

UNIVERSIDADE SÃO FRANCISCO
CURSO DE ENGENHARIA ELÉTRICA

QUALIDADE DE VOZ EM REDES IP

Área de Engenharia Elétrica

por

Rafael P. H. Moreno

José Sindi Yamamoto, Prof. Dr.
Orientador

Itatiba (SP), Outubro de 2004

UNIVERSIDADE SÃO FRANCISCO
CURSO DE ENGENHARIA ELÉTRICA

QUALIDADE DE VOZ EM REDES IP

Área de Engenharia Elétrica

por

Rafael P. H. Moreno

Relatório apresentado à Banca Examinadora do
Trabalho de Conclusão do Curso de Engenharia
Elétrica para análise e aprovação.
Orientador: José Sindi Yamamoto, Prof. Dr.

Itatiba (SP), Outubro de 2004

SUMÁRIO

LISTA DE ABREVIATURAS	iv
LISTA DE FIGURAS	v
LISTA DE TABELAS.....	vi
RESUMO.....	vii
ABSTRACT	viii
1. INTRODUÇÃO	9
1.1. OBJETIVOS.....	9
1.1.1. Objetivo Geral.....	9
1.1.2. Objetivos Específicos.....	9
1.2. METODOLOGIA.....	9
1.3. ORGANIZAÇÃO DO TRABALHO.....	10
2. FUNDAMENTAÇÃO TEÓRICA	10
2.1. OS DIFERENTES TIPOS DE TELEFONIA IP.....	10
2.1.1. Cenário 1: Conexão PC - PC.....	10
2.1.2. Cenário 2: Conexão Fone – Fone sobre IP	11
2.1.3. Cenário 3: PC – Fone ou Fone – PC.....	13
2.2. QUALIDADE DA VOZ.....	14
2.2.1. Problemas na rede telefônica convencional.....	16
2.2.2. Problemas em redes de pacotes	23
2.3. ACEITABILIDADE DE UMA CHAMADA TELEFÔNICA COM ECO E	
ATRASSO.....	31
2.3.1. Curva G.131.....	31
2.3.2. TELR.....	32
2.3.3. Interatividade	33
2.3.4. Modelo – E.....	33
2.3.5. Conseqüências para uma rede de telefonia IP.....	34
3. CONSIDERAÇÕES FINAIS	37
REFERÊNCIAS BIBLIOGRÁFICAS	38

LISTA DE ABREVIATURAS

IP	<i>Internet Protocol</i>
VoIP	<i>Voice over IP (Voz sobre IP)</i>
ISP	<i>Internet Service Provider</i>
ATM	<i>Asynchronous Transfer Mode</i>
QoS	<i>Quality of Service</i>
TELRL	<i>Talker Echo Loudness Rating</i>
ITSP	<i>Internet Telephony Service Provider</i>
ERL	<i>Echo Return Loss</i>
MSC	<i>Mobile Switching Center</i>
FIR	<i>Finite Impulse Response</i>
EEC	<i>Electric Echo Canceller</i>
AEC	<i>Acoustic Echo Canceller</i>
ERLE	<i>Echo Return Loss Enhancement</i>
RED	<i>Random Early Detection</i>
LAN	<i>Local Area Network</i>
WAN	<i>Wide Area Network</i>
SLR	<i>Send Loudness Rate</i>
RLR	<i>Receive Loudness Rate</i>
MOS	<i>Mean Opinion Score</i>
%GoB	<i>% Good or Better</i>
%PoW	<i>% Poor or Worse</i>
VAD	<i>Voice Activity Detection</i>

LISTA DE FIGURAS

Figura 1. Exemplo de uma comunicação de voz utilizando dois PCs.....	11
Figura 2. Exemplo de uma comunicação de voz sobre IP utilizando-se uma rede controlada	12
Figura 3. Utilização de dispositivos de adaptação para estabelecimento de conversação via Internet	13
Figura 4. Modelo de transmissão de voz em uma rede IP.....	15
Figura 5. Integração de rede IP com telefonia convencional e a utilização de uma híbrida para a interface do usuário	15
Figura 6. Híbrida	18
Figura 7. Pontos para medida do eco elétrico	19
Figura 8. Modelo funcional de um cancelador de eco	19
Figura 9. Forma de utilização de um cancelador de eco	20
Figura 10. Utilização de um cancelador de eco por um MSC.....	21
Figura 11. Utilização dos canceladores de eco acústico e elétrico e sentido da realimentação acústica.....	22
Figura 12. Atrasos em telefonia IP.....	24
Figura 13. Tempo para partida e chegada de um pacote composto por L e J (<i>jitter</i>).....	28
Figura 14. Acumulação do fluxo de dados	29
Figura 15. Curva G.131 que apresenta valores de TELR em função do atraso unidirecional.....	31
Figura 16. Circuito de referência para cálculo de TELR	32

LISTA DE TABELAS

Tabela 1. Classe de qualidade de acordo com o atraso de transmissão	24
---	----

RESUMO

MORENO, Rafael P. de H.. **Qualidade de voz em redes IP**. Itatiba, 2004. n° f. 39 Trabalho de Conclusão de Curso, Universidade São Francisco, Itatiba, 2004.

Este trabalho apresenta alguns conceitos básicos da nova tendência da tecnologia atual, VoIP (*Voice over IP*). Os principais problemas da telefonia convencional e de sua integração com uma rede de dados, ou seja, a telefonia IP também são abordados de forma a identificar algumas formas de mensurá-los e também solucioná-los.

Palavras-chave: Telefonia IP, VoIP, qualidade de voz, convergência.

ABSTRACT

This document presents some basic concepts of the new technology called VoIP. The reasons for the convergence tendency of conventional telephony networks to the larger data networks which exists nowadays, the Internet, are also discussed. This evolution has been taking place in a very quickly way, although there are many crucial issues to be studied and passed over specially on the voice quality which is introduced on the network.

The main problems as its solutions even on conventional or IP telephony will be held on this document.

Keywords: *IP Telephony. VoIP. Quality of Service (QoS), Convergence.*

1. INTRODUÇÃO

O suporte às comunicações de voz utilizando o protocolo Internet (IP), mais comumente chamado de VoIP, ou *Voice over IP* (Voz sobre IP), torna-se um atrativo ao usuário de telefonia pois uma ligação telefônica é realizada a um custo muito menor do que o método convencional. De fato, a busca por uma boa qualidade de telefonia em redes IP é um dos passos chave em direção à convergência das indústrias de voz, vídeo e comunicação de dados. VoIP pode ser definida como a habilidade de se fazer chamadas e enviar FAX em redes de dados baseadas em IP com uma qualidade de serviço (QoS) aceitável e um custo/benefício superiores.

1.1. OBJETIVOS

1.1.1. Objetivo Geral

Apresentar a tecnologia de voz sobre IP, seus principais problemas e formas de corrigi-los ou ameniza-los adequadamente, assim como maneiras de medir a qualidade da voz em uma conversação telefônica.

1.1.2. Objetivos Específicos

Os objetivos específicos deste trabalho são apresentados a seguir:

- Introduzir a tecnologia e topologia de uma rede de voz sobre IP apresentando seus componentes e suas funções;
- Identificar e discutir os principais problemas que comprometem a qualidade da voz tanto em uma rede convencional como em uma rede IP, como eco, *jitter*, influências de *codecs*, sistemas operacionais e agrupamentos de quadros como formas de gerar atrasos na rede, perda de pacotes e outros itens;
 - Apresentar métodos utilizados para limitação de ecos em interfaces analógicas;
 - Discutir a aceitabilidade de uma chamada telefônica com eco e atraso, apresentando conceitos da curva g.131, TELR e modelo – E.

1.2. METODOLOGIA

A metodologia utilizada para a realização deste trabalho foi a identificação de bibliografias adequadas e o estudo da tecnologia VoIP e os problemas decorrentes de sua utilização.

1.3. ORGANIZAÇÃO DO TRABALHO

Este trabalho é dividido em três seções as quais apresentam os diferentes tipos de telefonia IP e topologias possíveis atualmente (seção 2.1). Já a seção 2.2 trata da qualidade de voz, identificando os principais problemas acarretados pela utilização de uma rede de dados para trafegar voz e também problemas gerados a partir de interfaces acústicas, como é o caso do eco. Na seção 2.3, são descritas formas de medir a qualidade de voz em uma rede de dados, seja esta forma subjetiva (como é o caso do modelo-E) ou analítica, no caso da TELR ou Curva G.131.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. OS DIFERENTES TIPOS DE TELEFONIA IP

Dependendo da natureza da rede IP utilizada, existem das duas maiores categorias para transmissão de voz sobre redes IP. A primeira é essencialmente baseada na Internet, que é vista como a interconexão da grande maioria de computadores de redes públicas e privadas em uma escala global. O segundo é oferecido por operadores de serviços utilizando-se de redes IP controladas, nas quais um número pré-instalado de mecanismos sejam eles algoritmos de roteamento ou codificação, servem para garantir uma qualidade de serviço aceitável para uma certa conversação.

Existem atualmente três cenários de utilização da rede de voz sobre IP de acordo com os equipamentos utilizados e os tipos da rede:

2.1.1. Cenário 1: Conexão PC - PC

Neste caso, tanto o chamador quanto o usuário chamado possuem computador que permitem a eles conectar-se à Internet geralmente via rede de um provedor (ISP – *Internet Service Provider*). Ambos conseguem estabelecer uma comunicação de voz somente se houver um agendamento prévio, uma vez que os dois devem estar conectados à rede ao mesmo tempo. Os usuários devem possuir um software compatível com a tecnologia de voz sobre IP. Além disso, o usuário chamador deve conhecer o endereço IP do outro usuário.

