



**Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica e de  
Computação**

**Edgard Luciano Oliveira da Silva**

**Medidas de Qualidade de Experiência baseada no  
Modelo-E durante uma chamada VoIP**

**CAMPINAS  
2016**



Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica e de  
Computação

**Edgard Luciano Oliveira da Silva**

**Medidas de Qualidade de Experiência baseada no Modelo-E durante  
uma chamada VoIP**

Tese de doutorado apresentada a Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática (AG).

**Orientador: Prof. Dr. Yuzo Iano**

Este exemplar corresponde à versão final da Tese defendida por Edgard Luciano Oliveira da Silva e orientada pelo Prof. Dr. Yuzo Iano.

---

Orientador

CAMPINAS  
2016

**Agência(s) de fomento e nº(s) de processo(s):** Não se aplica.

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Elizangela Aparecida dos Santos Souza - CRB 8/8098

Si38m Silva, Edgard Luciano Oliveira da, 1967-  
Medidas de qualidade de experiência baseada no Modelo-E durante uma chamada VoIP / Edgard Luciano Oliveira da Silva. – Campinas, SP : [s.n.], 2016.

Orientador: Yuzo Iano.  
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Telefonia pela Internet. 2. Codificador de voz - Avaliação. 3. Qualidade de serviço (Redes de computadores. 4. Sistemas de processamento da fala. I. Iano, Yuzo, 1950-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

#### Informações para Biblioteca Digital

**Título em outro idioma:** Measures of quality of experience based on the E-model during a VoIP call

**Palavras-chave em inglês:**

Internet telephony

Vocoder - Rating

Quality of service (Computer networks)

Speech processing systems

**Área de concentração:** Telecomunicações e Telemática

**Titulação:** Doutor em Engenharia Elétrica

**Banca examinadora:**

Yuzo Iano [Orientador]

Adão Boava

Vicente Idalberto Becerra Sablón

Rangel Arthur

Luiz César Martini

**Data de defesa:** 30-05-2016

**Programa de Pós-Graduação:** Engenharia Elétrica

## COMISSÃO JULGADORA - TESE DE DOUTORADO

**Candidato:** Edgard Luciano Oliveira da Silva **RA:** 945020

**Data da Defesa:** 30 de maio de 2016.

**Ttulo da Tese:** Medidas de Qualidade de Experiência baseada no Modelo-E durante uma chamada VoIP.

Prof. Dr. Yuzo Iano (Presidente, FEEC/UNICAMP)  
Prof. Dr. Adão Boava (/UFSC)  
Prof. Dr. Vicente Idalberto Becerra Sablón (/UNIMEP)  
Prof. Dr. Rangel Arthur (FT/UNICAMP)  
Prof. Dr. Luiz César Martini (FEEC/UNICAMP)

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no processo de vida acadêmica do aluno.

# Dedicatória

*Aos meus filhos  
Gabriela, Ana Carolina,  
Luís Felipe e Pedro Luciano*

*Aos meus pais  
Alair e Rosa Maria*

*Aos meus sogros  
Duroal (in memoriam) e Honória*

*A minha esposa  
Aracélis*

*[...] O que podia ser,  
sempre vai ser melhor que é.  
As possibilidades  
vão perseguir o que a gente escolher. [...]*

*Topaz*

# Agradecimentos

Ao meu orientador Yuzo Iano pela paciência, conselhos e ajuda durante a minha longa jornada do doutorado e pela oportunidade de chegar até Campinas, onde finalmente pude encontrar o meu verdadeiro caminho.

Aos meus colegas mais próximos de Laboratório GRM Felipe Santos, Gabriel Montenegro, Anderson Rodrigues, Bruno Auzier, Diego Soares e Andresa Helena pela infinita ajuda nas simulações, escrita de rotinas, apoio e incentivo a minha pesquisa e pelos litros de café consumidos durante todo o trabalho. Ao Hermes José Loschi, do Laboratório de Comunicações Visuais (LFP) do Departamento de Comunicações (DECOM), pela enorme ajuda nas questões burocráticas e nas infinitas conversas e conselhos. A todos, meu muito obrigado.

A Celi Aparecida Pavanatti (DECOM-UNICAMP) e Noêmia Benatti (PPG-UNICAMP) por todos esses anos de amizade sempre acreditando que este dia chegaria. A Elienai Nogueira (ICOMP-UFAM) sempre prestativa e amiga e conselheira.

Aos professores Leandro Galvão e Edjair Mota do Grupo de Redes e Multimídia (GRM) do Instituto de Computação (ICOMP) da Universidade Federal do Amazonas (UFAM) pelas valiosas discussões e contribuições para este trabalho.

Ao Laboratório de Redes e Multimídia (GRM) do Instituto de Computação (ICOMP) da Universidade Federal do Amazonas (UFAM) pelo espaço e equipamentos cedidos para a realização parcial dos experimentos.

Agradeço aos professores do Núcleo de Computação (NuComp) da Escola Superior de Tecnologia (EST) da Universidade do Estado do Amazonas (UEA), que sempre me auxiliaram, apoiaram e incentivaram a terminar esse trabalho.

Gostaria de agradecer também ao Departamento de Comunicações da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas (UNICAMP) pela ótima estrutura que oferece aos estudantes e pesquisadores e CAPES pelo apoio financeiro no desenvolvimento deste trabalho.

Ao grupo de Pesquisa Aplicada em Processamento Digital de Sinais (PAPDS) do CPqD Telebrás, na figura dos Engenheiros e Pesquisadores, Guilherme Baird, Antonio Claudio França Pessoa (*in Memoriam*), Fernando Runstein, Edson José Nagle, Ituo pelos primeiros passos na área de avaliação de Qualidade de Sinais da Fala ainda no século passado.

Aos Professores Doutores Patrícia Crosseti (CEFET-RJ), Agnelo Marotta Cassula (UNESP- Faculdade de Eng. Guaratinguetá), Classius Ferreira da Silva (UNIFESP) e Hani C. Yehia (UFMG-CEFALA) pelo apoio, conselhos e incentivo a não desistir deste trabalho.

À minha família, minha esposa e filhos pelo apoio durante todo o período de dedicação.

À Universidade do Estado do Amazonas (UEA), pelo suporte e incentivo.

Enfim, a todos que de alguma forma contribuíram com o meu progresso como aluno e como Ser.

# Resumo

A tecnologia de voz sobre protocolo de internet (voz sobre IP) ou simplesmente, VoIP (*Voice Over Internet Protocol*) ganha novos usuários decorrente do cenário atual de mercado de convergência de redes de dados e telecomunicações. Entretanto, o sucesso desta tecnologia depende fortemente da qualidade do sinal da fala percebida pelo usuário, uma medida subjetiva, também conhecida como Qualidade de Experiência (*Quality of Experience - QoE*). Esta Tese propõe e valida um mecanismo de observação de variáveis de ambiente durante uma comunicação VoIP tendo como saída um número que representa a QoE não referenciada (sem o sinal transmitido, somente o sinal recebido) ao utilizar o Modelo-E com o propósito de selecionar os melhores parâmetros disponíveis de configuração que afetam o fluxo de voz e ao mesmo tempo obter a melhor qualidade de chamada possível dentro de um determinado cenário de rede. O resultado obtido fez a ligação entre uma medida objetiva, o parâmetro R gerado pelo Modelo-E, e uma medida subjetiva estimada, o MOS (*Mean Opinion Score*), durante uma chamada VoIP, não se limitando a medida em si. Todavia, cobriu-se todo o cenário para a medição e a comparação com sistemas padrões de medição de qualidade, formando uma base de conhecimento com os resultados obtidos. O método utilizado de estimativa de qualidade da fala foi comparado em diferentes codecs de voz padrão ITU-T (PCMU, GSM, G.723, G.729, G.726-32), testados em uma topologia de rede que sofreu distorções, como diferentes situações de perdas de dados (0,0%, 1,0%, 2,0%, 2,5%, 3,0%, 5,0%, 7,5%, 10,0%). Uma análise de regressão foi utilizada para permitir uma melhor compreensão do impacto das condições de rede e codecs sobre a QoE do serviço VoIP medido. Foi utilizado um suíte de testes padronizados para medição da QoE nos arquivos de voz recebidos e transmitidos durante os testes baseados em testes referenciados (com os sinais transmitidos e recebidos) nos padrões ITU-T P.863 (*Perceptual Objective Listening Quality Assessment - POLQA*) e ITU-T P.862 (*Perceptual Evaluation of Speech Quality - PESQ*) e os resultados foram comparados com os obtidos pelo método não referenciado proposto para medida de QoE. Para os resultados do codec testado foi aplicado o método de regressão linear sendo a variável independente as medidas de QoE obtidas pelo método proposto e a variável dependente foram os resultados obtidos pelos algoritmos PESQ e POLQA. Para todos os codecs testados, o Coeficiente de Determinação ( $R^2$ ) entre o método proposto e os resultados obtidos pelo algoritmo PESQ foram superiores a 0,90 indicando uma forte correlação linear. Já entre o método proposto e o algoritmo POLQA, para os codecs PCMU, GSM e G.723, os resultados de  $R^2$  foram superiores a 0,973, indicando uma correlação muito forte.  $R^2$  para o codec G.726-32 foi de 0,88 indicando uma correlação forte. Já para o codec G.729,  $R^2$  ficou em 0,67 indicando que o modelo linear pode não ser o mais adequado para explicar a relação entre os resultados do método proposto e os valores obtidos pelo algoritmo POLQA.

**Palavras-chaves:** qualidade de experiência; QoE; voice over internet protocol; VoIP; métricas de qualidade de voz; ITU-T Rec. P.862; PESQ; ITU-T Rec. P.863; POLQA.

# Abstract

The technology of Voice over Internet Protocol (Voice over IP or simply VoIP) is present in our personal and professional lives. The number of VoIP users increases day after day due to the current scenario of convergence of data and telecommunications networks. However, the success of this technology depends on the speech signal quality as perceived by the user, a subjective measure as function of the user's point of view, also known as the Quality of Experience (QoE). This thesis proposes and validates an environment variable observation mechanism during a VoIP communication having as output a number that represents the QoE not referenced (without the transmitted signal, only the received signal) of the call, using the E-Model, in order to select the best available parameter settings that affect voice flow of the current VoIP call and at the same time gets the best call quality as possible within a given network scenario. The result relates an objective measurement, the R parameter generated by the E-Model, to the estimated subjective measurement, MOS, during a VoIP call, not limited to the measurement itself. However, it covered the whole scenario for measurement and comparison with quality measurement standards systems, forming a knowledge base with the results. The method of speech quality estimation was compared in different standard voice codec's ITU-T (PCMU, GSM, G.723, G.729, G.726-32) tested in a network topology that has suffered distortions, as different situations of data loss (0.0%, 1.0%, 2.0%, 2.5%, 3.0%, 5.0%, 7.5%, 10.0%). A regression analysis was used to allow a better understanding of the impact of network conditions and codec's on the VoIP service QoE measured. In this thesis it was used a suite of standardized tests for measuring QoE in voice files received and transmitted during testing based on referenced tests (with the transmitted and received signals) in the ITU-T P.863 standard (Perceptual Objective Listening Quality Assessment - POLQA) and ITU-T P.862 (Perceptual Evaluation of Speech Quality - PESQ) and the results were compared with those obtained by the no referenced measure method of QoE proposed in this thesis. The linear regression was applied in order to analyze the results of the tested codec. The independent variable was the QoE measurements obtained by the proposed method to measure QoE and the dependent variable were the results of PESQ and POLQA algorithms. For all tested codec's, the Coefficient of Determination ( $R^2$ ) between the proposed method and the results of the PESQ algorithm was higher than 0.90 indicates a strong linear correlation. For PCMU, G.723, GSM codec's  $R^2$  was greater than 0.973 indicating a strong correlation between the results of proposed method and the results of POLQA algorithm.  $R^2$  for G.726-32 codec was 0.88 that indicates a high correlation. For G.729 codec,  $R^2$  was 0.67 that indicates the linear model may not be the most appropriate to explain the relationship between the results of the proposed method and values obtained by POLQA algorithm.

**Keywords:** quality of experience; QoE; voice over internet protocol; VoIP; speech quality metrics; ITU-T Rec. P.862; PESQ; ITU-T Rec. P.863; POLQA; MOS; ITU-T objective models; speech quality degradation; test speech signal; Objective Speech Quality.

# Trabalhos Publicados

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. *Mean Opinion Score Measurements Based on E Model During a VoIP call*. In: Telecommunications, 2015. AICT 2015: The Eleventh Advanced International Conference on Telecommunications. Bruxelas, Bélgica: IARIA, 2015. p. 132-135. ISSN:2308-4030, ISBN:978-1-61208-411-4.

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. *Medidas de Mean Opinion Score baseada no Modelo-E durante uma chamada VoIP*. CACSI 2015: I Congresso Amazônico de Computação e Sistemas Inteligentes -. Manaus: [s.n.], 2015. p. 110-118. ISSN:2447-0414.

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. *QoE-Based Adaptive Control of Speech Quality in a VoIP Call*. In: Telecommunications, 2015. AICT 2015: The Eleventh Advanced International Conference on Telecommunications. Bruxelas, Bélgica: IARIA, 2015. p. 136-137. ISSN:2308-4030, ISBN:978-1-61208-411-4

# Lista de Figuras

2.1	Esquema de comunicação entre dois usuários. . . . .	23
2.2	Diferença entre os métodos intrusivos e não intrusivos. . . . .	30
2.3	Diagrama de blocos de um sistema de avaliação PESQ /POLQA. . . . .	32
2.4	Esquema das contribuições no QoS de ponta a ponta. . . . .	34
2.5	Comparação entre os componentes de QoE e QoS. . . . .	37
2.6	Resumo das diferentes aspectos entre QoE e QoS em chamadas VoIP. . .	37
4.1	Cenário de teste. . . . .	51
4.2	Fluxo de teste. . . . .	57
5.1	MOS instantânea em função do tempo para diferentes perdas de pacotes.	62
5.2	MOS médio em função do tempo para diferentes perdas de pacotes. . .	62
5.3	MOS médio <i>versus</i> perda de pacotes. . . . .	63
5.4	Cenário de teste - codec PCMU. . . . .	66
5.5	Cenário de teste - codec GSM. . . . .	68
5.6	Cenário de teste - codec G.723. . . . .	69
5.7	Cenário de teste - codec G.729. . . . .	71
5.8	Cenário de teste - codec G.726-32. . . . .	72
5.9	Cenário de teste - taxa de perdas 0%. . . . .	73
5.10	Cenário de teste - taxa de perdas 1%. . . . .	75
5.11	Cenário de teste - taxa de perdas 2%. . . . .	76
5.12	Cenário de teste - taxa de perdas 2,5%. . . . .	77
5.13	Cenário de teste - taxa de perdas 3%. . . . .	78
5.14	Cenário de teste - taxa de perdas 5%. . . . .	79
5.15	Cenário de teste - taxa de perdas 7,5%. . . . .	80
5.16	Cenário de teste - taxa de perdas 10%. . . . .	81

# Lista de Tabelas

2.1	Características dos codecs de voz de Banda Estreita ( <i>Narrowband</i> ) mais conhecidos. . . . .	27
2.2	Pontuação MOS. . . . .	31
2.3	Classes de sensibilidade ao atraso para diferentes casos de uso. . . . .	40
4.1	Arquivos de áudio utilizados nos testes. . . . .	50
4.2	Codecs utilizados nos experimentos. . . . .	51
4.3	Perdas de Pacotes utilizados nos experimentos. . . . .	51
4.4	Tela do PJSUA se tudo estiver funcionando corretamente. . . . .	55
4.5	Interpretação do Coeficiente de Correlação de Pearson. . . . .	59
5.1	Qualidade PESQ e POLQA dos arquivos originais - <i>Full Reference</i> . . . . .	64
5.2	Média e desvio das medidas para o codec PCMU para cada perda em teste. . . . .	65
5.3	Correlação entre as medidas do codec PCMU. . . . .	66
5.4	Média e desvio das medidas para o codec GSM para cada perda em teste. . . . .	67
5.5	Correlação entre as medidas do codec GSM. . . . .	67
5.6	Média e desvio das medidas para o codec G.723 para cada perda em teste. . . . .	68
5.7	Correlação entre as medidas do codec G.723. . . . .	69
5.8	Média e desvio das medidas para o codec G.729 para cada perda em teste. . . . .	70
5.9	Correlação entre as medidas do codec G.729. . . . .	70
5.10	Média e desvio das medidas para o codec G.726-32 para cada perda em teste . . . . .	71
5.11	Correlação entre as medidas do codec G.726-32. . . . .	72
5.12	Média e desvio das medidas para perda de 0% para cada codec em teste. . . . .	73
5.13	Correlação entre as medidas de perda 0,0%. . . . .	73
5.14	Média e desvio das medidas para perda de 1% para cada codec em teste. . . . .	74
5.15	Correlação entre as medidas de perda 1,0%. . . . .	74
5.16	Média e desvio das medidas para perda de 2% para cada codec em teste. . . . .	75
5.17	Correlação entre as medidas de perda 2,0%. . . . .	76
5.18	Média e desvio das medidas para perda de 2,5% para cada codec em teste. . . . .	76
5.19	Correlação entre as medidas de perda 2,5%. . . . .	77
5.20	Média e desvio das medidas para perda de 3% para cada codec em teste. . . . .	78
5.21	Correlação entre as medidas de perda 3,0%. . . . .	78
5.22	Média e desvio das medidas para perda de 5% para cada codec em teste. . . . .	79
5.23	Correlação entre as medidas de perda 5,0%. . . . .	79
5.24	Média e desvio das medidas para perda de 7,5% para cada codec em teste. . . . .	80
5.25	Correlação entre as medidas de perda 7,5%. . . . .	80
5.26	Média e desvio das medidas para perda de 10% para cada codec em teste. . . . .	81

5.27	Correlação entre as medidas de perda 10,0%. . . . .	81
5.28	Parâmetros do modelo de Regressão Linear para resultados do PESQ . .	82
5.29	Parâmetros do modelo de Regressão Linear para resultados do POLQA	83
5.30	Recomendação de seleção de codecs de acordo com os resultados do Modelo-E. . . . .	84
5.31	Recomendação de seleção de codecs de acordo com os resultados do PESQ.	84
5.32	Recomendação de seleção de codecs de acordo com os resultados do POLQA. . . . .	84

# Lista de Abreviaturas e Siglas

<b>ACELP</b>	Algebraic Code Excited Linear Prediction
<b>ADPCM</b>	Adaptive Differential Pulse Code Modulation
<b>AMR</b>	GSM Adaptive Multi-Rate
<b>AMR-WB</b>	AMR Wideband
<b>ANOVA</b>	Analysis of Variance
<b>BER</b>	Bit Error Rate
<b>BV</b>	Broad Voice
<b>CELP</b>	Code Excited Linear Prediction
<b>CS-ACELP</b>	Conjugate Structure – Algebraic Code Excited Linear Prediction
<b>ETSI</b>	European Telecommunications Standards Institute
<b>FR</b>	Full Reference
<b>GSM</b>	Global System for Mobile Communications
<b>GSM-EFR</b>	GSM – Enhanced Full Rate
<b>GSM-FR</b>	GSM – Full Rate
<b>GSM-HR</b>	GSM – Half Rate
<b>HTTP</b>	Hypertext Transport Protocol
<b>IEEE</b>	Institute of Electric and Electronic Engineers
<b>iLBC</b>	Internet Low Bit Rate Codec
<b>IP</b>	Internet Protocol
<b>ITU</b>	International Telecommunications Union
<b>ITU-T</b>	ITU Telecommunication Standardization Sector
<b>LAN</b>	Local Area Network
<b>LD-CELP</b>	Low Delay – Code Excited Linear Prediction
<b>LPC</b>	Linear-Predictive Coding
<b>MOS</b>	Mean Opinion Score
<b>MOS-CQE</b>	Mean Opinion Score – Conversational Quality, Estimated
<b>MOS-LQO</b>	Mean Opinion Score – Listening Quality Objective
<b>MOS-LQS</b>	Mean Opinion Score – Listening Quality Subjective
<b>MP-MLQ</b>	Multi-Pulse – Maximum Likelihood Quantizer
<b>NB</b>	Narrow band
<b>PCM</b>	Pulse Code Modulation
<b>PER</b>	Packet Error Rate
<b>PESQ</b>	Perceptual Evaluation of Speech Quality
<b>PLB</b>	Packet Loss Behavior
<b>PLC</b>	Packet Loss Concealment
<b>POLQA</b>	Perceptual Objective Listening Quality Assessment
<b>PSNR</b>	Peak signal-to-noise ratio
<b>PSTN</b>	Public Switched Telephone Network
<b>QoE</b>	Quality of Experience
<b>QoS</b>	Quality of Service

<b>RFC</b>	Request for Comments
<b>RPE-LTP</b>	Regular Pulse Excitation – Long Term Prediction
<b>RTP</b>	Real-time Transport Protocol
<b>RTCP</b>	RTP Control Protocol
<b>RTCP SR</b>	RTCP Sender Report
<b>RTCP RR</b>	RTCP Receiver Report
<b>RTCP XR</b>	RTCP eXtended Reports
<b>SIP</b>	Session Initiation Protocol
<b>TCP</b>	Transmission Control Protocol
<b>TSNFC</b>	Two-Stage Noise Feedback Coding
<b>UDP</b>	User Data Protocol
<b>UWB</b>	Ultra-Wide band
<b>UMTS</b>	Universal Mobile Telecommunications System
<b>VAD</b>	Voice Activity Detector
<b>VoIP</b>	Voice over Internet Protocol
<b>WB</b>	Wide band

# Sumário

<b>1</b>	<b>Introdução</b>	<b>19</b>
1.1	Problema	19
1.2	Objetivos	21
1.3	Estrutura do documento	22
<b>2</b>	<b>Referencial Teórico</b>	<b>23</b>
2.1	Voice over Internet Protocol	23
2.1.1	Influência do codificador em chamadas VoIP	24
2.1.2	Influência das condições da rede em chamadas VoIP	24
2.1.3	Fluxos de dados em uma chamada VoIP	25
2.2	Codificador-Decodificador - Codec	26
2.2.1	Tempo de atraso no Codec	26
2.2.2	Técnicas de Codificação	27
2.3	Sistemas de Controle Adaptativo	28
2.4	Conceitos de Qualidade em VoIP	29
2.4.1	Qualidade do Sinal da Fala	29
2.4.2	Qualidade Subjetiva do Sinal da Fala	30
2.4.3	Qualidade Objetiva do Sinal da Fala	31
2.4.4	Qualidade Princípios e Definições	33
2.5	Qualidade de Serviço (QoS)	34
2.6	Qualidade de Experiência (QoE)	35
2.7	Considerações finais sobre QoE e QoS	36
2.8	Modelo-E	37
2.8.1	Relação sinal-ruído básica, $R_0$	38
2.8.2	Fator de perdas simultâneas, $I_s$	39
2.8.3	Fator de perdas associadas ao atraso, $I_d$	39
2.8.4	Fator de perdas associadas ao equipamento, $I_{e-eff}$	41
2.8.5	Fator de Vantagem, $A$	42
2.8.6	Considerações sobre o Modelo-E	42
<b>3</b>	<b>Trabalhos Relacionados</b>	<b>43</b>
<b>4</b>	<b>Materiais e Métodos</b>	<b>46</b>
4.1	Metodologia	46
4.2	Protocolo VoIP	46
4.3	Agente VoIP	47
4.4	Biblioteca Intel	47
4.5	Sistemas Operacionais Utilizados	47
4.5.1	Ferramentas Auxiliares	48

4.6	Montagem do cenário . . . . .	49
4.7	Cenário de Teste . . . . .	49
4.8	Áudios Selecionados . . . . .	49
4.8.1	Nomenclatura dos arquivos . . . . .	50
4.9	Instalação e Configuração do Ambiente . . . . .	53
4.9.1	Instalação do IPP . . . . .	53
4.9.2	Exportação das variáveis de ambiente . . . . .	53
4.9.3	Instalação do IPP Codecs (Samples) . . . . .	53
4.9.4	Instalação do PJSIP . . . . .	54
4.9.5	Verificação . . . . .	54
4.9.6	Modificações . . . . .	55
4.9.7	Utilização . . . . .	55
4.10	Realização dos Experimentos . . . . .	57
4.11	Teste Piloto . . . . .	58
4.12	Teste de Massa . . . . .	58
4.13	Teste Estatístico . . . . .	58
4.14	Medidas . . . . .	60
<b>5</b>	<b>Análises dos Resultados</b>	<b>61</b>
5.1	Teste Piloto . . . . .	61
5.2	Teste de Massa . . . . .	63
5.2.1	Resultados por Codec . . . . .	65
5.2.2	Resultados por Perdas . . . . .	72
5.3	Relação entre as medidas obtidas, PESQ e POLQA . . . . .	82
5.4	Classificação dos Codecs . . . . .	83
<b>6</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>85</b>
	<b>Referências Bibliográficas</b>	<b>88</b>
	<b>Anexo</b>	<b>94</b>
	<b>Anexo</b>	
<b>A</b>	<b>Senteças de Harvard</b>	<b>95</b>
A.1	OSR_us_000_0010_8k.wav Feminina 1 . . . . .	95
A.2	OSR_us_000_0011_8k.wav Feminina 1 . . . . .	95
A.3	OSR_us_000_0012_8k.wav Feminina 1 . . . . .	96
A.4	OSR_us_000_0013_8k.wav Feminina 1 . . . . .	96
A.5	OSR_us_000_0014_8k.wav Feminina 1 . . . . .	97
A.6	OSR_us_000_0015_8k.wav Feminina 2 . . . . .	97
A.7	OSR_us_000_0016_8k.wav Feminina 2 . . . . .	98
A.8	OSR_us_000_0017_8k.wav Feminina 2 . . . . .	98
A.9	OSR_us_000_0018_8k.wav Feminina 2 . . . . .	99
A.10	OSR_us_000_0019_8k.wav Feminina 2 . . . . .	99
A.11	OSR_us_000_0030_8k.wav Masculina . . . . .	100
A.12	OSR_us_000_0031_8k.wav Masculina . . . . .	100
A.13	OSR_us_000_0032_8k.wav Masculina . . . . .	101
A.14	OSR_us_000_0034_8k.wav Masculina . . . . .	101

A.15 OSR_us_000_0035_8k.wav Masculina	102
A.16 OSR_us_000_0036_8k.wav Masculina	102
A.17 OSR_us_000_0037_8k.wav Masculina	103
A.18 OSR_us_000_0038_8k.wav Masculina	103
A.19 OSR_us_000_0039_8k.wav Masculina	104
A.20 OSR_us_000_0040_8k.wav Masculina	104
A.21 OSR_us_000_0057_8k.wav Masculina	105
A.22 OSR_us_000_0059_8k.wav Masculina	105
A.23 OSR_us_000_0060_8k.wav Masculina	106
A.24 OSR_us_000_0061_8k.wav Masculina	106

# Capítulo 1

## Introdução

A tecnologia Voz sobre Protocolo de Internet (*Voice over Internet Protocol* - VoIP<sup>1</sup>) permite a transmissão de sinais de voz, através da rede de computadores, por meio do *Internet Protocol* (IP<sup>2</sup>) na forma de pacotes de dados. A tecnologia VoIP é a combinação da tecnologia de telecomunicações com a tecnologia de rede de computadores, em particular as redes IP. Sua evolução resultou na Telefonia IP, que consiste no fornecimento de serviços de telefonia utilizando a rede IP para o estabelecimento de chamadas e comunicação de voz (KARAPANTAZIS; PAVLIDOU, 2009).

A necessidade ou desejos de se comunicar com alguém fisicamente distante, por motivos pessoais ou profissionais, estimulam o uso de serviços baseados em VoIP. A telefonia IP apresenta-se no cotidiano das pessoas de forma direta, por meio de aplicativos como o Skype, ou indireta, quando uma operadora de telefonia encaminha as diversas chamadas de seus usuários por meio de uma infraestrutura baseada em VoIP, uma vez que com o uso da rede IP, a localização do usuário é transparente. Assim, o custo de uma chamada usando essa tecnologia não é calculado pela distância entre os pontos chamado e chamador como na comunicação de longa distância (*Discagem Direta à Distância* - DDD/*Discagem Direta Internacional* - DDI) pelas redes convencionais dos sistemas públicos de telefonia comutada (*Public Switched Telephone Network* - PSTN), tornando-a mais atraente ao usuário.

Na comunicação VoIP, os sinais analógicos da fala, como em uma chamada, são transformados em dados digitais e transferidos por meio da rede IP. Também são implementados protocolos adicionais de sinalização de chamadas e de transporte de voz, uma vez que permitem a comunicação com qualidade equivalente ou superior, em alguns casos, àquela fornecida pelas PSTN.

### 1.1 Problema

De acordo com Karapantazis e Pavlidou (2009), a tecnologia para transmissão de voz sobre IP data da década de 1990, quando começou mais como uma brincadeira entre os entusiastas de computador. Mesmo com limitações de largura de banda na época, a

---

<sup>1</sup>doravante denominada VoIP

<sup>2</sup>doravante denominada IP

tecnologia VoIP desencadeou um grande interesse em ambas as comunidades industriais e científicas por ter o potencial de mudar radicalmente as comunicações telefônicas. Embora a tecnologia VoIP desfrute dos avanços realizados nas duas últimas décadas, um dos problemas que ainda não foram solucionados é a provisão de garantias de qualidade. Para manter e atrair novos usuários, a qualidade da prestação de serviços VoIP precisa ser medida e otimizada para garantir a satisfação do usuário (RODRIGUES *et al.*, 2009).

Nos últimos anos, os mecanismos de controle de Qualidade de Serviço (*Quality of Service* - QoS) têm sido desenvolvidos para otimizar a utilização dos recursos da rede e do terminal do usuário de modo a minimizar os efeitos de degradações sobre a qualidade da fala. Alguns desses mecanismos procuram adaptar o fluxo de voz ou de outros parâmetros relacionados à VoIP, de acordo com mudanças significativas na rede, preferência dos usuários finais, ou de exigências das operadoras prestadoras de serviços.

A transmissão de sinais da fala sobre redes de computadores (VoIP) está sujeita a diversas degradações, tanto na camada de aplicação quanto na camada de rede. Como as redes de dados baseadas no protocolo TCP/IP não foram projetadas para entregar pacotes em tempo real, os valores de atraso e de taxa de perda de pacotes podem comprometer a interatividade e a integridade da chamada VoIP. Entre as degradações, encontram-se a compressão do codificador, o atraso de fim-a-fim e a perda de pacotes. Além disso, a convergência (telefônica, dados, etc.) incentiva projetos de arquitetura de rede que atende diferentes tipos de serviços em uma única rede. As aplicações do Skype e do Google *Talk* permitem que os usuários não apenas façam "ligações telefônicas" pela Internet, mas também as melhorem com videoconferências.

