qualidade de voz”.

O sexto capítulo apresenta trabalhos correlacionados com o tema desta dissertação.

O sétimo capítulo descreve o objetivo principal desta dissertação, a proposta de uma metodologia para sistematizar a avaliação de redes para o tráfego de voz sobre IP de acordo com as possibilidades disponíveis no cenário a ser avaliado.

E por fim, o oitavo capítulo apresenta as conclusões e sugestões de trabalhos futuros.

2 PROTOCOLOS DE VOZ SOBRE IP

Os protocolos utilizados em voz sobre IP podem ser divididos em dois grandes grupos: aqueles que tratam do transporte e aqueles que tratam da sinalização e controle. Os dois grupos trabalham simultaneamente na comunicação entre dois terminais na telefonia IP (BERNAL 2003).

No processo de transporte da voz, existem duas tarefas distintas : o controle do transporte e o transporte propriamente dito.

No processo de sinalização e controle se realizam o estabelecimento, o acompanhamento e a finalização da chamada.

Por sua vez, os protocolos de sinalização e controle também podem ser divididos em dois grupos : os protocolos “mestre/escravo” e os protocolos “*peer to peer*”.

Os protocolos “mestre/escravo”, (MGCP e MEGACO), normalmente são empregados quando elementos inteligentes controlam elementos sem inteligência.

Os protocolos “*peer to peer*”, (H.323 e SIP), são utilizados em interações entre elementos inteligentes.

A figura 2.1 apresenta um diagrama em blocos desses protocolos.

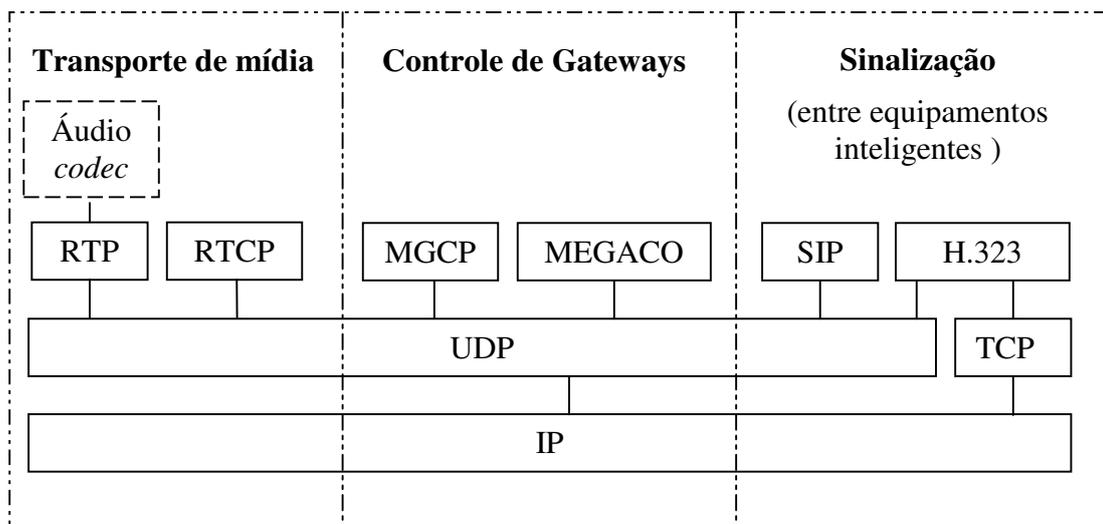


Figura 2.1: Principais Protocolos de voz sobre IP
(BERNAL 2003)

A seguir são apresentados os protocolos de transporte RTP, RTCP/RTCP XR, os protocolos de Sinalização SIP e H.323, e os protocolos de Controle de Gateway MGCP e H.248/MEGACO, com o intuito de explicitar o seu funcionamento para posterior entendimento de suas influências nos ambientes onde pretende-se avaliar a qualidade da voz.

2.1 Real-time Transport Protocol – RTP

O RTP está definido na RFC 3550 (SCHULZRINNE 2003) . O protocolo RTP utiliza o UDP como protocolo de transporte. A garantia de entrega e integridade dos dados, fornecida pelo protocolo TCP, não é adequada para aplicações do tipo multimídia em tempo real, pois gera *overhead* e atrasos não tolerados por estas categorias de aplicação. Não existem portas UDP padrão para transmissão de voz pelas sessões RTP. Estas portas são negociadas dinamicamente pelos protocolos de sinalização durante o processo de estabelecimento de chamada. O RTP apresenta mecanismos como número de seqüência e *timestamping* de pacotes, a fim de garantir que eventuais perdas serão detectadas pelo receptor. A implementação usando UDP/RTP não implementa a confirmação ou retransmissão de pacotes perdidos, visto que as mesmas não se adaptam às características das aplicações de tempo real. Além do RTP, a RFC 1889 descreve também o *RTP Control Protocol* (RTCP), responsável pela troca de mensagens relativas à qualidade do tráfego RTP entre os participantes de uma sessão. Informações trocadas incluem número de pacotes RTP perdidos, atraso e *jitter* observados por cada participante. A análise dessas informações seria por si só suficientes para aferir a qualidade de uma ligação VoIP, tornando desnecessária a análise separada do tráfego RTP. Entretanto, a implementação do RTCP não é obrigatória. Desta forma, muitos equipamentos VoIP disponíveis no mercado não implementam o RTCP, ou o implementam de forma incorreta, diminuindo assim a credibilidade das informações contidas no mesmo (DAL PIVA 2004). A Figura 2.2 apresenta o formato de um pacote RTP.

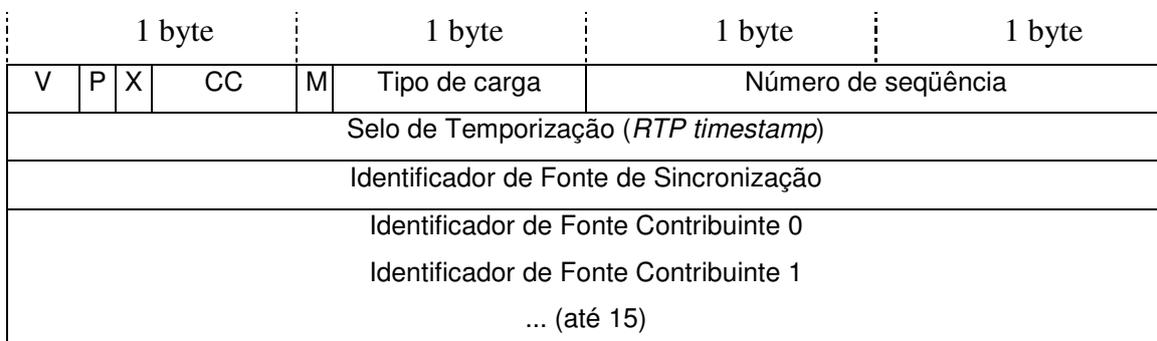


Figura 2.2: Formato de um pacote RTP

A descrição dos campos do pacote RTP é apresentada abaixo (MEDEIROS 2003).

V – versão, (2 bits). Especifica a versão do RTP.

0:refere-se ao primeiro protocolo utilizado na ferramenta de áudio “vat”.
1:especifica a primeira versão do RTP utilizada como teste.
2:identifica a versão do RTP especificada na RFC 1889.

P – Preenchimento/padding, (1 bit). Sinaliza a adição de octetos de enchimento adicionais ao conteúdo da carga (payload) sem fazer parte da mesma. O último octeto do preenchimento contém a informação de quantos octetos foram inseridos. Este preenchimento adicional é normalmente utilizado para uso de algoritmos de criptografia de tamanho de blocos fixos ou para transmissão de pequenos conteúdos.

X – Extensão/eXtension, (1 bit). Com esse bit marcado, é acrescentado uma extensão ao cabeçalho original.

CC - Contador de fonte Contribuinte, (4 bits). Este campo contém o número de identificadores de fonte contribuinte.

M – marcador/marker, (1 bit). Usado para identificar as fronteiras de um quadro numa corrente de pacotes.

PT – Tipo de carga/payload type, (7 bits). Este campo identifica o formato da carga do pacote RTP como também a determinação de sua interpretação pela aplicação.

Numero de seqüência (16 bits) – A numeração seqüenciada põe em ordem os diversos pacotes de RTP. A cada novo pacote, a numeração é incrementada de uma unidade. Basicamente, esse ordenamento serve para o receptor detectar os pacotes perdidos e restaurar a seqüência de pacotes.

Selo de temporização/RTP timestamp, (16 bits). Esse campo reflete o instante de amostragem do primeiro octeto no pacote RTP.

Identificador de fonte de Sincronização , (32 bits). Esse campo identifica a fonte de sincronismo. Esta identificação foi escolhida aleatoriamente tencionando-se que duas fontes de sincronismo com a mesma sessão RTP não teriam o mesmo identificador SSRC.

Identificador de fonte Contribuinte (32 bits cada identificador, podendo ter de o a 15 identificadores). A lista SCRC identifica a contribuição da fonte no conteúdo da carga (*payload*) de cada pacote. O número de identificadores é dado pelo campo CC. Se houver mais de 15 fontes contribuintes, somente 15 serão identificadas.

2.2 RTP Control Protocol – RTCP

Apesar das características vantajosas para multimídia do RTP, ele ainda não satisfaz nenhuma das necessidades de controle e QoS. O protocolo RTCP (SCHULZRINNE 2003), por sua vez, implementa funções de controle na troca de informações entre as fontes e os destinos. No caso específico de VoIP, sua função é informar ao transmissor como a conexão está sendo vista pelo receptor em termos de QoS. Existem cinco tipos diferentes de mensagens RTCP (MEDEIROS 2003):

SR (Sender Reports) – São mensagens geradas pelos usuários que estão enviando os pacotes. Elas descrevem além da quantidade dos dados transmitidos, as informações de sincronismos entre os diferentes tipos de transmissão.

RR (Receiver Reports) – São emitidos pelos receptores das informações revelando a qualidade. A qualidade na recepção do fluxo. Isso faz com que a fonte possa fazer alterações na transmissão baseando-se nas mensagens RR que recebem dos destinos.

SDES (Source Description) – Nessa mensagem seguem informações adicionais sobre cada participante de uma sessão RTP (*e-mail*, telefone, localização geográfica, etc.) visando exclusivamente sua identificação.

BYE – Enviado por uma das fontes quando está saindo da sessão RTP.

REPORTING INTERVAL – Mensagem fornecida a todos os participantes de uma sessão RTP contendo suas informações sobre a qualidade do fluxo de dados recebidos e enviados.

Tanto o RTP e como RTCP foram projetados para serem independentes das camadas de rede e transporte.

Diferentes tipos de pacotes RTCP podem ser encapsulados em um mesmo pacote UDP, criando um pacote RTCP composto. Desta forma, há uma redução do *overhead* causado pela inserção dos cabeçalhos IP e UDP.

O RTCP adiciona informações sobre perda de pacotes, *jitter*, atraso e nível de sinal.

O protocolo RTCP XR (FRIEDMAN 2003) possui todas as funções do RTCP mais o fator R e o MOS. O fator R está explicado em detalhes no item **4.6 Modelo E** e o MOS se encontra detalhado no item **4.1 MOS**.

2.3 Session Initiation Protocol – SIP

O SIP (ROSENBERG 2002) é um protocolo de sinalização, que estabelece, modifica e termina sessões multimídia e/ou ligações. Essas sessões podem incluir conferências de multimídia, educação a distância, chamados de telefonia IP e distribuição de arquivos com multimídia (CLAUDE 2002).

Em combinação com outros protocolos, ele é usado para descrever as características de uma sessão para seus potenciais participantes. Embora o SIP não faça nenhuma especificação quanto ao tipo de protocolo de transporte, o tráfego de voz é geralmente feito através do *Real-Time Transport Protocol* (RTP) .

O SIP suporta cinco categorias para estabelecimento e terminação das chamadas: pela localização, pela característica e disponibilidade, início de chamadas e modificação das características das chamadas.

O SIP considera separadamente a sinalização da informação transmitida. Tal separação é importante, pois a sinalização pode passar por um ou mais *proxy servers* ou *redirect servers*, enquanto que a informação propriamente dita toma um caminho mais direto. Os componentes de um ambiente de rede SIP serão apresentadas na subseção a seguir.

2.3.1 Componentes de um ambiente de rede SIP

Os componentes de um ambiente de rede SIP são: o *User Agent*, o *Proxy Server*, o *Redirect Server*, o *Registrar Server*, e o *User agente Server*. O conjunto destes componentes atuando em uma rede IP é definido como ambiente de “rede” SIP. O SIP define duas classes básicas de entidades de rede: clientes e servidores. Um cliente (também conhecido por *User Agent*) é uma aplicação que envia requisições SIP, enquanto o servidor é a entidade responsável por responder tais requisições. Os cinco componentes SIP estão descritos a seguir:

User Agent : cliente da arquitetura, ou ponto final da comunicação multimídia.

Proxy Server : Servidor que centraliza requisições e respostas SIP. Passa a realizar a sinalização como se fosse o originador da chamada, e quando a resposta lhes é enviada, ela é reenviada para o originador real. Atua de forma similar a um servidor *proxy* utilizado para acesso a *web* a partir de redes locais.

Redirect Server : recebe requisições SIP, fazendo o mapeamento para um ou mais endereços de destino e retornando a informação para seu originador. Dessa forma, o originador pode enviar a requisição diretamente para o destino desejado. Diferente do *proxy server*, o *redirect* apenas provê a informação de destino, não fazendo o envio da mensagem ao endereço final.

Registrar Server : Servidor que guarda informações de registro dos usuários, recebidas através das mensagens REGISTER. Estas informações são utilizadas para indicar em qual endereço de rede determinado cliente pode ser encontrado. O conceito de registro possibilita ao usuário a mobilidade sem a perda do serviço.

User Agent Server : recebe requisições SIP e contata o usuário de destino, gerando uma resposta ao cliente que iniciou a requisição. Na RFC 3261 (ROSENBERG 2002), apenas as funcionalidades de armazenamento e consulta de registros de usuários SIP neste servidor são descritas, ficando a critério do implementador da solução SIP a escolha da melhor tecnologia para esta finalidade.

Implementações do protocolo SIP geralmente combinam um *registrar server* com um *proxy* ou um *redirect server*, sendo pouco provável a implementação das entidades acima de forma separada. As entidades clientes também apresentam a implementação de um *user agent server (UAS)* de forma a serem capazes de responder a requisições SIP e, conseqüentemente, fazer e receber ligações. Esta característica também possibilita a conexão ponto a ponto entre clientes sem o uso de *proxy servers*.

2.3.2 Acesso ao ambiente de rede SIP

O ambiente de “rede” SIP pode ser acessado via Internet usando uma URI (Uniform Resource Identifier). A URI é uma *string* compacta para endereçar os recursos físicos ou abstratos dentro da rede. Uma SIP URI é descrita no formato *user@host* de forma similar a um endereço de e-mail. Exemplos de endereçamentos SIP são “*alias*” (ou apelido) como esta URI <sip://usuário@servidor> ou pode ser um número de telefone, como <tel://123456@inf.ufrgs.br>. A parte do host na identificação URI pode ser um domínio internet alfanumérico válido ou um endereço IP numérico.

2.3.3 Requisições e respostas SIP

O SIP funciona numa arquitetura cliente/servidor, e suas operações envolvem apenas métodos de requisição e respostas. Uma requisição SIP inicia com uma *request-line* contendo um método SIP e uma *Request-URI*. O método SIP indica o tipo de requisição que se deseja fazer enquanto que a *Request-URI* contém o endereço da entidade que se deseja contatar.

Os métodos de requisição do SIP são os seguintes: *INVITE*, *ACK*, *OPTIONS*, *BYE*, *CANCEL* e *REGISTER*. O comportamento destes métodos está descrito abaixo:

INVITE : Indica que o usuário está sendo convidado a participar de uma sessão multimídia. O corpo da mensagem pode conter uma descrição da sessão, utilizando-se o protocolo de descrição de sessão SDP (*Session Description Protocol*) (HANDLEY 1998) .

ACK : Mensagem recebida como resposta final a um *INVITE*. A requisição *ACK* pode conter o SDP de descrição da sessão negociada entre ambos os clientes. Se não contiver o SDP, o usuário chamado pode assumir a descrição dada pelo primeiro *INVITE*, se houver.

OPTIONS : Faz uma pergunta sobre quais métodos e extensões são suportados pelo servidor e pelo usuário descrito no campo de cabeçalho <To:> . O servidor pode responder a esta pergunta com o conjunto de métodos e extensões suportado pelo usuário e por ele mesmo.

BYE : Usado para liberar os recursos associados a uma ligação e forçar a desconexão da mesma.

CANCEL : Cancela uma requisição que ainda esteja pendente, ou seja, em andamento. Uma requisição é considerada pendente, se e somente se, ela não foi atendida com uma resposta final.

REGISTER : Um cliente usa este método para registrar o "alias" (apelido) do seu endereço em algum servidor SIP, que, por aceitar registro de usuários, chamamos de serviço REGISTRAR.

Dentro da arquitetura SIP, muitas vezes temos a figura de um servidor de localização, onde ficam os registros de usuários. Normalmente, para a localização destes nomes, são usadas bases de dados locais ou servidores LDAP (*Lightweight Directory Access Protocol*), onde é possível montar diretórios de usuários e seus perfis.

Para cada requisição ou resposta, temos um grupo de cabeçalhos, divididos em: cabeçalhos gerais, com informações importantes sobre a chamada; cabeçalhos de entidade, com meta-informação sobre o corpo da mensagem; e os cabeçalhos específicos, que permitem passar informações adicionais, que não couberam na linha de *status* da requisição ou da resposta.

Quando requisições são atendidas, as respostas enviadas são identificadas por números, que significam a classe da resposta. Pode-se enviar diversas mensagens provisórias antes de se enviar uma resposta definitiva. Existem seis classes possíveis de resposta: Classe 1XX, respostas temporárias ou informativas (180 *Ringing*); Classe 2XX, resposta final de sucesso (200 *OK*); Classe 3XX, redirecionamento da requisição (301 *Moved Permanently*); Classe 4XX, erros no cliente (407 *Proxy Authentication Required*); Classe 5XX, erros do servidor (501 *Not Implemented*); e Classe 6XX, erros globais na rede (600 *Busy Everywhere*).

2.3.4 Processo de Execução de Conexões SIP

Um exemplo de execução de uma conexão SIP, retirado de (GOMIDE 2004), é apresentado a seguir, mostrando as características deste processo.

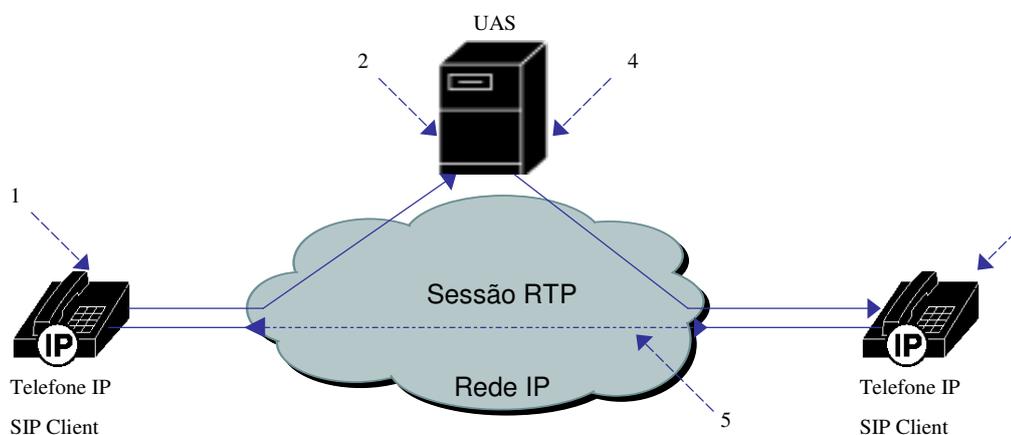


Figura 2.3: Processo de execução de uma conexão SIP

A seqüência de passos representada na figura 2.3 é explicada a seguir :

1 – O terminal chamador insere uma solicitação “*invite*” na direção do servidor de direcionamento com base na URL (requer DNS) e informando o endereço do terminal do destino;

2 – O servidor de direcionamento responde com o endereço válido do servidor “*user agent server – UAS*”;

3 – O terminal chamador estabelece outro comando “*invite*” em direção ao servidor “*UAS*”;

4 - Em direção ao terminal destinatário que confirma com “*ok*” em direção ao terminal chamador, abrindo assim a conexão IP entre os dois terminais;

5 – O terminal chamador confirma o sucesso da conexão com um comando “*acknowledgment*”, estabelecendo a sessão RTP.

A transação SIP consiste numa solicitação SIP e uma resposta apropriada. Para facilitar o casamento da solicitação e resposta encontram-se vários campos com valores idênticos , tais como: o identificador de chamado “*call ID*”, o número de seqüência do comando, o campo “*to*”, o campo “*from*” e “*tag*” (GOMIDE 2004).

O protocolo SIP utiliza algumas características dos outros protocolos, como por exemplo, os cabeçalhos, erros e regras de codificação do protocolo HTTP. O protocolo de sinalização SIP inicia uma sessão entre usuários.

2.4 O Padrão H.323

A Recomendação H.323 (ITU-TH.323 1999) é o padrão publicado pelo *International Telecommunications Union – Telecommunications Sector* (ITU-T) para comunicações multimídia em tempo real e para conferência em redes comutadas por pacotes como a Internet.

A base da recomendação é construída pelos seguintes documentos:

H.323 – Documento que define o uso do H.225.0, H.245 e outros documentos para a entrega de serviços de conferência multimídia sobre redes comutadas por pacotes.

H.225.0 – Descreve três protocolos de sinalização (RAS, Call Signaling e “Anexo G”). Este documento é derivado da recomendação Q.931 do ITU-T.

H.245 – Multimedia Control Protocol (comum ao H.310, H.323 e H.324)

E pelos seguintes documentos adicionais:

H.235 – Segurança em sistemas baseados no H.245

H.245 – Conexão com redes PSTN

H.450.x – Serviços suplementares

H.460.x – Várias extensões à H.323

H.501 – Protocolo para gerenciamento móvel e comunicação inter/intra domínio

H.510 – Usuário, terminal e serviços móveis

H.530 – Especificações de segurança para H.510

O H.323 ainda incorpora um grande número de anexos (partes do padrão), apêndices (documentos apenas para informação) e uma variedade de formatos de mídia e estruturas de dados para as aplicações, como os codificadores de áudio G.711, G.722, G.723.1, G.728, G.729, GSM e os codificadores de vídeo H.261, H.262 e H.263.

2.4.1 Componentes de um Sistema H.323

A seguir a descrição dos componentes de um Sistema H.323 :

Terminal : é o cliente da arquitetura, ou ponto final onde um fluxo de dados e sinalização H.323 é originado ou terminado. Pode ser um PC, um telefone IP, um sistema de correio de voz, etc. Um terminal deve suportar fluxos de áudio, suporte à vídeo e dados é opcional.

Gatekeeper : é um componente opcional, porém muito funcional em um sistema H.323, sendo considerado o “cérebro” deste tipo de sistema por acumular funções de gerenciamento central e controle de serviços. O *Gatekeeper* proporciona tradução de endereços, admissão e controle de acesso de terminais, gerenciamento de banda, roteamento e demais serviços de valor agregado.

MCU : Uma MCU (*Multipoint Control Unit*) habilita conferência entre três ou mais terminais, proporcionando um ponto centralizado de para este tipo de chamada. O MCU contém um MC (*Multipoint Controller*) que gerencia a sinalização de chamada e pode ter MP's (*Multipoint Processors*) para manipular comutação e processamentos multimídia.

Gateway : é uma interface entre um sistema H.323 e outras redes como PSTN, ISDN, etc, e outros sistemas H.323 (*proxy*). Um *gateway* é composto por um MGC (*Media Gateway Controller*) que manipula a sinalização da chamada e outras funções não relativas a mídia e por um MG (*Media Gateway*) que manipula a funções de mídia como tradução de codecs.

2.4.2 Protocolos H.323

A camada de rede e de transporte é transparente para a pilha de protocolos H.323, não fazendo parte da sua recomendação (ITU-TH.323 1999).

A Figura 2.4 mostra toda a pilha de protocolos H.323 numa uma rede IP.

Aplicações de Áudio e Vídeo	Gerenciamento e Controle de Terminal				Aplicações de Dados
G.7xx H.26x	RTCP	H.225.0 RAS	H.450.x Serviços Suplemen	H.245 Controle de Mídia	T.124
RTP			H.225.0 Controle de Chamada		T.125
UDP			TCP		T.123
Camada de Rede (IP)					
Camadas Física e de Enlace					

Figura 2.4: Pilha de Protocolos H.323
(DAVID 2003)

2.4.2.1 H.225 RAS

O protocolo de sinalização entre *endpoints* (terminais e *gateways*) e *gatekeepers* é o RAS (*Registration, Admission, Status*). O RAS permite o registro de *endpoints* no *gatekeeper*, o controle de admissão pelo *gatekeeper*, status e a desconexão com o *gatekeeper*. Um canal RAS é usado para a troca de mensagens do tipo RAS. Este canal de sinalização é aberto entre o *endpoint* e o *gatekeeper* antes de qualquer outro canal de comunicação ser estabelecido.

2.4.2.2 H.225 Call Signaling

O protocolo de sinalização H.225 *call signaling* é utilizado para o estabelecimento de sessão entre *endpoints* ou *endpoints* e *gatekeeper*, através da qual os dados de mídia serão transportados. A troca de mensagens H.225 é feita usando um canal de comunicação confiável. No caso de redes IP, o TCP é adotado como protocolo de camada de transporte.

As mensagens H.225 são trocadas diretamente entre os *endpoints*, caso não esteja sendo usado um *gatekeeper*. O *gatekeeper* pode funcionar como interface entre a troca de mensagens dos *endpoints*, neste caso atuando como *proxy*.

