

Determinação do comportamento de codificadores em diferentes condições de rede em um cenário experimental

Demóstenes Zegarra Rodriguez, Miguel Arjona Ramírez

Resumo - Perda de pacotes e atrasos influenciam na qualidade de uma conversa interativa em uma rede IP. Este artigo mostra uma metodologia para prever o comportamento de diferentes codificadores, utilizando modelos objetivos como o PESQ e a Recomendação E-Model do ITU-T. A metodologia apresentada é útil para medir a performance de novos codificadores, ou de codificadores que ainda não possuam parâmetros definidos em diferentes cenários de perda de pacotes e atrasos, em uma comunicação bidirecional utilizando uma rede emulada com open source software, sendo o cenário facilmente reproduzido.

Palavras-Chave - Voz sobre IP, MOS, codificador, E-Model, PESQ.

Abstract - Packet losses and delays have influence in the quality of an interactive voice conversation in an IP network. This article exhibits a methodology to predict the behavior of different codecs using objective models as PESQ and E-Model Recommendation from ITU-T. This methodology is useful to measure the performance of new codecs, or that do not have defined parameters in different scenarios of packet losses and delays in a bidirectional communication using a network emulated with software open source, where the scenario can be easily reproduced.

Keywords - VoIP, MOS, codec, E-Model, PESQ.

I. INTRODUÇÃO

NO projeto de redes de pacotes devem-se considerar diversos parâmetros como: perda de pacotes, atrasos fixos e variáveis, taxa de codificação, taxa total de transmissão e o eco [1], [2], [3] e [4] que afetam diretamente a qualidade da transmissão e, conseqüentemente, a qualidade do sinal no receptor.

É necessário que o projetista de rede conheça em que medida cada parâmetro afeta a qualidade em uma comunicação. Este trabalho está orientado ao estudo do fator de degradação do codificador definido na recomendação ITU-T G.107 [5], mais conhecida como: E-model. A recomendação ITU P.862

[6] geralmente conhecida como PESQ (perceptual evaluation of speech quality), em termos gerais, descreve um método objetivo para avaliar a qualidade subjetiva de codificadores. É importante considerar que este algoritmo estabelece uma avaliação precisa em comunicações unidirecionais, mas não permite medir a qualidade da transmissão em uma comunicação bidirecional, sendo possível obter resultados altos de qualidade mesmo com uma conexão global ruim, assim, por exemplo, o atraso ponto a ponto em uma rede de transmissão de pacotes afeta a qualidade de voz somente nas comunicações bidirecionais.

No caso de uma transmissão em uma rede ideal (sem perdas e nem atrasos) ou que apresente somente perda de pacotes, os resultados do PESQ são confiáveis e estes podem ser utilizados junto com a Recomendação ITU-T G.107 para calcular o parâmetro que corresponde às características próprias de cada codificador, e que é denominado nesta recomendação como I_e , fator de degradação do codificador. Conhecendo este parâmetro, pode-se estimar o comportamento do codificador em estudo em diferentes cenários de atraso e perdas de pacotes.

Na Recomendação ITU G.108 [7] é apresentada uma explicação de como se devem interpretar as equações da Recomendação ITU-T G.107 apresentando-se alguns casos práticos que mostram como este método é corretamente utilizado. Na recomendação ITU-T G.113 Anexo I [8] são apresentados os valores de I_e para diferentes codificadores que já foram estudados pela ITU-T, estes valores são obtidos de maneira experimental e são atualizados periodicamente.

Portanto, neste artigo se propõe fazer medidas experimentais com os codificadores de voz Rec. ITU-T G.723.1 [9] e iLBC [10] em um cenário de rede emulada e utilizando a ferramenta PESQ para estimar o índice MOS (mean opinion score) descrito na Recomendação ITU-T P.800 [11] para diferentes cenários de probabilidade de perda de pacotes (incluindo perda de pacotes igual a zero). Com estes valores obtidos e empregando as equações do E-Model que relacionam o valor do índice MOS com o valor de determinação do índice de transmissão R , pode-se calcular o parâmetro I_e do codificador em estudo.

