

Análise da Qualidade de Voz em uma Rede Ad Hoc Comunitária

Rafael B. Z. Antas, Natalia C. Fernandes, Danilo M. Taveira
Miguel Elias M. Campista, Luis Henrique M. K. Costa e Otto Carlos M. B. Duarte¹

¹Grupo de Teleinformática e Automação
PEE/COPPE-DEL/POLI
Universidade Federal do Rio de Janeiro
<http://www.gta.ufrj.br>

Abstract. *Different voice encoders were proposed for voice transmission over packet-switched networks. This article analyzes, based on the E model defined by the Resolution ITU-T G.107, the quality of these algorithms when applied to wireless ad hoc networks. Simulations evaluated the voice quality for different encoders considering the packets losses and the end-to-end delay. The results showed that the iLBC encoder is the most adaptive under these conditions, being viable to be applied into a community network for IP telephony.*

Resumo. *Diversos codificadores de voz foram propostos para a transmissão da voz em redes de comutação de pacotes. Este artigo analisa, baseado no modelo E, definido na Recomendação ITU-T G.107, a qualidade de alguns destes algoritmos quando aplicados em redes ad hoc sem fio. Para isso, resultados de simulações avaliam a qualidade da voz para diferentes codificadores, levando em consideração a perda de pacotes e o atraso fim-a-fim. Os resultados mostram que o codificador iLBC é o que melhor se adapta às condições adversas de perdas e atrasos.*

1. Introdução

A demanda por telefonia vem crescendo ao longo das últimas décadas, incentivando uma busca constante por novas tecnologias de transmissão de voz. A telefonia começou baseada na comutação de circuitos, onde existe a dedicação exclusiva de recursos para cada chamada. Com o advento da comutação de pacotes e o surgimento da Internet, foi criado um novo paradigma de comunicação. Isto gerou uma demanda por codificadores, para que fosse possível transmitir a voz digitalmente.

Com a popularização da Internet em banda larga e a convergência da telefonia com a Internet, surgiu o serviço de voz sobre IP. A voz sobre IP possui custos muito inferiores aos da telefonia convencional. No entanto, essa tecnologia ainda está restrita à Internet cabeada, cujos gastos com infra-estrutura, impossibilitam a aquisição do serviço pela parcela da população de mais baixo poder aquisitivo. Uma solução plausível para a redução dos custos da rede é o corte nos gastos com infra-estrutura.

As redes ad hoc são redes sem infra-estrutura e auto-configuráveis, que se adaptam bem a ambientes comunitários cuja alta densidade de usuários propicia alta conectividade.

A proposta deste trabalho é avaliar a utilização de voz sobre IP em redes ad hoc instaladas em comunidades carentes. Para tanto, foi simulada a utilização de alguns

dos codificadores de voz mais utilizados, o G.711 [ITU-T 1988], G.726 [ITU-T 1990], G.723 [ITU-T 1996a], e iLBC [Andersen et al. 2004].

Para avaliar a qualidade de cada um dos codificadores, foi utilizado o Modelo E, definido pelo ETSI (*European Telecommunications Standards Institute*). O Modelo E foi posteriormente normatizado pela ITU-T (*ITU Telecommunication Standardization Sector*) na recomendação G.107 [ITU-T 2003a].

A maioria dos trabalhos que aborda a utilização de codificadores e voz sobre IP limita-se a testes de desempenho em redes de pequeno porte, avaliando apenas a vazão e perdas de pacotes [Takaki et al. 2005] [da Silva e Rubinstein 2005]. Em contrapartida, ao aplicar o Modelo E, o presente trabalho leva em consideração não só os parâmetros da rede, mas também características inerentes aos próprios codificadores, como as técnicas de codificação. Lustosa propõe um Modelo E estendido, que avalia a rede segundo fatores que não são considerados no Modelo E básico do ETSI [Lustosa 2005]. Os requisitos necessários para avaliar o tráfego de voz em redes ad hoc são vistos em Velloso [Velloso et al. 2003].

Este trabalho foi dividido da seguinte forma. Na Seção 2, os conceitos sobre os codificadores utilizados e os métodos de avaliação da qualidade da voz, em especial o modelo E são discutidos. Na Seção 3, o ambiente em que foram feitas as simulações é exposto. Os resultados de simulação obtidos são mostrados e avaliados na Seção 4. A Seção 5 conclui este trabalho.

