



ADILSON TORRES

**ESTUDO DA QUALIDADE DA VOZ EM SISTEMAS DE
COMUNICAÇÃO MÓVEL ATRAVÉS DO CODEC EVS
UTILIZANDO AS MÉTRICAS DE QUALIDADE P.862 E
E-MODEL.**

**LAVRAS – MG
2023**

ADILSON TORRES

**ESTUDO DA QUALIDADE DA VOZ EM SISTEMAS DE COMUNICAÇÃO MÓVEL ATRAVÉS
DO CODEC EVS UTILIZANDO AS MÉTRICAS DE QUALIDADE P.862 E E-MODEL.**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para a obtenção do título de Mestre.

Prof. DSc. Demóstenes Zegarra Rodríguez
Orientador

**LAVRAS – MG
2023**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Torres, Adilson

Estudo da Qualidade da Voz em Sistemas de Comunicação Móvel através do Codec EVS utilizando as métricas de qualidade P.862 e E-Model. / ADÍLSON TORRES - 2023.. – Lavras : UFLA, 2023.

105 p. :

Dissertação (Mestrado) –Universidade Federal de Lavras, 2023.

Orientador: Prof. DSc. Demóstenes Zegarra Rodríguez.

Bibliografia.

1. Avaliação da Qualidade da Voz. 2. P.862. 3. E-Model. I. Rodríguez, Demóstenes Zegarra. II. Título.

ADILSON TORRES

**ESTUDO DA QUALIDADE DA VOZ EM SISTEMAS DE COMUNICAÇÃO MÓVEL ATRAVÉS
DO CODEC EVS UTILIZANDO AS MÉTRICAS DE QUALIDADE P.862 E E-MODEL.**

**VOICE QUALITY ASSESSMENT FOR COMMUNICATION SYSTEMS WITH EVS CODEC
USING P.862 AND E-MODEL ALGORITHMS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para a obtenção do título de Mestre.

APROVADA em 30 de Junho de 2023.

Prof. DSc. Demóstenes Zegarra Rodríguez	UFLA
Prof. DSc. Renata Lopes Rosa	UFLA
Profa. Esp. Dante Coaquira Begazo	USP

Prof. DSc. Demóstenes Zegarra Rodríguez
Orientador

**LAVRAS – MG
2023**

Sou grato a Deus por todas as bênçãos que Ele tem me concedido, e também à minha querida esposa Rosália e minha preciosa filha Giovanna, meus pais, irmãs e sobrinhos pelo apoio e compreensão.

AGRADECIMENTOS

Inicialmente, expresso minha gratidão a Deus e à minha família pelo suporte e oportunidades em momentos cruciais desta jornada. A realização deste trabalho foi possível graças ao apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à Universidade Federal de Lavras, que me proporcionou as condições necessárias. Aos membros da banca, os professores Dante e Renata Rosa, que gentilmente se dispuseram a participar deste trabalho e contribuíram consideravelmente com seu conhecimento. É com grande satisfação que reconheço o trabalho do professor Demóstenes, meu orientador, a quem admiro muito, e também dos demais professores Luiz Henrique, Mayron, Tales, André Pimenta e Renata Rosa, pelo apoio e contribuição de seu conhecimento. Não poderia deixar de agradecer aos meus amigos, Diego e Marcelo, pela ajuda e apoio fundamentais na elaboração deste trabalho. Enfim, a todos que participaram desta caminhada, direta ou indiretamente, meu sincero agradecimento. Todo o esforço foi recompensado.

*"O mundo é um livro, e aqueles que não viajam lêem apenas uma página."
(Santo Agostinho.)*

RESUMO

A avaliação da qualidade da voz em sistemas de telecomunicação é de extrema importância para garantir que as comunicações sejam eficazes e satisfatórias para os usuários. Nesta pesquisa, foram utilizados os algoritmos da Recomendação ITU-T P.862 e do E-Model ITU-T G.107 como referência padrão para analisar a qualidade da voz em sistemas de comunicação. A metodologia usada visou garantir a precisão e relevância dos resultados. Foram realizados testes objetivos em cenários de comunicação móvel usando os codecs EVS (*Enhanced Voice Services*) e AMR-WB (*Adaptive Multi Rate Wide Band*). O propósito deste estudo consistiu em analisar e propor alternativas que avaliem percepção do usuário e controlem a degradação por meio de métodos imparciais, para a análise perceptual de sinais de áudio com qualidade. Focando na Qualidade de Experiência (QoE), das comunicações de voz em sistemas de comunicação móvel, os resultados obtidos foram de grande importância. Durante os testes, foram realizadas codificações e decodificações de sinais de voz utilizando o codec EVS, tanto no modo interoperável com AMR-WB quanto no modo nativo do EVS, usando os parâmetros DTX (*Discontinuous Transmission*) e -q. Os resultados mostram que o desempenho da transmissão de voz foi beneficiado por taxas de bits mais elevadas e pela ativação de parâmetros específicos das condições de transmissão. A utilização dos parâmetros, como o DTX e o -q, mostrou-se eficaz para melhorar o desempenho da transmissão de voz pelo codec EVS. O DTX é um recurso que economiza largura de banda interrompendo temporariamente a transmissão de voz quando não há fala detectada. O parâmetro -q, refere-se à utilização de um algoritmo de cancelamento de ruído ou supressão de silêncio durante a medição da qualidade de voz. Já o E-Model avalia a qualidade da voz, considerando fatores como largura de banda disponível, taxa de bits, perda de pacotes e latência da rede, também desempenha um papel fundamental na determinação da qualidade percebida pelos usuários. Concluindo, a pesquisa demonstrou a importância de avaliar a qualidade da voz em sistemas de telecomunicação e propôs soluções para aprimorar essa qualidade em comunicações em tempo real usando os codecs EVS e AMR-WB. Os resultados obtidos são valiosos para o aprimoramento da qualidade das comunicações de voz em sistemas de comunicação móvel, garantindo uma melhor experiência para os usuários.

Palavras-chave: VoIP. VoLTE. MOS. 3GPP. E-model. Métricas de Avaliação da Qualidade da Voz. Codificadores de Voz. Qualidade de Experiência (QoE). EVS. Redes Móveis.

ABSTRACT

The evaluation of voice quality in telecommunications systems is of utmost importance to ensure effective and satisfactory communication for users. In this research, the algorithms of ITU-T Recommendation P.862 and ITU-T G.107 E-Model were used as standard references to analyze voice quality in communication systems. The methodology used aimed to ensure the accuracy and relevance of the results. Objective tests were conducted in mobile communication scenarios using the EVS (Enhanced Voice Services) and AMR-WB (Adaptive Multi-Rate Wideband) codecs. The objective of this study was to examine and suggest alternatives that assess user perception and control degradation through unbiased methods for perceptual analysis of audio signals' quality. Prioritizing Quality of Experience (QoE), the results obtained were of great importance in improving the quality of voice communications in mobile communication systems. During the tests, voice signal encodings and decodings were performed using the EVS codec, both in interoperable mode with AMR-WB and in native EVS mode, using parameters such as DTX (Discontinuous Transmission) and -q. The results showed that voice transmission performance benefited from higher bit rates and the activation of specific transmission condition parameters. The use of parameters such as DTX and -q proved effective in improving voice transmission performance using the EVS codec. DTX is a feature that saves bandwidth by temporarily interrupting voice transmission when no speech is detected. The -q parameter refers to the use of a noise cancellation or silence suppression algorithm during voice quality measurement. The E-Model evaluates voice quality, considering factors such as available bandwidth, bit rate, packet loss, and network latency, also playing a fundamental role in determining the perceived quality by users. In conclusion, the research demonstrated the importance of evaluating voice quality in telecommunications systems and proposed solutions to enhance this quality in real-time communications using EVS and AMR-WB codecs. The results obtained are valuable for improving the quality of voice communications in mobile communication systems, ensuring a better experience for users.

Keywords: VoIP. VoLTE. MOS. 3GPP. E-model. Voice Quality Assessment Metrics. Voice Coders. Quality of Experience (QoE). EVS. Mobile networks.

LISTA DE FIGURAS

Figura 2.1 – Métodos de avaliação da qualidade da voz	22
Figura 2.2 – VoIP entre computador e telefone comum	24
Figura 2.3 – Serviço VoIP	25
Figura 2.4 – Tecnologia 5G e Internet das Coisas	27
Figura 2.5 – Configuração básica de referência do <i>E-model</i>	30
Figura 2.6 – Algoritmo PESQ	31
Figura 2.7 – Estimadores da Qualidade da Voz	32
Figura 3.1 – Transcodificação	42
Figura 3.2 – Comparação EVS e AMR	43
Figura 5.1 – Fluxograma de Execução Teste do EVS	49
Figura 5.2 – Fluxograma de Auto Nível de Codificação e Decodificação do EVS	51
Figura 5.3 – Funções de áudio - Transmissor	53
Figura 5.4 – Funções do áudio - Receptor	54
Figura 5.5 – Interface Gráfica do Codificador	58
Figura 5.6 – Interface gráfica do decodificador	58
Figura 5.7 – Diagrama de blocos do Multi EVS-TDOA	60
Figura 5.8 – Codificação do T00.INP no modo AMR e EVS no modo interoperável.	61
Figura 5.9 – Amostra decodificação T00.INP nas taxas de bits do modo AMR-WB e EVS inter- operável.	62
Figura 5.10 – Menu da Ferramenta PESQ	62
Figura 5.11 – Avaliação objetiva do sinal original T00.INP com o sinal degradado T00_6600_16.OUT usando PESQ.	63
Figura 5.12 – E-model Calculator	65
Figura 5.13 – Correlação de Pearsen	69

LISTA DE GRÁFICOS

Gráfico 6.1 – ppl %0	81
Gráfico 6.2 – ppl %1	82
Gráfico 6.3 – ppl%2	83
Gráfico 6.4 – ppl%3	84
Gráfico 6.5 – ppl%5	85
Gráfico 6.6 – ppl%10	86
Gráfico 6.7 – ppl%20	87
Gráfico 6.8 – ppl em vários valores de perda	88
Gráfico 6.9 – ppl em vários valores de perda (SWB)	91
Gráfico 6.10 – ppl em vários valores de perda (AMR-WB)	96
Gráfico 6.11 – MOS - P.862	96
Gráfico 6.12 – MOS - Modelo E	97

LISTA DE TABELAS

Tabela 2.1 – Escala de pontuação MOS	26
Tabela 6.1 – Valores MOSLLQ do arquivo T04.INP modo interoperável com taxa do AMR-WB usando a frequência +16000	71
Tabela 6.2 – Valores MOSLLQ do arquivo T04.INP modo interoperável com AMR-WB usando a taxa de bits na frequência de +16000 mais o parâmetro (dtx) ativado	71
Tabela 6.3 – Valores MOS do arquivo T04.INP modo interoperável com AMR-WB usando a frequência de +16000 mais o parâmetro (-q) na codificação e decodificação	72
Tabela 6.4 – Valores MOS do modo interoperável com taxa do AMR-WB - T19.INP baseado na frequência +16000	73
Tabela 6.5 – Valores do MOS no modo interoperável com taxa de AMR-WB e com (dtx) ativado - T19.INP	73
Tabela 6.6 – Valores do MOS no modo interoperável com taxa do AMR-WB e com (-q) na codificação e decodificação - T19.INP utilizando a frequência +16000	74
Tabela 6.7 – Valores MOS modo nativo do codec EVS através do arquivo - T04.INP	75
Tabela 6.8 – Valores MOS do modo nativo do codec EVS com (dtx) ativado através do arquivo - T04.INP e sinais degradados	75
Tabela 6.9 – Valores MOS do modo nativo com (-q) através do arquivo - T04.INP e sinais degradados	76
Tabela 6.10 – Valores MOS do modo nativo e sinais degradados	76
Tabela 6.11 – Valores MOS do modo dtx nativo através do arquivo T19.INP do codec EVS e sinais degradados	77
Tabela 6.12 – Valores MOS do modo nativo do codec EVS através do arquivo T19.INP COM O -q ativo no EVS	78
Tabela 6.13 – Cálculo do valor de R e MOS Banda Larga (WB) variando o ppl em 0%	81
Tabela 6.14 – Cálculo do valor de R Banda Larga (WB) variando em ppl 1.0 %	82
Tabela 6.15 – Cálculo do valor de R Banda Larga (WB) variando em ppl 2%	83
Tabela 6.16 – Cálculo do valor de R Banda Larga (WB) variando ppl 3%	84
Tabela 6.17 – Cálculo do valor de R Banda Larga (WB) variando ppl 5%	85
Tabela 6.18 – Cálculo do valor de R Banda Larga (WB) variando ppl 10%	86
Tabela 6.19 – Cálculo do valor de R Banda Larga (WB) variando ppl 20%	87
Tabela 6.20 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 0.0%	88
Tabela 6.21 – Cálculo do valor de R Super Banda Larga (SWB) variando 1.0%	89
Tabela 6.22 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 2%	89

Tabela 6.23 – Cálculo do valor de R Super Banda Larga (SWB) variando bpl 3%	89
Tabela 6.24 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 5%	90
Tabela 6.25 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 7%	90
Tabela 6.26 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 10.0%	90
Tabela 6.27 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 20.0%	91
Tabela 6.28 – Cálculo do valor de R Banda Larga (WB) variando ppl a 0.0%	92
Tabela 6.29 – Cálculo do valor de R Banda Larga (WB) variando ppl a 1.0%	92
Tabela 6.30 – Cálculo do valor de R Banda Larga (WB) variando ppl a 2.0%	93
Tabela 6.31 – Cálculo do valor de R Banda Larga (WB) variando ppl a 3.0%	93
Tabela 6.32 – Cálculo do valor de R Banda Larga (WB) variando ppl a 5.0%	94
Tabela 6.33 – Cálculo do valor de R Banda Larga (WB) variando ppl a 10.0%	94
Tabela 6.34 – Cálculo do valor de R Banda Larga (WB) variando ppl a 20.0%	95

LISTA DE QUADROS

Quadro 3.1 – Características Do Codec AMR-NB	37
Quadro 3.2 – Características Do Codec AMR-WB	38
Quadro 3.3 – Especificações Técnicas Do Codec EVS	41
Quadro 3.4 – Características Do Codec EVS	46
Quadro 5.1 – Descrição e Valores dos Elementos da Declaração de Codificação.	52
Quadro 5.2 – Descrição e Valores dos Elementos da Declaração de Decodificação.	53
Quadro 5.3 – Quadros de Fala	55
Quadro 5.4 – Fluxo de Bits	55
Quadro 5.5 – Taxas de bits do codec AMR-WB	57
Quadro 5.6 – Recomendação ITU-T G.107 - Valores padrão e intervalos permitidos para os parâmetros para uso com o cálculo do fator R do modelo E	66
Quadro 5.7 – Recomendação ITU-T G.107 - Valores padrão e intervalos permitidos para os parâmetros para uso com o cálculo do fator R do modelo E	67
Quadro 6.1 – Valores pré-estabelecidos e faixa de valores permitidos para o Fator R	79
Quadro 6.2 – Valores padrão e intervalos permitidos para os parâmetros	80

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Contextualização e Motivação	16
1.2	Justificativa	18
1.3	Objetivo	18
1.3.1	Objetivos Específicos	18
1.4	Organização do Trabalho	19
2	REFERENCIAL TEÓRICO	20
2.1	Avaliação da Qualidade da Fala	20
2.2	Avaliação Subjetiva de Qualidade da Voz	20
2.3	Avaliação Objetiva da qualidade da fala	20
2.4	Métodos para Avaliação da Qualidade de Voz	21
2.5	Métodos Subjetivos	22
2.6	Método Objetivo	23
2.7	Conceitos e Definição de QoE	23
2.8	Conceitos e Definição de QoS	23
2.9	Comunicação Básica do VoIP	24
2.10	O Serviço de VoIP	24
2.11	E-model Banda Larga	28
2.12	Avaliação da Qualidade da Voz por Percepção (PESQ)	30
3	CODIFICADORES DE VOZ	34
3.1	G.711	34
3.2	G.726	34
3.3	G.729	35
3.4	Codificador Adaptativo Multi-Rate (AMR)	35
3.5	Codificador Enhanced Voice Service (EVS)	38
3.5.1	Especificações Técnicas do Codec do EVS	41
3.5.2	Recursos do codec EVS	43
3.5.3	Característica do EVS	44
4	TRABALHOS RELACIONADOS	47
5	METODOLOGIA	49
5.1	Descrição das Ferramentas Usadas nos Experimentos	50
5.1.1	Execução do Codec EVS	50
5.2	Funções de Áudio	53

5.3	Propriedade do Arquivo	54
5.3.1	Visão Geral do Formato de Áudio AMR	57
5.3.2	Características do Arquivo	58
5.3.3	Formato de Arquivo Teste	59
5.4	Cenário - 1	59
5.5	Cenário - 2	64
5.5.1	E-MODEL AVALIAÇÃO	64
5.6	Correlação de Pearson	69
6	RESULTADOS	70
6.1	Valores obtidos através da Rec. ITU-T P.862	70
6.2	TESTES E-MODEL	78
6.2.1	Avaliação da Probabilidade de Perdas de Pacotes	80
6.2.2	Avaliação Codec Adaptive Multi-Rate Wideband (AMR-WB) com E-model	92
6.3	Gráfico Rec. P.862 e E-model	96
7	CONCLUSÃO	98
	REFERÊNCIAS	99

1 INTRODUÇÃO

Na estrutura deste trabalho está incluída uma visão geral dos sistemas de comunicação, seguida de uma descrição mais detalhada da evolução dos serviços de voz.

1.1 Contextualização e Motivação

O estudo da qualidade da voz em sistemas de comunicação móvel é de suma importância, pois a voz é um dos principais meios de comunicação nas chamadas telefônicas entre outras formas de comunicação em tempo real. A utilização de codecs eficientes, como o *Enhanced Voice Services* (EVS), é fundamental para garantir uma experiência de comunicação satisfatória em sistemas de comunicação móvel (3GPP, 2011)

O codec EVS é um formato de compressão de áudio amplamente utilizado para melhorar a qualidade da voz em comunicações móveis e serviços de voz sobre IP. Ele é empregado em aplicativos de comunicação em tempo real e utiliza técnicas avançadas de codificação para fornecer uma experiência de áudio superior. O EVS é padronizado pelo 3GPP, assegurando sua adoção e interoperabilidade em diferentes dispositivos e redes (3GPP, 2011)

No entanto, é de extrema importância conduzir avaliações objetivas da qualidade de voz de sistemas de comunicação móvel que utilizam o codec EVS, devido à falta de estudos relacionados. Neste contexto, avaliar a Qualidade de Experiência (QoE) torna-se um conceito fundamental quando se trata da qualidade da voz em sistemas de comunicação móvel. A QoE se refere à percepção subjetiva do usuário sobre a qualidade geral do serviço oferecido.

Portanto, o objetivo deste estudo foi avaliar a qualidade da voz, através de sistemas de comunicação móvel, que utilizam o codec EVS através de dois métodos padronizados pela ITU-T (*International Telecommunication Union*) para avaliar a qualidade da voz: o P.862 e o E-Model. Ambos os métodos são amplamente utilizados na avaliação de codecs e desempenham um papel fundamental nesse contexto.

A recomendação P.862, também conhecida como Avaliação Perceptual de Qualidade da Fala (PESQ), é um método intrusivo que descreve a predição da qualidade subjetiva de sinais de voz de banda estreita (0,3 - 3,4 kHz). A PESQ avalia um sinal ao compará-lo com um sinal degradado resultante da passagem por um sistema de comunicação. Essa metodologia fornece uma estimativa da qualidade percebida por meio de um teste subjetivo, utilizando uma pontuação de qualidade do áudio ou Mean Opinion Score Listening Quality Objective (MOS-LQO), similar ao MOS (GOUVEIA; ROSA; RODRÍGUEZ, 2018).

O MOS-LQO é uma medida objetiva que representa a qualidade do áudio escutado por um indivíduo durante o teste subjetivo. Essa pontuação é calculada com base na comparação entre o sinal original e o sinal degradado, considerando os efeitos do sistema de comunicação. O MOS-LQO permite

uma quantificação da qualidade percebida de forma padronizada e é amplamente utilizado como uma métrica confiável para avaliar a qualidade de voz (GOUVEIA; ROSA; RODRÍGUEZ, 2018).

O E-Model é um modelo matemático que leva em consideração fatores como atraso, jitter, perda de pacotes e níveis de ruído para prever a qualidade percebida da voz em uma chamada. Ele utiliza uma abordagem paramétrica para estimar a qualidade da voz com base em diferentes parâmetros que afetam a experiência de comunicação (GOUVEIA; ROSA; RODRÍGUEZ, 2018).

Portanto, a recomendação ITU-T Rec. G.107, também conhecida como E-Model, é amplamente utilizada como um modelo paramétrico para avaliar a qualidade de voz, levando em consideração os parâmetros da rede de transporte e as condições acústicas do ambiente. Essa abordagem ajuda a compreender e prever a qualidade percebida pelos usuários em diversas situações de comunicação.

Após a avaliação realizada pelo E-Model e P.862, foi conduzida uma análise de correlação de Pearson para avaliar a relação entre a taxa de bits utilizada pelo codec e o escore MOS gerado. A correlação de Pearson é uma medida estatística que avalia a força e a direção da relação linear entre duas variáveis, sendo neste caso a taxa de bits e o escore MOS. Essa análise permitiu investigar o grau de influência da taxa de bits na qualidade percebida da voz, fornecendo uma compreensão mais detalhada de como as variáveis estão relacionadas.

A demanda crescente por serviços de dados e voz em redes móveis motivou este estudo. A comunicação por voz desempenha um papel fundamental em setores como telecomunicações, call centers, suporte ao cliente e telemedicina. A expectativa dos usuários é de uma qualidade de sinal de voz excelente, garantindo uma comunicação clara e sem interrupções.

A qualidade do sinal de voz é essencial para uma comunicação eficaz. Problemas como ruídos, eco, cortes ou distorções podem comprometer a compreensão entre os interlocutores, levando a dificuldades e possíveis erros de comunicação. Além disso, a qualidade do sinal de voz está diretamente relacionada à satisfação do usuário. Baixa qualidade pode causar frustração e insatisfação, prejudicando a imagem da empresa prestadora do serviço.

Em resumo o método consistiu em realizar testes objetivos com o codec EVS em cenários de comunicação móvel, avaliando a qualidade da voz através do MOS e considerando parâmetros relevantes, como DTX e taxa de bits. O uso do E-model permitiu uma avaliação abrangente dos fatores de qualidade de voz.

Os resultados mostram que o desempenho da transmissão de voz é beneficiado por taxas de bits mais elevadas e pela ativação dos parâmetros das condições de transmissão utilizadas

1.2 Justificativa

O mundo está cada vez mais conectado e a procura por serviços de voz está crescendo cada vez mais, especialmente com a popularidade dos serviços de comunicação pela internet, como Voz sobre IP (VoIP). O VoIP é uma tecnologia inovadora que está mudando a maneira das pessoas se comunicarem. Com a crescente demanda por serviços de voz e a disponibilidade cada vez maior de conexões de internet de alta velocidade, é provável que o uso de VoIP continue crescendo no futuro.

Apesar de ser um serviço de melhor esforço e não oferecer garantias de entrega pontual ou completa dos pacotes de áudio, podem ocorrer problemas de qualidade, como interrupções no áudio, perda de qualidade sonora ou atrasos na transmissão. É importante ressaltar que, apesar da falta de garantias de qualidade na rede IP, essa tecnologia é amplamente utilizada atualmente para a transmissão em tempo real de voz, videoconferências e transmissões ao vivo, devido à sua eficiência e popularidade. No entanto, existem técnicas e soluções disponíveis para minimizar os problemas que afetam a qualidade de voz, buscando garantir a melhor qualidade possível dentro das limitações existentes. (SHIGUEOKA, 2016).