2.4.2.3 H.245 Control Signaling

O protocolo de sinalização H.245 é feito entre os *endpoints*. Permite a abertura e o fechamento de canais lógicos de comunicação para o transporte

de mídia, além da troca de características dos terminais como *codecs* suportados, por exemplo.

2.4.3 Processo de Execução de Conexões H.323

Os passos de estabelecimento de uma chamada H.323, troca de mídia e término da chamada são descritos a seguir. O exemplo assume o estabelecimento de sessão entre terminais H.323 (T1 e T2) conectados através de um *gatekeeper*.

O estabelecimento de uma chamada H.323, baseado em (DAVID 2003) e ilustrado na Figura 2.5, é descrito a seguir.

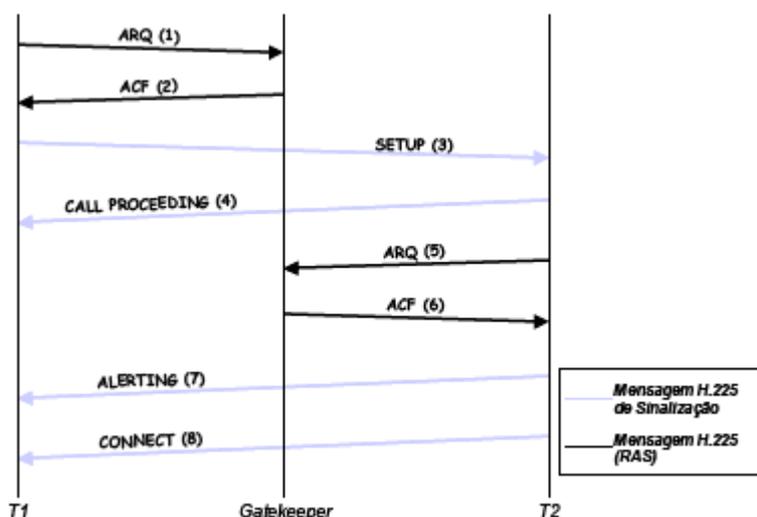


Figura 2.5: Estabelecimento da chamada
(DAVID 2003)

1 - T1 envia uma mensagem RAS ARQ para o canal RAS do *gatekeeper*, solicitando o seu registro. Neste exemplo, T1 solicita sinalização direta com o outro *endpoint* (T2).

2 - O *gatekeeper* confirma a admissão de T1, enviando uma mensagem ACF para T1.

3 - T1 então envia uma mensagem H.225 *call signaling* para estabelecimento de sessão (*setup*) com T2.

4 - T2 responde com uma mensagem de estabelecimento de sessão em andamento.

5 - Em seguida, T2 solicita registro ao *gatekeeper*, através de uma mensagem ARQ.

6 - este registro é confirmado pela mensagem ACF enviada pelo *gatekeeper*.

7 - T2 alerta T1 sobre o estabelecimento de conexão.

8 - Finalmente, T2 confirma o estabelecimento da chamada.

A Figura 2.6 mostra um exemplo de sinalização H.245, com troca de características entre *endpoints* e negociação de portas de mídia.

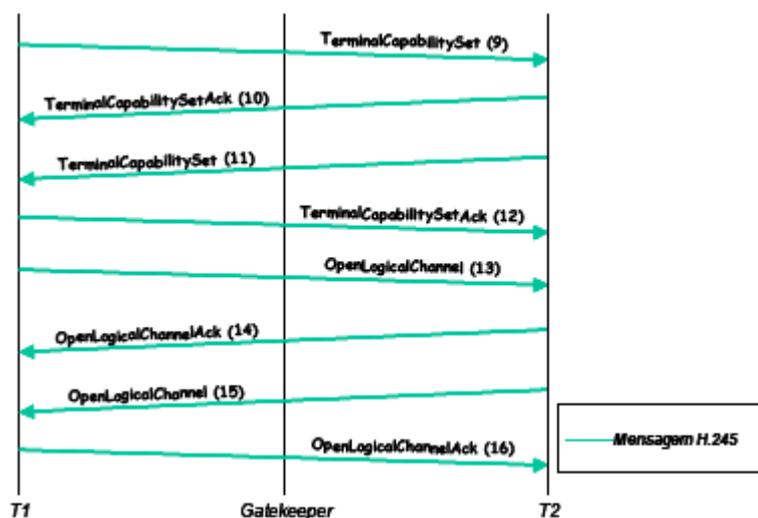


Figura 2.6: Sinalização de Controle
(DAVID 2003)

9 - Um canal de controle H.245 é estabelecido entre T1 e T2. T1 envia uma mensagem para T2 informando suas características,

10 - com confirmação de T2 para T1 em seguida,

11 - T2 envia para T1 as suas características,

12 - com confirmação posterior de T1.

13 - T1 abre um canal lógico de comunicação de mídia, enviando uma mensagem H.245 *OpenLogicalChannel*, informando também o canal de comunicação do RTCP.

14 - T2 confirma o estabelecimento do canal, informando qual a porta que T2 estará pronto para receber pacotes RTP de T1.

15 - De forma análoga, T2 abre o seu canal de comunicação com T1,

16 - informando a sua porta de recebimento de pacotes RTCP. T1 informa em seguida qual a porta estará recebendo pacotes RTP vindos de T2 .

Após esta troca de mensagens, um canal bidirecional de mídia é estabelecido entre T1 e T2, com pacotes RTP e RTCP passando a trafegar entre os dois, transportando a mídia e informações de controle de QoS.

O término de uma sessão é ilustrado na Figura 2.7. Quando um dos *endpoints* decide encerrar a comunicação, uma mensagem de encerramento de sessão é enviada.

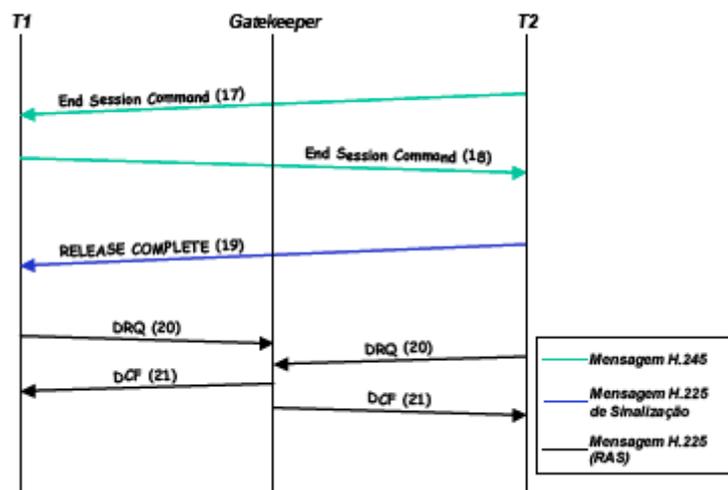


Figura 2.7: Término da Chamada
(DAVID 2003)

17 - T2 inicia o encerramento da chamada.

18 - T1 libera a chamada e confirma o seu término através de mensagem H.245 para T2.

19 - T2 envia então uma mensagem de sinalização H.225 informando que a conexão foi desfeita.

20 - Posteriormente, T1 e T2 solicitam ao *gatekeeper* sua desconexão, através de mensagens RAS DRQ.

21 - O *gatekeeper* confirma finalmente a liberação dos *endpoints*, através de mensagens DCF.

Funcionalmente o H.323 e o SIP são similares. Ambos são protocolos “*peer-to-peer*” para o transporte de mídia e provêm o controle da sinalização das chamadas.

A seguir são apresentados os protocolos “mestre-escravo” de controle de *Gateway*.

2.5 O Padrão MGCP

O MGCP (*Media Gateway Control Protocol*) é um protocolo do controle de *gateway* (ARANGO 1999) que segue a filosofia de uma arquitetura de interconexão aberta de módulos de software que controlam chamadas de voz e dados em uma rede convergente. Esta interconexão é conhecida como

Softswitch. Isto facilita o controle de cada *gateway* de VoIP como uma entidade separada (GOMIDE 2004).

MGCP é um protocolo do tipo *master-slave* que coordena as ações de *gateways* (Figura 2.8). O controlador do *gateway*, na nomenclatura de MGCP, recebe consultas dos *gateways*, sendo chamado também de agente da chamada. O agente da chamada controla a inteligência sinalizando *call-related* do controle, quando o *gateway* informar o agente da chamada de eventos do serviço. O agente da chamada instrui o *gateway* para criar e encerrar conexões quando as chamadas são geradas. Na maioria dos casos, o agente da chamada informa aos *gateways* para iniciar uma sessão de RTP entre dois *endpoints* (GOMIDE 2004).

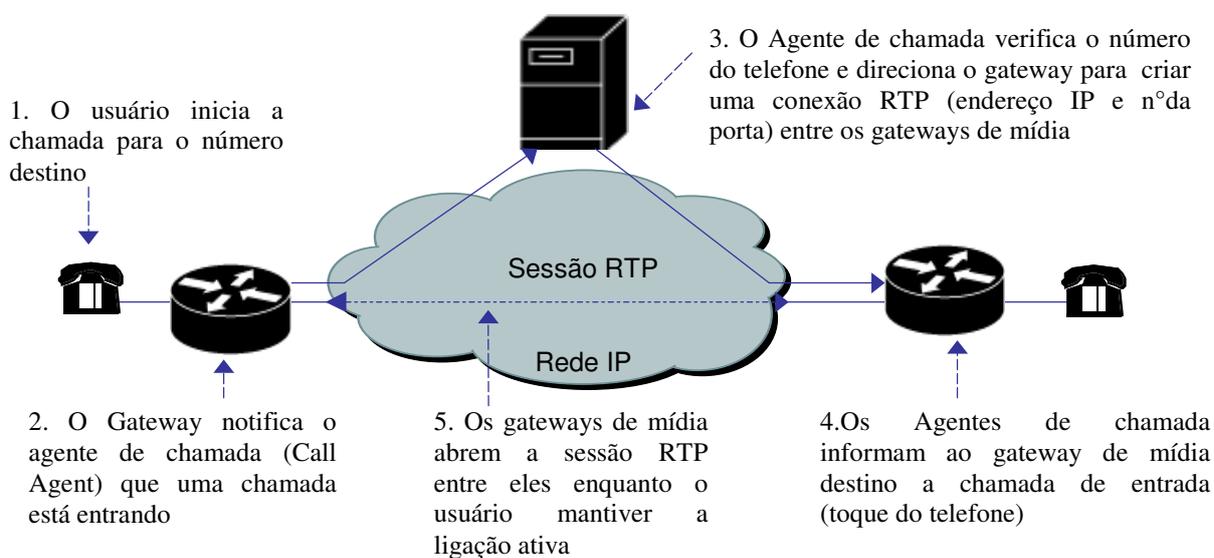


Figura 2.8: Controle de ações do MGCP
(GOMIDE 2004)

As sinalizações executadas pelo agente e pelos *gateways* de chamada estão no formulário de mensagens estruturadas dentro dos pacotes do UDP. O agente da chamada e os *gateways* de mídia têm recursos de retransmissão para estas mensagens.

2.6 O Padrão H.248/MEGACO

O H.248/MEGACO é resultado de um desenvolvimento conjunto entre o IETF e o ITU, resultando num melhoramento do MGCP (ITU-TH.248 2002) (GROVES 2003).

A arquitetura de MEGACO define os *gateways* de mídia que fornecem a conversão dos meios e as origens das chamadas, quando os controladores do *gateways* fornecerem o controle da chamada.

O MEGACO possui as mesmas exigências que MGCP e, em conseqüência, há algum esforço em compatibilizar os protocolos. O MEGACO define uma série das transações coordenadas por um controlador de *gateway* de mídia para o estabelecimento de sessões da chamada. O foco preliminar do MEGACO é ser o promotor de estabilização para o equipamento de telefonia IP.

O H.248/MEGACO segue um modelo centralizado, concentrando a inteligência, ao contrário do SIP, que usa um modelo distribuído, utilizando a inteligência dos equipamentos das pontas. Justamente por ter esta forma centralizada, o H.248 possui uma ótima compatibilidade com o sistema de sinalização da telefonia convencional (IEC 2005).

O H.248/MEGACO é uma evolução do MGCP e portanto ocupa o mesmo lugar do MGCP no controle de *gateways*. Na figura 2.9 é mostrado um exemplo de rede onde o *Media Gateway Controller* utiliza MGCP ou MEGACO para controlar os *Gateways* de Mídia.

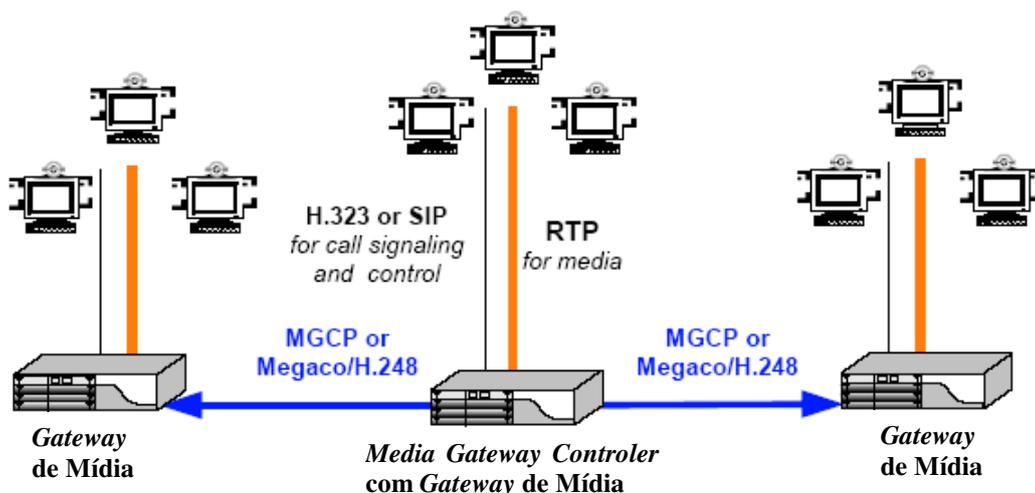


Figura 2.9: MGCP ou H.248/MEGACO no controle de Gateways
(RADVISION 2002)

O objetivo desse tópico foi justamente apresentar as principais características do protocolo H.248/MEGACO. Não faz parte desse estudo, entretanto, uma descrição aprofundada sobre o referido protocolo. Informações detalhadas podem ser encontradas em (ITU-TH.248 2002) e em (GROVES 2003).

2.7 Conclusões do capítulo 2

Este capítulo teve como propósito apresentar os principais protocolos abertos utilizados em voz sobre IP. Cabe ressaltar que ainda existem soluções particulares com especificações não acessíveis. Como exemplo pode-se citar o aplicativo *peer-to-peer* chamado Skype. Maiores esclarecimentos sobre o Skype podem ser obtidos em (BASED 2004).

Como visto o protocolo RTP, responsável pelo transporte da mídia, não apresenta informações sobre qualidade de voz. Os protocolos RTCP e RTCP-XR apresentam estas informações, mas não são de uso obrigatório.

Segundo (DALGIG 1999), numa comparação entre o padrão H.323 e o SIP, observa-se que o H.323 possui uma definição mais rigorosa dos serviços suplementares e uma melhor compatibilidade e interoperabilidade com a telefonia convencional. Observa-se também que, por outro lado, o SIP traz como vantagens a flexibilidade para adicionar novas características e a maior facilidade de implementação e depuração.

Os protocolos de controle de *Gateway* nasceram pela necessidade das redes IP se interconectarem com sistemas de telefonia convencional. Os protocolos de controle de *Gateway* fornecem um controle remoto dos fluxos de mídia para o trânsito entre redes IP e redes de telefonia convencional.

A diferença fundamental entre os protocolos de controle de *Gateway* e os protocolos de sinalização é que estes últimos iniciam e gerenciam as chamadas, e os primeiros definem como os fluxos de mídia destas chamadas são iniciados e conduzidos entre redes IP e outras redes.

Todas estas características dos protocolos de VoIP são levadas em consideração no método a ser apresentado no capítulo 8.

3 FATORES DE QUALIDADE DE VOZ

A qualidade de uma conexão de voz é afetada diretamente por diversos fatores. O tráfego de voz é extremamente sensível a atrasos e um pouco mais tolerante à perda de pacotes do que o transporte de dados. Também pode-se observar que certos fatores são independentes da tecnologia adotada como por exemplo, aspectos do comportamento humano ou ruídos externos durante a transmissão ou recepção da ligação.

São apresentados a seguir os principais fatores dependentes da tecnologia adotada para o transporte de voz.

3.1 Perda de Pacotes

As duas principais causas de perdas de pacotes em redes IP são os descartes de pacotes por roteadores e *switch routers*, provocados por erros, congestionamentos, etc, e por perdas ocorridas na camada 2 (PPP - *Point – to – Point protocol*, *Ethernet*, *Frame Relay*, ATM, etc.) durante o transporte dos mesmos (SAWASHIMA 1997).

Outro ponto importante é que, para aplicações multimídia em tempo real, pacotes que chegam muito atrasados em relação ao instante de tempo em que deveriam ser reproduzidos no lado receptor, tornam-se inúteis, e conseqüentemente são descartados, sendo considerados pacotes perdidos sob o ponto de vista do receptor. Portanto, ao se avaliar a qualidade de voz, não basta computar os pacotes perdidos sob o ponto de vista da rede de transmissão, mas também em relação à camada de aplicação (MIRAS 2002).

Em aplicações como VoIP, a perda de pacotes impacta na qualidade das ligações, sendo caracterizada pelo surgimento de cortes ou períodos de silêncio na fala. Os efeitos da perda de pacotes variam de acordo com a frequência e a distribuição da perda, além do tipo de *codec* utilizado. A perda de apenas 1% dos pacotes de uma ligação com *codec* G.711 pode degradar significativamente a qualidade da mesma. *Codecs* que utilizam técnicas de compressão podem apresentar uma degradação ainda maior com a mesma perda de pacotes (THORELL 2005).

No caso de uma pequena ocorrência de perda de pacotes a maneira mais simples de tratamento é reproduzir a amostra anterior. Para perdas maiores, foram desenvolvidas outras técnicas, como a técnica chamada *Packet Loss*

Concealment (PLC), que faz uma interpolação entre as últimas amostras de som recebidas prevendo assim a amostra perdida.

3.2 Perda em Rajadas

A ocorrência de determinados fatores na infra-estrutura de redes durante um período de tempo longo podem causar a perda de vários pacotes consecutivos. A perda de pacotes em rajada provoca a degradação na voz, de modo que até mesmo *codecs* com algoritmos de correção de perdas não conseguem minimizar seu efeito. Cox e Perkins compararam o impacto da perda de pacotes distribuída randomicamente ao longo de uma ligação em relação à perda em rajadas, utilizando os *codecs* G.711 (COX 1999).

Foi constatado que, para taxas de perda muito baixas, a qualidade das ligações com perda de pacotes em rajadas foi mais bem avaliada subjetivamente do que as ligações com perdas randomicamente distribuídas. Tal efeito ocorre pois uma rajada de curta duração tem um impacto não muito maior do que a perda de um único pacote, além de existirem grandes períodos de transmissão normal entre as perdas. No entanto, quando a perda ultrapassa o limite de 3%, pode-se observar que a perda de pacotes em rajadas tem um impacto muito superior do que a perda de pacotes distribuída aleatoriamente, como pode ser visto na Figura 3.1 (DAVID 2003).

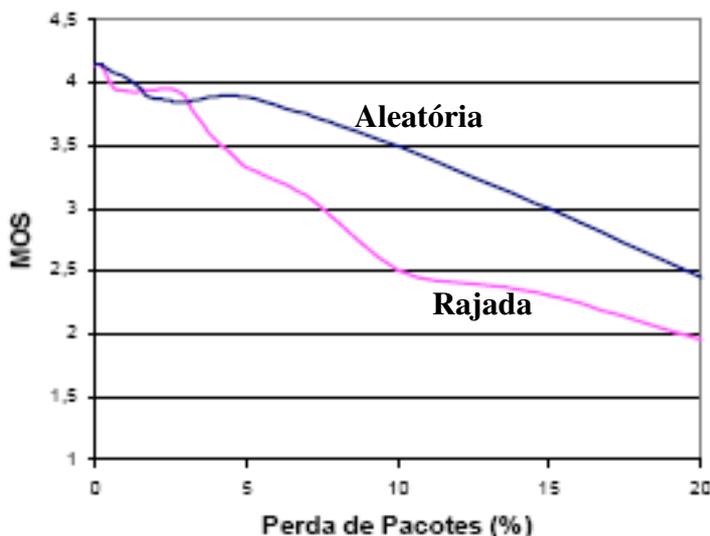


Figura 3.1: Perda por Rajada versus perda Aleatoriamente Distribuída (DAVID 2003)

A figura 3.2 apresenta um exemplo de um gráfico com medições de qualidade de voz, utilizando o método PESQ de avaliação de qualidade de voz, em um ambiente real usando *codecs* G723.1 e G729. Este exemplo foi retirado de (FERNANDES 2003). O método PESQ é apresentado em detalhes no item **4.2.4 PESQ**.

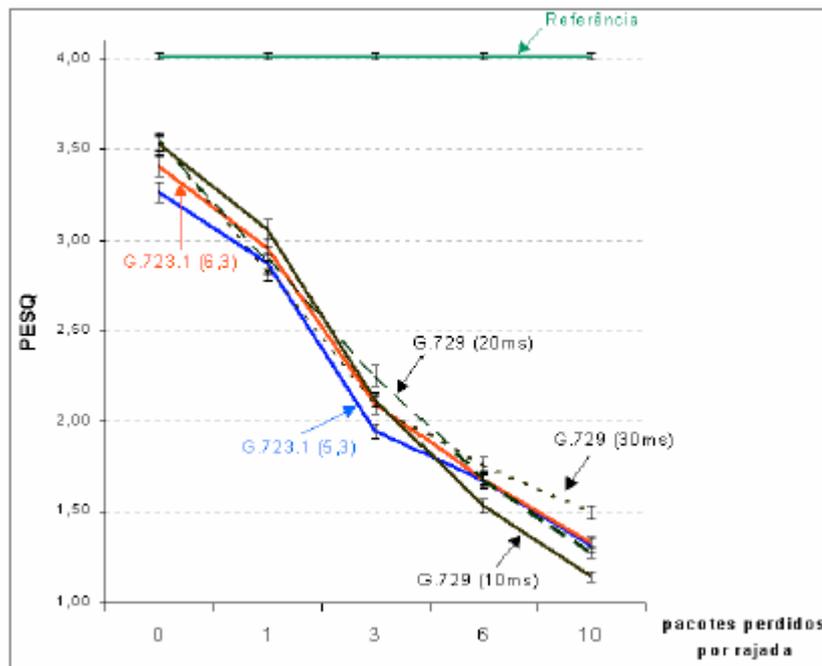


Figura 3.2: Valores PESQ para comparação entre *codecs* (FERNANDES 2003)

Independente de qual *codec* adotado, nota-se a degradação da qualidade do sinal na medida que aumenta a quantidade de pacotes perdidos por rajada.

3.3 Atraso

Atraso ou latência refere-se ao tempo que o pacote leva para percorrer todo o percurso de uma rede do ponto de partida do pacote ao ponto de destino. Em telefonia, a latência é o tempo médio que a voz leva para chegar até o ouvinte. Baixos níveis de latência são imperceptíveis em uma conversação, no entanto valores maiores proporcionam desconforto, tornando as falas desencontradas, dependendo do grau de interatividade do diálogo. Os grandes valores de latência não degradam a qualidade do som de uma chamada telefônica, mas podem causar uma falta da sincronização entre os clientes.

O padrão ITU-T G.114 (ITU-TG114 2000) indica os limites para os tempos de transmissão fim-a-fim dividindo-os em três categorias :

- De 0 até 150 ms – Aceitável para a maioria das aplicações;
- De 150 até 400 ms – Deve-se estar atento ao impacto do tempo de transmissão na qualidade da aplicação. A comunicação entre dois pontos que dependa de um enlace satélite se encontra contida nesta faixa.

- Acima de 400 ms – De forma geral é inaceitável para aplicações em rede. Entretanto, casos excepcionais com a necessidade de dois enlaces satélites podem estar nesta faixa.

O padrão ITU-T G.114 (ITU-TG114 2000) sugere um atraso máximo aceitável de 150ms em um sentido (*one-way delay*). Entretanto, o impacto que o atraso causa na qualidade da voz depende de outros fatores, como o grau de interatividade da própria conversa. Para garantir que a latência permaneça abaixo de 150 ms, deve ser feita uma verificação sobre as causas da latência (KAMIENSKI 2000).

Ao projetar uma rede de multi-serviço, o total de atraso sofrido por um sinal ou pacote é a somatória de todas as latências em todos os pontos da rede. Este atraso é formado por um componente fixo e outro variável (CISCO 2005).

A amostragem e codificação dos sinais de voz analógicos em pacotes de dados, seguido opcionalmente de compactação, gasta um tempo conhecido como *packetization delay*. Este tempo é fixo e depende diretamente do tipo de codificação e *codec* utilizado, variando conforme sua complexidade. Por exemplo, um *codec* G711 a 64Kbits/s gasta 20 ms para o empacotamento de 240 bytes. Já um *codec* G723.1 a 5,3Kbits/s gasta 30ms para o empacotamento de 20bytes (BOGER 2005). O tempo decorrente da descompactação e decodificação no lado receptor é sempre muito inferior, mas também deve ser contabilizado (BOGER 2005).

Uma outra fonte de latência vem da inserção dos bits dos pacotes de dados nos links físicos nos equipamentos de interconexão (PERCY 1999). Este atraso, também conhecido como *serialization delay*, é inversamente proporcional à velocidade do link. Ou seja, quanto mais rápidos os links, mais baixa é a latência. Este valor depende da tecnologia de conexão, bem como o seu método de acesso. Por o exemplo, leva-se o tempo de 125 microssegundos para inserir 1 byte em um circuito de 64Kbits/s. O mesmo byte colocado em um circuito Oc-3/stm-1 (155Megabits/s) é transmitido em 0,05 microssegundos. Embora este atraso não seja variável (independente da largura de banda usada), mantendo o número de link pequeno e usando taxas mais altas, o atraso é reduzido.

