

Utilização do Modelo E para avaliação da qualidade da fala em sistemas de comunicação baseados em voz sobre IP

Leandro C. G. Lustosa², Leandro S. G. Carvalho¹, Paulo H. de A. Rodrigues²,
Edjair de S. Mota¹

¹Laboratório de Voz sobre IP – Departamento de Ciência da Computação –
Universidade Federal do Amazonas (UFAM)[†]

²Laboratório de Voz sobre IP – Núcleo de Computação Eletrônica – Universidade
Federal do Rio de Janeiro (NCE/UFRJ)[‡]

{leandro,aguilar}@nce.ufrj.br, {galvao,edjair}@pop-am.rnp.br

Abstract. *An objective method for VoIP quality evaluation based on the ITU-T E-model is presented. This method incorporates extensions proposed by Alan Clark, later adopted by ETSI, and it corrects some expressions necessary for calculating internal extended E-model parameters. This revised method was implemented as an analytical evaluation module and incorporated to an existing in-house developed VoIP monitoring tool. Results from experimental measurements demonstrate the adequacy of our suggestions and considerations.*

Resumo. *Este artigo descreve um método de avaliação da qualidade da fala em aplicações Voz sobre IP (VoIP), baseado no Modelo E da ITU-T. É também descrita uma extensão a esse modelo, proposta por Alan Clark e adotada pela ETSI, e algumas correções são apresentadas à mesma. Para comparar o Modelo E estendido e revisado com o Modelo E original, foi implementado um módulo de avaliação analítica de qualidade de fala para ligações VoIP baseado em uma ferramenta de um trabalho anterior. Da comparação dos resultados observa-se o aprimoramento do método de avaliação na abordagem estendida e revisada.*

1. Introdução

Com a popularização de equipamentos e soluções de comunicação baseados em Voz sobre IP (VoIP) nos mais diversos ambientes (corporativo, acadêmico, doméstico, entre outros), surge a necessidade de se medir a qualidade da fala fim-a-fim em tais sistemas.

[†] Parcialmente suportados com recursos do projeto UFAM/INdT.

[‡] Parcialmente suportados com recursos do GT-VOIP, grupo de trabalho mantido pela RNP. Paulo H. de A. Rodrigues é também professor do Departamento de Computação/IM da UFRJ.

Uma das primeiras respostas a tal necessidade foi a utilização da Pontuação de Opinião Média, ou MOS, método definido na Recomendação ITU-T P.800 [16], pelo qual avaliadores ouvintes atribuem uma pontuação de 1 (pobre) a 5 (excelente) à qualidade da fala reproduzida pelo sistema de comunicação em teste. Como expressa diretamente a opinião média dos usuários, o MOS representa um índice de referência para avaliação da qualidade da fala em sistemas de comunicação.

Contudo, como a obtenção da pontuação MOS é um procedimento de difícil reprodução, tem-se optado pelo desenvolvimento e implementação de métodos objetivos para estimá-la em sistemas de comunicação que utilizam a tecnologia de Voz sobre IP [2]. Entre esses métodos objetivos, destaca-se o Modelo E, definido na Recomendação G.107 [9].

O Modelo E implementa um mecanismo baseado na soma de termos que representam distorções na qualidade da fala, tais como atrasos de transmissão, eco, distorções introduzidas pelos equipamentos utilizados, entre outros fatores. O resultado do modelo é o fator escalar R , mapeável para a escala de pontuação MOS.

Apesar de ter sido desenvolvido para a fase de planejamento de sistemas de comunicação de voz [9], o Modelo E tem sido objeto de estudos para aplicá-lo como ferramenta de monitoração da qualidade da fala no cotidiano de redes VoIP [3,4].

Neste artigo, o Modelo E é revisado e as diferentes abordagens sugeridas na literatura para computar as diversas variáveis que compõem o modelo são criticadas. Sugestões para aperfeiçoar estas abordagens são apresentadas e uma ferramenta que implementa estas abordagens é desenvolvida e utilizada para melhor estudá-las, discutí-las e validá-las.

O trabalho está organizado da seguinte forma: na seção 2, abordamos os principais fatores que entram no cômputo do fator R . Na seção 3, discutimos abordagens de implementação realizadas por alguns pesquisadores da área e algumas sugestões são apresentadas. Na seção 4, apresentamos uma ferramenta analítica para avaliação de qualidade de voz em sistemas de comunicação baseados em VoIP e alguns resultados de avaliação de uma ligação VoIP de um cenário real, utilizando as abordagens discutidas no trabalho e, por fim, na seção 5, apresentamos nossas conclusões e sugestões para trabalhos futuros.

2. O Modelo E

O princípio fundamental do Modelo E, definido na Recomendação ITU-T G.107 [9], baseia-se no conceito de que “fatores psicológicos numa escala subjetiva são aditivos” [9], ou seja, cada contribuição ocasionada por um fator de perda na degradação de um sistema de comunicação de voz pode ser computada separadamente, embora isso não implique que tais fatores não estejam correlacionados [4]. O resultado final do cômputo dos fatores de perda é um fator escalar R , que varia de 0 (pior caso) a 100 (excelente).

O fator R pode ser convertido para escala de pontuação MOS através da seguinte expressão de terceiro grau [9]:

$$\begin{aligned} \text{Para } R < 0: \text{ MOS} &= 1 \\ \text{Para } 0 \leq R \leq 100: \text{ MOS} &= 1 + 0,035 R + 7 \cdot 10^{-6} R (R-60) (100-R) \\ \text{Para } R > 100: \text{ MOS} &= 4,5 \end{aligned} \tag{1}$$

Normalmente, o fator R é descrito em categorias de valores, tal como pode ser consultado na Tabela 1. Sistemas cuja qualidade da fala seja avaliada em $R \leq 60$ não são recomendáveis [4], sendo desejável obter $R \geq 70$.