2. Codificadores e Qualidade da Voz

A telefonia é uma das aplicações mais populares que existe. Para aumentar a eficiência do compartilhamento de diversas comunicações telefônicas em uma mesma banda passante, diversos codificadores (*codecs*) de voz foram desenvolvidos. Os *codecs* podem ser classificados em duas grandes categorias: de forma de onda e paramétricos. Os codificadores de forma de onda fornecem um sinal codificado o mais próximo possível do sinal analógico original. Por sua vez, os codificadores paramétricos modelam o sistema que gera o sinal de voz original e enviam apenas parâmetros deste modelo. Muitos dos codificadores paramétricos se baseiam em predição linear (*Linear Prediction Code* - LPC) para alcançar altas taxas de compressão. Em contrapartida às altas taxas de compressão, os codificadores paramétricos são mais complexos e requerem um armazenamento para o processamento da voz codificada.

Um algoritmo de compressão de voz é normalmente avaliado em quatro quesitos: a eficiência, a complexidade, o atraso e a qualidade. A eficiência é normalmente representada pela taxa binária necessária para a transmissão da voz. A complexidade é medida em milhões de instruções por segundo (MIPS) necessárias na codificação do sinal. O atraso corresponde ao tempo em milissegundos que é necessário armazenar a voz para que o algoritmo de compressão seja aplicado.

O método mais usado para a avaliação da qualidade da voz é um método subjetivo conhecido como MOS (*Mean Opinion Score*). Definido nas Recomendações ITU-T P.800 [ITU-T 1996b] e ITU-T P.830 [ITU-T 1996d], o método se baseia na análise subjetiva de um conjunto de avaliadores que atribuem uma pontuação entre 1 (pobre) e 5 (excelente) à qualidade da fala ouvida. A média desses valores representa um índice de referência para a avaliação da qualidade da fala.

Atualmente, uma grande variedade de codificadores [ITU-T 1988, ITU-T 1996a, ITU-T 1990, iLBC 2005, 3GPP 2002] com taxas de transmissão estão disponíveis. Os codificadores avaliados neste trabalho serão brevemente descritos a seguir.

O codificador PCM (*Pulse Code Modulation*) é um codificador de forma de onda simples e de excelente qualidade de reprodução da voz. O sinal de voz é amostrado na frequência de 8 kHz (0,125 ms) e cada amostra é quantizada com 8 bits, fornecendo assim uma taxa binária de 64 kbps. Este *codec* não requer nenhum armazenamento para sua codificação e sua complexidade não atinge 1 MIPS. Este *codec* foi padronizado pelo ITU-T (*International Telecommunication Union - Telecommunication standardization sector*) na sua recomendação G.711 [ITU-T 1988]. Os quadros são compostos de 160 amostras (160 Bytes) o que implica em atraso de 20 ms para formação do pacote.

A Recomendação G.726 [ITU-T 1990] especifica um codificador de forma de onda cuja técnica de codificação é o ADPCM (*Adaptive Differential Pulse Code Modulation*). Assim como no PCM, o sinal também é amostrado a uma taxa de 8 kHz, mas há uma maior compactação do sinal de voz, o que lhe permite alcançar taxas de 40, 32, 24 e 16 kbps. Possui atraso de armazenamento em memória desprezível e seu atraso máximo de codificação não atinge 0,25 ms. A complexidade é considerada mediana, pois apresenta uma demanda de processamento de 10,54 MIPS. Para a taxa de 32 kbps e quadros de 80 bytes o tempo necessário para se formar um quadro é de 20 ms.

A Recomendação G.723.1 [ITU-T 1996a] define dois codificadores paramétricos que se servem das técnicas ACELP (*Algebraic Code Excited Linear Prediction*) para uma taxa de 5,3 kbps e MP-MLQ (*Multi Pulse-Maximum Likelihood Quantization*) para a taxa de 6,3 kbps. Assim como no PCM e ADPCM a frequência de amostragem é de 8 kHz. A complexidade dos algoritmos de compressão é média, correspondendo entre 10 e 11 MIPS. Para a taxa de 6,3 kbps são usados quadros de 24 Bytes o que corresponde a um tempo de aproximadamente 30 ms. A técnica de predição linear usada analisa 7,5 ms do quadro seguinte, o que é chamado tempo de *look-ahead*, implicando em um atraso de armazenamento de 37,5 ms.

O iLBC (*Internet Low Bit rate Codec*) [iLBC 2005] é um codificador de voz paramétrico, voltado para comunicações de voz sobre IP, numa rede de tráfego intenso como a Internet. É utilizado pelo Gnomemeeting [Gnome Meeting 2005] e pelo Skype [Skype 2005]. A técnica de codificação utilizada consiste em um algoritmo LPC independente de bloco [Andersen et al. 2004]. O iLBC alcança uma taxa de 13,3 kbps ou 15,2 kbps com atraso de *buffering* de 30 ms ou 20 ms, respectivamente. O atraso máximo de codificação é 60 ms no caso da taxa de 13,3 kbps e 40 ms para 15,2 kbps. Este *codec* oferece boa qualidade de voz na ocorrência de perdas de pacotes e apresenta uma complexidade mediana, já que sua implementação em DSP exige cerca de 8 a 9 MIPS.