O tempo de propagação de um sinal elétrico para percorrer o comprimento de um meio físico também constitui um componente fixo de atraso, conhecido como atraso de propagação. A velocidade de propagação é sempre mais lenta do que a velocidade de luz. Há sempre um atraso na propagação; entretanto, o atraso somente é considerado quando o sinal (ou o pacote) viaja a uma longa distância. A fórmula para cálculo do tempo de propagação é a seguinte:

$$\text{Tempo de propagação} = L / (299300 \times 6)$$

Onde “L” é o comprimento do enlace em quilômetros e o resultado é dado em microsegundos.

3.4 Variação de Atraso ou *Jitter*

Componentes variáveis de atraso são relacionados a problemas de congestionamento na rede, ou mesmo a mudanças na sua topologia. Roteadores e comutadores no caminho entre a origem e o destino da conexão geram atrasos variáveis, dependendo da velocidade de comutação de pacotes dos equipamentos, bem como dos retardos decorrentes de enfileiramento e armazenamento dos pacotes durante a sua transmissão. As cargas do tráfego da rede resultam num enfileiramento variável. O tamanho dos *buffers* usados em uma fila é geralmente um parâmetro configurável; quanto menor o tamanho do *buffer*, menor será a latência. A implementação de um esquema de QoS reduz o fator de atraso no *buffer*, dando prioridade aos pacotes de voz e, por consequência, reduzindo o seu tempo de espera em filas de transmissão. Entretanto, o atraso também pode ser originado pela quantidade de tráfego que o roteador está tentando passar em um determinado enlace, e conseqüentemente aumentará com a carga da rede. Com isso, torna-se necessário reservar a largura de banda adequada e os recursos para o tráfego de voz. Se a fila usada para o tráfego da voz não estiver suportando a demanda e essa fila estiver crescendo muito, o resultado é um aumento da latência.

O tempo de transferência do pacote num roteador ou num *buffer de switch* para a devida interface de saída pode ser outra origem de latência. Embora este atraso seja geralmente pequeno, a arquitetura do roteador ou do *switch* deve ser um fator a se considerar. Se um pacote deve ser armazenado como parte de seu processamento, isso irá repercutir em uma maior latência.

Como já visto, o atraso de um pacote de voz da origem até o destino não é constante, pois possui componentes fixos e variáveis. Esta variação de atraso, conhecida como *jitter*, pode ser definida como o valor médio de todas as diferenças de atrasos incrementais, do início da medição até o pacote atual. Os pacotes mais recentes têm um peso maior do que o dos pacotes menos recentes (ROWE 1999).

O *jitter* é causado principalmente pela variação do tempo de enfileiramento devido a mudanças dinâmicas nas cargas de tráfego da rede.

Esta variação de atraso geralmente é solucionada através da utilização de *buffers* no lado receptor de compensação de *jitter*. Antes de serem passados para reprodução pelo *codec*, os pacotes são reordenados e seu compasso reajustado de forma a receberem os tempos relativos originais do momento da transmissão. *Buffers* maiores permitem compensar variações maiores, mas aumentam o atraso total da reprodução. Para minimizar esse problema de atraso, a maioria dos equipamentos implementa um *jitter buffer* adaptativo cujo tamanho é ajustado de acordo com o *jitter* observado na rede, operando entre uma faixa de valores mínimos e máximos (BOGER 2005).

3.5 Eco

A qualidade de uma ligação telefônica também é influenciada pelo efeito de eco. O eco não é considerado problema quando ocorre com defasagem de 16 a 20 ms em relação ao som original. Pelo contrário, fornece ao transmissor a

sensação que a conexão ainda está estabelecida, e que está sendo recebida pela outra ponta. No entanto, para valores superiores a 32 ms, surge uma sensação de desconforto para o usuário (MIRAS 2002).

A ocorrência de eco pode ter causa acústica e/ou elétrica.

A causa acústica é devida ao retorno de sinal de um alto-falante ou fone de ouvido para um microfone, através do ambiente. Este tipo de eco não recebe tratamento (FERNANDES 2003).

O eco elétrico ocorre apenas em ambientes analógicos. Várias podem ser as causas do eco elétrico. A principal delas é a existência de híbridas de acoplamento e separação de sinais entre trechos com 2 (a linha do assinante) e 4 fios (trancos telefônicos com conexão analógica de 4 fios E&M). Nestas híbridas ocorre retorno do sinal que chega, devido a imperfeições de acoplamento (CHONG 2005). A híbrida também é encontrada dentro dos aparelhos telefônicos analógicos, para acoplar a linha de dois fios ao sistema de quatro fios, de cápsulas para fala e escuta. Outras fontes de eco por causa elétrica, podem ser a indução entre os condutores internos do fio que liga o aparelho telefônico ao fone ou o inadequado casamento de impedâncias entre a interface analógica da central telefônica e o aparelho (FERNANDES 2003).

O eco elétrico pode receber tratamento através de supressores ou de canceladores de eco. Os supressores são dispositivos que após detectarem sinal sendo transmitido num sentido, atenuam no sentido inverso o retorno deste sinal, praticamente inibindo a percepção do sinal de eco. Os canceladores de eco são mais sofisticados que os supressores, e normalmente se encontram na porção digital da rede. Sua função é subtrair do sinal recebido o sinal coincidente que foi transmitido anteriormente (FERNANDES 2003).

3.6 Momento da falha

O momento da ocorrência da falha também influencia na avaliação da qualidade de voz de uma determinada conexão. O cérebro humano dá maior prioridade a informações mais recentes, (BADDELEY 1997). Assim ligações com falhas ocorridas no seu final tendem a ser mais mal avaliadas do que outras. O laboratório de pesquisas da AT&T também realizou um estudo que comprova este efeito : a qualidade de uma mesma mensagem foi considerada pior na medida que um mesmo ruído era introduzido cada vez mais próximo do final desta mensagem (ROSENBLUTH 1998).

Outro estudo, realizado nos laboratórios da France Telecom demonstra que melhorias nas condições da ligação não são percebidas de imediato. É necessário um período de aproximadamente 5 segundos de diferença em relação ao momento de transições da qualidade da ligação de boa para ruim e de 15 segundos em transições de qualidade ruim para boa (FRANCE 2000).

3.7 Peculiaridades dos codecs

Os codificadores de voz, normalmente conhecidos como vocoders, no lado transmissor, convertem os sinais analógicos em sinais digitais e transformam estes sinais em pacotes de dados, utilizando um determinado algoritmo de codificação. No lado receptor eles realizam a função inversa. Alguns algoritmos de *codecs* reduzem a banda necessária para a transmissão através da supressão de informações redundantes ou menos relevantes. A taxa de compressão obtida é determinada por fatores como, por exemplo, qualidade da voz requerida, capacidade computacional de equipamento para a codificação e atraso máximo tolerado. Quanto maior a redução de largura de banda obtida, maior o custo computacional do *codec*, conseqüentemente, maior o atraso gerado pelo processo de codificação. *Codecs* que ao invés de transmitir uma representação do formato da onda do sinal de voz, transportam uma representação subjetiva do som original, para obter taxas reduzidas de transmissão, tem maior sensibilidade a fatores de degradação como perda de pacotes, por exemplo. Ao se avaliar a qualidade do serviço VoIP, tais fatores devem ser levados em consideração.

Alguns codecs têm a capacidade de reconhecer períodos de silêncio (VAD :*Voice Activity Detection*) e não transmitem nada neste período, gerando economia na banda consumida durante a transmissão de até 50% (BOGER 2005)(PRACHT 2001). O *codec*, após detectar um período de silêncio, pode perder o início de uma palavra. Este efeito é conhecido como FEC (*front end clipping*) (ROSENBLUTH 1998). O tempo de espera para se identificar um período de silêncio é conhecido como HOT (*hold-over time*). O HOT, quando é excessivo, reduz a economia de banda obtida. Por outro lado, quando o HOT é muito pequeno, a reprodução gerada é muito picotada, pois qualquer pausa entre palavras é considerada um período de silêncio (ROSENBLUTH 1998).

O VAD, quando há longos períodos de silêncio, pode causar a sensação de perda de conexão por parte do usuário receptor. Para minimizar este efeito, o *codec* do lado receptor pode gerar um ruído de conforto, dando a sensação ao usuário que a conexão permanece ativa (ROSENBLUTH 1998).

4 MÉTODOS DE AVALIAÇÃO DA QUALIDADE DE VOZ

Este capítulo trata dos métodos de avaliação da qualidade de voz oferecida por uma rede IP. Num primeiro momento os métodos são apenas descritos, sem nenhuma avaliação crítica. No final do capítulo, após a apresentação de todos os métodos, é realizada uma análise crítica e comparativa sobre os mesmos.

Estes métodos podem ser divididos em dois grupos : os métodos subjetivos, baseados na avaliação de pessoas através da audição, e métodos objetivos, baseados em modelos matemáticos. Os métodos objetivos podem ser intrusivos ou não intrusivos (SUN 2004). A figura 4.1 mostra a classificação destes métodos.

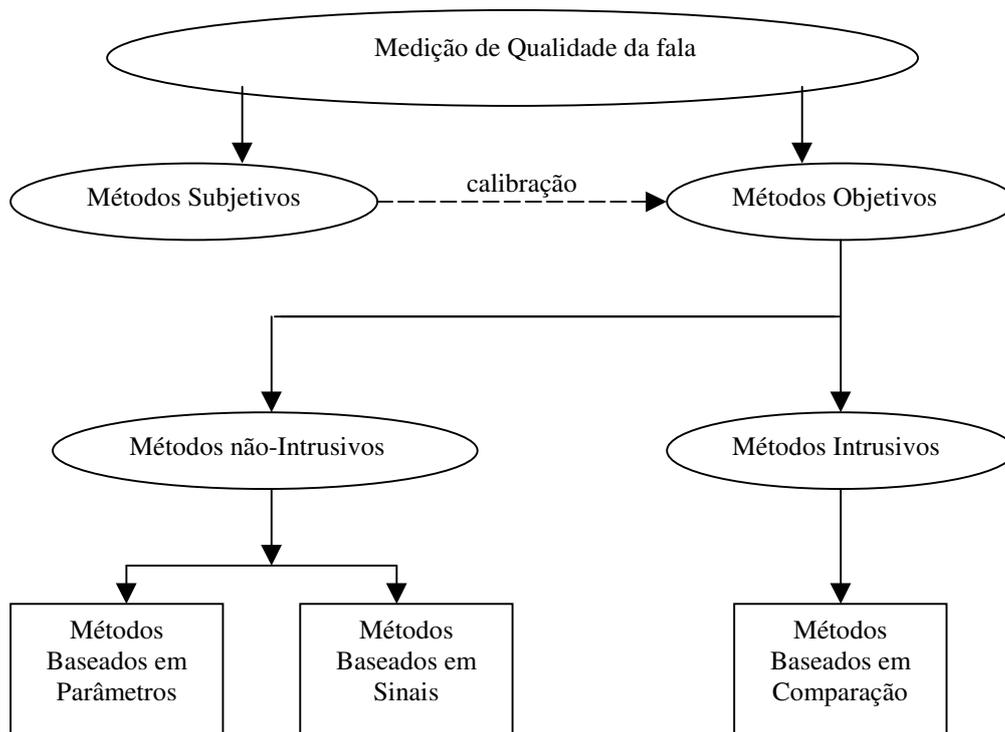


Figura 4.1: Classificação de Métodos de Medição
(SUN 2004)

Como se vê na figura 4.1, os métodos não intrusivos podem ser baseados em parâmetros ou em sinais, e os métodos intrusivos são baseados em

comparação. A figura 4.2 mostra a forma de coleta de dados destas três categorias de métodos objetivos.

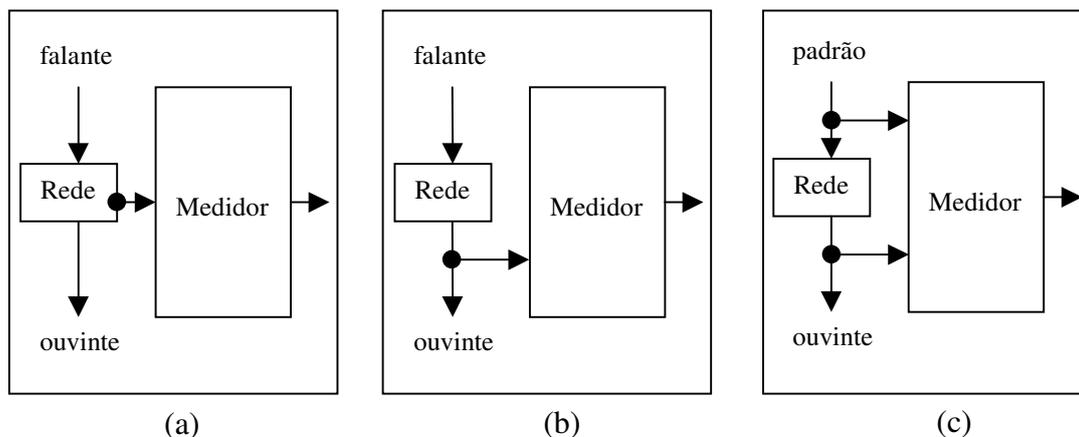


Figura 4.2-Métodos baseados em:(a) Parâmetros, (b) Sinais, (c) Comparação (SUN 2004)

O método subjetivo mais comum é o teste ACR (*Absolut Category Rating*) de onde surge o MOS (*Mean Opinion Score*). Em alguns casos é utilizado o DCR (*Degradation Category Rating*) que gera o DMOS (*Degradation Mean Opinion Score*). Ainda existe o CCR (*Comparison Category Rating*). Este último, o CCR, compara pares de arquivos e produz o CMOS (*Comparison Mean Opinion Score*) (SUN 2004). Os métodos ACR e DCR são apresentados em detalhe nos itens 4.1.1 e 4.1.2 respectivamente.

4.1 Mensuração Subjetiva da Qualidade da Fala

Os métodos subjetivos servem como referência para os métodos objetivos.

A recomendação P.800 do ITU-T (ITU-TP800 1996) descreve vários métodos e procedimentos subjetivos de avaliação de qualidade da fala. Segundo (SUN 2004), dos métodos descritos na recomendação P.800 os mais comumente usados são o ACR e o DCR, apresentados a seguir.

4.1.1 ACR – MOS

Para um teste de escuta com taxa de categoria absoluta (*Absolute Category Rating* - ACR), ouvintes não treinados escutam mensagens pré-gravadas e opinam quanto a qualidade delas. A médias da opinião destes ouvintes dá origem ao MOS.

Assim, o MOS, *Mean Opinion Score*, é um método subjetivo, baseado na opinião de um conjunto de avaliadores, que escutam individualmente, e várias vezes, ligações com mensagens pré-gravadas. O valor médio destas

avaliações, que fica entre 1 e 5, representa a qualidade da ligação. A recomendação ITU P.800 (ITU-TP800 1996) descreve este método e a recomendação P.830 (ITU-TP830 1996) apresenta os testes mais específicos para uma análise subjetiva de codificadores de áudio.

A quantidade mínima recomendada de avaliadores é de 100 pessoas e a lista de sentenças (de 2 a 3 segundos cada) recomendada é de no mínimo 2 e no máximo 5 (FERNANDES 2003).

A escala de valores para a avaliação de qualidade da conexão e da clareza de voz obtida é a seguinte :

- 1 = péssima : impossível de entender, independente do esforço realizado
- 2 = Ruim : esforço considerável exigido
- 3 = Razoável : esforço moderado exigido
- 4 = Bom : necessário prestar atenção, esforço não significativo exigido
- 5 = Excelente : completo relaxamento, nenhum esforço exigido

Para ligações VoIP, os valores considerados satisfatórios ficam entre 3,5 e 4,2 (ITU-TP830 1996).

Os testes de MOS, justamente por serem subjetivos, apresentam algumas desvantagens. Fatores como o estado de espírito do avaliador, o idioma utilizado, o perfil do diálogo (coloquial, técnico, com números, etc) influenciam consideravelmente no resultado final (ITU-TP830 1996). Para minimizar estes fatores, é necessário um grande número de avaliadores, o que torna o processo caro e complexo. Isto despertou a busca por métodos objetivos e automatizados.

4.1.2 DCR – DMOS

Quando amostras de voz de boa qualidade são avaliadas, O ACR tende a ser ineficiente para detectar pequenas diferenças. Nestes casos pode se usar então a taxa de categoria de degradação (*Degradation Category Rating - DCR*). O procedimento DCR usa uma escala de incômodo e uma referência de qualidade. Os ouvintes avaliam o nível de incômodo ou degradação comparando as amostras de voz com uma referência. Os níveis de degradação correspondem a escala mostrada a seguir (SUN 2004):

- 5 = som inaudível
- 4 = som audível mas não incômodo
- 3 = som levemente incômodo
- 2 = som incômodo
- 1 = som muito incômodo

A média de opiniões dos ouvintes no teste DCR é chamada de DMOS (*Degradation Mean Opinion Score*).

4.2 Mensuração Objetiva Intrusiva da Qualidade da Fala

Os métodos objetivos intrusivos normalmente usam dois sinais de entrada no Sistema Medidor, um sendo o sinal de referência e o outro o sinal degradado retirado da saída da rede em teste. Estes métodos são chamados intrusivos porque injetam sinais de teste (o sinal de referência) na rede. Estes métodos são mais acurados para medidas de qualidade percebida de voz fim-a-fim e são inadequados para monitoração de tráfego da rede.

A seguir são apresentados os métodos objetivos intrusivos PSQM, PSQM+, MNB, PAMS e PESQ.

4.2.1 PSQM e PSQM+

O método PSQM (*Perceptual Speech Quality Measurement*), tem por objetivo reproduzir a percepção humana dos sons para o julgamento da qualidade de voz. Este método de apuração foi considerado consistente e passou a ser conhecido como a recomendação P.861 do ITU-T (ITU-TP861 1996) (BEERENDS 1994).

O PSQM faz a comparação entre os sinais da fonte e o resultante da saída do *codec*. Ele é um processo matemático que consegue prever resultados de testes subjetivos, como os do padrão P.830. A escala de valores é diferente da adotada pelo MOS. O esquema de funcionamento do modelo PSQM, usa a gravação de mensagens com as mesmas características específicas para testes MOS do padrão ITU P.830. Esta mensagem é processada por um codificador. O sinal original e o sinal codificado são submetidos ao algoritmo de comparação do PSQM. As comparações são realizadas em pequenos segmentos sincronizados, levando em consideração não apenas o espectro, como também a sensibilidade auditiva humana, como frequência e *loudness*. A comparação assume valores de zero a infinito, representando a distância entre os sinais de entrada e saída dos codificadores. O valor zero representa sinais idênticos, sem nenhuma degradação. Quanto mais alto é o valor gerado pelo PSQM, maior é o nível de distorção. Na prática os valores máximos de PSQM variam entre 15 e 20 (ANDERSON 2002).

O PSQM é adequado para avaliar o tipo de *codec* e propriedades específicas da mensagem, tais como idioma.

O método PSQM não é adequado para medir a qualidade da voz em situações que envolvem ruídos de fundo, fontes simultâneas de mensagens, atraso, *jitter* ou perda de pacotes.

Como o PSQM não leva em consideração fatores que causam distorção do sinal original, como a perda de pacotes, ele tende a reportar uma qualidade melhor que uma interpretação humana faria. Esta característica deu origem ao aparecimento do PSQM+ (BEERENDS 1997).

O PSQM+ incorporou alterações que o tornam sensível a falhas relativas a perda de pacotes e variação de atraso, dando resultados mais próximos aos do MOS obtidos em testes subjetivos (ANDERSON 2002).

Em 1999, a *KPN Research* aprimorou o PSQM+, para melhorar seu funcionamento em testes sob condições reais de rede. Esta nova versão recebeu o nome de PSQM99 (FERNANDES 2003).

4.2.2 MNB

O MNB (*Measuring Normalizing Blocks*) é uma técnica alternativa ao PSQM, baseada num relatório do grupo de estudo 12 do ITU-T (ATKINSON 1997). O método MNB se adequa melhor para medir o impacto na clareza da voz porque leva em consideração outros fatores não tratados pelo PSQM original, tais como erros nos canais de comunicação ou *codecs* com taxas de transmissão inferiores a 4 Kbits/s (VORAN 1999).

4.2.3 PAMS

O método PAMS (*Perceptual Analysis Measurement System*) foi desenvolvido pelo mesmo centro de pesquisa que fez o PSQM, o *KPN Research*, e foi aprimorado pela *British Telecommunications* (BEERENDS 2000).

O PAMS é semelhante em alguns aspectos ao PSQM, porém utiliza técnicas diferentes para processamento do sinal e simulação da percepção humana (BEERENDS 2000). O sinal de entrada utilizado segue o especificado no padrão P.830. O PAMS exige volume constante do sinal de entrada e inexistência de ruído de fundo. A perda de pacotes, o *jitter* e a distorção causada pelo codificador são considerados e refletem no resultado da avaliação. O atraso e a variação lenta do atraso não são medidas pelo PAMS.

A escala de resultados do PAMS é a mesma utilizada pelo MOS.

4.2.4 PESQ

O modelo PESQ (*Perceptual Evaluation of Speech Quality*) foi desenvolvido conjuntamente entre a *British Telecommunications* e a *KPN Research*, e posteriormente submetido ao ITU-T, passando a ser o padrão P.862 (BEERENDS 2000).

O PESQ é adequado para todos os tipos de *codecs*, bem como para avaliar atraso variável, perda de pacotes ou células e erros em canais de transmissão.

PESQ combina a consideração dada ao *jitter* no método PAMS com uma apurada modelagem da percepção humana oriunda do modelo perceptual do PSQM99 (ANDERSON 2002)(FERNANDES 2003).

O PESQ faz a comparação entre um sinal original e um sinal degradado, oriundo de um sistema de comunicação. O Resultado é uma predição da percepção de qualidade que seria obtida num teste subjetivo de escuta (FERNANDES 2003).

Os atrasos entre os sinais são tratados por um algoritmo de alinhamento de tempo. O sinal degradado é alinhado no tempo e no nível de potência (HOLLIER 2002), e tratado por um modelo perceptual (RIX 2002) para depois ser comparado com o sinal original, conforme mostra a figura 4.3 .

A modelagem perceptual transforma os sinais de origem e degradado em representações internas de acordo com a percepção humana, comparando a diferença entre os sinais. Esta modelagem inclui mapeamento tempo-freqüência (similar ao PSQM), deformação no domínio da freqüência, compressão na escala da intensidade com uma deformação não-linear proporcional à escala em dBm para refletir a sensibilidade humana e trata da distorção que o sinal degradado apresenta em relação ao original (FERNANDES 2003).

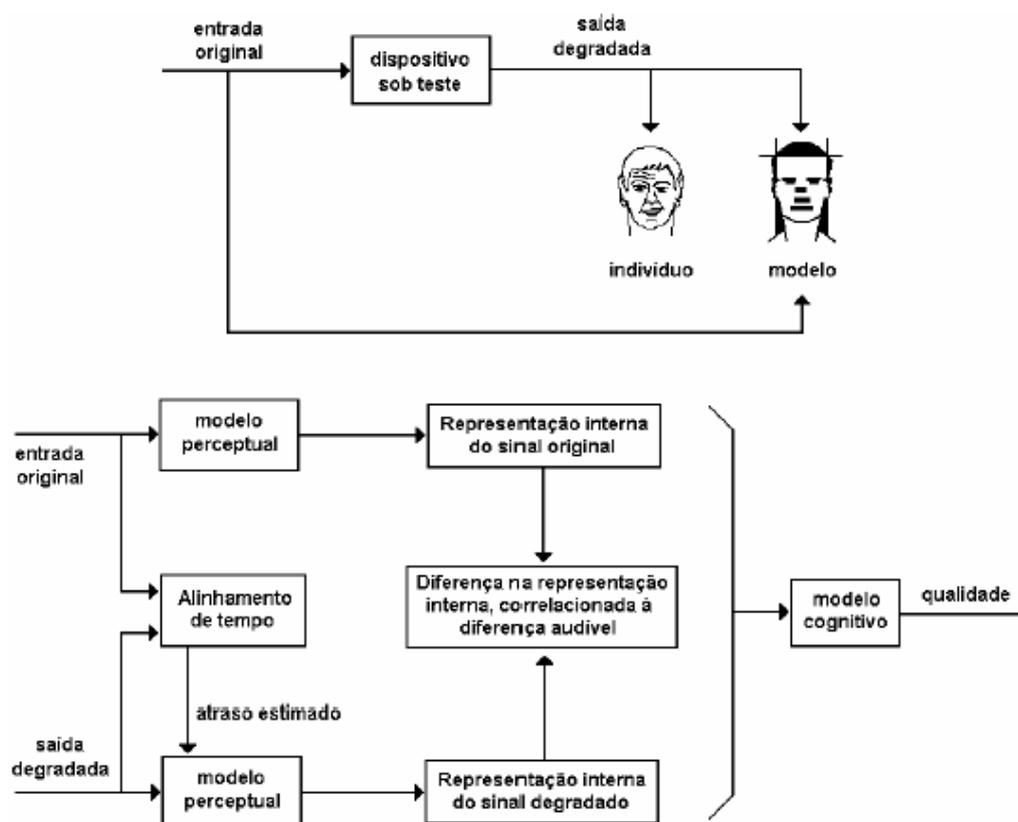


Figura 4.3: Filosofia básica do PESQ
(FERNANDES 2003)

O PESQ permite avaliar a interferência da perda de pacotes e da variação de atraso no sinal recebido. O atraso propriamente dito não é avaliado, por conta do alinhamento temporal dos sinais a serem comparados. Este alinhamento deixa passar apenas as variações (FERNANDES 2003).