A metodologia seguida para estes codificadores é aplicável para os codificadores novos que ainda não tenham parâmetros definidos e a partir destes, fazer uma estimativa de sua qualidade em diferentes cenários de probabilidade de perda de pacotes e em diferentes tempos de atraso em comunicações bidirecionais.

Este artigo é inovador pelo fato da utilização de ferramentas

Laboratório de Processamento de Sinais da Universidade de São Paulo, Brasil, E-mails: demostenes@lps.usp.br, miguel@lps.usp.br

Este trabalho foi financiado parcialmente pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) através dos processos 2007/08288-2 e 2004/13206-7 e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) através de uma Bolsa de Mestrado.

open source que podem ser facilmente reproduzidas e assim possibilitar a comprovação de resultados e permitir que se obtenha os parâmetros I_e de novos codificadores que estão sendo implementados, estimulando o surgimento destes.

O artigo está dividido respectivamente em: Seção 2, onde é feito um resumo teórico das ferramentas utilizadas. Seção 3 que apresenta o cenário onde são realizados os testes indicando a metodologia seguida e os parâmetros avaliados. Seção 4, onde são apresentados e discutidos os resultados obtidos, e por fim, a Seção 5 que expõe a importância da determinação do fator de degradação do codificador descrito no E-Model em um cenário de rede experimental.

II. REVISÃO TEÓRICA

Os métodos utilizados neste trabalho foram o método intrusivo, ou seja, que requer dois sinais de referência, PESQ e o método não intrusivo, que só requer de um sinal, E-Model.

O PESQ é uma ferramenta objetiva na avaliação da qualidade de voz, utilizando como entradas no seu algoritmo, dois sinais: o sinal original e o sinal degradado.

O E-model é um modelo computacional que mede os efeitos da variação dos parâmetros de uma rede de transporte e do sinal na qualidade de voz transmitida.

O E-Model é definido pela seguinte equação:

$$R = R_o - I_s - I_d - I_e + A \quad (1)$$

Onde:

- R: fator de determinação do índice de transmissão, representa o parâmetro que está relacionado com o valor de qualidade e tem uma correspondência com a pontuação MOS UIT-T P.800.
- R_o : relação sinal-ruído.
- I_s : fator de degradações simultâneas, representa todas as degradações que se produzem simultaneamente com o sinal de fala, como são os níveis dos sinais de fala de maior intensidade (OLR não ótima), o efeito local não ótimo (STMR) e o ruído de quantificação.
- I_d : degradação da qualidade causada devido ao atraso existente na rede.
- I_e : degradação da qualidade por efeito do equipamento (codificador).
- A: fator de melhora, não tem relação com os demais parâmetros da transmissão.

O valor padrão de R_o é 93.2 [5], que é obtido ao se colocar todas as entradas do modelo com seus valores padrões. O parâmetro I_s não será considerado nos cálculos, já que este descreve as condições que são relacionados com o sinal, não dependendo da rede de transporte. O fator A, segundo [5] tem o valor 0 (zero) para redes com fio e é o que corresponde ao cenário de emulação deste trabalho.

O fator I_d é definido como:

$$I_d = I_{dte} + I_{dle} \quad (2)$$

Os fatores I_{dte} e I_{dle} correspondem ao atraso causado pelo eco no emissor e no receptor, respectivamente. Estes fatores

não serão considerados neste trabalho já que para o cenário de testes se assumirá uma supressão perfeita do eco.

O I_{dte} representa os atrasos produzidos na rede e no codificador. O atraso ocasionado pelo codificador se define em [3] como:

$$I_{dte} = 2N + \text{Pré-análise} \quad (3)$$

Onde N é o tamanho da amostra de cada codificador. Valores de N e dos tempos de pré-análise para diferentes codificadores podem ser encontrados em [12]. Com estas considerações, pode-se calcular o parâmetro R em função dos parâmetros: atraso (I_d) e o fator correspondente ao codificador (I_e):

$$R = R_o - I_d - I_e \quad (4)$$

III. CENÁRIO EXPERIMENTAL

O seguinte diagrama de blocos, Figura 1, ilustra o cenário experimental utilizado para encontrar o parâmetro de degradação do codificador I_e , empregando as recomendações descritas anteriormente.