Um resumo dos principais parâmetros que foram levados em conta na análise é apresentado na Tabela 2 [ITU-T 2003c], onde o quadro é o bloco gerado na saída do codificador. A taxa de codificação é dada pela divisão do quadro em bytes pelo quadro em milissegundos. O atraso de codificação corresponde a duas vezes o tamanho do quadro, devido à codificação e decodificação, somado ao tempo de *look-ahead*.

Para realizar testes de qualidade subjetivos, como o MOS, é necessário um ambiente especial, livre de ruídos. Para contornar este inconveniente, métodos objetivos

Tabela 1. Parâmetros dos codificadores.

Codificador	Taxa de Codificação (kbps)	Quadro (ms)	Quadro (Bytes)	Atraso de Codificação (ms)	MOS
G.711	64	0,125	8	0,25	4,1
G.726	32	0,125	4	0,25	4
G.723	6,3	30	24	67,5	3,85
	5,3	30	20	67,5	3,65
iLBC	15,2	20	38	40	3,9
	13,3	30	50	60	3,8

de medição da qualidade da voz foram propostos. Os modelos objetivos, tais como o PESQ (*Perceptual Evaluation of Speech Quality*) [ITU-T 2001], o PSQM (*Perceptual Speech Quality Measure*) [ITU-T 1996c], o PAMS (*Perceptual Analysis Measurement System*) [Rix e Hollier 2000] e o Modelo E, utilizam recursos computacionais para obter a qualidade da voz que está sendo transmitida em um sistema de comunicação. Os três primeiros algoritmos são classificados como algoritmos perceptuais, ou seja, utilizam técnicas de processamento de sinais para comparar um trecho de voz gravado com um trecho de voz degradado. A partir das diferenças, constrói-se um modelo de avaliação da qualidade da voz. Já o modelo E baseia-se apenas em medidas de tráfego da rede e nas especificidades dos codificadores.

Como os parâmetros deste modelo são medidas da rede, pode-se tanto fazer uma medição em tempo real quanto contabilizar separadamente cada um dos fatores que influenciam na perda da qualidade da voz. Isto permite que o diagnóstico e a correção dos problemas do tráfego sejam feitos com mais precisão pelo gerente da rede, representando a maior vantagem do modelo E em relação aos métodos perceptuais. Devido à esta particularidade, o modelo E foi selecionado para a avaliação da voz neste trabalho.

Para assegurar a qualidade de voz, é necessário avaliar requisitos que garantam a qualidade do sinal e a interatividade da conversação. Uma transmissão de voz em tempo real deve apresentar, por exemplo, uma boa relação sinal-ruído, baixas perdas consecutivas de pacotes, pequenos atrasos fim-a-fim e uma baixa variação desses atrasos. Dentre esses requisitos, os atrasos são os maiores responsáveis pela perda de interatividade. Assim, é preciso controlar atrasos de codificação, de decodificação e de filas na rede.

A variação de atraso (*jitter*) também é responsável pela degradação da voz, visto que há necessidade de se respeitar uma cadência de reprodução da voz para que se mantenha a inteligibilidade. Para minimizar os efeitos do *jitter*, utilizam-se memórias (*buffers de jitter*) para garantir a cadência do fluxo de voz. Estas memórias introduzem atrasos que devem ser considerados na avaliação da qualidade da conversação.

O modelo E baseia-se na idéia de que fatores psicológicos numa escala subjetiva são aditivos [ITU-T 2003a], ou seja, é possível somar fatores diferenciados para obter um grau da qualidade da voz em um sistema de comunicação. O resultado desse somatório é computado na variável R, que varia de 0 a 100, onde 100 é excelente, o que corresponde à voz real, e 0 é considerado péssimo. A variável R pode ser convertida para a escala MOS, utilizando a Expressão 1. No entanto, existe uma degradação inerente à conversão

analógico-digital da voz, e posteriormente a conversão digital-analógico. Isto reduz o valor teórico máximo de R para 93,2, e um valor MOS máximo de 4,41. Por essa razão, a norma G.107 [ITU-T 2003a] determina um MOS máximo de 4,5 para aplicações de voz sobre IP.

$$MOS = 1 + 0,035R + 7.10^{-6}R(R - 60)(100 - R), \quad (1)$$

onde $0 \leq R \leq 100$. O grau de satisfação do usuário pode ser descrito segundo intervalos do valor R como visto na Tabela 2. O fator R é dado pela Expressão 2 que contém os

Tabela 2. Categorias de transmissão de voz.