4.3 Mensuração Objetiva Não-Intrusiva da Qualidade da Fala

Nos métodos objetivos intrusivos o tráfego normal da rede tem que ser interrompido durante o teste. Nos métodos não intrusivos não é necessário nem parar o tráfego normal, nem introduzir um sinal de referência. Por isso estes métodos são apropriados para a monitoração do tráfego normal (SUN 2004).

Há duas categorias de métodos não-intrusivos. Um deles, mostrado na figura 4.4 como “(Método 2)”, retira a predição da qualidade da fala diretamente das variações dos parâmetros da rede IP (perda, *jitter* e atraso) e de parâmetros não IP (*codec*, eco e idioma). Seu propósito é estabelecer uma relação entre a qualidade percebida e parâmetros IP e não-IP (SUN 2004). Exemplos deste tipo de método são o Modelo E e o Modelo de Rede Neural Artificial (ANN – *Artificial Neural Network*) apresentados nos itens 4.3.1 e 4.3.2 respectivamente.

Outra abordagem é prever a qualidade da fala diretamente do sinal de voz degradado, usando métodos de processamento digital de sinais, mostrado na figura 4.4 como “(Método 1)”. O P.SEAM, apresentado no item 4.3.3, é um exemplo deste tipo de método.

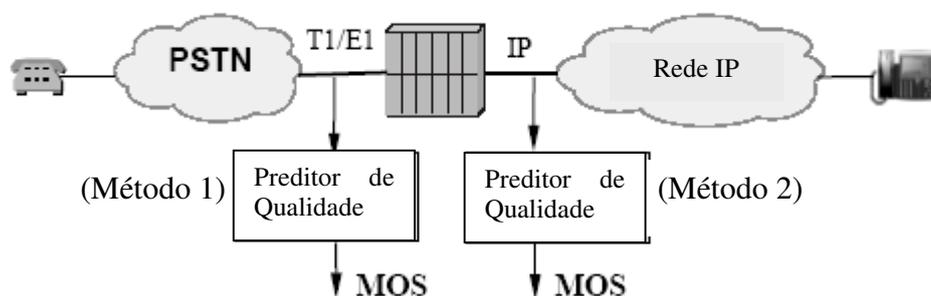


Figura 4.4 : Métodos Não-Intrusivos
(SUN 2004)

A seguir são apresentados os métodos objetivos não-intrusivos : o Modelo E, o Modelo E Extendido e o Modelo de Rede Neural Artificial.

4.3.1 Modelo E

O Modelo E foi proposto pelo ETSI (1996) e depois transformado em recomendação ITU-T (ITU-TG107 2005). Este modelo baseia-se no conceito de que “fatores psicológicos numa escala subjetiva são aditivos” (ITU-TG107 2005), ou seja, os fatores de perda de qualidade podem ser computados separadamente, mesmo estando correlacionados (COLE 2001). O resultado

do modelo é um fator escalar R (R -factor), que varia de 0 (pior caso) a 100 (excelente). O Fator R se relaciona com o MOS conforme a tabela 4.1 .

Em termos práticos pode-se considerar o intervalo típico do fator R entre 50 e 94 pois valores abaixo de 50 são considerados inaceitáveis e ligações telefônicas não ultrapassam o valor 94 (ANDERSON 2002). A telefonia convencional obtém um MOS no pior caso de 3.6, sendo que a meta a ser alcançada é um MOS superior a 4,0 (ANDERSON 2002).

Tabela 4.1: Relação entre fator R e MOS

Fator R	Nível de satisfação	MOS
90 a 100	Muito satisfeitos	4.3 a 4.5
80 a 90	satisfeitos	4.0 a 4.3
70 a 80	Alguns usuários insatisfeitos	3.6 a 4.0
60 a 70	Muitos usuários insatisfeitos	3.1 a 3.6
50 a 60	Praticamente todos insatisfeitos	2.8 a 3.1
0 a 50	Não recomendado	1.0 a 2.8

O Fator R é calculado a partir da seguinte equação (ITU-TG107 2005):

$$R = (R_o - I_s) - I_d - I_e + A$$

Onde :

R_o é a relação sinal ruído. Segundo a G.107 o valor padrão é de 94,77;

I_s é o fator de degradação oriundo da quantização da voz. Valor padrão de 1,41 segundo a G.107;

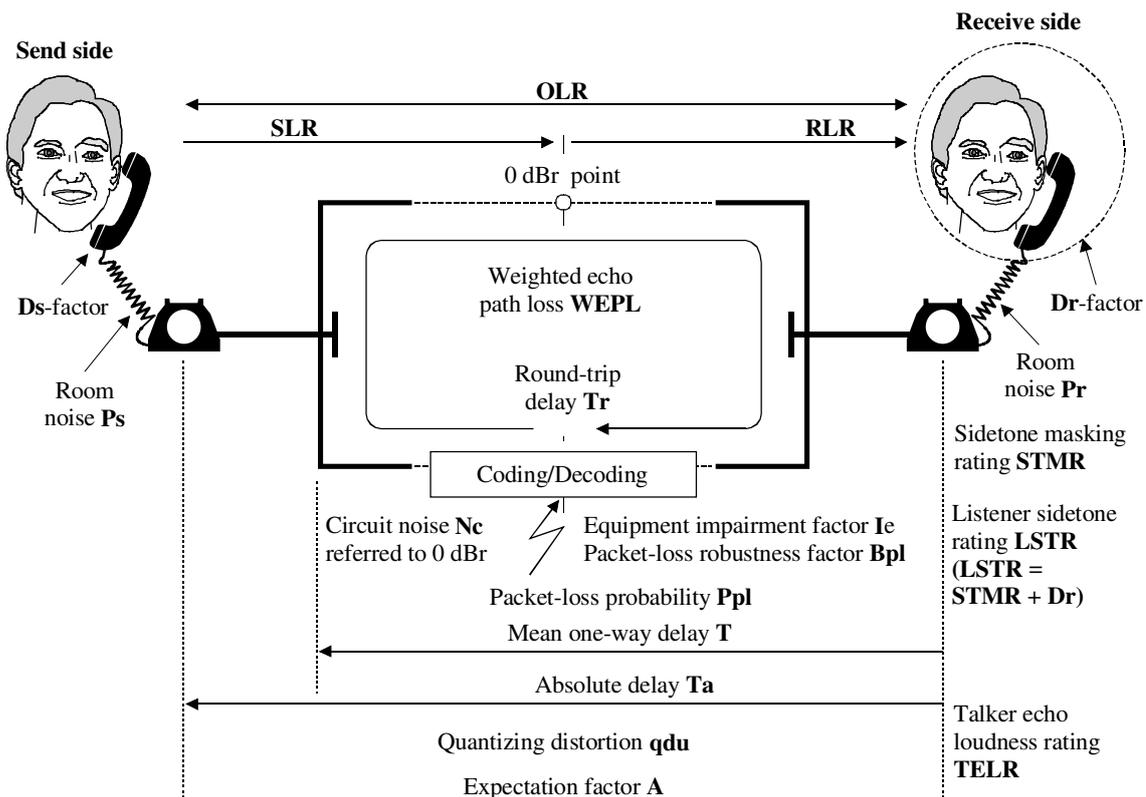
I_d corresponde a degradação causada pelo atraso, e pode assumir valores típicos de 0 a 60, segundo a G.107;

I_e representa a distorção no sinal devido aos *codecs*, perda de pacotes, e *buffers* de reprodução e compensação de *jitter*. A G.113 fornece valores provisórios de I_e para alguns *codecs*, por exemplo : G.771= 0 , G.729-A= 11 , G.723.1-5.3= 19;

A corresponde a quanta degradação o usuário está disposto a aceitar frente as vantagens do uso de determinada solução. Este fator varia de 0 a 20.

Os fatores **R_o** e **I_s** são intrínsecos do próprio sinal de voz e não dependem do meio de transmissão, conseqüentemente são considerados irrelevantes quando se compara telefonia convencional com voz sobre IP (ANDERSON 2002). O fator **A** (*advantage factor*), para efeito de comparação com PSTN, também pode ser ignorado (ANDERSON 2002).

O Modelo E segue uma conexão de referência onde são apresentados os parâmetros de transmissão usados como entrada de um modelo computacional conforme mostra a figura 4.5.



G.107_F01

Figura 4.5 : Conexão referência do Modelo E
(ITU-TG107 2005)

A descrição dos termos utilizados na figura 4.2, assim como seus valores *default* e a faixas permitidas, se encontram na tabela 4.2, a seguir :

Tabela 4.2 : Descrição dos parâmetros da figura 4.2
(ITU-TG107 2005)

Parâmetro	Abrev.	Unid.	Valor <i>default</i>	Faixa Permitida
Send Loudness Rating	SLR	dB	+8	0 ... +18
Receive Loudness Rating	RLR	dB	+2	-5 ... +14
Sidetone Masking Rating	STMR	dB	15	10 ... 20
Listener Sidetone Rating	LSTR	dB	18	13 ... 23
D-Value of Telephone, Send Side	Ds	-	3	-3 ... +3
D-Value of Telephone Receive Side	Dr	-	3	-3 ... +3
Talker Echo Loudness Rating	TELR	dB	65	5 ... 65
Weighted Echo Path Loss	WEPL	dB	110	5 ... 110

Parâmetro	Abrev.	Unid.	Valor default	Faixa Permitida
Mean one-way Delay of the Echo Path	T	ms	0	0 ... 500
Round-Trip Delay in a 4-wire Loop	Tr	ms	0	0 ... 1000
Absolute Delay in echo-free Connections	Ta	ms	0	0 ... 500
Number of Quantization Distortion Units	qdu	–	1	1 ... 14
Equipment Impairment Factor	Ie	–	0	0 ... 40
Packet-loss Robustness Factor	Bpl	–	1	1 ... 40
Random Packet-loss Probability	Ppl	%	0	0 ... 20
Burst Ratio	BurstR	–	1	1 ... 2
Circuit Noise referred to 0 dBr-point	Nc	dBm0p	–70	–80 ... –40
Noise Floor at the Receive Side	Nfor	dBmp	–64	–
Room Noise at the Send Side	Ps	dB(A)	35	35 ... 85
Room Noise at the Receive Side	Pr	dB(A)	35	35 ... 85
Advantage Factor	A	–	0	0 ... 20

As fórmulas de cálculo dos componentes da equação do fator R, utilizando os parâmetros da tabela 4.2, são apresentados em detalhes em (ITU-TG107 2005).

O Modelo E é utilizado atualmente como ferramenta de apoio para planejamento de capacidade e monitoração de desempenho de serviços de voz (ANDERSON 2002).

4.3.2 Extended E-Model

O Extended E-Model (ETSI 2002) é uma extensão proposta ao Modelo E que leva em consideração a distribuição do atraso e das perdas, e também aspectos humanos como por exemplo o efeito memória (CLARK_1 2001).

Em testes de comparação realizados pela empresa Telchemy (Telchemy 2003), comparando o Extended E-Model com o E-Model verifica-se que os valores estimados pelo Extended E-Model são muito mais próximos aos do MOS do que os valores calculados através do E-Model como pode ser visto na figura 4.6. O Extended E-Model usado foi implementado através do software Vqmon (CLARK_2 2001), da Telchemy.

O Extended E-Model (CLARK_2 2001) estabeleceu quatro novos conceitos :

- Qualidade instantânea : leva em conta a perda de pacotes ou outros fatores de degradação.
- Qualidade percebida : é determinada pelo usuário no momento da escuta.

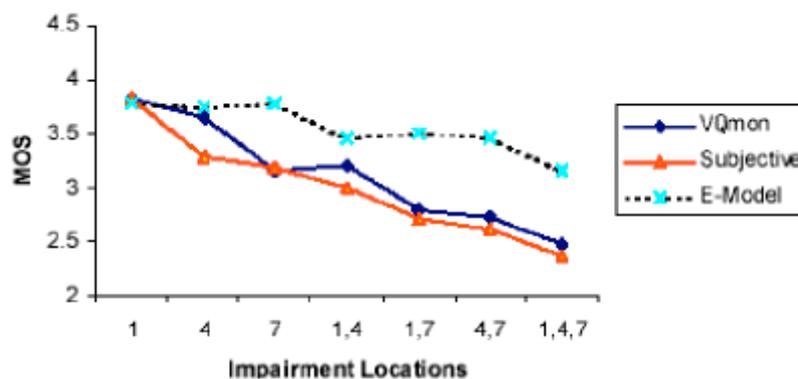


Figura 4.6 : Comparação entre o Extended E-Model (VQmon), o MOS (Subjective) e o Modelo E (Telchemy 2003)

- Comportamento da perda de pacotes alternantes : as perdas de pacotes normalmente ocorrem em períodos de rajadas (*burst*) que se alternam com períodos de perdas isoladas (*gap*) (CLARK_2 2001). Mecanismos de compensação de perdas (PLC) conseguem amenizar tais distorções em períodos de poucas perdas, mas não quando estas ocorrem em rajadas (LUSTOSA 2004).

- Memória recente : Durante pesquisas conduzidas para determinação de pontuações MOS, notou-se que a qualidade percebida de uma chamada telefônica varia de acordo com a localização, no tempo, das perdas. Notou-se que distorções ocorridas no final de uma chamada têm um efeito psicológico negativo maior sobre os avaliadores do que distorções ocorridas no início (LUSTOSA 2004)(ETSI 2002).

Experimentos em (CLARK_2 2001) verificou-se que o usuário demora um pouco para perceber a mudança de qualidade instantânea de “bom” para “ruim”. Esse efeito pode ser aproximado por uma curva exponencial de constante de tempo de 5s na transição de “bom” para “ruim” e de 15s na transição de “ruim” para “bom” (LUSTOSA 2004)(CLARK_2 2001).

A diferença fundamental entre o E-Model e o Extended E-Model está no cálculo do fator de perda l_e , como apresentado na figura 4.7. O Extended E-Model leva em consideração a taxa de perda em rajada e a taxa de perdas isoladas. Estes fatores passam por um modelo de percepção que simula o comportamento humano e o efeito memória recente a estas manifestações.

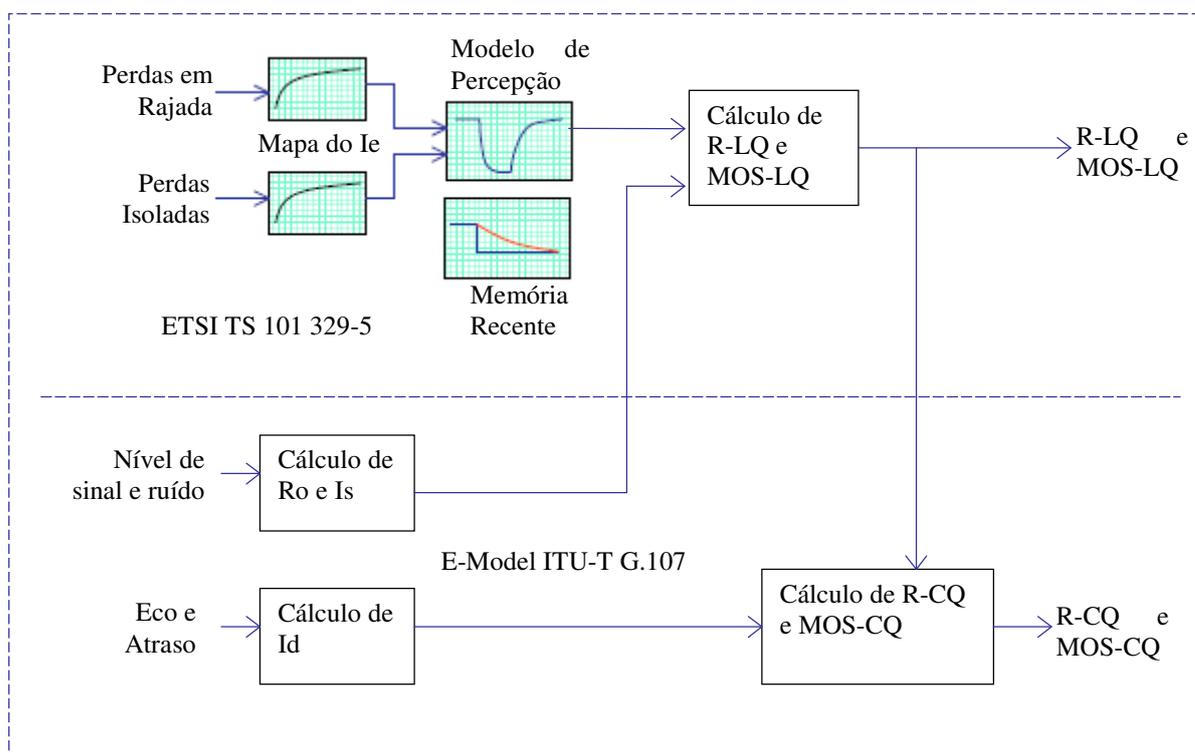


Figura 4.7 : Modelo computacional do Extended E-Model
(CLARK-VDC 2005)

A sigla “LQ” que aparece na figura 4.7 refere-se a Qualidade de Escuta (*Listening Quality*) e a sigla “CQ” refere-se a Qualidade da Conversação (*Conversational Quality*). A Qualidade da Escuta e a Qualidade da Conversação estão explicadas nos itens **5.1.1 Qualidade da Escuta** e **5.1.2 Qualidade da Conversação**.

4.3.3 Modelo de Rede Neural Artificial

Ao contrário do modelo E (um modelo matemático/computacional estático) , o modelo de Rede Neural Artificial (ANN) pode se adaptar dinamicamente ao ambiente de rede IP, por ter a capacidade de aprender.

Um modelo ANN pode ser construído por aprendizado de relações não-lineares entre qualidade percebida da voz (por exemplo MOS) e uma variedade de parâmetros de rede ou de parâmetros relativos a voz (SUN 2004).

Redes neurais artificiais tem sido utilizadas com sucesso em problemas de engenharia tais como reconhecimento de voz e imagem, controles adaptativos, detecções, estimativas, e áreas de telecomunicações tais como controle de fluxo e de admissão de chamadas ATM (NORDSTROM 1995), (NEVES 1995), previsão de tráfego para serviços multimídia (DOULAMIS 2003), (PARLOS 2003) e controle de tráfego (YOUSEFIZADEH 2002), (YOUSEFIZADEH 2003).

As principais aplicações de ANN podem ser classificadas em três categorias: classificação e reconhecimento de padrões, predição, e controle e otimização.

A estrutura básica de um modelo ANN é ilustrada na figura 4.8. Os parâmetros de entrada podem ser taxa de pacotes perdidos, tamanho dos pacotes, atrasos, tipo de codec, idioma, eco, ruído, etc. O *jitter* normalmente é transformado em pacotes perdidos ou atraso (SUN 2004).

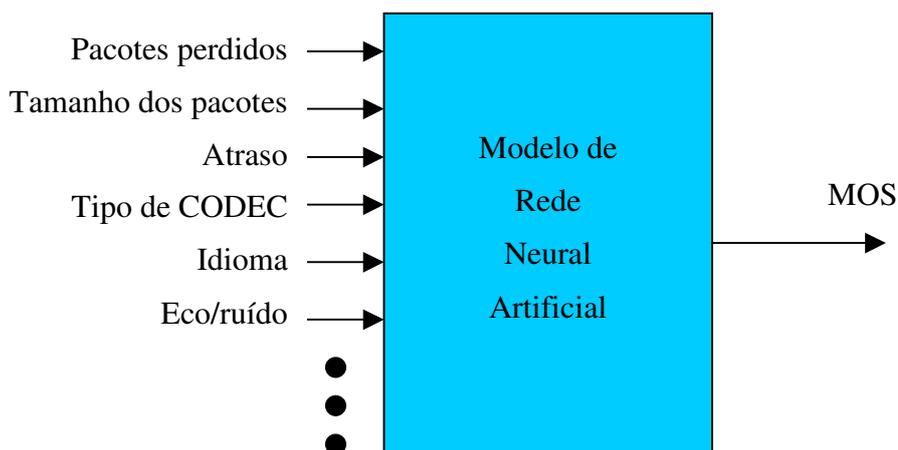


Figura 4.8 : Diagrama conceitual do modelo ANN
(SUN 2004)

Os modelos ANN em (MOHAMED 2002) prometem prever a qualidade de voz e vídeo através de parâmetros de rede e parâmetros não-IP, mas dependem de testes subjetivos para criar os contextos de treino. Lamentavelmente os testes subjetivos são caros e consumidores de tempo assim os resultados dos cenários de treino são limitados e não cobrem todos os possíveis cenários reais de rede. O impacto da variedade de parâmetros de rede e de parâmetros não-IP na qualidade percebida da voz permanece nebuloso. Além disto, o desenvolvimento dos modelos ANN foram baseados em um número limitado de codecs e podem prever somente a qualidade de escuta num sentido, de acordo com (SUN 2002). Há uma necessidade não coberta ainda, de modelos de qualidade de conversação para contabilizar a interatividade entre falantes, segundo (SUN 2004).

4.3.4 P.SEAM

Percebendo a necessidade de um padrão não-intrusivo de medição de qualidade de voz baseado diretamente no sinal de voz degradado, o ITU-T anunciou em 2002 uma competição aberta em busca do algoritmo o mais preciso possível para esta medição, que trabalhasse sob uma extensa faixa de condições de rede. Em 2003 as empresas Psytechnics, Opticom e SwissQual fizeram um acordo combinando suas experiências com seus modelos individuais para criar o melhor algoritmo possível. Cada companhia contribuiu com seu algoritmo, a Psytechnics com o NiQA (*Non-Intrusive speech-Quality*

Assessment) (GRAY 2000), a SwissQual com o NiNA (*Non-INtrusive speech quality Assessment*) e a Opticom com o P3SQM (*Perceptual Single Sided Speech Quality Measure* - desenvolvido em parceria com a KJPN Research). A combinação dos três algoritmos não-intrusivos resultou na recomendação P.563 do ITU-T (HOLUB 2004).

O modelo P.563 (ITU-TP563 2004) ou P.SEAM (*Single Ended Assessment Model*) é um método não-intrusivo baseado em sinais, fazendo medidas no lado do ouvinte, sem a necessidade de nenhum sinal de referência. Em contrapartida ao fato de perder alguma precisão se comparado a métodos intrusivos como o PESQ, o novo método baseado em sinais tem a enorme vantagem de poder realizar medições em praticamente qualquer ponto da rede com qualquer sinal de voz numa situação real de uso (OPTICOM 2004).

O P.563 é baseado nos mesmos princípios fundamentais da percepção humana usados em modelos intrusivos, como o PSQM, o PAMS e o PESQ.

O método P.563/P.SEAM pode ser dividido em 4 blocos, mostrados na figura 4.9 : pré-processamento, análise de distorções/extração de parâmetros, escolha da distorção dominante, e determinação da qualidade estimada.

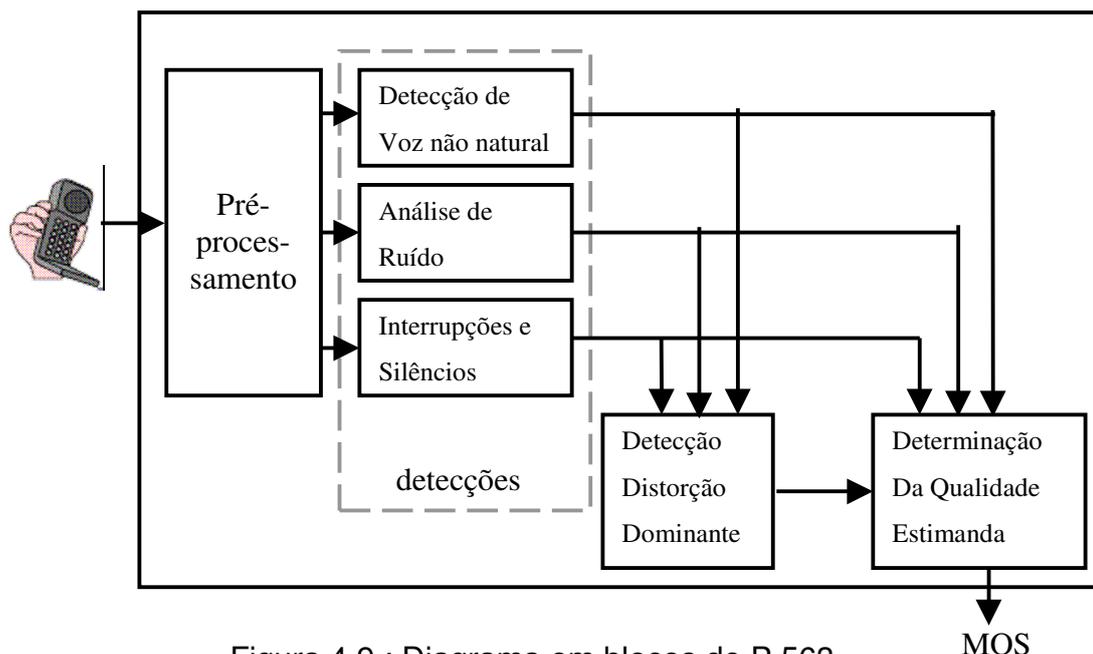


Figura 4.9 : Diagrama em blocos do P.563
(OPTICOM 2004)

No pré-processamento o sinal de voz é capturado e separado do sinal de áudio, e o nível de amplitude é adequado.

Na análise de distorções/extração de parâmetros é realizada uma análise de trato vocal, detectado “voz de robô”, analisada a relação sinal/ruído estática e por segmento de voz, e detectada interrupções no sinal e períodos de silêncio.

Na escolha da distorção dominante verifica-se qual distorção tem maior peso no segmento analisado. Isto é feito porque se aproxima do julgamento humano da qualidade da voz.

Na determinação da qualidade estimada correlacionam-se todas as distorções levando-se em conta a distorção dominante.