Neste cenário, o ISP é geralmente acessado via rede telefônica pública por meio de uma simples chamada telefônica. Este meio de acesso ainda predomina, até mesmo em países

desenvolvidos. Soluções alternativas, conhecidas como ‘banda larga’ e baseadas em redes telefônicas ADSL, rede de TV a cabo ou acesso wireless (tecnologia LDMS), estão ainda em seus estágios iniciais de desenvolvimento, mas ainda não estão sendo amplamente utilizados até mesmo em países que já são bem equipados. [2]

O papel do ISP neste cenário é limitado a uma simples provisão de acesso à rede, que por sua vez possibilita ao usuário o acesso à Internet. A aplicação de voz utilizada por este cliente é transparente à ISP, que não toma medidas específicas para garantir a qualidade do serviço de voz. O protocolo utilizado neste tipo de comunicação é o H.323, definido pela ITU-T; Contudo o protocolo SIP (IETF) poderia ser mais amplamente utilizado no futuro. Esta solução é ilustrada na figura 1.

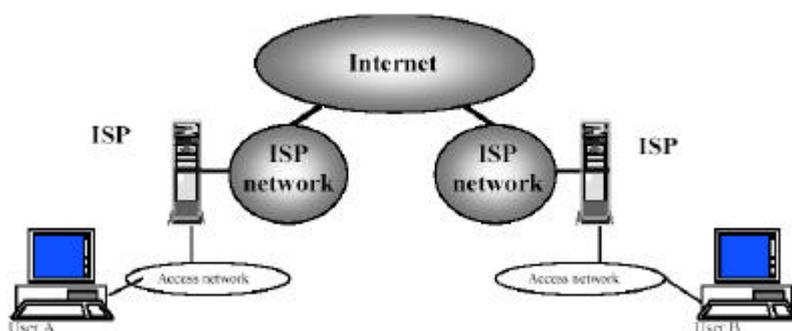


Figura 1. Exemplo de uma comunicação de voz utilizando dois PCs

2.1.2. Cenário 2: Conexão Fone – Fone sobre IP

Neste caso, os usuários chamado e chamador são ambos assinantes da rede de telefonia pública, seja ela fixa ou móvel, e utilizam seus aparelhos telefônicos para comunicação de voz de forma comum. Existem dois métodos para comunicação por meio de dois aparelhos comuns via rede IP ou a Internet.

2.1.2.1. Utilização de *gateways*

Isto significa que um ou mais fornecedores de serviço de telecomunicações estabeleceram gateways que possibilitam a transmissão de voz sobre uma rede IP de forma a ser transparente a usuários de telefone. O que temos neste caso não é a Internet mas uma rede IP controlada, isto é, uma rede que foi dimensionada de tal forma a permitir que a voz seja trafegada com uma qualidade de serviço aceitável. A figura 2 ilustra este cenário.

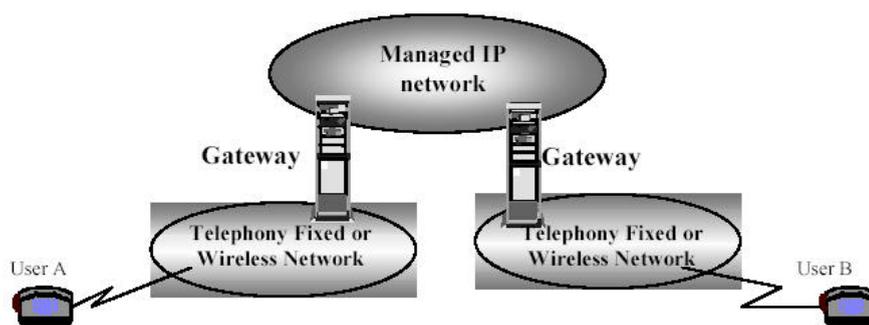


Figura 2. Exemplo de uma comunicação de voz sobre IP utilizando-se uma rede controlada

Neste cenário, os gateways e a rede IP controlada poderia pertencer a diferentes empresas fornecedoras de serviços, dependendo do ponto de vista:

- a) o uso puramente interno do VoIP na rede de uma única operadora de telefonia, que controla toda a operação, suportando ambos usuários A e B;
- b) a provisão de um serviço de voz de longa distância por uma operadora de longa distância utilizando tecnologia VoIP (usuários A e B neste caso pertencem a diferentes redes), nesta situação toda a operação pertence e é controlada pela operadora de longa distância.

2.1.2.2. Utilização de caixas de adaptação

Um número grande de companhias comercializa equipamentos que lembram *modems* e que são instalados entre o telefone do usuário e sua conexão à rede telefônica.

Para este arranjo operar de forma adequada, cada um dos dois usuários envolvidos na chamada devem ter uma assinatura com um provedor de acessos à Internet, cujos parâmetros de acesso foram pré-programados no dispositivo.

O usuário chamador inicia sua chamada da mesma forma que em uma rede convencional, e a primeira fase da chamada é de fato realizada na rede comum; contudo, imediatamente após isto, os dispositivos trocam informações requisitadas pela segunda fase. A chamada convencional é então terminada e os dispositivos, baseados nos dados trocados e através dos parâmetros pré-programados, estabelecem uma conexão entre cada um dos dois correspondentes e seus provedores. Uma vez que a chamada foi estabelecida, os dispositivos convertem localmente os sinais de voz em pacotes IP que serão transportados pela Internet como ilustrado na figura 3. Este cenário é, à princípio, muito similar ao cenário 1, exceto pelo fato que os usuários não necessitam de um PC e a necessidade de se deparar com a Internet é facilitada pelo procedimento ter sido iniciado na forma

de uma chamada comum. Entretanto, este tipo de arranjo não possui muito sucesso em seu uso pois necessita, como no caso da comunicação PC – PC, que os dois correspondentes estejam equipados com o mesmo tipo de equipamento.

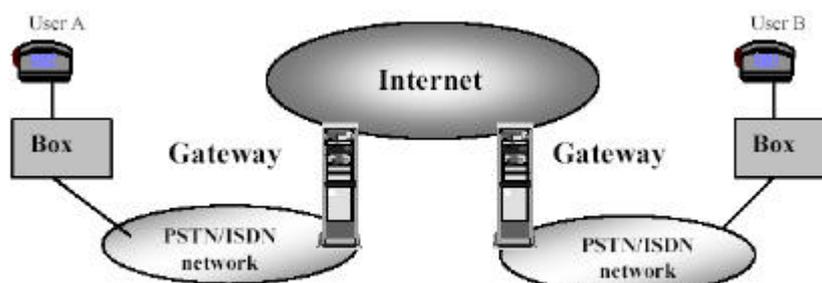


Figura 3. Utilização de dispositivos de adaptação para estabelecimento de conversação via Internet

Os dois métodos neste cenário utilizam dois tipos de rede para estabelecer uma chamada telefônica, ou seja, a Internet ou rede IP controlada (Intranet) e a rede pública.

2.1.3. Cenário 3: PC – Fone ou Fone – PC

Neste cenário, um dos usuários possui um computador pelo qual conecta-se à Internet via uma rede de acesso ou um provedor (similamente ao cenário 1), enquanto o outro usuário é um assinante comum de uma rede telefônica fixa ou móvel.

2.1.3.1. PC – Fone

Quando o usuário que possui o PC deseja chamar um usuário em seu aparelho telefônico, este deve iniciar conectando-se à Internet da forma tradicional via a rede de seu provedor. Uma vez conectado, ele utiliza os serviços de um provedor de serviços de telefonia IP (ITSP), que opera um *gateway* que garante acesso ao ponto que é mais próximo à rede do usuário chamado. É este *gateway* que gerencia toda a chamada e todas as sinalizações relacionadas à chamada telefônica na ponta do usuário chamado.

Deve-se notar que a ITSP oferece um serviço de PC a telefone em apenas uma direção, não permitindo gerenciar assinantes do serviço; de fato, o usuário de um PC utiliza os serviços da ITSP somente para chamadas de saída. Deve-se também notar que a ITSP possui uma rede IP controlada, garantindo assim uma certa qualidade de serviço para voz, tanto quanto o *gateway* mais perto do usuário chamado, e a ITSP também gerencia a interconexão com o operador do telefone. Apesar do

uso da tecnologia VoIP, ITSPs são consideradas provedores de serviços de telefonia, fornecendo seus serviços aos clientes de modo convencional, ou seja, com uma tarifa por minuto.

2.1.3.2. Fone – PC

Neste caso, o usuário que chama é um assinante de telefonia convencional e o usuário chamado possui um PC. Como um usuário da telefonia pode essencialmente discar um número E.164 para chamar uma outra pessoa, então de alguma forma o usuário do PC também possui um número E.164

- indiretamente: no caso de sua interconexão com a rede por trás de um *switch* de tecnologia IP.
- diretamente: neste caso, o endereço E.164 do lado do assinante IP possui um operador de telefonia.

2.2. QUALIDADE DA VOZ

Ao realizar uma chamada via IP para outra pessoa que não utiliza um *headset* (conjunto de fone de ouvido e microfone integrados), muitos problemas podem ocorrer.

Tal fato ocorre pela falta de maturidade existente em termos de qualidade de serviço (QoS) para aplicações deste tipo.

De acordo com [1], uma boa maneira de descobrir porque realmente existe um problema com a telefonia IP é realizar um simples teste quando na realização de uma chamada. Este método é conhecido por “contagem colaborativa”, que possui uma regra simples: ao ouvir a pessoa dizer ‘n’, diz-se imediatamente ‘n+1’. Para realizar uma comparação da telefonia clássica com a telefonia IP, primeiro uma pessoa faz uma chamada comum para alguém que conhece e diz ‘1’, a pessoa do outro lado da linha diz ‘2’ etc. Marca-se então o tempo gasto para contarem até 25, Em seguida, faça uma chamada telefônica IP e repita o mesmo jogo. Em todas as tentativas, vai demorar muito mais tempo, caso consiga atingir 25 alguma vez.

Os problemas descritos acima – eco e atraso – são muito conhecidos por planejadores de redes telefônicas desde o início da telefonia, e as redes telefônicas atuais foram projetadas para manter esses efeitos imperceptíveis para a maioria dos clientes. Ao transportar voz sobre IP, torna-se muito difícil controlar o eco e o atraso. Isso exigirá tecnologia de ponta e ajuste dos componentes para tornar o serviço aceitável para todos os clientes.

Podemos modelar uma conexão para transmissão de voz de algumas formas. A primeira forma abaixo mostra um modelo que não apresenta terminais PSTN ou ISDN envolvidos.