Fator R	MOS	Satisfação do Usuário
$90 \leq R \leq 100$	4,34-4,50	Muito satisfeito
$80 \leq R \leq 90$	4,03-4,34	Satisfeito
$70 \leq R \leq 80$	3,60-4,03	Alguns insatisfeitos
$60 \leq R \leq 70$	3,10-3,60	Muitos insatisfeitos
$50 \leq R \leq 60$	2,60-3,10	Quase todos insatisfeitos
$0 \leq R \leq 50$	1 - 2,60	Não recomendado

parâmetros necessários para medir a qualidade da voz:

$$R = R_o - I_s - I_d - I_{e,eff} + A \quad (2)$$

onde R_o representa o nível de ruído básico; I_s , as perdas simultâneas à fala, quantificando o ruído de quantização; I_d , as perdas associadas ao atraso fim-a-fim; $I_{e,eff}$, as perdas associadas aos equipamentos utilizados, o que inclui a perda de pacotes na rede, o descarte do *buffer* de compensação de *jitter* e o codificador utilizado; e A, um fator de vantagem, que para a telefonia fixa e VoIP vale 0.

O fator R_o é obtido a partir do ruído gerado pelos circuitos de transmissão, do ruído ambiente no receptor e no emissor e do teto de ruído correspondente à sensibilidade auditiva humana. Esses valores podem ser considerados constantes e são determinados na norma ITU-T G.107 [ITU-T 2003a], resultando em um R_o igual a 94,77.

A variável I_s é calculada considerando as perdas e interferências simultâneas à voz do emissor, tais como um volume demasiadamente alto, interferência da voz do locutor sobre o fone de ouvido ou distorções de quantização gerada pela digitalização do sinal [Lustosa 2005]. Esse valor também pode ser considerado constante e igual a 1,41 [ITU-T 2003a].

O fator I_d , que está relacionado às perdas associadas ao atraso fim-a-fim, tem uma estimativa mais complexa dada pela Expressão 3. Na Expressão 3, os termos I_{dte} e I_{dle} representam as perdas geradas pelo eco no transmissor e no receptor, respectivamente. Já o I_{dda} representa as perdas relacionadas ao atraso absoluto (T_a), quando o atraso da voz é superior a 100 ms.

$$I_d = I_{dte} + I_{dle} + I_{dda} \quad (3)$$

onde:

$$I_{dte} = f(T, N_o, RLR, TELR), \quad (4)$$

$$I_{dle} = f(RLR, TELR, T, WELP, T_R) \text{ e} \quad (5)$$

$$I_{dd} = \begin{cases} 0, & T_a \leq 100 \text{ ms} \\ f(T_a), & T_a > 100 \text{ ms}. \end{cases} \quad (6)$$

Cabe observar que I_{dd} , por medir perdas relacionadas ao atraso absoluto da rede (T_a), representa diretamente a perda de interatividade. Uma vez que um $T_a < 100\text{ms}$ não causa um atraso perceptível na conversação, neste caso, I_{dd} passa a valer zero. Para valores superiores a esse nível de atraso, é utilizada uma função crescente para quantificar a perda de interatividade. As constantes RLR , $TELR$, $WELP$, N_o e R_o estão definidas na Tabela 2. T_a , T_R e T são atrasos característicos da rede, e estão representados na Figura 1, onde T corresponde ao atraso médio percorrido. A variável T_a corresponde ao atraso absoluto no sistema sem eco e, T_r , ao tempo de ida e volta no circuito a quatro fios, ou seja, o dobro do intervalo representado na Figura 1. As expressões 4, 5 e 6 detalhadas e os cálculos dos seus parâmetros são encontrados em [ITU-T 2003a].

Tabela 3. Parâmetros utilizados no cálculo de I_{dle} , I_{dle} e I_{dd} .

Grandeza	Significado	Valor padrão
N_o	Ruído total em dBm0p	-61,18 dBm0p
R_o	Relação sinal-ruído básica	94,77
RLR	Nível de intensidade no receptor	2,00 dB
$TELR$	Nível de intensidade do eco no emissor	65,00 dB
$WELP$	Caminho do eco ponderado	110,00 dB

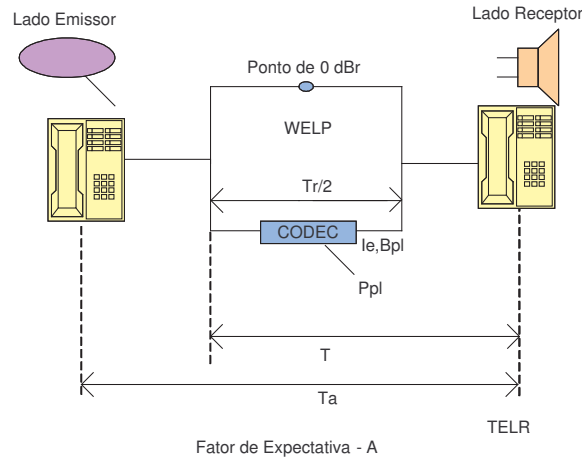


Figura 1. Representação dos atrasos em um sistema de comunicação de voz.