Procurar controlar a qualidade de uma chamada em tempo real, buscando a melhor qualidade possível tem sido um desafio. Os sistemas de controle adaptativo em geral respondem às mudanças de estado em seu ambiente interno ou externo com a orientação de um sistema de controle subjacente. Os sistemas de VoIP são particularmente susceptíveis de exigir adaptação dinâmica de uma solução para lidar com o dilema complexo entre qualidade de voz e entaves, consequência à natureza do controle descentralizado das redes IP e da natureza estocástica da entrega de pacotes de dados. Enquanto as soluções existentes para o controle de QoS para VoIP apresentam alguma melhoria de desempenho e algum tipo de realimentação, elas não fornecem foco explícito na malha de controle (MÜLLER *et al.*, 2008). Entretanto, a QoS não é como um usuário percebe a qualidade de um serviço. O termo Qualidade de Experiência (QoE) é mais adequado. Medidas de QoS relacionam-se diretamente com o estado da rede, enquanto que medidas de QoE relacionam-se diretamente a percepção do nível de qualidade que os usuários têm. Essa percepção irá desempenhar um papel fundamental na tomada de decisão em uma aplicação VoIP para se obter sucesso ou fracasso (CANO; CERDAN, 2012).

Um tradicional sistema de VoIP, sem qualquer comportamento adaptativo, consiste em três componentes principais: um transmissor (fonte de tráfego VoIP), um receptor (ponto de destino do fluxo de voz) e a rede IP subjacente. No lado do transmissor, a fala é codificada em quadros de voz, que são agrupados em pacotes (empacotamento).

Este fluxo de voz é tipicamente transportado através da rede IP juntamente com outros tipos de tráfego ou através de uma rede VoIP dedicada. No destino, o *buffer* de compensação de *jitter* elimina variações de atraso, e o decodificador compensa a perda de pacotes antes de reproduzir o fluxo de voz reconstituído para o receptor.

Para tornar um sistema adaptativo (ou seja, para gerenciar a qualidade da fala em tempo de execução de acordo com alguma política pré-estabelecida), é necessário projetar dois componentes, de acordo com Myakotnykh e Thompson (2009) : (1) mecanismos objetivos de avaliação da qualidade da fala em tempo real, e (2) algoritmos de adaptação de parâmetros.

## 1.2 Objetivos

O Objetivo Geral desta Tese é propor um método para se medir a Qualidade de Experiência (QoE) durante uma chamada VoIP utilizando-se do Modelo-E modificado, comparando os resultados obtidos a partir destas medidas com os resultados obtidos por outros mecanismos objetivos de avaliação da qualidade da fala como o PESQ (ITU-T, 2001) e POLQA (ITU-T, 2014). O intuito é de formar uma base de conhecimento com o propósito de sugerir os melhores parâmetros disponíveis de configuração que afetam o fluxo de voz da chamada VoIP em curso, afim de obter-se a maior Qualidade de Experiência de chamada possível dentro de um determinado cenário de rede.

Os Objetivos Específicos desta Tese são:

- combinar e refinar os métodos para se medir a Qualidade de Experiência, agregando contribuições dos métodos propostos por Carvalho (2011), Rodríguez (2009) e Myakotnykh e Thompson (2009) para medidas de Qualidade de Serviço (QoS) modificando-os com o foco em Qualidade de Experiência;
- desenvolver e validar uma rotina para avaliar a qualidade instantânea e acumulada da Qualidade de Experiência de uma chamada VoIP na estação receptora do fluxo de dados;
- efetuar chamadas VoIP em diversas configurações de chamada e diferentes condições de rede;
- medir a Qualidade de Experiência do sinal resultante da transmissão por meio de mecanismos objetivos padronizados de medida de qualidade de fala como o PESQ (ITU-T, 2001) e POLQA (ITU-T, 2014);
- analisar e comparar os resultados obtidos em tempo real com os obtidos a partir de outros mecanismos objetivos de avaliação da qualidade da fala como o PESQ (ITU-T, 2001) e POLQA (ITU-T, 2014);
- classificar os resultados obtidos pelas diferentes formas de medidas de Qualidade de Experiência; e

- formar uma base de conhecimento com as informações das chamadas VoIP em certos cenários, de modo a recomendar o melhor codec candidato disponível (nem sempre o melhor codec estará disponível) para a substituição do codec em uso, e/ou os novos parâmetros de operação sob esta nova condição.

### **1.3 Estrutura do documento**

Esta Tese está dividida em seis capítulos dos quais este foi o Capítulo 1 . A estrutura é a seguinte: Capítulo 2 apresenta o referencial teórico deste trabalho. O Capítulo 3 apresenta os resultados da revisão sistemática. O Capítulo 4 apresenta os materiais utilizados, a configuração utilizada no ambiente de testes, a metodologia seguida para se obter as medidas de interesse e detalha os cenários para testes. O Capítulo 5 apresenta os resultados das medições e discute as análises dos resultados. O Capítulo 6 apresenta as considerações finais, conclusões da investigação bem como as perspectivas de continuidade deste trabalho.

## Capítulo 2

### Referencial Teórico

Este Capítulo apresenta os conceitos necessários ao melhor entendimento desta Tese e foram divididos em Seções. Cada Seção engloba um conceito que pode ser lido individualmente e não necessariamente na ordem apresentada.

#### 2.1 Voice over Internet Protocol

Voz sobre IP (*Voice over Internet Protocol - VoIP*), telefonia IP, telefonia Internet, telefonia em banda larga e voz sobre banda larga é o roteamento de conversação humana que usa a Internet ou qualquer outra rede de computadores baseada no Protocolo de Internet (IP), tornando a transmissão de voz mais um dos serviços suportados pela rede de dados (KARAPANTAZIS; PAVLIDOU, 2009).

Durante uma chamada de voz em um sistema de telefonia IP, o fluxo de mídia percorre diversos dispositivos e/ou meios físicos, desde sua origem até o destino. Em todos eles, a passagem não ocorrerá sem que haja perdas ou atraso (CARVALHO, 2004).

A Figura 2.1 apresenta o caminho percorrido pela voz em uma conversação em um único sentido (da esquerda para a direita / do transmissor até o receptor / do microfone até o alto-falante) em uma aplicação VoIP sem a existência de um terminal da PSTN ou da Rede de Telefonia Móvel (*Public Land Mobile Network - PLMN*).

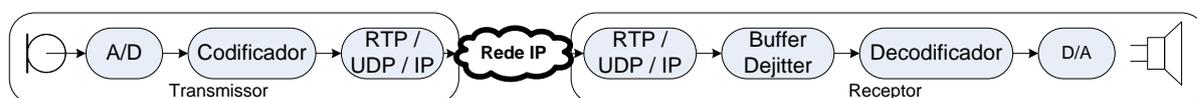


Figura 2.1: Esquema de comunicação entre dois usuários.

Fonte: Adaptado de (CARVALHO, 2011)

Na Figura 2.1, o sinal de voz do remetente (onda mecânica) é captada por um microfone, a qual é convertida em um sinal elétrico analógico contínuo no tempo. Em seguida, este sinal é digitalizado por um conversor Analógico-Digital (A/D) e quantizado em amplitude e discretizado no tempo. O sinal discreto (em amplitude e tempo) resultante é então codificado e comprimido por um algoritmo CODEC (CODificador / DECODificador) em quadros de voz. Um ou mais quadros são encapsulados em um pacote de voz por meio da adição de cabeçalhos RTP (*Real Time Protocol*), UDP (*User*

*Data Protocol*) e IP (*Internet Protocol*). Em seguida, estes pacotes de voz são enviados para a rede IP, onde são submetidos alguma restrição, como o atraso de transmissão, *jitter* e perdas devido a erros de congestionamento ou de transmissão.

No receptor, os pacotes que chegam são inseridos em um *buffer dejitter* (ou *buffer* de reprodução), na qual são armazenados temporariamente para serem reproduzidos de forma síncrona. Se os pacotes não chegarem a tempo ao *buffer dejitter* a tempo de serem reproduzidos, estes pacotes são considerados atrasados e geralmente são descartados e, portanto, considerados perdidos no aplicativo. Depois de deixar o *buffer dejitter*, os quadros de fala são decodificados. Se um quadro for perdido, o decodificador preenche a lacuna por meio da aplicação de um algoritmo de ocultação de perda de pacotes (*Packet Loss Concealment* - PLC). Finalmente, o sinal digital é transformado em um sinal acústico, reproduzido para o ouvinte. Todavia, o sinal analógico decodificado reproduzido será apenas uma aproximação do sinal de áudio original e a qualidade de som poderá ser nitidamente degradada (KUROSE *et al.*, 2013; ASSEM *et al.*, 2013a).

A transmissão VoIP está sujeita a degradações na camada de aplicação e na camada de rede. Na camada de aplicação, têm-se degradações na compressão do codificador, por exemplo. Na camada de rede, tem-se o atraso fim-a-fim e na perda de pacotes, no *jitter* e nos níveis de largura de banda. Estas degradações serão explicadas com mais detalhes nas Seções 2.1.1 e 2.1.2.

### 2.1.1 Influência do codificador em chamadas VoIP

Existe uma relação entre degradação do sinal codificado, compactação dos bits, tempo de codificação, qualidade do codificador. Diferentes codificadores apresentam diferentes desempenhos e diferente tolerância à perda de pacotes.

### 2.1.2 Influência das condições da rede em chamadas VoIP

Nesta Seção é apresentada a interação de qualidade de voz e as condições de rede. Serão considerados os fatores que definem a condição de rede atual:

1. perda de pacotes;
2. latência;
3. *jitter*; e
4. largura de banda disponível.

Optou-se por esses parâmetros porque eles influenciam as comunicações VoIP. Juntos esses parâmetros determinam a Qualidade de Serviço (QoS) (TANENBAUM; WETHERALL, 2011).

Um pacote é uma unidade formatada de dados de tamanho limitado (PETERSON; DAVIE, 2004) e é utilizado para transmitir as informações na rede em formato de pequenas unidades gerenciáveis. A perda de pacote ocorre quando um pacote de informação transmitido não chegam ao seu destino a tempo de ser utilizado e ocorre o

descarte deste pacote. A perda de pacotes influencia a qualidade do serviço VoIP por causa da natureza do *streaming* de tempo real no tráfego VoIP; pacotes de VoIP perdidos normalmente não são retransmitidos. Assim, sabendo como cada codec pode lidar com a perda de pacotes é crucial escolher um codec adequado sob determinadas condições de rede instáveis. Diferentes codificadores apresentam diferentes desempenhos.

Segundo Peterson e Davie (2004), em redes, latência é o tempo que leva para uma mensagem atravessar de uma ponta da rede até a outra. Em outras palavras, latência da rede, também conhecido como atraso ou tempo de propagação, é o tempo que leva para um pacote de informação deixar o computador de origem e chegar ao seu destino.

Da mesma forma, também afeta a comunicação VoIP a prevalência de flutuação, ou *jitter* que segundo Tanenbaum e Wetherall (2011) é a variação estatística no atraso na entrega de dados em uma rede ou no tempo médio de chegada do pacote (latência). Pode ser definida como a medida de variação do atraso entre os pacotes sucessivos de dados, isto é, a variação do atraso de chegada entre os pacotes de fim-a-fim. Mesmo que os pacotes sejam despachados a intervalos constantes entre eles no transmissor, eles chegarão ao receptor com tempos relativamente diferentes (TANENBAUM; WETHERALL, 2011). As propriedades em tempo real de VoIP, que permitem a interatividade e comunicação entre os usuários de modo que duas ou mais pessoas possam conversar como se estivessem usando o sistema telefônico tradicional (PSTN), exigem a chegada de pacotes de forma constante. Se a variação do atraso de chegada for muito alta, os pacotes que chegam muito atrasados são descartados. Normalmente, as implementações de *streaming* incluem um *buffer* de *jitter* que ajuda a reduzir o impacto deste efeito, ao armazenar os pacotes por um curto período antes de reproduzi-los. Efetivamente, negocia uma redução na perda de pacotes por um atraso maior, mas o atraso deve ser mantido baixo para a VoIP. Esta técnica é limitada, e se o *jitter* aumentar para além dos limites do *buffer*, influenciará negativamente a qualidade da voz. Considera-se o *jitter* um importante fator de condição de rede, uma vez que as habilidades dos codecs lidarão com as diferentes variações de atraso de pacotes.

Finalmente, a largura de banda, também chamada de *throughput*, é dada pelo número de bits que pode ser transmitidos pela rede em um certo período de tempo (PETERSON; DAVIE, 2004). É importante, porque diferentes codecs têm diferentes requisitos de largura de banda devido ao fato de que eles empregam diferentes quantidades de compressão de dados. Os codecs de taxa alta de bits têm requisitos de largura de banda maiores, mas geralmente oferecem maior qualidade.

### 2.1.3 Fluxos de dados em uma chamada VoIP

Uma típica chamada VoIP compreende basicamente três fluxos de dados (CARVALHO, 2011):

1. *Fluxo de Sinalização*: permite que os terminais da chamada VoIP (*endpoints*) se comuniquem uns com os outros, configura e derruba chamadas, e re-negocia os parâmetros da sessão, durante a chamada;

2. *Fluxo de mídia*: ele transporta a voz codificada por algum algoritmo codec do remetente até o destinatário. É transmitida pelo *Real-time Transport Protocol* (RTP) (SCHULZRINNE *et al.*, 2003); e
3. *Fluxo de controle de mídia*: realizado pelo *RTP Control Protocol* (RTCP), quando fornece o controle e monitoramento da qualidade do serviço de transporte RTP. Três relatórios RTCP são de grande importância para fins adaptativos: relatório remetente (*sender report - SR*); relatório receptor (*receiver report - RR*), que realiza a transmissão e recepção estatísticas básicas do(s) remetente(s) ativo(s) e receptor(es) em uma sessão; e relatório estendido (*extended report - XR*) (FRIEDMAN *et al.*, 2003), que carrega informações como tempo de atraso no *buffer de dejitter* e informações de qualidade como o MOS (*Mean Opinion Score*) que será explicado detalhadamente na Seção 2.4.1.

## 2.2 Codificador-Decoder - Codec

O Codificador-Decoder (codec) é um algoritmo para reduzir a utilização largura de banda, a fim de que a transmissão possa ter uma gama de sinais ou ampliar os sinais (CARVALHO, 2011). A transmissão destes sinais possui custos associados, dentre eles o custo do meio de transmissão quando se procura transmitir a maior quantidade possível de dados, utilizando uma técnica de compressão.

O foco desta Tese é apenas os codificadores da fala, apresentados nas próximas Subseções.

### 2.2.1 Tempo de atraso no Codec

Os modernos codificadores de voz operam em um conjunto de amostras de fala conhecidos como quadros. Cada bloco de entrada de amostras de fala é transformado em um quadro comprimido. O quadro de voz codificada não é gerado até que todas as amostras de fala deste bloco de entrada sejam recolhidas pelo codificador ( $N_{frame}$ ). Existe um atraso de um quadro antes da transformação começar ( $T_{frame}$ ). Além disso, muitos codificadores também olham para o próximo quadro (quadro sucessivo) para melhorar a eficiência de compressão. O comprimento deste aspecto de antecipação é conhecido por tempo de antecipação do codificador ( $T_{look}$ ). O tempo necessário para processar um quadro de entrada é assumido por ser o mesmo que o comprimento do quadro, dado que a utilização eficaz dos recursos de processador será realizada quando um par codificador/decoder (ou vários pares de codificador / decoder operando em paralelo na entrada de vários fluxos) utilizar a capacidade total de processamento disponível (uniformemente distribuídos no domínio do tempo) (ITU-T, 2003). Assim, o atraso mínimo devido ao processamento relacionados com o codec de VoIP é dado pela seguinte Expressão 2.1:

$$T_{codec} = 2 \times (N_{frame} + T_{frame}) + T_{look} \quad (2.1)$$

A Tabela 2.1 lista o atraso necessário para gerar um quadro de voz ( $T_{codec}$ ). Também apresenta um resumo dos parâmetros: taxa de codificação, duração do quadro ( $T_{frame}$ ) e tempo de antecipação de codificadores de voz ( $T_{look}$ ).

Tabela 2.1: Características dos codecs de voz de Banda Estreita (*Narrowband*) mais conhecidos.

Codec	Taxa de Bits (kbit/s)	$T_{frame}$ (ms)	$T_{look}$ (ms)	Algoritmo (tipo)	Licença
G.711	64	0,125 <sup>a</sup>	—	PCM	Livre
G.723.1	6,3	30	7,5	MP-MLQ	Proprietário
G.723.1	5,3	30	7,5	ACELP	Proprietário
G.726	16, 24, 32, 40	0,125 <sup>a</sup>	—	ADPCM	Livre
G.728	16	0,625 <sup>a</sup>	—	LD-CELP	Proprietário
G.729	8	10	5	CS-ACELP	Proprietário
G.729A	8	10	5	CS-ACELP	Proprietário
G.729D	6,4	10	5	CS-ACELP	Proprietário
G.729E	11,8	10	5	CS-ACELP LPC	Proprietário
GSM-FR (6.10)	13	20	—	RPE-LTP	Livre
GSM-HR (6.20)	5,6	20	4,4	VSELP	Proprietário
GSM-EFR (6.60)	12,2	20	—	ACELP	Proprietário
AMR-NB	4,75 – 12,2	20	5	ACELP	Proprietário
Speex (NB)	2,15 – 24,6	20	10	CELP	Livre (codigo-aberto)
iLBC	13,33	30	10	LPC	Livre
iLBC	15,2	20	5	LPC	Livre
BV16	16	5	—	TSNFC	Proprietário

<sup>a</sup> Normalmente, as implementações em tais codecs geram pacotes com um comprimento mínimo de 10ms.

Fonte: Adaptado de (CARVALHO, 2011)

## 2.2.2 Técnicas de Codificação

As técnicas de codificação de fala são divididas em três categorias: codificadores de forma de onda, vocoders e codificadores híbridos.

### Codificadores de forma de onda

Os codificadores de forma de onda exploram a correlação no domínio do tempo e no domínio da frequência. Além disso, buscam preservar a forma geral da forma de onda do sinal. Estes codificadores são de baixa complexidade e de atraso menor, caracterizados por seguirem o sinal amostra-a-amostra, com base em suas características temporais, sem considerar em detalhes a natureza da produção e percepção de fala humana, obtendo altas qualidades do sinal, introduzindo um pequeno retardo na voz e utilizando uma largura de banda maior (RODRÍGUEZ, 2009; CARVALHO, 2011; SUN, 2004).

A forma mais simples de codificação de formato de onda é a Modulação por Código de Pulso (*Pulse Code Modulation* - PCM). Tem-se ainda o método de Modulação Diferencial por Códigos de Pulsos (*Differential Pulse Code Modulation* - DPCM) que trabalha apenas com a diferença dos valores entre amostras consecutivas do sinal de voz,

fazendo uma previsão do valor da próxima amostra de voz, a partir dos valores anteriores, e codifica apenas a diferença entre o valor previsto e o real. Se o método de previsão for bom, essas diferenças serão pequenas, podendo ser codificadas com um número de bits menor que o da amostra original. Já a Modulação Diferencial Adaptativa por Códigos de Pulsos (*Adaptive Differential Pulse Code Modulation* - ADPCM) tem uma maior eficiência, quando comparado aos métodos anteriores, uma vez que o quantizador se adapta às estatísticas na mudança do resíduo estimado. Neste grupo, somente o algoritmo PCM é de domínio público, ou seja, pode ser implementado livremente, sem preocupações com patentes (SUN, 2004; CARVALHO, 2004; LAMARE, 1999; CARVALHO, 2011; RODRÍGUEZ, 2009). Na Tabela 2.1 pode-se observar os algoritmos Codec G.711 PCM (64 kb/s) e G.726 ADPCM (40/32/24/16 kb/s) que representam este tipo em questão.

### **Codificadores de voz paramétricos (vocoders)**

Os codificadores de voz paramétricos (vocoders) são baseados no modelo de produção da fala e não são feitas tentativas para preservar a forma de onda de voz original. Estes codificadores produzem alguns parâmetros que descrevem a forma como o sinal foi gerado, considerando a natureza da produção e percepção de fala humana. O resultado é um discurso sintético de qualidade inferior, além do aumento do atraso e da complexidade computacional. Neste sentido, os sinais podem ser transmitidos com uma taxa de bits muito baixa, como por exemplo, o algoritmo Codec LPC (2,4 / 1,2 kb/s) apresentados na Tabela 2.1 na página 27.

### **Codificadores híbridos**

Os codificadores híbridos incorporam as vantagens de codificadores de forma de onda e vocoders para alcançar uma boa qualidade de voz com taxas entre 4,8 a 16 kb/s. Eles enviam um certo número de parâmetros, bem como algumas informações codificadas por forma de onda, fornecendo um compromisso razoável entre a qualidade e eficiência de codificação de voz. Englobam os codecs modernos e são largamente utilizados em sistemas VoIP de hoje. Pode-se citar como exemplo: G.729 CSACELP (8 kb/s), G.723.1 MP-MLQ / ACELP (6,3 / 5,3 kb/s), AMR (*Adaptive Multi-Rate*, ACELP) e iLBC (*Internet Low Bit Rate Codec*) apresentados na Tabela 2.1 na página 27 .

## **2.3 Sistemas de Controle Adaptativo**

A Qualidade de Experiência (QoE) depende, ao longo de uma transmissão de voz, de fatores como perda de pacotes, *jitter*, largura de banda além do fator de comprometimento do codec. É possível medir os valores e transformá-los em uma medida que servirá de parâmetro de decisão para a troca de codec ou de outro parâmetro do sistema de controle adaptativo.

Os sistemas de controle adaptativo em geral respondem a mudanças de estado em seu ambiente interno ou externo com a orientação de um sistema de controle sub-

jacente. Os sistemas de VoIP são particularmente susceptíveis de exigir adaptação dinâmica de uma solução para lidar com o dilema complexo entre qualidade de voz e entaves, devido à natureza do controle descentralizado das redes IP e da natureza de chegada dos pacotes de dados ao destino. Enquanto as soluções existentes para o controle de QoS para VoIP mostram alguma melhoria de desempenho e têm algum tipo de realimentação, o foco explícito não é fornecido na malha de controle (MÜLLER *et al.*, 2008). A Qualidade de Serviço (QoS) relaciona-se diretamente com o estado da rede, enquanto que a Qualidade de Experiência (QoE) relaciona-se diretamente a percepção do nível de qualidade que os usuários têm. Essa percepção desempenhará um papel fundamental na tomada de decisão de uma aplicação multimídia para VoIP para se obter sucesso ou fracasso (CANO; CERDAN, 2012).

## 2.4 Conceitos de Qualidade em VoIP

Nesta Seção apresenta-se os conceitos necessários referentes a Qualidade. Desta forma, inicialmente é apresentado o conceito de qualidade, bem como os tipos de qualidade e os métodos de medida das mesmas. Finalmente, comparações entre os diferentes tipos são descritas para o melhor entendimento desta Tese.

### 2.4.1 Qualidade do Sinal da Fala

Em sistemas de comunicação, as métricas de qualidade da fala fazem-se necessárias para projetar, testar, manter e expandir tais sistemas. Isto também vale para redes de comunicação voz sobre IP, uma vez que em sistemas baseados em VoIP, o receptor irá sempre obter uma mensagem aproximada do que foi dito na fonte. Isto ocorre devido à compressão com perdas realizada por codecs de voz, uma vez que depois de digitalizada, compactada e empacotada, a voz perde clareza, dependendo do algoritmo codificador utilizado. Além disso, redes TCP/IP não foram desenvolvidas para atender a aplicações de telefonia, dificultando a interatividade dos usuários, uma vez que os pacotes IP, que transportam voz, devem ser entregues em tempo real. Pacotes IP estão susceptíveis a congestionamento em nós da rede e podem se perder entre um usuário e outro. Por conseguinte, são necessários métodos de avaliação para determinar a qualidade da mensagem recebida.

A qualidade de voz pode ser medida subjetiva, objetiva ou estimada. O procedimento preciso para a aferição da qualidade é a avaliação que utiliza experimentos psicofísicos (subjetivos) com observadores humanos, tendo como resultado, por exemplo, o  $MOS_{LQS}$  (*Mean Opinion Score Listening Quality Subjective*). Embora precisa, esta metodologia demanda um longo período de tempo e um elevado custo operacional. Alternativamente, é possível utilizar métricas objetivas para obter uma estimativa da qualidade do sinal de fala. As métricas objetivas podem ser classificadas segundo a quantidade de informação necessária para estimar a qualidade do sinal: (1) Referência Completa (*Full Reference - FR*), (2) Referência Reduzida (*Reduced Reference - RR*), e (3) Sem Referência (*No-Reference - NR*). Como resultado, por exemplo, temos o  $MOS_{LQO}$

(*Mean Opinion Score Listening Quality Objective*). Temos ainda as Métricas Objetivas Estimadas, onde a partir das observações do canal de teste, um valor de Medida de Qualidade é estimado. Um exemplo de resultado deste tipo de métrica é o  $MOS_{CQE}$  (*Mean Opinion Score Computational Quality Estimation*). A síntese destes conceitos é apresentada na Figura 2.2. No caso, na Figura 2.2, o Canal de Teste, pode ser uma rede de computadores ou de telecomunicações, o ar, enfim, um meio por onde o sinal de teste atravessa e possa sofrer algum tipo de degradação até chegar ao seu destino.

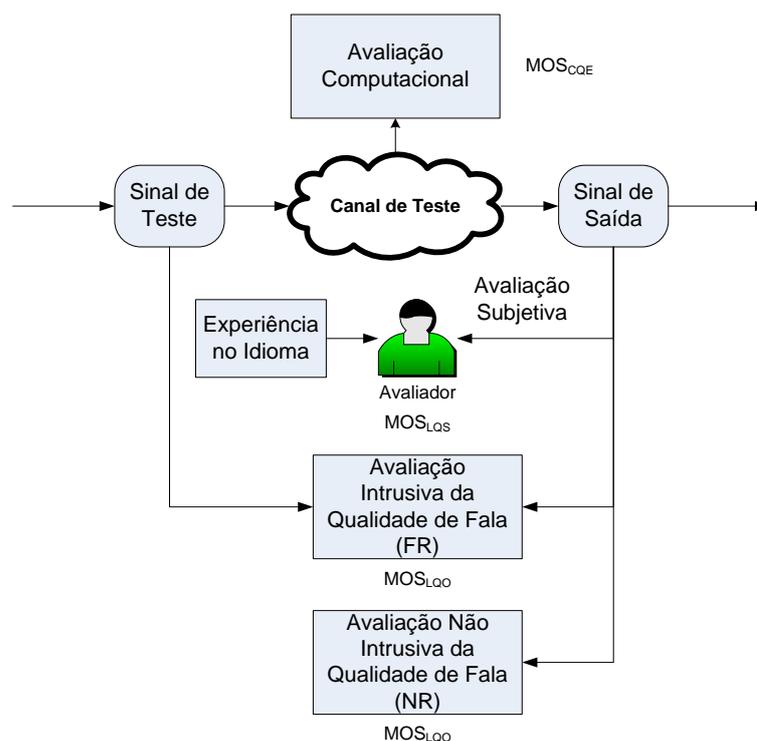


Figura 2.2: Diferença entre os métodos intrusivos e não intrusivos.  
Adaptado de (ITU-T, 1996)

Esta Subseção está subdividida em três partes: a primeira fornece uma visão geral e análise de Metodologias Subjetivas (Subseção 2.4.2). A segunda Subseção apresenta uma visão geral e análise de Metodologias Objetivas atuais de Medição de Qualidade da Fala (Subseção 2.4.3). A última Subseção 2.4.4 fornece uma descrição das formas bem como sintetiza os diversos tipos de medidas de qualidade.

## 2.4.2 Qualidade Subjetiva do Sinal da Fala

Os métodos de avaliação da qualidade subjetiva medem a qualidade percebida a partir da perspectiva do usuário. Um dos primeiros métodos para se medir a qualidade da fala foi a utilização da Pontuação de Opinião Média, ou (*Mean Opinion Score - MOS*), definido na Recomendação ITU-T P.800 (ITU-T, 1996). 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.

Avaliação subjetiva é muitas vezes consideradas como o método mais preciso para prever qualidade com base na percepção humana. A desvantagem é, no entanto, que

a realização de tais experiências é muito cara e demorada. Além disso, os resultados são válidos somente dentro do contexto de um experimento com amplo controle da expectativa dos envolvidos. Pode-se comparar, por exemplo, o desempenho de dois codecs, desde que a experiência realizada seja a mesma. Pode-se realizar o experimento em diferentes laboratórios, mas isso requer um enorme conhecimento (*know-how*) de avaliação subjetiva e o teste realizado pelos diferentes laboratórios deve ser muito bem equilibrado para evitar que um desequilíbrio no teste afete os resultados. Além disso, a reprodutibilidade dos resultados é fortemente dependendo do dia-a-dia dos indivíduos envolvidos no teste, bem como o número de indivíduos e testes no experimento. As experiências subjetivas são amplamente utilizados hoje em dia, a fim de avaliar e selecionar codecs ou para avaliar a precisão dos métodos de medição objetiva. É geralmente impossível, ou pelo menos não é prático, usar experiências subjetiva no campo, diariamente no laboratório ou, por exemplo, em testes ao dirigir um veículo.

A avaliação subjetiva por envolver de 12 a 24 participantes ouvindo individualmente um fluxo de áudio de vários segundos, é dispendiosa, consome tempo, não pode ser feita em tempo real e é de baixa reprodutibilidade. Neste método, os avaliadores ouvintes classificam a qualidade de áudio em uma escala, atribuindo uma pontuação de 1 (pobre) a 5 (excelente), conforme apresentado na Tabela 2.2, sobre qualidade da fala reproduzida pelo sistema de comunicação em teste e esta avaliação também é chamada de MOS direto.

Tabela 2.2: Pontuação MOS.

MOS	Qualidade	Descrição
5	Excelente	Um sinal de voz perfeito gravado em um local silencioso
4	Bom	Qualidade de uma chamada telefônica de longa distância (PSTN)
3	Razoável	Requer algum esforço na escuta
2	Pobre	Fala de baixa qualidade e difícil de entender
1	Ruim	Fala não clara, quebrada.