Tabela 1: Categorias de transmissão da fala. Fonte: [4].

Fator R	MOS	Satisfação do usuário
$90 \leq R < 100$	4,34 – 4,50	Muito satisfeitos
$80 \leq R < 90$	4,03 – 4,34	Satisfeitos
$70 \leq R < 80$	3,60 – 4,03	Alguns insatisfeitos
$60 \leq R < 70$	3,10 – 3,60	Muitos insatisfeitos
$0 \leq R < 60$	1,00 – 3,10	Quase todos insatisfeitos

O fator R é obtido pela seguinte fórmula:

$$R = R_o - I_s - I_d - I_e + A, \quad (2)$$

onde R_o representa os efeitos da relação sinal-ruído (SNR); I_s representa as perdas simultâneas ao sinal de voz; I_d representa as perdas associadas ao atraso fim-a-fim; I_e representa as perdas associadas ao equipamento utilizado; e A corresponde ao fator de vantagem, ou fator de expectativa.

Vejam os fatores e grandezas que influenciam cada uma dessas parcelas e quais delas podem ser reduzidas a uma constante.

2.1. Relação sinal-ruído básica, R_o

Na determinação da relação sinal-ruído básica, são computados o ruído acrescido pelos circuitos de transmissão, o ruído ambiente no lado do receptor e no lado do emissor e o teto de ruído correspondente à sensibilidade do sistema auditivo humano. [9] apresenta as expressões matemáticas e os valores padrões para cada um desses ruídos, resultando num valor padrão de 94,77 para R_o .

2.2. Fator de perdas simultâneas, I_s

Entre as perdas que ocorrem mais ou menos simultaneamente ao sinal de voz, estão a queda na qualidade devido a uma conexão de volume demasiadamente alto, perdas causadas pela interferência da própria voz do locutor sobre o fone de ouvido do mesmo *handset* que utiliza para falar e a distorção de quantização causada pela digitalização do sinal de voz. As perdas causadas pelos codecs de baixa taxa de transmissão de bits são computadas separadamente pelo fator de perda I_e . [9] também apresenta as expressões e os valores padrões para as perdas que contribuem no cômputo do fator I_s .

A mais recente versão da Recomendação ITU-T G.107 [9] apresenta uma modificação no cálculo de I_s em relação às versões anteriores. Porém, a atualização não alterou substancialmente o valor padrão de I_s , que reduziu de 1,43 para 1,41.

2.3. Fator de perdas associadas ao atraso, I_d

O fator I_d é determinado pela seguinte expressão [9]:

$$I_d = I_{dte} + I_{dle} + I_{dd}, \text{ onde:} \quad (3)$$

$Idte$ representa as perdas devido ao eco no lado transmissor; $Idle$, as perdas devido ao eco no lado receptor; e Idd as perdas relacionadas ao longo atraso absoluto da voz (mais que 100 ms).

Para o cálculo do fator Id , é necessária a medida de três diferentes tipos de atraso, ilustrados na Figura 1.

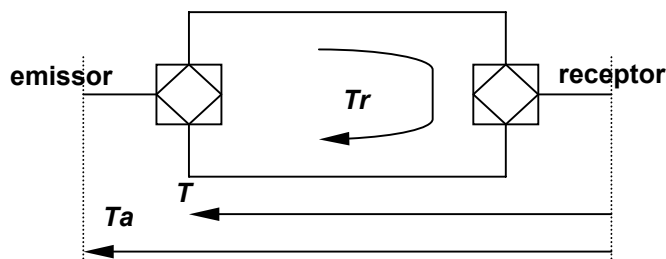


Figura 1 – Representação de atrasos num sistema de comunicação de voz [19].

T_a é o atraso absoluto no sistema sem eco. T é o atraso médio percorrido pelo eco. T_r é o atraso de ida-e-volta no circuito a quatro fios. Como não é possível medir separadamente esses três atrasos [19], usamos a aproximação [4]:

$$T = T_a = T_r/2 \quad (4)$$

Na seção 3 são apresentadas considerações na determinação do fator Id .

2.4. Fator de perdas associadas ao equipamento, I_e

O fator de perda I_e é um meio flexível de se computar as perdas causadas pelos codecs de baixa taxa de transmissão de bits. Seu valor para cada codec depende dos resultados de exaustivos testes de pontuação MOS sob diversas taxas de perdas de pacotes.

O Apêndice I da Recomendação ITU-T G.113 [13] fornece valores provisórios do I_e para alguns codecs (G.723.1-A, G.729-A, GSM-EFR e G.711) sob condições de perdas aleatórias de pacotes e de perdas em rajada. Na seção 3 também serão apresentadas algumas abordagens para o cálculo do fator de perda I_e .

2.5. Fator de vantagem, A

O fator de vantagem A é empregado para definir o grau de tolerância que um usuário espera pela conveniência de utilização de uma determinada tecnologia. Segundo os valores provisórios apresentados em [9], este fator varia de 0 (para telefonia fixa) a 20 (para localidades de difícil acesso que necessitam de enlace de satélite, por exemplo).

Para a telefonia celular é recomendado o uso de A igual a 5 para redes *in-door* ou igual a 10 para redes geográficas. Para VoIP, [19] recomenda zero. No entanto, consideramos que este fator é uma variável cultural e que, para países que possuem serviço de telefonia precário e caro, como o Brasil, este fator pode ser estabelecido em um valor positivo. Para efeitos gerais, consideramos neste trabalho o fator A como zero.

Assim, temos que a equação (2) pode ser reduzida a:

$$R = 93,36 - Id(T_a) - I_e(\text{codec, perda}), \quad (5)$$

onde Id é uma função do atraso fim-a-fim e Ie é, simplificada, uma função que depende do tipo de codec utilizado e da taxa de perda de pacotes. Na seção a seguir veremos detalhadamente como esses dois fatores de perda podem ser determinados.