Os codificadores de voz que foram avaliados em um cenário sem perdas e sem atrasos foram: G723.1 e iLBC.

Para o codificador G.723.1 foram feitos testes com probabilidades de perdas de pacotes de 0%, 5%, 10% e 20% com atraso ponto a ponto igual a zero.

Este cenário é composto por uma sub-rede (PC1) conectado a uma outra sub-rede (PC3), e estes são interconectados por um roteador GNU/Linux (PC2). Sendo utilizada uma rede com fio para este cenário.

Os softwares aplicativos utilizados nos testes foram:

- PC1, PC3: clientes com softphone MyPhone 0.2b10 [13], analisador de pacotes Wireshark [14] e software para gravação de áudio .wav VRS Recording System [15].

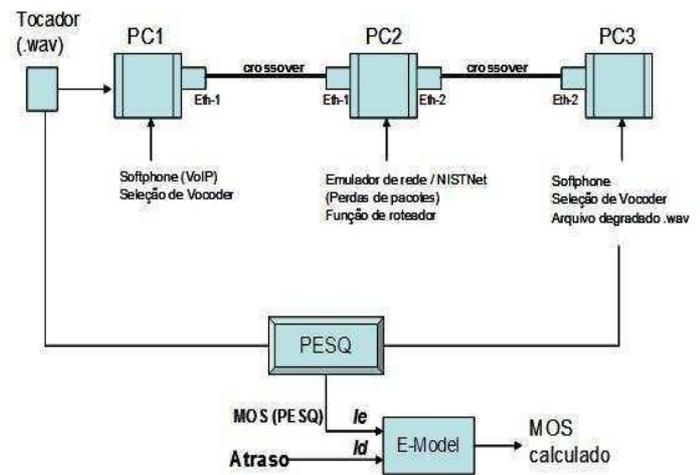


Fig. 1. Diagrama de blocos do cenário de testes utilizando ferramentas open source

- PC2: roteador com o emulador de redes Nistnet [16] utilizado para simular perda de pacotes, atrasos, variações de atrasos e largura de banda e o software ITU-T P.862 utilizado para encontrar o índice MOS da qualidade de voz.

O som transmitido, foi gerado por um reprodutor de arquivo .wav que é conectado a entrada do microfone do PC1 mediante um cabo de áudio. Este arquivo possui uma duração de 8 segundos e foi amostrado a 8 kHz e 16 bits.

A Metodologia seguida nos testes para se obter o valor MOS utilizando a ferramenta PESQ é a seguinte:

- Inicialmente, se inicia uma comunicação entre o PC1 e o PC3 mediante os softphones que cada PC possui instalado e que permite escolher o codificador de voz utilizado para cada chamada VoIP, onde esta é feita de computador a computador, e a sinalização utilizada foi a recomendação H.323.
- Para cada cenário o emulador de rede é configurado com os parâmetros requeridos do teste a executar.
- O reprodutor transmite o áudio (arq-orig.wav) ao PC1 onde este som é gravado (arq-orig2.wav), como a chamada já está ativa, a voz é codificada e transmitida para o PC3 através do PC2.
- Enquanto os dados são transmitidos, o programa Wireshark que está sendo executado no PC1 e no PC3, salva a informação da rede como: as mensagens de sinalização para estabelecimento, manutenção e finalização das chamadas, as mensagens do protocolo RTP, o tamanho médio do pacote (Bytes), número médio de pacotes transferidos por segundo e largura de banda média.
- No PC3, o áudio recebido é gravado em um arquivo (arq-deg.wav). Este arquivo e o arquivo 'arq-orig2' são comparados mediante o programa PESQ, que é executado no PC2, tendo como resultado uma pontuação MOS. Foram feitos dez testes para cada cenário, tomando-se o valor médio como resultado válido de cada cenário. O desvio padrão obtido pelo programa PESQ foi de 0,05 para cada codificador.

É importante mencionar que antes de realizar cada teste, o programa PESQ foi configurado adequadamente. Neste sentido, verificaram-se todos os arquivos de teste que existentes no software PESQ, obtendo-se resultados satisfatórios.