Fonte: Adaptado de Recomendação ITU-T P.800 (ITU-T, 1996)

### 2.4.3 Qualidade Objetiva do Sinal da Fala

Medida Objetiva não é tão precisa quanto medida subjetiva, mas a Medida Objetiva é mais fácil e barata de se aplicar. Em muitos casos, a Medida Objetiva é a única maneira viável para avaliar a qualidade, como foi no caso deste trabalho. Sua maior vantagem é a repetibilidade da medida. Sempre o mesmo resultado será obtido a partir dos mesmos dados de entrada, independentemente do dia, localização ou qualquer outro fator. Métodos objetivos são livres de contexto. No entanto, nunca se deve confiar nos resultados cegamente. Resultados objetivos só são válidos se o método objetivo foi validado para o tipo de prejuízos esperados e eles podem ter uma precisão limitada em algumas aplicações.

Diante do apresentado na Subseção 2.4.2, têm sido propostas várias técnicas para estimar objetivamente a MOS (sem percepção humana), também conhecido por MOS medida ou indireta, tal como a Avaliação Perceptiva da Qualidade da Fala (*Perceptual Evaluation of Speech Quality* - PESQ) especificada na Rec.ITU-T P.862 (ITU-T, 2001) e a Avaliação Perceptiva Objetiva da Qualidade de Audição (*Perceptual Objective Listening Quality Assessment* - POLQA) especificada na Rec.ITU-T P.863 (ITU-T, 2014), sucessora do algoritmo PESQ. Estes algoritmos são medidas de referência completa (FR) descritas nos padrões ITU-T que permitem a previsão da qualidade da fala comparando-se um sinal de referência a um sinal recebido conforme apresentado na Figura 2.2 (Avaliação Intrusiva) na página 30.

O algoritmo PESQ é uma família de padrões, compreendendo um método de ensaio para a avaliação automática da qualidade de voz como o experimentado por um usuário de um sistema de telefonia permitindo calcular a pontuação MOS aproximada.

A chave do processo é a transformação de ambos os sinais originais e degradados em uma representação intermédia, que é análoga à representação psicofísica (modelo perceptual) dos sinais de áudio no sistema auditivo humano conforme apresentado na Figura 2.3. PESQ combina o modelo perceptual com um algoritmo de alinhamento temporal que lida com atrasos variáveis, obtendo a representação interna do sinal original e sinal degradado. As diferenças na representação interna do sinal original e sinal degradado determinam a diferença audível. Estas diferenças são comparadas ao serem processados pelo modelo cognitivo psicoacústico. A saída dada pelo algoritmo PESQ é uma previsão do valor de MOS da qualidade percebida  $MOS_{LQO}$ , isto é, uma estimativa do resultado dos testes subjetivos reais  $MOS_{LQS}$ .

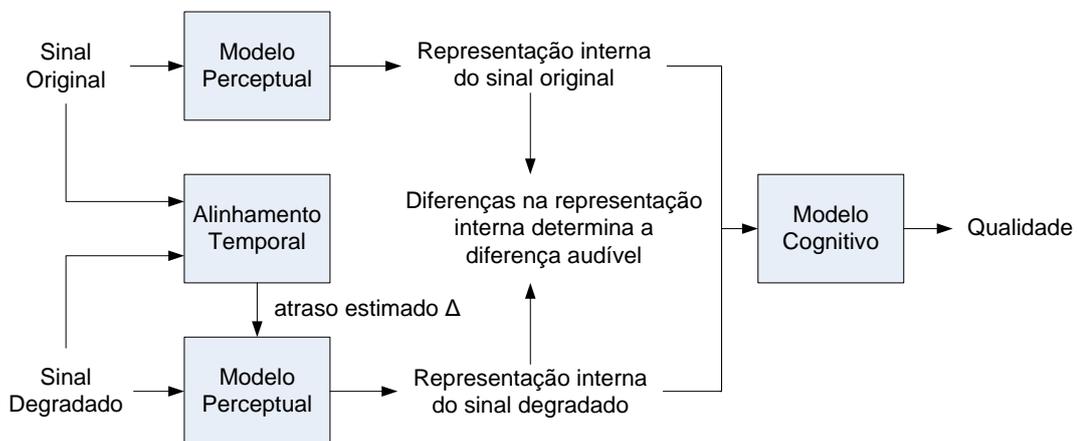


Figura 2.3: Diagrama de blocos de um sistema de avaliação PESQ /POLQA.

Um problema do algoritmo PESQ é que ele usa o mesmo modelo cognitivo tanto para o sinal original quanto para o sinal degradado. Isso implica que se houver pouca ou nenhuma degradação no sinal degradado, o valor de  $MOS_{LQO}$  será próximo 4,5 independente da qualidade do sinal original, mesmo que este não tenha esta qualidade.

O algoritmo POLQA foi especialmente desenvolvido para Voz HD, reder 3G e 4G/LTE, VoIP e aborda outras limitações do PESQ. POLQA estima a qualidade tanto de sinais de voz de banda estreita (*narrow band*) (300 Hz até 3400 Hz), quanto sinais de

banda larga (*wide band*) (50 Hz até 7000 Hz) e ultra banda larga (*ultra wide band*) (50Hz até 14000 Hz) ao contrário do algoritmo PESQ, que foi desenvolvida para dar uma estimativa objetiva da qualidade de voz de banda estreita (*narrow band*) (300 Hz até 3400 Hz).

Há várias diferenças algorítmicas significativas em todas as fases: alinhamento temporal, escalonamento e filtragem do nível do sinal e calibração dos modelos perceptivos e cognitivos, impactando no mapeamento da pontuação POLQA na escala MOS (COTANIS, 2011).

Além de melhorar a correlação geral com pontuações subjetivas MOS (HINES *et al.*, 2013), POLQA utiliza modelos cognitivos construídos de forma diferente entre si, ao contrário do PESQ, de modo que mesmo que não haja degradação no sinal degradado, a qualidade do sinal original será levada em consideração, evitando este problema presente no algoritmo PESQ (OPTICOM, 2007). O uso do POLQA ainda não é generalizado, ou livremente disponível para testes, razão por que o número limitado de publicações do desempenho deste algoritmo (POLQA) fora dos próprios testes de desenvolvimento e de conformidade.

No entanto, como PESQ, POLQA tem seu próprio conjunto de limitações, conforme descrito na ITU-T Rec. P.863.2 (ITU-T, 2014), tais como sequências de conversação longas e ambientes sonoros ruidosos. Além disso, por enquanto, devido à falta de bases de dados de fala, POLQA permanece insuficientemente testados em alguns idiomas como o russo, árabe, espanhol, vários idiomas asiáticos (COTANIS, 2011) e português, razão pela qual optamos por utilizar uma base de conversação em inglês.

Tanto o algoritmo PESQ (ITU-T, 2001) quanto o algoritmo POLQA (ITU-T, 2014) são medidas de referência completa (FR) descritos nos padrões ITU, pois permitem a previsão da qualidade da fala comparando dois sinais para gerar a MOS medida: um sinal de referência (por exemplo, capturado no transmissor) e um sinal degradado (por exemplo, capturado no receptor). Isto é feito de forma *off-line*. Esta exigência torna esta abordagem inadequada para monitoramento durante uma chamada. Para contornar este problema, utiliza-se uma medida de qualidade estimada, o Modelo-E (ITU-T, 2015), apresentado na Subseção 2.8.

#### 2.4.4 Qualidade Princípios e Definições

Esta Subseção contém algumas definições gerais de termos que são utilizados ao longo desta Tese. Definir o uso destes termos é necessário, uma vez que alguns deles são muitas vezes utilizados como sinônimos para descreverem diferentes aspectos de qualidade.

Os termos de medição "Referência Completa", "teste intrusivo" e "ensaio ativo", bem como seus opostos, muitas vezes são usados como sinônimos. No entanto, todos eles têm o seu significado distinto e não devem ser misturados.

Referência Completa (*Full Reference* - FR) / sem Referência (*No-Reference* - NR) refere-se ao tipo do algoritmo aplicado:

- Um algoritmo de Referência Completa requer acesso ao sinal de referência não distorcida, bem como para o sinal degradado. A medida é baseada na análise de

diferenças dos dois sinais.

- Um algoritmo com nenhuma referência (sem referência) por sua vez, precisa de acesso a apenas o sinal degradado, como consequência das características do sinal original serem praticamente desconhecidas.

Assim, a medida é baseada na suposição de que resulta apenas de artefatos detectados a partir da rede em teste e não representam características de um sinal de fonte original.

Em ambos os casos de avaliação objetiva, Referência Completa e Sem Referência, a saída será o  $MOS_{LQO}$  (*Mean Opinion Score - Listening Quality Objective*) por se tratar de medidas objetivas. Quando trabalhamos com medidas subjetivas o resultado será o  $MOS_{LQS}$  (*Mean Opinion Score - Listening Quality Subjective*). Quando estimamos a qualidade a partir de observações do Sistema, temos o  $MOS_{CQE}$  (*Mean Opinion Score - Conversational Quality, Estimated*). A Figura 2.2, na página 30, sintetiza as diferentes formas de se medir a Qualidade de uma comunicação.

## 2.5 Qualidade de Serviço (QoS)

O conceito de Qualidade de Serviço (*Quality of Service - QoS*) é utilizado para descrever a qualidade do serviço em uma rede. O termo QoS, originalmente, destinava-se a ser orientado ao usuário (ITU-T, 2008), passando a ser considerada como uma medida a partir da perspectiva de rede (DAI, 2011; HOSSFELD *et al.*, 2007).

A QoS foi definida pela Nokia (NOKIA, 2004) como "capacidade da rede para prestação de um serviço com um nível de serviço assegurado". Enquanto que Rahrer (RAHRER *et al.*, 2006) a definiu como "uma medida de desempenho na camada do pacote a partir da perspectiva de rede". QoS também se refere a um conjunto de tecnologias "que permite que o administrador de rede possa gerenciar os efeitos do congestionamento no desempenho do aplicativo, bem como proporcionar atendimento diferenciado ao fluxo de tráfego de rede selecionado ou para usuários selecionados." (RAHRER *et al.*, 2006).

A QoS de ponta a ponta depende das contribuições dos componentes como ilustrado na Figura 2.4.

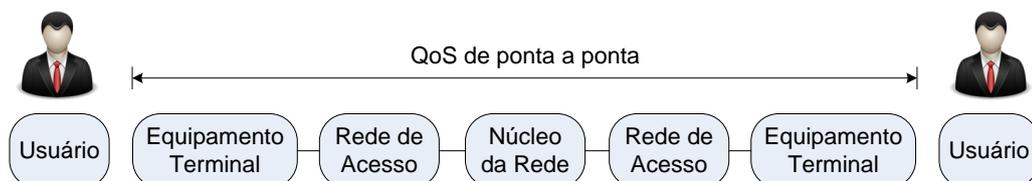


Figura 2.4: Esquema das contribuições no QoS de ponta a ponta. (ITU-T, 2008)

Para especificar a QoS fim-a-fim, é necessário mencionar as condições de operação especificadas em que um serviço é suportado através de uma conexão (conexão ou

orientado a conexão) que ocorre. A QoS pode também ser alterada para um dado conjunto de condições de funcionamento especificados pelas condições ambientais, tais como o tráfego e encaminhamento.

A QoS compreende tanto o desempenho relacionado à rede, como também o desempenho não relacionado à rede. Exemplos de desempenho relacionado à rede (*Network performance* - NP) são a taxa de bit errados (*Bit Error Ratio* - BER), latência, etc. Para o QoS não relacionado a rede, tem-se o tempo de fornecimento (operação), tempo de reparo, a gama de tarifas e tempo de resolução das reclamações, etc. A lista de critérios de QoS, para um determinado serviço, seria dependente do serviço e a relevância desta lista de critérios pode variar entre os segmentos da população de usuários.

As métricas de QoS da rede geralmente incluem perda de pacotes, atraso e *jitter* que são as deficiências principais que afetam a qualidade de voz e vídeo em aplicações VoIP.

## 2.6 Qualidade de Experiência (QoE)

A contínua mudança da metodologia para avaliação do desempenho da rede tem lugar na Internet. No passado, o desempenho da rede era principalmente avaliado usando medidas de valores concretos como perda de pacote, de atraso e *jitter* (medidas objetivas - meio). A tendência atual vai em direção à qualidade da aplicação específica que julga mais as experiências subjetivas do usuário final (ponta da rede) do que o parâmetro de rede (HOSSFELD *et al.*, 2007; BERG *et al.*, 2009).

O conceito de Qualidade de Experiência (QoE) relaciona diretamente a percepção do nível de qualidade que os usuários têm por um serviço prestado, por exemplo, VoIP. Isto representa uma medida subjetiva de experiências de um usuário para os serviços prestados, refletindo o grau de satisfação / insatisfação dos usuários finais com um determinado serviço, além de representar o quão bem o serviço atende às expectativas do usuário. Este termo é também referido como Qualidade de Serviço Percebida pelo Usuário (*Perceived Quality of Service* - PQoS), a fim de diferenciar da Qualidade de Serviço (QoS), que reflete o desempenho da rede.

O conceito de QoE pode apresentar diversas definições por diferentes comunidades. Uma definição unificada de QoE ainda é necessária. A primeira introdução do termo QoE ocorreu no início do século (XXI) e foi predominantemente promovida pela indústria. Por exemplo, Nokia (NOKIA, 2004) introduziu o conceito da percepção dos usuários finais sobre a qualidade do serviço e declarou: "QoE é como um usuário percebe a usabilidade de um serviço, ao utilizá-lo e o seu grau com este serviço. O termo QoE se refere à percepção do usuário sobre a qualidade de um serviço particular ou rede". Um ano mais tarde, o fórum de banda larga (*Broadband Forum*) definiu como um indicador de um sistema no atendimento dos requisitos dos usuários a medida de QoE, em seu relatório técnico, TR-126 (RAHRER *et al.*, 2006), explica que: "QoE é uma medida de desempenho de ponta a ponta (fim-a-fim) do nível de serviço a partir da perspectiva do usuário e uma indicação de quão bem o sistema atende às necessidades

do usuário.”

Com base nas descrições anteriores, a QoE é definido em Dai (2011) como uma descrição orientada para o utilizador da qualidade do serviço. Ele se refere à avaliação “subjetiva” dos usuários finais sobre a qualidade do serviço “global” e reflete as experiências dos usuários durante/ao longo de todo o serviço utilizado. A métrica chave do QoE é o MOS (*Mean Opinion Score*) (SUN *et al.*, 2013). Para dimensionar / avaliar a QoE de um serviço, as percepções do usuário, tais como “Bom”, “Satisfatória”, “Pobre” ou “Ruim” são frequentemente utilizados. QoE pode ser medido a partir de diferentes abordagens subjetivas, tais como as expectativas do usuário e experiências, podendo determinar a percepção do usuário (geral) final que são baseados em pesquisas, enquanto as medidas objetivas são principalmente com base em parâmetros de desempenho de QoS da rede, relacionadas com o serviço entrega (de transporte de rede). Geralmente incluem perda de pacotes, atraso e *jitter* que são as deficiências principais que afetam a voz e qualidade de vídeo em aplicações VoIP (SUN *et al.*, 2013). Os métodos existentes utilizados para avaliação da Qualidade de Experiência (QoE) baseiam-se principalmente em pesquisa com usuários e nos escores destes usuários, que são muito subjetivos, precisam de muito tempo de processamento e apresentam custo elevado (GONG *et al.*, 2009), porém os resultados são mais precisos (CANO; CERDAN, 2012).

A percepção positiva de um usuário sobre um serviço muitas vezes implica no bom desempenho da QoE cobrindo todo o período de utilização de todo o serviço, razão da QoE ser, às vezes, um conceito de “longo prazo”. No entanto, QoE também pode ser um conceito de “curto prazo”. Ele reflete as experiências do usuário, particularmente, a percepção negativa durante um curto período. Por exemplo, a má experiência sobre a qualidade de voz, mesmo em um curto período de uma aplicação de voz, é provavelmente grave o suficiente para permitir ao usuário terminar a utilização do serviço (DAI, 2011).

Kuipers *et al.* (KUIPERS *et al.*, 2010) concluíram uma visão geral de várias técnicas para medir a QoE, tendo o foco em ferramentas e metodologias livremente disponíveis.

A Qualidade de Experiência do usuário (QoE) está ganhando impulso, juntamente com os esforços de fornecer arquiteturas apropriadas, capazes de fornecer o nível esperado de Qualidade de Serviço (QoS) ao usuário (CANO; CERDAN, 2012).

## 2.7 Considerações finais sobre QoE e QoS

Diante do exposto nas seções anteriores para manter e atrair novos usuários, a qualidade da prestação de serviços de VoIP precisa ser medida e otimizada para garantir a satisfação do usuário. Aspectos de QoE, diferentemente dos aspectos de QoS apresentados na Figura 2.4 na página 34 englobam não só os usuários mas toda a infraestrutura envolvida, conforme Figura 2.5.

A provisão de garantia da qualidade é um dos problemas que não foi resolvido, embora a tecnologia VoIP desfrute dos avanços realizados desde a década de noventa (KARAPANTAZIS; PAVLIDOU, 2009). Os mecanismos de controle de Qualidade de

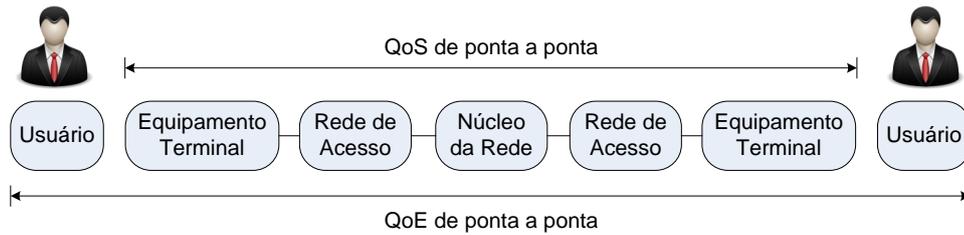


Figura 2.5: Comparação entre os componentes de QoE e QoS.

Serviço (*Quality of Service - QoS*) foram desenvolvidos para otimizar o uso dos recursos de rede e de terminais de usuário, a fim de minimizar os efeitos da degradação na qualidade da voz. Alguns destes mecanismos procuram adaptar o fluxo de voz ou outros parâmetros de VoIP, de acordo com mudanças significativas na preferência de usuários finais de rede, ou o fornecimento de requisitos de prestadores de serviços.

No futuro, a área de serviços da internet seguirá um novo paradigma no qual a inteligência do controle de rede é gradualmente transferida para a ponta da rede. Isso afetará tanto a Qualidade de Serviço (QoS) da conexão fim-a-fim, bem como a qualidade subjetiva de Experiência (QoE) percebida pelo usuário final (HOSSFELD; BIN-ZENHÖFER, 2008).

A Figura 2.6 sintetiza as diferenças entre QoE e QoS.

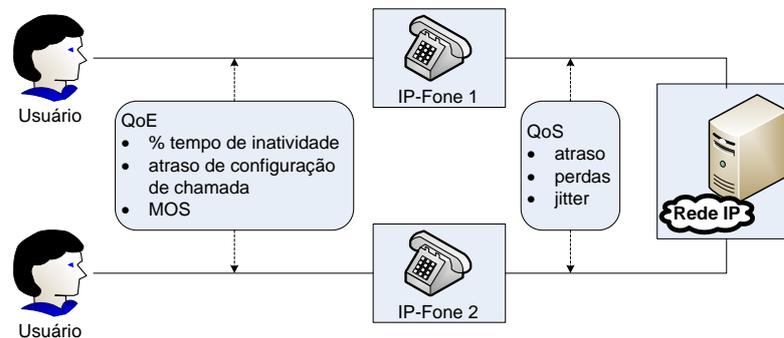


Figura 2.6: Resumo das diferentes aspectos entre QoE e QoS em chamadas VoIP.

## 2.8 Modelo-E

O Modelo-E foi proposto inicialmente pelo Instituto Europeu de Padronização em Telecomunicações (*European Telecommunications Standards Institute - ETSI*) no Relatório Técnico 250 (ETR250, 1996). Posteriormente foi atualizado pela União Internacional de Telecomunicações (*International Telecommunication Union - ITU*) e definido na Recomendação ITU-T G.107 (ITU-T, 2015). Ele baseia-se na premissa de que "fatores psicológicos, numa escala subjetiva, são aditivos" (ITU-T, 2015; ITU-T, 2007), em outras palavras, cada fator que contribui para a degradação de um fluxo de voz pode ser computado independentemente em um conjunto de equações, embora isso não implique que tais fatores não estejam correlacionados.

O Modelo-E fornece uma estimativa da qualidade da conversação com base em vários parâmetros de rede, ambiente e codificação, por isso, não precisa do sinal original. Concebida originalmente para fins de planejamento, exatamente por não precisar do sinal original, pode ser utilizada para a avaliação da qualidade, com certo cuidado.

O procedimento de medição no Modelo-E consiste na combinação pertinente de todos os parâmetros de transmissão do fluxo de voz para uma dada ligação, que retornam como resultado o Fator R, cujo valor varia de 0 (pior) a 100 (melhor) como uma expressão da medida da qualidade avaliada (ITU-T, 2015; ITU-T, 2007; LUSTOSA *et al.*, 2004; COSTA; NUNES, 2009).

O Fator R representa a soma de todos os fatores de degradação na comunicação e é obtido utilizando a expressão 2.2:

$$R = R_0 - I_s - I_d - I_{e-eff} + A \quad (2.2)$$

em que:

- $R_0$  representa em princípio, os efeitos da relação sinal-ruído (SNR) básica, como o ruído do circuito e ruído ambiente ;
- $I_s$  representa uma combinação de todas as deficiências que ocorrem mais ou menos simultaneamente com o sinal de voz;
- $I_d$  representa os prejuízos provocados associados ao atraso fim-a-fim;
- $I_{e-eff}$  representa as perdas associadas ao equipamento utilizado (fator de comprometimento do equipamento). Representa as perdas causadas por codecs de baixa taxa de bits;
- $A$  corresponde ao fator de vantagem ou fator de expectativa. Permite a compensação de fatores de comprometimento quando o usuário se beneficia de outros tipos de acesso, por ex. rede cabeada, rede celular dentro de uma construção; ou móvel em um veículo em movimento; ou acesso difícil, por multi saltos satelitais. Diferente dos outros parâmetros,  $A$  é um valor que é adicionado a  $R_0$ .

Nas Subseções 2.8.1, 2.8.2, 2.8.3 e 2.8.4 são apresentados detalhadamente os fatores e grandezas que influenciam cada uma dessas parcelas e quais delas podem ser reduzidas a uma constante.

### 2.8.1 Relação sinal-ruído básica, $R_0$

Na determinação da relação sinal-ruído básica ( $R_0$ ) é computado o ruído acrescido pelas diversas fontes de ruído, como circuitos de transmissão, ruído ambiente no lado do emissor, ruído ambiente no lado do receptor e ruído de fundo correspondente à sensibilidade do sistema auditivo humano. A recomendação ITU-T G.107 (ITU-T, 2015) apresenta as expressões matemáticas e os valores padrões para todos os parâmetros que não variam durante o cálculo do Fator R, resultando em um valor padrão de 93,2 para  $R_0$ .

### 2.8.2 Fator de perdas simultâneas, $I_s$

O Fator de perdas simultâneas ( $I_s$ ) é a soma de todas as deficiências que possam ocorrer mais ou menos simultaneamente com a transmissão da voz. Entre estas deficiências estão a queda na qualidade devido a uma conexão com o volume do sinal de fala demasiadamente alto, perdas causadas pela interferência da própria voz do locutor sobre o microfone/fone de ouvido do mesmo *handset* (microfonia) utilizado na transmissão e a distorção de quantização causada pela digitalização do sinal de voz. A Recomendação ITU-T G.107 (ITU-T, 2015) 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 (ITU-T, 2015) 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 (LUSTOSA *et al.*, 2004).

### 2.8.3 Fator de perdas associadas ao atraso, $I_d$

O fator de comprometimento de perdas associadas ao atraso ( $I_d$ ) representa todos os prejuízos devido ao atraso de sinais de voz. De acordo com a ITU-T G.107 (ITU-T, 2015), ele é dividido em três fatores de acordo com a Equação 2.3:

$$I_d = I_{dte} + I_{dle} + I_{dd} \quad (2.3)$$

em que:

- $I_{dte}$  representa as degradações devido ao eco no lado transmissor;
- $I_{dle}$  representa as degradações devido ao eco no lado receptor; e
- $I_{dd}$  representa as degradações causadas pelos atrasos  $T_a$  produzidos na rede e no codificador.

O valor de  $T_a$  representa o atraso absoluto de ponta-a-ponta e é dado pela Equação 2.4 (CARVALHO, 2011):

$$T_a = T_{net} + T_{buffer} + T_{codec} \quad (2.4)$$

em que:

- $T_{net}$  é o atraso da rede, referente ao primeiro pacote recebido no período de fala;
- $T_{buffer}$  é o atraso introduzido pelo *buffer* de compensação de *jitter* (*dejitter*); e
- $T_{codec}$  é o atraso inerente ao codec utilizado que possui valor fixo conhecido dada pela Equação 2.1.

O *buffer* de compensação de *jitter* elimina a variação de  $T_a$ , uma vez que todos os pacotes são atrasados de modo que possuam um  $T_a$  correspondente ao atraso máximo

tolerado (ITU-T, 2003). Isso assegura a cadência correta na reprodução do sinal da fala no dispositivo receptor.

O  $I_{dd}$  representa a degradação causada pelos atrasos ( $T_a$ ) produzidos na rede e no codificador. O atraso ocasionado pelo codificador é definido na Rec.G.107 (ITU-T, 2015) de acordo com o valor de  $T_a$ .

Se  $T_a \leq mT$  tem-se  $I_{dd} = 0$  e para  $T_a > mT$  tem-se:

$$I_{dd} = 25 \left\{ (1 + X^{6-sT})^{\frac{1}{6-sT}} - 3 \left( 1 + \left[ \frac{X}{3} \right]^{6-sT} \right)^{\frac{1}{6-sT}} + 2 \right\} \quad (2.5)$$

$$X = \frac{\log \left( \frac{T_a}{mT} \right)}{\log 2} \quad (2.6)$$

Na versão revista publicada em 05/2015 da Rec. ITU-T G.107 (ITU-T, 2015) do Modelo-E, a sensibilidade de atraso ( $sT$ ) e o mínimo atraso perceptível ( $mT$ ) foram introduzidos. Isto foi necessário, pois com base em diversos testes de conversação, demonstrou-se que os valores de atraso mesmo longos podem não afetar a qualidade de voz percebida, isto é, em termos de atribuição do efeito de atraso do sistema. Em tais casos, as previsões realizadas por versões anteriores do Modelo-E podem ser mais pessimistas do que a opinião real do usuário. Como consequência, em casos específicos, pode-se procurar previsões da qualidade da fala que são mais adequados a alguns requisitos de atraso menos rigorosos. Com uma escolha adequada destes dois parâmetros, ambas as chamadas padrão e menos sensíveis a atrasos podem ser abordadas. É aconselhável usar esses novos parâmetros cuidadosamente de acordo com a classe de atraso esperado de acordo com a Tabela 2.3.

Tabela 2.3: Classes de sensibilidade ao atraso para diferentes casos de uso.

Classe de sensibilidade de atraso	sT	mT (ms)	Casos de uso
Padrão	1	100	Aplicável a todos os tipos de conversas telefônicas. Deve ser usado para: - Operadoras de telefonia fixa ou móvel; - Quando grupos alvo de usuários e as exigências de atraso são desconhecidas
Baixo	0,55	120	Aplicável apenas nos casos em que é conhecido que os usuários têm baixa sensibilidade ao atraso, por exemplo, cenários de conversação não sensíveis ao tempo.
Muito Baixo	0,4	150	Aplicável apenas nos casos em que é conhecido que os usuários têm sensibilidade muito baixa ao atraso, por exemplo, em casos principalmente não-interativos, como ouvir essencialmente uma conversa ou para uma palestra.

O fator de comprometimento atraso,  $I_d$ , no Modelo-E inclui três componentes: (1) o atraso em si e os efeitos de eco entre (2) locutor e (3) ouvinte. Um grande atraso de fim-a-fim é um dos principais problemas de transmissão de voz sobre redes IP. Quando este atraso se torna significativo, longas pausas em conversas ocorrem. Para fornecer

serviços de voz de alta qualidade, a rede VoIP deve ser capaz de garantir baixa latência. A Recomendação ITU-T G.114 (ITU-T, 2003) define o tempo aceitável do atraso de ida e volta entre dois *gateways* VoIP: 150 ms de atraso em um sentido. Atrasos entre 150 ms e 400 ms podem tornar as conversas possíveis, mas a demora excessiva torna-se irritante. Os atrasos de mais de 400 ms são inaceitáveis para fins de planejamento de rede geral. A Modelo-E tradicional de banda estreita utiliza uma equação complexa para se calcular o efeito do atraso na qualidade da voz (MYAKOTNYKH; THOMPSON, 2009).

### 2.8.4 Fator de perdas associadas ao equipamento, $I_{e-eff}$

Os valores para o Fator de Comprometimento do Equipamento ( $I_e$ ) usando codecs de baixa taxa de bit não estão relacionados com outros parâmetros de entrada. Eles dependem de resultados médios de testes de opinião de avaliação subjetivas sob diversas taxas de perdas de pacotes.

O Fator de Comprometimento Eficaz de Equipamento a perdas de pacotes,  $I_{e-eff}$ , depende do fator de comprometimento equipamentos  $I_e$  e da taxa de perda de pacotes e eficácia dos algoritmos de ocultação de perda de pacotes (*Packet Loss Concealment - PLC*) do codec em uso. A perda de pacotes inclui a perda causada pela rede e também pela variação excessiva ao atraso. Pacotes perdidos são restaurados usando-se mecanismos de ocultação de perda de pacotes utilizados por diferentes codecs. A Equação 2.7 assume que a perda de pacotes é aleatória e em rajadas (*Burst*).

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}}, \quad (2.7)$$

em que:

- $P_{pl}$  é a taxa de perda de pacotes como percebido pelo aplicativo no receptor, que inclui a perda de pacotes de rede ( $P_{pl_{net}}$ ) e perda de descarte no buffer de de-jitter ( $P_{pl_{buffer}}$ );
- $B_{pl}$  é a robustez a perda de pacotes (que é a eficácia de algoritmos específicos do codec a ocultação de perda de pacotes);
- $BurstR$  expressa o ambiente de perda de pacotes (*Packet Loss Behavior - PLB*) e é o fator de robustez específico do codec à perda de pacotes também conhecido como coeficiente de ruptura, que é definido como a razão entre o comprimento médio de rajadas observados em uma sequência de chegada para um comprimento médio de rajadas esperados sob perda aleatória. Quando a perda de pacotes é aleatório (ou seja, independente)  $BurstR = 1$ ; e quando a perda de pacotes é em rajadas (ou seja, dependente)  $BurstR > 1$  (ITU-T, 2007).