Este tipo de algoritmo exige uma capacidade significativa de processamento, da ordem de 100MIPs por chamada analisada, além de uma grande quantidade de memória. Para muitas aplicações isto é impraticável e nestes casos uma abordagem de análise baseada somente em parâmetros é aconselhada (CLARK-VQM 2005).

4.4 Métodos de Avaliação em fase de Padronização

A seguir são apresentadas breves descrições sobre dois métodos em fase de padronização no ITU-T : o P.AAM e o P.VTQ.

4.4.1 P.AAM

O método P.AAM (*Acoustic Assessment Model*), ainda não terminado, é uma extensão do P.862/PESQ tendo como novidade a inclusão de interfaces acústicas. O PESQ provê avaliação de qualidade fim-a-fim de sinais da banda de voz em interfaces elétricas de componentes de rede. A extensão P.AAM irá suportar a funcionalidade de avaliação acústica para terminais incluindo telefones, fones de ouvido, e conjuntos fones-de-ouvidos/microfone (*hands-free*). É possível que a versão acústica baseada no modelo P.862/PESQ original torne-se um novo padrão do ITU, o padrão P.863 (OPTICOM 2004).

Após vários estudos e artigos, Antony Rix, co-autor do PAMS e do PESQ, juntamente com John Beerends, co-autor do PSQM e do PESQ, e ainda Thomas Goldstein, apresentaram um esboço para a recomendação P.AAM para o ITU-T em julho de 2003 (GOLDSTEIN 2003). Em maio de 2004, no ICASSP (*International Conference on Acoustics, Speech, and Signal Processing- Quebec, Canadá*), Goldstein e Rix apresentaram um artigo atualizado sobre o P.AAM (GOLDSTEIN 2004).

Jens Berger afirmou na apresentação "*Future work on objective speech quality measurements in ITU*" no *Workshop on Wideband Speech Quality in Terminals and Networks : Assessment and Prediction-Mainz-Alemanha*, em junho de 2004, que os trabalhos no P.AAM foram suspensos em 2004 por conta de desacordos nos cenários de teste a serem usados (BERGER 2004).

4.4.2 P.VTQ

O ITU-T está trabalhando na padronização de uma metodologia de medição de qualidade objetiva baseada somente em informações de pacotes IP (sem avaliar a fala contida dentro do pacote) para uso em monitorações de qualidade em tempo real. Esta padronização está sendo chamada provisoriamente de P.VTQ (*Voice Transmission Quality*). Segundo (TAKAHASHI 2004), na fase de seleção do algoritmo, foram escolhidos dois candidatos : o PsyVoIP (BROOM 2003) e o VQmon (CLARK_2 2001). Alan Clark, criador do VQmon, em (CLARK-IT 2005), informa que o VQmon foi incorporado como *subset* do

P.VTQ. Ole Lauridsen em (LAURIDSEN 2004) e Mike Hollier, da Psytechnics, criadora do PsyVoIP, em (HOLLIER 2005), apresentam o PsyVoIP como sendo o próprio P.VTQ. O ITU-T ainda não terminou a recomendação.

A seguir é apresentada uma descrição apenas do PsyVoIP, visto que o VQmon já foi apresentado em **4.3.2 Extended E-Model**.

O PsyVoIP é um conjunto de componentes de Software para avaliação e gerenciamento de Qualidade de fala em VoIP. O PsyVoIP opera monitorando chamadas VoIP em tempo real para determinar a qualidade da fala em diversos pontos da rede (BROOM 2003).

O núcleo do PsyVoIP foi concebido como uma grande *probe*, um grande monitor. Este monitor combina componentes de software que capturam pacotes, extraem fluxos de chamada, parâmetros de degradação de VoIP, e predições de qualidade de fala. O monitor foi projetado para rodar dentro de equipamentos de teste de rede ou elementos de rede VoIP (BROOM 2003).

A arquitetura do monitor está mostrada na figura 4.10. Pacotes são capturados da rede e passam para o módulo de identificação de chamadas que mapeia os pacotes de uma chamada específica. O bloco de pré-processamento extrai informações necessárias ao resto do monitor e o resto dos pacotes pode ser descartado. Pacotes fora da seqüência são enviados para um buffer de “resseqüenciamento” permitindo que os pacotes sejam processados na sua ordem original. O bloco de detecção de atividade de voz (VAD- Voice Activity Detection) permite a marcação de pacotes sem voz. Isto permite maior precisão na predição da qualidade da fala porque pacotes com voz perdidos tem um maior impacto na qualidade da fala do que pacotes com silêncio. Finalmente descritores estatísticos são extraídos do fluxo de pacotes e os valores dos parâmetros internos são atualizados (BROOM 2003).

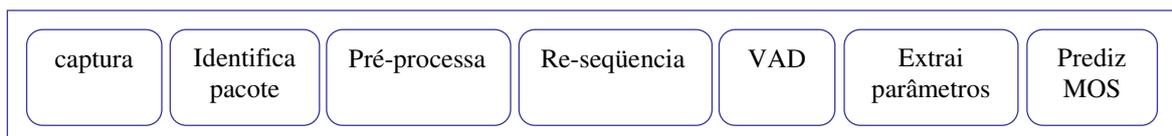


Figura 4.10 : Arquitetura da PROBE PsyVoIP
(BROOM 2003)

A predição de qualidade da fala corrente (MOS) para uma chamada pode então ser calculada sempre que requisitada, a partir dos valores dos parâmetros internos (BROOM 2003).

A descrição do P.VTQ a seguir foi baseada em (TAKAHASHI 2004), que apresentou o que havia sido desenvolvido até então.

O P.VTQ pode ser dividido em dois estágios.

O primeiro estágio do procedimento do P.VTQ calcula os parâmetros de qualidade intermediários como taxa de pacotes perdidos, padrão de perda de pacotes, e jitter (do RTP ou através do RTCP). Então o segundo estágio do P.VTQ estima o MOS de escuta. Os parâmetros de qualidade intermediários formam um *subset* de entidades definidas no RTCP-XR (*eXtended Report*)

proposto no IETF. Fica fácil de construir um sistema de gerenciamento de qualidade se o RTCP-XR estiver implementado no sistema VoIP: a partir de um terminal que provê as informações necessárias para o RTCP-XR, a qualidade de escuta pode ser estimada pela aplicação do segundo estágio do algoritmo P.VTQ. O RTCP-XR pode também ser usado para transportar a qualidade estimada de escuta. Em caso onde os terminais não podem manipular o RTCP-XR, os parâmetro de qualidade intermediários podem ser estimados do RTP e/ou dos pacotes RTCP (o primeiro estágio do algoritmo P.VTQ). O segundo estágio do algoritmo é então usado para estimar a qualidade de escuta. Neste caso a taxa de pacotes perdidos pode ser diferente, dependendo de como foi implementado o *jitter buffer* do terminal. Psytechnics (BROOM 2003) resolve este problema preparando um arquivo de calibração para cada tipo de terminal em adiantado, contendo a descrição das características do *jitter buffer* do terminal(TAKAHASHI 2004).

Até o momento da escrita desta dissertação o ITU-T ainda não havia divulgado informações sobre a especificação definitiva do P.VTQ.

4.5 Análise sobre os Métodos de Avaliação de Qualidade da Voz

Aqui são apresentadas considerações sobre os diferentes métodos de avaliação descritos neste capítulo.

4.5.1 Comparação entre Medições Ativas e Medições Passivas

Medições ativas são utilizadas em situações de teste. São utilizadas nas avaliações iniciais de uma rede, antes de permitir o tráfego real de voz. Outra situação de uso de medições ativas ocorre quando surgem problemas na rede e é preciso criar tráfego artificial para a avaliação dos mesmos (CLARK-ISP 2005).

Medições ativas permitem o controle explícito na geração de pacotes, e testam o que se quer , no momento desejado. Por outro lado injetam tráfego artificial (COTTRELL 2005).

Problemas transientes podem passar despercebidos em monitorações ativas já que elas não são constantes (CLARK-ISP 2005).

Os equipamentos utilizados para monitoração ativa normalmente apresentam mais de um método intrusivo de medida. Atualmente o método intrusivo de melhor precisão é o PESQ-P.862. É necessário um poder de processamento grande para executá-lo, como mostrado na figura 4.11.

Medições passivas ocorrem na monitoração de chamadas reais, na operação em regime normal de tráfego de voz (CLARK-ISP 2005).

Estas medições podem monitorar todas as chamadas, permitem capturar informações de problemas transientes, e podem fornecer dados para análises posteriores (CLARK-ISP 2005). A coleta constante de amostras de dados necessita de grande espaço de armazenagem.

A monitoração passiva tem as vantagens de não injetar tráfego extra e de monitorar o tráfego real (COTTRELL 2005).

Para transformar o método PESQ (intrusivo) num método não intrusivo foi criado o P.SEAM/P.563. Como visto no item 4.3.4 o P.SEAM faz a avaliação do sinal de voz num ponto de chegada sem a necessidade de introdução de um sinal de referência no ponto de entrada da rede. O poder de processamento necessário é bem maior que o do PESQ, e a precisão da medida é bem menor, como se pode observar nas figuras 4.11 e 4.12. O ideal é que os métodos de medição gastem o menor processamento possível, para que sejam facilmente incluídos nos equipamentos de rede.

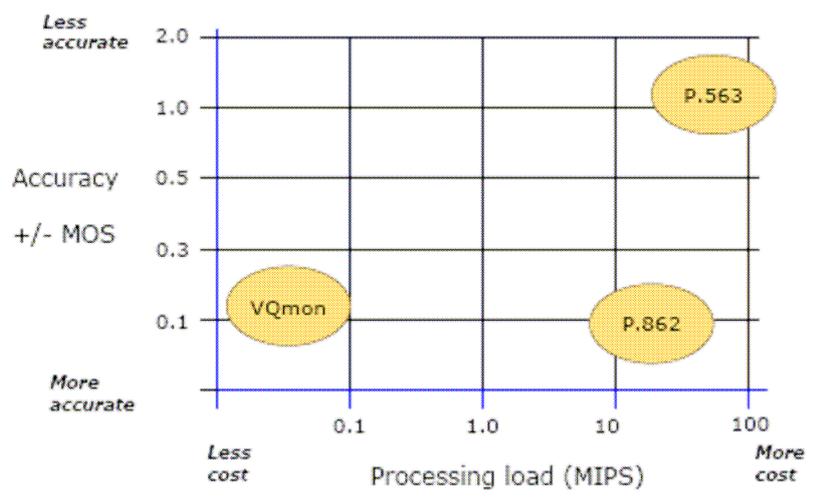


Figura 4.11 : Comparação de processamento e acuracidade entre métodos (CLARK-ISP 2005)

A necessidade de alto poder de processamento e a baixa acuracidade do P.SEAM/P.563 me levam a acreditar que este método não será utilizado na prática. Como ilustração é mostrado na figura 4.12 uma comparação entre uma medida subjetiva usando 16 ouvintes com a estimativa do P.563.

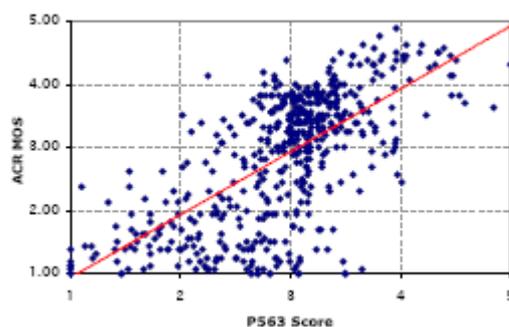


Figura 4.12 : Comparação entre P.563 e ACRMOS (CLARK-IT 2005)

4.5.2 Considerações sobre Extração de Parâmetros

A extração de parâmetros está diretamente relacionada às facilidades oferecidas pela implementação de VoIP existente para obtenção dos mesmos. Mais especificamente, a extração fica fácil se existir protocolo de controle de RTP, que hoje não é obrigatório. Podemos classificar em 3 tipos de extração : sem RTCP, com RTCP e com RTCP-XR.

Extração sem RTCP: na ausência de protocolo de controle do RTP é necessária a implementação de aplicações para obtenção dos parâmetros. A informação do codec utilizado pode ser obtida do campo PT (*Payload Type*) do cabeçalho RTP. As informações de atraso podem ser obtidas medindo a latência no protocolo de sinalização (THORELL 2005). O *jitter* médio pode ser calculado através da média das diferenças entre timestamps de dois pacotes RTP consecutivos seguindo a fórmula recomendada pela RFC3550 (SCHULZRINNE 2003). O campo *sequence* do cabeçalho RTP pode ser usado para calcular a perda de pacotes, (com o cuidado de levar em conta o limite de 65536 do campo de 16 bits) (THORELL 2005). Os métodos de monitoração perceptuais utilizam o conteúdo do pacote RTP(sinal de voz) para realização da análise de qualidade.

Extração com RTCP: sendo utilizado o protocolo RTCP, as informações sobre perda de pacotes, jitter, delay, nível de sinal da chamada podem ser obtidas diretamente dele como já visto no item **2.2 RTP Control Protocol – RTCP**.

Extração com RTCP-XR : com o uso do protocolo RTCP XR, além das informações fornecidas pelo RTCP, obtém-se diretamente também o fator R e o MOS de cada chamada.

5 DIMENSÕES DA AVALIAÇÃO DA QUALIDADE DE VOZ SOBRE IP

A qualidade da Voz sobre IP pode ser avaliada sob diversos pontos de vista. Neste capítulo é apresentada uma classificação, individualizando as diferentes dimensões do termo qualidade da chamada. Posteriormente é apresentada a dimensão total da qualidade percebida pelo usuário.

5.1 Definição de Qualidade da Chamada

A qualidade de uma chamada IP pode ser afetada por ruído, distorção, volume muito alto ou muito baixo, eco, falhas na fala e uma variedade de outros problemas. A medida da qualidade de uma chamada, pode assumir três dimensões (CLARK-VQM 2005) : qualidade da Escuta, qualidade da Conversação e qualidade da Transmissão.

5.1.1 Qualidade da Escuta

A qualidade de escuta refere-se a como os usuários avaliam o que eles ouviram durante a chamada.

Os métodos mais comumente usados para medida de qualidade de escuta são os testes subjetivos, normalmente processos dispendiosos e consumidores consideráveis de tempo. Uma das melhores metodologias conhecidas de testes subjetivos é a ACR (*Absolute Category Rating*), já descrita no capítulo 4.

5.1.2 Qualidade da Conversação

A qualidade da Conversação refere-se a como os usuários avaliam a qualidade no todo de uma chamada baseado na qualidade de escuta e na facilidade de conversar durante a chamada. Aqui é incluído qualquer eco ou atraso que podem afetar a conversação.

Testes de Qualidade da Conversação são mais complexos, e por conta disso, usados com menor frequência. Num teste de conversação, um conjunto de pessoas são colocadas com atores em cenários de comunicação interativa (sistemas telefônicos ou de voz sobre IP). São introduzidos efeitos com atraso

e eco, e os atores posteriormente fornecem suas opiniões sobre a qualidade da conversação.

O efeito do atraso na qualidade da conversação é muito dependente da tarefa. Para tarefas não muito interativas, atrasos início-fim(num único sentido) de várias centenas de milissegundos podem ser tolerados; para tarefas altamente interativas, pequenos atrasos podem causar dificuldades de conversação.

A interpretação sobre a qualidade da conversação depende do tipo de tarefa que está sendo executada. Por exemplo, dois sistemas VoIP idênticos tem 300ms de atraso num sentido. Num deles é realizada uma negociação comercial altamente interativa. No outro acontece um bate-papo informal entre amigos. No primeiro exemplo, os usuários podem considerar a qualidade da conversação muito ruim. No segundo caso, os usuários podem até nem ter percebido os atrasos (CLARK-VQM 2005).

5.1.3 Qualidade da Transmissão

A qualidade da Transmissão refere-se a qualidade da conexão de rede usada para transporte do sinal de voz. Os parâmetro aqui presentes não se referem a uma chamada específica e sim a infra-estrutura para o transporte da voz (TELCHEMY 2005).

Neste tipo de avaliação entram medições específicas como a banda oferecida, atrasos, priorização de tráfego, atraso, *jitter*, perda de pacotes,etc. (MONFORT 2003). Além destes temos o eco, o ruído, entre outros (HARDY 2001).

5.2 Métricas Primárias de Qualidade de Voz sobre IP

Como visto no item anterior, dois dos maiores desafios dos sistemas de voz sobre IP são a qualidade de escuta e a qualidade da conversação. Os usuários finais frequentemente não fazem distinção entre estes dois tipos de qualidade. No entanto, é razoável que se faça a distinção entre eles. Claramente, a qualidade da escuta é diretamente impactada por ruídos e outros tipos de distorção. É também claro que o sinal de voz distorcido impacta negativamente a conversação telefônica. E alguns fenômenos telefônicos, intensificados pelos processos de voz sobre IP, afetam a conversação sem afetar a qualidade do som como um todo, sendo por isso referenciados a qualidade da transmissão.

Em ambientes de voz sobre IP, três elementos, mostrados na figura 5.1, emergem como os fatores primários a afetar a qualidade da escuta e da conversação (IEC-VQ 2005).

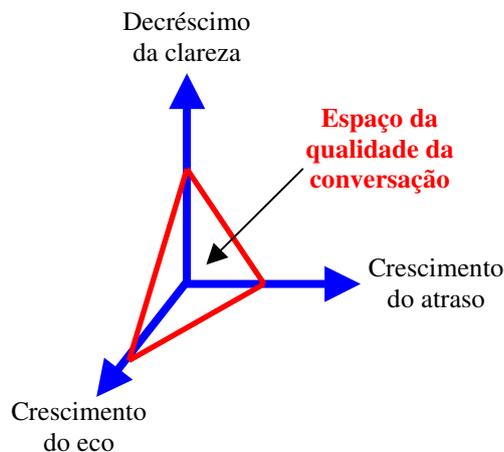


Figura 5.1 : Métricas Primárias de Qualidade de Voz
(HARDMAN 2003)

Clareza e atraso podem ser vistos como grandezas ortogonais já que normalmente uma não afeta diretamente a outra. Eco, por outro lado, afeta a percepção de clareza e, em muitos casos, pode ser mais notado com o crescimento do atraso (HARDMAN 2003). A figura 5.1 mostra uma relação rudimentar entre clareza, atraso e eco, a relação estritamente matemática não existe. Por outro lado nota-se que quando clareza é boa, atraso é pequeno, e eco é reduzido, a qualidade da voz como um todo melhora. Frequentemente se opta pela melhoria de um parâmetro em detrimento do outro. Por exemplo, para diminuição do atraso, os projetistas de VoIP podem usar esquemas de codificação menos complexos (e por isso de qualidade inferior), que acabam afetando a clareza do sinal de voz (HARDMAN 2003).

5.2.1 Clareza

Clareza genericamente refere-se a reprodução fiel do sinal de voz, livre de distorção, com inteligibilidade. Há primariamente uma métrica de qualidade de som onde a presença de ruído e distorção joga o papel mais importante. Clareza é uma métrica muito subjetiva e é desafiadora para medir, particularmente em aplicações de voz sobre IP. A clareza de um sinal de voz ou de um canal de voz tem sido medida subjetivamente de acordo com a recomendação P.800 do ITU-T (ITU-TP800 1996) resultando num escore médio de opinião (MOS), já visto no item 4.1.1. Em ambientes VoIP, problemas de clareza são frequentemente causados por perda de pacotes, jitter fora de controle e ruído nos circuitos analógicos. A clareza também é influenciada pelos *codecs* usados nos canais de voz (HARDMAN 2003).

5.2.2 Atraso e *Jitter*

O atraso fim-a-fim é o tempo gasto para o sinal ir do emissor da voz até o ouvinte. O atraso do sinal de voz é resultado da soma do processo de codificação e empacotamento da voz sobre IP e do transporte destes pacotes de voz. O atraso afeta a qualidade da conversação sem afetar o som real do sinal de voz – atraso não introduz ruído ou distorção no canal de voz (HARDMAN 2003).

Quando o atraso fim-a-fim chega por volta de 250ms, os participantes de uma conversa telefônica começam a perceber seus efeitos. Entre 300 e 500ms, uma conversação normal fica difícil. Acima de 500ms fica impossível estabelecer uma conversação normal. Em redes telefônicas públicas o atraso fim-a-fim é tipicamente abaixo de 10ms. Em redes IP o limite inferior fica entre 50 e 100ms por causa das operações dos *codecs*, do empacotamento e da compressão de voz (HARDMAN 2003).

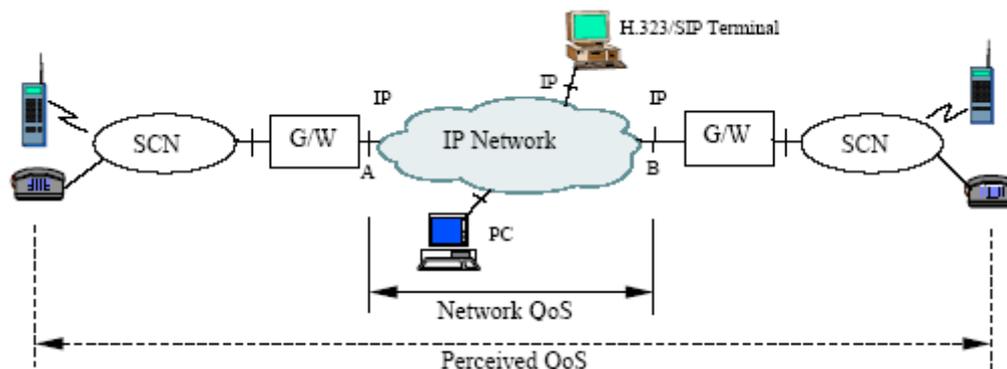
Outro aspecto de atraso, a sua variação (*jitter*), pode causar distorção no sinal de voz. *Jitter* pode ser visto como a variação temporal de chegada de pacotes individuais de voz aos *gateways* de voz. Para redes de dados o *jitter* é o menor dos problemas porque os pacotes que chegam podem ser armazenados por longos períodos de tempo. Para aplicações de tempo real como voz, entretanto, algum *jitter* pode ser tolerado, mas limites superiores rigorosos devem ser impostos. Quando pacotes chegam com *jitter* fora do limite superior, são descartados ou ignorados sendo contabilizados como pacotes perdidos. Pacotes perdidos afetam diretamente a distorção do sinal de voz, e devem ser controlados ou gerenciados em sistemas *VoIP* para reduzir os efeitos negativos (HARDMAN 2003).

5.2.3 Eco

O Eco, já visto no item 3.5, é o som da voz dos falantes que retorna aos seus ouvidos. Eco, como o atraso, influencia na qualidade da conversação mais do que na qualidade do som (HARDMAN 2003). Entretanto, eco pode afetar significativamente a percepção da qualidade do som. Em contextos de Voz sobre IP, o eco (já existe na rede de telefonia pública mas raramente percebido) tem seu efeito aumentado por causa dos atrasos dos processos de Voz sobre IP (FERNANDES 2003).

5.3 Qualidade de Serviço Percebida (QoE)

A qualidade percebida da fala em comunicações de voz é a métrica de qualidade de serviço mais importante, estando diretamente relacionada a percepção do ouvinte. Em aplicações de voz sobre IP o termo usado para isso é QoS (*Quality of Service*) percebido ou QoE (*Quality of Experience*): Qualidade Experimentada (SUN 2004). Como pode ser visto na figura 5.2, A Qualidade de Serviço Percebida (*Perceived QoS* na figura) é mais abrangente que a qualidade de serviço da Rede (*Network QoS* na figura).



SCN = *Switched Circuit Network* ou rede comutada por circuito

G/W = *Gateway*

Figura 5.2 : Dimensão da Qualidade Percebida
(SUN 2004)

Vários fatores influenciam a qualidade percebida da fala, como mostrado na figura 5.1. Os fatores da rede incluem perda de pacotes, *jitter* e atraso. O *jitter buffer* e o *codec* localizados no lado terminal consistem nos fatores de aplicação. Do ponto de vista fim a fim, a totalidade dos pacotes perdidos inclui os pacotes perdidos na rede e também aqueles perdidos por chegarem tarde demais no *jitter buffer*. O atraso total consiste do atraso na rede e do tempo gasto dentro do *jitter buffer*. Além da perda total de pacotes e do atraso total, a qualidade percebida da fala fim a fim (QoE) depende ainda do *codec* e da estratégia de dissimulação de pacotes perdidos do *codec* (assumindo que nenhuma dissimulação externa de pacotes perdidos é usada) (MONFORT 2003).

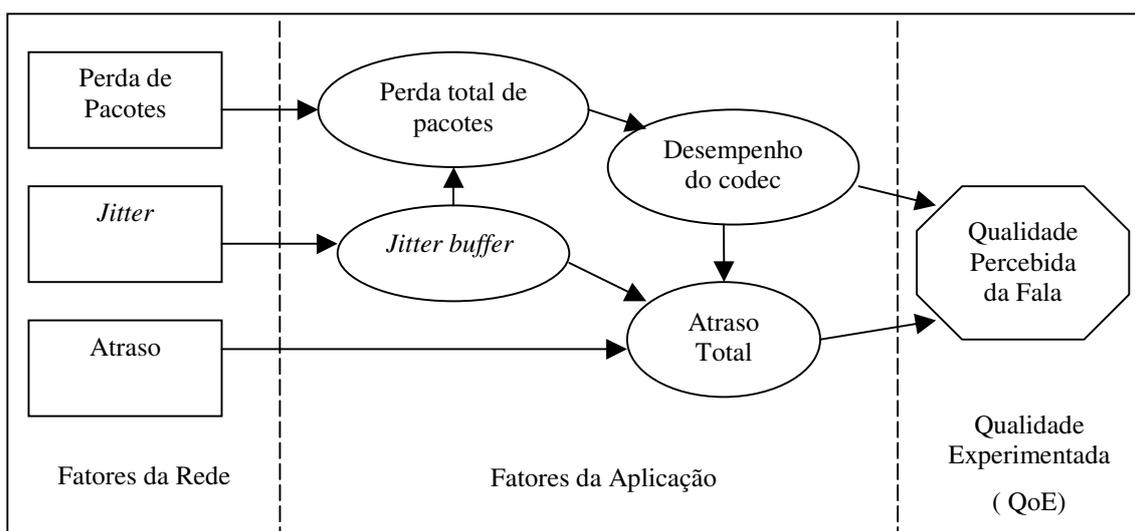


Figura 5.3 : Inter-relacionamento de fatores de QoS
(MONFORT 2003)

Outros fatores também afetam a qualidade percebida da fala fim a fim e não são mostrados na figura 5.3. São eles o eco, o ruído, o *cross-talk* (interferência de um canal de voz no outro, que ocorre pela proximidade dos pares de fios de canais diferentes na rede telefônica física), baixo (e/ou alto) volume etc.(HARDY 2001).