3. Determinação dos fatores de perda Id e Ie

3.1. Determinação do fator de perda Id

Segundo ITU-T G.107 [9] e versões anteriores, o fator de perda Id é determinado pela equação (3), cujas parcelas são determinadas pelas seguintes expressões:

Parcela correspondente às perdas devido ao eco no lado transmissor ($Idte$):

$$Idte = \left[(Roe - Re) / 2 + \sqrt{(Roe - Re)^2 / 4 + 100} - 1 \right] \cdot (1 - e^{-T}), \text{ onde:} \quad (6)$$

$$Roe = -1,5 \cdot (No - RLR) \quad (7)$$

$$Re = 80 + 2,5 \cdot \left(TELR - 40 \cdot \log_{10} \left(\frac{1 + T/10}{1 + T/150} \right) + 6 \cdot e^{-0,3T^2} - 14 \right) \quad (8)$$

Parcela correspondente às perdas devido ao eco no lado receptor ($Idle$):

$$Idle = (Ro - Rle) / 2 + \sqrt{(Ro - Rle)^2 / 4 + 169}, \text{ onde:} \quad (9)$$

$$Rle = 10,5 \cdot (WELP + 7) \cdot (Tr + 1)^{-0,25} \quad (10)$$

Parcela correspondente às perdas relacionadas ao atraso absoluto da voz (Idd):

Para $Ta \leq 100$ ms: $Idd=0$; Para $Ta > 100$ ms:

$$Idd = 25 \cdot \left[\sqrt[6]{1 + (\log_2(Ta/100))^6} - 3 \cdot \sqrt[6]{1 + ((\log_2(Ta/100))/3)^6} + 2 \right] \quad (11)$$

Na Tabela 2 encontram-se os valores recomendados por [9] para os termos das equações de 6 a 11:

Tabela 2: Valores recomendados para o cálculo de Id .

Grandeza	Significado	Valor padrão segundo [9]
No	Ruído total em dBm0p	-61,18 dBm0p
Ro	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

Uma interpolação pode ser utilizada para aproximar as equações (3), (6), (7), (8), (9) e (10), combinadas com a consideração feita pela equação (4). Cole e Rosenbluth propõem em [4] uma expressão interpolada para Id , dependente apenas de Ta :

$$Id = 0,024 Ta, \quad Ta \leq 177,3 \text{ ms}; \quad (12)$$

$$Id = 0,134 Ta - 18,103, \quad Ta > 177,3 \text{ ms}$$

Em outro trabalho apresentado por Clark [3] é proposta uma expressão que está nitidamente incorreta, pois penaliza o fator Id para valores de $Ta \leq 175$ ms:

$$\begin{aligned}
 Id &= 4 Ta, Ta \leq 175 \text{ ms;} \\
 Id &= 0,111 Ta - 15,444, Ta > 175 \text{ ms}
 \end{aligned}
 \tag{13}$$

Este fato demonstra possivelmente um erro tipográfico. Calculando o ponto de inflexão entre as duas curvas que deveriam ser expressas pela expressão (13) de Clark, concluímos que a equação correta deveria ser expressa por:

$$\begin{aligned}
 Id &= 0,023 Ta, Ta \leq 175 \text{ ms;} \\
 Id &= 0,111 Ta - 15,444, Ta > 175 \text{ ms}
 \end{aligned}
 \tag{14}$$

que apresenta grande semelhança com a expressão (12).

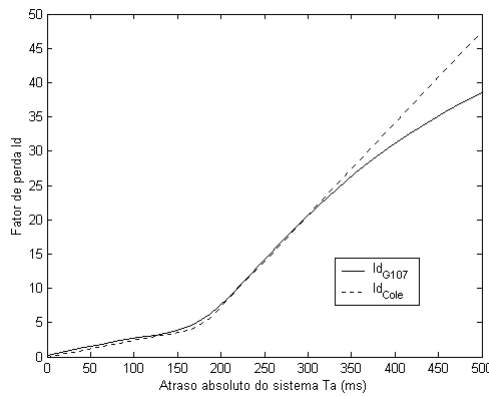


Figura 2 – Comparação entre a aproximação paramétrica sugeridas por [4] e a curva de Id recomendada por [9].

Na Figura 2, temos representadas a curva obtida a partir da Recomendação ITU-T G.107 [9] e a aproximação sugerida por Cole e Rosenbluth em [4]. A abordagem adotada em [4], embora matematicamente muito mais simples, retorna valores suficientemente precisos de Id para $Ta \leq 330$ ms. Para situações onde $Ta > 330$ ms, a aproximação piora progressivamente. Entretanto, para situações de Ta tão elevado, a qualidade da ligação já estará seriamente comprometida, exceto em casos específicos onde um codec de baixa compressão em condições mínimas de perdas e valor elevado de fator de vantagem A é utilizado.

No presente trabalho, utilizamos a abordagem da Recomendação ITU-T G.107 [9] para determinação de Id . Entretanto, para uma monitoração em tempo real, o uso da aproximação feita em [4], devido a sua baixa complexidade, é recomendado.

3.2. Determinação do atraso absoluto (Ta)

Baseado em [14], Ta pode ser obtido pela equação:

$$Ta = T_{codec} + T_{rede} + T_{buffer}
 \tag{15}$$

onde T_{codec} é o atraso inerente ao codec utilizado, T_{rede} é o atraso da rede e T_{buffer} é o atraso adicionado pelo buffer de compensação de *jitter*.