O Apêndice I da Recomendação ITU-T G.113 (ITU-T, 2007) fornece valores recomendados de  $I_e$  e  $B_{pl}$  para alguns codecs específicos (G.723.1-A, G.729-A, GSM-EFR e G.711) sob condições de perdas aleatórias de pacotes e de perdas em rajada. A qualidade do codec é medida em relação à qualidade do codec G.711 de banda estreita (ou seja, para este codec,  $I_e = 0$ ). Codecs piores têm valores mais elevados de  $I_e$ . Estes

valores para diferentes codecs estão baseados em diferentes resultados experimentais da ITU (MYAKOTNYKH; THOMPSON, 2009).

### 2.8.5 Fator de Vantagem, $A$

Devido ao significado específico do Fator de Vantagem,  $A$ , não existe relação alguma com quaisquer um dos outros parâmetros de transmissão.

A Recomendação Rec.G.107 (ITU-T, 2015) em sua Tabela 2 apresenta os seguintes valores de  $A$ :

- Sistema de comunicação convencional (cabado),  $A = 0$ ;
- Sistema de comunicações móveis em edifício (sem fio),  $A = 5$ ;
- Sistema de comunicações móveis em veículos (sem fio),  $A = 10$ ;
- Sistema de comunicações por satélites,  $A = 20$ .

### 2.8.6 Considerações sobre o Modelo-E

O Modelo-E é um método não invasivo, por não necessitar do sinal original de fala, que utiliza as métricas de rede monitoradas localmente e o fator de comprometimento do equipamento para estimar a qualidade de experiência, de modo que pode ser utilizada para monitoramento durante uma chamada. Um problema com o Modelo-E é que apenas a ITU-T fornece o fator de comprometimento dos equipamentos especificados pela recomendação ITU-T Rec. G.113 (ITU-T, 2007). Para uma série de outros codecs comumente utilizados não especificados pela ITU-T Rec. G.113, o fator de comprometimento do equipamento não é fornecido.

## Capítulo 3

# Trabalhos Relacionados

Estudos mostram que a adaptação de parâmetros durante uma chamada VoIP (por exemplo, codec, empacotamento e redundância) pode melhorar significativamente a qualidade de voz ou até mesmo fazer com que os usuários envolvidos não se apercebam de quaisquer lacunas ou falhas na entrega de fluxo de voz (KARAPANTAZIS; PAVLIDOU, 2009; MANOUSOS *et al.*, 2005; COSTA; NUNES, 2009; QIAO *et al.*, 2004; MYAKOTNYKH; THOMPSON, 2009; VIANNA *et al.*, 2006). No entanto, essas obras muitas vezes se concentram em parametrização e pouco ou nada têm por base os avanços da investigação em sistemas adaptativos (CARVALHO, 2011).

Aktas *et al.* (2012) compararam a qualidade de voz de um conjunto de codecs VoIP padrão dado diferentes condições da rede e propuseram um codec adaptável fim-a-fim mudando o esquema baseado em perda de pacotes, *jitter*, largura de banda disponível como os fatores que definem a condições da rede em determinado momento.

Costa e Nunes (2009) descreveram uma técnica adaptável de comutação de codec que começa a monitorar e analisar a qualidade da voz, mudando para um codec de taxa de transmissão inferiores ou superiores de acordo com valores de limite pré-definidos para cada codec. Os autores propuseram uma solução adaptável para proporcionar Qualidade de Serviço de voz sobre IP de comunicação. Esta solução baseou-se em três componentes que interagiram com o objetivo de alcançar maior qualidade na comunicação de voz. As duas primeiras consistiram em mudar o codec e o protocolo de transporte em tempo real durante uma conversa. O terceiro consistiu na utilização de um mecanismo de correção de erros para se recuperar de perdas de pacotes. Para demonstrar a qualidade de voz obtida por esta solução, um aplicativo cliente de VoIP foi desenvolvido, compatível com outros clientes de VoIP, para implementar a qualidade proposta de algoritmo de serviço e controlar a qualidade de voz durante uma conversa. Os resultados das medições experimentais e simulações realizadas demonstraram que esta solução era viável e aumentava significativamente a qualidade de voz na comunicação VoIP, segundo os autores.

Assem *et al.* (2013a) apresentaram e avaliaram um algoritmo que realizou a seleção do codec de áudio mais adequada, dadas as condições prevaletentes no caminho de rede entre os pontos inicial e final de uma chamada de voz.

Assem *et al.* (2013b) focaram em otimização de QoE no ambiente de VoIP. Diferentes codecs de VoIP apresentaram diferente desempenho e diferente tolerância a perda

de pacotes. Os autores compararam o desempenho de diferentes codecs, identificado diferentes pontos teóricos de comutação de codecs. No entanto, a troca de codecs nesses pontos nem sempre era possível, na prática, por causa sobrecarga em termos da lacuna na substituição de codecs (ou seja, o tempo de resposta ao alternar entre codecs). Os autores apresentaram e avaliaram um algoritmo adaptativo que executava durante uma chamada de voz, a seleção do codec de áudio mais adequado dadas as condições prevalecentes no caminho de rede entre os pontos finais desta chamada de voz. Métricas QoE de áudio existentes foram usadas para avaliar a solução. Os resultados mostraram que o algoritmo de adaptação foi capaz de coincidir com a maior pontuação MOS possível, uma vez que selecionou o melhor codec de áudio disponível para a comutação em determinado ponto da chamada.

Jiang e Huang (2011) introduziram um sistema de monitoramento de qualidade de voz baseado no protocolo SIP, usando estatísticas RTP para obter pontuação MOS por intermédio do Modelo-E simplificados. Kim e Choi (2010) propuseram um método de monitoramento de desempenho de rede que utiliza estatísticas RTCP para monitorar os serviços multimídia como VoIP e IPTV. Silva e Lins Silva e Lins (2006) analisaram os resultados de QoS fornecidos pelo SIP para tráfego de voz por meio da medição do atraso, *jitter* e perdas de pacotes. Carvalho *et al.* (2005) propuseram três correções para o Modelo-E, a fim de dar resultados mais precisos, indicando a QoS esperada para o usuário final; Carvalho *et al.* (2005) também descreveram uma ferramenta de mensuração com base nessas correções. Gong *et al.* (2009) propuseram um modelo pentagrama para medir a QoE com base na integralidade do serviço, capacidade de retenção de serviço, disponibilidade do serviço, instantaneidade do serviço e usabilidade do serviço. As experiências mostraram que o modelo proposto por Gong *et al.* (2009) ofereceram um melhor desempenho em termos de medição da qualidade do que o método convencional. Devido também a falta de sistemas de monitoramento da QoE, Hershey *et al.* (2009) propuseram uma abordagem que agregou observações de aplicações em tempo real que funcionaram em sistemas corporativos centrados em redes (*net-centric*) e os autores apresentaram seus resultados em vários cenários VoIP, incluindo um evento Negação de Serviços (*Denial-of-Service*) causando atraso perceptível da aplicação.

Jammeh *et al.* (2012) apresentaram um esquema de adaptação QoE orientada a voz e vídeo sobre IP com o objetivo de resolver o problema de otimização para fornecer QoE ideal para voz em rede e aplicações de vídeo. A arquitetura de adaptação VoIP foi implementada e testada tanto no NS-2 e em uma rede aberta de serviços de mídia para Internet (*Internet Media Services - IMS*) para permitir a simulação extensa e um banco de ensaio de avaliação. Os resultados mostraram que o esquema era perfeitamente sensível a largura de banda disponível e congestionamento para voz e vídeo e otimizado para QoE entregue para diferentes condições de rede, e era amigável ao tráfego TCP.

Enquanto para as aplicações atuais os dispositivos de medição da qualidade da fala em VoIP no mercado são muito caros, Huntgeburth *et al.* (2010) propôs uma solução baseada em componentes de código aberto permitindo a determinação do valor da medida da qualidade de fala (MOS) de acordo com a metodologia de teste de Avaliação

Perceptiva da Qualidade de Fala (*Perceptual Evaluation of Speech Quality* - PESQ).

Assem *et al.* (2013b) descreveram um *framework* de testes que era capaz de fornecer estimativas *on-line* de qualidade da chamada de áudio em caminhos pré-determinados de rede, sem a necessidade de qualquer envolvimento do usuário final ou a disponibilidade prévia de sequências de áudio ou rastreamento de rede. A estrutura incluiu uma ferramenta que emula o tráfego de áudio de chamadas IP e utiliza um Modelo-E estendido para medir a qualidade de áudio. Além disso, Assem *et al.* (2013b) pôde emular deficiências de rede para executar experiências em diferentes condições de rede. Os resultados experimentais apresentados mostraram que as medições de qualidade adquiridas usando o *framework* se aproximaram das medições padrão da indústria mais comumente aplicada para testes *off-line* de avaliação objetiva de voz e vídeo por meio de PESQ e Pico da Relação Sinal-Ruído (*Peak signal-to-noise ratio* - PSNR) respectivamente.

Carvalho (2011) trabalha nas transições das trocas de codecs em um sistema adaptativo, baseado em medidas de QoS, utilizando um softphone próprio e frases em português não padronizadas para os testes, efetuando uma análise externa baseado somente em PESQ ao contrário desta Tese que foca não na adaptação do sistema, mas tem como foco as medidas baseadas em QoE. Além disto, esta Tese utiliza frases em inglês padronizadas, um softphone de código livre e valida o sistema não só com PESQ mas também com o POLQA.

Mohamed (2013) abordou as comunicações em tempo real usando a tecnologia Voice-Over-IP (VoIP). O foco foi a avaliação da Qualidade de Experiência (QoE) do ponto de vista do usuário final e o desenvolvimento de algoritmos e técnicas para melhorar a QoE geral. Uma das principais contribuições foi uma ferramenta de teste genérico que pôde ser usada para qualquer aplicação de voz e vídeo sobre IP (VVoIP) em qualquer ambiente. Utilizando técnicas de emulação de rede para fornecer estimativas para a qualidade de voz e vídeo percebida em diferentes caminhos da rede, a ferramenta funcionou sem a necessidade de utilização de técnicas de avaliação de qualidade tradicionais, que são conhecidos pelo tempo e os recursos que consomem, exigindo o envolvimento do usuário final para coletar sequências de áudio / vídeo e rastreamentos de rede. A ferramenta simulou o trace de áudio e vídeo e empregou o Modelo-E de opinião de qualidade para estimar qualidade de áudio e vídeo, respectivamente, com a vantagem de emular várias condições da rede para executar experiências em vários cenários. Além disso, apresentou um algoritmo adaptativo genérico para mudança de codecs de áudio em todo o andamento da chamada. Codecs são conhecidos por terem comportamentos diferentes em diferentes condições de rede; Foi apresentado o comportamento de cinco dos codecs mais utilizados (incluindo alguns codecs não ITU), derivando modelos para eles, para que o Modelo-E pudesse ser utilizado para avaliar a Qualidade de Experiência. Além disso, foi analisada, da perspectiva do usuário final, a mudança de codec, de forma que esse impacto negativo pudesse ser minimizado tanto quanto possível. O autor descreveu os resultados dos testes do algoritmo em diferentes cenários de rede; estes resultados sugeriram que o algoritmo poderia oferecer melhor Qualidade de Experiência do que deveria ter sido realizada por intermédio do emprego de apenas um codec durante a chamada.

# Capítulo 4

## Materiais e Métodos

### 4.1 Metodologia

Nesse Capítulo são descritos os experimentos realizados com o objetivo de se obter as medidas de QoE, bem como são apresentadas as ferramentas de hardware e software utilizados para o desenvolvimento e a execução dos experimentos.

O método utilizado neste trabalho parte inicialmente do módulo de diagnóstico proposto por Carvalho (CARVALHO, 2011), adaptado conforme Seção 2.8 no Capítulo 2. O módulo de diagnóstico é composto por agentes de monitoramento e análise. Ele usa os relatórios RTCP XR (*Control Protocol Extended Reports - RTP*) (FRIEDMAN *et al.*, 2003). Estes relatórios contêm informações sobre a qualidade instantânea de uma chamada VoIP, tais como perda, atraso e ruído ambiente, fator de comprometimento do codec. Tudo isso foi utilizado na função que calcula o Fator R baseado no Modelo-E para o cálculo da qualidade instantânea da chamada.

### 4.2 Protocolo VoIP

Existe uma série de protocolos em uso no mercado para a comunicação entre clientes (por exemplo: SIP, H.323) e para a sinalização das chamadas VoIP (por exemplo, RTP e RTCP) cada um com algumas particularidades. Para a escolha do protocolo VoIP a ser utilizado, foram considerados maturidade da tecnologia, disseminação, facilidade de uso, disponibilidade de ferramentas existentes e o conhecimento do autor nesta tecnologia. Diante destes fatores, adotou-se o Protocolo de Iniciação de Sessão (*Session Initiation Protocol - SIP*). O SIP é um protocolo da camada de aplicação para a sinalização do tráfego multimídia. Funciona independentemente do tipo de mídia e de aplicação utilizados. Segue um modelo de requisições e respostas, especificado pelo documento de Pedido para Comentários (*Request For Comments - RFC*) 3261 de 2003 (ROSENBERG *et al.*, 2003). Sua principal característica, além da independência da camada de transporte, é a simplicidade.

### 4.3 Agente VoIP

O agente é um módulo que fornece a estrutura SIP necessária ao projeto. Sua função, além de encaminhar o áudio na estação transmissora e gravar o áudio recebido na estação receptora é enviar, receber e tratar toda a sinalização do protocolo SIP.

Para cumprir essas funcionalidades, o agente possui um serviço que além de realizar as chamadas na máquina transmissora, aguarda pelo recebimento de novas chamadas na máquina receptora. Este serve como autenticação no servidor SIP e está especificado na RFC do protocolo número 3261 (ROSENBERG *et al.*, 2003; ROSENBERG, 10/2002).

O PjSIP é uma biblioteca Open Source de comunicação multimídia e de código aberto escrito na linguagem C que implementa os protocolos baseados em padrões como SIP, SDP, RTP, entre outros (PRIJONO; ISMANGIL, 2003), disponibilizando as principais operações necessárias para um aplicativo que utilize o protocolo SIP para realizar chamadas VoIP. Possui como vantagem o desempenho considerável e grande portabilidade, uma vez que está disponível para várias arquiteturas.

Já o PJSUA, também utilizado, é o *User Agent* para VoIP que utiliza a pilha implementada pelo PJSIP. Possui funcionalidades mais complexas do que o PJSIP como gerenciamento de contas e fluxo de chamadas, alcançando uma abstração maior das camadas de baixo nível.

### 4.4 Biblioteca Intel

As primitivas integradas de desenvolvimento Intel IPP (*Integrated Performance Primitives*) é uma extensa biblioteca de funções de software que auxiliam o desenvolvimento para processamento de dados e aplicações de comunicações.

Estas funções estão prontas para uso, livre de royalties e são altamente otimizadas usando as extensões de streaming Intel® SSE (*Streaming SIMD Extensions*) e um conjunto de extensões de instruções avançadas Intel® (Intel® AVX, Intel® AVX2), que muitas vezes superam o que um compilador otimizado pode produzir sozinho segundo a própria Intel (INTEL, 2012).

### 4.5 Sistemas Operacionais Utilizados

Para a realização dos experimentos foi utilizado o ambiente Linux. Foi empregada a distribuição Ubuntu 10.04, por questão de compatibilidade com o *softphone* utilizado, já explicado nas seções anteriores. As rotinas de operação do *softphone* de seleção de arquivo, perdas e codec foram feitas em Pearl. Para a realização da análise dos resultados e escrita do texto foi utilizado o ambiente Windows 7, uma vez que a ferramenta PEXQ possui somente versão para este sistema operacional. Na fase de análise, as rotinas desenvolvidas foram na linguagem Pearl e R (R Core Team, 2015).

### 4.5.1 Ferramentas Auxiliares

Neste trabalho além do ambiente apresentado para a obtenção dos dados, foram necessárias ferramentas auxiliares para extração dos dados como a Format Factory (HAO, 2007–2014), Direct Wave MP3 Splitter (SOFTWARE, 2004-2016) e PEXQ (OPTICOM, 2007).

A ferramenta Format Factory (HAO, 2007–2014) é um conversor de multimídia multi-funcional. Provê funcionalidades para converter os seguintes tipos de vídeo do tipo MP4, 3GP, MPG, AVI, WMV, FLV, SWF. Ela é capaz também de converter os seguintes tipos de áudio MP3, WMA, AMR, OGG, AAC e WAV além de converter os seguintes tipos de imagens JPG, BMP, PNG, TIF, ICO, GIF, TGA. Nesta Tese utilizou-se apenas o recurso de conversão de áudio. Esta conversão de áudio foi necessária, pois os arquivos resultantes no ambiente experimental, embora no formato WAV, não puderam ser lidos pelo ambiente Windows nas máquinas de análise. Assim, uma conversão de WAV (Linux) para WAV (Windows) foi necessária. Essa conversão apenas ocorre no cabeçalho do arquivo resultante e os dados de forma de onda permaneceram inalterados.

Direct WAV MP3 Splitter (SOFTWARE, 2004-2016) possibilita fatiar/dividir qualquer arquivo MP3 ou WAV em pedaços menores, tornando os arquivos resultantes mais gerenciáveis. Permite também remover trechos de silêncio ou outras partes indesejadas. Direct WAV MP3 Splitter prove o fatiamento de áudio e não faz qualquer mudança na qualidade ou formato. Durante este trabalho, esta ferramenta foi utilizada para efetuar o fatiamento de arquivos provenientes do ambiente experimental. Ao final, os arquivos resultantes foram comparados aos originais e provou-se que as fatias obtidas estão exatamente iguais aos trechos fatiados, exceto pelo acréscimo de um cabeçalho WAV.

PEXQ (OPTICOM, 2007) é uma ferramenta tudo em um, um suíte no estado da arte para testes de medição de qualidade perceptual. PEXQ possui uma série de algoritmos para a medição objetiva da qualidade percebida de canais de comunicação e um conjunto de software para usá-los. Em particular, os algoritmos de medição são:

- PEAQ (*Perceptual Evaluation of Audio Quality*) para medidas de qualidade de áudio, por exemplo para codecs sem perdas;
- PEVQ (*Perceptual Evaluation of Video Quality*) para algoritmos de vídeo;
- PESQ (*Perceptual Evaluation of Speech Quality*) para transmissão de sinais da fala;
- POLQA (*Perceptual Objective Listening Quality Assessment*) para medidas de qualidade de sinais da fala; e
- PEDQ (*Perceptual Evaluation of Data-Download Quality*) para transferência de dados.

Nesta Tese foram utilizados apenas os algoritmos para análise da qualidade de sinais da fala (PESQ e POLQA).

## 4.6 Montagem do cenário

## 4.7 Cenário de Teste

Para a montagem do cenário de teste foram utilizadas máquinas de 64 bits com sistema operacional Ubuntu 10.04 (versões mais recentes revelaram-se ineficientes nos testes preliminares e incompatíveis com o Intel IPP 7.0.4, a biblioteca de codec utilizada nos experimentos) para implementar o PJSIP 1.10. O fator limitante para este cenário foi a biblioteca Intel IPP 7.0.1 (INTEL, 2012), como citada anteriormente em 4.4.

Nesta Tese, tem-se inicialmente uma máquina (transmissor) diretamente ligada à outra máquina (receptor) como prova de conceito para os testes locais e condições controladas. O objetivo foi verificar e validar o mecanismo de medida interno de Qualidade de Experiência (QoE) baseado no Modelo-E, além da formação da base do conhecimento.

Em todas as fases do experimento registrou-se arquivos de log de transmissão (arquivo enviado) e log de recepção (arquivos recebidos) para análise no Capítulo 5. Foi transmitido o mesmo arquivo de áudio dez vezes em diferentes situações de perdas de dados (0,0%, 1,0%, 2,0%, 2,5%, 3,0%, 5,0%, 7,5%, 10,0%) e testado vinte e quatro arquivos diferentes de fala, com o objetivo de se comparar as medições dos sistemas (medições internas) e o resultado das medições globais (POLQA e PESQ) apresentadas pelo software PEXQ.

## 4.8 Áudios Selecionados

Sentenças de Harvard são um conjunto de exemplos de frases que são utilizados para testes padronizados de Voz sobre IP, celular e outros sistemas de telefonia. Eles são sentenças foneticamente equilibradas que usam fonemas específicos com a mesma frequência que eles aparecem na língua inglesa.

Práticas recomendadas pelo IEEE para medições de qualidade de voz (ROTHAUSER *et al.*, 1969) estabelecem 72 listas de 10 frases, descritas como o "1965 Lista Revisada de sentenças foneticamente equilibradas (frases de Harvard)". Eles são amplamente utilizados em pesquisas sobre telecomunicações, fala e acústica, onde são necessárias sequências padronizadas e repetitivas do discurso. O *Open Speech Repository* (PROJECT, 2015) fornece alguns arquivos WAV pré gravadas, livremente utilizáveis de Sentenças de Harvard em Inglês americanos e britânicos, com vozes masculinas e femininas. Neste trabalho foram utilizados todos os 24 arquivos de teste com frases em Inglês americano de ambos os sexos disponíveis na base OSR.

A Tabela 4.1 apresenta os nomes dos arquivos de áudio de testes e o número do conjunto de sentenças, cuja transcrição dos respectivos conteúdos estão presentes no Anexo A. Essa transcrição dos arquivos de áudio é particularmente útil, quando se aplica codificadores mais sensíveis a perdas de pacotes tornando muitas vezes o sinal de fala no receptor incompreensível.

Tabela 4.1: Arquivos de áudio utilizados nos testes.

Ord.	Arquivo OSR (.wav)	Voz	Sent. Harv.	Duração (s)	Parâmetro	
					FF	TT
0	OSR_us_000_0010_8k	Fem. 1	1 A.1	33,623	10	34
1	OSR_us_000_0011_8k	Fem. 1	2 A.2	32,784	11	33
2	OSR_us_000_0012_8k	Fem. 1	3 A.3	33,268	12	34
3	OSR_us_000_0013_8k	Fem. 1	4 A.4	32,542	13	33
4	OSR_us_000_0014_8k	Fem. 1	5 A.5	31,122	14	32
5	OSR_us_000_0015_8k	Fem. 2	1 A.6	41,424	15	42
6	OSR_us_000_0016_8k	Fem. 2	2 A.7	35,128	16	36
7	OSR_us_000_0017_8k	Fem. 2	3 A.8	35,505	17	36
8	OSR_us_000_0018_8k	Fem. 2	4 A.9	34,578	18	35
9	OSR_us_000_0019_8k	Fem. 2	5 A.10	32,312	19	33
10	OSR_us_000_0030_8k	Masc.	57 A.11	46,927	30	47
11	OSR_us_000_0031_8k	Masc.	59 A.12	42,129	31	43
12	OSR_us_000_0032_8k	Masc.	23 A.13	52,484	32	52
13	OSR_us_000_0034_8k	Masc.	72 A.14	35,547	34	36
14	OSR_us_000_0035_8k	Masc.	42 A.15	38,214	35	39
15	OSR_us_000_0036_8k	Masc.	31 A.16	45,086	36	45
16	OSR_us_000_0037_8k	Masc.	32 A.17	38,877	37	39
17	OSR_us_000_0038_8k	Masc.	33 A.18	39,287	38	39
18	OSR_us_000_0039_8k	Masc.	34 A.19	39,846	39	40
19	OSR_us_000_0040_8k	Masc.	35 A.20	47,548	40	48
20	OSR_us_000_0057_8k	Masc.	25 A.21	39,259	57	40
21	OSR_us_000_0059_8k	Masc.	27 A.22	55,712	59	56
22	OSR_us_000_0060_8k	Masc.	28 A.23	58,257	60	59
23	OSR_us_000_0061_8k	Masc.	30 A.24	53,591	61	54

Nota 1: As transcrições estão disponíveis no Anexo A.

Nota 2: Arquivos disponíveis em (PROJECT, 2015).

A coluna Ord., na Tabela 4.1, representa a ordem da codificação utilizada nas planilhas resumo dos resultados obtidos.

#### 4.8.1 Nomenclatura dos arquivos

Conforme Figura 4.1, foram utilizadas duas máquinas no sistema, denominadas respectivamente de transmissora e receptora. O sistema proposto foi alimentado pelos arquivos da base de dados OSR conforme Tabela 4.1.

As saídas do sistema possuem as seguintes nomenclaturas baseados nos seguintes fatores:

1. nome do arquivo de entrada;
2. codec utilizado na transmissão;
3. taxa de perdas de pacotes em %; e

4. número da amostra dentro de um ciclo de repetições para uma dada configuração.

Os codecs selecionados para os testes pertencem a biblioteca IPP e estão presentes na Tabela 4.2. Foi adotado como critério de seleção, os codecs que possuíam os valores dos parâmetros  $I_e$  e  $B_{pl}$  conhecidos conforme a Rec.ITU-T G.113(ITU-T, 2007).

Tabela 4.2: Codecs utilizados nos experimentos.

Ord.	Nome	Parâmetro CCC	enum codec.h	$I_e$	$B_{PL}$
0	PCMU	PCMU	0	0,0	25,1
1	GSM	GSM	3	5,0	10,0
2	G.723	G723	4	15,0	16,1
3	G.729	G729	18	11,0	19,0
4	G.726-32	G726-32	110	7,0	23,0

Fonte: Adaptado de Recomendação ITU-T G.113 (ITU-T, 2007)

As perdas dos testes selecionadas são apresentadas na Tabela 4.3.

Tabela 4.3: Perdas de Pacotes utilizados nos experimentos.

Ord.	Perda	Parâmetro PPP
0	0,0 %	000
1	1,0 %	010
2	2,0 %	020
3	2,5 %	025
4	3,0 %	030
5	5,0 %	050
6	7,5 %	075
7	10,0 %	100

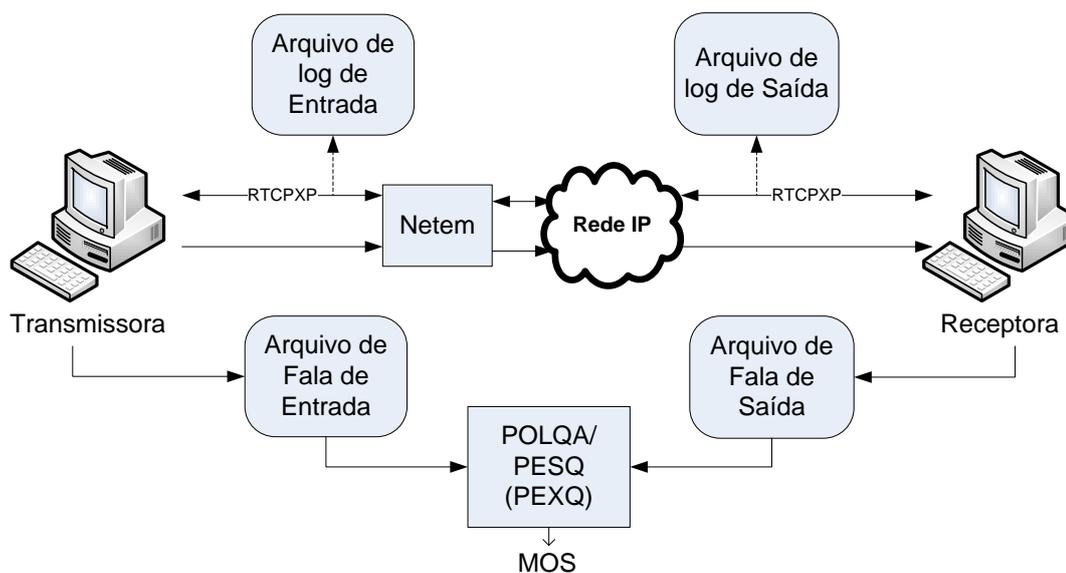


Figura 4.1: Cenário de teste.

O Sistema proposto foi alimentado pelos arquivos da base de dados OSR apresentados na Tabela 4.1.

O arquivo de entrada possui a seguinte nomenclatura: `OSR_us_000_00FF_8k.wav`, onde FF é um parâmetro que representa o nome do arquivo conforme coluna FF da Tabela 4.1.

Após a transmissão do arquivo de entrada pela máquina transmissora (*caller*) para a máquina receptora (*listener*) obtêm-se os seguintes arquivos resultantes:

`OSR_us_000_00FF_8k-cCCC-pPPP-M-A.TTT`, onde os parâmetros/siglas representam:

- *FF*: o nome do arquivo conforme coluna FF da Tabela 4.1;
- *CCC*: codificador utilizado na transmissão conforme coluna CCC da Tabela 4.2;
- *PPP*: perda de pacotes conforme coluna PPP da Tabela 4.3;
- *M*: máquina transmissora (c) ou máquina receptora (l);
- *A*: número do teste da série de amostras entre 0 e 4 para 5 repetições;
- *TTT*: tipo de arquivo podendo ser wav para arquivos de áudio ou txt para arquivos de log.

Os arquivos foram armazenados nas seguintes pastas:

- *origem*: arquivos baixados do site OSR e transmitidos pela máquina transmissora (*caller*);
- *loteFF00*: Testes efetuados com os arquivos FF;
- *00\_original*: arquivos wav gravados na máquina transmissora (*caller*) que opera em Linux após chamada;
- *01\_originalFF*: arquivos wav convertidos de Linux para padrão Windows pelo ferramenta Format Factory;
- *02\_sliceFF*: arquivos wav convertidos e depois fatiados em diversos arquivos pela ferramenta Direct Wave MP3 Splitter;
- *03\_caller*: arquivos de log na máquina transmissora;
- *04\_saida*: arquivos de log na máquina receptora;
- *05\_saidaSlice*: arquivos de log na máquina receptora fatiados de modo que cada fatia represente um teste.