Uma aproximação realista para o cálculo de T , T_a e T_r , que corresponde ao caso prático de um cancelador de eco quase ideal [ITU-T 2006], é dado por:

$$T = T_a = \frac{T_r}{2}. \quad (7)$$

O cálculo de T_a é dado pela Expressão 8. É visto que o atraso absoluto corresponde ao atraso de codificação (T_{codec}) somado ao atraso do *buffer* para compensar o *jitter* (T_{buffer}) mais o atraso de rede (T_{rede}). Este último representa o atraso de propagação e o tempo de espera em fila.

$$T_a = T_{codec} + T_{rede} + T_{buffer} \quad (8)$$

O fator $I_{e,eff}$ representa um meio para se calcular a degradação da qualidade do sinal de codificadores que geram baixas taxas de transmissão, mas possuem uma grande complexidade. Seu valor é dado por uma média de exaustivos testes de pontuação MOS, considerando diversas taxas de perda de pacotes. Os valores de I_e para os codificadores considerados neste trabalho, no caso de haver perda de pacotes, estão dispostos na Tabela 2 [ITU-T 2003b, Laboratories 2005].

O valor efetivo de I_e ($I_{e,eff}$) é dado por

$$I_{e,eff} = I_e + \frac{(95 - I_e)P_{pl}}{P_{pl} + B_{pl}}, \quad (9)$$

onde B_{pl} é o fator de robustez à perda de pacotes e P_{pl} é a probabilidade de perda de pacotes [ITU-T 2003a].

O valor de P_{pl} deve ser substituído pelo percentual de perda de pacotes obtido pela observação do sistema em questão. Isto deve ser feito após o *buffer de jitter* (variação de atraso), pois pacotes podem ser entregues à aplicação VoIP tarde demais para que sejam utilizados. O valor do fator de robustez à perda de pacotes (B_{pl}) depende diretamente da técnica de codificação inerente a cada *codec*, e deve seguir o que está especificado na Tabela 2. O fator de vantagem A é utilizado para determinar o grau de tolerância do

Tabela 4. Valores de I_e e B_{pl} específicos.

Codificador	I_e	B_{pl}
G.711	0	10
G.723.1	15	20
G.726	12	24
iLBC	10	34

usuário, medindo a conveniência esperada para determinada tecnologia. Essa variável vale 0 para telefonia fixa e VoIP, 10 para celular e 20 para enlaces de satélite. No entanto, esta não é uma métrica fixa, já que esse fator se trata de uma variável dependente do grau de aceitação do usuário.

3. Ambiente de Simulação

O simulador utilizado foi o Network Simulator 2 (ns-2) [Fall e Varadhan 2002]. O cenário analisado neste trabalho leva em consideração valores típicos de dispositivos que implementam o padrão de redes locais sem fio IEEE 802.11g, no modo Ad Hoc. Assim, foi atribuída aos nós uma potência de transmissão de 18 dBm, ou 63,1 mW [Bengtsson et al. 1999], e uma sensibilidade de recepção de -65 dBm [IEEE 2003]. A taxa de transmissão da camada física é de 54 Mbps.

O cenário escolhido simula uma comunidade carente com alta densidade de residências. Conseqüentemente, há uma baixa probabilidade de enlaces em visada direta. Além disso, espera-se que os dispositivos de comunicação estejam estaticamente situados no interior das residências. Logo, o modelo de propagação aproxima-se mais de um ambiente interno (*indoor*), com maior perda do que um ambiente externo (*outdoor*). Estas características justificam o emprego do modelo de propagação de perda de percurso [Rappaport 2001].

Os nós foram posicionados em uma grade. Nesta grade, um dos vértices é o *gateway* para Internet. O nó *gateway* é o receptor de todas as chamadas telefônicas e é responsável por encaminhá-las a seus destinos. A simulação foi realizada com 81 nós a uma densidade de 12100 residências/ km^2 , representando uma região como a favela da Rocinha, localizada no Rio de Janeiro.

O tamanho da fila de interface (*Interface Queue - IFQ*) da subcamada MAC é de 50 pacotes. Este parâmetro é importante porque o descarte de pacotes pela subcamada MAC, devido ao transbordo da fila IFQ, é um dos motivos de perdas de pacotes.

Como protocolo de roteamento ad hoc, escolheu-se o protocolo AODV (*Ad-hoc On-demand Distance Vector*), um protocolo reativo, que só busca rotas quando existe um pedido de conexão. Assim, como o cenário é estático, rotas construídas inicialmente não sofrerão alterações por quebra de enlace devido à mobilidade dos nós, o que irá gerar pouco tráfego de roteamento.