T_{rede} foi considerado simétrico ($RTT/2$), já que para uma medição mais precisa, seria necessário que mantivéssemos os relógios das máquinas de origem e destino

sincronizados, o que exigiria uma infra-estrutura mais elaborada, como por exemplo, a utilização de GPS.

Uma observação importante é que a variação de T_a é eliminada pelo buffer de compensação de *jitter*, já que todos os pacotes são atrasados de modo que possuam um T_a correspondente ao atraso máximo tolerado [14]. Isso é necessário para se assegurar a cadência correta na reprodução da fala no dispositivo receptor. Conseqüentemente, quando um buffer fixo é utilizado, a influência do tempo é mantida constante durante toda a ligação. Entretanto, caso seja utilizado um buffer dinâmico, a influência do tempo pode oscilar no momento em que o tamanho do buffer se altera, o que só deve ocorrer em períodos de silêncio, para não degradar a reprodução da fala.

Dessa forma, usamos:

$$T_a = T_{\text{codec}} + T_{\text{primeiro pacote}} + (T_{\text{buffer total}}/2) \quad (16)$$

onde T_{codec} é o atraso inerente ao codec utilizado, que possui valor fixo conhecido [14], $T_{\text{primeiro pacote}}$ é o T_{rede} referente ao primeiro pacote recebido no período de fala (*talkspurt*) e $T_{\text{buffer total}}$ é o tamanho total do buffer de compensação de *jitter*.

3.3. Determinação do fator de perda I_e

A determinação do fator de perda I_e depende de testes subjetivos (pontuação MOS) executados pelo Grupo de Estudo 12 da ITU-T, cujos resultados são publicados periodicamente no Apêndice I da Recomendação ITU-T G.113 [11,12,13].

Segundo [9], o fator efetivo de perda associada ao equipamento, $I_{e,eff}$, é calculado a partir da expressão:

$$I_{e,eff} = I_e + (95 - I_e) \cdot Ppl / (Ppl + Bpl) \quad (17)$$

onde, I_e é o fator de perda associada ao equipamento a 0% de perda de pacotes, Bpl é o fator de robustez à perda de pacotes e Ppl é a probabilidade de perda de pacotes.

Os valores de I_e para perda de pacotes nula e de Bpl estão tabelados para alguns codecs em [13]. O valor de Ppl deve ser de conhecimento do projetista ou substituído pela informação de perda de pacote obtida a partir da monitoração do sistema de comunicação VoIP em estudo. Para determinar os valores de I_e e Bpl , [13] leva em consideração que as condições de perdas são aleatórias, ou seja, que “a probabilidade de perda de um pacote independe da probabilidade de perda de qualquer outro pacote”. Porém, [13] alerta que essa premissa é, na prática, insatisfatória para redes VoIP.

Além disso, deve-se ressaltar que a determinação do $I_{e,eff}$ não consiste de apenas uma simples consulta de uma tabela, pois alguns fatores estão envolvidos:

- **Porcentagem de perda de pacotes.** Para que o resultado do Modelo E se aproxime da pontuação MOS, essa porcentagem deve ser medida após o buffer de *jitter*, pois pacotes podem ser entregues à aplicação VoIP demasiadamente tarde para que sejam tocados. [13] alerta ainda que “implementações de buffers de *jitter* podem variar consideravelmente entre fabricantes e ainda mesmo entre versões diferentes de um mesmo dispositivo”. Ou seja, a taxa de perda de pacotes depende, além das condições da rede, do algoritmo implementado no buffer de compensação de *jitter*.

- **Número de quadros (*frames*) de voz por pacote de dados.** Após digitalizar e comprimir a voz, o codec entrega quadros à camada de transporte, que pode, ou não, aguardar a chegada de mais quadros antes de formar um pacote. Por isso, afetam mais ou menos trechos de voz, conforme a lógica do aplicativo VoIP utilizado. No entanto, as tabelas constantes em [13] apresentam os valores de I_e e B_{pl} para tamanhos fixos de quadros, conforme o codec consultado.

- **Uso de algoritmos compensadores de perda (PLC).** O uso de PLC torna o codec mais robusto a perdas de pacotes. Dessa forma, o projetista deve tomar conhecimento de qual mecanismo de compensação de perda o aplicativo VoIP utiliza, pois em [13] os valores de I_e e B_{pl} são dependentes do tipo de PLC empregado.

- **Uso de detetores de atividade de voz (VAD).** A detecção da atividade de voz no aplicativo VoIP não é imediata, ou seja, a fala está presente antes do início da execução da função de detetor e pode ocorrer o corte nas primeiras sílabas da locução. Esse efeito é denominado *clipping* e seu impacto sobre a qualidade da voz varia de um codec a outro e influencia no valor de I_e , mesmo em condições de perda de pacote nula.

Uma importante contribuição foi apresentada por Clark em [3] para o cálculo do fator R , tendo sido adotada pelo grupo TIPHON da ETSI na especificação técnica ETSI TS 101 329-5 v1.1.1 [6] e revisão [7]. Trata-se de uma versão estendida do Modelo E apresentado na Recomendação ITU-T G.107 e que estabelece alguns novos conceitos:

- **Qualidade instantânea.** É aquela medida ou calculada utilizando-se a perda de pacotes ou outros fatores de degradação.

- **Qualidade percebida.** É aquela percebida pelo usuário em algum instante durante a locução.

- **Comportamento da perda de pacotes alternantes.** Normalmente, as perdas de pacotes nas redes de dados IP ocorrem em períodos de rajadas (*burst*) que se alternam com períodos de perdas isoladas (*gap*) [3]. Mecanismos de compensação de perdas (PLC) conseguem amenizar tais distorções em períodos de poucas perdas, mas não quando estas ocorrem em rajadas.