## 4.9 Instalação e Configuração do Ambiente

Nesta Seção descrevem-se os procedimentos para a configuração do ambiente de trabalho. Inicialmente foi necessário preparar as máquinas de 64 bits, instalando o sistema operacional Ubuntu 10.04. Versões posteriores ou anteriores são incompatíveis com a biblioteca Intel IPP 7.0.4. Optou-se por utilizar a biblioteca Intel IPP 7.0.4, por essa ser livre de pagamento de Royalties. Foi baixada também a versão do PJSIP 1.10. Esta versão é compatível com a IPP 7.0.4, compila corretamente e não apresentam *warnings*. Qualquer versão diferente de software do que foi apresentado, o sistema não funcionava corretamente. Antes de compilar o PJSIP junto com biblioteca Intel, foi necessário acrescentar o código do Modelo-E modificado e ajustar as variáveis do ambiente, como habilitar medidas de qualidade, espaço entre medidas, etc. Estas modificações estão descritas na Subseção 4.9.2 e Subseção 4.9.6 .

Para instalação e uso correto do PJSIP instalou-se primeiro a biblioteca IPP da Intel, a qual contém uma variedade maior de codecs (Arquivo IPPSamples) que podem ser utilizados pelo PJSIP. Para instalação das ferramentas referidas anteriormente utilizou-se uma máquina rodando Sistema Operacional Linux 64 bits, pois os arquivos disponíveis só funcionam em sistemas operacionais de 64 bits.

Outra consideração a ser feita é quanto à compilação dos arquivos. O PJSIP usa preferencialmente o C++, portanto, a versão mais atual do g++ foi necessária antes de executar os comandos de instalação.

### 4.9.1 Instalação do IPP

Para a instalação do IPP Intel, após a instalação do Ubuntu 10.04, foram necessários os seguintes passos:

1. Descompactou-se a IPP em uma pasta, conforme o exemplo.  
Ex.: # /home/usuario/l\_ipp.7.0.4.220\_intel64/
2. Executou-se o comando: # ./install.sh
3. As instruções apresentadas foram seguidas. Utilize-se o número serial fornecido.

### 4.9.2 Exportação das variáveis de ambiente

Foi necessário exportar as variáveis de ambiente por meio dos seguintes comandos:

```
# export LD_LIBRARY_PATH=/opt/intel/composerxe-2011.5.220/compiler/lib/intel64/  
/opt/intel/composerxe-2011.5.220/ipp/lib/intel64  
# export IPPROOT=/opt/intel/composerxe-2011.5.220/ipp  
# export IPPSAMPLES=/home/usuario/ipp-samples
```

### 4.9.3 Instalação do IPP Codecs (Samples)

Foi necessária a instalação dos componentes de amostras dos codecs. Sem estes componentes, o sistema não funcionou.

1. Descompactou-se a IPP Codecs em uma pasta conforme o exemplo.

Ex.: # /home/usuario/ipp-samples

2. Selecionou-se na seguinte pasta: # `ipp-samples/speech-codecs`
3. Executou-se # `./build.intel64.sh`
4. As instruções apresentadas foram seguidas.

#### 4.9.4 Instalação do PJSIP

Para a Instalação do PJSIP os seguintes passos foram necessários:

1. Descompactou-se o PJSIP em uma pasta conforme exemplo.

Ex.: # /home/usuario/

2. Modificou-se o arquivo "`config_site_sample.h`":

```
# gedit /home/usuario/pjproject-1.10/pjlib/include/pj/config-site.h
```

Acrescentou-se a linha neste arquivo: `#define PJMEDIA_HAS_INTEL_IPP 1`

O arquivo foi salvo e fechado.

3. Executou-se os seguintes comandos na pasta `/home/usuario/pjproject-1.10`

```
# ./configure --enable-ipp --with-ipp=$IPPROOT --with-ipp-samples=$IPPSAMPLES
```

```
# make dep
```

```
# make
```

#### 4.9.5 Verificação

1. Estando na pasta `pjproject-1.10`, foi executado o seguinte comando:

```
# ./pjsip-apps/bin/pjsua-x86z-64-unknown-linux-gnu
```

2. Seguindo os passos anteriores, apareceu um menu de opções em modo texto (PJSUA) como apresentado na Tabela 4.4.
3. Caso não aparecesse um menu de opções em modo texto (PJSUA) como apresentado na Tabela 4.4, era necessário verificar se o `pjsua` foi criado com um nome diferente de:

```
pjsua-x86_64-unknown-linux-gnu.
```

Tabela 4.4: Tela do PJSUA se tudo estiver funcionando corretamente.

Call Commands:	Buddy, IM & Presence:	Account:
m Make new call	+b Add new buddy	+a Add new acct
M Make multiple calls	-b Delete buddy	-a Delete acct.
a Answer call	i Send IM	!a Modify acct.
h Hangup call (ha=all)	s Subscribe presence	rr (Re-)register
H Hold call	u Unsubscribe presence	ru Unregister
v re-inVite (release hold)	t ToGgle Online status	> Cycle next ac.
U send UPDATE	T Set online status	< Cycle prev ac.
,  Select next/prev call		
x Xfer call	Media Commands:	Status & Config:
X Xfer with Replaces	cl List ports	d Dump status
# Send RFC 2833 DTMF	cc Connect port	dd Dump detailed
Send DTMF with INFO	cd Disconnect port	dc Dump config
dq Dump curr. call quality	V Adjust audio Volume	f Save config
S Send arbitrary REQUEST	Cp Codec priorities	f Save config
q QUIT L ReLoad sleep MS echo [0—1—txt]		n: detect NAT type

You have 0 active call

>>>

#### 4.9.6 Modificações

Para emitir relatórios RTCP XR, foi necessário editar o arquivo de configurações do *pjmedia*:

```
# gedit /home/usuario/pjproject-1.10/pjmedia/include/pjmedia/config.h
```

Editou-se as seguintes linhas no arquivo de configurações do *pjmedia*:

```
define PJMEDIA_STREAM_ENABLE_XR 1
define PJMEDIA_HAS_RTCP_XR 1
define PJMEDIA_RTCP_INTERVAL 1000 /* msec*/
```

#### 4.9.7 Utilização

Para executar as chamadas, utilizaram-se duas máquinas. A máquina transmissora (*caller*) foi utilizada para criar a chamada e a máquina receptora (*listener*) para receber a chamada. O nome do *softphone* implementado pelo PJSIP denomina-se PJSUA (*user agent* - UA). A interface do PJSUA, apresentada na Tabela 4.4, foi exibida quando o PJSUA estava instalado corretamente.

Executou-se primeiro na máquina receptora (*listener*) o seguinte comando:

```
# ./pjsip-apps/bin/pjsua-x86.64-unknown-linux-gnu --auto-answer=200 --auto-rec --rec-file=
$FILEROOT-1-$i.wav --duration=$DURACAOL --clock-rate=8000 --play-file=$PLAYFILE.wav --auto-play
```

```
--log-file=$FILEROOT-l-$i.txt --log-level=5 --quality=9;
```

Em seguida, na máquina transmissora *caller* executou-se o seguinte comando:

```
# ./pjsip-apps/bin/pjsua-x86.64-unknown-linux-gnu --play-file=$PLAYFILE.wav --auto-play
--duration=$DURACAO --log-file=$FILEROOT-c-$i.txt --log-level=9 --add-codec=$CODEC sip:$IPADDR
--quality=9;
```

em que:

```
pjsua [opcoes] [SIP URL a ser chamada]
```

- `--play-file=$PLAYFILE.wav` Nome do arquivo (`$PLAYFILE.wav`) a ser reproduzido durante a chamada;
- `--auto-play` reproduz automaticamente o arquivo apontado em `play file` (Apenas para as chamadas recebidas)
- `--duration=$DURACAO` Duração máxima da chamada em segundo. Setado para a duração do arquivo reproduzido conforme coluna TT da Tabela 4.1.
- `--log-file=FILE` Salva todos os eventos de log no arquivo FILE;
- `--log-level=5` Definir o nível máximo de detalhamento (5) do registro de log.
- `--add-codec=$CODEC` Definir o nome do codec com maior prioridade de uso. O `$CODEC` é o nome padrão do codec conforme especificado pelas RFCs relevantes. Os nomes utilizados neste trabalho estão descritos na Tabela 4.2 coluna Paramêtro CCC.
- `sip:$IPADDR` Especifica o endereço (`$IPADDR`) destino da chamada;
- `--quality=N` Especifica a qualidade de mídia.

Na máquina transmissora (*listener*) o *softphone* respondia automaticamente qualquer chamada de entrada (*auto-answer*), gravando-a em um arquivo chamado "`$FILEROOT-l-$i.wav`". Ele também reproduzia o arquivo "`$PLAYFILE.wav`" e registrava os eventos e as estatísticas da chamada em um arquivo de log "`$FILEROOT-l-$i.txt`", com nível de detalhamento 9.

O *softphone* na máquina receptora (*caller*) efetuou uma chamada para o outro *softphone* localizado no número SIP "`numero_sip`" que é o endereço na máquina *listener*. Após a conexão efetuada, o *softphone* na máquina Receptora (*caller*) reproduzia o arquivo "`$PLAYFILE.wav`" durante `$DURACAO` segundos que era a duração da chamada. O tempo de reprodução do arquivo (duração da chamada) deveria ser o mais próximo possível da duração real do arquivo a ser reproduzido, truncado para o valor inteiro superior. Isto foi necessário devido à fase de análise dos resultados. O arquivo de referência deveria ter a duração do arquivo enviado. Se a duração do arquivo enviado fosse menor do que a duração do arquivo de referência, este seria preenchido com

silêncio e a ferramenta de análise **POLQA** apresentaria uma resposta incorreta, sempre próximo de 1, indicando uma qualidade extremamente ruim. Por outro lado, se a duração da chamada fosse maior do que o tempo do arquivo transmitido, o arquivo de entrada seria reproduzido novamente até atingir a duração programada. Novamente, a análise na ferramenta **POLQA** seria prejudicada. Se o arquivo de entrada apresentasse duração maior do que TT s, este era interrompido. Se fosse menor, era repetido até esgotar os TT s. A chamada era realizada preferencialmente com o codec \$CODEC e foi registrada as estatísticas da chamada em um arquivo de log "\$FILEROOT-l-\$i.txt", com nível de detalhamento 9. O detalhamento de cada opção do PJSUA está disponível em (LTD, 2007).

Os testes foram compostos por 24 arquivos da Tabela de Harvard conforme apresentado na Tabela 4.1. Embora existam mais de 72 frases, foram selecionadas apenas as frases presentes no repositório OSR. Cada arquivo possui duração de S.MMM segundos conforme coluna Duração da Tabela 4.1. Como o PJSIP necessita de um parâmetro de duração inteiro para segundos, a duração de transmissão era truncada para o inteiro imediatamente superior.

## 4.10 Realização dos Experimentos

A realização dos experimentos em si, após o cenário implementado, seguiu o fluxo-grama descrito na Figura 4.2. Não houve preparação dos arquivos de áudio selecionados. Eles foram utilizados da maneira que se encontram na internet. A codificação foi realizada durante a fase de transmissão. Ambas as máquinas, transmissora (*caller*) e receptora (*listener*) possuíam os codecs selecionados. As perdas foram programadas na máquina transmissora (*caller*) configurando o software Netem (FOUNDATION, 2009). Os valores utilizados nos parâmetros foram os apresentados nas Tabelas 4.3 e 4.2.

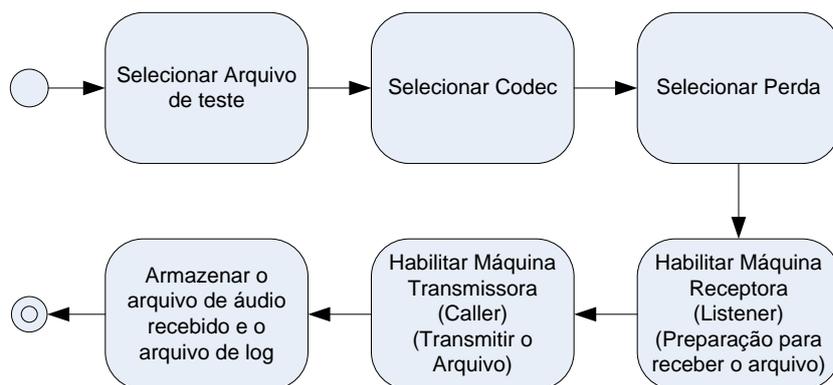


Figura 4.2: Fluxo de teste.

## 4.11 Teste Piloto

Um teste piloto com todo o fluxo de metodologia foi realizado como prova de conceito. O fluxo de metodologia era composto de um arquivo (OSR\_us\_000\_0010\_8k.wav), um codec (PCMU), 5 tipos diferentes de perdas (0%, 2,5%, 5%, 7,5% e 10%), 5 repetições de cada situação e a análise do algoritmo PESQ foi realizada como prova de conceito. Os resultados obtidos neste teste foram publicados em (SILVA *et al.*, 2015b). Posteriormente o teste foi expandido e a análise foi realizada com o algoritmo POLQA (SILVA *et al.*, 2015a). A aplicação dos resultados em uma proposta de modelo de Sistema Adaptativo foi publicada em (SILVA *et al.*, 2015c).

## 4.12 Teste de Massa

Após a validação do método de Teste proposto, um teste com cada um dos arquivos da base OSR de dados foi conduzido. Os testes de massa foram realizados com os 24 arquivos de testes disponíveis no site da OSR (PROJECT, 2015). As características dos arquivos de testes estão apresentadas no Capítulo 4, Tabela 4.1.

Para cada arquivo em teste, cinco codecs da base ITU-T foram utilizados (PCMU, GSM, G.723, G.729, G.726-32), com uma combinação maior de Taxa de Perdas (0%, 1%, 2%, 2,5%, 3%, 5%, 7,5% e 10%) e um número maior de repetições de 10 vezes para cada combinação. Os resultados do teste estão presentes no Capítulo 5, Seção 5.2.

Todas as conversações entre máquina transmissora e máquina receptora foram realizadas em tempo real. A duração de cada transmissão de cada arquivo está representado na coluna Parâmetro TT da Tabela 4.1 na página 50. Entre cada transmissão, um pausa de 25 segundos foi inserida. No total, para uma situação (um codec e uma taxa de perda) contendo todos os 24 arquivos, foram necessários 25 minutos e 26 segundos para se completar o ciclo. Considerando-se todos os cinco codecs, todas as 8 Taxas de Perdas e 10 repetições, foram necessários 203 horas e 28 minutos de transmissões. Essas conversas resultaram em 11.520 arquivos de áudio e 23.040 arquivos de log nas duas máquinas (transmissora e receptora). Os arquivos de áudio resultantes foram analisados pelo PEXQ, resultando em um arquivo com 11.520 registros.

## 4.13 Teste Estatístico

Como passo final da metodologia foi realizada a análise estatística dos resultados dos testes. Os resultados da análise estatística são apresentados no Capítulo 5 Seção 5.2. O objetivo foi a identificação da relação entre os resultados obtidos pelo método proposto (Modelo-E modificado para medida de Qualidade de Experiência) e os resultados da avaliação objetiva dos sinais em teste (PESQ e POLQA), além da significância estatística da diferença entre os valores médios de desempenho das combinações codec-taxa de perdas. Para isso, foi calculado o coeficiente de correlação de Pearson.

O coeficiente de correlação de Pearson, também chamado de "coeficiente de correlação produto-momento ou simplesmente de " $\rho$ , de Pearson" mede o grau da correlação

(e a direção dessa correlação - se positiva ou negativa) entre duas variáveis. Ele quantifica a força de associação linear entre os conjuntos de duas variáveis, e portanto descreve quão bem uma linha reta se ajustaria através de nuvem de pontos.

Este coeficiente, normalmente representado por  $\rho$ , assume apenas valores entre -1 e 1.

- $\rho = 1$  Significa uma correlação perfeita positiva entre as duas variáveis. Isto é, se uma aumenta, a outra sempre aumenta.
- $\rho = -1$  Significa uma correlação perfeita negativa entre as duas variáveis. Isto é, se uma aumenta, a outra sempre diminui.
- $\rho = 0$  Significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir uma dependência não linear. Assim, o resultado  $\rho = 0$ , deve ser investigado por outros meios.

O critério utilizado nesta Tese para interpretação da magnitude dos coeficientes é apresentado na Tabela 4.5 adaptado de Mukaka (2012).

Tabela 4.5: Interpretação do Coeficiente de Correlação de Pearson.

$\rho$	Interpretação
Entre 0,90 e 1,00 (-1,00 e -0,90)	Correlação positiva (negativa) muito forte.
Entre 0,70 e 0,90 (-0,90 e -0,70)	Correlação positiva (negativa) forte.
Entre 0,50 e 0,70 (-0,70 e -0,50)	Correlação positiva (negativa) moderada.
Entre 0,30 e 0,50 (-0,50 e -0,30)	Correlação positiva (negativa) fraca.
Entre 0,00 e 0,30 (-0,30 e 0,00)	Correlação positiva (negativa) desprezível.

Seu cálculo é determinado pela Equação 4.1:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 - \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4.1)$$

onde  $x_1, x_2, \dots, x_n$  e  $y_1, y_2, \dots, y_n$  são os valores medidos de ambas as variáveis. Além disso, a Equação 4.2 e a Equação 4.3:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (4.2)$$

e

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i \quad (4.3)$$

são as médias aritméticas de ambas as variáveis ( $\bar{x}$  e  $\bar{y}$ ) envolvidas.

## 4.14 Medidas

Existem basicamente dois tipos de momentos de medições: (1) Durante uma chamada VoIP, medida ponto a ponto a Qualidade de Experiência a cada intervalo pré-determinado de tempo, chamada de pontual ou instantânea e (2) a média acumulada ao final de uma chamada VoIP, onde a Qualidade de Experiência é expressa por um valor global.

Durante o teste piloto (SILVA *et al.*, 2015a), os resultados das medições realizadas durante uma chamada VoIP em diversos cenários, o procedimento e os resultados obtidos são apresentados de forma resumida nos gráficos no Capítulo 5.

Em uma chamada VoIP, existem duas aplicações das medidas pontuais, que são resultados das medições fornecidos pela rotina do Modelo-E modificada implementada no PJSIP via protocolo RTCP-XR:

- **Desempenho da comunicação:** valor instantâneo da QoE durante um intervalo de análise de chamada VoIP.
- **Controle:** (No caso específico deste estudo, o aplicativo de controle não está habilitado);

Tem-se o Modelo-E (ITU-T Rec. G.107 (ITU-T, 2015)) entre os métodos de medida oportuna de qualidade da fala mais utilizados. O seu procedimento de medição consistiu na coleta de parâmetros de fluxo de voz, que servem como entrada para um conjunto de equações que retornam como resultado o Fator R, cujo valor varia de 0 (pior) a 100 (melhor) como uma expressão da medida da qualidade avaliada. O mapeamento do Fator R (0-100) em MOS (1-5) é representado pelo resultado da Equação 4.4 presente na ITU-T Rec. G.107 (ITU-T, 2015).

$$MOS = \begin{cases} 1 & \text{if } R < 0 \\ 1 + 0.035R + 7 * 10^{(-6)} & \text{if } 0 \geq R \geq 100 \\ 4.5 & \text{if } R > 100 \end{cases} \quad (4.4)$$

Tipicamente, o resultado do Modelo-E é transformado para uma escala de MOS (ITU-T, 1996), como apresentado na Tabela 2.2 do Capítulo 2, na página 31.

Ao final da chamada VoIP, um arquivo (arquivo recebido) de áudio e um arquivo de log para um determinado arquivo transmitido (arquivo enviado) em uma dada situação (Arquivo Original-Codec-Perda) foram gerados. Estes arquivos foram utilizados para se medir a qualidade global da comunicação VoIP durante a fase de análise apresentado no Capítulo 5, e uma pontuação de 1 a 5 foi gerada, incluindo neste caso, um intervalo de confiança.

## Capítulo 5

# Análises dos Resultados

Durante a transmissão do arquivo de fala da máquina transmissora (*caller*) para a máquina receptora (*listener*), uma medida de Qualidade de Experiência (QoE) foi efetuada a cada um segundo pelo Modelo-E. Esta medida periódica foi denominada de medida pontual (pontual ou de instante em instante). Ao final da transmissão, um arquivo contendo o registro com as medidas de Qualidade de Experiência da transmissão foi gerado. A média destas medidas de QoE denominou-se de média global da transmissão. Esta média global foi comparada com os resultados obtidos ao efetuar-se as medidas de Qualidade pelos algoritmos POLQA e PESQ utilizando-se o sinal original transmitido pela Máquina Transmissora (*caller*) como referência e o sinal registrado recebido pela Máquina Receptora (*listener*) como sinal em teste.

São apresentados na Seção 5.1 os resultados do Teste Piloto realizado somente com um arquivo, 5 repetições deste arquivo, um único codec(PCM-U-G.711), perdas de 0%, 2,5%, 5%, 7,5% e 10%. Esses resultados estão presentes no artigo (SILVA *et al.*, 2015a).

Na Seção 5.2 são apresentados os resultados dos testes exaustivos formadores do Base de Dados do conhecimento, objetivo desta Tese.

### 5.1 Teste Piloto

O objetivo deste teste piloto foi validar todo o processo de medição e comparação. Especificamente para o teste piloto, foi utilizado um arquivo de áudio relacionado com a base *Open Speech Repository* (OSR) (PROJECT, 2015). As Figuras 5.1, 5.2 e 5.3 estão relacionadas com o arquivo enviado, `osr_us_000_0010_8k.wav`, formato PCM 16bits, taxa de amostragem de 8 kHz, voz feminina da OSR.

A fim de emular a perda de pacotes em nosso cenário de rede, utilizou-se o software Netem (*Network Emulation*) (FOUNDATION, 2009), configurado na máquina transmissora.

Medidas pontuais (contínuas) foram tomadas durante o intervalo aleatório entre 0,5 e 1,5 segundo. À medida que as perdas de pacotes aumentam, a qualidade MOS pontual tornou-se mais instável, tal como apresentado na Figura 5.1.

Isso ocorre devido aos pacotes RTP de dados que trafegam em uma rede não conseguem chegar ao seu destino, ou seja, a estação receptora.

As Figuras 5.1, 5.2 e 5.3 apresentam o MOS no tempo considerando alguma deficiência. Eles mostram que as deficiências resultam na redução de MOS.

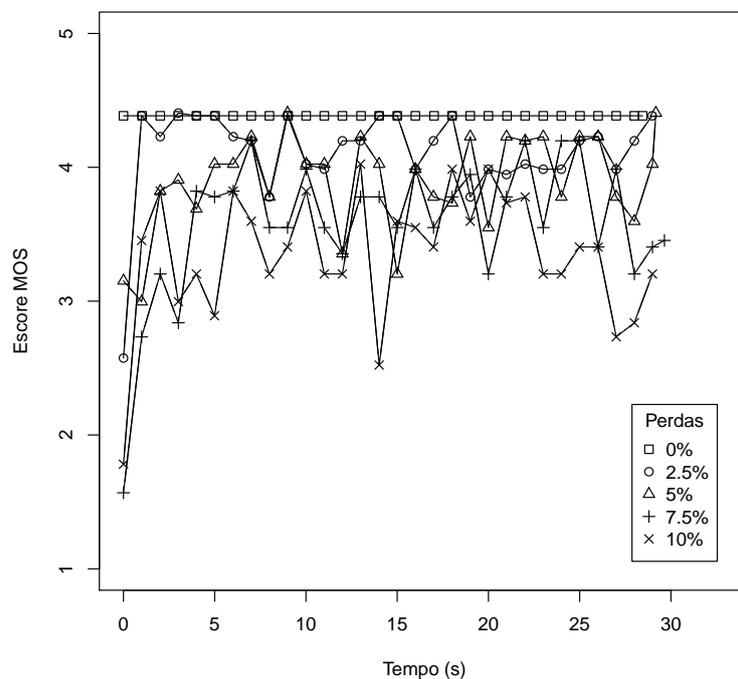


Figura 5.1: MOS instantânea em função do tempo para diferentes perdas de pacotes.

O Fator R é diretamente relacionado aos parâmetros de rede, sendo um deles a perda de pacotes. Variando-se o valor do fator de R, o valor instantâneo do MOS também varia. Uma variação do fator de R implica necessariamente uma variação do MOS. Note-se que o melhor dos casos para o MOS é quando não há nenhuma perda de pacotes.

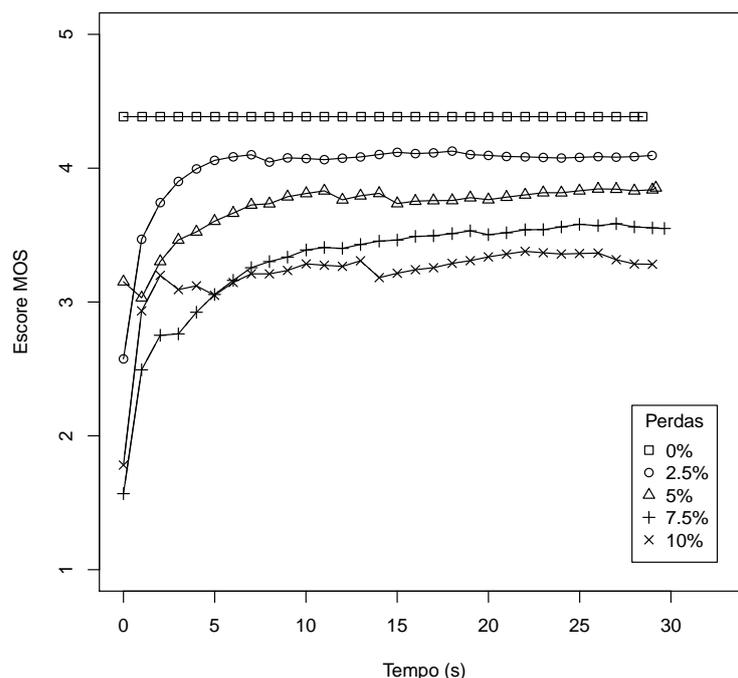


Figura 5.2: MOS médio em função do tempo para diferentes perdas de pacotes.

O valor instantâneo médio da Qualidade de Experiência, medido em MOS, se aproxima do valor global de MOS com o passar do tempo. Mais uma vez, note-se que na Figura 5.3, o valor MOS para uma dada situação de perda de pacotes piora à medida que as perda aumentam.

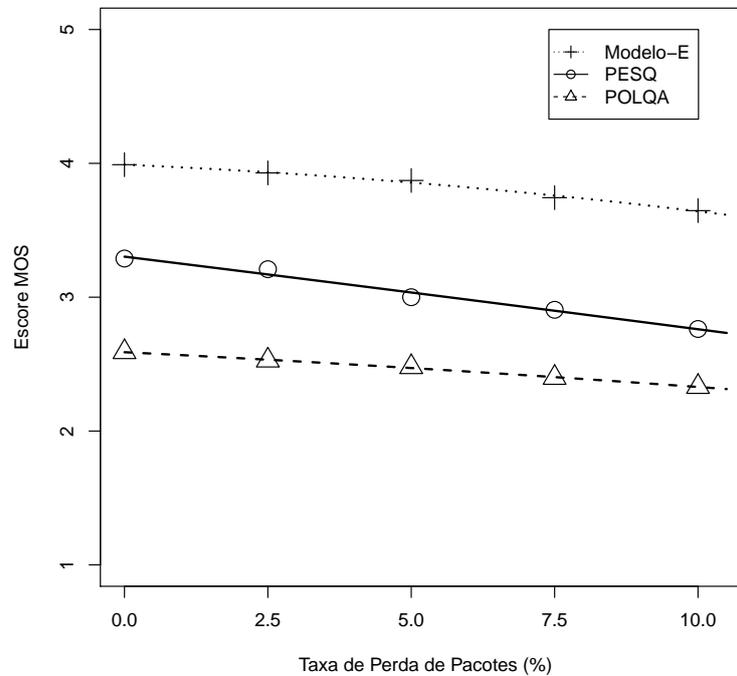


Figura 5.3: MOS médio *versus* perda de pacotes.

Na Figura 5.2, a média da qualidade MOS apresenta uma pequena diferença antes de 5 segundos entre os valores de perda de pacotes de 5% a 10%. Os valores iniciais têm oscilações de largura que são minimizadas pela média com o passar do tempo.

Com o aumento da perda, a qualidade MOS diminui, como apresentado na Figura 5.3.

## 5.2 Teste de Massa

Os testes de massa foram realizados com os 24 arquivos de testes disponíveis no site da OSR (PROJECT, 2015). As características dos arquivos de testes estão apresentadas no Capítulo 4, Tabela 4.1, na página 50. A qualidade dos arquivos de teste foi analisada pelos algoritmos PESQ e POLQA e os resultados das medidas do teste estão presentes na Tabela 5.1. Note que devido ao algoritmo PESQ utilizar o mesmo modelo perceptual para analisar ambos os sinais (sinal original e sinal degradado) conforme apresentado na Seção 2.4.3 do Capítulo 2 na página 31, os resultados obtidos não poderiam ser diferentes de 4,5, uma vez que o sinal original e o sinal degradado são os mesmos. Lembre-se que ao contrário do algoritmo POLQA, o algoritmo PESQ não leva em consideração a qualidade do sinal original.

Os valores apresentados na Tabela 5.1 foram os valores máximos considerados para as medidas. Como foi um teste *Full Reference* onde arquivos de entrada e arquivos de saída foram os mesmos, não existe razão para se supor que haveria me-

Tabela 5.1: Qualidade PESQ e POLQA dos arquivos originais - *Full Reference*.

Arquivo em Teste	$MOS_{LQO}$ P.862.1 (PESQ)	$MOS_{LQO}$ (POLQA)
OSR_us_000_0010_8k.wav	4,538	3,283
OSR_us_000_0011_8k.wav	4,549	3,324
OSR_us_000_0012_8k.wav	4,549	3,335
OSR_us_000_0013_8k.wav	4,548	3,324
OSR_us_000_0014_8k.wav	4,539	3,292
OSR_us_000_0015_8k.wav	4,549	3,262
OSR_us_000_0016_8k.wav	4,534	3,251
OSR_us_000_0017_8k.wav	4,549	3,277
OSR_us_000_0018_8k.wav	4,476	3,209
OSR_us_000_0019_8k.wav	4,549	3,248
OSR_us_000_0030_8k.wav	4,549	3,285
OSR_us_000_0031_8k.wav	4,549	3,319
OSR_us_000_0032_8k.wav	4,479	3,320
OSR_us_000_0034_8k.wav	4,549	3,358
OSR_us_000_0035_8k.wav	4,549	3,330
OSR_us_000_0036_8k.wav	4,481	3,315
OSR_us_000_0037_8k.wav	4,549	3,359
OSR_us_000_0038_8k.wav	4,549	3,341
OSR_us_000_0039_8k.wav	4,547	3,341
OSR_us_000_0040_8k.wav	4,481	3,335
OSR_us_000_0057_8k.wav	4,549	3,352
OSR_us_000_0059_8k.wav	4,549	3,304
OSR_us_000_0060_8k.wav	4,549	3,295
OSR_us_000_0061_8k.wav	4,535	3,324

dida com  $MOS_{LQO}$  maior dos que os valores apresentados, uma vez que o arquivo em teste não sofreu degradação por perdas de pacote ou pelo processo de codificação e decodificação.