Para garantir uma experiência de chamada satisfatória, é fundamental manter uma avaliação constante da qualidade da transmissão de voz.

Sem dúvida, a análise da eficácia de diferentes codecs, como o codec EVS, em diversos ambientes de rede, desempenha um papel fundamental na compreensão de seu comportamento e desempenho. Esses estudos são de extrema importância para identificar quais codecs são mais adequados para cada tipo específico de ambiente de rede e dispositivos móveis, o que contribui diretamente para aprimorar a qualidade das comunicações de voz em serviços VoIP.

Os estudos para melhorar a qualidade da voz em comunicações em tempo real são fundamentais para garantir uma comunicação clara e eficaz em diferentes ambientes de rede, permitindo que as pessoas se comuniquem sem problemas, independentemente das condições tecnológicas.

1.3 Objetivo

O objetivo principal é avaliar a qualidade da voz em sistemas de comunicação móvel utilizando os algoritmos da Recomendação ITU-T P.862 (MOS) e o E-Model ITU-T G.107. Ambos são amplamente aceitos como padrões para análise da qualidade da voz. Essa avaliação é importante pois está relacionada com a qualidade de experiência.

1.3.1 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- a) codificar codecs com extensões (.INP) utilizando o EVS *Enhanced Voice Services* em ponto flutuante e variar diferentes tipos de parâmetros, como taxa de bits, e perda de pacotes gerando o MOS;
- b) avaliar a qualidade do codec decodificado por meio da execução do algoritmo ITU-T P862;
- c) aplicação do E-model para análise comparativa do codec EVS e AMR-WB;
- d) usar o coeficiente de correlação de Pearson para analisar a relação entre diferentes métricas e dados coletados em testes objetivos;
- e) avaliar o desempenho do sistema em diferentes cenários e condições de rede.

1.4 Organização do Trabalho

Este trabalho está dividido em: Capítulo 2: O referencial teórico, no qual são apresentados diferentes temas relacionados a este trabalho. O capítulo se inicia com um breve estudo sobre a qualidade da fala, codecs e alguns destaques sobre a tecnologia LTE.

Capítulo 3: Neste capítulo, é realizada uma breve análise sobre os codificadores de voz, com ênfase no codec EVS.

Capítulo 4: Os trabalhos que se relacionam com o caso estudado nesta dissertação são apresentados neste capítulo.

Capítulo 5: Neste capítulo, é descrita a metodologia proposta e as diversas ferramentas que auxiliaram no desenvolvimento do trabalho, com informações detalhadas sobre o ambiente de simulação e os parâmetros de configuração utilizados.

Capítulo 6: Os resultados obtidos são descritos neste capítulo.

Capítulo 7: Conclusão

2 REFERENCIAL TEÓRICO

Neste capítulo, apresentamos os conceitos relacionados à avaliação da voz que envolvem a utilização de métricas objetivas, como a Rec. P.862 e o E-Model, além da consideração dos conceitos de QoE e QoS. Essas avaliações são fundamentais para garantir uma experiência de comunicação adequada e identificar possíveis problemas na rede ou nos serviços de voz.

2.1 Avaliação da Qualidade da Fala

A avaliação do sinal de voz pode ser realizada por meio de diversas técnicas e métricas, incluindo medições objetivas e subjetivas. Ambos os métodos possuem suas vantagens e desvantagens. Os métodos objetivos são ágeis, automatizados e podem ser aplicados em grande escala, mas podem não abranger todos os aspectos da percepção humana. Já os métodos subjetivos são mais precisos em refletir a percepção humana, mas são mais lentos, caros e exigem a participação de um grande número de indivíduos. Por isso, é comum utilizar uma combinação de métodos objetivos e subjetivos para obter uma avaliação mais completa da qualidade de voz (ITU-TP.800, 1996).

A análise subjetiva e objetiva são complementares e importantes para avaliar a qualidade da voz de um sistema. A combinação dessas análises pode fornecer uma visão mais completa e precisa da qualidade da voz, ajudando a garantir uma experiência positiva para o usuário final (SOUSA, 2007).

Em específico, para fins de projeto de rede, é muito importante avaliar o impacto dos sistemas MIMO¹ na melhoria da qualidade do sistema de transmissão (RODRÍGUEZ et al., 2019).

2.2 Avaliação Subjetiva de Qualidade da Voz

Vários aspectos importantes devem ser analisados em um teste subjetivo de qualidade de voz. O grau de importância que deve ser dado a cada um dependerá do tipo de aplicação desejada. Entre os fatores mais importantes, destacam-se a variação do nível de entrada do sinal a ser codificado, a ocorrência de erros no canal de operação do *codec*, a ocorrência de múltiplas transcodificações e diversidade de locutores, a presença de ruído em ambiente de níveis variados e o atraso introduzido pela codificação (ANDRADA L. & MARINHO, 2007).

2.3 Avaliação Objetiva da qualidade da fala

Existem dois métodos de avaliação objetiva: não intrusivos e intrusivos. Os métodos intrusivos requerem um sinal de referência para comparar com o sinal final e garantir a qualidade da pontuação.

¹ MIMO é a sigla em inglês para *Multiple Input Multiple Output* que, em uma tradução literal, significa “Múltiplas Entradas Múltiplas Saídas”. Trata-se de um sistema que visa alcançar maiores taxas de transmissão em redes sem fios. A tecnologia usa várias antenas para transmitir o sinal e os dados em uma rede.

Por essa razão, os métodos intrusivos são considerados mais confiáveis e servem como referência na avaliação objetiva. Por outro lado, os métodos não intrusivos são os que precisam apenas do sinal no receptor, ou em um dado ponto onde o sinal deve ser avaliado e assim são mais rápidos, o que possibilita seu uso em serviços de tempo real. (EMANUEL et al., 2016).

Os métodos e propriedades de avaliação da voz, sendo eles subjetivos ou objetivos, são bastante estudados, debatidos e submetidos a rigorosos testes de validação. A Recomendação ITU-T P.862 2001, mais conhecida como Avaliação Perceptual da Qualidade da Fala (PESQ, *Perceptual Evaluation of Speech Quality*), é um método objetivo intrusivo que estima uma Pontuação de Opinião Média (MOS, *Mean Opinion Score*), para a avaliação da qualidade de voz fim a fim nas redes telefônicas de banda estreita (NB, Narrow Band) (GOUVEA et al., 2018).

2.4 Métodos para Avaliação da Qualidade de Voz

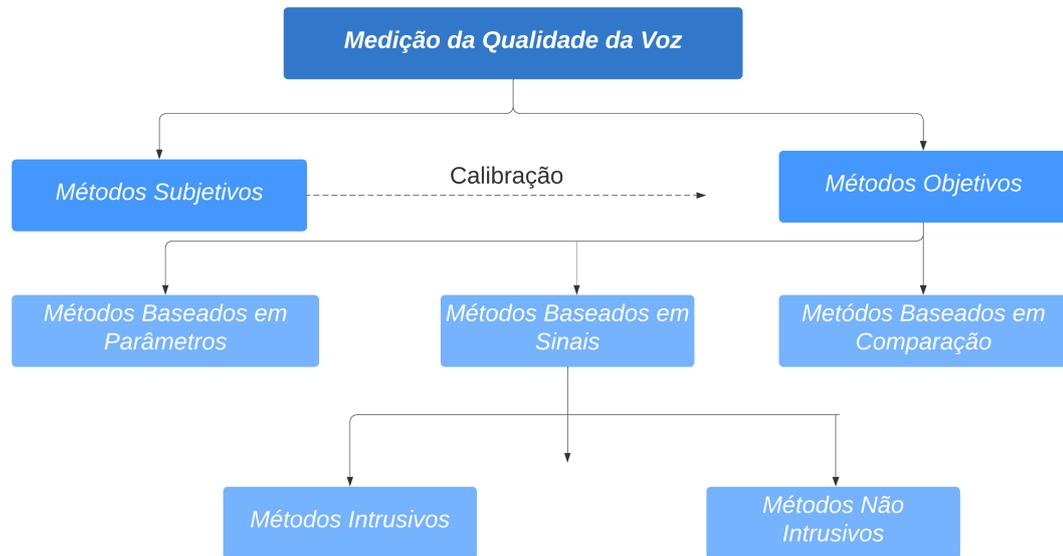
Para avaliar a qualidade da voz, existem dois métodos: os subjetivos e objetivos, que adotam abordagens diferentes em vários campos para obter, analisar e interpretar informações.

Os métodos subjetivos incluem opiniões, sentimentos e interpretações pessoais, e dependem da perspectiva e da experiência do observador. Por exemplo, em um estudo de satisfação no trabalho, um método subjetivo pode consistir em pedir aos funcionários que classifiquem seu nível de satisfação em uma escala e escrevam uma reflexão pessoal sobre suas experiências de trabalho (ITU-TP.800, 1996).

Já os métodos objetivos são aqueles que se baseiam em dados verificáveis, quantificáveis e observáveis. Eles são amplamente utilizados em várias áreas do conhecimento, como ciências naturais, matemática, estatística, economia, entre outras. Por exemplo, em um estudo de qualidade no trabalho, um método objetivo pode incluir a coleta e análise de dados sobre taxas de rotatividade de funcionários, absenteísmo e métricas de desempenho. (EMANUEL et al., 2016).

Os métodos subjetivos e objetivos apresentam vantagens e desvantagens distintas, e a escolha do método a ser empregado geralmente é influenciada pela questão de pesquisa e pelos objetivos específicos do estudo. Em muitos casos, é viável utilizar uma combinação de ambos os métodos.

Figura 2.1 – Métodos de avaliação da qualidade da voz



Fonte: Do Autor (2023)

2.5 Métodos Subjetivos

Os métodos subjetivos são essenciais para determinar o nível dos métodos objetivos. A Recomendação P.800 descreve vários métodos e procedimentos de avaliação da qualidade da transmissão (TYMCHENKO et al., 2020).

Os métodos subjetivos são fundamentados na opinião do usuário para medir a qualidade dos serviços. A avaliação de áudio pode ser realizada diretamente via índice de categoria absoluta² ou comparando o sinal de áudio no receptor com o sinal do áudio de referência. Ambas as abordagens têm suas vantagens e limitações. A avaliação absoluta é útil quando se deseja obter uma visão geral da qualidade de um único áudio, sem a necessidade de comparação direta. Já a avaliação comparativa é útil quando se deseja avaliar a qualidade relativa entre diferentes áudios ou comparar a qualidade de diferentes serviços. Com avaliações do tipo ACR, o áudio é classificado com valores entre 1 e 5, sendo 5 "Excelente" e 1 "Ruim". O MOS representa a média de ACR medido entre um grande número de usuários (RODRÍGUEZ, 2018).

Os métodos subjetivos são aqueles, em que se utilizam seres humanos, para avaliar um determinado conjunto de sinais, enquanto os métodos objetivos utilizam algoritmos implementados em máquinas. É de grande importância saber qual a qualidade percebida de sinais de fala que foram originados ou modificados por algum sistema eletrônico. Geralmente um teste subjetivo utiliza um grupo de pessoas, para avaliar sinais de fala e, no final do processo, uma nota média é associada para cada sinal (PREGO, 2012).

² Índice de categoria absoluta (ACR - *Absolute Category Rating*) é um método utilizado em avaliações subjetivas para medir a qualidade de áudio, vídeo ou outros sinais de comunicação.

No método subjetivo de análise, um grupo de pessoas utiliza critérios pré-estabelecidos para mensurar a qualidade de algo com base em suas expectativas e experiências pessoais (SOUSA, 2011).

2.6 Método Objetivo

No método objetivo, a métrica pode ser calculada, a começar de modelos obtidos da própria rede de comunicação, um exemplo, nível de recepção, taxa de erro de transmissão e qualidade de recepção, entre outros. Outra forma de obter uma avaliação objetiva consiste no sinal de voz da mensagem original com o sinal da mensagem recebida, após ter sofrido degradações inseridas na chamada, causadas pelo processo de codificação e decodificação e pela sua transmissão na rede (SOUSA, 2011).

2.7 Conceitos e Definição de QoE

A noção de Qualidade da Experiência (QoE), tem sido um dos temas de pesquisa na comunidade científica relacionada às telecomunicações. Há diferentes abordagens sugestivas para os termos de QoS e QoE. Embora esses dois termos tenham significados diferentes, podem estar relacionados de alguma forma. É evidente que a definição de QoS está focada, nos aspectos técnicos dos serviços de telecomunicações, enquanto a definição de QoE está focada no usuário (FAGHIHI; BEHDADFAR, 2020).

A Qualidade de Experiência é um conceito multidimensional que engloba aspectos financeiros, sociais, comportamentais e de preferências do cliente. É um diferencial competitivo para as empresas e está se tornando cada vez mais relevante em diversas áreas, impulsionando a realização de pesquisas e estudos para aprimorar a compreensão e a entrega de experiências excepcionais aos clientes. Um bom exemplo é o dos provedores de serviços de telecomunicações, em que a fidelização dos clientes depende do nível de QoE do usuário. Portanto é de extrema importância que os provedores de serviços conheçam esses usuários para oferecer um serviço cada vez melhor aos clientes (RODRIGUEZ et al., 2016).

A Qualidade da Experiência (QoE) é um campo multidisciplinar fundamentado em psicologia social, ciência cognitiva e ciência da engenharia, que se concentra em entender como as pessoas percebem e avaliam a qualidade de um produto ou serviço, geralmente relacionado a tecnologias de comunicação e informação, como videoconferência, streaming de vídeo, jogos online, etc (LAGHARI et al., 2012).

2.8 Conceitos e Definição de QoS

Qualidade de Serviço (QoS), em português Qualidade de Serviço, é a capacidade da rede de oferecer um nível de serviço satisfatório aos seus usuários/clientes, com garantia que algumas aplicações

ou serviços tenham prioridade sobre outros e que a rede possa suportar as demandas de tráfego sem comprometer a qualidade do serviço prestado (SANTOS, 2019).

Em chamadas VoIP é possível melhorar a QoS para controlar os parâmetros de atraso, jitter e perda de pacotes dentro dos limites recomendados pelo ITU-T. Isso proporciona uma melhor experiência para os usuários de VoIP e garante uma comunicação mais confiável e eficiente (SANTOS, 2019).

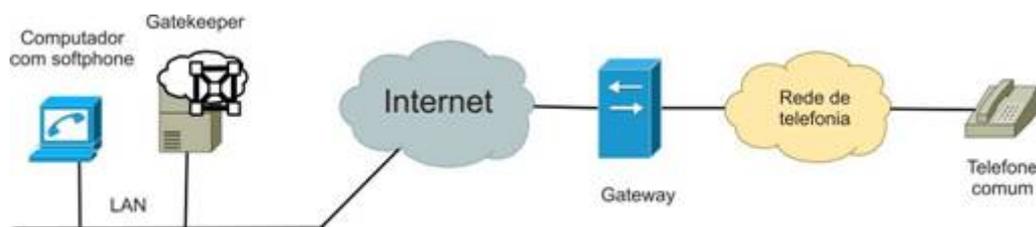
Qualidade de Serviço (QoS), é uma medida de quão bem uma rede está e é uma tentativa de definir as características e a natureza de um serviço. A QoS é especialmente importante para serviços críticos, como telemedicina, VoIP e videoconferência, onde a qualidade da rede pode ter um impacto direto na qualidade do serviço prestado. (TARUK et al., 2018).

2.9 Comunicação Básica do VoIP

A ideia básica é estabelecer uma comunicação entre origem e destino, por meio de uma rede IP e fazer a troca de pacotes de dados, em tempo real, com a informação de áudio (voz), de forma bidirecional. A arquitetura de rede com VoIP pode empregar diversos protocolos e tecnologias, sendo uma delas a recomendação H.323 (T. ZOURZOUVILLYS. & RESCORLA, 2010).

Na Figura 2.2, é exibida uma arquitetura típica de uma rede comum entre computador e telefone. Observe que nessa arquitetura de rede com VoIP, são utilizados equipamentos típicos da recomendação H.323, mas uma rede VoIP também pode ser implementada com o SIP (TELECO, 2019).

Figura 2.2 – VoIP entre computador e telefone comum



Fonte: Adaptada de: Teleco (2019)

2.10 O Serviço de VoIP

As soluções VoIP corporativas estão se tornando cada vez mais populares nas pequenas e médias empresas, oferecendo vantagens competitivas importantes para essas empresas em relação às empresas que ainda utilizam sistemas telefônicos convencionais, principalmente pela disponibilidade de serviços VoIP, baseados em nuvem (SINIARSKI et al., 2017).

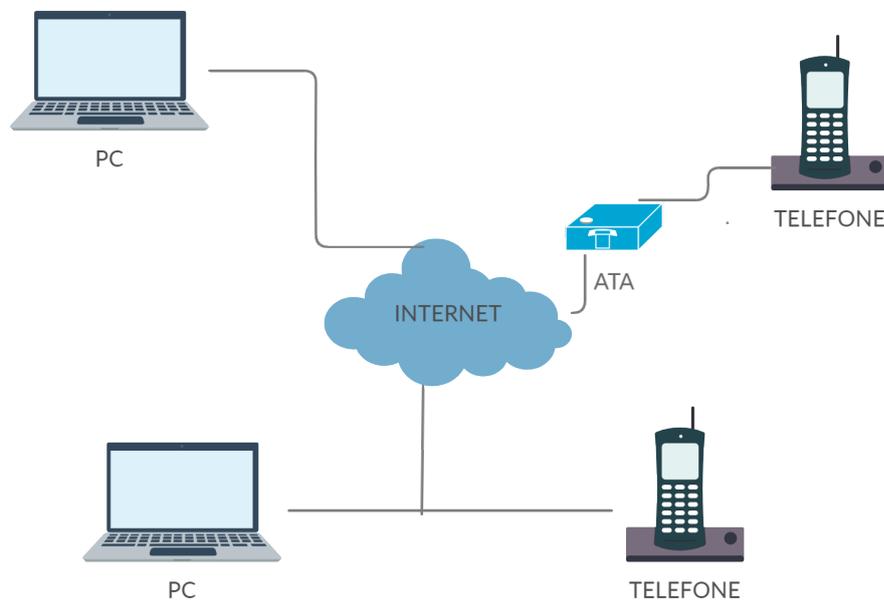
VoIP é uma tecnologia sensível à latência, perda de pacotes e jitter, que podem afetar a qualidade das chamadas de voz. Por isso, a equipe de TI das empresas precisam ter conhecimentos sobre as melhores práticas de configuração de rede para VoIP, incluindo o uso de VLANs, QoS (*Quality of Service*)

e outras tecnologias que ajudam a garantir a qualidade das chamadas de voz. Alguns departamentos estão evoluindo, para o paradigma DevOps ³, no qual o gerenciamento de serviços é integrado às funções operacionais do negócio, portanto é interessante explorar o uso de ferramentas de monitoramento de Qualidade de Experiência (QoE) VoIP para esses departamentos (SINIARSKI et al., 2017).

O VoIP é um serviço de chamadas que se refere a um conjunto de tecnologias, permitindo que as pessoas se comuniquem por voz pela Internet. O VoIP deverá substituir muito em breve grande parte da rede telefônica pública comutada. Entre os serviços VoIP mais populares estão incluídos o Skype, Google Hangouts Apple FaceTime (HARTE, et al., 2015).

Na Figura 2.3, é mostrado um exemplo convencional de Serviço VoIP.

Figura 2.3 – Serviço VoIP



Fonte: Do Autor (2023)

A tecnologia VoIP (*Voice over Internet Protocol*), lida com a conversão de sinais de áudio analógicos, como por telefone em dados digitais que podem ser transmitidos pela Internet, usando-se o protocolo da Internet (SONASKAR, S & GIRIPUNJE, S, 2011).

Os provedores de redes e serviços veem a tecnologia VoIP como uma solução capaz de reduzir seus custos e ofertar serviços existentes baseados em voz e novos serviços de multimídia. Os provedores de serviços também veem a infraestrutura VoIP como uma fonte de economia capaz de erguer novos serviços promotores de receita. À proporção que a tecnologia VoIP é implantada, ela se populariza e participa de um cenário competitivo e compartilhado; esse segundo objetivo torna-se mais importante, dessa

³ A palavra "DevOps" é a combinação dos termos "desenvolvimento" e "operações". No entanto, ela representa um conjunto de ideias e práticas que ultrapassam o significado desses dois termos.

forma, os provedores de serviços trabalham para aumentar suas bases de mercado (AHUJA; ENSOR, 2004).

É considerado um sistema com perdas, quando a saída decodificada não é exatamente igual à entrada original. O sistema é considerado sem perdas quando a saída e entrada forem iguais. Para cada codec, pode ser associada uma medida de qualidade, que é uma resposta subjetiva de um ouvinte. Uma medida subjetiva, comumente usada para determinar a qualidade do som produzido pelos *codecs*, é o *Mean Opinion Score*, que pode variar, em uma escala de 1(um) a 5, em que 1(um) indica voz inaceitável e de baixa qualidade, enquanto um valor - 5 indica alta qualidade de voz, sem problemas perceptíveis (CAVALCANTE et al., 2017).

Por meio do indicador *Mean Opinion Score (MOS)*, usado para o estabelecimento da qualidade dos serviços de telecomunicações, é possível avaliar qualidade da fala no Serviço de Atendimento *Over The Top (OTT)*. O serviço de chamadas integra outros serviços como multimídia, chat e vídeo. Com a demanda grande por parte do usuário para usar o serviço, faz-se necessário usar um indicador para avaliar a qualidade da fala, nesse caso o MOS. Na Tabela 4.1, de acordo com a recomendação ITU-T, há vários modos de indicar o valor MOS, seja objetiva ou subjetivamente e, um deles, é por meio da recomendação Rec. ITU-T P.862 (M. R. HIKMATULLAH & S. HARYADI, 2017).

Tabela 2.1 – Escala de pontuação MOS

Nível (MOS)	COMPREENSÃO
5	Excelente
4	Boa
3	Regular
2	Ruim
1	Péssimo

Fonte: Adaptado de ITU-TP.800 (1996)

Este trabalho contribui para a avaliação perceptual da qualidade da fala, fundamentada na recomendação a ITU-T Rec. ITU-T P.862 o valor do MOS⁴ pode ser medido precisamente, empregando-se o procedimento de avaliação da qualidade da conversação de ponta a ponta das redes telefônicas de banda estreita e codecs de fala. Esse procedimento é o desenvolvimento de uma avaliação subjetiva de valor MOS baseada na recomendação P.800 da ITU-T. O procedimento⁵ PESQ parte de uma avaliação de valor MOS subjetiva e, em seguida, os resultados são combinados com os resultados da avaliação objetiva (M. R. HIKMATULLAH & S. HARYADI, 2017).

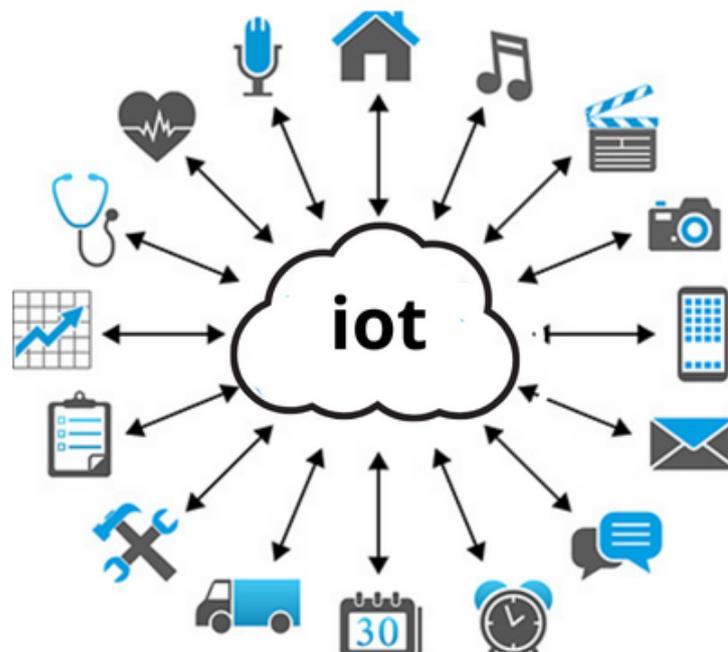
⁴ O Mean Opinion Score (MOS) é a mais conhecida medida da qualidade de voz. É um método subjectivo de teste de qualidade.