- **Memória recente.** Durante pesquisas conduzidas para determinação de pontuações MOS, notou-se que a qualidade percebida de uma chamada telefônica varia de acordo com a localização, no tempo, das perdas. Dessa forma, distorções que ocorrem ao final de uma chamada têm um efeito psicológico negativo maior sobre os avaliadores do que distorções que acontecem no início de uma chamada [3,7].

Segundo [3], se a qualidade instantânea muda de “bom” para “ruim” em algum momento durante a chamada, o usuário não irá instantaneamente perceber tal mudança, mas aos poucos ficará incomodado com a degradação da qualidade da fala. Experimentos descritos em [3] mostram esse efeito, aproximado por uma curva exponencial de constante de tempo de 5 s na transição de “bom” para “ruim” e de 15 s na transição de “ruim” para “bom”.

Para representar o comportamento alternante da perda de pacotes, [3] utiliza uma cadeia de Markov com 4 estados, os quais “representam as condições de recebimento e de perda de pacotes, nas condições de perda em rajada e de perdas isoladas”. O valor de I_e [13] pode ser obtido para a condição de perdas em rajada (I_{eb}) e

para a condição de perdas isoladas (I_{eg}). Ainda de acordo com [3], considere-se as seguintes grandezas:

I_1 – nível de qualidade na mudança de condição de I_{eb} para I_{eg}

I_2 – nível de qualidade na mudança de condição de I_{eg} para I_{eb}

b – duração do estado de perdas em rajadas (em segundos)

g – duração do estado de perdas isoladas (em segundos)

Elas se relacionam da seguinte forma (Figura 3):

$$I_1 = I_{eb} - (I_{eg} - I_2) e^{-b/t_1} \quad (18)$$

$$I_2 = I_{eg} + (I_1 - I_{eg}) e^{-g/t_2} \quad (19)$$

onde tipicamente $t_1=5$ s e $t_2=15$ s. Clark [3] integra as expressões de I_1 e I_2 para obter um I_e médio, que corresponderia ao $I_e(perda)$ da equação (5):

$$I_e(av) = (bI_{eb} + gI_{eg} - t_1(I_{eb} - I_2)(1 - e^{-b/t_1}) + t_2(I_1 - I_{eg})(1 - e^{-g/t_2})) / (b + g) \quad (20)$$

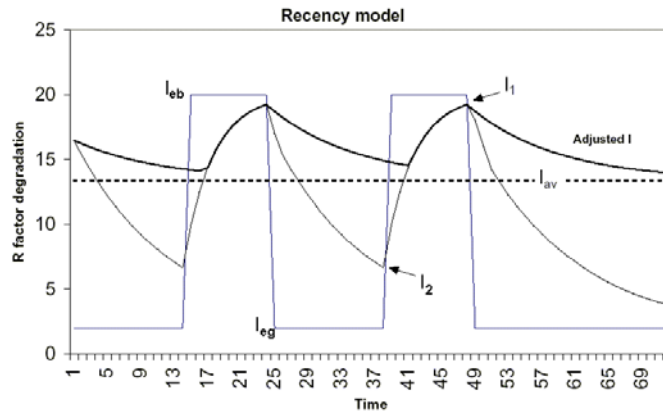


Figura 3 - Representação do efeito da memória recente [5].

De acordo com [3], o efeito da memória recente para determinação de $I_e(av)$ (20), pode ser modelado considerando que a qualidade percebida decai exponencialmente até I_{eb} , com constante de tempo t_1 , na transição entre um período de perdas isoladas para perdas em rajada, e aumenta exponencialmente até I_{eg} , com constante de tempo t_2 , na transição entre um período de perdas em rajada para um período de perdas isoladas, a partir de um valor de “saída” I_1 e I_2 , respectivamente. Já para a determinação da qualidade final da ligação, I_e (fim da chamada) (21), I_1 representa o valor de “saída” a partir da última perda significativa de pacotes em rajada, que aumenta exponencialmente, com constante de tempo t_3 , de valor típico de 30 s. y representa o tempo transcorrido desde a última perda em rajada e k é uma constante de valor nominal de 0,7.

$$I_e(\text{fim da chamada}) = I_e(av) + (k(I_1 - I_e(av))) e^{-y/t_3} \quad (21)$$

Todavia, refazendo os passos descritos em Clark [3], notamos uma incorreção na equação (18), que é desfeita nas equações (20) e (21), mas que pode causar problemas na determinação de $I_e(av)$ e de $I_e(\text{fim da chamada})$. Considerando-se que tal

incorreção foi transportada para a Especificação Técnica da ETSI TS 101 329-5 [6] e revisões posteriores [7,8], sugerimos que a equação (18) seja alterada para a expressão:

$$I_1 = I_{eb} - (I_{eb} - I_2) e^{-b/t_1} \quad (22)$$

4. Implementação dos Modelos

O Laboratório de Voz sobre IP da Universidade Federal do Rio de Janeiro desenvolveu uma infra-estrutura [1] de geração, coleta e monitoramento da qualidade de ligações VoIP baseada na biblioteca OpenH323 [17] capaz de extrair, por meio de monitoração ativa, dados estatísticos referentes ao nível de transporte de ligações e apresentar graficamente informações de perdas de pacotes, *jitter* e RTT ao longo de uma ligação.

Com o intuito de avaliar as abordagens do cômputo do Modelo E, estendemos essa infra-estrutura, implementando um módulo de avaliação analítica de qualidade de voz capaz de extrair, além das informações referentes ao nível de transporte, informações referentes ao nível de aplicação, tais como descarte de pacotes no buffer de compensação de *jitter* e atraso total fim-a-fim (T_a). Este módulo também oferece: análise da distribuição de descartes de pacotes no buffer e de perdas na rede, utilizando o modelo baseado em cadeia de Markov [3] capaz de distinguir períodos de perdas em rajadas de períodos de perdas isoladas; utilização de buffer de compensação de *jitter* com algoritmo adaptativo para ajuste dinâmico às condições da rede; e cômputo do Modelo E segundo as abordagens comentadas neste trabalho.