5.4 Análise sobre as Dimensões da Qualidade da Voz sobre IP

Aqui são apresentadas considerações sobre as dimensões da qualidade descritas neste capítulo.

Estas diferentes dimensões sobre a avaliação da qualidade da voz devem ser consideradas na execução de um procedimento para avaliação de qualidade de uma rede VoIP.

5.4.1 A diferença entre os Métodos de Medida de QoS e QoE

Um ponto bastante relevante na monitoração de qualidade é a clara definição do aspecto que se está querendo medir. Aqui dividiu-se o tipo de monitoração em dois aspectos : Qualidade do Serviço IP (QoS) e Qualidade do Experimento (QoE).

A Qualidade do Serviço IP é obtida dos parâmetros coletados da rede IP não levando em consideração os aspectos de fora da rede IP que influenciam a qualidade de voz.

A Qualidade do Experimento, também definida como qualidade percebida pelo usuário, leva em consideração todos os aspectos envolvidos desde a saída da voz no emissor até a chegada ao ouvinte.

5.4.2 Considerações sobre os Métodos de medida de Escuta, de Conversação e de Estrutura

É preciso definir claramente qual é o objetivo da medição para a escolha dos parâmetros adequados para aquele tipo de avaliação.

Se a medida for qualidade da escuta, o atraso fim-a-fim não é levado em consideração, já que a interação entre os atores não está sob avaliação. Outro parâmetro desconsiderado é o eco que retorna ao emissor, pois a avaliação está focada no ponto de vista do ouvinte.

Os principais pontos a serem considerados na qualidade da escuta são a distorção, o ruído, a perda percebida de segmentos de voz e a inclusão errada de segmentos de voz não pertencentes àquela chamada (também chamado *crosstalk*) (IEC-VQ 2005).

Quando a medida for a qualidade da conversação é necessário considerar, além dos parâmetros considerados na escuta, o eco, o atraso fim-a-fim, o desmepenho dos supressores de silêncio e o desmepenh dos canceladores de eco (IEC-VQ 2005).

Se a avaliação for da estrutura da rede, serão considerados os parâmetros primários de qualidade de serviço, quais sejam, atraso, jitter, banda e perda de pacotes.

6 TRABALHOS CORRELATOS

Neste capítulo são apresentados quatro trabalhos relacionados a metodologias de monitoração de qualidade de Voz sobre IP. Para cada um dos trabalhos é feita uma breve descrição incluindo suas aplicações, cenários possíveis e suas limitações.

6.1 Emissão de CDRs por *Gateways* de Voz

Os registros de detalhamento de chamada (CDR – *Call Detail Record*), originalmente criados para contabilização do tempo de início e término de uma ligação telefônica (VIANA 2004), para posterior registro e cobrança, também podem ser gerados por *gateways* de voz. Neste caso estes registros podem possuir alguns indicativos de qualidade. Estes indicativos são mais simples que, por exemplo, os fornecidos pelo modelo E estendido (LUSTOSA 2005). Estes CDRs emitidos pelos *gateways* de voz podem servir de base para uma infraestrutura de coleta e monitoração de qualidade de chamadas.

Aplicações :

Esta metodologia pode ser utilizada para cálculo da economia do uso de VoIP e demonstração dos benefícios deste tipo de implementação comparado à telefonia convencional (LIN 2003).

A coleta de CDRs em tempo real habilita os administradores de rede a monitorar o último tráfego de voz, para que possam atuar rapidamente na configuração da qualidade de serviço da rede, se necessário for (LIN 2003).

Cenários :

Em (LUSTOSA 2005) encontram-se dois exemplos de cenários de coleta de CDRs : a infraestrutura de monitoração e contabilização de chamadas da rede de ensino e pesquisa da República Tcheca, com coleta de CDRs emitidos por *gateways* de voz (UBIK 2001) e a arquitetura chamada ECAS – *Enterprise Call Analysis System* (LIN 2003).

Na figura 6.1 é mostrada a arquitetura ECAS, com os *gateways* de voz enviando os CDRs para os coletores de CDR, que por sua vez analisam os CDRs e os enviam para uma base de dados. Esta base de dados é acessada por um aplicativo analisador de chamadas (CAE – *Call Analysis Engine*) que transforma os registros em informações significantes para as empresas (total

de ligações, duração,...) além de análise estatística das características das chamadas (LIN 2003).

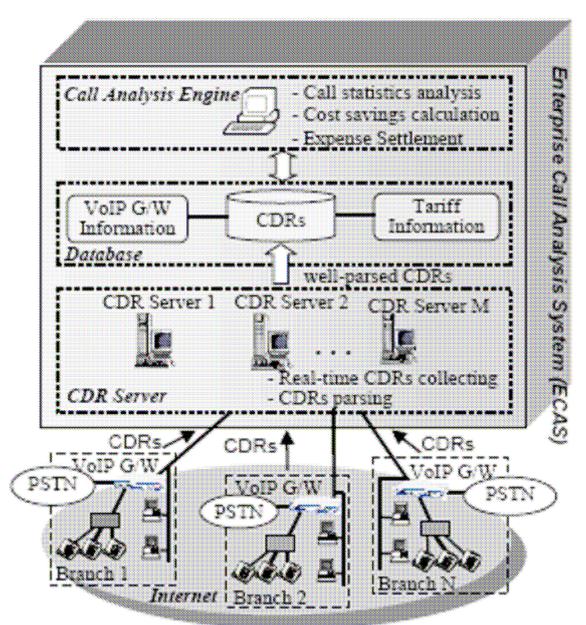


Figura 6.1 : Arquitetura do Sistema ECAS
(LIN 2003)

Limitações :

A coleta de CDRs de *gateways* de voz não tem condições de fornecer informações da qualidade percebida pelo usuário, pois só coleta informações de pontos intermediários da rede, deixando de fora os pontos finais, os telefones IP.

6.2 Monitoração Ativa

Monitoração Ativa é um método intrusivo baseado em geradores de chamadas colocados na rede, *probes* ativas trabalhando aos pares. Estas *probes* geram chamadas e as analisam (BERNEX 2005).

Este tipo de monitoração gera tráfego extra na rede.

Aplicações :

Este método geralmente é utilizado para testes de viabilidade da rede para posterior instalação de serviços de voz (BERNEX 2005).

Depois de instalados os serviços de voz, este método pode ser usado para análises periódicas da qualidade da rede.

Cenários :

Um cenário típico com esta metodologia pode ser visto na figura 6.2 onde 4 *probes* intrusivas trocam tráfego VoIP entre si e geram informações de qualidade de serviço enviando-as a um servidor de QoS.

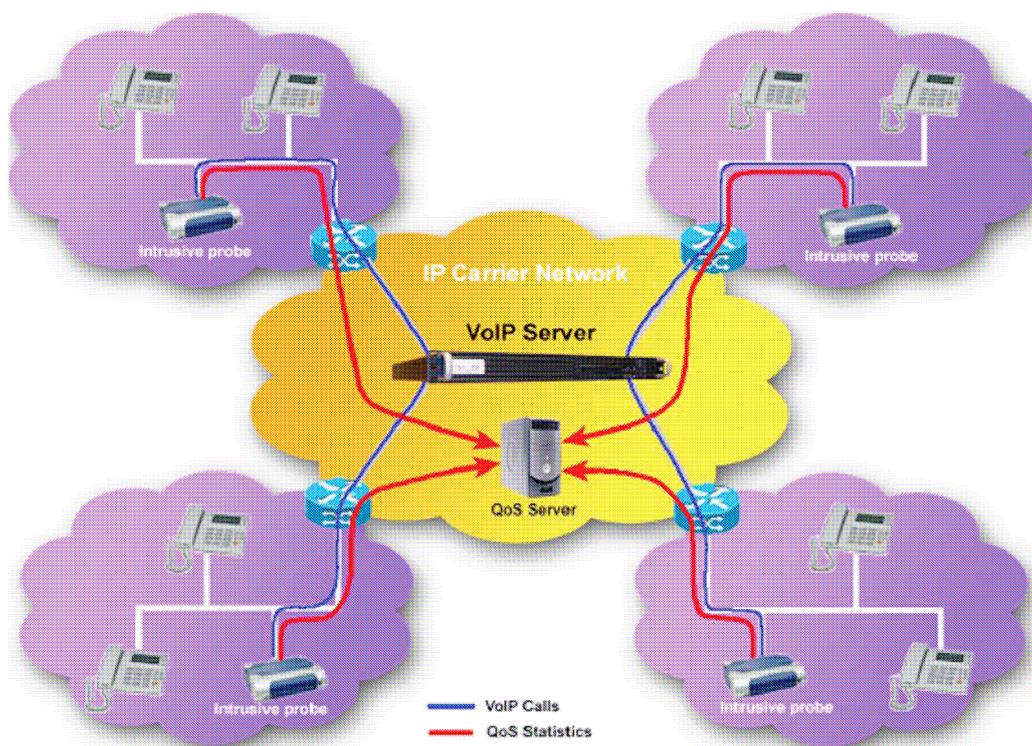


Figura 6.2 : Monitoração Ativa da Qualidade de Voz
(BERNEX 2005)

Um exemplo deste tipo de monitoração ativa pode ser encontrado em (HUANG 2003), onde um conjunto de simuladores de chamadas distribuídos pela rede geram tráfego VoIP e disponibilizam informações de perda de pacotes, atraso e *jitter* via SNMP (LUSTOSA 2005).

Limitações :

Uma das limitações para o uso deste método é o tráfego adicional gerado pelos instrumentos de supervisão que podem causar prejuízos na qualidade se a rede estiver sendo usada no limite de sua capacidade (BERNEX 2005).

Outra limitação é o fato das informações não serem baseadas em chamadas reais. Desta forma não se tem estatísticas das chamadas reais, e a amostragem obtida pode não refletir a realidade (LUSTOSA 2005).

6.3 Monitoração Passiva em Pontos Intermediários da Rede

Monitores são colocados em pontos intermediários da rede para capturar e analisar todos os fluxos de voz daquela rede. Após a análise pode-se gerar relatórios sobre a qualidade de voz oferecida (LUSTOSA 2005).

No caso de um conjunto de monitores não-intrusivos, uma chamada é analisada simultaneamente e independentemente por todos os monitores instalados na rota da chamada (BERNEX 2005).

Aplicações :

Entre dois pontos de medida poderão existir um ou mais elementos que podem degradar a qualidade da chamada. É possível obter estatísticas de qualidade de serviço e estatísticas de qualidade de voz para cada seção localizada entre duas *probes* (BERNEX 2005). Sendo assim, a distribuição de *probes* em pontos estratégicos da rede possibilita a rápida detecção de segmentos com qualidade degradada, ou por defeito em algum elemento, ou por limitação da construção daquele segmento da rede.

Cenários :

Um exemplo de cenário de monitoração passiva é apresentado na figura 6.3. O tipo de aplicativo de monitoração pode variar. Exemplos de aplicativos para este tipo de monitoração são o PsyVoIP (BROOM 2003), o 3SQM (OPTICOM 2004), e o Brix Test Suite (BRIX 2005).

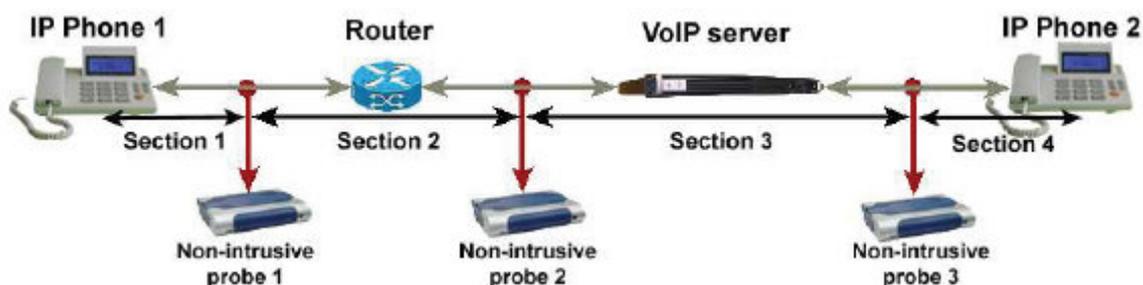


Figura 6.3 : Monitoração Passiva da Qualidade de Voz com várias *probes* (BERNEX 2005)

Limitações :

Esta solução não é aplicável quando os pacotes são criptografados, ou quando as chamadas ponto a ponto não passam por nenhum monitor (LUSTOSA 2005).

Dependendo do aplicativo usado, as medidas no meio da rede podem apresentar um fator de qualidade de voz pior do que o percebido no ponto final da chamada, já que os *softfones* podem ser providos de *jitter-buffers* que

melhoram a condição de *jitter* (BROOM 2003). O PsyVoIP leva isto em consideração e procura imitar a compensação dos jitter-buffers para se aproximar a medida de qualidade intermediária da medida de qualidade no ponto final (BROOM 2003).

Neste tipo de monitoração não é possível medir a qualidade percebida pelo usuário, já que os pontos finais não são monitorados.

6.4 Monitoração nos Pontos Iniciais e Finais da Chamada

Neste tipo de monitoração os indicadores de qualidade são coletados diretamente nos telefones IP. Cada telefone avalia a qualidade da chamada sob seu ponto de vista, em cima da mídia recebida. Assim capta-se dados mais confiáveis e abrangentes, avaliando-se todas as chamadas (QUINELLATO 2005).

Aplicações :

Esta forma de obtenção da qualidade de voz se aplica em situações onde há o uso de criptografia na rede não sendo possível retirar os parâmetros de monitoração em pontos intermediários na rede.

Esta forma também se aplica para avaliação de ligações ponto a ponto e em situações onde se deseja obter a qualidade percebida do usuário.

Cenários :

O primeiro cenário a ser apresentado é um novo formato de CDR proposto por Leandro Lustosa (LUSTOSA 2005). O denominado VQCDR (*Voice Quality CDR*) é composto por um CDR estendido, incluindo a monitoração de qualidade de chamadas VoIP. A metodologia proposta por Lustosa utiliza CDRs coletados por RADIUS (RIGNEY 2000), vindos de telefones IP e de *gateways* de voz. Foi desenvolvida uma biblioteca, *Voice Quality*, para ser usada como ferramenta de medições de qualidade de voz. Esta biblioteca é composta por três módulos, como pode ser visto na figura 6.4. O primeiro módulo, PM (*Parser Module*), recebe e interpreta os parâmetros vindos da pilha de protocolos VoIP.

O segundo, VQEM (*Voice Quality Evaluation Module*), aciona o PM para obter parâmetros, e realiza a avaliação da qualidade. O terceiro, VGM (*VQCDR Generation Module*), aciona o VQEM, monta o VQCDR e envia-o para o servidor (LUSTOSA 2005).

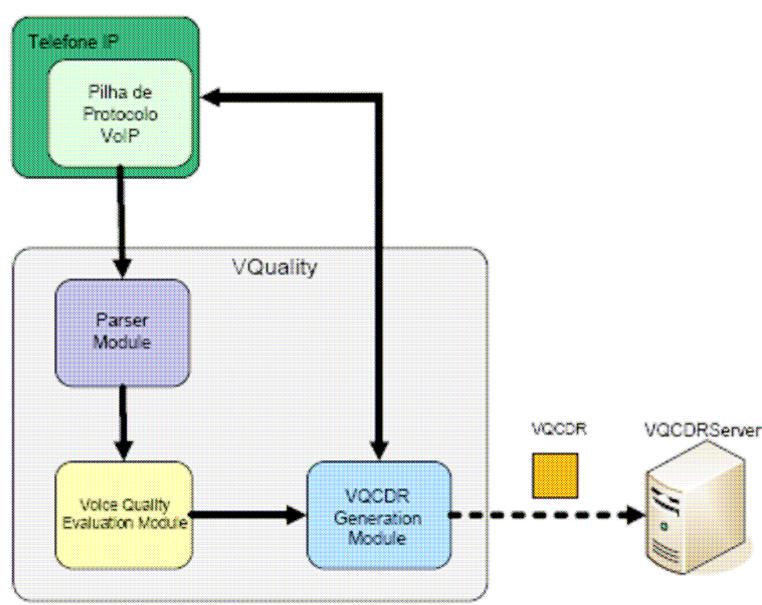


Figura 6.4 : Arquitetura da Biblioteca VQuality
(LUSTOSA 2005)

O segundo cenário a ser apresentado trata-se de um sistema de monitoração fim-a-fim em tempo real da qualidade da chamada, conhecido por VQmon (CLARK_2 2001). Vide figura 6.5 abaixo.

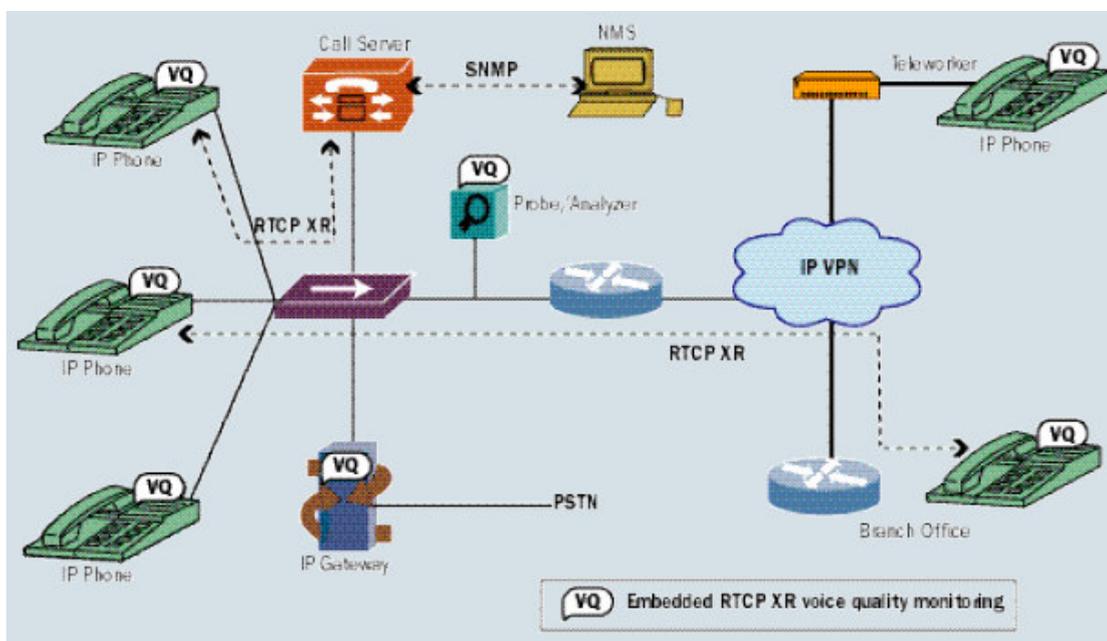


Figura 6.5 : Uso de VQmon em pontos finais de chamada
(CLARK 2004)

O VQmon fornece informações de falhas e desempenho. Neste modelo é utilizado o protocolo RTCP-XR para o reporte de perda de pacotes, atraso, e

várias métricas relacionadas a voz, incluindo nível de sinal, nível de ruído e eco (CLARK 2004).

O terceiro cenário de monitoração fim-a-fim envolve controle de admissão da chamada. Neste método a chamada só se efetivará se houver condições mínimas de qualidade na rede, para a inclusão de mais este fluxo de pacotes. Este método é chamado EMBAC (*End-to-end Measurement Based Admission Control*)(MASE 2002). O EMBAC apresenta duas formas de funcionamento : um método passivo e um ativo. O método passivo é chamado de EMBAC-CP (*Constant Probing*) e se baseia na monitoração passiva e constante da rede para estabelecer as condições mínimas para a permissão da chamada. O método ativo chama-se EMBAC-DP (*on Demand Probing*).

Na figura 6.6 observa-se o estabelecimento de uma chamda utilizzando o método EMBAC-DP (Ativo). Antes de estabelecer uma chamada o ponto originador e o ponto receptor trocam uma certa quantidade de pacotes para monitorar as condições da rede para a execução da chamada. Esta troca leva de um a dois segundos conforme (MASE-ICC 2004). Se houver condições a chamda é estabelecida, senão é abortada.

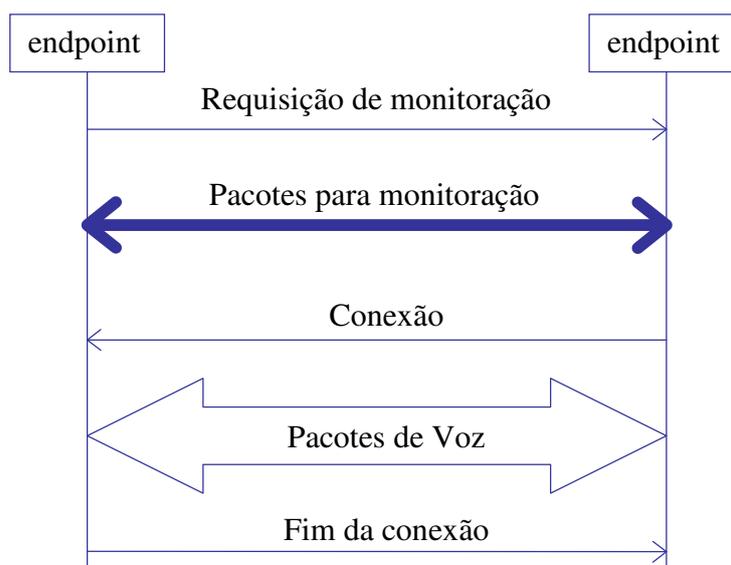


Figura 6.6 : Estabelecimento de chamada com EMBAC-DP
(MASE 2002)

O método ativo tem a vantagem de ter uma funcionalidade mais simples, fazendo a monitoração apenas no momento necessário. E tem como desvantagem a adição de tempo antes de estabelecer a chamada (MASE-ICM 2004).

Limitações :

O uso de monitores baseados nos *endpoints*, pelo fato de serem soluções particulares, e ainda não haver um padrão, torna obrigatória a seleção de uma das soluções possíveis, limitando a escolha dos componentes da rede.

A solução baseada em controle de admissão passiva, EMBAC-CP, há uma carga de processamento maior, visto que a monitoração é constante. No EMBAC-DP, o processamento é mais leve, há um pequeno atraso no estabelecimento da chamada.

Neste capítulo foram apresentadas quatro formas de avaliação de qualidade, suas aplicações e limitações. É fácil verificar que cada solução tem sua aplicação específica, podendo não servir para alguma aplicação diferente.

No próximo capítulo é proposta uma metodologia flexível onde podem ser utilizadas as mais diversas formas para a obtenção da avaliação da qualidade de uma rede para o tráfego de voz sobre IP.

7 METODOLOGIA PROPOSTA PARA AVALIAÇÃO DE REDES DE VOZ SOBRE IP

O objetivo da metodologia é sistematizar o procedimento de avaliação de redes para o tráfego de voz sobre IP levando em conta as condições do cenário a ser avaliado.

7.1 O Procedimento em 7 Blocos

Inicialmente dividimos o procedimento em sete partes: Objetivo, Ambiente, Ferramentas, Abordagem, Aplicação, Análise e Replanejamento. Estes sete grandes blocos são detalhados nos itens a seguir.

7.1.1 Bloco OBJETIVO

Neste bloco se descreve o que se quer ver ou medir, se descreve o objetivo ou o problema a ser analisado. A definição clara do OBJETIVO é primordial para a escolha do tipo de FERRAMENTA de monitoração e para a definição da ABORDAGEM a ser usada, como será visto no item **7.1.4 Bloco ABORDAGEM**.

A lista não exaustiva a seguir apresenta quatro possíveis objetivos dentro de uma avaliação de qualidade de voz sobre IP :

- Análise de uma rede de dados para inclusão de voz sobre IP : esta é uma situação real facilmente encontrada, onde já existe uma rede de dados funcionando a contento, na qual pretende-se incluir o tráfego de voz.
- Monitoração de um enlace específico : esta é outra situação real, onde já trafega voz sobre IP, e se deseja analisar algum problema numa parte específica da rede.
- Supervisão de uma rede VoIP : ocorre quando já se tem uma rede IP com tráfego de voz e pretende-se supervisionar esta rede para controlar ou ter informações de sua qualidade.
- Monitoração num período específico : nesta situação a questão é a ocorrência de problemas de qualidade de voz num determinado momento, voltando a normalidade após o término deste período.

7.1.2 Bloco AMBIENTE

Depois do OBJETIVO definido, busca-se tomar conhecimento sobre o cenário, a arquitetura e/ou as restrições do ambiente a ser analisado. Este novo bloco recebe o nome de AMBIENTE. A seguir são listadas situações e condições de ambientes possíveis de serem encontradas.

7.1.2.1 *Redes Internas e/ou Redes Externas*

Em redes internas a monitoração é facilmente realizada, com todos os pontos normalmente disponíveis. A autorização para intromissão está mais perto do responsável pela implementação do sistema de voz sobre IP, ao contrário da situação com Redes Externas.