A aplicação voz sobre IP se serve da pilha de protocolos constituída pelos protocolos RTP (*Real-time Transport Protocol*), UDP (*User Datagram Protocol*), IP e IEEE 802.11. Assim, a voz é encapsulada pelo RTP, que adiciona 12 Bytes de cabeçalho, depois pelo UDP, que acrescenta mais 8 Bytes de cabeçalho e pelo IP, adicionando mais 20 Bytes ao pacote. É importante ressaltar que a influência dos cabeçalhos é significativa no desempenho das aplicações voz sobre IP, pois os 40 bytes de cabeçalho correspondem a um acréscimo de até 160% no tamanho dos quadros de voz codificada.

Com o objetivo de simular fluxos de voz, as fontes foram modeladas segundo uma cadeia de Markov de dois estados (*on - off*), que representam períodos de atividade e silêncio. Nos períodos de atividade o tráfego de voz gerado é modelado por uma fonte com taxa binária constante (*Constant Bit Rate - CBR*). Já nos períodos de silêncio nenhum tráfego é gerado. O tempo de permanência em cada estado é representado por uma variável aleatória com distribuição exponencial de média 1,2 s para o estado ativo e 1,8 s para o estado de silêncio [Velloso et al. 2003].

Os codificadores avaliados foram o G.711, o G.726, na taxa de 32 kbps, o G.723, na taxa de 6,3 kbps e o iLBC, na taxa de 13,3 kbps. A taxa de transmissão no período ativo e o tamanho dos pacotes correspondentes são funções dos codificadores utilizados nas simulações, conforme visto na Tabela 2. O número de conversações, definidos como pares de fluxos de voz, foi variado entre 10 e 80. Foi ainda empregado um tráfego de fundo constante de 2 Mbps para simular a agregação de fluxos concorrentes aos tráfegos de voz. Para todas as medidas, foram calculados intervalos de confiança de 90% relativos à média das amostras. Estes intervalos são representados por barras de erros verticais.

4. Resultados

A primeira métrica utilizada para comparar os codificadores foi a taxa de entrega de pacotes, que indica a relação entre o número de pacotes de voz enviados pela fonte e o número de pacotes de voz recebidos pelo destino. Deve ser ressaltado que são descartados, por transbordo, pacotes que encontram cheios os *buffers* da camada MAC e o *buffer* de *jitter* da aplicação. Este último caso ocorre quando o atraso na rede é maior do que o tempo configurado no buffer de *jitter*.

Na Figura 2, percebe-se que os codificadores que impõem maior carga à rede

possuem maior perda de pacotes. Pode-se observar, também, que os codificadores G.711 e G.726 saturam a rede antes que os *codecs* de menor taxa de dados, G.723 e iLBC.

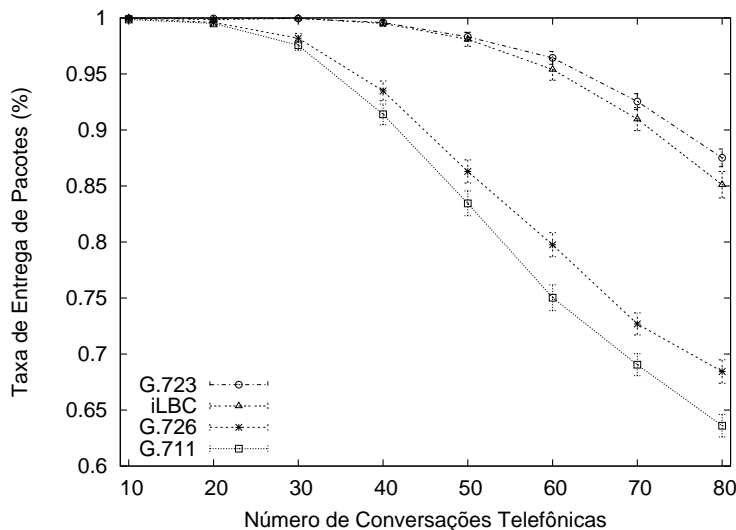


Figura 2. Taxa de entrega de pacotes.

A garantia de interatividade da aplicação voz sobre IP impõe que o atraso fim-a-fim dos pacotes na rede seja limitado. A Figura 3 mostra que o atraso não ultrapassa o valor de 150 ms. Assim, não acontece perda perceptível de interatividade em até cinquenta comunicações telefônicas simultâneas, qualquer que seja o *codec* de voz utilizado. Ainda na Figura 3, pode ser observado que o algoritmo G.723 é o que apresenta o melhor comportamento em relação ao atraso dentre os codificadores testados. Isto também se deve a baixa carga que o codificador agrega a rede, o que diminui o atraso nas filas.