⁵ O algoritmo PESQ foi projetado para prever pontuações subjetivas de opinião de uma amostra de áudio degradada.

A rede pública de telefone (PSTN) com fio tradicionalmente suporta apenas o codec G.711. Ele usa uma taxa de amostragem de 8 kHz e uma profundidade de bits de 8 bits para codificar o sinal de voz analógico em um sinal digital. Embora esse codec forneça uma qualidade de áudio razoável para chamadas telefônicas, ele tem limitações em termos de largura de banda e capacidade de transmissão de áudio de alta qualidade (T. ZOURZOUVILLYS. & RESCORLA, 2010).

IoT ou internet das coisas e as redes 5G são duas tecnologias complementares que abrem novas possibilidades de conectividade e interação entre dispositivos. A combinação dessas duas tecnologias pode levar a inovações significativas em diversos setores, melhorando a eficiência, a segurança e a qualidade de vida. A conectividade IoT se baseia na comunicação entre dispositivos que podem ser sensores, atuadores, câmeras, entre outros, que coletam e transmitem dados através da Internet. A tecnologia VoIP *Voice over Internet Protocol* permite a transmissão de voz em tempo real através da Internet. Com a integração do VoIP nos gateways IoT, os usuários podem interagir com esses dispositivos por meio da voz, o que abre novas possibilidades de interação entre humanos e dispositivos conectados à Internet. Essa integração tem o potencial de tornar a vida das pessoas mais fácil e mais conveniente, permitindo o controle remoto de dispositivos e a comunicação de voz com eles (COMER, 2016).

Figura 2.4 – Tecnologia 5G e Internet das Coisas



Fonte: Adaptada de Lins (2015)

A comunicação de voz do VoIP (Voice over Internet Protocol) é realizada por meio da rede IP, que busca oferecer uma qualidade de voz semelhante às redes de telefonia baseadas na comutação de circuitos. No entanto, devido às características inerentes da tecnologia IP e das redes de dados, manter a mesma qualidade de serviço (QoS) e experiência do usuário (QoE) das redes de comutação de

circuitos pode ser desafiador. Diante dessas dificuldades, as aplicações VoIP de alguma forma procuram compensar os usuários com novos serviços, tentando diminuir os problemas e induzindo a redução dos custos operacionais e de infraestrutura (FONSECA, 2014).

A tecnologia de Voz sobre IP (VoIP) tem se tornado cada vez mais popular devido aos avanços na qualidade da internet. A qualidade de transmissão de voz através da internet melhorou significativamente nos últimos anos, permitindo que as pessoas usem VoIP em diferentes tipos de conexões de internet, incluindo cabo, fibra óptica, internet móvel 3G, 4G, 5G e 6G (UHL, 2018,).

Apesar dos desafios, é provável que o 5G se torne cada vez mais comum nos próximos anos, à medida que mais dispositivos e redes são atualizados para suportar essa tecnologia. Isso pode oferecer muitas oportunidades para empresas e consumidores que desejam aproveitar ao máximo as vantagens do 5G (ISP BLOG, 2020).

Talvez não seja possível garantir a qualidade do serviço VoIP devido a fatores externos, existem algumas medidas que podem ser tomadas para minimizar problemas e melhorar a experiência do usuário. Fazer testes de qualidade de chamada: existem ferramentas disponíveis na internet para testar a qualidade de chamada do seu serviço VoIP. Fazer esses testes irá ajudar a identificar problemas e tomar medidas para resolvê-los (UHL, 2018,).

É muito importante tomar várias precauções antes de uma implementação VoIP. Primeiro, precisa-se analisar e avaliar a estrutura da rede, conhecer as necessidades do usuário, ter um bom plano estruturado e saber escolher o codec certo. A implementação VoIP pode assumir a forma de um "hardphone", um dispositivo de comunicação física dedicado, como um telefone IP, ou um "softphone", um programa baseado em software, como skype. Hoje em dia, é mais comum utilizar softphones pela facilidade de uso e comodidade para o usuário (A. ALHAYAJNEH et al. , 2018).

2.11 E-model Banda Larga

O *E-Model* ou modelo E em português, é uma técnica de avaliação objetiva de qualidade da voz conhecida pela ITU-T G.107, é um modelo matemático bastante usado para estimar a qualidade de uma chamada telefônica em uma rede de comutação de circuitos. O *E-Model* leva em conta vários fatores que podem afetar a qualidade de uma chamada, como o tipo de codec usado, a taxa de bits e a quantidade de atraso e perda na rede. O modelo E fornece um único número que representa o score médio de opinião esperado (MOS) para uma chamada, que é uma classificação subjetiva da qualidade da fala em uma escala de 1 a 5. Porém, ao avaliar a qualidade de voz em uma rede de telecomunicações é importante considerar outros fatores, como a experiência do usuário e as expectativas do serviço (REC. ITU-T, 2013).

O E-Model é uma ferramenta útil para engenheiros de rede em uma ampla variedade de aplicações, permitindo que eles tomem decisões mais informadas e eficazes em relação ao projeto, implantação e gerenciamento de redes de comunicação (REC. ITU-T, 2013).

"O E-Model prevê a qualidade de uma conversa e não a qualidade de escuta"(Michael,et. al., 2020)

O E-Model se fundamenta na ideia de que os aspectos psicológicos em uma escala subjetiva são aditivos. Dessa forma, cada fator que degrada a qualidade da voz pode ser computado separadamente, mesmo que isso não signifique que esses fatores não estejam correlacionados (VIANNA, 2007).

Com o Modelo-E é possível fazer uma análise subjetiva da qualidade de voz, como se o próprio usuário estivesse avaliando. O E-Model é um modelo computacional que avalia as variações dos diversos fatores que combinados geram atrasos e degradação da voz, afetando a qualidade (BORGES, 2012).

O E-Model é amplamente utilizado devido à sua capacidade de prever a qualidade de transmissão percebida pelo usuário final com precisão razoável. O modelo E é baseado na modelagem dos resultados de um grande número de testes subjetivos feitos no passado em uma ampla gama de parâmetros de transmissão. O modelo E é baseado, em um algoritmo matemático, que transforma esses parâmetros de transmissão em "fatores de comprometimento" individuais que representam a degradação percebida causada por cada parâmetro. Esses fatores de deficiência são considerados aditivos em uma escala psicológica, o que significa que são combinados para calcular um valor geral de deficiência (REC. ITU-T, 2008).

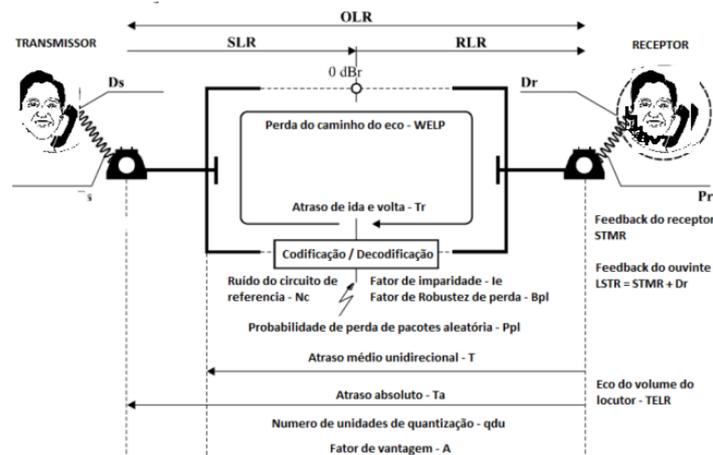
A relação entre as diferentes condições de deterioração R é dada pela equação: $R = R_o - I_s - I_d - I_{e-eff} - I_{recência} + A$

R_o – Valor da relação sinal-ruído (SNR) que é determinado pelos níveis de ruído, sonoridade, etc., incluindo ruído do circuito, ruído da sala e ruído da linha do assinante. É o maior valor de R Stasinszky (2022). **I_s** – Todas as deficiências se manifestam simultaneamente com o sinal de voz. Os principais fatores contribuintes são as classificações de volume do aparelho de telefone, classificação de volume do tom lateral e o número de unidades de distorção de quantização Stasinszky (2022). **I_d** – As deficiências introduzidas pelo atraso. Os fatores que contribuem para essas deficiências são a quantidade de atraso presente na rede, bem como os valores das classificações de sonoridade do eco do locutor e do ouvinte Stasinszky (2022). O valor de I_d é normalmente pequeno e não degrada significativamente o R-Factor. A degradação pode ser significativa caso o atraso se torne grande o suficiente. **I_{e-eff}** – Um componente de deficiência maior do que I_d que compreende deficiências resultantes de CODECs de baixa taxa de bits, perda de pacotes e rejeição de pacotes Stasinszky (2022). **I_{recência}** – As deficiências causadas pela perda significativa de pacotes – oito ou mais pacotes perdidos em uma linha Stasinszky (2022). **A** – O fator de vantagem, que permite compensar os fatores de comprometimento e representa

a tolerância do usuário à degradação da qualidade da voz. Para comunicação com fio, o componente A do fator de vantagem é sempre 0 porque o dispositivo de teste não tem conhecimento prévio desse componente Stasinszky (2022).

Na Figura 2.5, apresentamos o *E-Model*:

Figura 2.5 – Configuração básica de referência do *E-model*



Fonte: Adaptado de Chochol (2023)

2.12 Avaliação da Qualidade da Voz por Percepção (PESQ)

Existem outros meios de avaliação de qualidade de voz, como o *Mean Opinion Score* (MOS) e o *Perceptual Evaluation of Speech Quality* (PESQ). Ambos foram criados para avaliar a qualidade de voz em sistemas de telefonia, porém o MOS depende da opinião dos ouvintes, o que torna o processo caro e demorado. Já o PESQ utiliza um modelo matemático para prever a qualidade da voz percebida pelos usuários, sendo mais preciso e eficiente em diversas condições de rede, incluindo conexões analógicas, codecs, perda de pacotes e atraso variável (RIX et al., 2001).

Um dos algoritmos mais utilizados para a avaliação objetiva da qualidade de voz em sistemas de transmissão e/ou codificação de áudio é o MOS (Mean Opinion Score), que é uma medida subjetiva amplamente utilizada para avaliar a qualidade da voz. O MOS é obtido por meio de testes de avaliação subjetiva, nos quais os ouvintes classificam a qualidade da voz em uma escala de 1 a 5, sendo 1 a pior qualidade e 5 a melhor qualidade (SOUSA, 2011).

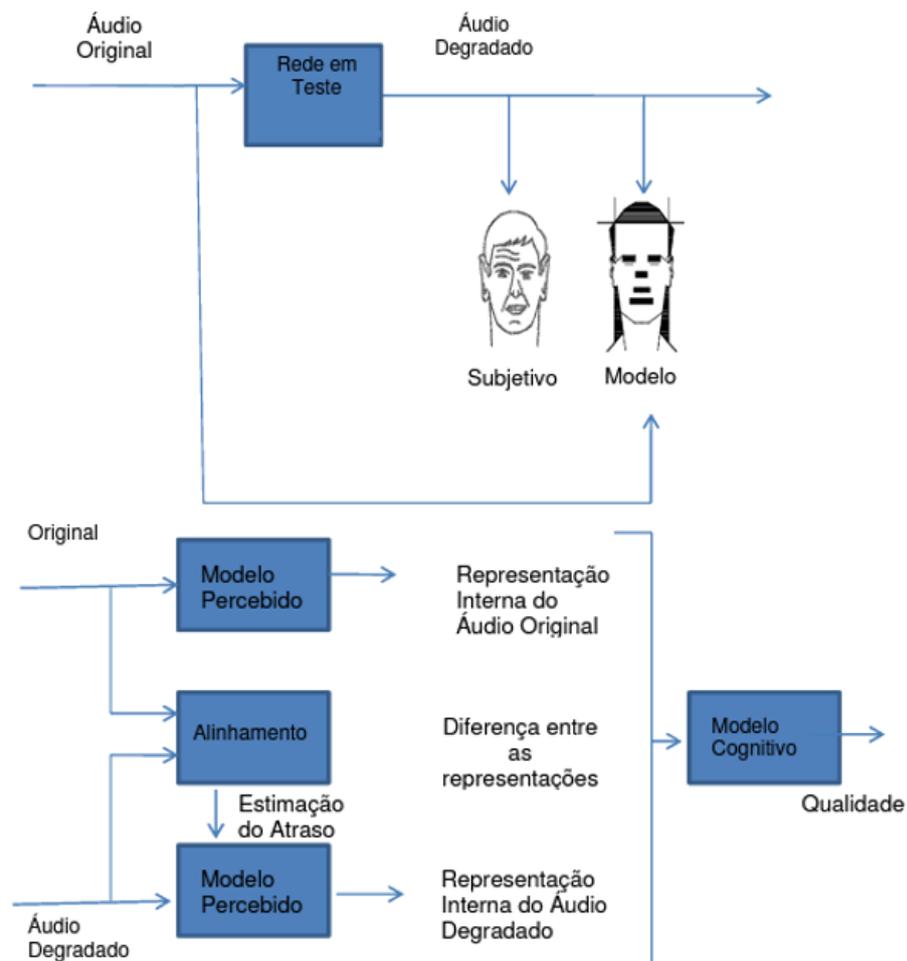
A recomendação ITU-T P.862, também conhecida como PESQ, é um método utilizado para avaliar a qualidade da fala percebida pelos seres humanos. Ele compara um sinal de voz original, representado por $X(t)$, com um sinal degradado resultante da transmissão desse sinal por um sistema de comunicações, representado por $Y(t)$. O PESQ utiliza um modelo perceptual para prever a qualidade da fala percebida pelos seres humanos com base no sinal degradado $Y(t)$. Esse modelo leva em considera-

ção diversos fatores perceptuais, como distorção, atraso, ruído, entre outros, que afetam a qualidade da fala. (GOUVEA et al., 2018).

Ao aplicar o PESQ, obtém-se uma previsão da qualidade percebida pelos ouvintes em um teste de escuta subjetiva, ou seja, uma estimativa da qualidade que os seres humanos atribuiriam ao sinal degradado $Y(t)$. Essa previsão é expressa em uma escala objetiva, como o MOS, que representa a média das avaliações subjetivas realizadas por um grupo de ouvintes GOUVEA et al. (2018).

O PESQ faz comparação de um sinal original $X(t)$ com um sinal degradado $Y(t)$ que é a resultância da passagem de $X(t)$, por um sistema de comunicações. A saída do PESQ é uma previsão da qualidade percebida que seria dada a $Y(t)$, por pessoas, em um exame de audição subjetivo (StudFiles, 2015). Na Figura 2.6, ilustra-se a Algoritmo PESQ:

Figura 2.6 – Algoritmo PESQ



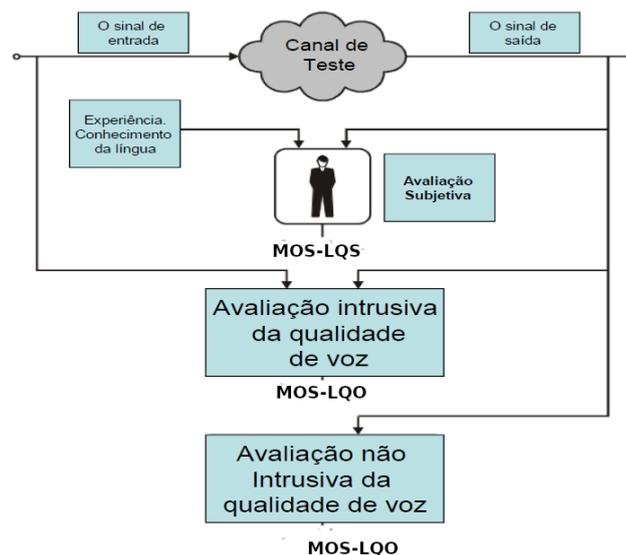
Fonte: Adaptado de ITU-TP.862 (2001)

Baseado no conjunto de atrasos encontrados, o PESQ compara o sinal original (entrada) com a saída degradada alinhada do dispositivo em teste, usando um modelo perceptual.

A chave do processo é a transformação de ambos os sinais originais e degradados em uma representação interna análoga à representação psicofísica (modelo perceptual) dos sinais de áudio, no sistema auditivo humano, conforme ilustração da Figura. 4.1. 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. Essas 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 MOSLQO "Mean Opinion Score - Listening Quality Objective"(Pontuação Média de Opinião - Qualidade de Audição Objetiva), isto é, uma estimativa do resultado dos testes subjetivos reais MOSLQS "Mean Opinion Score - Listening Quality Low". É uma metodologia subjetiva que avalia a qualidade de voz percebida pelo usuário (StudFiles, 2015).

Por meio de resultados experimentais, estabelece-se a magnitude do estudo da disposição das perdas de pacotes e sua relação com os índices MOS e, também, sua localidade temporal. O índice MOS pode ser determinado e, para isso, são utilizados dois procedimentos de avaliação objetivos e subjetivos de qualidade de voz, visto que a métrica objetiva se divide em intrusivas e não intrusivas com mostra a a figura 4.2 (PEREIRA; RODRIGUEZ, 2016).

Figura 2.7 – Estimadores da Qualidade da Voz



Fonte: Adaptado de Rodríguez D.Z (2009)

Segundo (L. Malfait et al., 2006) a conversação humana sobre Internet (VoIP), em especial, a análise e previsão da qualidade da fala, é uma área que vive em constante pesquisa. Vários métodos foram desenvolvidos, para a avaliação da qualidade da fala transmitida, entre esses métodos, podemos citar os intrusivos e não intrusivos. De modo geral, os procedimentos não intrusivos são adequados para

avaliar a qualidade da fala ao vivo. O padrão ITU-T, recentemente proposto, o P.563, é um método não intrusivo, desenvolvido para a avaliação da qualidade da fala. Porém não é adequado para condições de VoIP e produz resultados não confiáveis.

3 CODIFICADORES DE VOZ

Neste capítulo, serão apresentados alguns exemplos importantes de codificadores de voz, cada um com suas características próprias, vantagens e desvantagens, dependendo das necessidades específicas de aplicação.

Vale a pena mencionar que os codecs de áudio são extremamente importantes para quem trabalha com produção musical, sonorização de vídeos e transmissão de conteúdo online, pois permitem que arquivos de áudio sejam transmitidos com qualidade adequada em diferentes formatos e plataformas. Além disso, os codecs ajudam a economizar espaço de armazenamento e reduzem o tempo de carregamento de arquivos de áudio em dispositivos e redes.

O padrão G.701 é importante porque garante a interoperabilidade e compatibilidade entre diferentes equipamentos de telecomunicações de diferentes fabricantes. De acordo com a recomendação ITU-T G.701 a definição de *codec* é "uma combinação de um codificador e um decodificador que operam em direções opostas de transmissão no mesmo equipamento." Comumente pode-se dizer que a palavra *codec* é um acrônimo de compressão e descompressão, especialmente no que se refere aos formatos de arquivos de imagem, áudio e vídeo, usados em computadores e na internet, ao passo que, em telecomunicações, o conceito de codificação (origem) e decodificação (destino) (REC.G.107, 1993).

3.1 G.711

O G.711 é capaz de fornecer qualidade de áudio de alta fidelidade em redes locais, o que o torna uma escolha popular para sistemas de telefonia e outras aplicações de áudio em redes corporativas. O G.711 é, naturalmente, escolhido para redes locais. Usa uma escala semi logarítmica, chamada de PCM (*Pulse Code Modulation*), aumentando, dessa forma, os sinais de baixa amplitude, enquanto os de alta amplitude, são tratados de forma proporcional, opera de forma análoga ao ouvido humano (HARFF, 2008).

3.2 G.726

O G.726 usa uma técnica chamada *Adaptive Differential Pulse Code Modulation* (ADPCM) para comprimir e descomprimir sinais de voz. Ele fornece quatro taxas de bits diferentes (16 kbps, 24 kbps, 32 kbps e 40 kbps) para suportar vários requisitos de largura de banda e qualidades de sinal. A relação entre os sinais de frequência de voz e as leis de codificação/decodificação PCM é totalmente especificada para G.726 na Recomendação G.711 (TECHNOLOGIES, 2017).

O G.726 perdeu espaço na década de 1990 por causa de problemas de compatibilidade com sinais de modem e fax. Isso aconteceu porque esses sinais foram projetados para operar em uma faixa

de frequência específica, e o G.726 não conseguia processá-los adequadamente. Atualmente, o G.726 está sendo novamente aceito por causa do seu baixo consumo de largura de banda e baixo processamento exigido. Ele oferece a mesma qualidade de voz que o G.711, porém usando a metade da largura de banda necessária, o que sem dúvida é uma vantagem muito importante em sistemas de telefonia e comunicação digital onde a largura de banda é limitada (COELHO, 2008).

3.3 G.729

G.729 é um padrão de áudio bem conhecido capaz de fornecer som de alta qualidade por VoIP. Baseado no popular modelo de codificação *Code Excited Linear Prediction (CELP)*, o G.729 é um padrão ITU-T para telefonia de longa distância de alta velocidade e qualidade. O G.729 usa apenas a taxa de bits de 8 kbps e comprimento de quadro de 10 ms, permitindo *streaming* de baixa taxa de bits com atraso mínimo. O G.729 é um codec muito eficiente para comunicação em redes com largura de banda limitada, ou seja ele tem capacidade de equilibrar a qualidade de áudio com o consumo de largura de banda suavizando o ambiente e permitindo uma comunicação clara (KONG et al., 2008).

3.4 Codificador Adaptativo Multi-Rate (AMR)

O Codec AMR-NB (*Adaptive Multi-Rate Narrowband*) e o Codec AMR-WB (*Adaptive Multi-Rate Wideband*) foram de fato originalmente projetados para sistemas de rádio móveis comutados por circuito. Esses codecs foram desenvolvidos para otimizar a qualidade da voz em redes de comunicação com largura de banda limitada, com o objetivo de fornecer uma experiência de voz clara e robusta em ambientes de comunicação móvel (SJOBERG et al., 2007).

As principais vantagens de usar o codec AMR-WB é, sem dúvida, maior naturalidade da voz; percepção aprimorada de voz, melhor reconhecimento de alto-falante; e maior qualidade para chamadas em um ambiente barulhento (ABICHANDANI et al., 2017).

O codec de banda larga AMR é o mais novo codec de voz padronizado para sistemas GSM e WCDMA 3G. Embora todos os codecs anteriores em comunicação móvel operem em largura de banda de áudio limitada a 200–3400 Hz, o AMR-WB amplia a largura de banda de áudio para 50–7000 Hz, melhorando consideravelmente a qualidade. O codec AMR-WB opera em nove taxas de bits de codificação de voz entre 6,6 e 23,85 kbit/s. Como os outros codecs GSM e WCDMA 3G, o AMR-WB também possui um modo dependente da fonte de baixa taxa de bits para codificação de ruído de fundo (SALAMI et al., 2002).

O codificador AMR é adaptável a diversas situações de canal e pode selecionar automaticamente o modo de codificação mais apropriado para manter a melhor qualidade de fala possível. O AMR-NB

e o AMR-WB são baseados na tecnologia ACELP e utilizam diferentes modos de taxa de bits para se adaptar a diferentes condições de canal. (BYUN et al., 2006).