Para cada um dos 24 arquivos de teste da base OSR, todos os codecs disponíveis padrão ITU-T foram utilizados (PCMU, GSM, G.723, G.729, G.726-32) e para cada codec, diferentes perdas foram selecionadas (0,0%, 1,0%, 2,0%, 2,5%, 3,0%, 5,0%, 7,5%, 10,0%). Para cada caso, cinco repetições foram executadas.

Os resultados dos testes de massa são apresentados nas Tabelas 5.2 à 5.27 e nas Figuras 5.4 à 5.16. As informações foram organizadas da seguinte maneira. Separou-se as informações em dois grupos: (1) por codecs, formadas pelos elementos de teste (PCMU, GSM, G.723, G.729, G.726-32) apresentadas na Subseção 5.2.1 e (2) por Taxa de Perdas formada pelas elementos de teste (0,0%, 1,0%, 2,0%, 2,5%, 3,0%, 5,0%, 7,5%, 10,0%) apresentadas na Subseção 5.2.2.

Para cada elemento de cada grupo são apresentadas duas Tabelas e uma Figura. A primeira Tabela apresenta os valores médios e desvios padrões das medidas de qualidade obtidas pelo Modelo-E, PESQ e POLQA para cada elemento do outro grupo, associado a aquele elemento dentro do grupo. Por exemplo, a Tabela 5.2 no Grupo de Codec, apresenta as médias de qualidade para cada Taxa de Perda (%) do codec

PCMU. Já a Tabela 5.12 no Grupo de Taxa de Perdas, apresenta as médias de qualidade para cada codec com a Taxa de Perda de 0%. Essa informação foi particularmente útil ao determinar-se o melhor codec para uma determinada Taxa de Perdas em uma determinada medida de qualidade.

A segunda Tabela, de cada elemento em cada grupo, apresenta os resultados da Correlação de Pearson apresentada na Seção 4.13 para as medidas de qualidade obtidas entre Modelo-E x PESQ e entre Modelo-E x POLQA. A título de curiosidade, também foi calculado o coeficiente de Pearson entre as medidas PESQ e POLQA.

Na Subseção 5.2.1 são apresentados os resultados individualmente para cada codec testado e na Subseção 5.2.2 são apresentados os resultados para cada valor de Taxa de Perda.

## 5.2.1 Resultados por Codec

### Codec PCMU

As Tabelas 5.2 e 5.3 apresentam, respectivamente, os valores médios das medidas de qualidade e a correlação entre os diferentes tipos de medidas (MODELO-E, PESQ e POLQA) obtidos por todos os arquivos (24) transmitidos da base OSR (PROJECT, 2015) e todas as repetições (5) de cada um deles.

As médias e desvios medidos de qualidade para o codec PCMU apresentados na Tabela 5.2 estão coerentes, mas alguns fatos chamam a atenção. Para as medidas de 1%, 3% e 10% a variância foi maior que para as demais para todos os métodos (Modelo-E, PESQ e POLQA) indicando que pode eventualmente ter ocorrido algum outro tipo de problema na transmissão e na dispersão das perdas.

Tabela 5.2: Média e desvio das medidas para o codec PCMU para cada perda em teste.

Perdas	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
0,0%	4,022	0,351	3,324	0,037	2,665	0,409
1,0%	4,013	0,614	3,300	0,126	2,668	0,641
2,0%	3,979	0,365	3,300	0,117	2,629	0,508
2,5%	4,011	0,353	3,328	0,079	2,660	0,398
3,0%	4,031	0,658	3,353	0,163	2,667	0,628
5,0%	3,938	0,378	3,226	0,164	2,586	0,468
7,5%	3,889	0,368	3,206	0,150	2,583	0,409
10,0%	3,800	0,643	3,100	0,224	2,499	0,657

A Tabela 5.3 possui uma boa correlação em todas as medidas, entre 0,7 e 0,9, indicando uma correlação forte positiva entre PESQ e Modelo-E e POLQA e Modelo-E. Já a correlação entre PESQ e POLQA apresentam uma correlação muito forte positiva, maior do 0,9, exceto para a taxa de perda 5% onde todas as correlações ficaram entre 0,7 e 0,9. Mesmo assim, é possível afirmar que a correlação é forte positiva.

Já na Figura 5.4, o comportamento das três curvas foi bem próximo, exceto por uma polarização entre as medidas. Em 3% todas as medidas foram maiores que nas

Tabela 5.3: Correlação entre as medidas do codec PCMU.

Perda	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
0,0%	0,782	0,844	0,929
1,0%	0,834	0,870	0,940
2,0%	0,826	0,887	0,935
2,5%	0,806	0,856	0,948
3,0%	0,797	0,835	0,918
5,0%	0,787	0,756	0,875
7,5%	0,841	0,865	0,936
10,0%	0,817	0,867	0,930

vizinhas 2,5% e 5%, o que não deveria ocorrer. Merece um estudo mais aprofundado, mas nada que invalide o experimento.

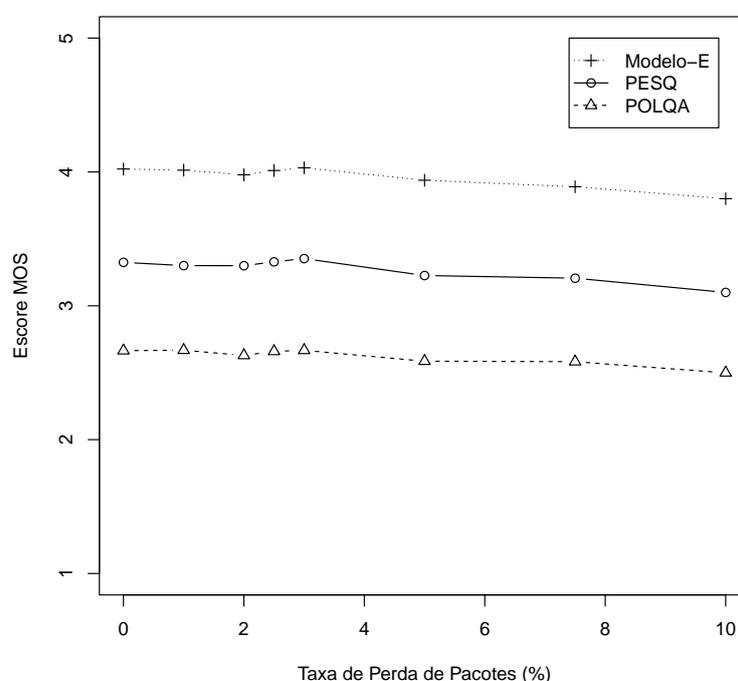


Figura 5.4: Cenário de teste - codec PCMU.

### Codec GSM

Para o codec GSM, a Tabela 5.4 apresentou a queda das médias de qualidade em função do aumento da Taxa de Perdas, conforme esperado em todos os métodos (Modelo-E, PESQ e POLQA), ocorrendo as menores variâncias para o método PESQ.

Observa-se na Tabela 5.5, que o coeficiente de correlação linear de Pearson para a Taxa de Perda de 1% foi próximo de zero e negativo. Tal fato indica que a correlação não é linear, entretanto, não exclui-se a possibilidade da correlação apresentar comportamento não-linear. Ainda na Tabela 5.5, pode-se observar que a medida que a Taxa de Perdas aumenta, o Coeficiente de Pearson aumenta. Ainda assim, não se pode

Tabela 5.4: Média e desvio das medidas para o codec GSM para cada perda em teste.

Perdas	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
0,0%	4,251	0,406	3,280	0,194	2,627	0,492
1,0%	4,190	0,373	3,254	0,184	2,599	0,413
2,0%	4,129	0,640	3,212	0,269	2,582	0,657
2,5%	4,101	0,401	3,209	0,249	2,562	0,490
3,0%	4,077	0,370	3,192	0,224	2,552	0,415
5,0%	3,989	0,615	3,157	0,304	2,513	0,652
7,5%	3,910	0,377	3,122	0,259	2,491	0,491
10,0%	3,861	0,384	3,088	0,329	2,449	0,447

considerar uma correlação forte. No melhor caso, na Taxa de Perdas de 10%, pode-se considerar moderada.

Tabela 5.5: Correlação entre as medidas do codec GSM.

Perda	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
0,0%	0,159	0,134	0,969
1,0%	-0,008	0,097	0,951
2,0%	0,176	0,351	0,940
2,5%	0,189	0,446	0,914
3,0%	0,262	0,481	0,936
5,0%	0,432	0,692	0,919
7,5%	0,558	0,749	0,942
10,0%	0,622	0,767	0,944

Na Figura 5.5 observa-se que o codec se comportou como o esperado, ou seja, quanto maior a Taxa de Perdas, menor a qualidade. Nem sempre isso será verdade como isto é verificado nas Tabelas seguintes para os outros codecs. É recomendável a realização de novos testes estatísticos para esse codec, por exemplo, ANOVA (*Analysis of Variance*) ou T-Student.

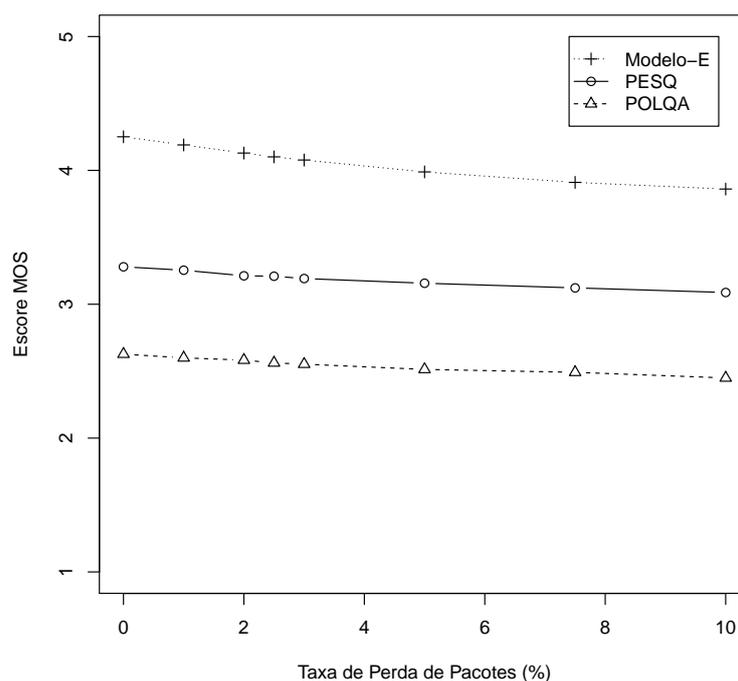


Figura 5.5: Cenário de teste - codec GSM.

### Codec G.723

Para o codec G.723, a Tabela 5.6 apresentou decaimento das médias de qualidade em função do aumento da Taxa de Perdas, conforme esperado em todos os métodos (Modelo-E, PESQ e POLQA), mas desta vez com a variâncias mais equilibradas quando comparadas ao codec GSM na Tabela 5.4.

Tabela 5.6: Média e desvio das medidas para o codec G.723 para cada perda em teste.

Perdas	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
0,0%	3,856	0,618	3,612	0,374	2,621	0,661
1,0%	3,823	0,404	3,561	0,321	2,579	0,519
2,0%	3,788	0,434	3,525	0,465	2,557	0,526
2,5%	3,772	0,661	3,502	0,457	2,527	0,728
3,0%	3,753	0,438	3,488	0,390	2,522	0,557
5,0%	3,699	0,463	3,443	0,589	2,476	0,522
7,5%	3,632	0,634	3,402	0,538	2,445	0,700
10,0%	3,563	0,417	3,357	0,442	2,404	0,554

No codec G.723, para a taxa de Perdas igual a 0% apresentado na Tabela 5.7 foi identificado o coeficiente de correlação linear de Pearson negativo muito próxima de zero tanto para o algoritmo PESQ quanto para o algoritmo POLQA quando confrontados com os resultados no Modelo-E. Novamente isso indica que o modelo linear pode não se ajustar perfeitamente para determinar se existe ou não uma correlação entre as medidas.

Observa-se ainda na Tabela 5.7 que a medida que a Taxa de Perdas aumenta, o

coeficiente de correlação linear de Pearson aumenta, chegando próximo à 1 a partir da taxa de 5%, indicando uma correlação muito forte positiva. Entre 2% e 3% tem-se uma correlação forte positiva e em 1% uma correlação positiva moderada.

Tabela 5.7: Correlação entre as medidas do codec G.723.

Perda	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
0,0%	-0,018	-0,028	0,911
1,0%	0,563	0,497	0,902
2,0%	0,785	0,737	0,935
2,5%	0,869	0,845	0,952
3,0%	0,849	0,855	0,966
5,0%	0,931	0,919	0,981
7,5%	0,940	0,949	0,980
10,0%	0,965	0,962	0,989

Pela Figura 5.6 observa-se que o codec G.723 comportou-se como esperado, ou seja, maior Taxa de Perdas apresentou menor qualidade. Mas quando comparado à Figura 5.5 observa-se que as medidas de PESQ estão mais próximas das medidas obtidas pelo Modelo-E no codec G.723 do que no codec GSM.

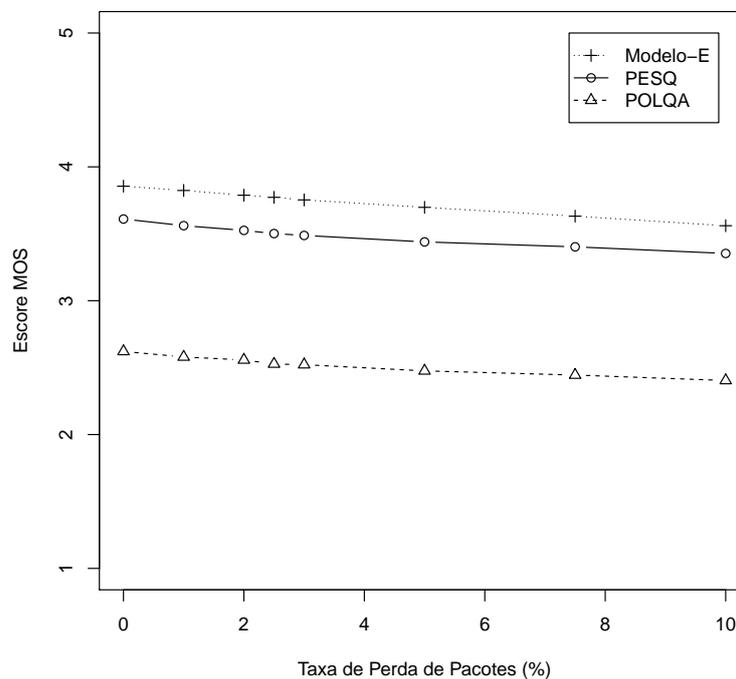


Figura 5.6: Cenário de teste - codec G.723.

### Codec G.729

O codec G.729 na Tabela 5.8, assim como o codec PCMU na Tabela 5.2, não apresentou um decaimento das médias de qualidade em função do aumento da Taxa de Perdas,

conforme esperado em todos os métodos (Modelo-E, PESQ e POLQA). Em dois momentos, em 1% e 2,5%, a medida apresentada foi maior do que a sua antecessora, o que não era esperado. Entretanto, esse comportamento foi observado nos três métodos de avaliação (Modelo-E, PESQ e POLQA) indicando que provavelmente o problema está na qualidade do sinal em teste. As medidas com POLQA indicam também que a qualidade do sinal medido é muito baixa. Os números são próximos a 2, indicando uma qualidade ruim para a maioria dos usuários.

Tabela 5.8: Média e desvio das medidas para o codec G.729 para cada perda em teste.

Perdas	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
0,0%	3,493	0,011	2,952	0,501	2,051	0,223
1,0%	3,539	0,453	2,994	0,755	2,059	0,403
2,0%	3,442	0,247	2,869	0,653	1,953	0,200
2,5%	3,553	0,120	3,008	0,482	2,086	0,229
3,0%	3,415	0,440	2,847	0,739	1,941	0,420
5,0%	3,461	0,249	2,916	0,628	2,004	0,212
7,5%	3,390	0,229	2,860	0,498	1,989	0,242
10,0%	3,358	0,464	2,825	0,722	1,979	0,413

Mesmo diante do ocorrido na Tabela 5.8, observa-se na 5.9 o coeficiente de correlação linear de Pearson extremamente forte (maior que 0,95) e positiva em todas as situações.

Tabela 5.9: Correlação entre as medidas do codec G.729.

Perda	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
0,0%	0,976	0,975	0,986
1,0%	0,968	0,956	0,981
2,0%	0,968	0,949	0,982
2,5%	0,970	0,948	0,976
3,0%	0,970	0,955	0,978
5,0%	0,975	0,963	0,986
7,5%	0,972	0,961	0,989
10,0%	0,961	0,946	0,987

A Figura 5.7 indica que o comportamento das curvas são semelhantes, sendo que um pico foi observado por volta de 3%. Isto pode indicar uma falha no teste de 2,5% ou um comportamento não detectado nas amostras como, por exemplo, falha na distribuição das perdas no Netem.

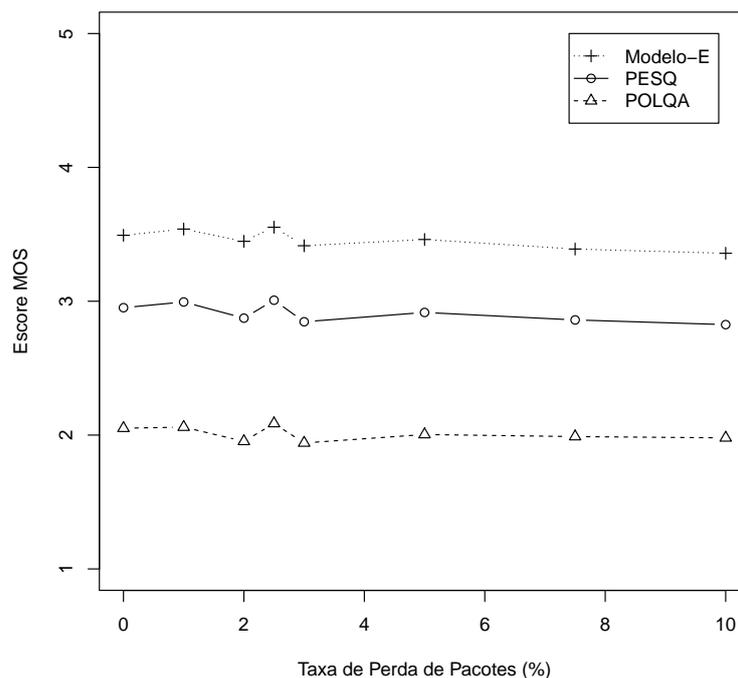


Figura 5.7: Cenário de teste - codec G.729.

### Codec G.726-32

O codec G.726-32 com suas médias e desvios padrões apresentados na Tabela 5.10 possui um comportamento oscilatório em relação a Taxa de Perdas entre 0% e 5%. Isso se deve parcialmente a sua robustez a perda de pacotes. Acima de 5% de perdas houve o decaimento da qualidade em todos os métodos conforme esperado.

Tabela 5.10: Média e desvio das medidas para o codec G.726-32 para cada perda em teste

Perdas	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
0,0%	3,806	0,267	3,095	0,624	2,427	0,232
1,0%	3,670	0,281	2,926	0,481	2,280	0,225
2,0%	3,708	0,451	2,984	0,726	2,318	0,426
2,5%	3,704	0,273	2,987	0,625	2,318	0,241
3,0%	3,742	0,332	3,033	0,505	2,367	0,247
5,0%	3,715	0,454	3,036	0,716	2,360	0,443
7,5%	3,633	0,283	2,939	0,618	2,296	0,247
10,0%	3,577	0,506	2,890	0,523	2,252	0,286

O coeficiente linear de Pearson apresentado na Tabela 5.11 apresenta forte correlação positiva em todos os cenários, com 0,87 de aderência.

As medidas do Modelo-E, PESQ e POLQA estão igualmente espaçadas como se observa na Figura 5.8 e também na Figura 5.4 na página 66 para o codec PCMU.

Tabela 5.11: Correlação entre as medidas do codec G.726-32.

Perda	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
0,0%	0,887	0,944	0,952
1,0%	0,869	0,914	0,917
2,0%	0,885	0,917	0,938
2,5%	0,883	0,916	0,944
3,0%	0,886	0,912	0,934
5,0%	0,897	0,945	0,958
7,5%	0,903	0,945	0,955
10,0%	0,900	0,933	0,961

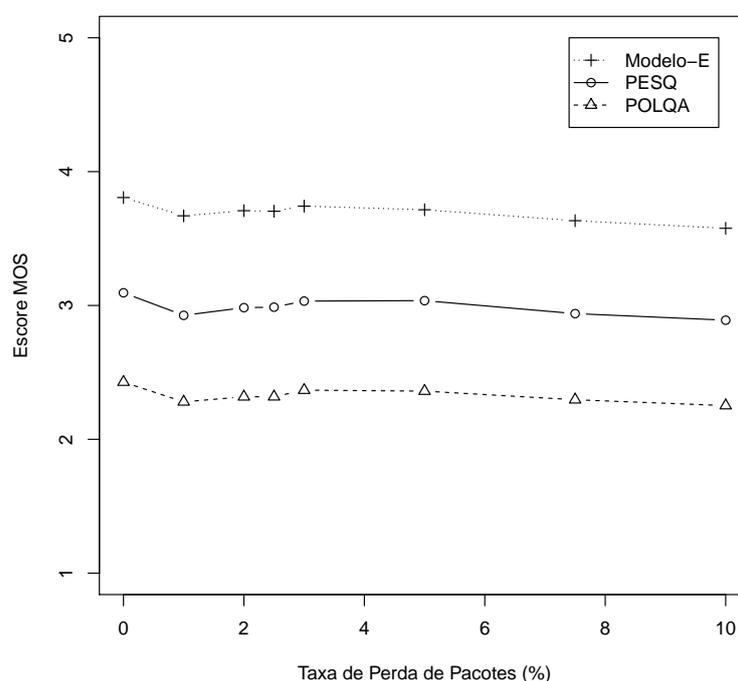


Figura 5.8: Cenário de teste - codec G.726-32.

## 5.2.2 Resultados por Perdas

Esta Subseção agrupa os resultados apresentados Tabelas 5.2 a Tabela 5.11 da Subseção 5.2.1 e as re-organiza de modo que possam servir como critério de seleção de codec dada uma determinada condição de Taxa de Perdas.

### Perda 0%

Em valores absolutos, ao se observar os resultados apresentados na Tabela 5.12, o codec GSM apresentou o melhor desempenho medido pelo Modelo-E. Para os algoritmos PESQ e POLQA, o codec de melhor desempenho nos teste foi o codec PCMU. Entretanto, ao considerar-se as incertezas das medidas, os codecs GSM, G.723 e PCMU tiveram desempenhos semelhantes. Em todos os casos, o codec G.729 apresentou o pior desempenho tanto pelo Modelo-E quanto por PESQ e POLQA.

A correlação entre os valores medidos de qualidade para a Taxa de Perdas de 0%, foi

Tabela 5.12: Média e desvio das medidas para perda de 0% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	4,022	0,351	3,324	0,037	2,665	0,409
GSM	4,251	0,406	3,280	0,194	2,627	0,492
G.723	3,856	0,618	3,612	0,374	2,621	0,661
G.729	3,493	0,011	2,952	0,501	2,051	0,223
G.726-32	3,806	0,267	3,095	0,624	2,427	0,232

extremamente forte e positiva para os codecs G.729 e G.726-32. Já para os codecs GSM e G.723, a correlação linear de Pearson foi fraca. Entretanto, não exclui-se a possibilidade da correlação entre as medidas de mesma perda apresentar comportamento não-linear.

Tabela 5.13: Correlação entre as medidas de perda 0,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,775	0,844	0,933
GSM	0,163	0,134	0,969
G.723	-0,016	-0,028	0,915
G.729	0,973	0,975	0,985
G.726-32	0,879	0,944	0,949

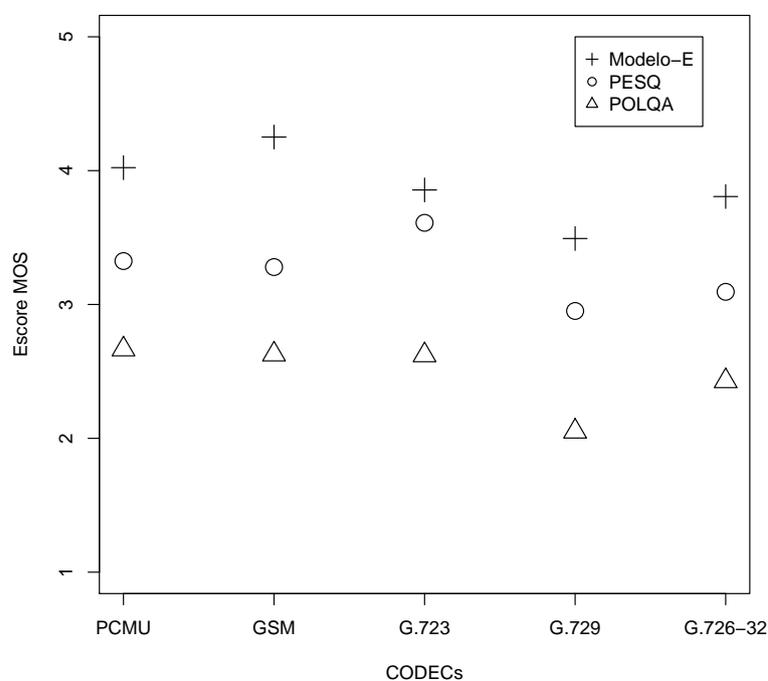


Figura 5.9: Cenário de teste - taxa de perdas 0%.

**Perda 1%**

Semelhante aos resultados apresentados na Subseção 5.2.2 na Tabela 5.14, o codec GSM apresentou o melhor desempenho nesta situação para as medidas do Modelo-E. Já o codec G.723 apresentou o melhor desempenho em PESQ e o codec PCMU teve o melhor desempenho em POLQA. Mas considerando-se os desvios padrões, GSM, PCMU e G.723 tiveram desempenhos semelhantes nas três métricas (Modelo-E, PESQ e POLQA).

Tabela 5.14: Média e desvio das medidas para perda de 1% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	4,013	0,614	3,300	0,126	2,668	0,641
GSM	4,190	0,373	3,254	0,184	2,599	0,413
G.723	3,823	0,404	3,561	0,321	2,579	0,519
G.729	3,539	0,453	2,994	0,755	2,059	0,403
G.726-32	3,670	0,281	2,926	0,481	2,280	0,225

O codec GSM apresentou uma correlação fraca indicando que o modelo linear de Pearson não está adequado. PCMU, G.726-32 e G.729 apresentam uma correlação forte positiva entre Modelo-E e PESQ e Modelo-E e POLQA conforme apresentado na Tabela 5.15.

Tabela 5.15: Correlação entre as medidas de perda 1,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,827	0,870	0,943
GSM	-0,005	0,097	0,951
G.723	0,554	0,497	0,902
G.729	0,965	0,956	0,981
G.726-32	0,863	0,914	0,915

Na Figura 5.10 a menor distância entre as medidas de Qualidade do Modelo-E e PESQ ocorreu com o codec G.723 e a maior ocorreu com o codec GSM.

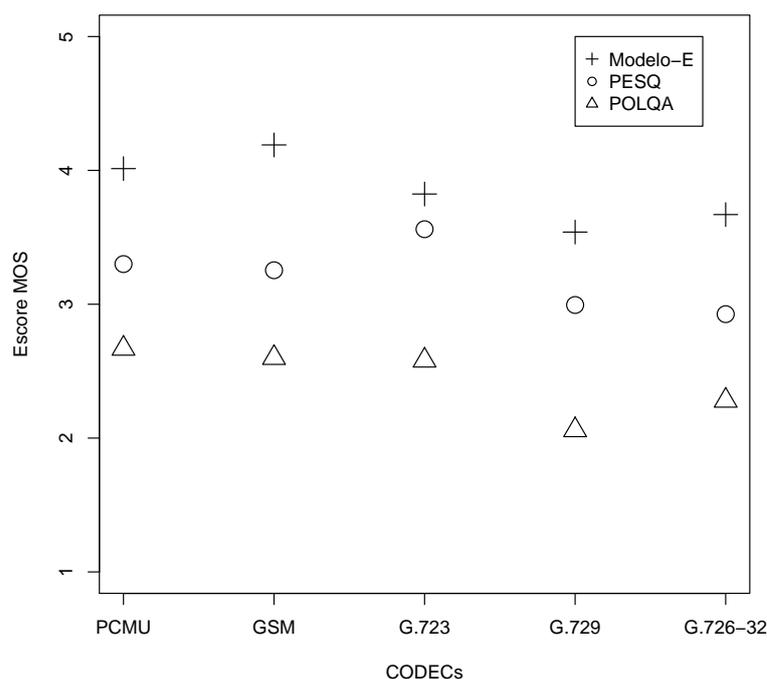


Figura 5.10: Cenário de teste - taxa de perdas 1%.

### Perda 2%

Na Tabela 5.16 o codec GSM apresentou o melhor desempenho para o Modelo-E, o codec G.723 para o algoritmo PESQ e o codec PCMU para a medida de qualidade do POLQA. Ao comparar-se os resultados obtidos com PESQ e POLQA com os resultados apresentados na Tabela 5.1, observa-se a queda de qualidade em relação as medidas sem perdas e não codificadas.

Tabela 5.16: Média e desvio das medidas para perda de 2% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	3,979	0,365	3,300	0,117	2,629	0,508
GSM	4,129	0,640	3,212	0,269	2,582	0,657
G.723	3,788	0,434	3,525	0,465	2,557	0,526
G.729	3,442	0,247	2,869	0,653	1,953	0,200
G.726-32	3,708	0,451	2,984	0,726	2,318	0,426

Observa-se na Tabela 5.17 a forte correlação positiva entre Modelo-E e PESQ e entre Modelo-E e POLQA para os codecs PCMU, G.723, G.729 e G.726-32 na Tabela 5.17. O codec GSM apresentou uma correlação moderada positiva.