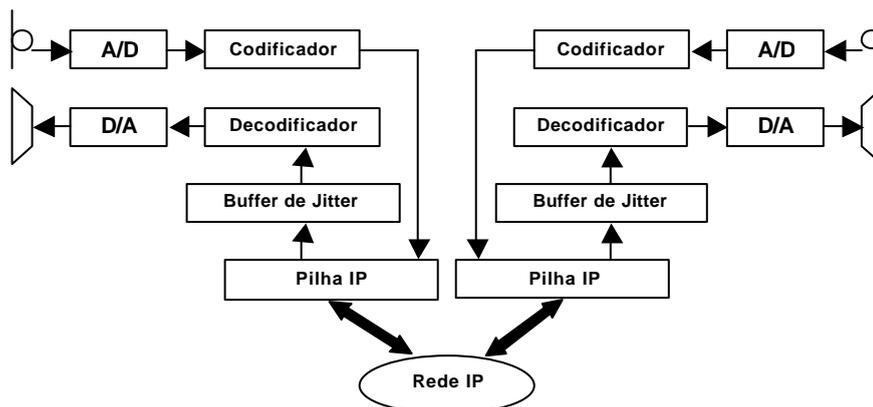


Figura 4. Modelo de transmissão de voz em uma rede IP

Em uma rede de telefonia analógica, a interface do usuário se dá, na maioria dos casos, usando-se apenas dois fios (sinais que entram e saem compartilhando o mesmo par), tornando necessária a utilização de um híbrido de quatro fios / dois fios. A figura 5 ilustra a integração de uma rede IP à rede de telefonia convencional e a utilização do híbrido.

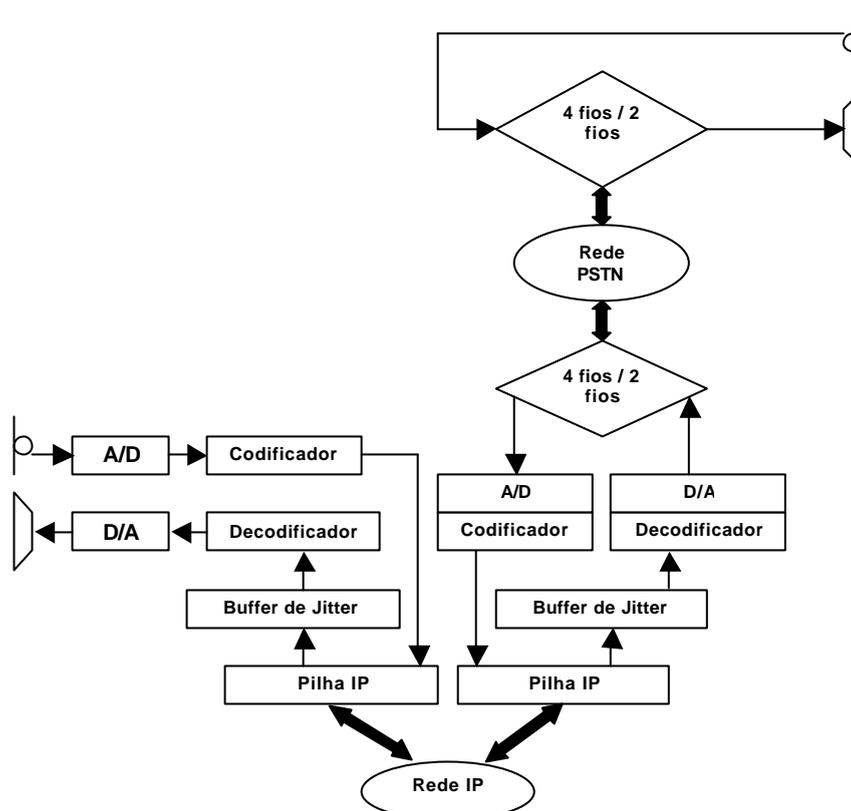


Figura 5. Integração de rede IP com telefonia convencional e a utilização de uma híbrida para a interface do usuário

Este último modelo inclui as fontes mais significativas de degradação da qualidade de voz:

- A rede IP introduz perda de pacotes, atrasos e *jitter*;
- Os *buffers* de *jitter* (JB) influenciam o atraso fim a fim e a perda de quadros;
- As interfaces acústicas introduzem eco acústico;
- As interfaces introduzem eco nas híbridas.

Os fatores acima são os que mais possuem influência sobre a percepção da qualidade de voz pelo usuário final. A maioria destes fatores são comuns tanto à telefonia convencional (comutação de circuitos) quanto à telefonia IP. No entanto, a telefonia IP possui algumas características únicas, como atrasos longos, *jitter* e perda de pacotes, portanto requer uma nova estrutura para avaliar a qualidade de voz.

2.2.1. Problemas na rede telefônica convencional

2.2.1.1. Eco

Eco pode ser descrito como o retorno de um sinal no lado oposto à transmissão. Generalizando, este problema aumenta no contexto das comunicações PC - fone, Fone - PC ou Fone - fone. É causado por componentes elétricos de sessões analógicas do sistema retornando uma parte do sinal processado.

Conforme descrito em [1], um eco menor que 50 ms não é crítico em uma conversação telefônica. Acima deste nível, o locutor ouvirá sua própria voz logo após falar. Onde o objetivo é prover serviços de telefonia IP, os *gateways* devem processar o eco elétrico gerado a partir da híbrida (transformação de dois para quatro fios). Se tal processamento não for levado em consideração, não é possível utilizar o serviço com aparelhos analógicos convencionais. Para resolver o problema, canceladores de eco são instalados no estágio do *gateway* da rede. A seguir são apresentados os tipos de ecos decorrentes da rede convencional de telefonia e alguns mecanismos de cancelamento e supressão.

2.2.1.1.1. Eco acústico

Pode-se chamar de ‘telefone amplificado’ ao telefone como amplificador e alto-falante sem cancelamento de eco acústico, e de ‘telefone viva-voz’ um telefone amplificado com cancelamento de eco acústico.

De acordo com [1], o eco acústico é simplesmente a parte do sinal acústico realimentada do alto-falante de um dispositivo para o microfone desse mesmo dispositivo. Geralmente o eco acústico é um sinal parasita de 10 a 15 dB (no caso de um 'telefone amplificado') abaixo do sinal acústico da pessoa que realmente fala ao telefone. Da mesma maneira que o eco híbrido, esse nível de eco acústico não será notado se o atraso de ida e volta estiver abaixo de 20ms. Acima de 40ms a pessoa do outro lado da linha tem a impressão de estar conversando em uma cisterna e as coisas pioram acima dos 40ms de atraso de ida e volta.

Uma maneira fácil de suprimir o eco acústico é usar um *headset*, mas com dispositivos apropriados é possível reduzir a potência do eco a cerca de 45 dB abaixo do sinal do locutor [1], mesmo no caso de um telefone viva-voz.

O desempenho de canceladores de eco compatíveis com a Recomendação G.165 do ITU podem não ser suficientes. Uma recomendação mais nova, a G.168, está disponível e já implementada por algumas empresas. Essa recomendação contém acréscimos importantes, tais como a capacidade de parar o cancelamento de eco ao detectar o tom de inversão de fase dos *modems* de alta velocidade.

Valores típicos para a atenuação de eco acústico nos dispositivos atuais são:

- telefones amplificados (80 % do mercado na França): 10 a 15 dB;
- telefones viva-voz: 35 a 40 dB;
- telefones com *handset* de boa qualidade: 35 a 40 dB.

Eco acústico gerado no terminal do usuário

O eco mais importante é o eco de locutor - ele é a percepção que o locutor tem de sua própria voz, porém atrasada. Esse eco pode ser causado por um eco elétrico (híbrido) ou eco acústico ocorrido no lado do ouvinte. Se o eco do locutor for refletido duas vezes, ele também poderá afetar o ouvinte. Nesse caso, o ouvinte escuta a voz do locutor duas vezes – um sinal alto primeiro e depois um sinal atenuado e muito atrasado.

Eco acústico gerado na híbrida

O eco híbrido, ou elétrico, é a fonte primária de eco gerado pela rede pública de telefonia (PSTN). Este eco resulta do desbalanceamento da híbrida criado à medida que sinais de voz são transmitidos pela rede através da conexão híbrida nos pontos de conversão dois fios/quatro fios, refletindo energia elétrica pelo circuito de quatro fios para o locutor.

O caminho do sinal entre dois telefones, envolvendo outro tipo de chamada que não seja local, requer amplificação utilizando um circuito de quatro fios. Embora não seja um fator importante na rede de telefonia celular, o eco híbrido torna-se um problema em chamadas originadas pela rede convencional. O custo e cabeamento necessários acabam afastando a possibilidade da adoção de um circuito quatro fios a partir da casa do assinante até a central telefônica, ou seja, ponto-a-ponto. Por esta razão, uma solução alternativa teve que ser encontrada e os circuitos dos entroncamentos de quatro fios foram convertidos para dois fios (cabeamento local), utilizando-se de um dispositivo chamado ‘híbrido’. A figura 6 apresenta este dispositivo.

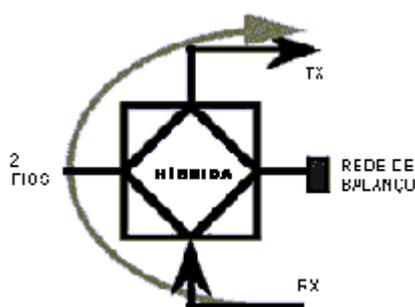


Figura 6. Híbrida

À medida que sinais de voz passam da porção da rede quatro fios para a de dois fios, a parte da energia na seção de quatro fios é refletida de volta, criando o eco da voz. Como a híbrida é dimensionada de tal forma que o atraso ponto-a-ponto ocorra apenas em alguns milissegundos (aproximadamente 28ms), ele gera uma sensação de que a chamada está ‘viva’, adicionando sinais complementares (*sidetones*), que contribuem positivamente para a melhoria da qualidade da chamada. [2]

Ainda segundo [2], em casos onde o atraso total da rede excede 36 ms, os benefícios desaparecem, resultando em ecos intrusivos. O volume real de sinal que é refletido de volta depende em quão balanceado está o circuito em relação à linha de dois fios. Na grande maioria dos casos, o casamento destas impedâncias é ruim, resultando em um nível considerável de sinal refletido. Isto é medido em perda de retorno por eco (*ERL – Echo Return Loss*), sendo que quanto maior o ERL, menor o sinal refletido de volta ao locutor, e vice-versa.