Atualmente existem vários tipos de *codecs*, mas alguns são mais usados como o *Global System for Mobile* (GSM), e o *Adaptive Multi-Rate* (AMR). Existem duas versões do AMR, uma versão banda larga (AMR-WB) e outra banda estreita (AMR-NB). O *codec* possui diversos modos diferentes, para AMR-NB, para AMR-WB, cada um com uma relação diferente entre a codificação da fonte e do canal, mas com uma taxa de bits combinada fixa. O que torna esse *codec* exclusivo é fato dele não possuir relacionamento fixo entre a codificação da fonte e o canal. Com uma taxa de bits combinada fixa, pode, assim, alternar dinamicamente entre taxas de bits diferentes para AMR-NB de 4,75, 5,15, 5,90, 6,70, 7,40, 7,95, 10,20 e 12,20 kbit/s e, para AMR-WB de 6,60, 8,85, 12,65, 14,25, 15,85, 18,25 e 19,85 kbit/s, de acordo com as condições da rede (BOŘIL, 2018).

O *codec* AMR-WB, conhecido como *Adaptive Multirate Wideband*, foi selecionado pela *Third Generation Partnership* para o sistema de comunicação móvel de terceira geração GSM e WCDMA para fornecer serviços de voz em banda larga. O *codec* AMR-WB também foi selecionado em julho de 2001 pelo Setor de Telecomunicações da União Internacional de Telecomunicações para trabalhos de padronização para codificação de voz em banda larga em torno de 16 kbps e foi aprovado em janeiro de 2002 como Recomendação G.722.2. O AMR-WB usa uma largura de banda de áudio estendida de 50Hz a 7kHz e oferece qualidade de fala e naturalidade de voz superiores em comparação com os sistemas de comunicação móvel de segunda e terceira geração existentes (BESSETTE et al., 2002).

A adoção do AMR-WB pelo ITU-T é super importante, pois, pela primeira vez, o mesmo *codec* é adotado, para serviços sem fio, bem como de telefonia fixa. Isso exclui a necessidade de transcodificação e facilita a implementação de aplicativos e serviços de voz de banda larga em uma ampla variedade de sistemas e plataformas de comunicação (BESSETTE et al., 2002).

Similar ao AMR, o *codec* AMR-WB também é um codificador de fala multimodo. Ele suporta nove modos de codificação de voz de banda larga com taxas de bits, variando de 6,6 a 23,85 kbps. A frequência usada nas amostras é 16000 Hz o que permite a captura de uma gama mais ampla de frequências da voz humana do que o *codec* AMR, que usa uma frequência de amostragem de 8000 Hz. e o processamento da fala é realizado em quadros de 20 ms. Isso caracteriza que a cada quadro codificado por AMR-WB representa 320 amostras de fala (SJOBERG et al., 2007).

Atualmente o sistema LTE são o AMR-WB e o AMR-NB, conforme definido pela em 2013 pela 3GPP (*3rd Generation Partnership Project*), que é uma colaboração entre várias organizações de desenvolvimento de padrões de telecomunicações. O *codec* AMR-WB é utilizado para fornecer alta qualidade de voz em chamadas de voz em HD, enquanto o AMR-NB é utilizado em chamadas de voz padrão. O codificador de voz AMR, de banda estreita (AMR-NB), trabalha com uma banda limitada de

[200; 3400] Hz. Opera à frequência de 8kHz, com uma resolução de 13 bits e, uma vez que o tempo de trama é de 20ms, cada trama é transmitida com 160 samples (PEDRO, 2011).

AMR (*Multi-Rate Adaptável*) é um codificador de voz que trabalha, em taxas de oito bits, na faixa de 4,75 a 12,2 kbps e foi programado, especificamente, para melhorar a potência do link. Essa tecnologia foi criada, no início, para sistemas GSM (Sistema Global para Comunicações Móveis), o padrão de telecomunicações móveis 2G mais implantado em todo o mundo (VOICEAGE, 2019). O codec de banda estreita AMR foi padronizado pelo *European Telecommunications Standards Institute* (ETSI) e adotado pelo 3GPP (Projeto de Parceria da 3ª Geração) obrigatório para sistemas sem fio 2.5G e 3G, baseado na evoluída rede principal GSM (Sistema Global para Comunicações Móveis); W-CDMA, (abreviação de Wide-Band Code-Division Multiple Access) é a tecnologia 3G líder e é usado em redes UMTS e FOMA; EDGE ou *Enhanced Data Rates For GSM Evolution* (Taxas de Dados Ampliadas para a Evolução do GSM em inglês); GPRS é a sigla de *General Packet Radio Services*, ou Serviços Gerais de Pacote por Rádio (VOICEAGE, 2019).

O codec AMR-NB é um algoritmo de compressão de áudio amplamente utilizado em telecomunicações móveis, como em chamadas de voz em redes GSM (*Global System for Mobile Communications*). O quadro a seguir apresenta algumas características importantes do codec AMR-NB:

Quadro 3.1 – Características Do Codec AMR-NB

CARACTERÍSTICAS	DESCRIÇÃO
Largura de banda codificada	200-3400 Hz
Padronizado	ETSI 1999
Tipo de codificação	ACELP (Previsão Linear Animada do Código Algebrico)
Taxa de bits	Varia de 4,75 kbps a 12,2 kbps, dependendo do modo de operação
Tamanho do quadro de atraso	20 ms
Qualidade	Pedágio
Complexidade MIPS	15-20
RAM (palavras)	4 K

Fonte: Adaptado de Järvinen (2014)

O AMR-WB é o codificador perfeito para aplicativos de voz em banda larga em redes convergentes com ou sem fio sendo o primeiro codificador a ser padronizado para aplicativos (3GPP) e sistemas a cabo (recomendação ITU-T G.722.2) (VOICEAGE, 2019).

O codificador AMR-WB surgiu como uma extensão do *codec* AMR-NB, para banda larga, [50,7000] Hz, operando em uma frequência de 16kHz com uma resolução de 14 bits. Duas bandas, [50,6400] Hz e [6400,7000] Hz, são codificadas, separadamente, de maneira que possam diminuir a complexidade algorítmica e o ritmo de processamento, concentrando a localização de bits na gama de frequência de maior relevância perceptiva (PEDRO, 2011).

Na Quadro 3.2, são apresentados nove modos de codificação de voz e outras características importantes do AMR-WB, como um modo para a codificação de ruído de fundo, semelhante ao existente no AMR-NB. Os modos de maiores débitos binários correspondem a modos do codec, para voz de alta qualidade, enquanto os dois modos de menor débito são usados apenas, em situações de carácter temporário, em situações extremas do estado do canal ou durante períodos de congestionamento da rede. No AMR-WB, o sistema de VAD baseia a sua decisão na divisão da banda dos [0; 6,4] kHz do sinal de voz em 12 sub-bandas e no cálculo do nível do sinal em cada banda (PEDRO, 2011).

Quadro 3.2 – Características Do Codec AMR-WB

CARACTERÍSTICAS	DESCRIÇÃO
Largura de banda codificada	50-7000 Hzz
Padronizado	ITU-T / 3GPP 2002/2001
Tipo de codificação	ACELP (Previsão Linear Animada do Código Algebrico)
Taxa de bits	23,85 / 23,05 / 19,85 / 18,25 / 15,85 / 14,25 / 12,65 / 8,85 / 6,6 kbps
Atraso (ms): Tamanho do quadro	20 ms
Qualidade	Bom desempenho de fala a taxas de 12,65 kbps e mais 15,85 > = G.722 @ 56 23.05 > = G.722 @ 64
Complexidade: RAM WMOPS (palavras)	38 (incluindo processamento VAD / CNG) 5,3 K
VAD - DTX - GNC	Incluído

Fonte: Järvinen (2014)

3.5 Codificador Enhanced Voice Service (EVS)

Já o *Enhanced Voice Services* (EVS), desenvolvido pelo *Fraunhofer Institute for Integrated Circuits* em colaboração com a 3GPP, é um codec de áudio de alta qualidade que suporta diversas taxas de bits e é capaz de fornecer uma melhor qualidade de áudio em comparação com o AMR-WB (VOICE-AGE, 2019).

No ano de 2010, o 3GPP (*3rd Generation Partnership Project*) terminou um estudo relacionado a como o 3GPP poderia fazer, para manter alto valor e competitividade, em seus serviços de voz e, se o novo pacote de sistemas evoluídos, o LTE (*Long Term Evolution*), poderia abrir novas possibilidades para o desenvolvimento de serviços, em termos de qualidade de voz e áudio. O estudo considerou também casos de uso móveis importantes ao acesso e uso de LTE, casos que podem vir a se beneficiar da qualidade de áudio aperfeiçoada. No estudo, também, pesquisou-se a necessidade de codecs aperfeiçoados, além do AMR-WB (*Adaptive Multi-Rate Wideband*), o codec agora implantado no fornecer voz HD

em sistemas móveis 3GPP. Um adicional importante é que o estudo examinou como os novos serviços de voz aperfeiçoados poderiam interagir com os serviços de voz HD existentes (SCHNELL, 2015).

O Codec 3GPP para EVS foi projetado especialmente para redes de voz móvel com comutação de pacotes, como Voz sobre LTE (VoLTE) e telefonia IP geral, este dispositivo oferece suporte para telefonia NB, WB, SWB (banda super larga, 50–14000 Hz) e FB (banda completa, até 20 kHz), além de minimizar a degradação (MITTAG et al., 2018).

O codificador de voz e áudio EVS, da próxima geração de telefonia (móvel) e comunicação, teve seu início, em 2007, mas o processo de padronização durou quatro anos, de 2010 a 2014. O codificador EVS é o primeiro codec de comunicação 3GPP, com largura de banda de áudio super larga, usado para melhorar a qualidade da fala de 9,6 kbps. Ao mesmo tempo, o desempenho do *codec*, em sinais como música ou conteúdo misto, é comparável aos codecs de áudio modernos (3GPP, 2014).

O codec *Enhanced Voice Services* substituiu o *codec* de voz AMR-WB e foi padronizado pelo *3rd Generation Project* em setembro de 2014. Esse codificador utiliza tecnologia de ponta para melhorar a qualidade, eficiência e versatilidade da comunicação dos sistemas de comunicação móvel 3GPP (VOICEAGE, 2019).

Várias empresas que representam a indústria, incluindo chipsets, dispositivos e infraestrutura, operadoras e tecnologia, contribuíram para o desenvolvimento do *codec* EVS (VOICEAGE, 2015).

O codec EVS é dotado de um conjunto completo de recursos para melhorar o desempenho geral dos serviços de voz em diferentes tipos de redes e em diferentes tipos de bandas (COHEN, 2016).

Os *codecs* de voz 3GPP (EVS) foram finalizados com o objetivo de fornecer qualidade aprimorada de voz e música para serviços de conversação sobre LTE. Estendem-se os tradicionais codificadores de voz de banda estreita (NB) e de banda larga (WB) para *superwideband* (SWB) e *fullband* (FB) e também melhora a codificação genérica de áudio complementando a codificação de voz com uma codificação de transformação. O sistema alterna-se entre os dois esquemas de codificação, dependendo do conteúdo do sinal de entrada. Sinais semelhantes à fala são transmitidos para um codificador de fala baseado em ACELP enquanto a música é manipulada por um codificador de áudio genérico, baseado em MDCT (FUCHS et al., 2015).

O *codec* EVS usa conceitos similares aos de seus antecessores, como AMR-WB, e mantém conexão com versões anteriores em todas as nove taxas de bits de 6,6 kbit/s a 23,85 kbit/s (NAHUZ et al., 2008).

O *Enhanced Voice Services (EVS)*, é o primeiro codec 3GPP capaz de oferecer um serviço de áudio de banda larga de 20 kHz; é um codificador de voz e áudio, com telefonia móvel, que fornece uma qualidade de voz compatível com outras entradas de áudio, como música e conteúdo misto; exhibe

simultaneamente grande força para atraso jitter e perda de pacotes o que o torna uma preferência para aplicativos de telefonia móvel (PRADO, 2009).

O EVS é o primeiro *codec* de conversação 3GPP de alta qualidade, que oferece uma largura de banda de áudio de até 20 kHz. O EVS é usado em várias aplicações, como chamadas de voz em alta definição, conferências telefônicas, serviços de mensagens de voz e vídeo, transmissão de música e muito mais. Além disso, o *codec* EVS está sendo continuamente aprimorado para fornecer ainda mais recursos e qualidade de áudio avançada (FLYNN, 2022).

O EVS é uma tecnologia promissora para melhorar a qualidade de voz nas redes móveis. É uma evolução do serviço de voz tradicional, projetado para fornecer uma experiência de chamada de alta qualidade aos usuários GSA (2016). Essa tecnologia proporciona uma faixa de frequência mais ampla, permitindo uma reprodução do som mais precisa e natural. Além disso, o EVS incorpora técnicas avançadas de codificação de voz, como a codificação de áudio de banda larga e a codificação adaptativa, com o objetivo de garantir uma qualidade de voz consistente mesmo em condições de rede desafiadoras (GSA, 2016).

A tecnologia EVS, oferece recursos avançados de ocultação de pacotes e capacidade anti-jitter para garantir a melhor qualidade de voz, mesmo quando o usuário está em movimento ou em diferentes configurações de rede. Essa tecnologia parece ser útil para garantir a qualidade de comunicação de voz em ambientes desafiadores, como em redes celulares em que a conexão pode ser instável (CHEWE, 2015).

Em setembro de 2014 o *codec* EVS foi padronizado e avaliado, em 3GPP, mas, para 4G, por causa do seu lançamento ter sido um pouco tarde, somente alguns terminais suportavam o codificador. Já que o EVS não só exercita voz HD, mas também tornava melhor a cobertura, por fim, ele foi incluído no padrão 3GPP, como uma escolha obrigatória para voz 5G, graças aos conjuntos de operadoras de telecomunicações e fornecedores de equipamentos de telecomunicações (MARTÃO, 2018).

A técnica de ajuste de taxa do *codec* EVS mencionada no contexto é conhecida como algoritmo de reconhecimento de QoE *Quality of Experience*. Esse algoritmo é utilizado pelo equipamento do usuário (UE) para determinar a taxa de bits ideal para operar com base na condição do link e na qualidade da experiência do usuário durante uma chamada EVS (PRASAD et al., 2019).

O codificador EVS fornece larguras de banda mais altas, taxas de bits mais baixas e maior resistência a rajadas. As operadoras de telefonia fixa poderiam obter benefícios semelhantes, se implantassem o EVS, no entanto o custo e a complexidade associados à implementação de um *codec*, como o EVS em redes de telefonia fixa e telefones, significa que ainda não existem tais implementações (HUMEIDA et al., 2018).

3.5.1 Especificações Técnicas do Codec do EVS

As referências são específicas (identificadas pela data de publicação, número da edição, número da versão etc). No Quadro 3.3, as Especificações Técnicas novas e atualizadas que oferecem suporte para EVS em 3GPP são: (3GPP, 2022)

Quadro 3.3 – Especificações Técnicas Do Codec EVS

CARACTERÍSTICAS	DESCRIÇÃO
3GPP TS 26.441	Visão geral
3GPP TS 26.442	código ANSI C (ponto fixo)
3GPP TS 26.443	código ANSI C (ponto flutuante)
3GPP TS 26.444	Sequências de teste
3GPP TS 26.445	Descrição algorítmica detalhada
3GPP TS 26.446	Funções retrocompatíveis com AMR-WB
3GPP TS 26.447	Ocultação de erro de pacotes perdidos
3GPP TS 26.448	Gerenciamento de buffer de jitter
3GPP TS 26.449	Aspectos de geração de ruído de conforto (CNG)
3GPP TS 26.450	Transmissão Descontínua (DTX)
3GPP TS 26.451	Deteção de Atividade de Voz (VAD)
3GPP TS 26.114	P Multimedia Subsystem (IMS); Telefonia multimídia; Manipulação de mídia e interação
3GPP TR 21.905	Vocabulário para especificações 3GPP.
3GPP TR 26.952	Codec para Serviços de Voz Avançados (EVS); Caracterização de Desempenho

Fonte: Järvinen (2014)

A Especificação Técnica 3GPP TS 26.441 apresenta uma visão geral do Codec EVS, descrevendo sua arquitetura e recursos, incluindo o esquema de taxa de fonte controlada, detector de atividade de voz e som, sistema de geração de ruído de conforto e mecanismo de mascaramento para lidar com erros de codificação (ETSI, 2020).

Já a Especificação Técnica 3GPP TS 26.442 trata-se de uma cópia eletrônica do código ANSI-C para o EVS. O 3GPP TS 26.442 é um padrão técnico desenvolvido pelo 3 GPP, uma organização de padronização que define as especificações para as redes móveis de próxima geração. O TS 26.442 faz parte do conjunto de especificações do 3GPP para a tecnologia de comunicação de voz sobre LTE (VoLTE) fixadas em 3GPP TS 26.444 (3GPP, 2017). Essa especificação é intitulada "Codec de áudio adaptável de baixa latência para Serviço de Comunicação de Voz sobre LTE (VoLTE). O objetivo principal do TS 26.442 é definir um codec de áudio de baixa latência e alta qualidade para permitir comunicações de voz confiáveis e de alta fidelidade em redes LTE (3GPP, 2017).

A especificação 3GPP TS 26.443 é importante para garantir a interoperabilidade entre dispositivos móveis e sistemas de rede de diferentes fabricantes e operadoras, permitindo que os usuários finais desfrutem de áudio de alta qualidade em seus dispositivos móveis, independentemente de sua localização ou operadora. Esta Especificação Técnica (TS) foi produzida pelo *ETSI 3rd Generation Partnership*

Project (3GPP) e trata-se de uma cópia eletrônica do código de ponto flutuante ANSI C, para os Serviços de Voz Avançados(EVS) *Codec* e que foi utilizada nos testes realizados neste trabalho (3GPP, 2022a).

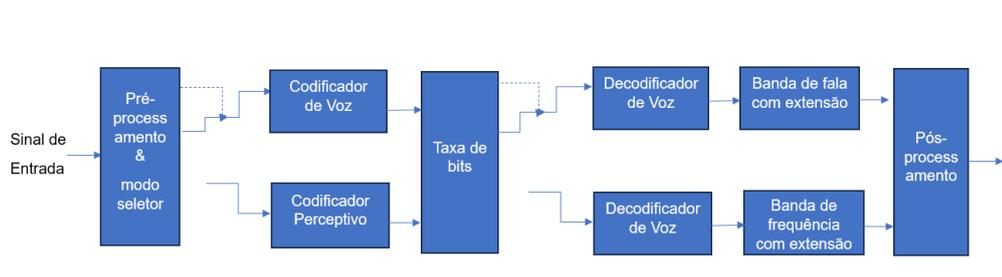
A Especificação Técnica 3GPP TS 26.444 a trata-se de Aspectos do Sistema; Codec para Serviços de Voz Aprimorados (EVS); Sequências de teste. São testes de sequências, para uma implementação de bits exatos do Codec EVS (3GPP TS 26.445); Detecção de Atividade de Voz (VAD)(3GPP TS.26.451); Geração de ruído de conforto (3GPP TS 26.449); Transmissão descontínua (DTX) (3GPP TS26.450); Ocultação de erro de pacotes perdidos (3GPP TS 26.447); Gerenciamento de buffer de jitter (JBM) (3GPP TS 26.448) e Função Interoperável AMR-WB (3GPP TS 26.446). Além do mais, o presente documento especifica a conformidade teste (3GPP, 2022a).

A Especificação Técnica 3GPP TS 26.445 é uma especificação técnica que define os codecs de áudio e vídeo usados em serviços multimídia em redes IMS (*IP Multimedia Subsystem*). A especificação fornece diretrizes para o uso de codecs, incluindo H.264, H.265 e VP9, em serviços de comunicação multimídia, como videochamadas e conferências (3GPP, 2022a).

A versão 17.0.0 Release 17 é a versão mais recente do 3GPP TS 26.445, publicada em dezembro de 2021. Esta versão inclui atualizações e melhorias nos codecs de áudio e vídeo, bem como novos recursos e funcionalidades para aprimorar os serviços de comunicação multimídia em redes IMS (3GPP, 2022b).

Na Figura 3.1, mostra partes de áudio, incluindo a conversão de analógico para digital e digital para analógico, estão incluídas, para mostrar o caminho completo da fala na entrada/saída de áudio no Equipamento do Usuário (UE) e a interface digital da rede. Além disso a figura apresenta uma visão geral das funções de processamento de áudio do lado transmissão e Lado de recepção.

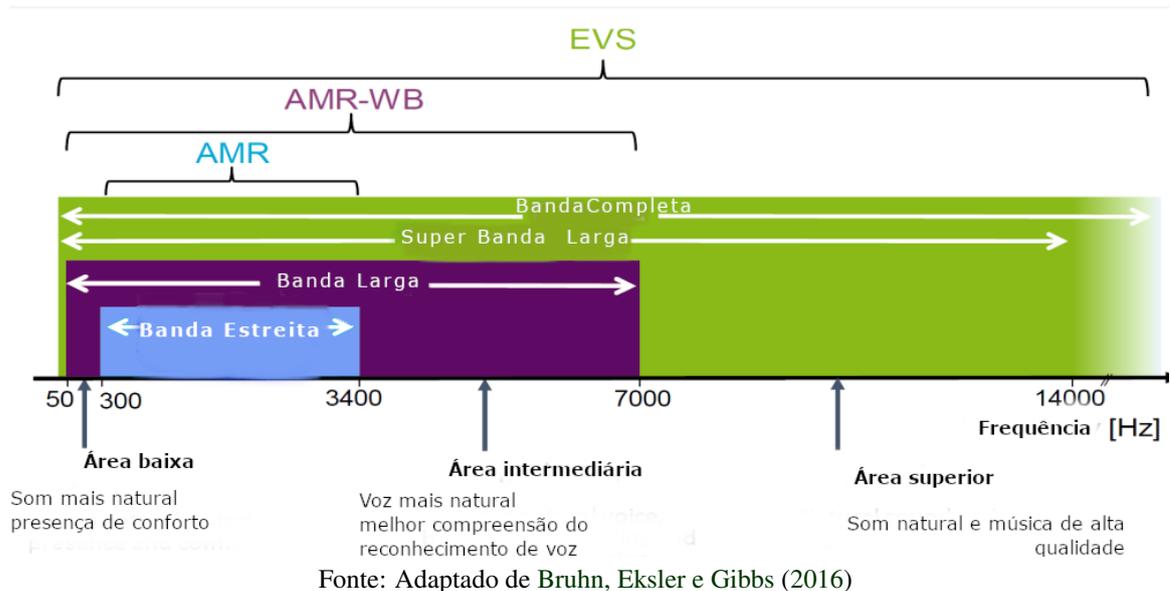
Figura 3.1 – Transcodificação



Fonte: Adaptado de Lutzky (2015)

Os codecs EVS e AMR-WB são usados para codificar áudio de alta qualidade em comunicações de voz sobre IP e telefonia móvel. Ambos foram projetados para melhorar a qualidade de voz e a eficiência de compressão em comparação com codecs mais antigos, como o G.711. A figura abaixo, ilustra um paralelo entre os padrões EVS e AMR-WB codecs.

Figura 3.2 – Comparação EVS e AMR



O codec EVS, é famoso por ser o primeiro codec de comunicação 3GPP capaz de fornecer superlargura de banda com excelente qualidade a sem intervalo de tempo a 9,6 kbps. Ao mesmo tempo, o desempenho do codificador, com tipos de sinal, como música ou misto, corresponde aos codecs de áudio modernos. A tecnologia subjacente do codec é um design versátil entre modos de codificação especializados para voz e música. O codec EVS é uma tecnologia de áudio avançada que foi desenvolvida por várias empresas líderes do setor de telecomunicações entre elas, temos: Ericsson, Fraunhofer IIS, Huawei Technologies Co. Ltd, NOKIA Corporation, NTT, NTTDOCOMO INC., ORANGE, Panasonic Corporation, Qualcomm Incorporated, Samsung Eletronics Co. Ltd, VoiceAge e ZTE Corporation como mostra a figura (BRUHN, 2016).