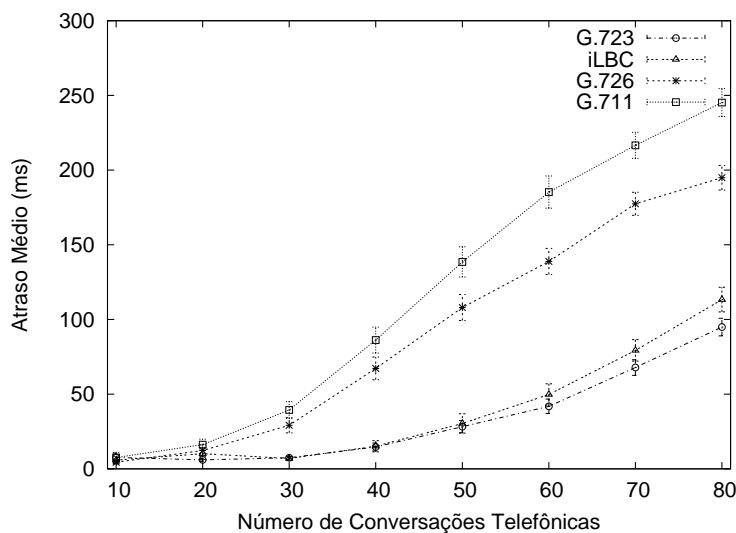


Figura 3. Atraso dos pacotes.

Com base nos resultados do atraso de rede (T_{rede}), na porcentagem de pacotes perdidos (P_{pl}), e nas singularidades de cada codificador, foi aplicado o Modelo E para

avaliar a qualidade da voz. A Figura 4 mostra o gráfico da qualidade da voz em função do número de chamadas telefônicas simultâneas.

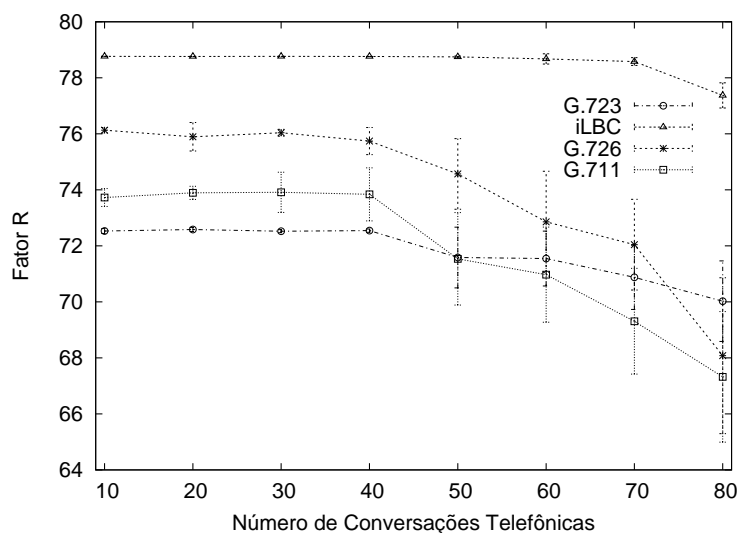


Figura 4. Análise da qualidade.

Em uma primeira análise dos resultados, nota-se que o codificador G.723 possui qualidade inferior aos demais para poucas conversações telefônicas. O modelo E nos permite observar que, apesar da superioridade do codificador G.723 nos quesitos de rede, seu atraso de codificação compromete o seu uso. Este fato é evidenciado na comparação do G.723 com os codificadores G.711 e G.726, após cinquenta pares de comunicação. Uma vez que o G.723 apresenta menor demanda de recursos de rede, com o aumento do número de pares de conversação, o congestionamento não influencia muito o seu valor R. Por outro lado, os algoritmos G.711 e G.726, apesar do menor atraso de codificação, são muito prejudicados pela saturação da rede. É visível o desempenho superior do codificador iLBC não importando o número de pares de comunicação. Isso se deve, em parte, ao equilíbrio entre a taxa de transmissão, ao atraso de codificação, e ao elevado fator de robustez (B_{pl}) do algoritmo. Este fator também sustenta a qualidade de transmissão da voz mesmo após a saturação da rede.

O valor alto do fator de robustez é devido ao uso de predição linear pelo codificador iLBC. Esta técnica, no caso de perdas de curta duração, interpola os estados dos coeficientes preditores lineares dos quadros antes e depois da perda, de forma que não é notada uma diferença significativa da qualidade da voz.

Nota-se também que quando se compara o valor R inicial de cada codificador com os respectivos MOS dados na Tabela 2 existe disparidade. Isto se deve ao escopo de cada um dos métodos de avaliação de qualidade. Enquanto o MOS verifica a qualidade do codificador quando posto em um ambiente ideal, o modelo E leva em consideração os efeitos da rede e dos ruídos adicionados pelo ambiente.