2.2.1.1.2. Como limitar o eco

Dois tipos de dispositivos são comumente usados para limitar o eco, são eles: os canceladores de eco e os supressores de eco.

O eco elétrico e a redução de eco são medidos no circuito de quatro fios com os pontos de referência mostrados na figura abaixo:

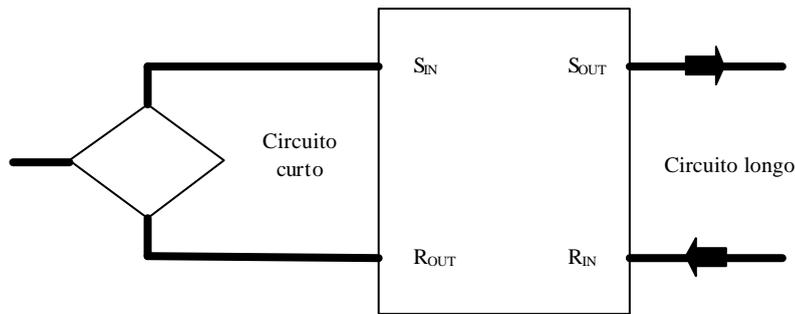


Figura 7. Pontos para medida do eco elétrico

Supressores de Eco

Os supressores de eco foram desenvolvidos na década de 1970. A idéia consiste em introduzir uma grande perda no caminho de envio quando a parte distante está falando. Essa técnica é amplamente utilizada nos telefones viva-voz mais simples, mas tende a atropelar a fala do locutor quando a parte distante fala ao mesmo tempo.

Canceladores de Eco

O modelo funcional de um cancelador de eco é mostrado abaixo:

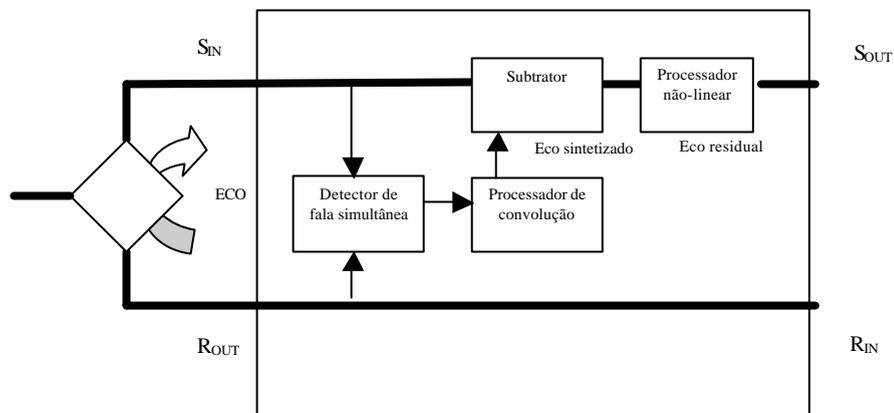


Figura 8. Modelo funcional de um cancelador de eco

Um cancelador de eco é muito mais complexo que um supressor de eco, pois ele realmente constrói uma estimativa de eco para removê-la do sinal que chega. O eco é modelado como uma soma de sinais, semelhantes ao sinal de entrada, porém atrasados e com menor amplitude (uma convolução do sinal de entrada); portanto, funciona apenas com modificações lineares do sinal entre Rin e Sin. A presença de saturação, por exemplo, comprometerá seriamente o desempenho de um cancelador de eco. O sinal de erro é medido e minimizado quando apenas a parte distante estiver falando, que é a razão pela qual o detetor de fala simultânea é usado.

O cancelador de eco precisa armazenar as amplitudes de todos os sinais atrasados que retornam para cada atraso possível entre zero e o maior atraso no circuito curto (esta é a resposta ao impulso do híbrido). Assim, os canceladores de eco capazes de lidar com atrasos elevados no circuito curto são mais caros que aqueles que lidam apenas com pequenos atrasos-, por isso é sempre melhor colocar o cancelador de eco mais perto o possível da fonte de eco.

Basicamente, os canceladores de eco elétrico são filtros digitais adaptativos (FIR – *Finite Impulse Response*: resposta finita ao impulso) colocados na rede, por exemplo, em um centro de comutação internacional para um link de satélite ou no centro de comutação móvel (MSC - *Mobile Switching Center*) para aplicações de celulares digitais. A próxima ilustração mostra a utilização de um cancelador de eco elétrico:

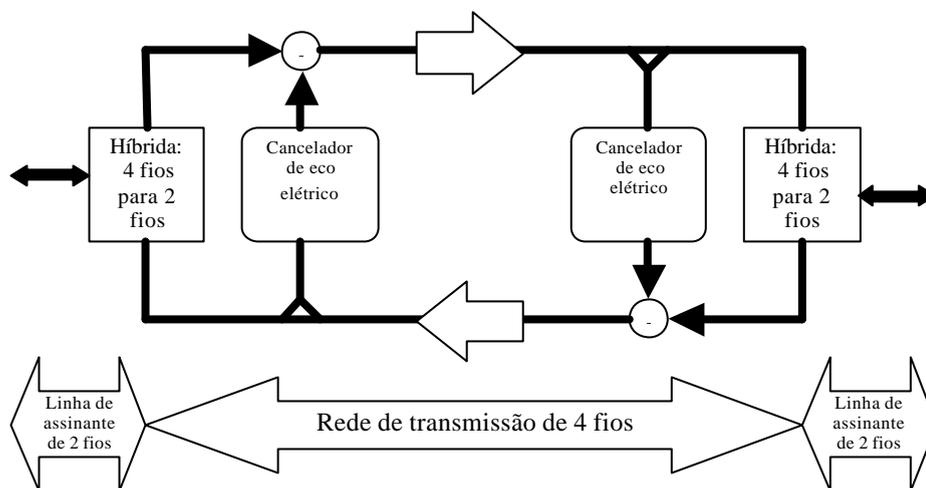


Figura 9. Forma de utilização de um cancelador de eco

Conforme descrito em [2], no sistema telefônico celular GSM, o atraso de ida é de aproximadamente 100ms devido a:

- o comprimento do quadro de 20 ms;

- o atraso de processamento de 20ms (pode ser reduzido usando-se um DSP poderoso);
- entrelaçamento para proteção de canais;
- bufferização e decodificação.

Desse modo, o cancelador de eco elétrico (EEC - *Electric Echo Canceller*) deve ser incluído no MSC, como mostrado a seguir:

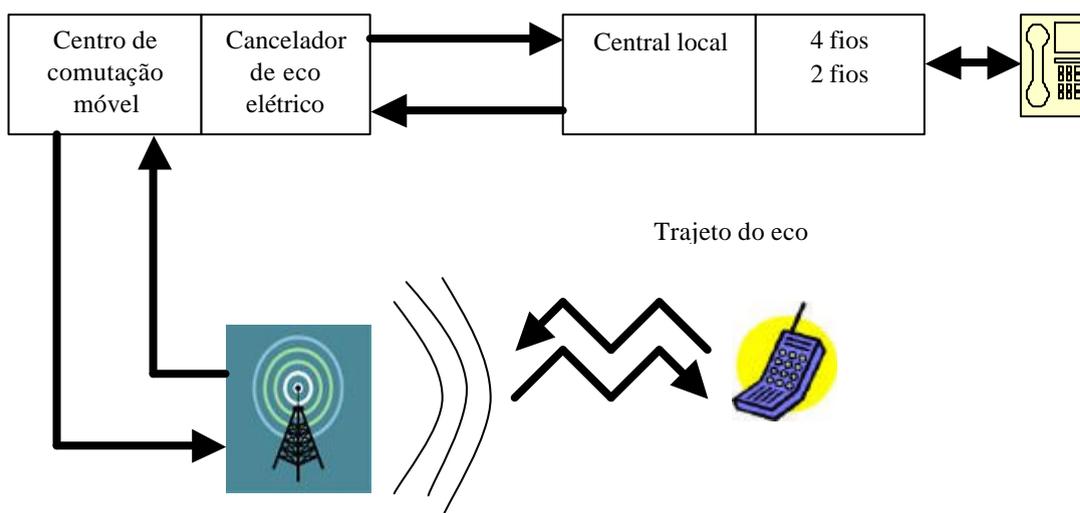


Figura 10. Utilização de um cancelador de eco por um MSC

No caso de voz em tempo real em uma rede de comutação de pacotes como VoIP, a situação de eco elétrico é bastante semelhante, e ele deve ser cancelado caso o *gateway* IP venha a estabelecer uma chamada entre um terminal PSTN de dois fios e um PC ou estação de trabalho multimídia. Uma das várias funções do *gateway* de telefonia IP é o cancelamento do eco elétrico gerado pela transformação de quatro fios para dois fios. Se o usuário de PC possuir um *headset* (que contém um microfone próximo à boca do locutor), não haverá eco acústico, portanto, não existirá a necessidade de um cancelador de eco acústico (AEC - *Acoustic Echo Canceller*) na placa de som de um PC. Como mostrado abaixo, é necessário ter um EEC no gateway e um AEC no ambiente do PC se o usuário do PC tiver um microfone e alto-falantes em uma situação de viva-voz. Se não houver nenhum EEC, o eco elétrico perturbará o lado do PC e, caso não exista nenhum AEC, o eco acústico perturbará o lado da PSTN.

Em uma rede IP o atraso geralmente é alto e imprevisível. Para uma rede comutada, como uma PSTN ou ISDN, este geralmente não é o caso; mas em países muito grandes, como os Estados Unidos (o caminho da costa leste à costa oeste também é longo para uma chamada telefônica), o atraso de transmissão é tão alto que um EEC faz-se necessário até mesmo para as chamadas terrestres (usando links de fibra ótica).