3.5.2 Recursos do codec EVS

Máxima transparência, mais qualidade e maior eficiência em serviços de banda estreita (NB) e banda larga (WB), para uma grande faixa de taxa de bits, iniciando em 7,2 kbps (3GPP, 2021).

- Voz banda larga transparente (a 24,4 kbps);
- Banda larga com conteúdo misturado e música transparente iniciando em (24,4kbps);
- Superior em todas as taxas de compreensão;
- Qualidade superior ao AMR-NB e AMR-WB em taxas parecidas;
- Atributos SWB sem taxas móveis prematuras iniciando em 9,6 kbps;
- Respectivamente oferece um passo expressivo na qualidade relativamente a telefonia tradicional NB e WB com operação super-WB SWB E Full-band iniciadas em 9,6 e 16,4 kbps, na ordem devida;

- g) Características preliminares pelo áudio SWB (largura de banda de 14 ou 16 kHz), em taxas de bits móveis de 9,6 kbps para cima, ótima qualidade de discurso, conteúdo misto e música;
- h) Competência para substituir a taxa de bits a cada quadro de 20 ms concede ao codec facilmente adaptação para mudanças na capacidade do canal;
- i) Funciona muito bem em conjunto com AMR-WB integrado às nove taxas de bits de 6,6 a 23,85 kbps;
- j) Outros recursos importantes como modo de taxa de bits variáveis (VBR) em operações NB e WB com taxa média de 5,9 kbps;
- k) Funciona com AMR-WB, em todas as taxas de bits, com excelente qualidade especialmente para ocultação e perda de pacotes;
- l) Transmissão descontínua (DTX) com algoritmos para detecção de atividade de voz/som (VAD) e geração de ruído de conforto (CNG);
- m) Recursos de ocultação de erros resistente para tornar menos intenso o impacto na qualidade de pacotes perdidos - adequado para Voz sobre WLAN;
- n) Sistema para gerenciamento de buffer de jitter (JBM - Jitter Buffer Management);
- o) Trabalho incessante em 3GPP para introduzir EVS em W-CDMA, trabalha em conjunto de forma transparente entre LTE e W-CDMA usando SRVCC.

3.5.3 Característica do EVS

O EVS é uma tecnologia avançada de codec de voz que oferece comunicação de voz de alta qualidade em várias redes, incluindo 2G, 3G, 4G e 5G. Aqui estão as principais características do EVS:

Qualidade de voz aprimorada: EVS fornece qualidade de áudio superior em comparação com codecs de voz anteriores, como AMR-WB (banda larga adaptativa de várias taxas). Ele oferece uma gama mais ampla de frequências de áudio e melhor inteligibilidade de fala, resultando em uma reprodução de voz mais natural e realista (FEASLEY, 2021).

Eficiência de baixa taxa de bits: O EVS foi projetado para garantir voz de alta qualidade ao mesmo tempo em que otimiza o uso da largura de banda. Ele consegue isso codificando dados de voz com eficiência e reduzindo a taxa de bits necessária para transmissão, resultando em um uso mais eficiente dos recursos de rede (FEASLEY, 2021).

Controle adaptável de taxa de bits: o EVS ajusta dinamicamente a taxa de bits com base nas condições da rede e na largura de banda disponível. Esse comportamento adaptativo garante que a qualidade de voz permaneça consistente mesmo em condições de rede desafiadoras, como conexões congestionadas ou não confiáveis (3GPP, 2014).

Supressão de ruído de fundo aprimorada: EVS emprega algoritmos avançados para reduzir o ruído de fundo e melhorar a inteligibilidade da fala. Ele pode efetivamente suprimir o ruído ambiente,

como sons de tráfego ou barulho de multidão, permitindo uma comunicação clara mesmo em ambientes ruidosos (3GPP, 2017).

Suporte de banda larga e banda super larga: EVS suporta uma ampla gama de frequências de áudio, incluindo áudio de banda larga (50 Hz a 7 kHz) e banda super larga (20 Hz a 14 kHz). Essa faixa de frequência estendida contribui para uma experiência de comunicação de voz mais natural e imersiva (3GPP, 2017).

Transição perfeita entre redes: o EVS foi projetado para oferecer suporte a transferências de voz perfeitas entre diferentes tipos de rede, como a transição de uma rede 4G LTE para uma rede VoLTE (Voice over LTE) ou alternar entre Wi-Fi e redes celulares.

Baixa latência: O EVS é uma tecnologia que contribui para a redução da latência na transmissão de voz, proporcionando uma experiência de comunicação mais suave e em tempo real em diferentes tipos de redes de comunicação (3GPP, 2019).

Compatibilidade e interoperabilidade: O EVS é compatível com codecs de voz existentes e pode ser implantado em redes que suportam padrões de codec anteriores. Além disso, ele oferece suporte à interoperabilidade com diferentes dispositivos e plataformas, garantindo uma comunicação perfeita entre dispositivos habilitados para EVS e dispositivos não EVS. Isso contribui para uma melhor experiência de comunicação e uma transição suave para a adoção do EVS em redes existentes (3GPP, 2014).

Através do Quadro 3.4 abaixo, é possível conhecer melhor as características do codec de voz *Enhanced Voice Services (EVS)* (VOICEAGE, 2021).

Quadro 3.4 – Características Do Codec EVS

PADRONIZADO	3GPP 2014
Taxas de amostragem	8, 16, 32 e 48 kHz
Modos e largura de banda codificada	NB (20-4000 Hz) WB (20-8000 Hz) SWB (20-16000 Hz) FB (20-20000 Hz)
Modos e largura de banda decodificada	ACELP/MDCT (comutação orientada por conteúdo entre compressão de voz e áudio) ACELP aprimorado estendido com modos especializados para diferentes conteúdos de fala) Codificação baseada em MDCT em diferentes variantes para codificação de áudio Codificação MDCT eficiente com baixo atraso/taxas de bits baixas Comutação perfeita e confiável entre a fala e os núcleos de áudio Extensões de largura de banda otimizadas para diversos conteúdos
Taxa de bits (kbps)	NB: 7,2, 8, 9,6, 13,2, 16,4, 24,4 WB: 7,2, 8, 9,6, 13,2, 16,4, 24,4, 32, 48, 64, 96, 128 SWB: 9,6, 13,2, 16,4, 24,4, 32, 48, 64, 96, 128 FB: 16,4, 24,4, 32, 48, 64, 96, 128 NB e WB VBR: 5,9 kbps AMR WB Modo Interoperável: 6,6, 8,85, 12,65, 14,25, 15,85, 18,25, 19,85, 23,05, 23,85
Atraso: tamanho do quadro à frente	20 ms 12 ms (atraso extra do codificador e do decodificador)
Qualidade	NB: Qualidade melhorada e robustez de perda de pacotes em AMR-NB WB: Qualidade melhorada e robustez de perda de pacotes em AMR-WB SWB: Qualidade SWB sem precedentes em taxas de bits móveis (a partir de 9,6 kbps) Conteúdo misto e música: Qualidade significativamente melhorada em AMR- WB
Gerenciamento de buffer de jitter	Incluído
AD-DTX-CNG	Incluído

Fonte: Adaptado de: Bruhn, Eksler e Gibbs (2016)

4 TRABALHOS RELACIONADOS

Neste capítulo, são apresentados alguns trabalhos importantes que merecem destaque. São trabalhos científicos complementares usados para comparação.

De acordo com (GOUVEA et al., 2018) a análise da qualidade de fala pode ser realizada por meio de métodos subjetivos e objetivos; os métodos subjetivos são frequentemente utilizados na avaliação da qualidade da voz, mas têm muitas limitações, como um número reduzido de participantes, o custo e o tempo necessários para a realização de testes. No entanto, é importante observar que os métodos objetivos também têm suas próprias limitações. Eles podem não capturar completamente as nuances da percepção humana da qualidade da fala e podem não levar em consideração fatores contextuais e individuais que afetam a percepção da qualidade da voz. Portanto, é comum que os métodos objetivos sejam usados em combinação com métodos subjetivos para obter uma avaliação mais completa e precisa da qualidade da fala.

Em (Rodríguez, et. al, 2009) o emprego de codificadores multitaxa nas comunicações de voz possibilita a implementação de algoritmos de determinação da taxa (RDA - *Rate Determination Algorithm*) que controlam a comutação entre taxas de codificação baseados em diferentes fatores de decisão como as características intrínsecas do sinal de voz a ser transmitido ou empregando a informação dos parâmetros de rede. O algoritmo de determinação da taxa (RDA) é responsável por selecionar a taxa de codificação adequada para cada situação específica. Ele pode levar em consideração várias informações para tomar essa decisão, incluindo as características intrínsecas do sinal de voz, como nível de ruído, presença de fala ou música, complexidade espectral, entre outros. Os testes de avaliação de qualidade de voz podem ser realizados seguindo métodos objetivos, em que o índice MOS é dado por software, ou subjetivos, nos quais há intervenção de indivíduos no resultado da avaliação (??). Em (Magro, et.al, 2005) o E-Model é considerado complexo, pois envolve uma série de cálculos e fatores a serem considerados. Ele é amplamente utilizado na indústria de telecomunicações para avaliar a qualidade de chamadas de voz em redes, ajudando a garantir uma experiência de chamada satisfatória para os usuários.

A qualidade da voz é influenciada por vários fatores e a percepção do usuário final é fundamental para determinar a satisfação com o serviço prestado. Embora seja possível medir alguns desses fatores objetivamente, como a taxa de bits, a largura de banda ou a clareza da transmissão, a qualidade da voz é, em última análise, uma experiência subjetiva (SOUSA, 2011).

O método proposto por Ribeiro et al. (2021) compara graficamente as diferenças entre o comportamento MOS e BER para uma amostra de áudio em 3 tipos de canais de desvanecimento: AWGN, Rician e Rayleigh.

Nesse trabalho o autor evidenciou que é possível determinar a qualidade do sinal de voz utilizando as características no domínio do tempo do mesmo sinal e técnicas de aprendizado de máquina.(JÚNIOR; CARLOS, 2019).

Neste estudo, o autor discute o problema da avaliação objetiva de qualidade de codecs de voz na faixa de telefonia (300-3400 Hz). São apresentados os procedimentos para as avaliações subjetivas e objetivas de qualidade e a modelagem matemática do ouvido humano. O método PSQM, padronizado pela “*International Telecommunication Union*”, é descrito com detalhes e serviu de base para a implementação de um programa de computador que apresenta uma série de recursos adicionais, dando origem a dois novos métodos, denominados Medida Objetiva de Qualidade de Voz 1 e 2 (MOQV1 e MOQV2) (BARBEDO, 2001).

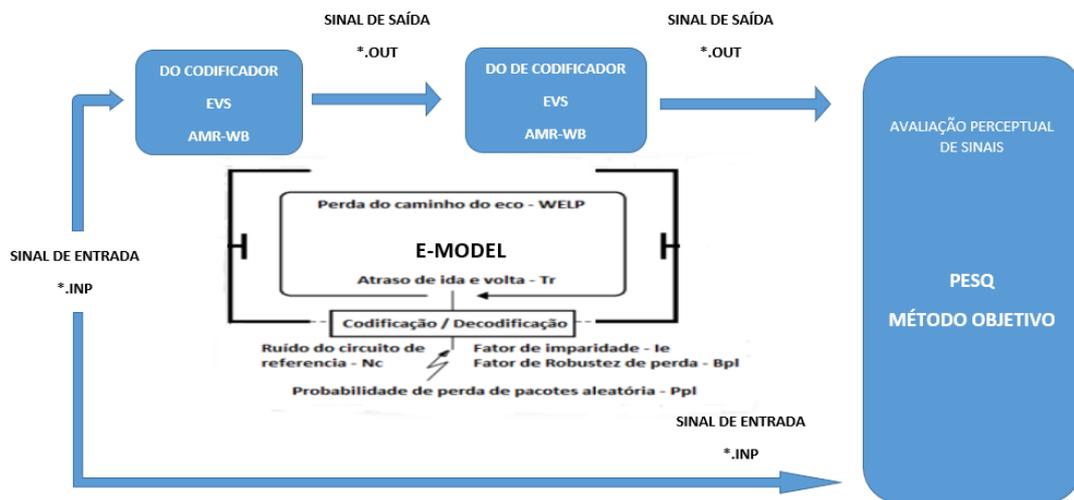
5 METODOLOGIA

Neste capítulo, apresentamos a metodologia utilizada nesta pesquisa. A pesquisa tem como foco analisar e sugerir soluções para aprimorar a qualidade da voz em comunicações em tempo real. Isso é realizado por meio de métodos objetivos de avaliação perceptual da qualidade de sinais de áudio. Para isso, foram utilizados os algoritmos da Rec P.862 da ITU-T e o E-Model G.107 da ITU-T com o objetivo de avaliar a qualidade de voz em uma chamada para um sistema de comunicação que usa o codec EVS.

Os detalhes dos testes aplicados, e os resultados obtidos e examinados serão mostrados posteriormente.

Na Figura 5.1, é possível observar como será a execução desses testes com o codec EVS.

Figura 5.1 – Fluxograma de Execução Teste do EVS



Fonte: Do Autor (2023)

A Figura 5.1, mostra os passos seguidos para avaliação do codec EVS realizado nesta pesquisa. Os sinais de entrada serão codificados e decodificados pelo codec EVS, respectivamente, para as avaliações objetivas utilizando o algoritmo da Rec. ITU-T P.862, que, como apresentado no referencial teórico, necessita de uma arquivo de entrada e outro de saída, para a avaliação do MOS, assim, utilizamos os arquivo (.INP) baixado diretamente da TS 26.443 que possui as especificações técnica desenvolvida pelo *Third Generation Partnership Project (3GPP)*. A codificação e decodificação do EVS é um processo que envolve várias etapas, incluindo pré-processamento, análise espectral, modelagem de fonte e filtro, quantização, codificação de entropia e pós-processamento. O processo de execução do EVS, para a codificação e decodificação, são apresentados na subseções abaixo.

Já na avaliação do codec EVS usando o E-Model, é necessário realizar testes específicos em condições controladas para determinar a pontuação de qualidade esperada em diferentes cenários. Esses testes podem ser realizados em laboratórios especializados ou por meio de simulações de rede avançadas.

5.1 Descrição das Ferramentas Usadas nos Experimentos

Os testes desta pesquisa foram realizados em uma máquina com sistema operacional Windows, na versão 11 de 64 Bits, com 8 GB de memória RAM, processador Intel i5 de décima geração e 250 GB de HD SSD, onde foram executados os algoritmos para a codificação e decodificação do codec EVS, e os algoritmos Rec. ITU-T P.862 e E-model para os testes de avaliação da qualidade da voz.

A versão 14.6.0/15.4.0/16.3.0 do codificador EVS em ponto flutuante foi a escolhida para esse processo, sendo uma das mais recentes, disponibilizada em (3GPP, 2022), utiliza o código ANSI-C, que divide a execução em dois programas, ou seja, uma para codificar e outro para decodificar. No processo de codificação, o sinal de entrada é um arquivo de áudio, com extensão (.INP), disponibilizado pela (3GPP, 2022), que será codificado nas seguintes frequências 8, 16, 32 ou 48 KHz e decodificado gerando uma saída em um áudio com extensão (.OUT). Depois será executado o algoritmo da rec. P.862, para se chegar ao MOS, onde, o sinal de entrada é o (.INP) e a saída o (.OUT).

Ao empregar a ferramenta de cálculo do E-model é possível fazer cálculos de modelo E online e obter o valor R e MOS como resultado. O E-model utiliza seus vastos conhecimentos e algoritmos sofisticados para realizar cálculos matemáticos, resolver equações, realizar análises estatísticas e muito mais (SW, 2020).

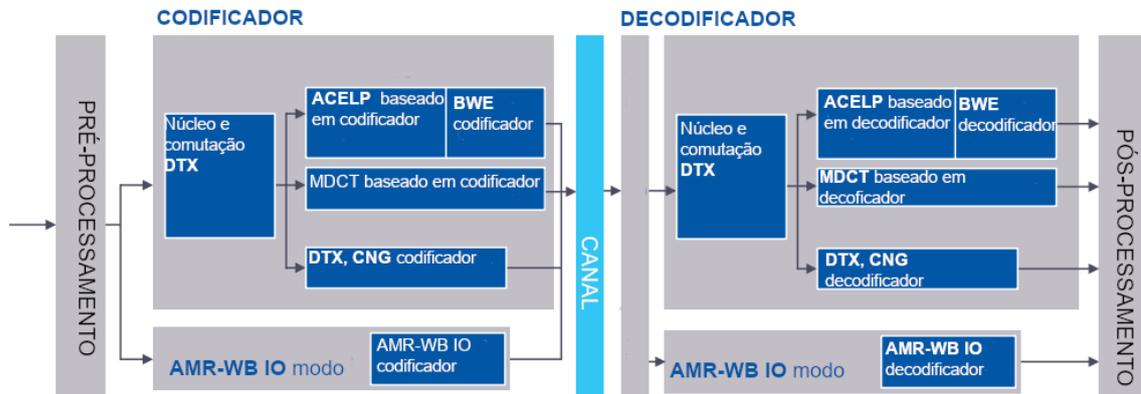
5.1.1 Execução do Codec EVS

Como já mencionado, o EVS é um codificador de áudio de fala, para Full Band¹, que foi desenvolvido para redes VoLTE², sendo apontado como um *codec* muito avançado e que possui vários recursos, tornando-se uma excelente escolha para a próxima geração de telefonia móvel. Os testes em laboratório envolvem a utilização de ferramentas de análise objetiva e subjetiva, como o PESQ (*Perceptual Evaluation of Speech Quality*) e o MOS (*Mean Opinion Score*), para avaliar a qualidade do áudio codificado pelo EVS em diferentes condições de sinal e taxas de bits. Apresentamos na Figura 5.2 um fluxograma de alto nível que mostra o processo de codificação e decodificação do EVS. O processo real envolve algoritmos e técnicas avançadas de processamento de sinal para produzir uma qualidade de som superior.

¹ Banda Completa

² Voice over Long-Term Evolution é um padrão de comunicação sem fio LTE de alta velocidade para telefones celulares e terminais de dados, incluindo dispositivos de Internet das coisas e wearables.

Figura 5.2 – Fluxograma de Auto Nível de Codificação e Decodificação do EVS



Fonte: Adaptada de Jordan (2016)

Breve explicação do fluxograma de alto nível do codec EVS:

- a) Entrada do sinal de áudio;
- b) Pré-processamento para preparar o sinal (codec EVS);
- c) Encoder - codificação do codec com taxa de bits e parâmetros DTX, transmissão;
- d) Decoder - decodificação com taxa de bits e parâmetro;
- e) pré-processamento sinal de chegada do (codec EVS);
- f) Saída do sinal.

O processo de codificação e decodificação utilizou os parâmetros apresentados no quadro 5.1 deste trabalho. Foram feitos alguns testes para medir o desempenho do codec EVS em diferentes taxas de bits, tanto em seu modo nativo quanto em taxas de bits do AMR-WB, já que é necessário para o modo interoperável com AMR-WB e também através do E-model.

Através do uso da plataforma Windows, descrita na seção 5.1 desse trabalho, o programa funcionará da seguinte forma: O código do EVS será processado no modo de administrador do computador, realizando o comando por meio do prompt de comando CMD.

O processo ou instrução para codificação é:

EVS_cod.exe [opções de codificação] R Fs <nome do arquivo de entrada> <nome do arquivo de saída codificado>

Para alcançar o objetivo desejado, é necessário verificar as opções de codificação registradas no quadro 5.1 abaixo, onde estão descritas as opções e seus valores.

Quadro 5.1 – Descrição e Valores dos Elementos da Declaração de Codificação.

PARÂMETRO	DESCRIÇÃO
R	Taxa de bits (bps)
	Para modos nativos EVS: R = (5900*, 7200, 8000, 9600, 13200, 16400, 24400, 32000, 48000, 64000, 96000, 128000). *Modo VBR (taxa de bits média). Para modos nativos de EVS e AMR-WB: R = (6600, 8850, 12650, 14250, 15850, 18250, 19850, 23050, 23850)
F	Taxa de amostragem de entrada (KHz). Fs = (8, 16, 32 ou 48)
Opções	
-q	Modo silencioso não conta quadros
-dtx D	Ativa o modo DTX, D = (0, 3-100) é a taxa de atualização de Silence Insertion Descriptor (SID) ³ , quando colocamos 0 é adaptável, enquanto 3-100 corrige um número de quadros. Por padrão está desabilitado.
-dtx	Ativa o modo Discontinuous Transmission (DTX) ⁴ , com uma taxa de atualização SID de 8 quadros.
-rf po	Ativa o modo “reconhecimento de canal” para sinais WB e SWB em 13,2 Kbps, em que p: LO ou HI indicador Forward Error Correction (FEC) ⁵ , e o: 2, 3, 5 ou 7 em número de quadros (deslocamento FEC). Alternativamente, p e o podem ser substituídos por um arquivo de configuração rf contendo o valor de p e o separados por um espaço. Por padrão está desabilitado.
-max_band B	-max_band B Ativar limitação de largura de banda
-no_delay_cmp	Desativa a compensação de atraso.
-mimo	Torna o formato do arquivo de saída mime. (TS26.445)

Fonte: 3GPP (2022)

O método utilizado na decodificação é o seguinte:

EVS_dec.exe [opções de decodificação] Fs <nome do arquivo de saída codificado> <nome do arquivo de saída>. No caso específico do nome do arquivo de saída codificado, refere-se ao nome do arquivo criptografado anteriormente ou o nome do pacote RTP se o modo VOIP estiver ativo. Os parâmetros e opções obrigatórios restantes estão descritos no quadro 5.2. abaixo:

Quadro 5.2 – Descrição e Valores dos Elementos da Declaração de Decodificação.

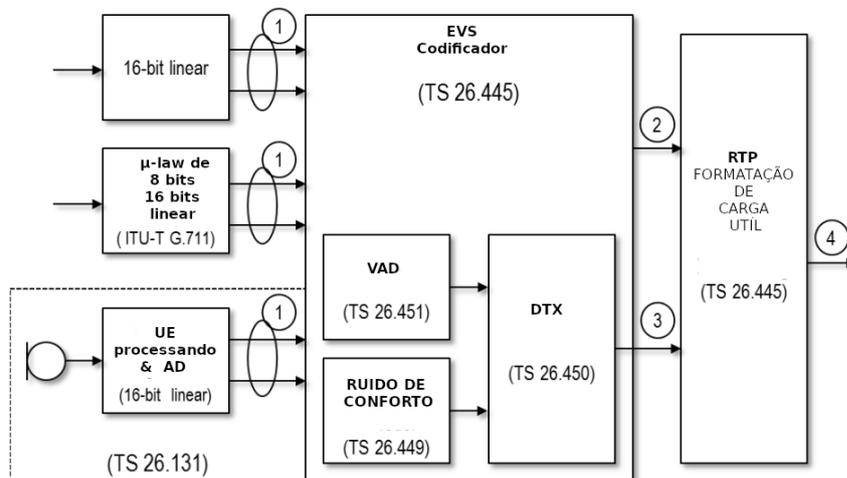
PARÂMETRO	DESCRIÇÃO
F	Taxa de amostragem de entrada (KHz). $F_s = (8, 16, 32 \text{ ou } 48)$
Funções	
-VOIP	Ative o modo VOIP.
Tracefile TF	Gera um arquivo de rastreamento chamado TF.
-no_delay_cmp	Desativar compensação de atraso.
-fec_cfg_file	Ótimo “reconhecimento de canal” com configuração calculada pelo gerenciamento de buffer de jitter (JBM) ⁶ , especificado na TS26.448. A saída é descrita em um arquivo.txt com o indicador FEC (HI/LO) e o deslocamento FEC ideal.
-mimo	Quando o sinal foi codificado com esta opção e o arquivo de entrada do decodificado tem este formato.
-q	Modo silencioso; não conta quadros. Por padrão está desabilitado.