A infra-estrutura utiliza o H.323 [15] como protocolo de sinalização de chamadas e possui sua arquitetura formada por dois programas: o *openam* (ou Open Answering Machine), um tipo de secretária eletrônica VoIP, e o programa *caller* (ou Call Generator), uma modificação do *openam* desenvolvida por [10]. O programa *caller* realiza ligações para o programa *openam*, enviando um arquivo pré-gravado de áudio durante toda a ligação. O *openam*, por sua vez, armazena o arquivo de áudio recebido, que poderia eventualmente ser submetido a outros métodos objetivos de avaliação, como o *Perceptual Evaluation of Speech Quality* – PESQ e o *Perceptual Analysis Measurement System* – PAMS [2], e registra em um log que possui, além de toda sinalização H.323, todos os eventos de envio e recepção de fluxos de mídia e relatórios RTCP e informações referentes à camada de aplicação, como por exemplo, eventos relacionados ao buffer de compensação de *jitter*.

Utilizando scripts desenvolvidos em Perl [18], o módulo é capaz de extrair dos arquivos de log das ligações as informações necessárias para computar o Modelo E e posteriormente apresentar graficamente os resultados obtidos.

4.1 Avaliação dos modelos

Com o objetivo de avaliar as diferentes abordagens do cômputo de Modelo E, utilizamos o módulo de avaliação para extrair a qualidade de várias ligações VoIP entre a Universidade Federal do Rio de Janeiro (UFRJ) e a Universidade Federal do Amazonas (UFAM), interconectadas através do *backbone* da Rede Nacional de Pesquisa (RNP2). Para o experimento, utilizamos o codec G.723.1 de 6,3 kbps + VAD com quadro de tamanho de 30 ms.

Dentre as ligações originadas, selecionamos uma que possuísse as características desejadas para observarmos o comportamento de perdas de pacotes, transição entre estados de perdas em rajada e perdas isoladas, efeito de memória recente, entre outros. Os resultados da avaliação da qualidade dessa ligação são apresentados a seguir.

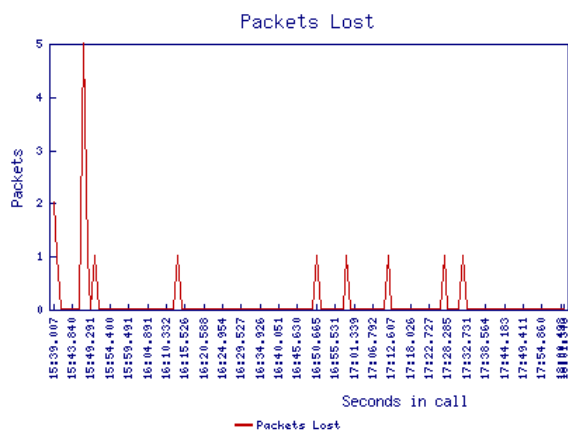


Figura 4 – Perda de pacotes na rede.

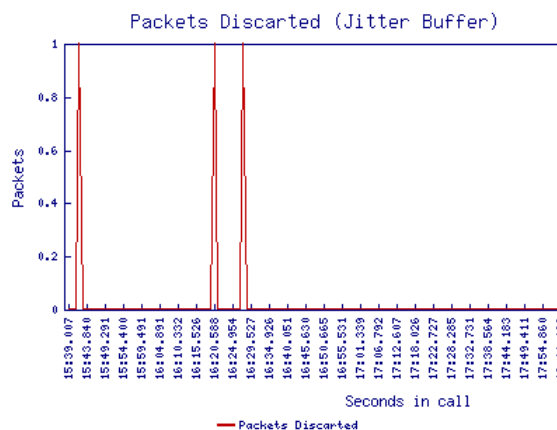


Figura 5 – Descarte de pacotes no buffer de compensação de *jitter*.

A Figura 4 ilustra a perda de pacotes na rede e a Figura 5 ilustra o descarte de pacotes no buffer de compensação de *jitter* ao longo da ligação selecionada. É importante notar que a análise da distribuição de perdas não distingue perdas de pacotes na rede de descartes de pacotes no buffer. Logo, é a soma destes dois fatores que influenciará na determinação do MOS da ligação.

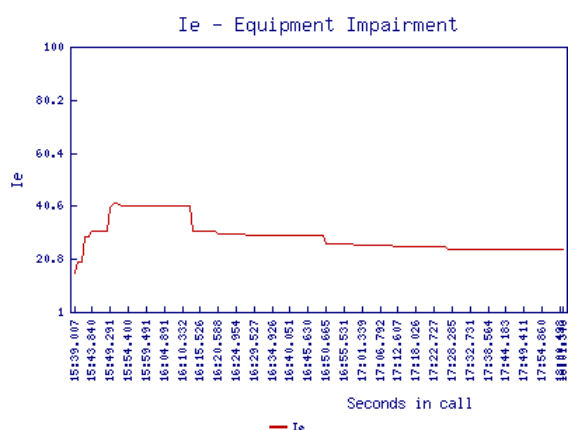


Figura 6 – I_e (uso da equação (22)).

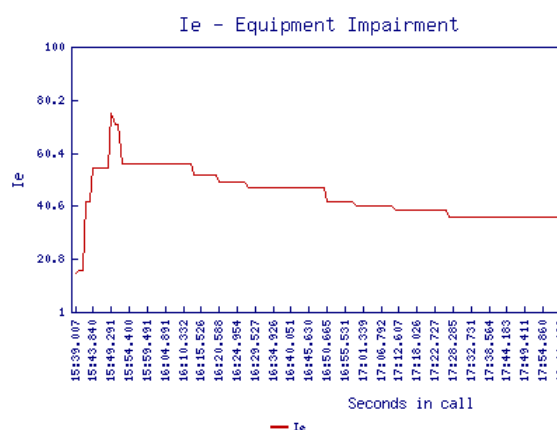


Figura 7 – I_e (abordagem de Clark [3]).