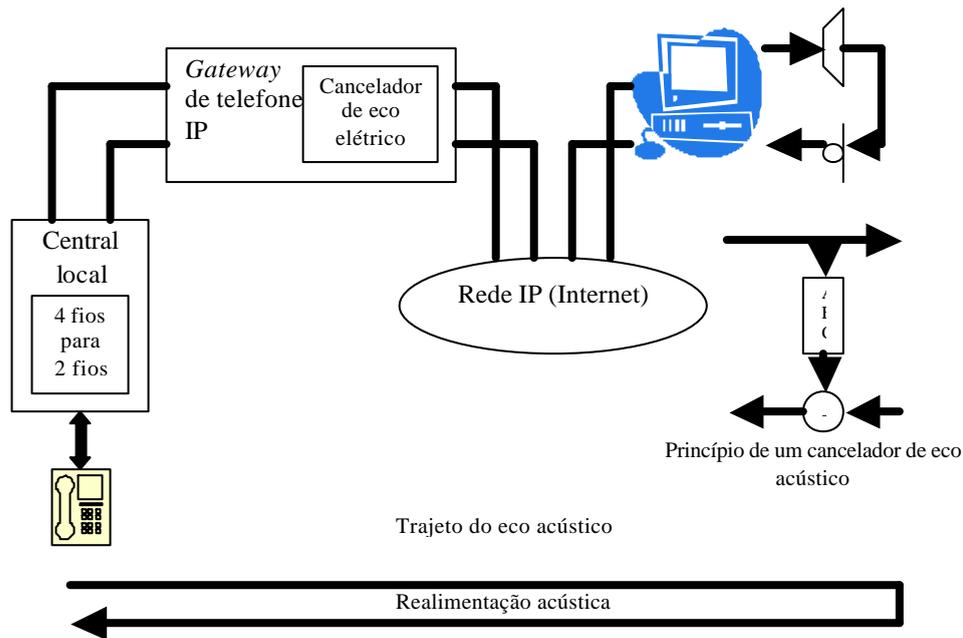


Figura 11. Utilização dos canceladores de eco acústico e elétrico e sentido da realimentação acústica

Mesmo quando se usa esquemas de codificação de baixo atraso para sistemas de telefonia celular digital, VoIP, voz em *frame relay* ou mesmo voz em ATM, a utilização de EECs é realmente obrigatória e isso causa impacto sobre o custo do sistema como um todo. Contudo, uma codificação de voz com um baixo atraso algorítmico e de processamento relaxará os requisitos para o desempenho do EEC (rejeição de eco). Para enfrentar o eco acústico, um AEC deve ser instalado no terminal viva-voz seja qual for a rede e o link de transmissão usados.

O AEC funciona de maneira semelhante ao EEC. Trata-se de um filtro digital adaptativo associado a alguns procedimentos de variação de ganho. Quanto maior o atraso total de um percurso ida e volta (*round trip*), mais difícil será sintonizar todos os parâmetros dos sistemas AEC (velocidade de convergência, faixa de variação de ganho etc.), de maneira a reduzir suficientemente o nível de eco e manter uma boa interatividade. O mais desejável em qualquer chamada telefônica ou vídeo conferência é um atraso baixo.

O desempenho de um cancelador de eco envolve muitos parâmetros. Os mais importantes são a melhoria da perda por retorno de eco (ERLE – *Echo Return Loss Enhancement* em dB) que é dada pela redução do nível de eco entre as portas Sin e Sout, e o tamanho da janela que modela a resposta ao impulso (alguns canceladores de eco são otimizados para cancelar todos os ecos que chegam com um atraso de 0 a T_{max} , outros são otimizados para modelar apenas ecos que chegam com um atraso de T_{min} a T_{max}). Outros parâmetros são o tempo de convergência e a qualidade da detecção de fala simultânea.

2.2.2. Problemas em redes de pacotes

2.2.2.1. Perda de pacotes

A perda de um pacote resulta em falta de informação quando o sinal de áudio é recebido. De acordo com o número de pacotes perdidos, a qualidade do som no ponto de recepção pode se alterar. Na abordagem do IP, perda de pacote forma a parte integral de um conceito: os roteadores cumprem sua função (com o algoritmo de detecção antecipada aleatória – RED – *Random Early Detection*) ao destruir pacotes com o objetivo de evitar possíveis congestionamentos.

Existem quatro principais causas de perdas de pacotes:

- ?? Tempo de vida do pacote esgotado (TTL = 0);
- ?? O atraso de recepção maior que o *buffer* de *jitter*;
- ?? Destruição por um módulo congestionado;
- ?? Invalidade do pacote por problemas na transmissão.

O protocolo UDP, assim como o IP e o RTP, são usados para transmitir voz sobre IP pelas vantagens de usar menos cabeçalho (overhead) e dependência sobre protocolos de camadas superiores (como RTCP/RTP) para fornecer controles de erro e fluxo ou onde aplicações de tempo real necessitam realizar retransmissões, como utilizado pelo protocolo TCP inapropriadamente.

A razão da perda de pacote dependerá da qualidade das linhas utilizadas, assim como no dimensionamento da rede. Se a qualidade da conversação deve ser aceitável, a taxa de perda de pacotes deve ser inferior a 20%, segundo [2].

Uma possível solução para redução de perdas de pacotes é implementar sistemas de correção de erros utilizando redundâncias e codificação adaptativa, ou seja, que varia de acordo com as perdas de pacotes observadas estatisticamente dentro da rede em um dado ponto no tempo. Usando-se destes sistemas, é possível atingir níveis muito altos de qualidade da voz, até mesmo sobre a Internet. Esta solução, contudo, eleva uma dificuldade adicional associada com o atraso total de transmissão, que deve ser dominada se a rede for utilizada para telefonia.

2.2.2.2. Atrasos em uma rede telefônica VoIP

O atraso é o tempo que se passa entre a transmissão da voz e sua reconstituição no lado receptor. Se for uma troca interativa, como em uma conversação normal por telefone, certas restrições de atraso devem ser aplicadas à transmissão da conversação. A tabela a seguir indica as

classes de qualidade e interatividade de acordo com o atraso de transmissão em uma conversação telefônica.

Tabela 1. Classe de qualidade de acordo com o atraso de transmissão

Classe	Atraso por direção	Comentários
1	0 - 150 ms	Aceitável para a maioria das conversações; apenas algumas tarefas altamente interativas podem sofrer degradações.
2	150 - 300 ms	Aceitável para chamadas de baixa interatividade.
3	300 - 700 ms	Praticamente uma chamada <i>half-duplex</i>
4	Acima de 700 ms	Impraticável ao menos que os locutores são acostumados com a arte da conversação <i>half-duplex</i> (como utilizado por militares).

Fonte: Adaptado de ITU/BDT (2002)

O atraso introduzido pela Internet [1] (de 50 ms até mais de 500 ms, de acordo com o estado da rede) é de uma ordem muito maior que o atraso encontrado em uma rede de telefonia convencional. Quantificar o atraso de transmissão sobre a rede com confiança nos dados obtidos é virtualmente impossível, vista ao grande número de fatores desconhecidos (tabelas de roteamento, congestionamento, falhas, enfileiramentos, etc.). Contudo, para o roteador que foi acessado por uma transmissão de voz, é possível detalhar certos tipos de atraso que são inerentes à rede, como ilustrado na figura 12, que mostra uma transmissão PC - fone via Internet. No caso abaixo, a rede IP do provedor do serviço de *gateways* está sendo considerada como ideal e como não contribuindo significativamente para o atraso de transmissão geral.

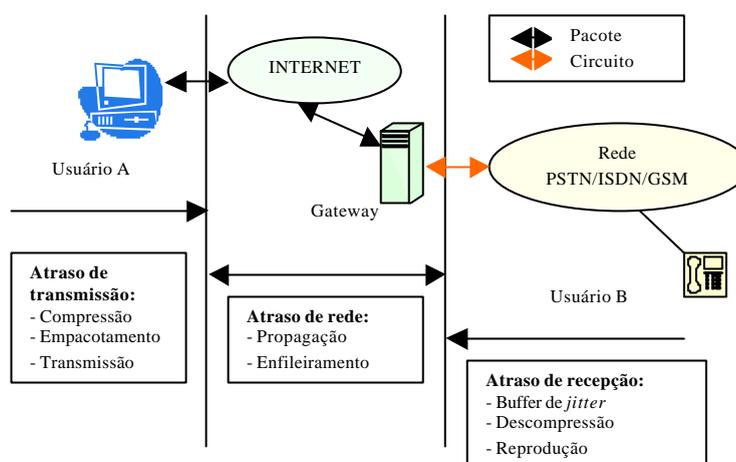


Figura 12. Atrasos em telefonia IP

2.2.2.2.1. Atraso de Transmissão

No ponto de transmissão, a fala é codificada e comprimida de forma a ser encapsulada em pacotes IP. O tamanho do pacote representa um compromisso entre a necessidade de redução do atraso de transmissão e a otimização da banda. Os componentes do atraso de transmissão são:

?? Digitalização e codificação, o tempo utilizado pelo gateway ou pela placa de som para digitalizar e codificar um sinal analógico.

?? Compressão, que podemos desmembrá-la em três partes:

- Atraso do quadro: diferente da digitalização do sinal, que é realizada de forma contínua, a compressão está relacionada com um comprimento específico do dado. A espera por esta informação pode acarretar em um tempo de processamento razoável.

- Atraso de codificação: este atraso, que envolve compressão por síntese baseada em predição, é requerido pelo codificador para conhecer, enquanto ativo, a evolução do sinal.

- Atraso de processamento: tempo levado pelo algoritmo para comprimir um quadro. Depende do processador e do algoritmo utilizado.

?? Formação dos pacotes: período de tempo durante o qual a aplicação agrupa os pacotes (criação do cabeçalho e inserção dos dados).

?? Transmissão: este período de tempo dependerá da configuração utilizada, ou seja, se a conexão é realizada via *modem* ou por acesso direto em uma LAN ou WAN.