Fonte: 3GPP (2022)

5.2 Funções de Áudio

Os transmissores digitais são utilizados para enviar sinais de áudio de alta qualidade e podem ser empregados para transmitir diversos canais de áudio simultaneamente. Eles trabalham convertendo o sinal de áudio em dados digitais e transmitindo esses dados por meio de ondas de rádio. Nas Figuras 5.3 e 5.4, é possível ver as funções de áudio do lado transmissor e do lado receptor, incluindo a conversão de analógico para digital e digital para analógico. Partes estas que estão incluídas, para mostrar o caminho de fala completo entre a entrada/saída de áudio no Equipamento do Usuário (UE) e a interface digital do rede (3GPP, 2017).

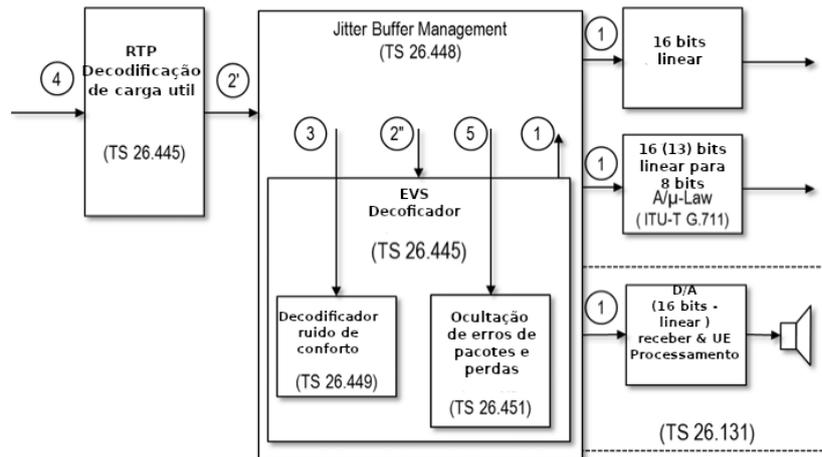
Figura 5.3 – Funções de áudio - Transmissor



Fonte: Adaptado de Qualcomm Technologies (2014)

- 1 - Amostras lineares de 16 bits PCM e taxa de amostragem (8, 16, 32 e 48 KHz).
- 2 - Quadro de áudio codificado, 50 quadros, número de bits/quadro dependendo do modo de codec EVS.
- 3 - Quadros descritores de silêncio codificados (taxa de quadros variável).
- 4 - Pacotes de carga rtp

Figura 5.4 – Funções do áudio - Receptor



Fonte: Adaptado de: Qualcomm Technologies (2014)

- 1 - Amostras PCM lineares de 16 bits e taxa de amostragem (8, 16, 32 e 48 KHz).
- 2 - Quadro de áudio recebido prejudicado, normalmente 50 quadros número de bits/quadro número de bits/quadro dependendo do modo de codec EVS.
- 3 - Quadros descritores de silêncio codificados (taxa de quadros variável).
- 4 - Pacotes de carga rtp
- 5 - Indicação de quadro ruim

De acordo com as figuras 5.3 e 5.4 o codificador de áudio recebe sua entrada, e produz uma saída no decodificador, na forma de um sinal suave de 16 bits *Pulse Code Modulated* (PCM) nas taxas de amostragem de 8 kHz, 16 kHz, 32 kHz ou 48 kHz. O áudio pode se originar e terminar no lado de áudio do UE (3GPP, 2022).

5.3 Propriedade do Arquivo

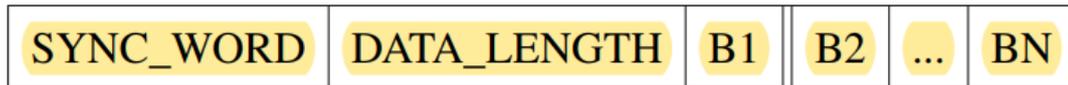
Os arquivos de fala analisados pelo codificador e gravados pelo decodificador são constituídos de uma amostra de fala/áudio de palavras de 16 bits. A ordem dos bytes vai depender da arquitetura do host. Tanto o codificador e o decodificador processam quadros completos (3GPP, 2017).

O codificador ocupará totalmente o último quadro para múltiplos inteiros de quadros de 20ms, ou seja, n quadros de fala serão produzidos a partir de um arquivo de entrada com um comprimento entre

$[(n-1)*20\text{ms}+1 \text{ amostra}; n*20\text{ms}]$. Os arquivos produzidos pelo decodificador terão sempre uma duração de $n*20\text{ms}$ (3GPP, 2017).

De acordo com ITU-T G.192, os arquivos produzidos pelo codificador de fala/áudio, esperados pelo decodificador de voz, no quadro 5.3 contêm um número arbitrário de quadros nos seguintes formatos disponíveis (G.192, 1996).

Quadro 5.3 – Quadros de Fala



Fonte: Adaptado de ITU-TP.800 (1996)

Cada célula do quadro corresponde a um valor Word16 no arquivo bitstream, para um total de $2 + nn$ palavras ou $4 + 2nn$ bytes por quadro, em que nn corresponde ao número de bits codificados. Na trama, cada bit codificado é representado da seguinte forma: Bit 0 = 0x007f, Bit 1 = 0x0081. O significado dos seguintes campos são: A SYNC_WORD do codificador é sempre 0x6b21. Se o decodificador receber SYNC_WORD como 0x6b20 indica que o frame atual foi recebido com erro (bad frame). Bons quadros (0x6b21) e quadros ruins (0x6b20) (3GPP, 2017).

DATA_LENGTH: Comprimento dos dados de voz. O parâmetro DataLen fornece o número de bits de dados de áudio no quadro. Por exemplo, usando DTX, DataLen para quadros NO_DATA é zero. Cada bit é apresentado da seguinte forma: Bit 0 = 0x007f, Bit 1 = 0x0081 (3GPP, 2017).

- O quadro 5.4 abaixo mostra o arquivo de fluxo de bits do parâmetro VoIP (entrada do decodificador).

Quadro 5.4 – Fluxo de Bits



Fonte: Adaptado de ITU-TP.800 (1996)

- Os campos tem o seguinte tamanho e significado;
 - 1 - TAMANHO DO PACOTE: inteiro sem sinal de 32 bits. ($= 12 + 2 + \text{DATA_LENGTH}$).
 - 2 - HORA DE CHEGADA: inteiro sem sinal de 32 bits em ms.
 - 3 - RTP HEADER RTP cabeçalho: 96 bits (ver RFC 3550), incluindo o Tempo RTP e (Synchronization Source Identifier (SSRC) ⁷).
 - 4 - Perfil de comutação de largura de banda (entrada do codificador).

⁷ É um número de identificação de 32 bits que é gerado pelo remetente de um fluxo de dados RTP para identificar exclusivamente a origem dos pacotes RTP. Esse número é incluído em cada pacote RTP enviado pelo remetente e é usado pelo receptor para identificar os pacotes que pertencem a esse fluxo de dados específico.

O arquivo é um arquivo de texto em que cada linha contém "nb_frames B". B especifica o largura de banda do sinal que é uma das quatro larguras de banda suportadas, ou seja, NB, WB, SWB ou FB e "nb_frames" é um número inteiro de quadros e caracteriza a duração da ativação da largura de banda B do sinal acompanhado (3GPP, 2017).

Arquivo de configuração "*Aware channel*" (entrada do codificador e saída do decodificador). O programa codificador pode opcionalmente ler em um arquivo de configuração que especifica os valores do sinalizador FEC ⁸ p e o deslocamento FEC o, em que o sinalizador FEC é denotado como p: LO ou HI e FEC offset, ou: 2, 3, 5 ou 7 em número de quadros. Cada linha do arquivo de configuração contém os valores de p e o separados por um espaço (RODRÍGUEZ, 2018).

Já o arquivo de configuração com reconhecimento de canal faz simulação do feedback do canal de um receptor, para um remetente, melhor dizendo, o decodificador geraria uma indicação FEC e valores de compensação FEC para feedback do receptor que correspondem às características atuais do canal de transmissão, autorizando assim a otimização da transmissão pelo codificador que aplica o offset FEC e a indicação FEC quando está no modo de detecção de canal (RODRÍGUEZ, 2018).

- Arquivo de rastreamento JBM (saída do decodificador).

Um arquivo de rastreamento JBM gerado pelo codificador com a opção -Tracefile como subproduto da operação do decodificador no caso da operação JBM (que acionado pela chave -VOIP no lado do decodificador (RODRÍGUEZ, 2018).

O arquivo trace começa com uma linha de cabeçalho, contendo os nomes das colunas em um rastreamento de arquivo CSV, com um ponto e vírgula como separador na seguinte ordem (RODRÍGUEZ, 2018).

rtpSeqNo;rtpTs;rcvTime;playtime;ativo. Segue a descrição dos termos acima:

- a) "rtpSeqNo" refere-se ao número de sequência RTP, um campo no cabeçalho RTP que ajuda a garantir a ordem correta e a integridade dos pacotes em um fluxo de áudio ou vídeo em tempo real;
- b) rtpTs: isso pode se referir ao carimbo de data/hora RTP, que é um valor usado no protocolo de transporte em tempo real (RTP) para carimbar pacotes de áudio e vídeo. O timestamp RTP fornece informações de tempo para sincronização e reprodução;
- c) rcvTime: Pode representar a hora de recebimento, indicando a hora em que um pacote foi recebido. Pode ser útil para medir latência ou calcular atrasos de rede;
- d) playtime: esse valor provavelmente representa o tempo de reprodução ou o tempo de reprodução, indicando a duração da reprodução de um arquivo de mídia ou fluxo. Pode ser medido em segundos ou outra unidade de tempo apropriado;

⁸ A técnica denominada FEC (Forward Error Correction), permite , através de alterações no sinal digital que está sendo transmitido, melhorias na performance sistêmica.

- e) ativo: este termo pode se referir ao status ou estado de algo, como se um media player ou dispositivo está ativo ou inativo.

Uma entrada é gravada no arquivo de rastreamento para cada quadro de fala produzido. O intervalo de valores de tempo de reprodução normalmente é de 20ms, mas pode ser diferente dependendo da operação do JBM. Cada entrada é uma linha no arquivo de rastreamento que contém valores conforme especificado (3GPP, 2020a).

- Formato de arquivo de teste

Os arquivos de sequência de teste na ordem de bytes do PC (little-endian) são fornecidos em um arquivo de (formato ZIP) em (3GPP, 2020b)

Depois da descompressão são fornecidos três tipos de arquivos:

- Arquivos para entrada do codificador de fala: *.INP;
- Arquivos para comparação com a saída do codificador e para entrada do decodificador de fala: *.COD;
- Arquivos para comparação com a saída do decodificador: *.OUT;
- Arquivos para entrada no decodificador de fala com JBM: *.RTP
- Arquivo de controle de um modo para teste de comutação de modo *.MOD;
- Instruções de como operar as sequências de teste: *.TXT.

5.3.1 Visão Geral do Formato de Áudio AMR

AMR-WB é um codec de áudio digital de alta qualidade muito usado em condições de rede variáveis e sua alta robustez contra erros de transmissão, graças à sua operação e adaptação multitaxa. Ao mesmo tempo oferece qualidade de voz superior aos padrões estreitos existentes (BESSETTE et al., 2002).

No Quadro 5.5, é possível identificar em quais taxas de bits trabalha o AMR-WB:

Quadro 5.5 – Taxas de bits do codec AMR-WB

Modos	Velocidade
AMR-WB_23.85	23,85 kbit/s
AMR-WB_23.05	23,05 kbit/s
AMR-WB_19.85	19,85 kbit/s
AMR-WB_18.25	18,25 kbit/s
AMR-WB_15.85	15,85 kbit/s
AMR-WB_14.25	14,25 kbit/s
AMR-WB_12.65	12,65 kbit/s
AMR-WB_8.85	8,85 kbit/s
AMR-WB_6.60	6,60 kbit/s
AMR-WB_SID	1,75 bits/s*

Fonte: Bruhn, Eksler e Gibbs (2016)

Dois programas são utilizados para o teste, o codificador e o decodificador, que são exibidos nas Figuras 5.5 e 5.6 abaixo. Para obter ajuda do codec, basta executar uma síntese sem argumentos de entrada, seja no codificador ou no decodificador (3GPP, 2019).

Na figura 5.5, ilustra-se a interface gráfica do codificador:

Figura 5.5 – Interface Gráfica do Codificador

```

Usage : coder (-dtx) (-itu | -mime) mode speech_file bitstream_file

Format for speech_file:
  Speech is read from a binary file of 16 bits data.

Format for bitstream_file (default):
  One word (2-byte) to indicate type of frame type.
  One word (2-byte) to indicate frame type.
  One word (2-byte) to indicate mode.
  N words (2-byte) containing N bits (bit 0 = 0xff81, bit 1 = 0x007f).

  if option -itu defined:
  One word (2-byte) for sync word (0x6b21)
  One word (2-byte) for frame length N.
  N words (2-byte) containing N bits (bit 0 = 0x007f, bit 1 = 0x0081).

  if option -mime defined:
  AMR-WB MIME/storage format, see RFC 3267 (sections 5.1 and 5.3) for details.

mode: 0 to 8 (9 bits rates) or
      -modefile filename
=====
mode  : (0) (1) (2) (3) (4) (5) (6) (7) (8)
bitrate: 6.60 8.85 12.65 14.25 15.85 18.25 19.85 23.05 23.85 kbit/s
=====

-dtx if DTX is ON, default is OFF

C:\amr>

```

Fonte: Do Autor (2023)

Figura 5.6 – Interface gráfica do decodificador

```

C:\amr>amrwb_dec.exe

=====
AMR Wideband Codec 3GPP TS26.190 / ITU-T G.722.2, Aug 25, 2003. Version 7.0.0.
=====

Usage : decoder (-itu | -mime) bitstream_file synth_file

Format for bitstream_file (default):
  One word (2-byte) to indicate type of frame type.
  One word (2-byte) to indicate frame type.
  One word (2-byte) to indicate mode.
  N words (2-byte) containing N bits (bit 0 = 0xff81, bit 1 = 0x007f).

  if option -itu defined:
  One word (2-byte) for sync word (good frames: 0x6b21, bad frames: 0x6b20)
  One word (2-byte) for frame length N.
  N words (2-byte) containing N bits (bit 0 = 0x007f, bit 1 = 0x0081).

  if option -mime defined:
  AMR-WB MIME/storage format, see RFC 3267 (sections 5.1 and 5.3) for details.

Format for synth_file:
  Synthesis is written to a binary file of 16 bits data.

C:\amr>

```

Fonte: Do Autor (2023)

5.3.2 Características do Arquivo

Os arquivos de fala lidos pelo codificador e gravados pelo decodificador consistem em palavras de 16 bits, cada uma contendo uma amostra de fala de 14 bits alinhada à esquerda. A ordem dos bytes

origina-se da arquitetura do *host*. Já à ordem dos bytes, existem duas convenções principais: a ordem *little-endian* e a ordem *big-endian*⁹. Na ordem *little-endian*¹⁰, os bytes menos significativos são armazenados primeiro, enquanto na ordem *big-endian*, os bytes mais significativos aparecem primeiro (G.192, 1996).

5.3.3 Formato de Arquivo Teste

Os arquivos de sequência de teste em ordem de byte PC *little-endian* são fornecidos em arquivos formato ZIP que acompanham o presente documento (3GPP, 2022). Após a descompressão, são fornecidos três tipos de arquivo:

- a) Arquivos para entrada no codificador de fala: *.INP
- b) Arquivos para comparação com a saída do codificador e para entrada no decodificador de fala: *.COD
- c) Arquivos para comparação com a saída do decodificador: *.OUT
- d) Arquivo de controle de um modo para o teste de mudança de modo T22.MOD

Cada arquivo *.INP inclui dois quadros de retorno, conforme (3GPP, 2019), no início da sequência de teste. A função desses *frames* é redefinida como variáveis do codificador de fala para seu valor inicial. No caso de uma instalação correta do Simulação ANSI C 26.173 em (3GPP, 2019), todos os quadros de saída de sinal de voz são mantidos ao quadro que corresponde ao arquivo *.COD. No caso de uma implementação de hardware correta na primeira saída de fala *frame*, não precisa de ser implementada ao primeiro *frame* no *.COD, mas todos os codificadores de fala e os de saída devem ser orientados aos quadros que corresponde no *.COD.

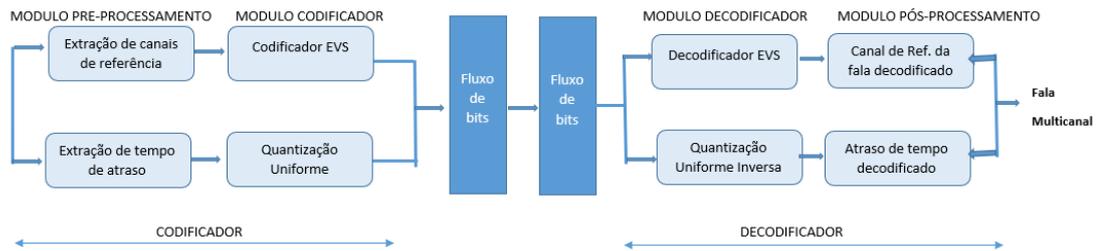
5.4 Cenário - 1

Nesse teste verificamos a capacidade do codec EVS em operar nas mesmas taxas de bits do codec AMR-WB. É importante lembrar que, para garantir a interoperabilidade, o codec EVS deve operar os modos suportados pelo codec AMR-WB, além dos modos de trabalho nativos do próprio EVS. Com base na figura 5.9 abaixo temos um diagrama de blocos com características daquilo que apresentamos na realidade.

⁹ Big-endian é uma ordem de byte ou uma maneira de organizar tipos de dados multi-byte, como inteiros e números de ponto flutuante na memória do computador. Em um sistema *big-endian*, o byte mais significativo (MSB) de um valor *multibyte* é armazenado no endereço de memória mais baixo, enquanto o byte menos significativo (LSB) é armazenado no endereço de memória mais alto

¹⁰ Little-endian é um termo usado para descrever a ordem de bytes usada em sistemas de computador para armazenar tipos de dados *multibyte*, como inteiros e números de ponto flutuante, na memória. Em um sistema *little-endian*, o byte menos significativo (LSB) de um valor *multibyte* é armazenado no endereço de memória mais baixo, enquanto o byte mais significativo (MSB) é armazenado no endereço de memória mais alto.

Figura 5.7 – Diagrama de blocos do Multi EVS-TDOA



Fonte: Cheng (2020)

Os arquivos de entrada usados estão disponíveis em formato .ZIP de (3GPP, 2020b). Estão divididos em: arquivos para o modo compatível com AMR-WB nas seguintes taxas $R = (6600, 8850, 12650, 14250, 15850, 18250, 19850, 23050, 23850)$ bps cuja informações são descritas em (3GPP, 2022), além de arquivos NB, WB, SWB e FB. Ao realizar este teste, consideramos somente os arquivos do modo interoperável, com conteúdo variado, incluindo vozes femininas, masculinas e infantis, além de ruídos de fundo adicionados em diferentes níveis. Os arquivos (.INP) são primeiro criptografados. Para habilitar o modo interoperável no codec EVS, basta codificar nas taxas de bits em que o codec AMR opera, que também estão disponíveis no codec EVS. Verifique no quadro 5.1 o parâmetro (R) referente a taxa de bits. Neste caso a frequência de amostragem será apenas 16kHz, pois os resultados serão comparados com os do codec AMR-WB funciona nesta taxa de amostragem e a ferramenta usada para avaliação de qualidade perceptual (PESQ) tem essa frequência como limite superior. A síntese abaixo mostra como codificar tais amostras:

EVS_cod.exe 6600 16 T00.INP T00_6600_16.COD

EVS_cod.exe: Nome do programa de software que está sendo executado, 6600 é taxa de bits usada no teste, 16 é frequência utilizada, T00.INP Arquivo de Entrada, T00_6600_16.COD: Este é o nome de um arquivo de saída que o programa irá gerar como resultado de seu processamento como mostra a figura 5.8, ilustrando a amostra de codificação:

Figura 5.8 – Codificação do T00.INP no modo AMR e EVS no modo interoperável.

```
C:\EVS_COD\testvec\bin>EVS_cod.exe 6600 16 T00.INP T00_6600_16.COD

=====
EVS Codec (Floating Point) 3GPP TS26.443 Nov 13, 2018. Version 12.11.0 / 13.7.0 / 14.3.0 / 15.1.0
=====

Input audio file:      T00.INP
Output bitstream file: T00_6600_16.COD

Input sampling rate:   16000 Hz
Bitrate:               6.60 kbps

----- Running the encoder -----
Frames processed:      91
Encoding finished

C:\EVS_COD\testvec\bin>
```

Fonte: Do Autor (2023)

Prosseguindo com o teste, o sinal será amostrado em diferentes modos:

EVS_cod.exe 8850 16 T00.INP T00_8850_16.COD

EVS_cod.exe 12650 16 T00.INP T00_12650_16.COD

Repetir o teste com todas as taxas de bits até: 23.850 kbps

EVS_cod.exe 23850 16 T00.INP T00_23850_16.COD

Em seguida, é necessário processar essas amostras restantes com a mesma taxa de amostragem de 16 kHz e diferentes modos de compressão, mudando o nome do sinal de entrada e saída para evitar sobrescrever o arquivo criptografado anteriormente.

EVS_cod.exe 6600 16 T01.INP T01_6600_16.COD

Seguindo o procedimento anteriormente mencionado:

EVS_cod.exe 23850 16 T22.INP T22_23850_16.COD

Utilizamos a instrução mencionada anteriormente para descriptografar os arquivos criptografados:

EVS_dec.exe 16 T00_6600_16.COD T00_6600_16.OUT

A Figura 5.9 apresenta apenas a frequência de amostragem ($F_s=16\text{KHz}$) e o nome do arquivo de entrada, que, neste caso, é o mesmo do arquivo de saída do codificador (T00_6600_16.COD) e o nome da saída do codificador, são passados como parâmetros obrigatórios do arquivo de saída decodificador (T00_6600_16.OUT). A Figura 5.9, ilustra a amostra de decodificação:

Figura 5.9 – Amostra decodificação T00.INP nas taxas de bits do modo AMR-WB e EVS interoperável.

```
C:\codigo_EVS\testvec\bin>EVS_dec.exe 16 T00_6600_16.COD T00_6600_16.OUT
=====
EVS Codec (Floating Point) 3GPP TS26.443 November 04, 2021.
Version 12.14.0 / 13.10.0 / 14.6.0 / 15.4.0 / 16.3.0
=====

Input bitstream file:  T00_6600_16.COD
Output synthesis file: T00_6600_16.OUT

Output sampling rate:  16000 Hz
Bitrate:               6.60 kbps

----- Running the decoder -----

Frames processed:      91

Decoding finished

C:\codigo_EVS\testvec\bin>
```

Fonte: Do Autor (2023)

Para as outras amostras será usado o mesmo procedimento para as taxas na decodificação.