A Figura 6 ilustra a variação do I_e ao longo da ligação. Os valores de I_e ilustrados nessa figura foram obtidos usando a equação (22), que substitui a equação (18) da abordagem de Clark [3] segundo nossas considerações (seção 3.3). A Figura 7 ilustra a variação de I_e ao longo da ligação segundo a abordagem original de Clark [3], que consideramos incorreta. Verificamos que Clark obtém uma avaliação da qualidade da voz mais pessimista, já que o valor de I_e aumenta mais bruscamente após um período de perdas em rajada e decresce mais lentamente na transição de um período de perdas em rajada para um período de perdas isoladas.

A Figura 8 ilustra a variação do I_d ao longo da ligação. Seus valores expressos na curva foram determinados segundo a abordagem em ITU-T G.107 [9].

É importante comentar que há uma ligeira queda da influência do atraso ao longo de toda ligação (Figura 8). Isso acontece porque o tamanho inicial do buffer de compensação de *jitter* é superestimado (em 116 ms). A aplicação percebe que esse valor é superior ao valor ideal e vai diminuindo-o progressivamente, o que conseqüentemente, reduz a influência do atraso na degradação da fala.

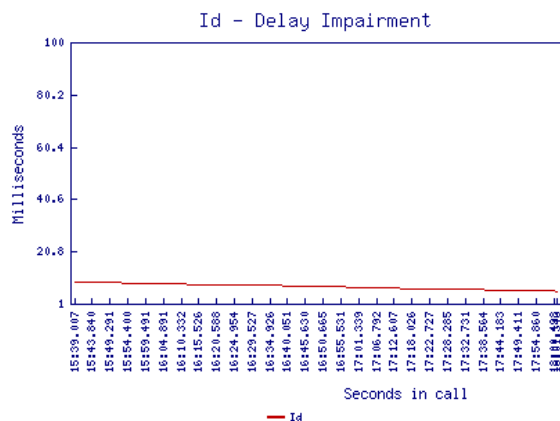


Figura 8 – I_d (ITU-T G.107 [9]).

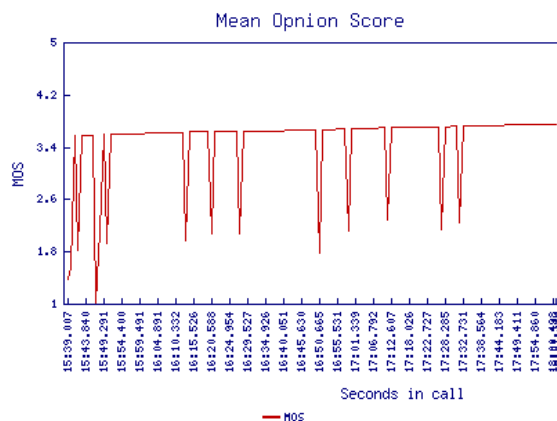


Figura 9 – MOS (ITU-T G.107 [9]).

A Figura 9 ilustra a variação do MOS computado pelo Modelo E segundo a abordagem em ITU-T G.107 [9]. Essa abordagem não considera o efeito de memória recente e não faz distinção entre os períodos de perdas em rajadas e perdas isoladas. Assim, cada ponto da curva ilustrada nessa figura representa um valor instantâneo de MOS, que espera-se que sejam menos fiéis à percepção humana que o resultado apresentado na Figura 10.

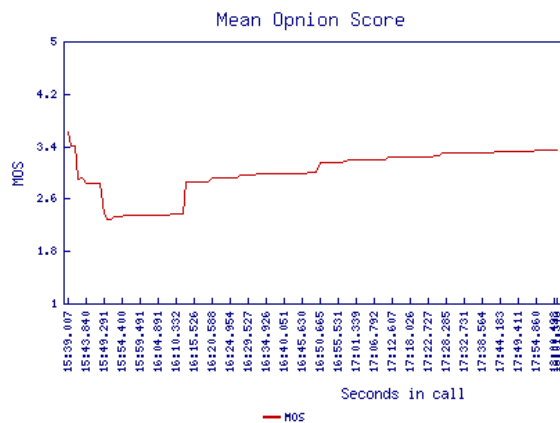


Figura 10 – Variação do MOS ao longo da ligação segundo nossas considerações (equação (22)).

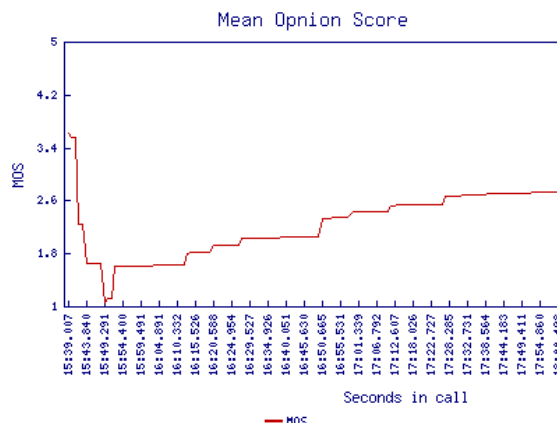


Figura 11 – Variação do MOS ao longo da ligação obtido segundo a abordagem de Clark [3].