Novamente o codec G.723 possuiu a menor distância entre os resultados apresentados pelas medidas de Qualidade do Modelo-E e PESQ, como pode ser observado na Figura 5.11.

Tabela 5.17: Correlação entre as medidas de perda 2,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,811	0,867	0,930
GSM	0,617	0,767	0,946
G.723	0,960	0,962	0,989
G.729	0,967	0,946	0,983
G.726-32	0,898	0,933	0,959

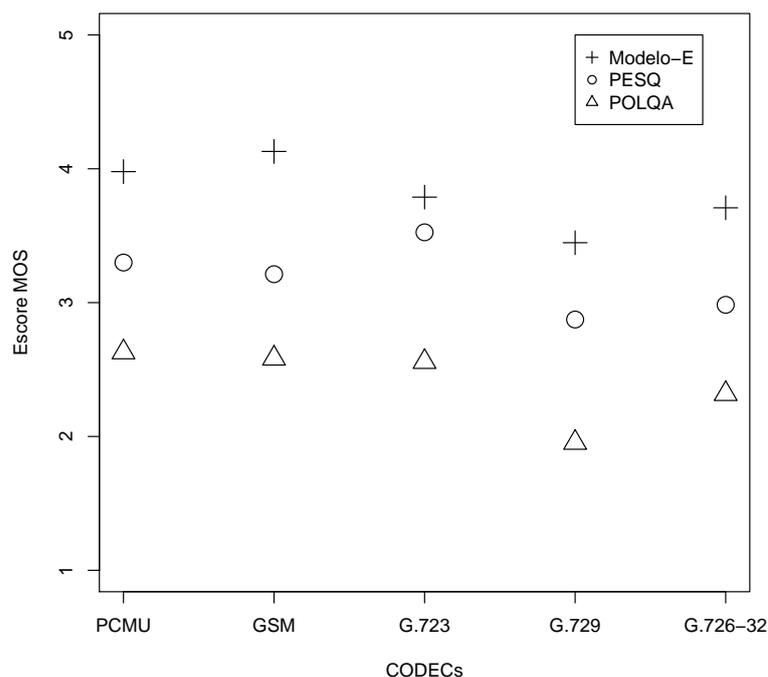


Figura 5.11: Cenário de teste - taxa de perdas 2%.

### Perda 2,5%

Na Tabela 5.18 observa-se resultado semelhante ao apresentado na Subseção 5.2.2: o codec GSM apresentou o melhor desempenho no Modelo-E. O codec PCMU apresentou o melhor desempenho segundo os algoritmos PESQ e POLQA. Mas na Tabela 5.19, as medidas do coeficiente de correlação entre Modelo-E e PESQ e entre Modelo-E e POLQA para o codec GSM foram fracas positivas indicando que o modelo de correlação linear não se ajusta bem aos resultados obtidos.

Tabela 5.18: Média e desvio das medidas para perda de 2,5% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	4,011	0,353	3,328	0,079	2,660	0,398
GSM	4,101	0,401	3,209	0,249	2,562	0,490
G.723	3,772	0,661	3,502	0,457	2,527	0,728
G.729	3,553	0,120	3,008	0,482	2,086	0,229
G.726-32	3,704	0,273	2,987	0,625	2,318	0,241

Tabela 5.19: Correlação entre as medidas de perda 2,5%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,819	0,887	0,939
GSM	0,178	0,351	0,941
G.723	0,775	0,737	0,938
G.729	0,966	0,949	0,982
G.726-32	0,879	0,917	0,937

Na Figura 5.12 os resultados foram semelhantes aos apresentados na Subseção 5.2.2, na Figura 5.11.

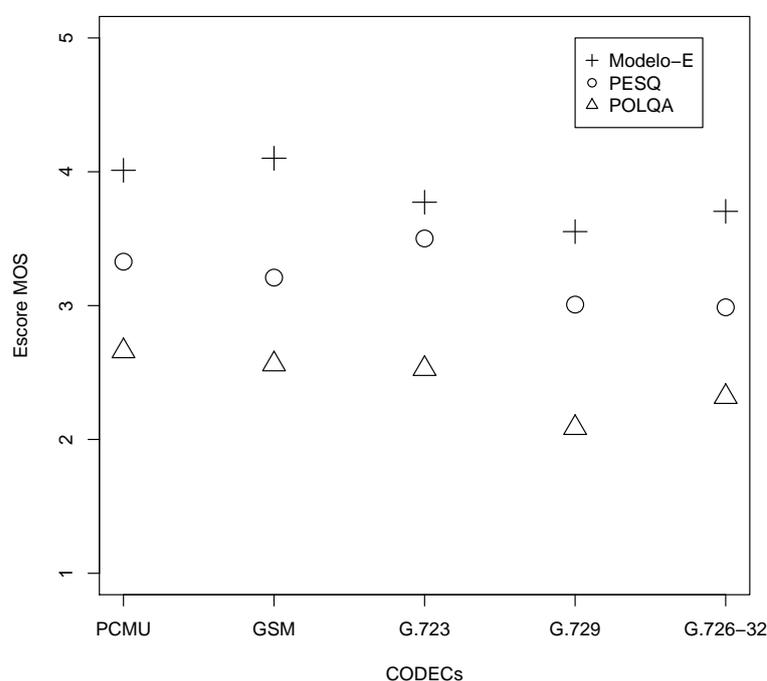


Figura 5.12: Cenário de teste - taxa de perdas 2,5%.

### Perda 3%

Para esta situação de taxa de perda de 3% apresentada na Tabela 5.20, o codec GSM leva pequena vantagem sobre o codec PCMU nas medidas de Qualidade de Experiência no Modelo-E. Esta vantagem desaparece ao considerar-se o número de casas decimais igual a 2, e apresentam um empate técnico se considerarmos os desvios padrões. Para as medidas de qualidade de PESQ e POLQA, já considerando os desvios padrões, tem-se um empate técnico entre os codecs G.723 e PCMU.

Novamente as medidas de qualidade do codec GSM apresentam correlação positiva fraca, indicando que o modelo linear não se aplica a correlação entre PESQ e Modelo-E para as medidas de mesma perda conforme apresentado na Tabela 5.21. Entretanto a correlação para estes mesmos codecs, embora não forte como nos outros codecs, apresentou uma correlação moderada quando comparada a correlação entre POLQA e Modelo-E.

Tabela 5.20: Média e desvio das medidas para perda de 3% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	4,031	0,658	3,353	0,163	2,667	0,628
GSM	4,077	0,370	3,192	0,224	2,552	0,415
G.723	3,753	0,438	3,488	0,390	2,522	0,557
G.729	3,415	0,440	2,847	0,739	1,941	0,420
G.726-32	3,742	0,332	3,033	0,505	2,367	0,247

Tabela 5.21: Correlação entre as medidas de perda 3,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,798	0,856	0,950
GSM	0,191	0,446	0,914
G.723	0,860	0,845	0,953
G.729	0,968	0,948	0,975
G.726-32	0,877	0,916	0,943

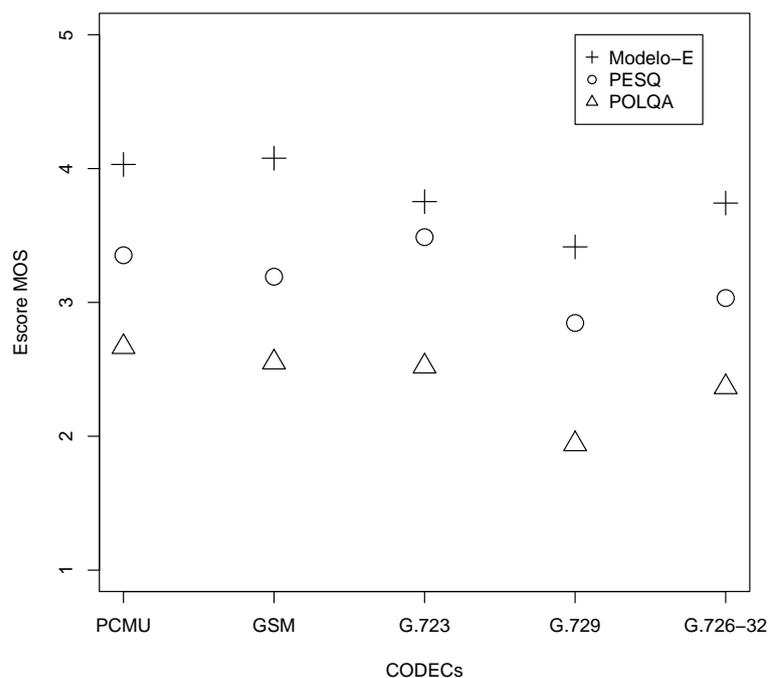


Figura 5.13: Cenário de teste - taxa de perdas 3%.

### Perda 5%

Outra vez, resultados semelhantes ao da Subseção 5.2.2 são apresentados nas Tabelas 5.22 e Tabela 5.23. Novamente, o codec GSM apresentou uma correlação menor do que as medidas dos demais codecs, mas superior ao resultado apresentado por este mesmo codec, GSM, na Tabela 5.19 e Tabela 5.21.

Tabela 5.22: Média e desvio das medidas para perda de 5% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	3,938	0,378	3,226	0,164	2,586	0,468
GSM	3,989	0,615	3,157	0,304	2,513	0,652
G.723	3,699	0,463	3,443	0,589	2,476	0,522
G.729	3,461	0,249	2,916	0,628	2,004	0,212
G.726-32	3,715	0,454	3,036	0,716	2,360	0,443

Tabela 5.23: Correlação entre as medidas de perda 5,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,788	0,835	0,921
GSM	0,263	0,481	0,936
G.723	0,843	0,855	0,968
G.729	0,969	0,955	0,976
G.726-32	0,881	0,912	0,933

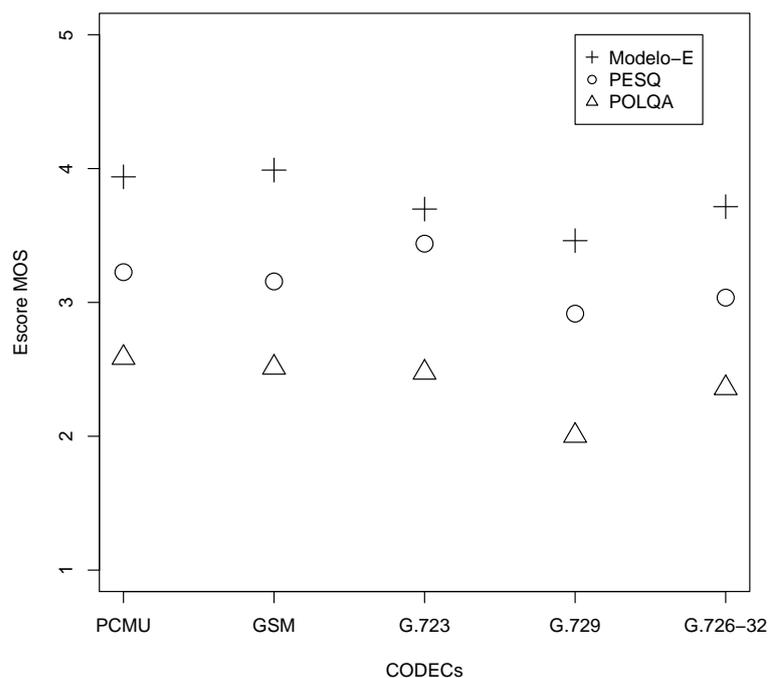


Figura 5.14: Cenário de teste - taxa de perdas 5%.

### Perda 7,5%

Resultados semelhantes ao da Subseção 5.2.2 são apresentados nas Tabelas 5.24 e Tabela 5.25. O codec GSM apresentou uma correlação menor do que as correlações dos demais codecs, mas uma correlação superior quando comparados com ele mesmo na Tabela 5.23.

Tabela 5.24: Média e desvio das medidas para perda de 7,5% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	3,889	0,368	3,206	0,150	2,583	0,409
GSM	3,910	0,377	3,122	0,259	2,491	0,491
G.723	3,632	0,634	3,402	0,538	2,445	0,700
G.729	3,390	0,229	2,860	0,498	1,989	0,242
G.726-32	3,633	0,283	2,939	0,618	2,296	0,247

Tabela 5.25: Correlação entre as medidas de perda 7,5%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,783	0,756	0,881
GSM	0,433	0,692	0,919
G.723	0,925	0,919	0,982
G.729	0,976	0,963	0,983
G.726-32	0,893	0,945	0,956

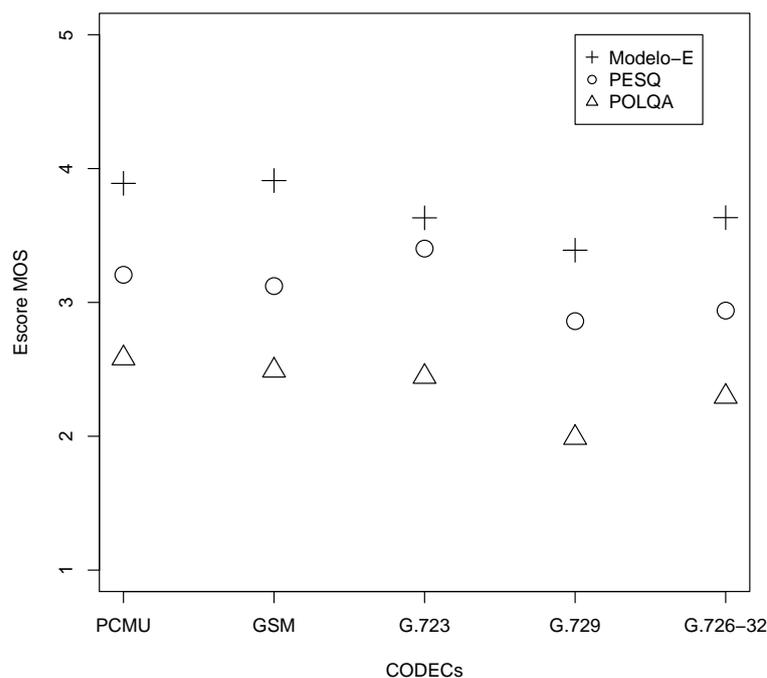


Figura 5.15: Cenário de teste - taxa de perdas 7,5%.

### Perda 10%

Com uma taxa de perda de pacotes de 10%, os resultados da Tabela 5.26 com as medidas de POLQA apontaram o codec PCMU com a maior média dentre os codecs testados. Entretanto, as medidas POLQA indicam para qualquer codec, uma qualidade, próximo a 2, ruim para a maioria dos usuários.

O coeficiente de correlação para o codecs GSM na Tabela 5.27 aumentou em relação ao mesmos coeficientes apresentados na Subseção 5.2.2, mas ainda estão na categoria de correlação moderada.

Os codecs presentes na Figura 5.16 apresentaram comportamento semelhante ao das Figuras 5.13, 5.14 e 5.15, exceto por um fator de escala. Como visto na Subseção 5.2.1, a medida que a Taxa de Perdas aumenta a partir de 5%, a qualidade diminui para todos os codecs.

Tabela 5.26: Média e desvio das medidas para perda de 10% para cada codec em teste.

Codec	Modelo-E		PESQ		POLQA	
	Media	Desvio	Media	Desvio	Media	Desvio
PCMU	3,800	0,643	3,100	0,224	2,499	0,657
GSM	3,861	0,384	3,088	0,329	2,449	0,447
G.723	3,563	0,417	3,357	0,442	2,404	0,554
G.729	3,358	0,464	2,825	0,722	1,979	0,413
G.726-32	3,577	0,506	2,890	0,523	2,252	0,286

Tabela 5.27: Correlação entre as medidas de perda 10,0%.

Codec	PESQ x Modelo-E	POLQA x Modelo-E	PESQ x POLQA
PCMU	0,835	0,865	0,937
GSM	0,558	0,749	0,942
G.723	0,935	0,949	0,979
G.729	0,974	0,961	0,987
G.726-32	0,901	0,945	0,952

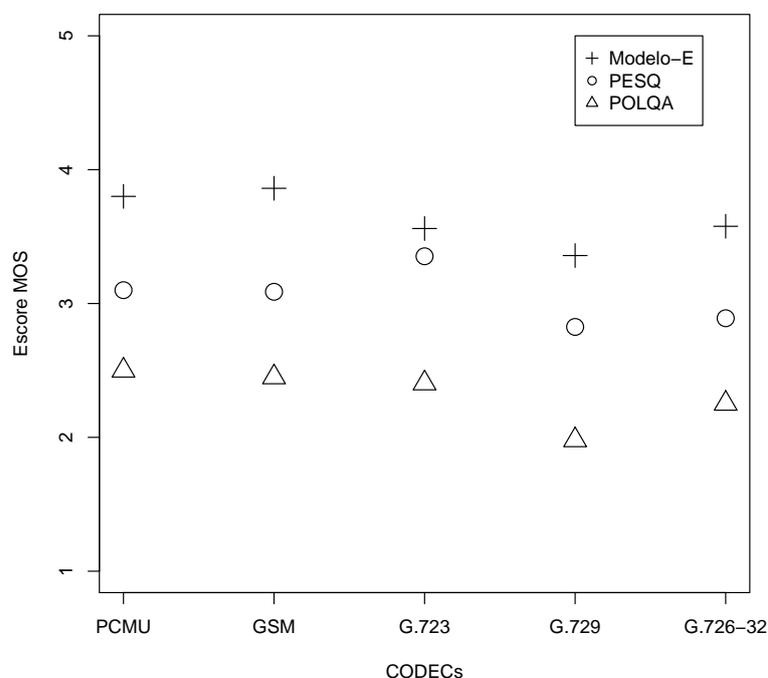


Figura 5.16: Cenário de teste - taxa de perdas 10%.

### 5.3 Relação entre as medidas obtidas, PESQ e POLQA

Na seção de 5.2, testes de Correlação Linear de Pearson foram aplicados entre os resultados de mesma categoria. Para um determinado codec, em uma determinada perda, foi calculado o coeficiente de correlação linear entre o método proposto e os resultados obtidos com o algoritmo PESQ e entre o método proposto e os resultados obtidos com o algoritmo POLQA. Nessa seção busca-se a equação que relaciona o modelo utilizado nas medidas online com os obtidos pelos algoritmos PESQ e pelo POLQA. A forma de análise regressiva aplicada foi o modelo de regressão linear para se determinar a equação de reta que melhor se ajusta aos dados obtidos. Isso ocorreu pois, modelos que dependem de forma linear dos seus parâmetros desconhecidos são mais fáceis de ajustar que os modelos não-lineares aos seus parâmetros, e porque as propriedades estatísticas dos estimadores resultantes são fáceis de determinar.

Para todas as situação de medidas foram aplicados o modelo de regressão linear. Este método foi o primeiro a ser testado por ser o mais simples. Entretanto, como será visto, em um dos testes específicos este método não se mostrou eficiente.

Para cada reta ajustada, foi calculado o Coeficiente de Determinação também chamado de  $R^2$  que é uma medida de como um modelo estatístico linear se ajusta aos valores observados e varia entre 0 e 1, indicando, em percentagem, o quanto o modelo consegue explicar os valores observados. Quanto maior o  $R^2$ , melhor ele se ajusta à amostra e mais explicativo é modelo.

A Equação 5.1 de reta teve seus termos  $A_{PESQ}$  e  $B_{PESQ}$  calculada pelos métodos dos mínimos quadrados.

$$V_{PESQ} = A_{PESQ} \times V_{Modelo-E} + B_{PESQ} \quad (5.1)$$

onde:

$V_{PESQ}$  = Valor estimado por PESQ

$A_{PESQ}$  = Coeficiente Angular da Reta ajustada ao PESQ

$V_{Modelo-E}$  = Valor medido no processo

$B_{PESQ}$  = Coeficiente Linear da Reta ajustada ao PESQ

$R^2_{PESQ}$  = Coeficiente de Determinação da Reta PESQ

Tabela 5.28: Parâmetros do modelo de Regressão Linear para resultados do PESQ

Codec	$A_{PESQ}$	$B_{PESQ}$	$R^2_{PESQ}$
PCMU	1,0284	-0,8057	0,9715
GSM	0,4774	1,2494	0,9939
G.723	0,8326	0,3758	0,9707
G.729	0,9799	-0,4779	0,9517
G.726-32	0,9245	-0,4291	0,9074

Observa-se que em todos os casos, o Coeficiente de Determinação foi superior a 0,90 indicando uma forte correlação linear para os resultados do modelo proposto e os apresentados pelo algoritmo PESQ.

De modo análogo, para a relação entre o modelo utilizado e o POLQA temos a Equação 5.2:

$$V_{POLQA} = A_{POLQA} \times V_{Modelo-E} + B_{POLQA} \quad (5.2)$$

onde:

$V_{POLQA}$  = Valor estimado por POLQA

$A_{POLQA}$  = Coeficiente Angular da Reta ajustado ao POLQA

$V_{Modelo-E}$  = Valor medido no processo

$B_{POLQA}$  = Coeficiente Linear da Reta ajustado ao POLQA

$R^2_{POLQA}$  = Coeficiente de Determinação da Reta POLQA

Tabela 5.29: Parâmetros do modelo de Regressão Linear para resultados do POLQA

Codec	$A_{POLQA}$	$B_{POLQA}$	$R^2_{POLQA}$
PCMU	0,7343	-0,2886	0,9726
GSM	0,4363	0,7741	0,9888
G.723	0,716	-0,1583	0,9731
G.729	0,6259	-0,1555	0,6817
G.726-32	0,7563	-0,467	0,8872

Nesta situação, observa-se que para os codecs PCMU, GSM e G.723, os resultados de  $R^2$  foram superiores a 0,973 indicando uma correlação muito forte.  $R^2$  para o codec G.726-32 foi de 0,88 indicando uma correlação forte. Já para o codec G.729,  $R^2$  ficou em 0,67 indicando que o modelo linear pode não ser o mais adequado para explicar a relação entre os resultados do método proposto e os valores obtidos pelo algoritmo POLQA.

## 5.4 Classificação dos Codecs

Ao analisar-se os dados das Tabelas 5.12 a 5.27, é possível classificar os codificadores de acordo com seus valores médios das medidas do Modelo-E, PESQ e POLQA respectivamente.

As Tabelas 5.30 a 5.32 apresentam os resultados classificados das medidas de Qualidade de Experiência do Modelo-E proposto e dos testes padrões PESQ e POLQA, respectivamente. Qualquer uma das Tabelas 5.30 a 5.32 pode ser adotado por um sistema adaptativo com o objetivo de se obter a melhor qualidade em determinada situação.

Tabela 5.30: Recomendação de seleção de codecs de acordo com os resultados do Modelo-E.

Perda	1º	2º	3º	4º	5º
0,0%	GSM	PCMU	G723	G726-32	G729
1,0%	GSM	PCMU	G723	G726-32	G729
2,0%	GSM	PCMU	G723	G726-32	G729
2,5%	GSM	PCMU	G723	G726-32	G729
3,0%	GSM	PCMU	G723	G726-32	G729
5,0%	GSM	PCMU	G726-32	G723	G729
7,5%	GSM	PCMU	G726-32	G723	G729
10,0%	GSM	PCMU	G726-32	G723	G729

Tabela 5.31: Recomendação de seleção de codecs de acordo com os resultados do PESQ.

Perda	1º	2º	3º	4º	5º
0,0%	G723	PCMU	GSM	G726-32	G729
1,0%	G723	PCMU	GSM	G729	G726-32
2,0%	G723	PCMU	GSM	G726-32	G729
2,5%	G723	PCMU	GSM	G729	G726-32
3,0%	G723	PCMU	GSM	G726-32	G729
5,0%	G723	PCMU	GSM	G726-32	G729
7,5%	G723	PCMU	GSM	G726-32	G729
10,0%	G723	PCMU	GSM	G726-32	G729

Tabela 5.32: Recomendação de seleção de codecs de acordo com os resultados do POLQA.

Perda	1º	2º	3º	4º	5º
0,0%	PCMU	GSM	G723	G726-32	G729
1,0%	PCMU	GSM	G723	G726-32	G729
2,0%	PCMU	GSM	G723	G726-32	G729
2,5%	PCMU	GSM	G723	G726-32	G729
3,0%	PCMU	GSM	G723	G726-32	G729
5,0%	PCMU	GSM	G723	G726-32	G729
7,5%	PCMU	GSM	G723	G726-32	G729
10,0%	PCMU	GSM	G723	G726-32	G729

## Capítulo 6

# Conclusões e Trabalhos Futuros

O sistema proposto mostrou-se eficiente em medir a Qualidade de Experiência durante umas chamadas VoIP. Para o codec G.723, as medidas obtidas pelo método proposto foram bem próximas as medidas obtidas pelo PESQ. Entretanto, todas chamadas tiveram um escore  $MOS_{LQO}$  baixo quando comparado ao medido pelo algoritmo POLQA.

O coeficiente linear de Pearson apresentou uma correlação forte a extremamente forte, entre Modelo-E e PESQ e entre Modelo-E e POLQA, para os codecs PCMU, G.729, G.726-32. Já para os codecs GSM e G.723, a correlação linear de Pearson apresentou, principalmente nas menores Taxas de Perdas (0% e 1%), uma correlação linear fraca, próxima de zero e alguns momentos, negativa. Tal fato indica que a correlação não é linear. Entretanto, não se exclui a possibilidade da correlação apresentar comportamento não-linear. Outros métodos podem ser aplicados como ANOVA ou T-Student em busca de uma correlação entre as medidas Modelo-E e PESQ e entre Modelo-E e POLQA.

Mesmo sendo métodos de medidas diferentes, os resultados dos algoritmos PESQ e POLQA apresentaram uma forte correlação linear, como pode ser observado nas Tabelas 5.12 a 5.27. Entretanto, a diferença entre as medidas dos algoritmos PESQ e POLQA deve ser levada em consideração. POLQA apresentou sempre as medidas de  $MOS_{LQO}$  menores que as medidas apresentadas pelo algoritmo PESQ. Em muitos casos, a diferença é de um ponto na escala de 1 a 5.

Pelos resultados apresentados na Subseção 5.2.2 do Capítulo 5 foi possível estabelecer um critério de seleção de codec baseado na médias das medidas de Qualidade. Entretanto, considerando-se os desvios padrões dos valores obtidos, qualquer critério adotado apresenta um empate técnico entre os primeiros codecs candidatos. Um critério de desempate poderia ser aplicado como largura de banda, por exemplo, se a base de dados contemplasse tal critério.

Dentre as restrições do sistema de ensaio VoIP (*test-bed*) tem-se as perdas de pacotes gerados durante os ensaios e testes do Modelo-E. Estas perdas de pacotes foram perdas aleatórias com distribuição uniformes gerados pelo Netem. Contudo, no mundo real, as perdas de pacotes são frequentemente não uniformes e podem ocorrer em rajadas. Sabe-se que a perda de pacotes em rajadas reduz a qualidade de voz mais do que a perda aleatória uniforme de pacotes. Portanto, no mundo real, a ocorrência de perdas de pacotes em rajada provavelmente irá resultar em valores menores de MOS.

A presente Tese propôs uma forma de se medir a Qualidade de Experiência (QoE) durante uma chamada VoIP utilizando-se do Modelo-E modificado. Modificações nos métodos de medida de Qualidade de Serviço propostos na literatura, se mostraram eficientes na determinação da Qualidade de Experiência em tempo real durante uma chamada VoIP.

Desenvolveu-se, implantou-se e validou-se uma rotina para avaliar a qualidade instantânea e acumulada da Qualidade de Experiência de uma chamada VoIP na estação receptora do fluxo de dados. As medidas instantâneas fizeram pouco sentido quando analisadas isoladamente, mas a medida de Qualidade de Experiência acumulada representa bem a medida de Qualidade quando comparado com outros métodos, por exemplo, os algoritmos PESQ e POLQA.

Foram efetuadas 11.560 chamadas VoIP em diversas configurações de chamada com os codecs em teste e em diferentes condições de perdas de pacotes na rede. Essas chamadas foram analisadas e testadas formando a base de conhecimento proposta nesta Tese. Esta base de conhecimento conta com mais de 200 horas de registros de chamadas, que pode sugerir, como apresentado no Capítulo 5 na Seção 5.4, um codec candidato que melhor se adéque a uma nova condição de rede. Para se chegar a isto, analisou-se e compararam-se os resultados obtidos em tempo real com os obtidos a partir de outros mecanismos objetivos de avaliação da qualidade da fala, neste caso, PESQ e POLQA, classificando e ordenando os resultados obtidos pelas diferentes formas de medidas de Qualidade de Experiência.

As ferramentas e métodos apresentados no Capítulo 4 se mostraram coerentes e eficientes na produção dos dados analisados no Capítulo 5.

As medidas pontuais (contínuas) médias representam a qualidade acumulada da chamada e representam uma medida capaz de alimentar o mecanismo de seleção de codec de um sistema adaptativo baseado na troca de codec.

O método proposto no Capítulo 4 apresenta condições de ser utilizado para outras restrições de rede, como por exemplo, alterações na largura de banda e *jitter*, aumentando a base de conhecimento.

Com a ferramenta PEXQ (OPTICOM, 2007), que é um suíte de testes para a medição da qualidade perceptual de voz, foi possível medir a Qualidade de Experiência do sinal resultante da transmissão por meio de mecanismos objetivos padronizados de medida de qualidade de fala como o PESQ (ITU-T, 2001) e POLQA (ITU-T, 2014). Cabe ressaltar que os artigos apresentados no Capítulo 3 de Trabalhos Relacionados não utilizam o POLQA, por ser um método mais novo e de maior custo.

## Trabalhos Futuros

Como sugestões para trabalhos futuros, é possível citar:

- repetir os testes com frases balanceadas em Português do Brasil. Para isso será necessário selecionar um conjunto de frases balanceadas. Isto se justifica, pois existem trabalhos que estudam a influência do idioma na Qualidade de Experiência de comunicações VoIP;

- realizar os mesmos testes com codecs não ITUT-T, estendendo o fator de comprometimento de equipamentos para codecs não ITU-T, tais como o codec Speex e iLBC, por exemplo;
- repetir os testes com codecs da Banda Larga (*wide-band*);
- adicionar à biblioteca de codificadores disponíveis, o OPUS, codec de Banda Larga programável;
- executar testes com perdas não uniformes de pacotes e perdas em rajadas para testar-se a qualidade da chamada medida pelo Modelo-E;
- executar os testes com limitação de banda ao invés de variação na Taxa de Perdas; e
- expandir a base de conhecimento para codecs de Banda Larga, não coberto neste trabalho.