Da mesma forma para os cenários de perdas de pacotes e atrasos, configurados no Nistnet, foram feitos testes para avaliar se estes atrasos e perdas eram reflexos de dados anteriores da rede a fim de se obter um resultado confiável.

IV. RESULTADOS

A. Cenário sem probabilidade de perda de pacotes

Para um cenário ideal: sem perdas, sem atrasos e com taxa de transmissão suficiente, conseguiram-se resultados semelhantes com a maioria da literatura como em [17], [18] e [19].

A Tabela I apresenta as médias dos valores obtidos nos testes.

Segundo [20] estes valores MOS podem ser convertidos em valores MOS-LQO (Mean Opinion Score - Listening Quality

TABELA I
ÍNDICE MOS PARA OS CODIFICADORES G.711, G.723.1 E iLBC EM UM CENÁRIO SEM PROBABILIDADE DE PERDAS DE PACOTES E SEM ATRASOS.

Codificador	Pontuação PESQ
G.711	4.09
G.723.1	3.52
iLBC	3.79

Objective), que é a pontuação do PESQ mais comparável com o índice MOS pela seguinte equação:

Sendo $MOS = x$ e para $x > 1.7$:

$$MOS-LQ = -0.156278x^3 + 1.386609x^2 - 2.504699x + 2.023345 \quad (5)$$

Para $MOS \leq 1.7$:

$$MOS-LQ = 1 \quad (6)$$

Com estes valores e utilizando a equação 6 definida em [9], pode determinar-se o valor de R a partir do índice MOS.

Para $MOS = 1.0$:

$$R = 0 \quad (7)$$

Para $1.0 < MOS < 4.5$:

$$MOS-LQ = 1 + 0.035 \times R + R \times (R - 60) \times 7 \times 10^{-6} \quad (8)$$

Para $MOS \geq 4.5$:

$$R = 100 \quad (9)$$

E para se obter o parâmetro I_e utilizou-se a equação 10, considerando $I_d=0$ e R_o igual ao valor obtido para o codificador da recomendação ITU-T G.711 [21] no cenário de rede de testes apresentado.

$$I_e = 86.5 - R \quad (10)$$

Com estas equações obtiveram-se os valores apresentados na Tabela II:

TABELA II
VALORES DE I_e , PESQ E PESQ-LQO PARA OS CODIFICADORES G.723 E iLBC.

Codificador	PESQ	PESQ-LQO	I_e
G.723.1	3.52	3.53	18
iLBC	3.79	3.88	13

Da Tabela II, pode-se observar que o valor do I_e obtido para o codificador iLBC apresenta um melhor resultado segundo a equação 2, onde um menor valor de degradação (I_e) reflete uma maior qualidade do sinal.

É importante ressaltar que o valor obtido para o $I_e = 18$ do codificador G.723.1 é muito próximo ao valor apresentado pela ITU-T no anexo I da Rec. G.113 que é igual a 19.

Com os valores da Tabela II para o codificador G723.1 obteve-se o gráfico da Figura 2, que apresenta a relação entre o MOSc (MOS conversacional) obtido do valor R do E-Model e os diferentes valores de atraso (0, 50, 100, 150, 200, 250, 300, 350, 400, 450 e 500 ms) para um cenário de perdas de pacotes igual a zero.

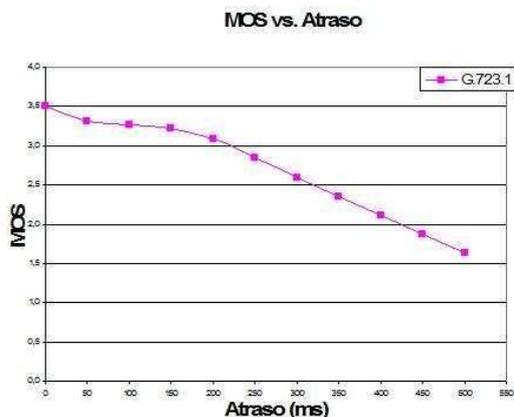


Fig. 2. Cálculo de MOS vs. Atraso obtido utilizando o E-model.