Existem três categorias maiores para transmissão de voz sobre IP, de acordo com a técnica de codificação utilizada:

- codificação do tempo (com velocidades entre 16 e 64 kbit/s);

- codificação paramétrica (com velocidades entre 2.4 e 4.8 kbit/s);

- codificação de análise-síntese (com velocidades entre 5 e 16 kbit/s).

Generalizando, as técnicas de codificação que fornecem baixas taxas necessitam de tempos de processamento maiores, conseqüentemente aumentando o tempo de trânsito. Atualmente é conhecido que o tempo médio para processamento da fala (compressão, descompressão e formação de pacotes) introduz um atraso de aproximadamente 50 ms para cada ponta do circuito de voz. [2]

Influência do Sistema Operacional

A maioria das aplicações de telefones IP são simplesmente programas executados sobre um sistema operacional com o MS-Windows. Elas têm acesso aos periféricos de som por meio de uma API (ex.: API WAVE para WINDOWS) e têm acesso à rede por meio da API soquete. À medida que se fala, a placa de som amostra os sinais do microfone e armazena o sinal amostrado em um *buffer* de memória (que também pode executar alguma compressão básica como a codificação G.711). Quando o *buffer* está cheio, a placa de som comunica ao sistema operacional, usando um pedido de interrupção (*interrupt*), que ele pode recuperar os dados do *buffer*, e as amostras seguintes são armazenadas em um novo *buffer*.

Interrupções param as atividades regulares do sistema operacional e disparam um programa muito pequeno chamado gerenciador de interrupções que, nesse caso, simplesmente armazena, para o programa que abriu o microfone, um ponteiro para o *buffer* de som.

No caso da API WAVE, o programa registra uma função *call-back* quando abre o microfone para receber os novos *buffers* de amostras, enquanto o sistema operacional simplesmente chama essa função para passar o *buffer* à aplicação de telefone IP. Quando a função *call-back* é chamada, ela verifica se há amostras suficientes para formar um quadro completo para o algoritmo de compressão e, em caso afirmativo, coloca o quadro comprimido resultante (empacotado com as informações RTP apropriadas) na rede usando a API soquete.

O fato de as amostras vindas do microfone serem enviadas ao sistema operacional em blocos usando uma interrupção introduz um atraso de acumulação mínimo, uma vez que a maioria dos sistemas operacionais não consegue acomodar muitos pedidos de interrupções por segundo.

A conclusão é que o sistema operacional é um parâmetro muito importante, o que deve ser levado em consideração ao tentar reduzir atrasos fim-a-fim para aplicações de telefonia IP. Para contornar essa limitação, alguns fornecedores de *gateways* de telefonia IP e de telefones IP usam sistemas operacionais de tempo real como o VxWorks (da Wind River Systems) ou o pSOS (da ISI), que são otimizados para tratar quantas interrupções forem necessárias para reduzir esse atraso de acumulação.

Uma outra maneira de contornar as limitações do sistema operacional é executar todas as funções de tempo real (aquisição de amostras, compressão e RTP) usando hardware dedicado, executando apenas as funções de controle a partir do sistema operacional, que não é de tempo real. Alguns fornecedores de placas de telefonia IP, como a Natural Microsystems, a Dialogic e a

AudioCodes usam esse tipo de abordagem para permitir que terceiros construam *gateways* de baixa latência com o Unix ou MS-Windows sobre seu equipamento.

2.2.2.2.2. Atraso de Rede

?? Propagação: em uma rede convencional (fios), a velocidade de propagação é 200.000 km/s [2], permitindo um tempo de propagação apreciável.

?? Roteamento e filas: de acordo com a natureza da rede, diferentes tempos podem ser vinculados.

No caso de uma rede IP bem controlada como uma Intranet ou equivalente, a transmissão de pacotes leva em torno de 50 a 100 ms (considerando a propagação e a compensação do *jitter*), com a introdução de atrasos pelos roteadores em torno de 50 ms. O atraso total resultante, portanto, situa-se entre 200 e 250 ms fim-a-fim para uma rede IP bem controlada (Intranet). Estes atrasos são significativamente maiores e até mesmo indeterminados no caso da Internet.

2.2.2.2.3. Jitter

Buffer de jitter: esta memória permite a resincronização de pacotes que chegam com atrasos variados. Este serve, portanto, para compensar defasagens de tempo e para realocar pacotes em suas ordens corretas.

?? Desmembramento dos pacotes;

?? Descompressão;

?? Decodificação e conversão digital – analógico.

O resultado disto, sob condições normais em termos de tecnologia utilizada pela Internet e seu dimensionamento, é que a telefonia IP seria viável apenas em uma rede IP controlada semelhante à Intranet, sendo seu comportamento muito imprevisível sobre a Internet.

Influência da política de buffer de jitter sobre o atraso

Um pacote IP precisa de algum tempo para ir de A até B através de uma rede de pacotes (Figura 13). Esse atraso $t_{AB} = t_{chegada} - t_{partida}$ é composto de uma parte fixa L característica dos atrasos de propagação e dos atrasos médios de enfileiramento (*average queuing delays*), e uma parte variável caracterizando o *jitter* causado pelo comprimento variável das filas nos roteadores, dentre outros fatores.

Os terminais usam um *buffer* de *jitter* para compensar os efeitos de *jitter*. Esse *buffer* segura os pacotes na memória até que $t_{\text{unbuffer}} - t_{\text{partida}} = L+J$. O instante de partida de cada pacote é determinado com o uso da informação de *timestamp* fornecida pelo RTP. Ao aumentar o valor de J, o terminal é capaz de re-sincronizar mais pacotes. Pacotes que chegam muito tarde ($t_{\text{chegada}} > t_{\text{unbuffer}}$) são descartados.

Os terminais usam heurísticas para ajustar o melhor valor de J: se J for muito pequeno, muitos pacotes serão descartados; se J for muito grande, o atraso adicional será inaceitável para o usuário. Para alguns terminais, a configuração do tamanho do *buffer* de *jitter* é estática; essa configuração não é ótima quando as condições da rede não são estáveis. Outros terminais podem redimensionar dinamicamente o tamanho de seus *buffers* de *jitter* usando heurísticas. Essas heurísticas podem tomar algum tempo, uma vez que o terminal precisa avaliar o *jitter* na rede, por exemplo, o terminal pode decidir começar inicialmente com um *buffer* muito pequeno e, progressivamente, aumentá-lo até que a porcentagem média de pacotes que chegam atrasados caia abaixo de 1%.

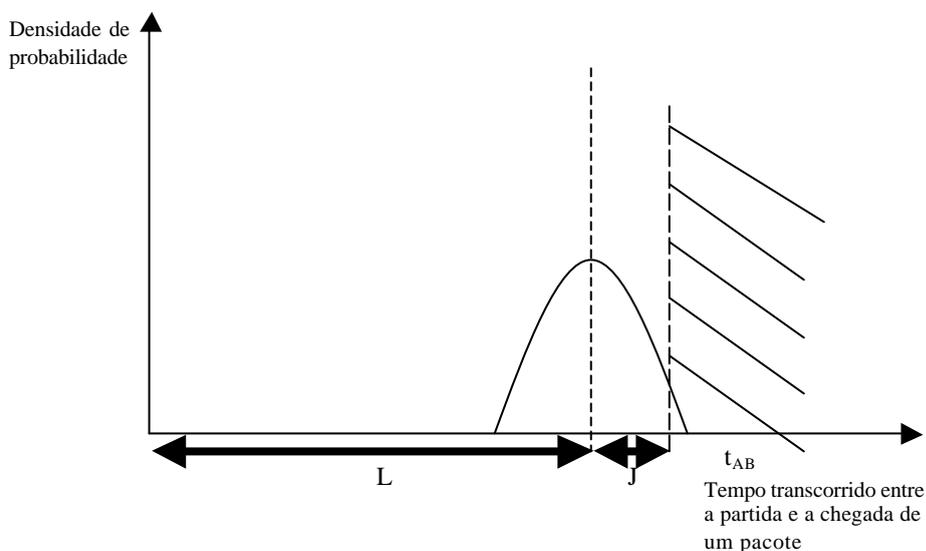


Figura 13. Tempo para partida e chegada de um pacote composto por L e J (*jitter*)

Influência do Codec, do agrupamento de quadros e da redundância

A maioria dos codificadores de voz se baseiam em quadros. Isso significa que eles comprimem blocos contendo um número fixo de amostras quantizadas linearmente, em vez de comprimirem amostra por amostra. Dessa maneira, o fluxo de dados de áudio precisa ser acumulado

até que ele atinja o tamanho do bloco antes de ser processado pelo codificador (Figura 14). Essa acumulação de amostras leva tempo, portanto, soma-se ao atraso de fim-a-fim. Além disso, alguns codificadores precisam conhecer mais amostras do que aquelas contidas no quadro que eles vão codificar (*lookahead*).

Uma rápida conclusão é que, para reduzir atrasos em uma rede ideal, o codec escolhido deve ter um comprimento de quadro pequeno.

Infelizmente, a regra geral é que, quanto mais se sabe sobre alguma coisa, mais fácil será modelá-la de maneira eficiente. Portanto, os codificadores com tamanhos de quadro maiores tendem a ser mais eficientes e a ter maiores taxas de compressão. Um outro fator é que o quadro não é transmitido através da rede sem que uma quantidade considerável de *overhead* seja acrescentada pelos próprios protocolos de transporte a cada pacote transmitido. Se cada quadro de voz comprimida for transmitido em um pacote próprio, esse *overhead* será acrescentado a cada um, e para alguns codificadores o *overhead* terá tamanho comparável, se não maior, aos dados úteis. Para diminuir o *overhead* a um nível aceitável, a maioria das implementações transmite vários quadros em cada pacote.