A monitoração em redes externas pode sofrer alguma dificuldade por falta de acesso a alguns pontos a serem monitorados. O problema de acesso aumenta na medida que se aproxima ou até se insere no *backbone* da rede externa.

A monitoração nos *backbones* realizada por pessoas ou entidades sem acesso privilegiado a rede pode inviabilizar a descoberta das causas reais dos fenômenos observados. Athina Markopoulou realizou vários trabalhos de medições de perda e de atraso em *backbones* da Internet (MARKOPOULOU 2002), (MARKOPOULOU 2003), (MARKOPOULOU 2005), mas por falta de permissão de acesso interno a rede Athina pôde apenas inferir as causas prováveis.

7.1.2.2 *Interconexão entre Redes de Voz*

A interconexão entre redes de Voz é realizada por elementos de intermediação, chamados *gateways*, que convertem a sinalização e os fluxos multimídia entre as duas diferentes soluções. Estas soluções podem ser de voz sobre IP ou de telefonia convencional.

Em julho de 2005 foi lançada a RFC4123 que descreve requisitos para a entidade lógica que intermedia SIP e H.323 (SCHULZRINNE 2005).

A interconexão entre um rede de voz sobre IP e a rede telefônica pública comutada é realizada por *gateways* de voz, que de um lado implementam uma solução de voz sobre IP (por exemplo, SIP ou H.323) e do outro simulam, ou aparelhos telefônicos, ou linhas telefônicas convencionais.

Nestes elementos de intermediação é possível inserir funções de monitoramento de qualidade.

7.1.2.3 *Permissão de Acesso a Rede*

Em redes cujo acesso é restrito a determinados períodos do dia pode haver dificuldade de monitoração de algum problema que ocorra fora do período de acesso permitido. Por exemplo: se o acesso só é permitido fora do horário de expediente, a monitoração de um problema que ocorre em determinado horário do expediente fica prejudicada. Será necessário disparar uma monitoração passiva não intrusiva antes do expediente e só poderá ser vista após o mesmo, sem nenhuma possibilidade de ação durante a ocorrência do problema.

7.1.2.4 Utilização de RTCP

O uso de RTCP, um protocolo não obrigatório, facilita a coleta de informações sobre a qualidade da conexão entre fontes e destinos das chamadas de voz sobre IP. Como visto no capítulo 2, o RTCP fornece informações sobre perda de pacotes, *jitter* e atraso, não fornecidos pelo RTP. E no caso de usar RTCP-XR tem-se ainda as informações de fator R e MOS. Infelizmente a não obrigatoriedade de uso deste protocolo restringe o seu uso, já que nem todos pontos iniciais e finais de chamada são implementados com RTCP.

7.1.2.5 Uso de criptografia nos enlaces

Em situações onde se usa criptografia dos dados, como por exemplo enlaces de Redes Privativas Virtuais, a monitoração dos pacotes VoIP não é possível, restringindo esta monitoração aos pontos iniciais e finais de chamada.

7.1.2.6 Uso de soluções de VoIP com especificação aberta ou fechada

Soluções de VoIP de especificação aberta, tais como SIP e H.323 já discutidas no capítulo 2, permitem um maior domínio sobre a solução para a escolha do tipo de monitoração. Soluções fechadas, como por exemplo o Skype (BASED 2004), não permitem a monitoração em estágios intermediários, permitindo apenas a monitoração baseada em sinais, após os pacotes de voz serem transformados novamente em sinais.

7.1.3 Bloco FERRAMENTAS

Neste bloco descreve-se o tipo de ferramentas disponíveis para as medições. A seguir são listadas ferramentas possíveis de serem encontradas.

7.1.3.1 Localização dos monitores

Os monitores podem ser colocados em pontos distribuídos pela rede, ou nos gateways de voz, ou nos pontos iniciais e finais da chamada.

A realização da monitoração em pontos distribuídos pela rede pode ajudar na obtenção de informações estatísticas e na geração de valores médios da qualidade do tráfego de voz. Além disto os monitores distribuídos também auxiliam na determinação de segmentos de rede com menor qualidade, entre um ponto de monitoração e outro. Se os pacotes forem criptografados não será possível a realização da monitoração.

A monitoração realizada nos *Gateways* de VoIP garante o uso das informações disponíveis no *Gateway* sobre o tráfego de voz controlado por ele e que passa por ele. Nesta situação perdem-se as informações que não passam pelo *gateway*.

A monitoração feita nos pontos iniciais e finais da chamada garante a aquisição de informações de todas as ligações recebidas e executadas nestes pontos.

7.1.3.2 *Monitores internos e externos*

Os monitores podem ser internos ou externos aos elementos já existentes na rede.

Monitores internos são aplicativos que, ou já fazem parte do elemento da rede, ou podem ser instalados no elemento. Eles trazem a vantagem de não necessitarem de obtenção de outro ponto de rede. Normalmente possuem restrições de precisão e abrangência de medida (podem não medir um dos fatores de qualidade, por exemplo atraso, ou *jitter*, ou perda,...), mas são úteis na maioria das situações.

Monitores externos são equipamentos com o fim específico de monitoração. Além de executar monitoração também podem ser geradores de tráfego artificial. Normalmente são equipamentos caros, e por isso sua aquisição ou aluguel nem sempre são justificáveis.

7.1.3.3 *Monitores intrusivos e não intrusivos*

Os monitores intrusivos geram tráfego artificial na rede. Têm a vantagem de criar uma situação mais controlável, mas por outro lado não estarão realizando a monitoração de um tráfego real. Podem ser usados para testes de viabilidade da rede para posterior instalação de serviços de voz (BERNEX 2005).

Os monitores não intrusivos tem a vantagem de analisar chamadas reais e não injetar tráfego extra na rede (COTTRELL 2005).

7.1.3.4 *Monitoração baseada em parâmetros, em sinais e em comparação*

- Equipamentos intrusivos
- Equipamentos não intrusivos

7.1.3.5 *Monitoração baseada em opinião, fala e pacotes*

Takahashi e Yoshino classificaram as medições de qualidade em modelos de opinião, modelos baseados na fala e modelos baseados nos pacotes IP (TAKAHASHI 2004). Os modelos de opinião correspondem aos métodos subjetivos de SUN, apresentados no capítulo 4. Os modelos baseados na fala correspondem aos métodos baseados em sinais de SUN. E os modelos baseados nos pacotes IP são equivalentes aos métodos baseados em parâmetros de SUN.

Uma monitoração baseada em opinião é subjetiva e dispendiosa, e para que se tenha um valor estatístico deve obter a opinião de várias pessoas para um mesmo momento monitorado. Não me parece o método mais apropriado para avaliação de casos práticos.

Uma monitoração baseada na fala implica o uso de monitores baseados em sinais, normalmente colocados nos pontos iniciais e finais da chamada. Como já visto no capítulo 4, pode –se utilizar monitoração intrusiva, caso do uso do PESQ, ou monitoração não intrusiva, caso do uso do P.SEAM. Relembrando, o P.SEAM é bem menos preciso que o PESQ.

Uma monitoração baseada em pacotes utiliza monitores baseados em parâmetros. São monitores mais fáceis de serem encontrados já disponíveis

nos elementos da rede, mas só podem ser usados desde que o tráfego analisado não esteja dentro de pacotes criptografados.

7.1.4 Bloco ABORDAGEM

Este bloco é responsável pela descrição da forma que será realizada a monitoração, do que será feito, como e quando será feito.

Este bloco usa as informações dos blocos OBJETIVO, AMBIENTE e FERRAMENTAS para definir a ABORDAGEM a ser utilizada, como mostra a figura 7.1.

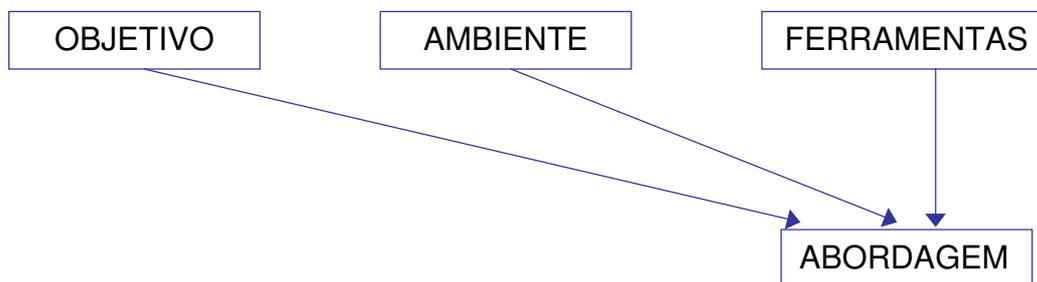


Figura 7.1 : Tomada de informações para definir a ABORDAGEM

A seguir são listadas vários tipos de Abordagem possíveis.

7.1.4.1 Utilização de Monitoração Ativa ou Passiva

A abordagem de monitoração ativa é pertinente em testes de viabilidade da rede para posterior instalação de serviços de voz , ou em testes interativos de segmento em segmento de rede, para isolamento de um segmento com falha.

A abordagem de monitoração passiva deve ser usada quando se deseja avaliar a qualidade de chamadas reais.

7.1.4.2 Utilização de uma ou de múltiplas Probes

A utilização de múltiplas *probes* proporciona a análise simultânea de várias chamadas, ou então de vários segmentos da rede. É indicada para análises periódicas da qualidade da rede.

Ocorre que nem sempre é possível ou necessário múltiplas *probes* para a avaliação de qualidade. Em situações onde a rede local não apresenta problemas e há um único ponto de acesso a rede externa, um único monitor localizado neste ponto de gargalo já é suficiente para a avaliação da qualidade de voz.

7.1.4.3 Local da inclusão da Ferramenta de Monitoração

A colocação dos monitores em pontos distribuídos pela rede podem ajudar na obtenção de informações estatísticas e geração de valores médios da qualidade do tráfego de voz. Além disto também auxiliam na determinação de

segmentos de rede com menor qualidade (entre um ponto de monitoração e outro). Se os pacotes forem criptografados não será possível a realização de monitoração.

A colocação dos monitores nos *Gateways* de VoIP garante o uso das informações disponíveis no *Gateway* sobre o tráfego de voz controlado por ele e que passa por ele. Nesta situação perde-se as informações que não passam pelo *gateway*.

A monitoração feita nos softfores garante a aquisição de informações de todas as ligações recebidas e executadas por ele.

7.1.4.4 Momento da Monitoração

Em redes de acesso restrito pode-se deparar com o problema de ter que realizar uma monitoração exatamente num período em que não se tenha acesso a rede. Para estes caso é necessário instalar monitores passivos no período de acesso permitido, e se possível programá-los para realizarem a monitoração automaticamente durante o período necessário. Nesta situação perde-se a possibilidade de alterações necessárias na configuração da rede ou do monitor no momento da ocorrência do problema.

7.1.4.5 Duração da Monitoração

Para períodos curtos de monitoração, quando o problema é facilmente reproduzível ou detectável, não é necessário maiores cuidados com a quantidade de informação sobre a monitoração a ser guardada.

Já em casos de longos períodos de monitoração, toma-se o cuidado para que os dados de monitoração sejam guardados em locais com capacidade suficiente. Se estes dados forem armazenados em local diferente que o elemento de rede onde estão sendo coletados, deve-se levar em consideração o fato de que está sendo gerado tráfego extra por conta da transferência dos dados, e assim alterando-se o fluxo de dados do ambiente analisado.

7.1.4.6 Formas de localização da fonte do problema

A localização da fonte do problema de qualidade, segundo (BERNEX 2005), pode ser realizada de três maneiras : análises simultâneas, análise interativa e análise estatística.

O uso de análises simultâneas é uma forma não-intrusiva que exige a colocação de monitores em vários pontos, atuando independente e simultaneamente. Para uma mesma chamada os monitores provêm diferentes resultados, de acordo com sua localização no caminho, identificando o segmento com problema.

A análise interativa é uma forma intrusiva que ao invés de analisar a chamada em vários pontos simultaneamente, insere monitores que geram e analisam chamadas de um ponto da rede a outros interativamente até descobrir o ponto de falha.

A análise estatística é uma forma intrusiva que gera tráfego artificial e obtém informações dos vários elementos de rede através, por exemplo, de

requisições SNMP, para posteriormente analisar os resultados e localizar os possíveis pontos de falha.

7.1.5 Bloco APLICAÇÃO

Após definida a ABORDAGEM a ser feita para atingir o OBJETIVO proposto aplica-se esta ABORDAGEM utilizando as FERRAMENTAS disponíveis e aplicáveis no AMBIENTE em questão. Esta fase do procedimento de Avaliação pertence ao bloco APLICAÇÃO.

7.1.6 Bloco ANÁLISE

Depois de realizada a APLICAÇÃO da ABORDAGEM escolhida no AMBIENTE a ser avaliado são verificados os resultados das monitorações da Qualidade da Voz. Se o OBJETIVO traçado não foi atingido o procedimento continuará através do bloco REPLANEJAMENTO, se o OBJETIVO foi atingido então o procedimento termina.

Numa situação onde o OBJETIVO é verificar se a rede possui condições para o tráfego de voz pode-se chegar a conclusão após a ANÁLISE que as condições podem ser melhoradas. Neste caso, o OBJETIVO inicial foi atingido, a verificação da condição de tráfego de voz, mas o procedimento não termina, seguindo para o bloco REPLANEJAMENTO, onde entre outras coisas, altera-se o OBJETIVO.

7.1.7 Bloco REPLANEJAMENTO

Neste bloco se redefinem ou se aprimoram os OBJETIVOS, podendo-se atuar no AMBIENTE para melhorar/modificar a condição da rede para resolver o problema definido no OBJETIVO e/ou buscando-se novas FERRAMENTAS, e reavaliando-se/aprimorando-se/mudando-se a ABORDAGEM para a APLICAÇÃO.

7.1.8 O Conjunto Final com todos os Blocos

A finalização do procedimento ocorre após a ANÁLISE, caso o OBJETIVO tenha sido plenamente realizado. Em caso negativo é necessário um REPLANEJAMENTO, iniciando um novo ciclo, com possíveis redefinições de OBJETIVO, AMBIENTE, FERRAMENTAS e ABORDAGEM. Depois disto é necessária uma nova APLICAÇÃO da ABORDAGEM no cenário em questão e uma nova ANÁLISE, como mostra a figura 7.2. O ciclo se repete até que na ANÁLISE verifique-se um resultado plenamente satisfatório, de acordo com o último OBJETIVO declarado.

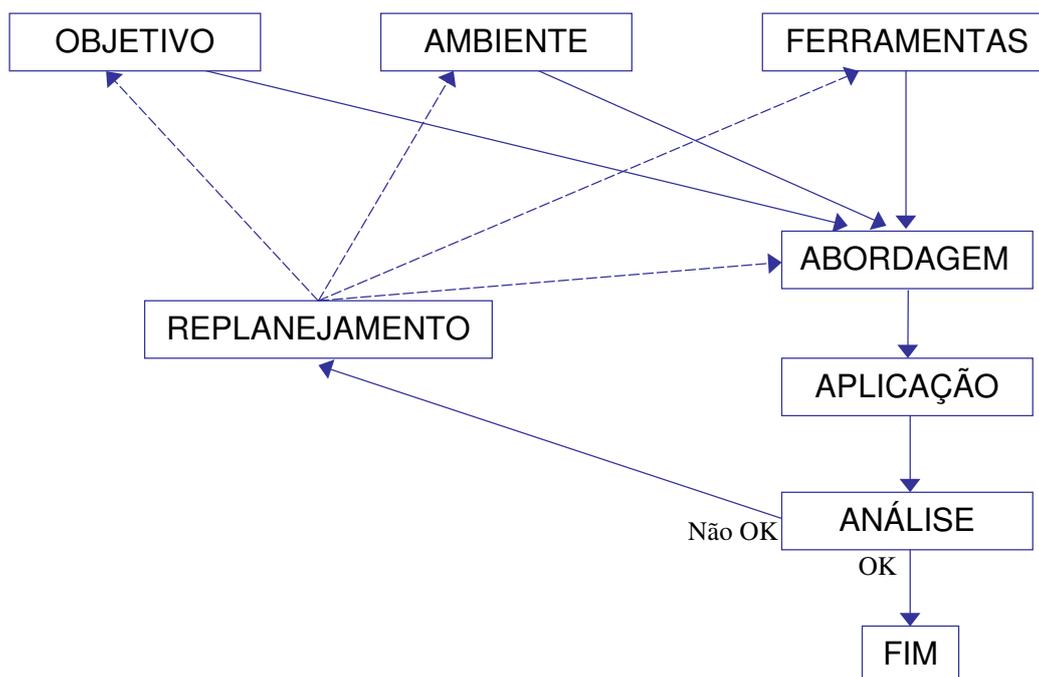


Figura 7.2 : Procedimento completo com os sete blocos

7.2 Validação do Método

O caso utilizado para a validação do método desta dissertação refere-se a avaliação da viabilidade de colocação de voz sobre IP num ponto específico da rede da UFRGS.

Os itens de 7.2.1 a 7.2.7 descrevem o procedimento seqüenciado nos sete blocos descritos no item 7.1.

7.2.1 Bloco OBJETIVO

Como já dito o OBJETIVO deste procedimento é a verificação de viabilidade de tráfego de voz sobre IP num enlace específico onde originalmente só trafega dados. O enlace será descrito no item AMBIENTE, a seguir.

7.2.2 Bloco AMBIENTE

O AMBIENTE a ser analisado situa-se na Estação Agronômica da UFRGS, na cidade de Eldorado, no Rio Grande do Sul. Trata-se de um ponto remoto, completamente fora das sub-redes internas da Universidade.

O acesso a Internet é realizado por um enlace PPP assíncrono através de um modem ADSL fornecido pela operadora Brasil Telecom, fornecendo uma banda de 400Kbits/s de *downstream* e 200Kbits/s de *upstream*. O acesso a Internet oferecido pela operadora Brasil Telecom é destinado para tráfego de dados e desprovido de Qualidade de Serviço.

A conexão de até 4 aparelhos telefônicos convencionais a rede de voz sobre IP é feita através de um *Gateway* de voz com interfaces FXS, da Digitel.

O protocolo de sinalização e controle a ser utilizado é o SIP. As requisições e respostas SIP são centralizadas num *Proxy Server* localizado no CPD da UFRGS, localizado no Campus Médico, em Porto Alegre. A implementação do Proxy Server foi realizada de um *SIP Express Router* (SER). O SER é um servidor de código aberto de alto desempenho que foi desenhado para implementar infraestruturas de telefonia IP em larga escala (VETTER 2006).

O SER faz parte do ambiente de VoIP da UFRGS, que por sua vez faz parte da rede fone@RNP da Rede Nacional de Pesquisa. A RNP utiliza H.323 para sinalização e controle. A interoperabilidade entre SIP e H.323, dentro da UFRGS, é feita através de um *gateway* SIP-H.323 implementado com o software aberto Asterisk.

A figura 7.3 ilustra todo o cenário para a validação do método.

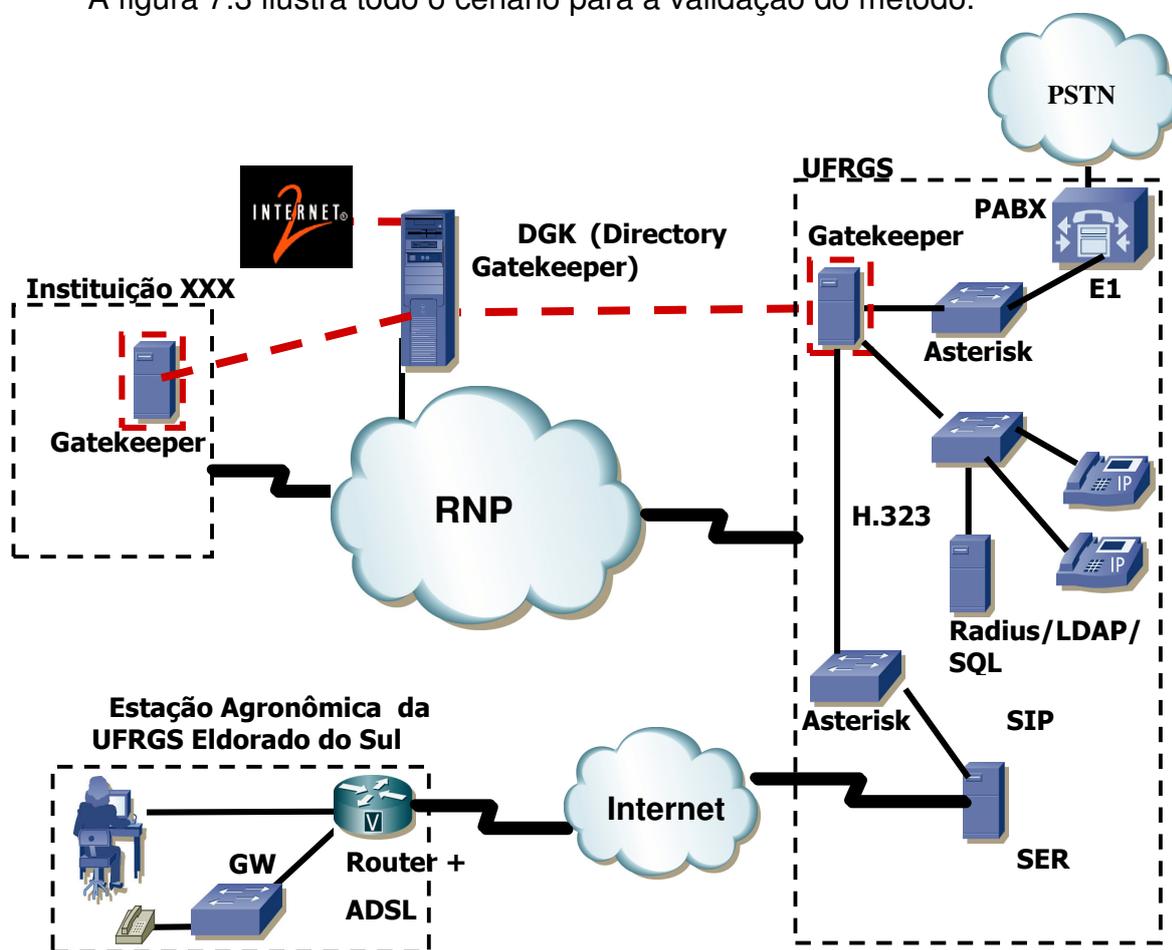


Figura 7.3 : Cenário para a Validação do Método

Para a interconexão do ambiente VoIP- h.323 com a telefonia convencional presente na UFRGS utiliza-se outro *gateway* implementado através do Asterisk.

Pode-se definir o AMBIENTE como sendo: uma interconexão de duas redes internas, uma em Eldorado e outra em Porto Alegre, através da Internet, sem qualidade de serviço. A intrusão é permitida, com acesso sempre permitido, e a rede não utiliza criptografia para a transferência dos pacotes de voz.

A seguir a descrição do tipo de ferramenta disponível e utilizada.

7.2.3 Bloco FERRAMENTAS

Para o OBJETIVO de analisar a rede para a inclusão de VoIP, temos como FERRAMENTA um aplicativo, nativo do próprio *Gateway* de Voz, que fornece informações de *jitter* máximo, *jitter* médio e desvio padrão, além de percentual de perdas e número de pacotes recebidos. Para nossas avaliações iniciais esta ferramenta nos pareceu suficiente.

7.2.4 Bloco ABORDAGEM

A ABORDAGEM usada para a verificação de viabilidade de tráfego de voz sobre IP no enlace especificado foi primeiramente verificar o funcionamento do *Gateway* de Voz modelo DVS2410 (4 portas FXS, com SIP) no ambiente de voz sobre IP do CPD da UFRGS para promover a familiarização dos profissionais com as formas de configurar o equipamento e com os recursos e características do mesmo.

7.2.5 Bloco APLICAÇÃO

A primeira APLICAÇÃO foi realizada no CPD da UFRGS onde instalou-se o *gateway* de voz DVS2410 configurando-o via scripts de configuração. Foi colocado um telefone convencional numa das portas FXS do equipamento e pôde-se verificar o funcionamento com o aparelho telefônico realizando e recebendo chamadas através do *gateway*.

7.2.6 Bloco ANÁLISE

Para a colocação de tráfego de voz sobre IP no local definitivo, na estação Agrônômica, considerou-se necessário o estabelecimento de uma Rede Privada Virtual até o CPD da UFRGS por motivo de segurança.. Por conta disto, far-se-á um REPLANEJAMENTO para a verificação de viabilidade de tráfego de voz no enlace em questão.

7.2.7 Bloco REPLANEJAMENTO

VPN ou Rede Privada Virtual pode ser feita de várias formas. A primeira forma sugerida foi solicitar que a operadora Brasil Telecom fornecesse a VPN. Esta forma foi descartada em função do acréscimo no custo da solução. A Segunda forma foi solicitar ao fabricante do *gateway* que o mesmo incluísse o protocolo de tunelamento PPTP no DVS2410. Desta forma, através de um servidor Windows no CPD, estabelecer-se-ia uma VPN até o site me Eldorado.

Como os blocos OBJETIVO, AMBIENTE e FERRAMENTA continuam os mesmos, o próximo passo fica sendo a nova ABORDAGEM.

7.2.8 Bloco (nova) ABORDAGEM

A nova ABORDAGEM não altera a estrutura física do local onde se quer inserir voz sobre IP, mas agrega uma nova função ao *Gateway* de Voz : o estabelecimento de uma VPN da Estação Agronômica até o CPD da UFRGS em Porto Alegre.

7.2.9 Bloco (nova) APLICAÇÃO

Depois de estabelecido o túnel PPTP e configurado o *gateway* adequadamente inseriu-se um telefone convencional numa de suas portas. O funcionamento da rede com voz sobre IP foi observado por uma semana, forçando a realização de chamadas em várias horas do dia.