Desta última figura, pode-se constatar que a qualidade de voz se conserva aceitável até 150 ms que é o atraso especificado na Recomendação ITU-T G.114 [22].

B. Cenário com probabilidade de perdas de pacotes

Para um cenário de probabilidade de perdas de pacotes e atrasos iguais a zero, e seguindo a metodologia explicada, mediu-se a qualidade de voz para os codificadores G.723.1 e iLBC.

As probabilidades de perdas consideradas nos testes realizados foram de 5%, 10% e 20%.

Os resultados obtidos são apresentados na Tabela III.

TABELA III

VALORES MOS E I_e PARA OS CODIFICADORES G.723.1 E iLBC EM UM

CENÁRIO COM PERDAS DE PACOTES.

	Codificador G.723.1	Codificador iLBC
Probabilidade de perda de pacotes	Probabilidade de perda de pacotes	Probabilidade de perda de pacotes
	5% - 10% - 20%	5% - 10% - 20%
MOS	2.85 - 2.60 - 2.14	3.26 - 3.02 - 2.61
I_e	37 - 44 - 60	26 - 32 - 44

O codificador iLBC em relação com o codificador G.723.1 teve uma melhor resposta ante perdas de pacotes, o que pode comprovar-se com os novos valores do fator I_e , obtidos para cada um dos codificadores nos diferentes cenários apresentados na Tabela III (Quanto menor I_e melhor resposta à perda de pacotes)

A resposta de um codificador ante diferentes cenários de perda de pacotes se denomina grau de Robustez do codificador e esta definida pela seguinte equação:

$$I_{e-eff} = I_e + (95 - I_e) \times (Ppl / (Ppl/BurstR + Bpl)) \quad (11)$$

Onde:

- I_{e-eff} : Valor de I_e para um determinado cenário de perda de pacotes.
- I_e : Valor de I_e num cenário ideal.
- Ppl: Probabilidade de perda de pacotes.
- BurstR: Relação de rajada.
- Bpl: Fator de Robustez contra perda de pacotes.

E considerando que:

- Lmc = Longitude média das rajadas observadas em uma seqüência de chegada;
- Lma = Longitude média das rajadas previstas na rede em condições de perdas arbitrárias.

Por sua vez o parâmetro BurstR é definido como:

$$BurstR = Lmc/Lma \quad (12)$$

Com perda de pacotes aleatória o valor de Burst será 1 e quando as perdas de pacotes obedecerem às rajadas, ou seja, que são dependentes, o valor de Burst será maior que 1. Para perda de pacotes que sigam o modelo de Markov com dos estados, se tem:

$$BurstR = 1/(p+q) = (Ppl/100) / p \quad (13)$$

Sendo p e q as probabilidades de transição entre pacotes localizados e pacotes de perda.

TABELA IV

VALORES DE ROBUSTEZ CONTRA PERDA DE PACOTES PARA OS CODIFICADORES G.723.1 E iLBC

	G.723.1	iLBC
I_e	18	13
Ppl	5	5
BurstR	1	1
I_{e-eff}	37	26
Bpl	15.26	26.54

A Tabela IV apresenta os valores de Robustez contra perda de pacotes para os codificadores G.723.1 e iLBC.

Observando-se um maior valor de Bpl para o codificador iLBC.

V. CONCLUSÕES

Este trabalho pretende ser útil para projetistas de rede na avaliação do valor de I_e de codificadores novos ou que não tenham parâmetros definidos pela ITU-T, e a partir deste prever o índice de qualidade MOS do codificador, em estudo para diferentes cenários de rede, ou seja, com diferentes probabilidades de perda de pacotes e atrasos tanto em conversações unidirecionais quanto em conversações bidirecionais.

Os resultados obtidos foram bastante aceitáveis considerando que são valores muito próximos aos descritos na documentação da ITU-T, dando um grau de confiabilidade ao método utilizado em estes testes e que podem ser facilmente reproduzíveis já que o cenário de teste foi baseado em *software open source*.