EVS_dec.exe 16 T00_8500_16.COD T00_8500_16.OUT

EVS_dec.exe 16 T22_23850_16.COD T22_23850_16.OUT

Para realizar a avaliação perceptiva do modo objetivo, o algoritmo PESQ Rec. ITU-T P.862, vai comparar os arquivos de áudio originais e degradados, através de um código executável via linha de comando por comandos MSDOS, e assim a pontuação MOS será atribuído. Usando a instrução de help PESQ, é exibido um menu que indica as opções e modos de operação desta ferramenta, conforme mostrado na figura 5.10, PESQ:

Figura 5.10 – Menu da Ferramenta PESQ

```
EACH COMPANY OFFERS OEM LICENSE AGREEMENTS, WHICH COMBINE OEM
IMPLEMENTATIONS OF THE PESQ ALGORITHM TOGETHER WITH A PESQ PATENT LICENSE
AGREEMENT. PESQ PATENT-ONLY LICENSE AGREEMENTS MAY BE OBTAINED FROM OPTICOM.
*****
* OPTICOM GmbH                               * Psytechnics Limited *
* Naegelsbachstr. 38,                       * Fraser House, 23 Museum Street, *
* D- 91052 Erlangen, Germany                * Ipswich IP1 1HN, England *
* Phone: +49 (0) 9131 53020 0                * Phone: +44 (0) 1473 261 880 *
* Fax: +49 (0) 9131 53020 20                * Fax: +44 (0) 1473 261 880 *
* E-mail: info@opticom.de,                  * E-mail: info@psytechnics.com, *
* www.opticom.de                             * www.psytechnics.com *
*****

Usage:
PESQ HELP          Displays this text
PESQ [options] ref deg
Run model on reference ref and degraded deg

Options: +8000 +16000 +swap +wb
Sample rate - No default. Must select either +8000 or +16000.
Swap byte order - machine native format by default. Select +swap for byteswap.
Default mode of operation is P.862 (narrowband handset listening). Select +wb
to use P.862.2 wideband extension (headphone listening).

File names may not begin with a + character.

Files with names ending .wav or .WAV are assumed to have a 44-byte header. All other file types are assumed to have no header.

C:\codigo_EVS\testvec\bin>
```

Fonte: Do Autor (2023)

Através da síntese abaixo compara-se o sinal de referência (T00.INP) com o sinal degradado (T00_6600_16.OUT) que é ilustrado por meio da figura:

Figura 5.11 – Avaliação objetiva do sinal original T00.INP com o sinal degradado T00_6600_16.OUT usando PESQ.

```

WILL REQUIRE A PESQ LICENCE AGREEMENT, WHICH MAY BE OBTAINED FROM EITHER
OPTICOM GMBH OR PSYTECHNICS LIMITED.

EACH COMPANY OFFERS OEM LICENSE AGREEMENTS, WHICH COMBINE OEM
IMPLEMENTATIONS OF THE PESQ ALGORITHM TOGETHER WITH A PESQ PATENT LICENSE
AGREEMENT. PESQ PATENT-ONLY LICENSE AGREEMENTS MAY BE OBTAINED FROM OPTICOM.

*****
* OPTICOM GmbH * Psytechnics Limited *
* Naegelsbachstr. 38, * Fraser House, 23 Museum Street, *
* D- 91052 Erlangen, Germany * Ipswich IP1 1HN, England *
* Phone: +49 (0) 9131 53020 0 * Phone: +44 (0) 1473 261 800 *
* Fax: +49 (0) 9131 53020 20 * Fax: +44 (0) 1473 261 880 *
* E-mail: info@opticom.de, * E-mail: info@psytechnics.com, *
* www.opticom.de * www.psytechnics.com *
*****

Reading reference file T00.INP...done.
Reading degraded file T00_6600_16.OUT...done.
Level normalization...
IRS filtering...
Variable delay compensation...
Acoustic model processing...

P.862.2 Prediction (MOS-LQ0): = 3.142

C:\codigo_EVS\testvec\bin>_

```

Fonte: Do Autor 2023

O resultado do teste pode ser visto através do MOS (MOS-LQ = 3.142) além do mais o PESQ gera um arquivo *.txt onde coleta todos os dados do MOS obtidos com várias informações, como os nomes dos sinais de referência e degradados, taxa de amostragem e modo usado. Essa ação deverá ser feita com os demais sinais restantes até obter o MOS dos sinais analisados. Seguindo o mesmo roteiro de testes em seguida faremos também este procedimento para os modos nativos EVS: R = (5900*, 7200, 8000, 9600, 13200, 16400, 24400, 32000, 48000, 64000, 96000, 128000 bps) apresentados no quadro 5.1 codificando e decodificando até que todos os valores MOS correspondentes sejam computados.

A razão para não ter usado a taxa de bits de 5900 é porque a taxa de bits de 5,9 kbps (quilobits por segundo) é bastante baixa e pode não ser adequada para muitas aplicações de áudio. O codec EVS é projetado para fornecer qualidade de áudio de alta definição em uma ampla gama de taxas de bits. Embora seja possível usar uma taxa de bits tão baixa quanto 5,9 kbps com o codec EVS, essa taxa resultaria em uma qualidade de áudio extremamente baixa e com perdas significativas.

Partindo da síntese:

EVS_cod.exe 7200 16 T00.INP T00_7200_16.COD

EVS_cod.exe 128000 16 T22.INP T22_128000_16.COD

EVS_dec.exe 16 T00_7200_16.COD T00_7200_16.OUT

EVS_dec.exe 16 T22_128000_16.COD T22_128000_16.OUT

Pesq.exe +16000 +wb T00.INP T00_7200_16.OUT

Pesq.exe +16000 +wb T22.INP T22_128000_16.OUT

O Pesq.exe é uma avaliação perceptiva da ferramenta de qualidade de fala usada para medir a qualidade dos sinais de fala. O comando:

Pesq.exe +16000 +wb T00.INP T00_7200_16.OUT

é uma instrução de linha de comando para executar Pesq.exe com determinados parâmetros e arquivos de entrada. Aqui está o que cada parte do comando significa:

"**Pesq.exe**" é um nome comum para um arquivo executável, geralmente usado para identificar um programa que será executado em sistemas Windows. A extensão ".exe" indica que se trata de um arquivo executável. "**+16000**": Este é um parâmetro que especifica a frequência de amostragem dos arquivos de entrada e saída. Neste caso, é definido para 16 kHz (16000 Hz). "**+wb**": Este é um parâmetro que especifica o tipo de codec usado nos arquivos de entrada e saída. Nesse caso, ele é definido como "banda larga", o que significa que o codec foi projetado para sinais de fala com largura de banda de 50 a 7.000 Hz. "T00.INP": Este é o nome do arquivo de entrada que contém o sinal de fala original. "T00_7200_16.OUT": Este é o nome do arquivo de saída que contém o sinal de fala processado.

5.5 Cenário - 2

Neste cenário o trabalho está concentrado no estudo do fator de degradação do codificador conforme definido na recomendação ITU-T G.107, que é mais conhecida como E-model.

5.5.1 E-MODEL AVALIAÇÃO

É importante deixar claro que o E-Model é uma ferramenta matemática e, embora possa fornecer uma estimativa da qualidade da chamada, nem sempre reflete totalmente a percepção subjetiva dos usuários.

Ao utilizar o *E-model calculator* para o codec EVS e AMR-WB, é preciso inserir os valores dos parâmetros I_e ¹¹, B_{pl} ¹² e P_{pl} ¹³ nas caixas de entrada correspondentes e realizar o cálculo para obter uma análise da qualidade da voz transmitida pelo codec como mostra a figura 5.12 abaixo: (Rec. ITU-T, 1996)

¹¹ Fator de deterioração do equipamento

¹² Fator de Robustez de Perda de Pacote

¹³ Probabilidade de Perda de Pacote

Figura 5.12 – E-model Calculator

E-Model (versão de março de 2005)				
Parâmetro	EU IA	Padrão	Valor	Dimensão
Ruído do Circuito Elétrico	Nc	(-70)	-70	dBm0p
piso de ruído	Nfor	(-64)	-64	dBmp
Ruído da Sala (Enviar)	PS	(35)	35	dB(A)
Ruído da Sala (Receber)	Pr	(35)	35	dB(A)
Enviar classificação de volume	SLR	(8)	8	dB
Receber classificação de sonoridade	RLR	(2)	2	dB
Classificação de mascaramento de tom lateral	STMR	(15)	15	dB
Fator D (Receber)	Dr	(3)	3	
Avaliação do Sidetone do Ouvinte	LSTR	STMR+Dr	18	dB
Fator D (Enviar)	Ds	(3)	3	
Atraso médio de sentido único	T	(0)	0	EM
Atraso absoluto de (S) para (R)	ta	(=T)	0	EM
Atraso de ida e volta	Tr	(=2T)	0	EM
Avaliação do volume do eco do locutor	TELR	(65)	65	dB
Perda de caminho de eco ponderada	WEPL	(110)	110	dB
Quantizando Unidades de Distorção	qdu	(1)	1	
Fator de deterioração do equipamento	ou seja	(0)	0	
Fator de Robustez de Perda de Pacote	Bpl	(1)	1	
Probabilidade de Perda de Pacote	peleas	(0)	0	%
taxa de explosão	Burst R	(1)	1	
Fator de Vantagem	A	(0)	0	
Resultados				
Fator R calculado	R	93.2	<input type="button" value="calcular"/>	
Pontuação de opinião média	MOS CQEn	4.41	<input type="button" value="reiniciar"/>	
	Opções		<input type="button" value="Opções"/>	T=Ta=Tr/2
Este front-end da web E-Model foi trazido a você por x-fabric.com				
Última atualização: 2008-06-02				

Fonte: Adaptado de ITU-T, G.107 (2015)

Para avaliar o codec EVS usando o E-Model, é preciso ter os seguintes dados: Através da Recomendação ITU-T G.107 — É possível obter os valores padrão e intervalos permitidos para os parâmetros para uso com o cálculo do fator R do modelo E (ITU-T, G.107, 2015).

Os valores padrão para todos os parâmetros de entrada usados no algoritmo do modelo E estão listados no quadro 5.4 logo abaixo: Recomenda-se fortemente utilizar os valores padrão para todos os parâmetros que permanecem constantes durante o cálculo de planejamento. Caso todos os parâmetros sejam definidos com esses valores, o resultado do cálculo será de alta qualidade com um fator de classificação de $R = 93,2$.

Quadro 5.6 – Recomendação ITU-T G.107 - Valores padrão e intervalos permitidos para os parâmetros para uso com o cálculo do fator R do modelo E

Nome do parâmetro	Parâmetro	Padrão	Un	Permitido
Enviar classificação de sonoridade	SLR	8	dB	0 a 18
Receber classificação de sonoridade	RLR	2	dB	-5 a 14
Classificação de máscara de tom lateral	STMR	15	dB	10 a 20
Avaliação do tom do lado do ouvinte	LSTR	18	dB	13 a 23
D – valor do telefone no lado de envio	Ds	3	—	-3 a 3
D – valor do telefone no lado de recebimento	Dr.	3	—	-3 a 3
Avaliação do volume do eco do locutor	TELR	65	dB	5 a 65
Perda de caminho de eco ponderada	WEPL	110	dB	5 a 110
Atraso médio unidirecional do caminho do eco	T	0	MS	0 a 500
Atraso de ida e volta	Tr	0	MS	0 a 1000
Atraso absoluto no eco – grátis conexão	T	0	MS	0 a 500
Número de quantização unidades de distorção	qdu	1	—	1 a 14
Fator de deterioração do equipamento	Ie	0	—	0 a 40
Fator de robustez de perda de pacotes	bpl	1	—	1 a 40
Probabilidade de perda aleatória de pacotes	ppl	0	%	0 a 20
taxa de explosão	Burst R	1	—	1 a 2
Ruído do circuito referido a 0-dBr apontar	Nc	-70	dBm0p	-80 a -40
Piso de ruído no lado de recepção	Nfor	-64	dBmp	-64
Ruído da sala no lado de envio	PS	35	dB(A)	35 a 85
Ruído da sala no lado de recepção	Pr	35	dB(A)	35 a 85
Fator de vantagem	A	0	—	0 a 20

Fonte: G107 (2007)

O E-Model é uma ferramenta utilizada em tecnologias de telecomunicações para medir objetivamente a qualidade de uma conversa telefônica. "Para fazer isso, são coletados diferentes parâmetros relacionados à comunicação, que combinados nas diferentes fórmulas dos algoritmos do modelo E, fornecem um valor numérico que representa a qualidade da comunicação. Esse valor é chamado Fator de taxa de transmissão, **R** (G107, 2007).

Sendo um modelo objetivo, desde que sejam introduzidos os mesmos parâmetros para fazer os cálculos, o resultado será o mesmo. Ao contrário do que poderia acontecer em um modelo subjetivo, que depende da opinião de um usuário. É por isso que é usado por muitos prestadores de serviços de telecomunicações, porque oferece um valor de qualidade preciso (G107, 2007).

Como os equipamentos da rede degradam a qualidade de voz, o valor R pode ser calculado através da equação, $R = R0 - Is - Id - Ie + A$ onde:

- Is, referente à soma de todas as degradações de sinal que ocorrem simultaneamente;
- Id, é o atraso fim-a-fim introduzido a chamada;
- Ie, trata-se da degradação inserida pelo próprio equipamento utilizado;
- A, se refere ao fator de vantagem, por exemplo, uma medida que os usuários aceitam uma qualidade menor devido as facilidades oferecidas.

O valor de R também pode ser relacionado com a classificação do modelo MOS para obter o nível de satisfação dos usuários conforme demonstra a figura 5.1 abaixo:

Quadro 5.7 – Recomendação ITU-T G.107 - Valores padrão e intervalos permitidos para os parâmetros para uso com o cálculo do fator R do modelo E

Fator de Classificação	Satisfação do Usuário
Maior que 90	Usuários muito satisfeitos
Entre 80 e 90	Entre 90 e 80
Entre 70 e 80	Alguns usuários não estão satisfeitos
Entre 60 e 70	Muitos usuários não satisfeitos
Entre 50 e 60	Quase nenhum usuário está satisfeito
Menor que 50	A qualidade é muito baixa e não é aceitável

Fonte: G.107 (2015)

Com esses dados, você pode montar um cenário ideal para avaliar o codec EVS usando o E-Model, que é um modelo matemático que calcula a qualidade da transmissão de voz em uma rede de comunicação. O modelo leva em consideração vários fatores, incluindo o atraso da rede, a largura de banda disponível e as perdas de pacotes.

Para usar o E-Model, é preciso ter as seguintes informações:

Nível de ruído de fundo (N) - Este é o nível de ruído de fundo em dB.

Relação sinal-ruído (SNR) - Este é o nível de sinal em dB menos o nível de ruído de fundo em dB.

Atraso de ida e volta (RTD) - Este é o tempo que leva para um pacote de voz ir de um ponto para outro na rede e voltar.

Fator de perda de pacotes (PLF) - Este é o número de pacotes de voz perdidos durante a transmissão.

Com essas informações, você pode calcular o MOS (*Mean Opinion Score*), que é uma medida da qualidade de áudio percebida pelo ouvinte. O MOS é calculado usando a fórmula do E-Model:

$$MOS = R - I - A - PL + S$$

onde:

R = nível de sinal de referência em dB' I = perda de pacotes interpolada em dB' A = atraso de ida e volta em milissegundos (ms) PL = perda de pacotes em dB S = fator de escala

Para um cenário ideal, você pode configurar a largura de banda disponível e o atraso de ida e volta de acordo com as especificações do codec EVS. Você também pode definir o nível de ruído de fundo e a relação sinal-ruído para testar o codec em diferentes condições de ruído.

Com todas essas informações, você pode avaliar a qualidade do codec EVS usando o E-Model e comparar seus resultados com outros codecs para determinar qual é o melhor codec para sua aplicação.

Ao comparar os resultados com o modelo *E-model* da ITU-T G.107, é possível avaliar a qualidade da experiência do usuário em diferentes condições de sinal, em cada uma das taxas de bits em que o codec EVS trabalha, e com algumas das opções que ele oferece (G107, 2007). Além disso, a avaliação objetiva do codec EVS é particularmente importante, alguns estudos anteriores foram baseados apenas em avaliações subjetivas (G107, 2007). Neste trabalho procurou as informações mais atualizadas sobre os valores disponíveis dos fatores de comprometimento do equipamento de banda completa, I_e, f_b , e fatores de robustez de perda de pacotes, B_{pl}, f_b , e fatores de robustez de rajada, B_{rf}, f_b . Destina-se a ser atualizado regularmente. Esses valores devem ser usados em uma escala estendida de classificação de transmissão (escala R), conforme definido em ITU-T G.107.2. Deve-se observar que os valores I_e, eff, f_b referem-se a um codec em uma taxa de bits especificada. No caso de codecs de taxa de bits variável, deve-se especificar se o valor correspondente foi obtido com uma configuração de taxa de bits constante ou variável, pois isso pode influenciar o comprometimento em condições de erro de transmissão (G107, 2007).

Além dos parâmetros básicos de codificação, como taxa de bits e taxa de amostragem, o codec AMR-WB possui dois parâmetros adicionais que podem ser ajustados para otimizar o desempenho do codec com base nas características do sinal de áudio. Esses parâmetros são:

I_e (Enhancement bit rate): É um parâmetro que controla a quantidade de bits adicionais usados para codificar informações de aprimoramento do sinal. Essas informações podem incluir a codificação de elementos como ruído de fundo, reverberação e supressão de eco. O ajuste do parâmetro I_e permite equilibrar a qualidade do sinal e a eficiência da taxa de bits (G107, 2007).

B_{pl} (Burst Packet Loss): É um parâmetro que permite o controle do comportamento do codec em relação à perda de pacotes. Pacotes perdidos podem ocorrer em redes de comunicação com perda de pacotes, como a internet. O parâmetro B_{pl} influencia o comportamento do codec ao lidar com a perda de

pacotes, adaptando a estratégia de codificação para minimizar os efeitos perceptíveis da perda de dados (G107, 2007).

5.6 Correlação de Pearson

O teste de correlação de Pearson é uma medida que avalia a relação estatística entre duas variáveis contínuas. Essa medida estatística indica o grau de relacionamento linear entre duas variáveis. Nesse caso, ela é empregada para verificar se as variações nos parâmetros técnicos da rede estão correlacionadas com as variações na qualidade da voz percebida (BERNARDINO, 2019).

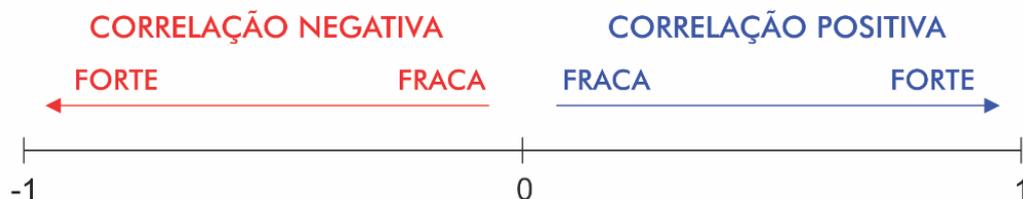
O coeficiente de Pearson pode variar de +1 a -1 e indicar a presença ou ausência de associação entre duas variáveis. Quando o valor é 0, não há associação. Quando é maior que 0, há uma associação positiva, ou seja, o valor de uma variável aumenta à medida que a outra também aumenta (BERNARDINO, 2019).

Para realizar a correlação de Pearson, é preciso seguir as seguintes condições:

- a) A escala de medição deve ser uma escala ou relação de intervalo;
- b) As variáveis devem ser aproximadamente distribuídas;
- c) A associação deve ser linear;
- d) Não deve haver valores atípicos nos dados.

O coeficiente de correlação de Pearson (PCC), também chamado de correlação linear ou r de Pearson, é um grau de relação entre duas variáveis quantitativas e exprime o grau de correlação através de valores situados entre -1 e 1, como mostrado na figura 5.13 abaixo:

Figura 5.13 – Correlação de Pearson



Fonte: Adaptado de Oliveira (2019)

Quando o coeficiente de correlação se aproxima de 1, percebe-se que o valor de uma variável aumenta quando a outra também aumenta, ou seja, há uma relação linear positiva. Quando o coeficiente se aproxima de -1, também é possível dizer que as variáveis estão relacionadas de forma linear negativa (OLIVEIRA, 2019).

6 RESULTADOS

O teste foi realizado com dois arquivos, sendo eles, o TO4.INP contendo voz feminina e ruído de ambiente. Enquanto T19.INP contém voz de criança com ruído ambiente também, conforme mencionado na metodologia codificados e decodificados com o EVS em ponto flutuante no modo interoperável para AMR-WB e no modo nativo do EVS. Já avaliação (para planejamento) da qualidade de áudio é realizada pelo G.107 E-Model, que leva em consideração diversos fatores, como atraso, perda de pacotes e degradação da qualidade de som.

6.1 Valores obtidos através da Rec. ITU-T P.862

Os valores apresentados nesta subseção, buscou, avaliar a capacidade do codec em transmissão de voz no modo de interoperabilidade, ou seja, com taxas de bits do AMR-WB e no modo de transmissão nativa. Os arquivos usados na realização do teste foram baixados da TS 26.443 da 3GPP e estão dentro de uma pasta (.ZIP). Assim, estão divididos em arquivos no modo compatível com AMR e também nos outros modos, NB, WB, SWB e FB.

Os arquivos de áudio selecionados para os testes possuem conteúdo misto, incluindo vozes femininas, masculinas e infantis, com diferentes tipos de ruído de fundo, como ruído branco adicionados em vários níveis. Esses áudios passaram pelo processo de codificação, onde foram criptografados primeiro e depois decodificados. O modo de interoperabilidade foi ativado ao codificar os arquivos em taxas de bits equivalentes ao AMR (*Adaptive Multi-Rate*). Ao realizar os testes com esses arquivos de áudio criptografados e codificados com taxas de bits equivalentes ao AMR, estaremos avaliando a capacidade de interoperabilidade desses arquivos em diferentes dispositivos e sistemas compatíveis com o formato AMR.

A frequência de amostragem, como mencionado acima, foi de 16 KHz, e os resultados apresentados, nas tabelas abaixo, possibilitam a comparação do MOS sobre os codecs em modo interoperável com o AMR e no modo nativo. Ambos os áudios também foram codificados ativando o parâmetro (dtx) e (-q) especificados no quadro 5.1 sendo que os áudios codificados com o (-q) também foram decodificados usando o mesmo parâmetro (-q). Nas tabelas seguintes é possível ver os valores obtidos:

Tabela 6.1 – Valores MOSLLQ do arquivo T04.INP modo interoperável com taxa do AMR-WB usando a frequência +16000

DEGRADED	MOSLLQ
T04_6600_16.OUT	2.847
T04_8850_16.OUT	3.358
T04_12650_16.OUT	3.747
T04_14250_16.OUT	3.839
T04_15850_16.OUT	3.911
T04_18250_16.OUT	3.959
T04_19850_16.OUT	3.983
T04_23050_16.OUT	3.994
T04_23850_16.OUT	4.032

Fonte: Do Autor (2023)

Na tabela 6.1, os dados apresentados são referentes ao processo de codificação e decodificação com taxas de bits no modo interoperável, para o áudio T04.INP, executado em modo normal, ou seja, sem a adição de nenhum parâmetro. A Rec. P.862 foi efetuada com o recebimento do arquivo de origem e do arquivo decodificado. Os resultados indicam que quanto maior a taxa de bits, melhor é a pontuação do MOS.

Tabela 6.2 – Valores MOSLLQ do arquivo T04.INP modo interoperável com AMR-WB usando a taxa de bits na frequência de +16000 mais o parâmetro (dtx) ativado

DEGRADED	MOSLLQ
T04_6600_16_dtx.OUT	3,544
T04_8850_16_dtx.OUT	3,881
T04_12650_16_dtx.OUT	4,145
T04_14250_16_dtx.OUT	4,216
T04_15850_16_dtx.OUT	4,239
T04_18250_16_dtx.OUT	4,245
T04_19850_16_dtx.OUT	4,279
T04_23050_16_dtx.OUT	4,275
T04_23850_16_dtx.OUT	4,275

Fonte: Do Autor (2023)

Na tabela 6.2, os dados apresentados são referentes ao processo de codificação e decodificação com taxas de bits no modo interoperável, para o áudio T04.INP, onde o parâmetro (dtx) foi ativado. A avaliação da Rec. ITU-T P.862 foi realizada com o arquivo de origem e o arquivo decodificado. Os resultados, conforme na tabela 6.1, indicam que a pontuação MOS aumenta à medida que a taxa de bits cresce. Outro aspecto notado é a melhora do MOS em relação aos testes realizados de forma normal, por exemplo, na tabela 6.1, o codec codificado e decodificado na taxa de 6600, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits mas com o parâmetro (dtx) ativado, fato esse que se repetiu para as demais taxas.