7.2.10 Bloco (nova) ANÁLISE

No período de avaliação da rede para o tráfego de voz foi verificado que a qualidade das ligações observada por um usuário era muito boa, praticamente não percebendo diferença entre uma ligação VoIP e uma via telefonia convencional.

Desta forma verificou-se ser viável o tráfego de voz sobre IP no enlace específico entre a estação Agronômica da UFRGS e a rede da UFRGS em Porto Alegre.

8 CONCLUSÃO

A presente dissertação apresentou uma proposta de metodologia para sistematizar a avaliação de qualidade de Redes para o tráfego de Voz sobre IP de acordo com as possibilidades disponíveis no cenário a ser avaliado.

O segundo capítulo descreveu os principais protocolos utilizados em voz sobre IP.

No terceiro capítulo foram mostrados os fatores que influenciam na qualidade de voz.

Uma apresentação das características e detalhes dos métodos de avaliação de qualidade de voz foi feita no quarto capítulo.

Nos estudos dos vários métodos de medição verificou-se que métodos baseados em sinais tem complexidade e necessidade de processamento maior que métodos baseados em parâmetros. Os estudos também revelam que monitoração ativa ou passiva, ou de outra forma, medições intrusivas ou não intrusivas, possuem seus prós e contras, havendo uso e necessidade de ambos. Os métodos ativos/intrusivos permitem situações mais controladas, mas não registrarão transientes que ocorram fora do momento da medição. Os métodos passivos/não-intrusivos por sua vez, podem monitorar continuamente, todas as chamadas, capturam transientes, provem dados para análises estatísticas posteriores, mas necessitam chamadas ativas para realizar monitorações.

Como o termo “qualidade de voz” é amplo e pode sofrer várias interpretações, o autor desta dissertação considerou necessário mostrar as várias abordagens diferentes e específicas para o termo. Isto foi mostrado no quinto capítulo.

No sexto capítulo, antes de abordar a proposta da dissertação, foram apresentados trabalhos correlatos.

O objetivo principal desta dissertação, a proposta de uma metodologia para sistematizar a avaliação de viabilidade de tráfego de voz em redes IP, foi abordado no sétimo capítulo. Para a caracterização dos sete blocos de informações utilizadas no método foram estudadas várias formas de monitoração e diversos aspectos e dimensões sobre qualidade de voz em redes IP. A coletânea disto propiciou uma categorização de informações e a sistematização de um procedimento para avaliação de qualidade. A validação desta metodologia foi executada com o estudo de um caso prático.

Como sugestões de trabalhos futuros podemos citar o uso da metodologia em vários casos reais e o conseqüente aprimoramento e evolução do método. Outra proposta é a criação de um software aplicativo para facilitar o uso do método.

REFERÊNCIAS

- ANDERSON, J. **Methods for measuring perceptual speech quality**. [S.l.]: Agilent Technologies, 2002. Disponível em: <<http://www.onenetworks.com/WhitePapers.asp>> . Acesso em: set. 2005.
- ARANGO, M. et al. **Media Gateway Control Protocol (MGCP) Version 1.0: RFC 2705**. [S.l.]: Internet Engineering Task Force, Network Working Group, 1999.
- ATKINSON, D.J. **Proposed Annex A to Recommendation P.861**. Genebra-Suíça: ITU-T, Study Group 12, 1997. (Contribution COM 12-24).
- BADDELEY, A. D. **Human Memory: Theory and Practice**. [S.l.]: Psychology Press: Hove Erlbaum Associates, 1997.
- BASET, S.; SCHULZRINNE, H. **An Analysis of the Peer-to-Peer Internet Telephony Protocol**. New York. Columbia University, 2004.
- BEERENDS, J. G.; STEMERDINK, J. A. A perceptual speech quality measure based on a psychoacoustic sound representation. **Journal of Audio Engineering Society**, [S.l.], v.14, p.115–123, Mar. 1994.
- BEERENDS, J. G.; MEIJER, E. J.; HEKSTRA, A. P. **Improvement of the P.861 Perceptual Speech Quality Measure**. Genebra-Suíça: ITU-T, Study Group 12, 1997. (Contribution COM 12-20).
- BEERENDS, J. G. et al. **Proposed Draft Recommendation P.862: Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs**. Genebra-Suíça: ITU-T, Study Group 12, 2000.
- BERGER, J. Future work on objective speech quality measurements in ITU. In: WORKSHOP ON WIDEBAND SPEECH QUALITY IN TERMINALS AND NETWORKS: ASSESSMENT AND PREDICTION, 2004, Mainz, Germany. **Proceeding...** [S.l.: s.n.], 2004.
- BERNAL FILHO, H. **Tutorial sobre Telefonia IP**. [S.l.], 2003. Disponível em: <<http://www.teleco.com.br/tutoriais/tutorialtelip/default.asp>> Acesso em: nov. 2005.
- BERNEX, E.; GATINEAU, A. **Quality of service in VoIP environments**. Lannion-França: NeoTIP, 2005.

BOGER, Y. **Fine-tuning Voice over Packet services**. Paramus, NJ: RADCOM, 2005. Disponível em: <<http://www.radcom.com/radcom/WhitePapers/voip2.htm>> . Acesso em: set. 2005.

BRIX NETWORKS. **Advanced VoIP Test Suite**. Shelmsford-MA-EUA, 2005. Disponível em: <<http://www.brixnet.com>>. Acesso em: dez. 2005.

BROOM, S.; HOLLIER, M. Speech Quality Measurement Tools for Dynamic Network Management. In: MEASUREMENT OF AUDIO AND VIDEO QUALITY IN NETWORKS, Praga, República Tcheca. **Proceeding...** [S.l.: s.n.], 2003.

CHONG H. et al. Rapid System Engineering Assessment of Voice-over-IP Deployment Decisions. **Technology Review Journal**, [S.l.], Spring/Summer 2005.

CISCO SYSTEMS. **Understanding Delay In Packet Voice Networks**. [S.l.], 2005. Disponível em: <<http://www.cisco.com/warp/public/788/voip/delay-details.html>>. Acesso em: set. 2005.

CISCO, SYSTEM. **Understanding Packet Voice Protocols**. [S.l.] 2005. Disponível em: <http://www.iec.org/online/tutorials/acrobat/packet_voice.pdf>. Acesso em: set. 2005.

CLARK, A. **Improving VoIP call quality with Embedded Monitoring**. [S.l.], 2001. Disponível em: <<http://www.commsdesign.com>>. Acesso em: set. 2005.

CLARK, A. Managing Enterprise VoIP: The Case for a New Standard. **Business Communications Review**, [S.l.], June 2004.

CLARK, A. **Extensions to the E Model to incorporate the effects of time varying packet loss and recency**. Suwanee,GA: Telchemy Incorporated, 2001. (T1A1.1/2001-037).

CLARK, A. Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality. In: IP – TELEPHONY WORKSHOP, 2001, New York, EUA. **Proceedings...** [S.l.: s.n.], 2004.

CLARK, A. QoS Tips and Tricks for VoIP Services : Delivering reliable VoIP Services. In: ISPCON, **ano de realização do evento**, Baltimore, USA. **Proceedings...** [S.l.: s.n.], 2005.

CLARK, A. Voice Performance Management. In: INTERNET TELEPHONY CONFERENCE & EXPO Fall 2005, Los Angeles, USA. **Proceedings...** [S.l.: s.n.], 2005.

CLARK, A. Fault and Performance Management for Next Generation IP Communication. In: VoIP DEVELOPER CONFERENCE, San Francisco, USA. **Proceedings...** [S.l.: s.n.], 2005.

CLARK, A. Voice quality measurement : understanding VoIP. **Electronic Engineerings Times**, [S.l.], oct. 2005. Disponível em: <http://www.eetasia.com/ARTICLES/2005OCT/B/2005OCT03_RFD_NETD_TA.pdf> . Acesso em: dez. 2005.

CLAUDE, F. **Protocolos e Aplicações para Rede de Computadores**. [S.l.]: Ed. Érica, 2002.

COLE, R. G.; ROSENBLUTH, J. H. Voice over IP performance monitoring. **Computer Communication Review**, New York, v.31, n. 2, Apr. 2001.

COTTRELL, L. **Internet Monitoring**. [S.l.], 2005. Disponível em: <<http://www.slac.stanford.edu/grp/scs/net/talk05/niit-measure-mar05.ppt>>. Acesso em: dez. 2005.

COX, R.; PERKINS, R. **Results of a Subjective Listening Test for G.711 with Frame Erasure Concealment**. [S.l.]: AT&T, 1999. (T1A1.7/99-016).

DAL PIVA, J. B. **QMON: Uma Ferramenta para Análise de Qualidade de Chamadas Telefônicas em uma Rede SIP**. 2004. Trabalho de Conclusão (Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre.

DAVID, F. **Ferramentas de Monitoração Ativa e Passiva para Avaliação da Qualidade de Redes VoIP**. 2003. 111 f. Dissertação (Mestrado em Informática) – Instituto de Matemática – IM, Núcleo de Computação Eletrônica – NCE, UFRJ, Rio de Janeiro.

DOULAMIS A. D. et al. An Adaptable Neural-Network Model for Recursive Nonlinear Traffic Prediction and Modeling of MPEQ Video Sources. **IEEE Transactions on Neural Networks**, New York, v. 14, p. 150–166, Jan. 2003.

ETSI. **Speech Communication Quality for Mouth to Ear for 3.1 kHz Handset Telephony across Networks**. [S.l.], 1996. (Technical Report ETR 250).

ETSI **TS 101 329-5 v1.1.2: Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 5: Quality of Service (QoS) measurement methodologies**. França, 2002.

FERNANDES, N. **Relação entre a Qualidade das Respostas das Recomendações G.723.1 e G.729, e o Comportamento da Rede IP de Suporte** 2003. 161 f. Dissertação (Mestrado em Engenharia de Sistemas e Computação) – COPPE/UFRJ, Rio de Janeiro.

FRANCE TELECOM. **Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of a recency effect**. [S.l.]: ITU Study Group 12 D.139, 2000.

FRIEDMAN, T.; CACERES, R.; CLARK, A. **RTP Control Protocol Extended Reports (RTCP XR)**: RFC 3611. [S.l.]: Internet Engineering Task Force, Network Working Group, 2003.

GOLDSTEIN, T.; BEERENDS, J. A.; RIX, A. W. **Draft Recommendation P.AAM: An objective method for end-to-end speech quality assessment of narrow-band telephone networks including acoustic terminal(s)**. Genebra-Suíça: International Telecommunications Union, 2003. (ITU-T Contribution COM 12-C64).

GOLDSTEIN, T.; RIX, A. Perceptual Assessment in acoustic and binaural applications. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 2004. **Proceedings...** [S.l.:s.n.], 2004.

GOMIDE, C. **Multiplexação de Sinais de Voz em Redes IP**. 2004. 110 f. Dissertação (Mestrado em Computação) – Centro Universitário Eurípides de Marília, Fundação de Ensino Eurípides Soares da Rocha, Marília.

GRAY, P. et al. Non-intrusive speech-quality assessment using vocal-tract models. **IEE Proc.-Vision Image Signal Process**, New York, v.147, n. 6, p. 493-501, Dec. 2000.

GROVES C. et al. **Gateway Control Protocol Version: RFC 3525**. [S.I.]: Internet Engineering Task Force, Network Working Group, 2003.

HANDLEY, M.; JACOBSON V. **SDP: Session Description Protocol: RFC 2327**. [S.I.]: Internet Engineering Task Force, Network Working Group, 1998.

HARDMAN, D. **Noise and Voice Quality in VoIP Enviroments**. Colorado Springs-USA: Agilent Technologies Inc., 2003.

HARDY, W. C. **QoS Measurement and Evaluation of Telecommunications Quality of Service**. [S.I.]: John Wiley & Sons, 2001.

HOLLIER P. et al. Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment, Part I – Time-Delay Compensation. **J. Audio Eng. Soc.**, [S.I.], v. 50, n. 10, 2002.

HOLLIER, M. Removing the "quality concerns" barrier to VoIP adoption in enterprises: a case study. In: IP Voice Meeting, 2005, Lisboa, Portugal. **Proceedings...** [S.I.:s.n.], 2005.

HOLUB, J. **Speech Transmission Quality Assessment Methods – Current Situation and Future Trends**. Praga-República Tcheca: Czech Technical University in Prague, Faculty of Electrical Engineering, 2004. Disponível em: <web.cvut.cz/rctu/51924/HP2004-02.pdf>. Acesso em: dez. 2005.

HUANG, C.; CHAO, C.; LIU, A. A Distributed Management Framework for H.323- Based VoIP System. In: INTERNATIONAL CONFERENCE ON COMMUNICATIONS IN COMPUTING, CIC, 2003, EUA. **Proceedings...** [S.I.:s.n.], 2003.

IEC. **Voice Quality (VQ) in Converging Telephony and IP Networks**, [S.I.], 2005. Disponível em: <http://www.iec.org/online/tutorials/acrobat/voice_qual.pdf>. Acesso em: set. 2005.

ITU-T. **The E-model, a computational model for use in transmission planning**: Recomendação G.107. Genebra-Suíça, 2005.

ITU-T. **General Characteristics of International Telephone Connections and International Telephone Circuits**: Recomendação G.114. Genebra-Suíça, 2000.

ITU-T. **Media Gateway Control Protocol**: Recommendation H.248.1 Genebra-Suíça, 2002. v.1.

ITU-T. **Packet-based multimedia communications systems**: Recommendation H.323. Genebra-Suíça, 1999.

ITU-T. **Single-ended method for objective speech quality assessment in narrow-band telephony applications**: Recomendação P.563. Genebra-Suíça, 2004.

ITU-T. **Methods for Subjective Determination of Transmission Quality**: Recomendação P.800. Genebra-Suíça, 1996.

ITU-T. **Subjective Performance Assessment of Telephone-Band And Wideband Digital Codecs**: Recommendation P.830. Genebra-Suíça, 1996.

ITU-T. **Objective Quality Measurement of Telephone Band (300-3400 Hz) Speech Codecs**: Recommendation P.861. Genebra-Suíça, 1996.

KAMIENSKI, C. Qualidade de Serviço na Internet. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, SBRC, 18., 2000, Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2000.

LAURIDSEN, O. Introduction to VoIP technology. In: VoIP CICT Conference, 2004, Lyngby, Dinamarca. **Proceedings...** [S.l.:s.n.], 2004.

LIN, M. et al. Design and Implementation of an Enterprise Call Analysis System for VoIP Deployments. In: AUSTRALIAN TELECOMMUNICATIONS, NETWORKS AND APPLICATIONS CONFERENCE, ATNAC, 2003, Australia. **Proceedings...** [S.l.:s.n.], 2003.

LUSTOSA, L. Utilização do Modelo E para avaliação da qualidade da fala em sistemas de comunicação baseados em voz sobre IP. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, SBRC, 22., 2004, Gramado. **Anais...** Gramado: UFRGS, 2004.

LUSTOSA, L. **Arquitetura de Monitoração de Qualidade de Chamadas Telefônicas IP**. 2005. 104f. Dissertação (Mestrado em Informática) – Instituto de Matemática – IM, Núcleo de Computação Eletrônica – NCE, UFRJ, Rio de Janeiro.

MARKOPOULOU, A. et al. Assessing the Quality of Voice Communications Over Internet Backbones. **IEEE/ACM Transactions on Networks**, New York, v.11, n.5, p.747 – 760, Oct. 2003.

MARKOPOULOU, A. et al, Loss and Delay Measurements of Internet Backbones. **Computer Communications**, New York, Special Issue on Monitoring and Measurements, May 2005.

MARKOPOULOU, A. **Assessing the Quality of Multimedia Communications over Internet Backbone Networks**. 2002. 230 f. Thesis (Doctor of Philosophy in Electrical Engineering) Department of Electrical Engineering - Stanford University, Stanford-CA-EUA.

MASE, K.; TOYAMA, Y. End-to-end Measurement Based Admission Control for VoIP Networks. In: IEEE INTERNATIONAL CONFERENCE ON COMMUNICATIONS, ICC, 2002. [S.l.: s.n.], 2002. **Proceedings...** [S.l.: s.n.], 2002. p.1194 – 1198.

MASE, K.; KOBAYASHI, H. An Efficient End-to-end Measurement Based Admission Control for VoIP Networks. IEEE INTERNATIONAL CONFERENCE ON COMMUNICATIONS, ICC, 2004. **Proceedings...** [S.l.: s.n.], 2004. p.1877 - 1881.

MASE, K. Toward Scalable Admission Control for VoIP Networks. **IEEE Communications Magazine**, New York, v.42, n.7, p. 42 – 47, July 2004.

MEDEIROS, M. S. **Trabalho sobre os protocolos de suporte a multimídia RTP, RTCP e RTSP**. Rio de Janeiro, 2003. Disponível em: <http://www.gta.ufri.br/grad/03_1/rtp/index.html>. Acesso em: ago. 2005.

MIRAS, D. **A Survey of Network QoS Needs of Advanced Internet Applications** [S.l.]:Internet2 QoS WG, 2002 Disponível em: <<http://qos.internet2.edu/wg/apps/fellowship/Docs/Internet2AppsQoSNeeds.pdf>>. Acesso em: set. 2005.

MOHAMED,S.; RUBINO, G. A Study of Real-Time Packet Video Quality Using Random Neural Networks. **IEEE Transactions on Circuits and Systems for Video Technology**, New York, v.12, p. 1071–1083, Dec. 2002.

MONFORT J. Basic Requirements to Quality of Service (IP centric). In: WORKSHOP ON STANDARDIZATION IN E-HEALTH, 2003,Geneva, Switzerland. **Proceedings...** [S.l.:s.n.], 2003.

NEVES, J. E. et al. Neural Networks in B-ISDN Flow Control: ATM Traffic Prediction or Network Modeling. **IEEE Communications Magazine**, New York, p. 50–56, Oct. 1995.

NORDSTROM E. et al. Neural Networks for Adaptive Traffic Control in ATM Networks. **IEEE Communications Magazine**, New York, p. 43–49, Oct. 1995.

NUNES, M.S. **Rede com Integração de Serviços** : 4ª parte - Voz sobre IP. Lisboa-PT: ITS-Universidade Técnica de Lisboa, 2003. Disponível em: <http://asterix.ist.utl.pt/ec-ris/textos-aulas/4a_parte_VoIP.pdf>. Acesso em: set. 2005.

OPTICOM GmbH. **3SQM Advanced non-Intrusive Quality Testing**, Erlangen-Alemanha, 2004. Disponível em: <<http://www.opticom.de>>. Acesso em: nov. 2005.

PARLOS, A. G. et al. Prediction of MPEQ-Coded Video Source Traffic Using Recurrent Neural Networks. **IEEE Transactions on Signal Processing**, New York,v.51, n.8, p. 2177–2190, Aug. 2003.

PERCY, A. **Understanding Latency in IP Telephony**. Needham – MA: Brooktrout Technology, 1999.

PRACHT, S.; HARDMAN, D. **Voice Quality in Converging Telephony and IP Networks**. [S.l.] : Agilent Technologies , 2001.

QUINELLATO, D. et al. Arquitetura de Monitoração de Qualidade de Chamadas Telefônicas IP. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, SBRC, 23., 2005, Fortaleza. **Anais...** Fortaleza: UFC, 2005.

RADVISION LTD. **Implementing Media Gateway Control Protocols**. [S.l.], 2002. Disponível em: <<http://www.radvision.com/ResourceLibrary/WhitePapers/>>. Acesso em: set. 2005.

RIGNEY, C. et al. **Remote Authentication Dial In User Service (RADIUS): RFC 2865.**[S.l.:s.n.], 2000.

RIX, W. et al. Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment, Part II – Psychoacoustic Model. **J. Audio Eng. Soc.**, [S.l.], v.50, n.10, 2002.

ROSENBLUTH, J. H. **Testing the Quality of Connections of Connections having Time Varying Impairments.** [S.l.]: Committee contribution, 1998. (T1A1.7/98-031).

ROWE, M. **Measure VoIP Networks for Jitter and Loss.** [S. l.]: Test & Measurement World, 1999. Disponível em: <<http://www.reed-electronics.com/tmworld/index.asp?layout=article&articleId=CA187534&rid=0&rme=0&cfd=1>>. Acesso em: ago. 2005.

SAWASHIMA, H. **Characteristics of UDP packet Loss: Effect of TCP Traffic.** [S.l.]: Inet, 1997. Disponível em: <http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/f3/f3_1.html>. Acesso em: ago.2005.

SCHULZRINNE, H. et al. **RTP: A Transport Protocol for Real-Time Applications: RFC 3550.** [S.l.]: Internet Engineering Task Force, Network Working Group, 2003.

SCHULZRINNE, H.; AGBOH, C. **Session Initiation PROTOCOL (SIP) – H.323 Internetworking Requirements: RFC 4123.** [S.l.] Internet Engineering Task Force, Network Working Group, 2005.

SIPCENTER. **MGCP/MEGACO Architecture.** [S.l.], 2005. Disponível em: <<http://www.sipcenter.com/sip.nsf/html/MGCP+Architecture>>. Acesso em: nov. 2005.

SUN, L.; IFEACHOR, E. Perceived Speech Quality Prediction for Voice over Ipbased Networks. In: IEEE INTERNATIONAL CONFERENCE ON COMMUNICATIONS, ICC, 2002, New York, USA. **Proceedings...** [S.l.:s.n.], 2002.

SUN , L. **Speech Quality Prediction for Voice over Internet Protocol Networks.** 2004. 198 f. Thesis (Doctor of Philosophy in Computing, Communications and Electronics) Faculty of Technology - University of Plymouth, Plymouth - Unided Kingdom.

TAKAHASHI, A.;YOSHINO, H. Perceptual QoS Assessment Technologies for VoIP. **IEEE Communications Magazine**, New York, v.42, n.7, p. 28 – 34, July 2004.

TELCHEMY. **Voice Quality.** [S.l.], 2003. Disponível em: <http://www.telchemy.com/references/voice_quality.html>. Acesso em: set. 2005.

TELCHEMY. **Voice Quality Measurement.** Suwanee-USA, 2005.

THORELL, C. **Avaliação da Qualidade de Ligações VoIP Baseada na Análise de Tráfego RTP**. 2005. Trabalho de Diplomação (Engenharia da Computação) - Instituto de Informática, UFRGS, Porto Alegre.

UBIK, S. IP Telephony Accounting and WAN Deployment Experience. In: IP TELEPHONY WORKSHOP, 2001, New York, USA. **Proceedings...** [S.l.:s.n.], 2001.

VETTER, M. et al. **VoIP@UFSC, uso de Voz sobre IP na Universidade**. [S.l.], 2006. Disponível em: < http://www.voip.ufsc.br/docs/artigo_voip.pdf>. Acesso em: jun. 2006.

VIANA, C.; OLIVEIRA, J. **Implementação de um Serviço de Tarifação para Serviços IP**. 2004. Trabalho de Conclusão (Engenharia Elétrica)- Faculdade de Tecnologia, Departamento de Engenharia Elétrica, UnB, Brasília.

VORAN, S. Objective estimation of perceived speech quality, Part I: Development of the measuring normalizing block technique; Part II: Evaluation of the measuring normalizing block technique. **IEEE Trans. on Speech and Audio Processing**, New York, v.7, n.4, p.383–390, July 1999.

YOUSEFIZADEH, H. A Neural-Based Technique for Estimating Self-Similar Traffic Average Queueing Delay. **IEEE Communications Letters**, [S.l.], v. 6, p. 419–421, Oct. 2002.

YOUSEFIZADEH, H. et al. Utilizing Neural Networks to Reduce Packet Loss in Self-Similar Teletraffic Patterns. In: IEEE INTERNATIONAL CONFERENCE ON Communications, ICC, 2003, Anchorage-USA. **Proceedings...** New York: IEEE, 2003. v. 3, p. 1942–1946.

GLOSSÁRIO

Call detail record (CDR): em voz sobre IP é um arquivo que contém informações sobre o uso recente do sistema assim como a identidade das fontes (pontos de origem) e a identidade dos destino (pontos de finalização), a duração da chamada, a contabilidade de cada chamada, o tempo total de uso no período, o tempo livre no período. O formato do CDR varia de acordo com o provedor de VoIP ou o programa de contabilidade.

ISDN: sigla de Integrated Service Digital Network. Trata-se de um Sistema telefônico digital que, mediante o uso de equipamentos especiais, permite enviar e receber voz e dados simultaneamente através de uma linha telefônica.

Media Gateway : termina chamadas de voz provenientes de entroncamentos PSTN, comprime, empacota e entrega a Rede IP os dados de voz. Para chamadas de voz originadas na Rede IP, realiza as funções inversas.

Media Gateway Controller: trata os recursos de registro e gerência no *media gateway(s)*, troca mensagens de supervisão com as centrais da PSTN através do *signaling gateway*.

PSTN: sigla de *Public Switching Telephone Network*, correspondente a sigla em português RTPC, descrita abaixo.

RTPC: sigla de Rede Telefônica Pública Comutada. Refere-se ao sistema telefônico baseado em pares de fios de cobre por onde trafegam sinais analógicos de voz.

Signaling Gateway: fornece interfuncionamento transparente de sinalização entre redes de circuitos comutados e redes IP. Pode terminar a sinalização ou traduzir/transferir para o *media gateway controller* através da rede IP ou outro *signaling gateway*.

Softswitch: é composto principalmente pelo Controlador de *Gateway* de Mídia (*MGC-Media Gateway Controller*) e pelo *Gateway* de Sinalização (*SG-Signaling Gateway*), que fazem interface com a rede SS7 da RTPC e administram o estado da chamada. Isto permite que operadoras direcionem qualquer tipo de tráfego por sistemas IP e PSTN e definem ofertas em telefonia IP.

Trunking Gateway: está localizado na rede provedora de serviços e atua como um mediador entre a rede IP e a rede de telefonia convencional. Ele provê a transcodificação entre a voz baseada em pacotes e a voz transportada por circuitos.