VI. TRABALHOS FUTUROS

Neste trabalho foi descrito uma metodologia para prever o comportamento de um codificador em diferentes cenários de redes; estes resultados podem ajudar na análise da rede de um modo estático e em momentos determinados com as modificações nos diversos elementos da rede, especialmente na eleição da taxa de codificação que só poderá ser feita manualmente. Como o comportamento das redes de dados é variável, a melhor solução é projetar um sistema que controle a qualidade da comunicação de uma forma dinâmica e automática. Considerando isto em um trabalho futuro pretende-se criar um algoritmo para a determinação da taxa de codificação (RDA) que ajude a controlar automaticamente a qualidade do sistema, ou seja, para condições de rede com sobrecarga de tráfego, a taxa a ser escolhida será sempre a menor e para condições de baixa carga da rede a taxa escolhida será a maior.

REFERENCES

- [1] M. Bathia, S. Kalindrini, *Voice over IP Fundamentals*, Second Edition, Cisco Systems, Aug. 2007.
- [2] J. Han, J. Chung, *Study of delay patterns of weighted voice traffic of end-to-end users on the VoIP Networks*, International Journal of Network Management, 2002.
- [3] A. P. Markopoulou, F. A. Tobagi, *Assessment of VoIP Quality over Internet*, in Proc. of IEEE Infocom, vol. 1, (New York, USA), pp. 150-159, June 2002.
- [4] D. Minoli and E. Minoli, *Delivering Voice over IP Networks*, 2nd Edition. John Wiley and Sons, 2002.
- [5] ITU-T Rec. G.107, *The E-model, a computational model for use in transmission planning*, Mar. 2005. <http://www.itu.int/rec/T-REC-G.107/en>.
- [6] ITU-T Rec. P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, <http://www.itu.int/rec/T-REC-P.862/en>.
- [7] ITU-T Rec. P.G.108, *Aplicación del modelo E: Directrices para la planificación*, Set. 1999. <http://www.itu.int/rec/T-REC-G.108/en>.
- [8] ITU-T Rec. G.113, *Degradaciones de la transmisión debido al tratamiento de las señales vocales*, May. 2002. <http://www.itu.int/rec/T-REC-G.113/en>.
- [9] ITU-T Rec. G.723.1, *General Aspects of Digital Transmission Systems - Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s*, May. 2006. <http://www.itu.int/rec/T-REC-G.723.1/en>.
- [10] IETF RFC 3951, *Internet Low Bit Codec Rate*, Dec. 2004. <http://www.ietf.org/rfc/rfc3951.txt>.
- [11] ITU-T Rec. P.800, *Methods for subjective determination of transmission quality*, Aug. 1996. <http://www.itu.int/rec/T-REC-P.800/en>.
- [12] M. Arjona, M. Minami, *Technology and Standards for Low-Bit-Rate Coding Methods*, The Handbook of Computer Networks. 1 ed. Hoboken, N.J.: John Wiley Sons, 2008, v. 2, p. 447-467.
- [13] *Softphone Myphone*, <http://myphone.sourceforge.net/>.
- [14] *Wireshark*, <http://www.wireshark.org/download.html>.
- [15] *VRS - Recording System*, <http://www.nch.com.au/vrs/index.html>.
- [16] *NISTNET*, <http://www-x.antd.nist.gov/nistnet/>.
- [17] M. Jiang, M. Schulzrinne, *Comparisons of fec and codec robustness on voip quality and bandwidth efficiency*, Columbia University, Department of computer Science, USA.
- [18] *Voice over packet, An assessment of voice performance on packet networks*, Nortel Networks. 2002.
- [19] M. Grant, S. Tonissen, *Voice Quality Monitoring for VoIP Networks*, Calyptech. 2003.
- [20] A. W. Rix, *Comparison between Subjective Listening Quality and P.862 PESQ Score*, in Proceedings of Online Workshop Measurement of Speech and Audio Quality in Networks, (Czech Republic), May 2003.
- [21] ITU-T Rec. G.711, *General Aspects of Digital Transmission Systems Terminal Equipments - Pulse Code Modulation (PCM) of Voice Frequencies*, 1972. <http://www.itu.int/rec/T-REC-G.711/en>.
- [22] ITU-T Rec. G.114, *One way transmission time*, May. 2003. <http://www.itu.int/rec/T-REC-G.114/en>.