A Figura 10 ilustra a variação do MOS computado segundo nossas considerações, ou seja, combinando os resultados de influência de I_e e I_d , apresentados na Figura 6 e na Figura 8, respectivamente. A curva ilustrada nessa figura representa a

qualidade perceptível ao usuário ao longo da ligação. A avaliação final da ligação, $Ie(\text{fim da chamada})$ (21), inferida pelo Modelo E, segundo esta abordagem, foi um MOS de 3,28.

A Figura 11 ilustra a variação do MOS computado segundo a abordagem de Clark [3], ou seja, utilizando a equação incorreta (18) na determinação dos valores de Ie (apresentados na Figura 7). A avaliação final da ligação, $Ie(\text{fim da chamada})$ (21), neste caso, foi um MOS de 2,66, bastante inferior a 3,28.

É notado que o erro introduzido pelo uso da equação (18) resulta em uma relevante discrepância na determinação da qualidade de voz em relação ao resultado obtido segundo nossas considerações.

5. Considerações finais e trabalhos futuros

Neste trabalho, analisamos o Modelo E [9], um método de avaliação da qualidade da fala em aplicações Voz sobre IP (VoIP), e uma extensão a esse modelo, formulada por Clark [3] e adotada pela ETSI [6,7,8]. Apresentamos algumas correções à forma estendida do Modelo E definida em Clark [3], formulando uma proposta revisada para o mesmo. Para validar essa revisão, realizamos ligações VoIP entre a UFRJ e a UFAM, através da RNP2, entre as quais selecionou-se uma cujas características de perdas de pacotes, transição entre estados de perdas em rajada e perdas isoladas, efeito de memória recente, entre outros, fossem mais nítidas para observação. A implementação de um módulo de avaliação analítica de qualidade de voz possibilitou a comparação entre as diferentes abordagens de cômputo do modelo, e os resultados apresentados confirmaram as revisões que fizemos sobre o trabalho apresentado por Clark [3] e se mostraram mais próximos aos resultados de um teste subjetivo MOS que aqueles obtidos pela aplicação do Modelo E original [9].

O desenvolvimento do módulo analítico de qualidade de voz possibilitará a realização de um leque de trabalhos futuros, como por exemplo, a validação da eficácia de mecanismos de priorização de tráfego de voz no *backbone* da RNP2 e a determinação da capacidade de ligações VoIP em redes Wireless 802.11, infra-estruturadas e *ad-hoc*, utilizando variados mecanismos de provisão de qualidade de serviço (QoS) [21,22].

Bibliografia

- [1] MARCONDES, C.A.C., RODRIGUES, P.H.A., DAVID, F., COSTA, J.C.P. *Ambiente para Simulação e Monitoração de Ligações Telefônicas IP*, in: Anais do XXI Simpósio Brasileiro de Redes de Computadores. Natal, maio 2003.
- [2] CARVALHO, Leandro S. G., MOTA, Edjair S. & QUEIROZ, Juliana M. *Análise comparativa de padrões para medida de qualidade de voz*, in: Anais do SUCESU'2003. Salvador, abr. 2003.
- [3] CLARK, Alan D. *Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality*. IP – Telephony Workshop. [USA], 2001.
- [4] COLE, Robert G. & ROSENBLUTH, J. H. *Voice over IP performance monitoring*. ACM SIGCOMM Computer Communication Review, vol.31, issue 2. abr. 2001.

- [5] Temporary Document 98. *Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON); Proposal for Passive Monitoring Methodology of Voice over IP*, [2001].
- [6] ETSI TS 101 329-5 v1.1.1. *Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 5: Quality of Service (QoS) measurement methodologies*. France, nov. 2000.
- [7] ETSI TS 101 329-5 v1.1.2. *Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 5: Quality of Service (QoS) measurement methodologies*. France, jan. 2002.
- [8] ETSI TS 102 024-5 v4.1.1. *Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 4; End-to-end Quality of Service in TIPHON systems; Part 5: Quality of Service (QoS) measurement methodologies*. France, set. 2003.
- [9] ITU-T Recommendation G.107. *The E-Model, a computational model for use in transmission planning*. Genève, mar 2003.
- [10] STOYCHEV, Milen. <http://kgb.cnsys.bg/voice/>. Acesso em nov. 2003.
- [11] ITU-T Recommendation G.113 Appendix I. *Provisional planning values for the equipment impairment factor Ie*. Genève, set. 1999.
- [12] ITU-T Recommendation G.113 Appendix I. *Provisional planning values for the equipment impairment factor Ie*. Genève, out. 2001.
- [13] ITU-T Recommendation G.113 Appendix I. *Provisional planning values for the equipment impairment factor Ie and packet-loss robustness factor Bpl*. Genève, maio 2002.
- [14] ITU-T Recommendation G.114. *One-way transmission time*. Genève, maio 2000.
- [15] ITU-T Recommendation H.323. *Packet-Based Multimedia Communications Systems*. Genève, nov 2000.
- [16] ITU-T Recommendation P.800. *Methods for subjective determination of transmission quality*. Genève, ago. 1996.
- [17] OPENH323 PROJECT. Disponível em <http://www.openh323.org/>. Acesso em nov. 2003.
- [18] The Perl Foundation. *The Perl Directory at Perl.org*. Disponível em <http://www.perl.org/>. Acesso em nov. 2003.
- [19] Psytechnics Group. *Estimating E-model Id within a VoIP network*. Psytechnics Technical note. United Kingdom, 2002.
- [20] RUBINSTEIN, M. G. & REZENDE, J. F. *Qualidade de Serviço em Redes 802.11*, XX Simpósio Brasileiro de Redes de Computadores (SBRC2002), Natal. maio 2002.
- [21] RUBINSTEIN, M. G. & REZENDE, J. F. *Quality of Service in Ad Hoc 802.11 Networks*, Journal of the Brazilian Computer Society, vol. 9, no. 1, pp. 32-43, ISSN 0104-6500, Brazilian Computer Society. jul 2003.