Como aplicação futura, esta Tese poderá ser a base do conhecimento para a tomada de decisão de um futuro sistema adaptativo de controle de qualidade da fala, baseado em *clusters* de codecs. Os resultados podem melhorar não somente os sistemas de monitoramento da qualidade de voz em redes corporativas, como também poderão ser utilizados em sistemas móveis, por exemplo. O futuro sistema poderá escolher um novo codec candidato de acordo com a tendência de variáveis monitoradas. Aprendizagem de máquina poderá ser utilizada a medida que novos codificadores aparecerem no mercado. Três fatores devem ser considerados: a) Quando mudar os codecs, b) o codec utilizado e o codec escolhido para substituí-lo e c) as razões que levaram à decisão da escolha do novo codec.

## Referências Bibliográficas

- AKTAS, I.; SCHMIDT, F.; WEINGÄRTNER, E.; SCHNELKE, C.-J.; WEHRLE, K. An adaptive codec switching scheme for sip-based voip. In: ANDREEV, S.; BALANDIN, S.; KOUCHERYAVY, Y. (Ed.). *Internet of Things, Smart Spaces, and Next Generation Networking*. Springer Berlin Heidelberg, 2012, (Lecture Notes in Computer Science, v. 7469). p. 347–358. ISBN 978-3-642-32685-1. Disponível em: [http://dx.doi.org/10.1007/978-3-642-32686-8\\_32](http://dx.doi.org/10.1007/978-3-642-32686-8_32). 43
- ASSEM, H.; ADEL, M.; JENNINGS, B.; MALONE, D.; DUNNE, J.; O’SULLIVAN, P. A generic algorithm for mid-call audio codec switching. In: *Integrated Network Management (IM 2013), 2013 IFIP/IEEE International Symposium on*. [S.l.: s.n.], 2013. p. 1276–1281. 24, 43
- ASSEM, H.; ADEL, M.; JENNINGS, B.; MALONE, D.; DUNNE, J.; O’SULLIVAN, P. Online estimation of vvoip quality-of-experience via network emulation. In: *IET. Signals and Systems Conference (ISSC 2013), 24th IET Irish*. [S.l.], 2013. p. 1–8. 43, 45
- BERG, H. van den; BOHNERT, T. M.; CABRAL, O.; MOLTCHANOV, D.; STAEHLE, D.; VELEZ, F. Performance evaluation and traffic modeling. In: SIRIS, V.; BRAUN, T.; BARCELO-ARROYO, F.; STAEHLE, D.; GIAMBENE, G. (Ed.). *Traffic and QoS Management in Wireless Multimedia Networks*. [S.l.]: Springer, 2009. v. 31, p. 89–150. 35
- CANO, M.-D.; CERDAN, F. Subjective qoe analysis of voip applications in a wireless campus environment. *Telecommun. Syst.*, Kluwer Academic Publishers, Norwell, MA, USA, v. 49, n. 1, p. 5–15, jan. 2012. ISSN 1018-4864. Disponível em: <http://dx.doi.org/10.1007/s11235-010-9348-5>. 20, 29, 36
- CARVALHO, L. S. G. *Implementação do Modelo-E para Avaliação Objetiva da Qualidade da Fala em Redes de Comunicação VoIP*. Dissertação (Mestrado) — Instituto de Ciências Exatas (ICE), Universidade Federal do Amazonas (UFAM), Manaus, AM, Brasil, 2004. Dissertação (Mestrado em Ciência da Computação), Programa de Pós-Graduação em Informática. 23, 28
- CARVALHO, L. S. G. *Adaptive Management of Speech Quality Between VoIP Terminals*. Tese (Doutorado) — Instituto de Computação (ICOMP), Universidade Federal do Amazonas (UFAM), Manaus, AM, Brasil, 2011. Tese (Doutorado em Ciência da Computação), Programa de Pós-Graduação em Informática. 21, 23, 25, 26, 27, 28, 39, 43, 45, 46
- CARVALHO, L. S. G.; MOTA, E.; AGUIAR, R.; LIMA, A. F.; BARRETO, A. *et al.* An e-model implementation for speech quality evaluation in voip systems. *2014 IEEE Symposium on Computers and Communications (ISCC)*, IEEE Computer Society, v. 0, p. 933–938, 2005. ISSN 1530-1346. 44

- COSTA, N.; NUNES, M. S. Adaptive quality of service in voice over ip communications. In: *Networking and Services, 2009. ICNS '09. Fifth International Conference on*. [S.l.: s.n.], 2009. p. 19–24. 38, 43
- COTANIS, I. Understanding the transition from pesq to polqa - an ascom network testing white paper. ASCOM, Vienna (Austria), 2011. Disponível em: <http://www.ascom.at/cn/understanding-the-transition-from-pesq-to-polqa.pdf>. 33
- DAI, Q. A survey of quality of experience. In: *Energy-Aware Communications*. [S.l.]: Springer, 2011. p. 146–156. 34, 36
- ETR250, E. *Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks*. 1996. Technical Report ETR 250. 37
- FOUNDATION, L. *Netem: Network Emulator*. 2009. <http://www.linuxfoundation.org/collaborate/workgroups/networking/netem>. [Online; accessed 02 2015]. 57, 61
- FRIEDMAN, T.; CACERES, R.; CLARK, A. *RTP Control Protocol Extended Reports (RTCP XR). Request for Comments (RFC) 3611*. 2003. [Online; accessed 28-Fevereiro-2015]. Disponível em: <https://tools.ietf.org/html/rfc3611>. 26, 46
- GONG, Y.; YANG, F.; HUANG, L.; SU, S. Model-based approach to measuring quality of experience. In: *IEEE. Emerging Network Intelligence, 2009 First International Conference on*. [S.l.], 2009. p. 29–32. 36, 44
- HAO, C. J. *Format Factory 3.6.0*. 2007–2014. <http://www.pcfreetime.com/>. 48
- HERSHEY, P. C.; PITTS, J. M.; OGILVIE, R. Monitoring real-time applications events in net-centric enterprise systems to ensure high quality of experience. In: *IEEE. Military Communications Conference, 2009. MILCOM 2009. IEEE*. [S.l.], 2009. p. 1–7. 44
- HINES, A.; SKOGLUND, J.; KOKARAM, A.; HARTE, N. Robustness of speech quality metrics to background noise and network degradations: Comparing visqol, pesq and polqa. In: *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. [S.l.: s.n.], 2013. p. 3697–3701. ISSN 1520-6149. 33
- HOSSFELD, T.; BINZENHÖFER, A. Analysis of skype voip traffic in umts: End-to-end qos and qoe measurements. *Computer Networks*, Elsevier, v. 52, n. 3, p. 650–666, 2008. 37
- HOSSFELD, T.; TRAN-GIA, P.; FIEDLER, M. Quantification of quality of experience for edge-based applications. In: *Managing Traffic Performance in Converged Networks*. [S.l.]: Springer, 2007. p. 361–373. 34, 35
- HUNTGEBURTH, B.; SCHUMANN, S.; LONDÁK, J. Voice over ip (voip) speech quality measurement with open-source software components. In: *ELMAR, 2010 PROCEEDINGS*. [S.l.: s.n.], 2010. 44
- INTEL. *Intel Integrated Performance Primitives Reference Manual*. 2012. [https://software.intel.com/sites/default/files/m/d/4/1/d/8/ipp\\_manual.zip](https://software.intel.com/sites/default/files/m/d/4/1/d/8/ipp_manual.zip). [Online; accessed 24 05 2015]. 47, 49

ITU-T. Rec. p.800: Methods for subjective determination of transmission quality. *ITU-T Recommendation*, Geneva (Switzerland), ago. 1996. 30, 31, 60

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. *ITU-T Recommendation*, Geneva (Switzerland), fev. 2001. Disponível em: <http://www.itu.int/rec/T-REC-P.862>. 21, 32, 33, 86

ITU-T. Rec. g.114: One-way transmission time. *ITU-T Recommendation*, Geneva (Switzerland), maio 2003. 26, 40, 41

ITU-T. Rec. g.113: Transmission impairments due to speech processing appendix i: Provisional planning values for the equipment impairment factor ie. *ITU-T Recommendation*, Geneva (Switzerland), nov. 2007. Disponível em: <http://www.itu.int/rec/T-REC-G.113-200711-I/en>. 37, 38, 41, 42, 51

ITU-T. Rec. e.800: Definitions of terms related to quality of service. *ITU-T Recommendation*, Geneva (Switzerland), set. 2008. 34

ITU-T. Rec. p.863: Perceptual objective listening quality assessment (polqa). *ITU-T Recommendation*, Geneva (Switzerland), set. 2014. 21, 32, 33, 86

ITU-T. Rec. g.107: The e-model, a computational model for use in transmission planning. *ITU-T Recommendation*, Geneva (Switzerland), jun. 2015. 33, 37, 38, 39, 40, 42, 60

JAMMEH, E.; MKWAWA, I.; KHAN, A.; GOUDARZI, M.; SUN, L.; IFEACHOR, E. Quality of experience (qoe) driven adaptation scheme for voice/video over ip. *Telecommunication Systems*, Springer, v. 49, n. 1, p. 99–111, 2012. 44

JIANG, C.; HUANG, P. Research of monitoring voip voice qos. In: IEEE. *2011 International Conference on Internet Computing and Information Services*. [S.l.], 2011. p. 499–502. 44

KARAPANTAZIS, S.; PAVLIDOU, F.-N. Voip: A comprehensive survey on a promising technology. *Computer Networks*, Elsevier North-Holland, Inc., New York, NY, USA, v. 53, n. 12, p. 2050–2090, ago. 2009. ISSN 13891286. Disponível em: <http://dx.doi.org/10.1016/j.comnet.2009.03.010>. 19, 23, 36, 43

KIM, H. L.; CHOI, S. G. Traffic quality monitoring system between different network providers. In: IEEE. *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*. [S.l.], 2010. v. 2, p. 1153–1158. 44

KUIPERS, F.; KOOIJ, R.; VLEESCHAUWER, D. D.; BRUNNSTRÖM, K. Techniques for measuring quality of experience. In: *Wired/wireless internet communications*. [S.l.]: Springer, 2010. p. 216–227. 36

KUROSE, J. F.; ROSS, K. W.; MARQUES, A. S.; ZUCCHI, W. L. *Redes de computadores e a Internet: Uma abordagem top-down*. 6. ed. São Paulo: Pearson Education do Brasil, 2013. 24

LAMARE, R. C. d. *CELP Coder for Speech Signals of Brazilian Portuguese*. Dissertação (Mestrado) — Poli/UFRJ, Rio de Janeiro, RJ, Brasil, 1999. 28

LTD, T. *PJSUA Manual Page*. 2007. <<http://www.pjsip.org/pjsua.htm>>. [Online; accessed 28-Fevereiro-2015]. 57

LUSTOSA, L. C.; CARVALHO, L. S.; RODRIGUES, P. H. d. A.; MOTA, E. d. S. Utilização do modelo-e para avaliação da qualidade da fala em sistemas de comunicação baseados em voz sobre ip. *Proceedings of XXII SBRC*, p. 603–616, 2004. 38, 39

MANOUSOS, M.; APOSTOLACOS, S.; GRAMMATIKAKIS, I.; MEXIS, D.; KAGKLIS, D.; SYKAS, E. Voice-quality monitoring and control for voip. *Internet Computing, IEEE*, v. 9, n. 4, p. 35–42, jul. 2005. ISSN 1089-7801. 43

MOHAMED, A. M. *Assessment and Improvement of the Quality of Voice-over-IP Communications*. Tese (Doutorado) — Waterford Institute of Technology, 2013. 45

MUKAKA, M. A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, Medical Association of Malawi, v. 24, n. 3, p. 69–71, 2012. 59

MÜLLER, H.; PEZZÈ, M.; SHAW, M. Visibility of control in adaptive systems. In: *Proceedings of the 2Nd International Workshop on Ultra-large-scale Software-intensive Systems*. New York, NY, USA: ACM, 2008. (ULSSIS '08), p. 23–26. ISBN 978-1-60558-026-5. Disponível em: <<http://doi.acm.org/10.1145/1370700.1370707>>. 20, 29

MYAKOTNYKH, E. S.; THOMPSON, R. A. Adaptive speech quality management in voice-over-ip communications. In: *Telecommunications, 2009. AICT '09. Fifth Advanced International Conference on*. [S.l.: s.n.], 2009. p. 64–71. 21, 41, 42, 43

NOKIA. Quality of experience (qoe) of mobile services: Can it be measured and improved? *White Paper*, Finland, ago. 2004. 34, 35

OPTICOM, T. P. Q. E. *PEXQ - perceptual quality by OPTICOM: all-in-one testing suite*. 2007. <<http://www.opticom.de/products/pexq.html>>. [Online; accessed 04 2015]. 33, 48, 86

PETERSON, L. L.; DAVIE, B. S. *Redes de Computadores (Uma abordagem de sistemas)*, 4ª edição. Rio de Janeiro: Campus (Elsevier), 2004. 24, 25

PRIJONO, B.; ISMANGIL, P. *PJSIP*. 2003. <<http://www.pjsip.org/>>. [Online; accessed 28-Fevereiro-2015]. 47

PROJECT, O. *Open Speech Repository*. 2015. <[http://www.voiptroubleshooter.com/open\\_speech/american.html](http://www.voiptroubleshooter.com/open_speech/american.html)>. [Online; accessed 24 02 2015]. 49, 50, 58, 61, 63, 65, 95

QIAO, Z.; SUN, L.; HEILEMANN, N.; IFEACHOR, E. A new method for voip quality of service control use combined adaptive sender rate and priority marking. In: *Communications, 2004 IEEE International Conference on*. [S.l.: s.n.], 2004. v. 3, p. 1473–1477. 43

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2015. 47

RAHRER, T.; FIANDRA, R.; WRIGHT, S. *et al.* Triple-play services quality of experience (qoe) requirements and mechanisms - for architecture & transport. In: *DSL-Forum. Technical Report TR-126*. [S.l.: s.n.], 2006. 34, 35

RODRIGUES, D.; CERQUEIRA, E.; MONTEIRO, E. Qoe assessment of voip in next generation networks. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. [S.l.: s.n.], 2009. v. 5842 LNCS, p. 94–105. ISBN 3642049931. ISSN 03029743. 20

RODRÍGUEZ, D. Z. *Algoritmo para determinação da taxa de transmissão em uma rede IP*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo (USP), São Paulo, SP, Brasil, 2009. 21, 27, 28

ROSENBERG, J. Rfc 3311: The session initiation protocol (sip) update method. *IETF, September, 10/2002*. 47

ROSENBERG, J.; SCHULZRINNE, H.; CAMARILLO, G.; JOHNSTON, A.; PETERSON, J.; SPARKS, R.; HANDLEY, M.; SCHOOLER, E. *RFC 3261: SIP: session initiation protocol*. [S.l.]: IETF, Tech. Rep., 2002.[Online]. Available: [www.ietf.org/rfc/rfc3261.txt](http://www.ietf.org/rfc/rfc3261.txt), 2003. 46, 47

ROTHAUSER, E.; CHAPMAN, W.; GUTTMAN, N.; NORDBY, K.; SILBIGER, H.; URBANEK, G.; WEINSTOCK, M. Ieee recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust*, v. 17, n. 3, p. 225–246, 1969. 49, 95

SCHULZRINNE, H.; CASNER, S.; FREDERICK, R.; JACOBSON, V. rfc3550. a transport protocol for real-time applications. *IETF*, jul. 2003. 26

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. Mean opinion score measurements based on e model during a voip call. In: *Telecommunications, 2015. AICT 2015, The Eleventh Advanced International Conference on*. Brussels, Belgium: [s.n.], 2015. p. 132–135. ISBN 978-1-61208-411-4. ISSN 2308-4030. Disponível em: [https://www.thinkmind.org/download.php?articleid=aict\\_2015\\_7\\_10\\_10047](https://www.thinkmind.org/download.php?articleid=aict_2015_7_10_10047). 58, 60, 61

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. Medidas de mean opinion score baseada no modelo-e durante uma chamada voip. In: *CACSI, 2015. Anais do I Congresso Amazônico de Computação e Sistemas Inteligentes*. Manaus, Brasil: [s.n.], 2015. p. 110–118. ISSN 2447-0414. Disponível em: <http://www.cacsi.com.br/anais/download.php>. 58

SILVA, E. L. O. d.; CARVALHO, L. S. G.; MOTA, E. S.; IANO, Y. Qoe-based adaptive control of speech quality in a voip call. In: *Telecommunications, 2015. AICT 2015, The Eleventh Advanced International Conference on*. Brussels, Belgium: [s.n.], 2015. p. 136–137. ISBN 978-1-61208-411-4. ISSN 2308-4030. Disponível em: [https://www.thinkmind.org/download.php?articleid=aict\\_2015\\_7\\_10\\_10047](https://www.thinkmind.org/download.php?articleid=aict_2015_7_10_10047). 58

SILVA, J. M. da; LINS, R. D. Analyzing the qos of voip on sip in java. In: *Proc. 2006 International Telecommunications Symposium*. [S.l.: s.n.], 2006. p. 576–581. 44

SOFTWARE, P. *Direct WAV MP3 Splitter Version 3.0.0.0*. 2004–2016. <http://www.pistonsoft.com/>. 48

SUN, L. *Speech quality prediction for voice over Internet protocol networks*. Tese (Doutorado) — School of Computing, Communications and Electronics, Faculty of Technology, University of Plymouth, 2004. 27, 28

SUN, L.; MKWAWA, I.-H.; JAMMEH, E.; IFEACHOR, E. Voip quality of experience (qoe). In: *Guide to voice and video over IP*. [S.l.]: Springer, 2013. p. 123–162. 36

TANENBAUM, A. S.; WETHERALL, D. J. *Redes de computadores*. 5. ed. São Paulo: Pearson Prentice Hall, 2011. ISBN 9788576059240. 24, 25

VIANNA, B. d. A.; MOURA, N. T.; ALBUQUERQUE, C. V. N. d.; REBELLO, V. E. F.; BOERES, C. adamos algoritmo mos-adaptativo para fontes voip. In: *WebMedia 2006, XII Simpósio Brasileiro de Sistemas Multímídia e Web*. Natal, RN, Brasil: [s.n.], 2006. p. 223–232. 43

# Anexo

## Anexo A

### Sentenças de Harvard

Este Anexo apresenta a transcrição de cada arquivo testado nesta Tese. Foi necessário ouvi-las e relaciona-las a Lista Revisada de sentenças foneticamente equilibradas “(frases de Harvard)” (ROTHAUSER *et al.*, 1969). No site da OSR (PROJECT, 2015) não foi encontrada tal transcrição das frases.

#### A.1 OSR\_us\_000\_0010\_8k.wav Feminina 1

Lista 1

1. The birch canoe slid on the smooth planks.
2. Glue the sheet to the dark blue background.
3. It's easy to tell the depth of a well.
4. These days a chicken leg is a rare dish.
5. Rice is often served in round bowls.
6. The juice of lemons makes fine punch.
7. The box was thrown beside the parked truck.
8. The hogs were fed chopped corn and garbage.
9. Four hours of steady work faced us.
10. Large size in stockings is hard to sell.

#### A.2 OSR\_us\_000\_0011\_8k.wav Feminina 1

Lista2

1. The boy was there when the sun rose.
2. A rod is used to catch pink salmon.

3. The source of the huge river is the clear spring.
4. Kick the ball straight and follow through.
5. Help the woman get back to her feet.
6. A pot of tea helps to pass the evening.
7. Smoky fires lack flame and heat.
8. The soft cushion broke the man's fall.
9. The salt breeze came across from the sea.
10. The girl at the booth sold fifty bonds.

### **A.3 OSR\_us\_000\_0012\_8k.wav Feminina 1**

#### Lista 3

1. The small pup gnawed a hole in the sock.
2. The fish twisted and turned on the bent hook.
3. Press the pants and sew a button on the vest.
4. The swan dive was far short of perfect.
5. The beauty of the view stunned the young boy.
6. Two blue fish swam in the tank.
7. Her purse was full of useless trash.
8. The colt reared and threw the tall rider.
9. It snowed, rained, and hailed the same morning.
10. Read verse out loud for pleasure.

### **A.4 OSR\_us\_000\_0013\_8k.wav Feminina 1**

#### Lista 4

1. Hoist the load to your left shoulder.
2. Take the winding path to reach the lake.
3. Note closely the size of the gas tank.
4. Wipe the grease off his dirty face.

5. Mend the coat before you go out.
6. The wrist was badly strained and hung limp.
7. The stray cat gave birth to kittens.
8. The young girl gave no clear response.
9. The meal was cooked before the bell rang.
10. What joy there is in living.

## **A.5 OSR\_us\_000\_0014\_8k.wav Feminina 1**

### Lista5

1. A king ruled the state in the early days.
2. The ship was torn apart on the sharp reef.
3. Sickness kept him home the third week.
4. The wide road shimmered in the hot sun.
5. The lazy cow lay in the cool grass.
6. Lift the square stone over the fence.
7. The rope will bind the seven books at once.
8. Hop over the fence and plunge in.
9. The friendly gang left the drug store.
10. Mesh mire keeps chicks inside.

## **A.6 OSR\_us\_000\_0015\_8k.wav Feminina 2**

### Lista 1

1. The birch canoe slid on the smooth planks.
2. Glue the sheet to the dark blue background.
3. It's easy to tell the depth of a well.
4. These days a chicken leg is a rare dish.
5. Rice is often served in round bowls.
6. The juice of lemons makes fine punch.

7. The box was thrown beside the parked truck.
8. The hogs were fed chopped corn and garbage.
9. Four hours of steady work faced us.
10. Large size in stockings is hard to sell.

## **A.7 OSR\_us\_000\_0016\_8k.wav Feminina 2**

### Lista 2

1. The boy was there when the sun rose.
2. A rod is used to catch pink salmon.
3. The source of the huge river is the clear spring.
4. Kick the ball straight and follow through.
5. Help the woman get back to her feet.
6. A pot of tea helps to pass the evening.
7. Smoky fires lack flame and heat.
8. The soft cushion broke the man's fall.
9. The salt breeze came across from the sea.
10. The girl at the booth sold fifty bonds.

## **A.8 OSR\_us\_000\_0017\_8k.wav Feminina 2**

### Lista 3

1. The small pup gnawed a hole in the sock.
2. The fish twisted and turned on the bent hook.
3. Press the pants and sew a button on the vest.
4. The swan dive was far short of perfect.
5. The beauty of the view stunned the young boy.
6. Two blue fish swam in the tank.
7. Her purse was full of useless trash.
8. The colt reared and threw the tall rider.
9. It snowed, rained, and hailed the same morning.
10. Read verse out loud for pleasure.

## **A.9 OSR\_us\_000\_0018\_8k.wav Feminina 2**

### Lista 4

1. Hoist the load to your left shoulder.
2. Take the winding path to reach the lake.
3. Note closely the size of the gas tank.
4. Wipe the grease off his dirty face.
5. Mend the coat before you go out.
6. The wrist was badly strained and hung limp.
7. The stray cat gave birth to kittens.
8. The young girl gave no clear response.
9. The meal was cooked before the bell rang.
10. What joy there is in living.

## **A.10 OSR\_us\_000\_0019\_8k.wav Feminina 2**

### Lista 5

1. A king ruled the state in the early days.
2. The ship was torn apart on the sharp reef.
3. Sickness kept him home the third week.
4. The wide road shimmered in the hot sun.
5. The lazy cow lay in the cool grass.
6. Lift the square stone over the fence.
7. The rope will bind the seven books at once.
8. Hop over the fence and plunge in.
9. The friendly gang left the drug store.
10. Mesh mire keeps chicks inside.

**A.11 OSR\_us\_000\_0030\_8k.wav Masculina**

Lista 57

1. Paint the sockets in the wall dull green.
2. The child crawled into the dense grass.
3. Bribes fail where honest men work.
4. Trample the spark, else the flames will spread.
5. The hilt. of the sword was carved with fine designs.
6. A round hole was drilled through the thin board.
7. Footprints showed the path he took up the beach.
8. She was waiting at my front lawn.
9. A vent near the edge brought in fresh air.
10. Prod the old mule with a crooked stick.

**A.12 OSR\_us\_000\_0031\_8k.wav Masculina**

Lista 59

1. Every word and phrase he speaks is true.
2. He put his last cartridge into the gun and fired.
3. They took their kids from the public school.
4. Drive the screw straight into the wood.
5. Keep the hatch tight and the watch constant.
6. Sever the twine with a quick snip of the knife.
7. Paper will dry out when wet.
8. Slide the catch back and open the desk.
9. Help the weak to preserve their strength.
10. A sullen smile gets few friends.

**A.13 OSR\_us\_000\_0032\_8k.wav Masculina**

## Lista 23

1. A pencil with black lead writes best.
2. Coax a young calf to drink from a bucket.
3. Schools for ladies teach charm and grace.
4. The lamp shone with a steady green flame.
5. They took the axe and the saw to the forest.
6. The ancient coin was quite dull and worn.
7. The shaky barn fell with a loud crash.
8. Jazz and swing fans like fast music.
9. Rake the rubbish up and then burn it.
10. Slash the gold cloth into fine ribbons.

**A.14 OSR\_us\_000\_0034\_8k.wav Masculina**

## Lista 72

1. A gold ring will please most any girl.
2. The long journey home took a year.
3. She saw a cat in the neighbor's house.
4. A pink shell was found on the sandy beach.
5. Small children came to see him.
6. The grass and bushes were wet with dew.
7. The blind man counted his old coins.
8. A severe storm tore down the barn.
9. She called his name many times.
10. When you hear the bell, come quickly.

**A.15 OSR\_us\_000\_0035\_8k.wav Masculina**

Lista 42

1. Nudge gently but wake her now.
2. The news struck doubt into restless minds.
3. Once we stood beside the shore.
4. A chink in the wall allowed a draft to blow.
5. Fasten two pins on each side.
6. A cold dip restores health and zest.
7. He takes the oath of office each March.
8. The sand drifts over the sill of the old house.
9. The point of the steel pen was bent and twisted.
10. There is a lag between thought and act.

**A.16 OSR\_us\_000\_0036\_8k.wav Masculina**

Lista 31

1. Slide the box into that empty space.
2. The plant grew large and green in the window.
3. The beam dropped down on the workmen's head.
4. Pink clouds floated with the breeze.
5. She danced like a swan, tall and graceful.
6. The tube was blown and the tire flat and useless.
7. It is late morning on the old wall clock.
8. Let's all join as we sing the last chorus.
9. The last switch cannot be turned off.
10. The fight will end in just six minutes.

**A.17 OSR\_us\_000\_0037\_8k.wav Masculina**

Lista 32 - OSR\_us\_000\_0037\_8k.wav Masculina

1. The store walls were lined with colored frocks.
2. The peace league met to discuss their plans.
3. The rise to fame of a person takes luck.
4. Paper is scarce, so write with much care.
5. The quick fox jumped on the sleeping cat.
6. The nozzle of the fire hose was bright brass.
7. Screw the round cap on as tight as needed.
8. Time brings us many changes.
9. The purple tie was ten years old.
10. Men think and plan and sometimes act.

**A.18 OSR\_us\_000\_0038\_8k.wav Masculina**

Lista 33

1. Fill the ink jar with sticky glue.
2. He smoke a big pipe with strong contents.
3. We need grain to keep our mules healthy.
4. Pack the records in a neat thin case.
5. The crunch of feet in the snow was the only sound.
6. The copper bowl shone in the sun's rays.
7. Boards will warp unless kept dry.
8. The plush chair leaned against the wall.
9. Glass will clink when struck by metal.
10. Bathe and relax in the cool green grass.

**A.19 OSR\_us\_000\_0039\_8k.wav Masculina**

## Lista 34

1. Nine rows of soldiers stood in line.
2. The beach is dry and shallow at low tide.
3. The idea is to sew both edges straight.
4. The kitten chased the dog down the street.
5. Pages bound in cloth make a book.
6. Try to trace the fine lines of the painting.
7. Women form less than half of the group.
8. The zones merge in the central part of town.
9. A gem in the rough needs work to polish.
10. Code is used when secrets are sent.

**A.20 OSR\_us\_000\_0040\_8k.wav Masculina**

## Lista 35

1. Most of the new is easy for us to hear.
2. He used the lathe to make brass objects.
3. The vane on top of the pole revolved in the wind.
4. Mince pie is a dish served to children.
5. The clan gathered on each dull night.
6. Let it burn, it gives us warmth and comfort.
7. A castle built from sand fails to endure.
8. A child's wit saved the day for us.
9. Tack the strip of carpet to the worn floor.
10. Next Tuesday we must vote.

**A.21 OSR\_us\_000\_0057\_8k.wav Masculina**

## Lista 25

1. On the islands the sea breeze is soft and mild.
2. The play began as soon as we sat down.
3. This will lead the world to more sound and fury.
4. Add salt before you fry the egg.
5. The rush for funds reached its peak Tuesday.
6. The birch looked stark white and lonesome.
7. The box is held by a bright red snapper.
8. To make pure ice, you freeze water.
9. The first worm gets snapped early.
10. Jump the fence and hurry up the bank.

**A.22 OSR\_us\_000\_0059\_8k.wav Masculina**

## Lista 27

1. The dark pot hung in the front closet.
2. Carry the pail to the wall and spill it there.
3. The train brought our hero to the big town.
4. We are sure that one war is enough.
5. Gray paint stretched for miles around.
6. The rude laugh filled the empty room.
7. High seats are best for football fans.
8. Tea served from the brown jug is tasty.
9. A dash of pepper spoils beef stew.
10. A zestful food is the hot-cross bun.

**A.23 OSR\_us\_000\_0060\_8k.wav Masculina**

## Lista 28

1. The horse trotted around the field at a brisk pace.
2. Find the twin who stole the pearl necklace.
3. Cut the cord that binds the box tightly.
4. The red tape bound the smuggled food.
5. Look in the corner to find the tan shirt.
6. The cold drizzle will halt the bond drive.
7. Nine men were hired to dig the ruins.
8. The junk yard had a mouldy smell.
9. The flint sputtered and lit a pine torch.
10. Soak the cloth and drown the sharp odor.

**A.24 OSR\_us\_000\_0061\_8k.wav Masculina**

## Lista 30

1. The mute muffled the high tones of the horn.
2. The gold ring fits only a pierced ear.
3. The old pan was covered with hard fudge.
4. Watch the log float in the wide river.
5. The node on the stalk of wheat grew daily.
6. The heap of fallen leaves was set on fire.
7. Write fast, if you want to finish early.
8. His shirt was clean but one button was gone.
9. The barrel of beer was a brew of malt and hops.
10. Tin cans are absent from store shelves.