5. Conclusão

Este trabalho avaliou uma arquitetura de rede capaz de ofertar, a baixos custos, aplicações VoIP a comunidades carentes. Sobre uma rede IEEE 802.11 comunitária em

modo ad hoc, foram testados quatro codificadores de voz diferentes. Estes codificadores foram submetidos a uma avaliação da qualidade da voz baseada no modelo E.

Após as simulações pôde-se concluir que, dentre os codificadores testados, o iLBC é o que melhor se adapta ao cenário, provendo um serviço de qualidade aceitável para até 81 conversações simultâneas. Além disso, a qualidade de serviço (QoS) se mostra estável mesmo para casos pouco prováveis, onde o número de conexões atinge o número de nós da rede.

O uso do modelo E mostrou que o fator que mais influencia na degradação da qualidade da voz em redes ad hoc para os codificadores G.711 e G.726 é o aumento da carga. Uma possível solução para este problema é a adoção de codificadores de baixa taxa de geração de pacotes como o iLBC. No entanto, isso deve ser ponderado com o atraso de codificação, para evitar perdas de interatividade, como ocorreu com o G.723.

Referências

- 3GPP (2002). Adaptive multi-rate (AMR) speech codec; transcoding functions. Relatório Técnico TS 26.090, 3GPP.
- Andersen, S., Duric, A., Astrom, H., Hagem, R., Kleijn, W. e Linden, J. (2004). Internet low bit rate codec (iLBC). RFC 3951, Global IP Sound.
- Bengtsson, H., Uhlemann, E. e Wiberg, P. (1999). Protocol for wireless real-time systems. *Em 11º Euromicro Conference on Real-Time Systems.*
- da Silva, M. W. R. e Rubinstein, M. G. (2005). Uma análise da capacidade de transmissão de voz em redes IEEE 802.11 no modo ad-hoc. *Em XXII Simpósio Brasileiro de Telecomunicações*, páginas 1230–1231.
- Fall, K. e Varadhan, K. (2002). *The ns Manual*.
- Gnome Meeting (2005). Gnome Meeting. <http://www.gnomemeeting.org/>.
- IEEE (2003). Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 2.4 GHz band. IEEE Standard 802.11g, IEEE.
- iLBC (2005). iLBC freeware. <http://www.ilbcfreeware.org/>.
- ITU-T (1988). Pulse code modulation (PCM) of voice frequencies. ITU-T Recommendation G.711, ITU-T.
- ITU-T (1990). 40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM). ITU-T Recommendation G.726, ITU-T.
- ITU-T (1996a). Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s. ITU-T Recommendation G.723.1, ITU-T.
- ITU-T (1996b). Methods for subjective determination of transmission quality. ITU-T Recommendation P.800, ITU-T.
- ITU-T (1996c). Objective quality measurement of telephone-band speech codecs. ITU-T Recommendation P.861, ITU-T.
- ITU-T (1996d). Subjective performance assessment of telephoneband and wideband digital codecs. ITU-T Recommendation P.830, ITU-T.

- ITU-T (2001). Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. ITU-T Recommendation P.862, ITU-T.
- ITU-T (2003a). The E-Model, a computational model for use in transmission planning. ITU-T Recommendation G.107, ITU-T.
- ITU-T (2003b). Estimates of I_e and B_{pl} parameters for a range of codec types. ITU-T Delayed Contribution 106, ITU-T.
- ITU-T (2003c). One-way transmission time. ITU-T Recommendation G.114, ITU-T.
- ITU-T (2006). Introduction to E Model. <http://www.itu.int/ITU-T/studygroups/com12/emodelv1/introduction.htm/>.
- Laboratories, C. T. (2005). Audio/video codecs. PacketCable 1.5 Specifications PKT-SP-CODEC1.5-I01-050128.
- Lustosa, L. C. G. (2005). Arquitetura de monitoração de qualidade de chamadas telefônicas IP. Master's thesis, IM/NCE/UFRJ.
- Rappaport, T. S. (2001). *Wireless Communications, Principles and Practice*. Prentice Hall Professional Technical Reference, 2ª Edição.
- Rix, A. W. e Hollier, M. P. (2000). The perceptual analysis measurement system for robust end-to-end speech quality assessment. Em *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, páginas 1515–1518.
- Skype (2005). Skype. <http://www.skype.com/>.
- Takaki, R., Santos, P. H. M., Vicentini, M. L. e Martins, J. A. (2005). Avaliação de codificadores de voz em uma rede ad hoc multihop sem fio. Em *XXII Simpósio Brasileiro de Telecomunicações*.
- Velloso, P. B., Rubinstein, M. G. e Duarte, O. C. M. B. (2003). Uma análise da capacidade da transmissão de voz em redes ad hoc. *XXI Simpósio Brasileiro de Redes de Computadores - SBRC'2003*, páginas 775–786.