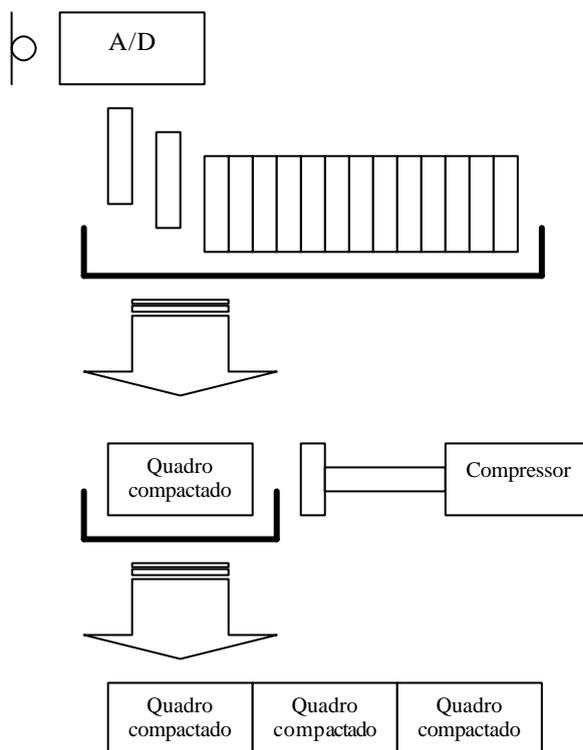


Figura 14. Acumulação do fluxo de dados

Se todos os quadros acumulados no pacote pertencerem ao mesmo fluxo de áudio, isso acrescentará mais atraso de acumulação. Na verdade, usar um codificador com um tamanho de quadro igual a f e três quadros por pacote é absolutamente equivalente, em termos de *overhead* e atraso de acumulação, a usar um codificador com um tamanho de quadro maior geralmente é mais eficiente, é provável que essa última solução também seja mais eficiente.

Observe se o sistema operacional dá acesso ao fluxo de áudio em blocos de tamanho C ms em vez de amostra por amostra, se for esse o caso, as amostras já terão sido acumuladas e, então, será melhor usar um codificador com um tamanho de quadro $F=C$ ou, caso C seja muito grande, utilizar *floor* (C/F) quadros por pacote (isto é, o maior número inteiro possível de quadros de duração F por bloco de duração F).

Uma maneira muito mais inteligente de empilhar vários quadros por pacote, sem causar nenhum impacto sobre o atraso, é colocar quadros de fluxos de áudio diferentes, mas com a mesma rede de destino. Este geralmente é o caso de troncos VoIP entre sites corporativos, ou entre *gateways* dentro de uma rede VoIP. Infelizmente, a maneira de se fazer essa multiplexação RTP (ou RTP-mux) ainda não foi padronizada no H.323, no SIP ou em outros protocolos VoIP, e os fabricantes precisam usar implementações proprietárias (como usar um canal lógico com um ‘codificador’ que de fato codifica muitos canais e permanece aberto permanentemente entre os *gateways*, usando, para cada nova comunicação, um codificador *dummy* ao abrir o canal lógico).

Um outro parâmetro que deve ser levado em consideração ao calcular o atraso fim-a-fim de uma implementação é a política de redundância. Uma rede real introduz perda de pacotes e os terminais podem usar redundância para conseguir reconstruir os quadros perdidos. Isso pode ser tão simples quanto copiar um quadro duas vezes em pacotes consecutivos ou bem mais complexo.

A redundância aumentará o tamanho do pacote, mas pode não piorar o congestionamento se este for devido a uma capacidade de comutação insuficiente. O terminal pode usar uma heurística para determinar a melhor estratégia de redundância.

A redundância influencia o atraso de fim-a-fim porque o receptor precisa ajustar o seu *buffer* de *jitter* para receber todos os quadros redundantes antes de transferir o quadro ao decodificador. Caso contrário, se o primeiro quadro for perdido, o *buffer* de *jitter* seria incapaz de esperar até que ele tivesse recebido as cópias redundantes e, então, elas seriam inúteis. Isso pode contribuir significativamente para o atraso de fim-a-fim, especialmente se os quadros redundantes forem armazenados em pacotes não-contíguos para ser robusto a perdas correlacionadas de pacotes.

2.3. ACEITABILIDADE DE UMA CHAMADA TELEFÔNICA COM ECO E ATRASO

2.3.1. Curva G.131

O grau de importunação devido ao eco de locutor depende tanto da qualidade de atraso quanto da diferença de nível entre os sinais de voz e eco. Essa diferença de nível é caracterizada pela TELR (*Talker Echo Loudness Rating* – Medida do Volume de Eco de Locutor) como descrito em [2].

Segundo a figura 15, obtida em [1], os requisitos mínimos de TELR como função do tempo médio de transmissão unidirecional T são mostrados. Conforme [2], em geral, uma curva aceitável é de 1% (1% significa que, em média, 1% dos usuários vão se queixar do problema de eco), e é a curva que deve ser seguida. A curva de 10% é um limite extremo, que deve ser permitido apenas em casos excepcionais.

Essa figura mostra claramente que o eco torna-se mais audível à medida que o atraso aumenta. Essa é a razão pela qual o eco é um problema tão importante em todas as tecnologias de telefonia que introduzem atrasos elevados, que incluem a maioria das tecnologias de voz comprimida, transmissões de satélite – e, naturalmente, telefonia IP.

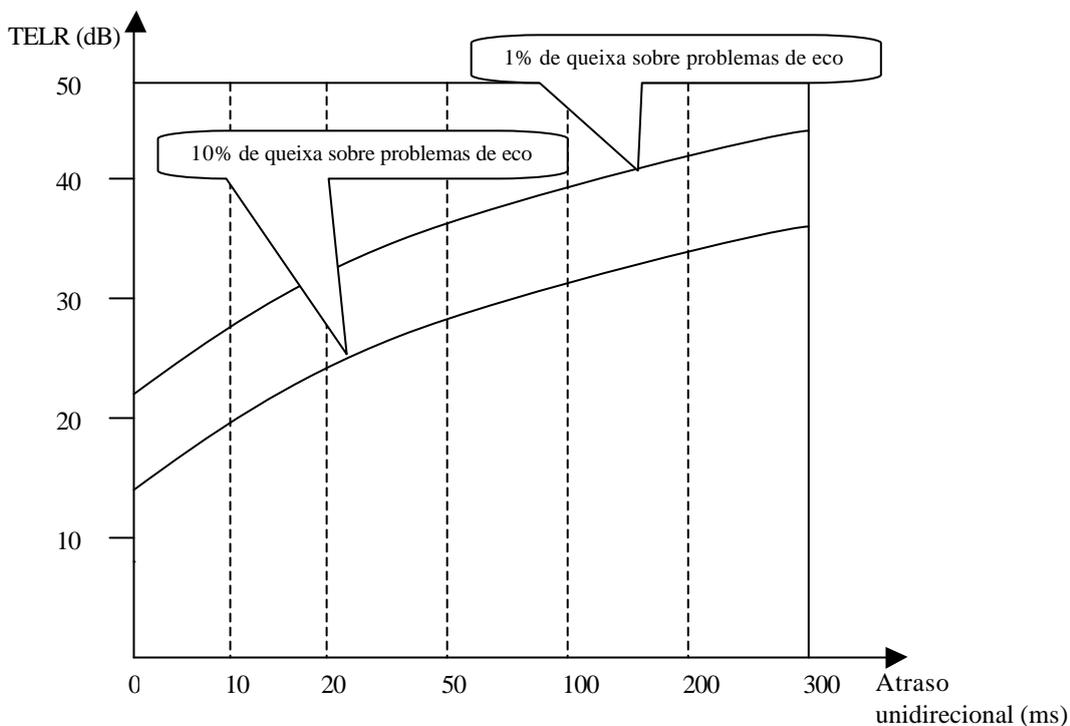


Figura 15. Curva G.131 que apresenta valores de TELR em função do atraso unidirecional

2.3.2. TELR

Foi utilizado o circuito de referência descrito na recomendação G.131 para o eco de locutor. Este circuito é descrito na figura 16.

Para telefones analógicos: $TELR = SLR + RLR + R + T + L_r$, onde R e T significam perda adicional introduzida no circuito analógico de maneira a se ter um ponto de 0 dBr na troca. A SLR (*Send Loudness Rating* – medida de volume de emissão) e a RLR (*Receive Loudness Rating* – Medida de Volume de Recepção) modelam a eficiência acústico/elétrica do receptor e do emissor respectivamente. Para telefones IP por *software*, esses valores podem ser afetados pelas configurações da placa de som e um bom *software* deve aplicar atenuação digital para assegurar-se de que as SLR e RLR resultantes ainda apresentem os valores recomendados.

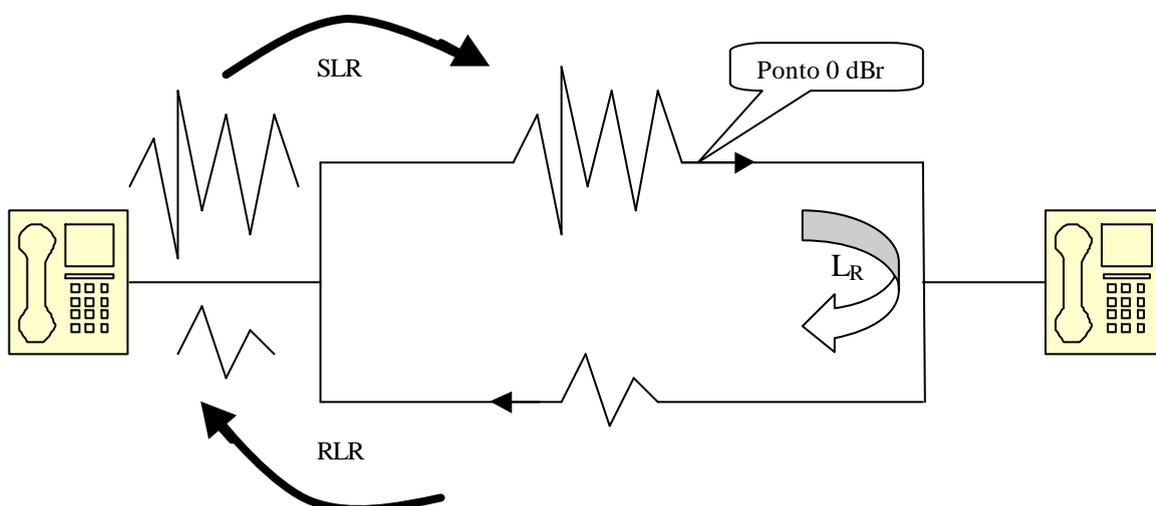


Figura 16. Circuito de referência para cálculo de TELR

Para telefones digitais: $TELR = SLR + RLR + TCL$

Valores típicos:

- Para telefones típicos, conforme descrito em [1], $SLR_{nom} = 7$ dB, $SLR_{min} = 2$ dB, $RLR_{nom} = 3$ dB, $RLR_{min} = 1$ dB. A margem de erro é de +/- 3 para telefones digitais e muito maior para telefones analógicos.

- A maioria das conexões de telefones analógicos têm um $L_r > 17$ dB para um comprimento médio do cabo do assinante; no entanto, em algumas redes ele pode ser de 14 dB com um desvio padrão de 3 dB.

- A maioria dos telefones digitais possui um TCL normalmente acima de 35 dB, em geral, na faixa de 40 a 46 dB.

Em muitas redes $R + T = 6$ dB (na França $T=0$ e $R=7$).

Exemplos:

Para $SLR=7$ dB, $RLR = 3$ dB, $L_r = 14$ dB, $R+T=6$ dB, obtemos um TELR igual a 30dB, o que leva a um limite aceitável para o atraso unidirecional de 18 ms (33 ms no caso limite). Para um aparelho telefônico 'alto' com $SLR = 2$ dB, $RLR = 1$ dB e um L_r igual a 8dB, obtemos um TELR igual a 17 dB e o atraso no caso limite agora é de 7ms. Ao ligar para um aparelho telefônico digital (TCL=45dB) com o telefone do locutor tendo um $SLR=7$ dB, $RR = 3$ dB, obtemos um TELR de 55dB, e o atraso unidirecional é 'aceitável' até o limite de 400ms.

2.3.3. Interatividade

No exemplo anterior, o termo 'aceitável' diz respeito apenas ao eco. A interatividade também deve ser considerada. Geralmente um atraso unidirecional abaixo de 100ms fornece uma boa interatividade. Um atraso unidirecional de 400ms deve ser excepcional (caso de dois enlaces via satélite) e é o limite acima do qual a conversa pode ser considerada *half duplex*.

Quando há uma demora elevada na linha, o locutor tende a pensar que o ouvinte não escutou ou não prestou atenção no que ele disse. O locutor então repetirá o que disse e será interrompido pela resposta atrasada, proveniente do lado de destino da chamada. Ambos vão parar de falar – e recomeçar simultaneamente. Com um pouco de treino, é possível comunicar-se corretamente, mas a conversa não é natural.

2.3.4. Modelo – E

O modelo-E foi desenvolvido originalmente na ETSI para atender às necessidades de planejamento de redes. Ele permite que a qualidade subjetiva de uma conversa seja avaliada como ela seria percebida por um usuário. O modelo-E associa cada fator de degradação na qualidade de voz percebida a um valor chamado 'fator de prejuízo'. Os fatores de prejuízo são então processados pelo modelo-E, que fornece como saída uma avaliação R entre 0 e 100. O valor de R pode ser mapeado em um valor MOS (*Mean Opinion Score* – Índice de Opinião Médio), ou em valores percentuais bom ou melhor (*% Good or Better* - %GoB) ou ruim ou pior (*% Poor or Worse* -

%PoW) usando-se tabelas. Um valor de R igual a 50 é muito ruim, enquanto um valor de R igual a 90 é muito bom.

O modelo-E leva em conta parâmetros que não são considerados na curva G.131, tais como a qualidade do codificador de voz. A maioria dos codificadores de voz foi avaliada com relação ao fator de prejuízo sem perdas de quadros e, conseqüentemente, o modelo-E (disponível como software comercial de vários fornecedores) pode ser usado facilmente para avaliar a qualidade de voz percebida através de uma rede de telefonia IP sem perda de pacotes e jitter baixo.

Os parâmetros do fator de prejuízo são avaliados a partir de testes subjetivos reais para calibrar o modelo. Portanto, a aplicabilidade do modelo-E para uma tecnologia em particular depende da quantidade de calibração realizada para essa tecnologia.

A telefonia IP, na maior parte dos casos, introduz perturbações que ainda não foram completamente avaliadas, tais como a influência da variação do atraso que pode ser introduzida em terminações que tentam ajustar dinamicamente o tamanho de seus *buffers* de *jitter*, saturação do sinal de voz introduzida por algoritmos VAD (*Voice Activity Detection*) ou perda correlacionada introduzida por agrupamento de quadros. Laboratórios de P&D especializados em qualidade de voz e planejamento de redes estão trabalhando nessas questões. Assim, novas versões de pacotes de software com o modelo-E e um melhor suporte a telefonia IP devem surgir em breve.

O modelo-E também leva em conta parâmetros psicológicos que não influenciam a qualidade absoluta da voz, mas influenciam a percepção do usuário. Por exemplo, a ‘expectativa’ do fator de prejuízo esperado leva em conta o fato de que a maioria dos usuários espera uma qualidade de voz degradada ao usar um telefone celular, portanto, serão mais indulgentes – e reclamarão menos – para o mesmo nível de qualidade do que se estivessem usando um telefone normal. Os fabricantes de telefones IP terão de encontrar um projeto reconhecível se quiserem beneficiar-se do ‘fator expectativa’.

2.3.5. Conseqüências para uma rede de telefonia IP

De certa maneira, as redes de telefonia IP e as redes de telefonia móvel (como o GSM) enfrentam muitas limitações semelhantes no que se refere à qualidade de voz. Do lado da telefonia IP, bem como do lado da telefonia ISDN, apenas o eco acústico é gerado uma vez que não existem híbridos. Com a tecnologia atual, mesmo 400ms parecem otimistas para o atraso unidirecional de fim-a-fim sobre algumas redes IP de grande extensão e qualquer atraso acima de 400 ms resulta em

uma perda real de interatividade entre as partes que estão conversando, mesmo com cancelamento perfeito de eco.

A conexão de telefonia IP está sujeita à mesma relação eco / atraso que qualquer outra conexão telefônica. Mesmo com um codificador de qualidade telefônica (*toll quality*), os tipos de atrasos unidirecionais encontrados na telefonia IP exigem um valor TELR de pelo menos 55 dB. Isso está próximo do mais alto valor alcançável para sinais de voz codificados com o G.711, devido ao ruído de quantização que é introduzido pelo G.711. Esse nível de cancelamento de eco pode ser reduzido a cerca de 30 dB sob condições de fala simultânea.

Telefone IP no PC para telefone IP no PC: se considerarmos que $SLR + RLR + 10$ dB, então a perda de eco em um telefone IP distante deverá ser no mínimo de $TCLw = 45$ dB. Isso poderia ser implementado nos periféricos de áudio (placa de som) ou no próprio software de telefonia IP (mas isso requer CPU de alto desempenho).

Telefone IP para um telefone comum: a maioria dos telefones ISDN possui um valor $TCLw$ de 45 dB, de maneira que o *gateway* de telefonia IP não precisa realizar cancelamento de eco no lado do telefone ISDN. Isso explica porque tantas demonstrações de *gateways* de telefonia IP propõem ligar para um telefone celular a partir de um telefone IP. As pessoas dirão que este é o pior caso porque, afinal de contas, esta é uma ligação para um telefone celular. Na verdade, muito frequentemente isso é feito de propósito, para ocultar a falta de dispositivos de cancelamento de eco no *gateway* de telefonia IP. O telefone celular possui um cancelador de eco poderoso, e a interface da rede de telefonia celular com a rede de telefonia convencional também é feita com canceladores de eco.

Obviamente, o telefone IP também precisa ter um cancelador de eco, ou o usuário ISDN ouvirá eco.

Telefone IP para usuário PSTN e vice-versa: o telefone PSTN gera eco híbrido e eco acústico. Como o tempo de propagação na PSTN geralmente é baixo, muitos *links* podem não implementar cancelamento de eco suficiente (se implementarem). Portanto, o *gateway* deve implementar cancelamento de eco. Novamente, isso se mostra difícil, uma vez que o *gateway* nem sempre está próximo do usuário PSTN, portanto, o atraso entre o sinal e o eco pode ser maior do que alguns canceladores de eco conseguem suportar. Na verdade, os canceladores de eco geralmente são caracterizados pela variação de atraso (*skew*) máxima entre os sinais que compõem

o eco. Esse sinal é uma superposição de sinais s_i que são uma cópia do sinal original mas atenuados por um fator A_i e atrasados por um fator $D + d_i$. Alguns canceladores de eco trabalham com $D=0$ e $0 < d_i < \text{Max}_{\text{skew}}$ (18 ms, por exemplo). Outros trabalham com valores de D de até 500 ms, e $0 < d_i < \text{Max}_{\text{skew}}$.

3. CONSIDERAÇÕES FINAIS

Embora a tendência tecnológica seja a convergência das redes telefônicas convencionais para a Internet, existem inúmeros fatores que devem ser relacionados e tratados adequadamente a fim de que se obtenha uma qualidade de voz satisfatória em uma conversação telefônica. Problemas como ecos podem ser facilmente resolvidos a partir da utilização de dispositivos de cancelamento e supressão. Entretanto, *jitter* e perdas de pacotes estão relacionados com o estado da rede em um momento específico, estando estes atrelados aos picos de tráfego na Internet ao longo do dia.

Ultimamente, VoIP possui muita aderência por parte de empresas corporativas devido a estas possuírem um ambiente de rede mais controlado, o que favorece a correta administração dos recursos de voz e conseqüentemente sua qualidade.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] G. David, P. Jean-Pierre, H. Oliver, “IP Telephony – Packet-based multimedia communications systems”, Pearson Education Limited 2000 – Great Britain

[2] ITU/BDT, “IP Telephony Report by the Group of Experts on Internet Protocol (IP) Telephony”, World Telecommunication Development Conference, 2002, Turkey