Quando o dtx está ativo, os valores de MOS são maiores porque ele permite que o codec interrompa a transmissão de voz quando não há atividade sonora detectada. Melhor esclarecendo, quando não há voz sendo falada, o codec não envia nenhum pacote de áudio, economizando assim a largura de banda. Essa interrupção na transmissão é conhecida como 'período de silêncio'. A razão pela qual os valores de MOS são maiores com o DTX ativado é porque ele reduz a quantidade de dados transmitidos durante os períodos de silêncio. Isso resulta em uma utilização mais eficiente da largura de banda disponível, permitindo que mais informações relevantes sejam transmitidas durante os períodos de atividade sonora.

Já com o DTX desabilitado, o codec continua transmitindo pacotes de áudio mesmo quando não há voz sendo falada. Isso leva a uma utilização desnecessária da largura de banda, pois pacotes de áudio adicionais estão sendo enviados sem contribuir com informações relevantes. Isso pode afetar a qualidade da transmissão de áudio e, conseqüentemente, reduzir o valor do MOS.

Tabela 6.3 – Valores MOS do arquivo T04.INP modo interoperável com AMR-WB usando a frequência de +16000 mais o parâmetro (-q) na codificação e decodificação

DEGRADED	MOSLLQ
T04_6600_16_q.OUT	3,536
T04_8850_16_q.OUT	3,886
T04_12650_16_q.OUT	4,144
T04_14250_16_q.OUT	4,197
T04_15850_16_q.OUT	4,235
T04_18250_16_q.OUT	4,249
T04_19850_16_q.OUT	4,278
T04_23050_16_q.OUT	4,276

Fonte: Do Autor (2023)

O processo de codificação e decodificação do áudio T04.INP foi realizado com taxas de bits no modo interoperável e utilizando o parâmetro (-q), conforme apresentado na tabela 6.2. Ao receber o arquivo de origem e o arquivo decodificado, a Rec. P.862 obteve resultados que podem ser vistos nas tabelas 6.1 e 6.2. É possível observar que a pontuação MOS melhora conforme a taxa de bits aumenta. Outro ponto observado, refere-se a melhora do MOS em relação aos teste realizados de forma simples, assim como no exemplo mencionado para a tabela 6.2, na tabela 6.1, o codec codificado e decodificado na taxa de 6600, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits utilizando o parâmetro (-q), para a codificação e decodificação, fato esse que se repetiu para as demais taxas.

O parâmetro -q, mencionado no contexto de avaliação de qualidade de voz, refere-se à utilização de um algoritmo de cancelamento de ruído ou supressão de silêncio durante a medição da qualidade de voz. Ao aplicar o parâmetro -q, o algoritmo de cancelamento de ruído ou supressão de silêncio é ativado durante a medição do MOS. Isso pode resultar em valores de MOS mais altos em comparação com a medição sem a utilização desse parâmetro. Existem motivos pelas quais isso pode acontecer:

Otimização do ruído de fundo: O algoritmo de cancelamento de ruído é desenvolvido para diminuir o ruído indesejado presente em uma gravação de voz. Ao reduzir o ruído, a voz se torna mais nítida e mais fácil de compreender, o que pode aprimorar a percepção.

Ocultação de silêncio: O algoritmo de supressão de silêncio identifica momentos de silêncio na gravação e os exclui. Isso pode resultar em uma gravação mais curta, com menos momentos de silêncio. Como os momentos de silêncio não contribuem para a qualidade da gravação, sua exclusão pode melhorar a experiência do usuário.

Tabela 6.4 – Valores MOS do modo interoperável com taxa do AMR-WB - T19.INP baseado na frequência +16000

DEGRADED	MOSLLQ
T19_6600_16.OUT	2,401
T19_8850_16.OUT	2,917
T19_12650_16.OUT	3,297
T19_14250_16.OUT	3,382
T19_15850_16.OUT	3,431
T19_18250_16.OUT	3,479
T19_19850_16.OUT	3,443
T19_23050_16.OUT	3,556
T19_23850_16.OUT	3,637

Fonte: Do Autor (2023)

Na tabela 6.4 os resultados, exibidos, são do processo de codificação e decodificação com taxas de bits no modo interoperável, para o áudio T19.INP, usando a frequência +16000 com execução simples, ou seja, sem a adição de parâmetros. Durante a execução da rec. P.862, foram utilizados o arquivo de origem e o arquivo decodificado. Os resultados mostram que quanto maior a taxa de bits melhor é a pontuação MOS, fato este que é condizente com os testes realizados para o áudio T04.INP apresentado na tabela 6.1.

Tabela 6.5 – Valores do MOS no modo interoperável com taxa de AMR-WB e com (dtx) ativado - T19.INP

DEGRADED	MOSLLQ
T19_6600_16_dtx.OUT	3,067
T19_8850_16_dtx.OUT	3,517
T19_12650_16_dtx.OUT	3,855
T19_14250_16_dtx.OUT	3,930
T19_15850_16_dtx.OUT	3,969
T19_18250_16_dtx.OUT	4,039
T19_19850_16_dtx.OUT	3,998
T19_23050_16_dtx.OUT	4,082
T19_23850_16_dtx.OUT	4,082

Fonte: Do Autor (2023)

Na tabela 6.5 os resultados, mostram o processo de codificação e decodificação com taxas de bits no modo interoperável, para o áudio T19.INP, onde o parâmetro (dtx) foi ativado. A execução da Rec. ITU-T P.862, utilizou o arquivo de origem e o arquivo decodificado. Os resultados, assim como na tabela 6.4 e demais tabelas acima, mostram que quanto maior a taxa de bits, melhor é o MOS. É importante ressaltar que houve uma melhora no parâmetro MOS em relação aos testes realizados de forma simples, conforme apresentado na tabela 6.4, em que o codec foi codificado e decodificado na taxa de 6600. De acordo com a tabela 6.2, o MOS para o codec codificado e decodificado na mesma taxa de bits, porém com o parâmetro (dtx) ativado, foi menor em comparação com as outras configurações testadas. Essa tendência foi observada nos testes realizados para o áudio T04.INP.

Tabela 6.6 – Valores do MOS no modo interoperável com taxa do AMR-WB e com (-q) na codificação e decodificação - T19.INP utilizando a frequência +16000

DEGRADED	MOSLLQ
T19_6600_16_q.OUT	3,041
T19_8850_16_q.OUT	3,510
T19_12650_16_q.OUT	3,861
T19_14250_16_q.OUT	3,939
T19_15850_16_q.OUT	3,966
T19_18250_16_q.OUT	4,013
T19_19850_16_q.OUT	4,023
T19_23050_16_q.OUT	4,082
T19_23850_16_q.OUT	4,083

Fonte: Do Autor (2023)

A tabela 6.6 exibe os resultados do processo de codificação e decodificação com taxas de bits no modo interoperável para o áudio T19.INP, usando a frequência +16000 onde o parâmetro (-q) foi utilizado tanto na codificação quanto na decodificação. Os arquivos de origem e decodificado foram utilizados na execução da Rec. ITU-T P.862. Os resultados obtidos, conforme as tabelas 6.5 e de mais acima, demonstram que a pontuação MOS melhora à medida que a taxa de bits aumenta. Outro ponto observado refere-se a melhora do MOS em relação aos teste realizados de forma simples, assim como no exemplo mencionado para a tabela 6.5, na tabela 6.4, o codec codificado e decodificado na taxa de 6600, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits utilizando o (-q), para a codificação e decodificação, fato esse que se repetiu para as demais taxas, e que é condizente com os testes realizados para o áudio T04.INP, apresentados na tabela 6.3.

Tabela 6.7 – Valores MOS modo nativo do codec EVS através do arquivo - T04.INP

DEGRADED	MOSLLQ
T04_7200_16.OUT	3,324
T04_8000_16.OUT	3,489
T04_9600_16.OUT	3,578
T04_13200_16.OUT	3,701
T04_16400_16.OUT	4,074
T04_24400_16.OUT	4,200
T04_32000_16.OUT	3,975
T04_48000_16.OUT	4,467
T04_64000_16.OUT	4,286
T04_96000_16.OUT	4,518
T04_128000_16.OUT	4,493

Fonte: Do Autor (2023)

Na tabela 6.7 os resultados, apresentados, são do processo de codificação e decodificação com taxas de bits no modo nativo do EVS, para o áudio T04.INP, usando a frequência +16000 com execução simples, ou seja, sem a adição de parâmetros. A execução da Rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados mostram que quanto maior a taxa de bits melhor é a pontuação MOS, fato esse que é equivalente ao que aconteceu nos testes com taxa no modo interoperável.

Tabela 6.8 – Valores MOS do modo nativo do codec EVS com (dtx) ativado através do arquivo - T04.INP e sinai degradados

DEGRADED	MOSLLQ
T04_7200_16_dtx.OUT	3,325
T04_8000_16_dtx.OUT	3,492
T04_9600_16_dtx.OUT	3,577
T04_13200_16_dtx.OUT	3,702
T04_16400_16_dtx.OUT	4,069
T04_24400_16_dtx.OUT	4,200
T04_32000_16_dtx.OUT	3,971
T04_48000_16_dtx.OUT	4,467
T04_64000_16_dtx.OUT	4,285
T04_96000_16_dtx.OUT	4,517
T04_128000_16_dtx.OUT	4,493

Fonte: Do Autor (2023)

Na tabela 6.8 os resultados, apresentados são do processo de codificação e decodificação com taxas de bits no modo nativo, para o áudio T04.INP, usando a frequência de +16000 onde o parâmetro (dtx) foi ativado. A execução da rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados, assim como, na tabela 6.7 mostram que quanto maior a taxa de bits melhor é a pontuação MOS, outro ponto observado, refere-se ao MOS em relação aos teste realizados de forma simples, por exemplo, na tabela 6.7, o codec codificado e decodificado na taxa de 7200, obteve MOS próximo ao codec codificado e decodificado na mesma taxa de bits mas com o parâmetro (dtx) ativado, sendo 3.324

na execução simples e 3.325 na execução com o parâmetro (dtx) ativado, assim observa-se que com o parâmetro (dtx) ativado no modo nativo não ha uma diferença na qualidade como apresentado para os testes realizados no modo interoperante apresentado na tabela 6.2.

Tabela 6.9 – Valores MOS do modo nativo com (-q) através do arquivo - T04.INP e sinais degradados

DEGRADED	MOSLLQ
T04_7200_16_q.OUT	3,898
T04_8000_16_q.OUT	4,000
T04_9600_16_q.OUT	3,984
T04_13200_16_q.OUT	4,105
T04_16400_16_q.OUT	4,264
T04_24400_16_q.OUT	4,336
T04_32000_16_q.OUT	4,261
T04_48000_16_q.OUT	4,456
T04_64000_16_q.OUT	4,391
T04_96000_16_q.OUT	4,521
T04_128000_16_q.OUT	4,525

Fonte: Do Autor (2023)

Na tabela 6.9 os resultados, apresentados são do processo de codificação e decodificação com taxas de bits no modo nativo, para o áudio T04.INP, usando a frequência +16000 onde se utilizou o parâmetro (-q), tanto na codificação quanto na decodificação. A execução da Rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados, assim, como, nas tabelas 6.8 e 6.7 mostram que quanto maior a taxa de bits melhor é a pontuação MOS, outro ponto observado, refere-se a melhora do MOS em relação aos teste realizados de forma simples e com o parâmetro (dtx) ativado, assim como no exemplo mencionado para a tabela 6.8 , na tabela 6.7, o codec codificado e decodificado na taxa de 7200, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits utilizando o parâmetro (-q), fato que se repete se compara o teste realizado com o parâmetro (dtx) ativado.

Tabela 6.10 – Valores MOS do modo nativo e sinais degradados

DEGRADED	MOSLLQ
T19_7200_16.OUT	2,690
T19_8000_16.OUT	2,812
T19_9600_16.OUT	2,985
T19_13200_16.OUT	3,197
T19_16400_16.OUT	3,745
T19_24400_16.OUT	3,920
T19_32000_16.OUT	3,346
T19_48000_16.OUT	4,315
T19_64000_16.OUT	4,200
T19_96000_16.OUT	4,472
T19_128000_16.OUT	4,481

Fonte: Do Autor (2023)

Na tabela 6.10 os resultados, apresentados, são do processo de codificação e decodificação com taxas de bits no modo nativo, para o áudio T19.INP, usando uma frequência de +16000 com execução simples, ou seja, sem a adição de parâmetros. A execução do Rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados mostram que quanto maior a taxa de bits melhor é a pontuação MOS, fato esse que equivalente ao que aconteceu, na execução do T04.INP em modo nativo, e nos testes com taxa no modo interoperável.

Tabela 6.11 – Valores MOS do modo dtx nativo através do arquivo T19.INP do codec EVS e sinais degradados

DEGRADED	MOSLLQ
T19_7200_16_dtx.OUT	3,246
T19_8000_16_dtx.OUT	3,269
T19_9600_16_dtx.OUT	3,570
T19_13200_16_dtx.OUT	3,611
T19_16400_16_dtx.OUT	3,998
T19_24400_16_dtx.OUT	4,101
T19_32000_16_dtx.OUT	3,689
T19_48000_16_dtx.OUT	4,286
T19_64000_16_dtx.OUT	4,273
T19_96000_16_dtx.OUT	4,400

Fonte: Do Autor (2023)

Na tabela 6.11 os resultados, apresentados são do processo de codificação e decodificação com taxas de bits no modo nativo, para o áudio T19.INP, usando uma frequência de +16000 onde o parâmetro (dtx) foi ativado. A execução da Rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados, assim como, na tabela 6.10 e demais acima, mostram que quanto maior a taxa de bits melhor é a pontuação MOS, outro ponto observado, refere-se ao MOS em relação aos teste realizados de forma simples, por exemplo, na tabela 6.10, o codec codificado e decodificado na taxa de 7200, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits mas com o parâmetro (dtx) ativado, sendo 2.690 na execução simples e 3.246 na execução com o parâmetro (dtx) ativado, assim observa-se que com o parâmetro (dtx) ativado no modo nativo, diferente do que acontece com os teste do T04.INP apresentados na tabela 6.8, o parâmetro (dtx) apresentou uma melhora em relação ao execução realizada de forma simples, fato este que não se repetiu na taxa de 96000, pois o MOS obtido através da execução simples obteve melhor resultado do que o MOS obtido com o parâmetro (dtx) ativado, sendo, 4.481 para execução simples e 4.400 para execução com o parâmetro (dtx) ativado. Assim, nota-se uma necessidade de realizar mais testes para verificar a ação do parâmetro (dtx) no modo nativo.

Tabela 6.12 – Valores MOS do modo nativo do codec EVS através do arquivo T19.INP COM O -q ativo no EVS

DEGRADED	MOSLLQ
T19_7200_16_q.OUT	3,229
T19_9600_16_q.OUT	3,574
T19_13200_16_q.OUT	3,609
T19_16400_16_q.OUT	4,029
T19_24400_16_q.OUT	4,137
T19_32000_16_q.OUT	3,667
T19_48000_16_q.OUT	4,325
T19_64000_16_q.OUT	4,310
T19_96000_16_q.OUT	4,514
T19_128000_16_q.OUT	4,493

Fonte: Do Autor (2023)

Na tabela 6.12 os resultados, apresentados são do processo de codificação e decodificação com taxas de bits no modo nativo, para o áudio T19.INP, onde se utilizou o parâmetro (-q), tanto na codificação quanto na decodificação. A execução do Rec. ITU-T P.862, recebeu o arquivo de origem e o arquivo decodificado. Os resultados, assim, como, nas tabelas 6.11 e demais acima, mostram que quanto maior a taxa de bits melhor é a pontuação MOS, outro ponto observado, refere-se a melhora do MOS em relação aos teste realizados de forma simples e com o parâmetro (dtx) ativado, assim como no exemplo mencionado para a tabela 6.11 , na tabela 6.10, o codec codificado e decodificado na taxa de 7200, obteve MOS inferior ao codec codificado e decodificado na mesma taxa de bits utilizando o (-q), fato que não se repete se compara o teste realizado com o parâmetro (dtx) ativado para essa taxa de bits, mas as demais taxas codificadas e decodificadas com o (-q) obteve melhor desempenho se comparado com os resultados apresentados na tabela 6.11 com o parâmetro (dtx) ativado, assim acredita-se que, o (-q) possui MOS melhor que com o parâmetro (dtx) ativado, uma vez que nos testes realizado co T04.INP, apresentado na tabela 6.9 o MOS foi superior em relação ao parâmetro (dtx) apresentado na tabela 6.8. Já ao comparar os testes do modo interoperável como os modos nativos, observa-se que o modo nativo possui melhor desempenho uma vez que suas taxas de bits são maiores que as do modo interoperável e assim obtém valores MOS mais altos.

6.2 TESTES E-MODEL

Valores padrão para todos parâmetros utilizados no algoritmo E-model estão listados no quadro 6.1 abaixo. É extremamente recomendável usar esses valores padrão para todos os parâmetros que não variam durante o cálculo do planejamento. Se todos os parâmetros forem definidos com valores padrão, o cálculo resulta em uma qualidade muito elevada com fator de classificação.(ITU-T, 2015).

Quadro 6.1 – Valores pré-estabelecidos e faixa de valores permitidos para o Fator R

Parâmetro	Abbr.	Unidade	Valor padrão	Faixa permitida	Observação
Enviar classificação de intensidade sonora	SLR	dB	+8	0 ... +18	(Nota 1)
Receber classificação de intensidade sonora	RLR	dB	+2	-5 ... +14	(Nota 1)
Classificação de mascaramento de sidetone	STMR	dB	15	10 ... 20	(Notas 2, 4)
Classificação do sidetone do ouvinte	LSTR	dB	18	13 ... 23	(Nota 2)
D-Valor do telefone, lado do envio	Ds	-	3	-3 ... +3	(Nota 2)
Valor D do telefone, lado de recepção	Dr	-	3	-3 ... +3	(Nota 2)
Classificação da intensidade do eco do locutor	TELR	dB	65	5 ... 65	
Perda ponderada do caminho do eco	WEPL	dB	110	5 ... 110	
Atraso médio unidirecional do caminho do eco	T	ms	0	0 ... 500	
Atraso de ida e volta em um loop de 4 fios	Tr	ms	0	0 ... 1000	
Atraso absoluto em conexões sem eco	Ta	ms	0	0 ... 500	
Sensibilidade do atraso	sT	-	1	0.4 ... 1	(Nota 7)
Atraso mínimo perceptível	mT	ms	100	20 ... 150	(Nota 7)
Número de unidades de distorção de quantização	qdu	-	1	1 ... 14	
Fator de redução ao valor recuperável de equipamentos	Ie	-	0	0 ... 40	(Nota 5)
Fator de robustez de perda de pacotes	Bpl	-	4.3	4.3 ... 40	(Notas 3 e 5)
Probabilidade aleatória de perda de pacotes	Ppl	%	0	0 ... 20	(Notas 3 e 5)

Fonte: Adaptado de G.107 (2015)

Quadro 6.2 – Valores padrão e intervalos permitidos para os parâmetros

Parâmetro	Abbr.	Unidade	Valor padrão	Faixa permitida	Observação
Ruído ambiente no lado de recepção	Pr	dB(A)	35	35 ... 85	
Fator de vantagem	A	-	0	0 ... 20	
<p>NOTA 1 - Valores totais entre o microfone ou o receptor e o ponto 0 dB.</p> <p>NOTA 2 - Relação fixa: LSTR = STMR + D.</p> <p>NOTA 3 - Atualmente em estudo.</p> <p>NOTA 4 - A Equação 7-24 também fornece previsões para STMR > 20 dB. No entanto, esses valores dificilmente podem ser medidos de forma confiável porque o dispositivo de medição cobrirá principalmente o acoplamento acústico, e não o elétrico.</p> <p>OBSERVAÇÃO 5 - Se Ppl > 0%, então o Bpl deve corresponder ao codec, ao tamanho do pacote e à ocultação de perda de pacotes (PLC) assumido.</p> <p>OBSERVAÇÃO 6 - As previsões do modelo E para valores de BurstR > 2 só são válidas se a porcentagem de perda de pacotes for Ppl < 2%.</p> <p>NOTA 7 - Somente configurações predefinidas são permitidas para sT e mT, refletindo casos de uso específicos, consulte o corpo principal da recomendação.</p>					

Fonte: Adaptado de G.107 (2015)

6.2.1 Avaliação da Probabilidade de Perdas de Pacotes

Para avaliar e projetar sistemas de comunicação com qualidade e confiabilidade, é necessário considerar aspectos como o comprometimento do equipamento, a robustez na perda de pacotes e a probabilidade de perda aleatória, que podem influenciar seu desempenho. Com base nos dados apresentados, as taxas de Probabilidade de Perda de Pacotes empregadas neste trabalho serão as seguintes: 0%, 1%, 2%, 3%, 5%, 10%, 20% iniciando com: Cálculo do valor de R Banda Larga (WB) com ppl iniciando a 0.0% esses valores foram tirados de (G., 2019)

Para o caso de banda estreita (NB), a escala de taxa de transmissão varia de R = 0 (qualidade mais baixa possível) a R = 100 (qualidade ótima). Nesta escala, um canal de transmissão NB padrão incluindo codificação logarítmica de Modulação de Código de Pulso (PCM) e um piso de ruído valores de parâmetro padrão de acordo com a figura 5.15, obtém uma classificação de R = 93,2. Para um canal de transmissão de fala WB, a qualidade geralmente é julgada melhor do que para um canal NB. Assim, este intervalo de escala foi estendida em aproximadamente 29% para refletir a melhoria da qualidade ao migrar de NB para WB. Esta escala R estendida é uma escala R "universal" e é aplicável para os canais de transmissão NB e WB (G.107, 2015)

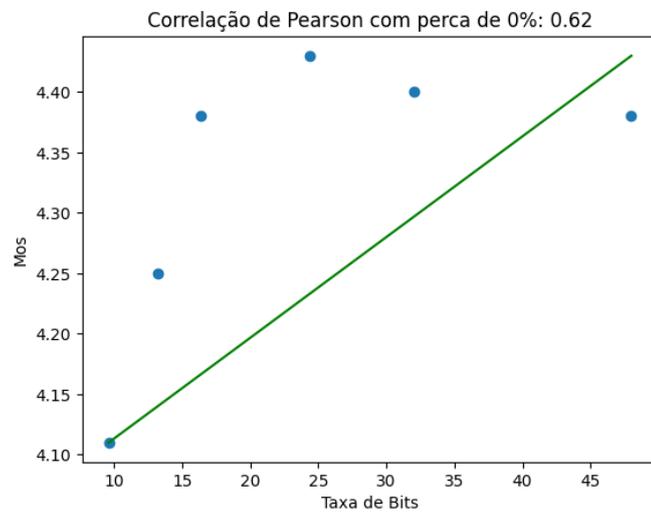
Tabela 6.13 – Cálculo do valor de R e MOS Banda Larga (WB) variando o ppl em 0%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOSCQEwb
1	9.6	13.0	22.7	106.1	4.11
2	13.2	11.7	17.1	111.7	4.25
3	16.4	10.8	10.3	118.5	4.38
4	24.4	11.4	7.2	121.6	4.43
5	32.0	8.7	9.3	119.5	4.40
6	48.0	9.6	10.2	118.6	4.38

Fonte: Do autor (2023)

Com a Probabilidade de Perda de Pacote (ppl) mantido em 0%, é possível perceber que os valores de bpl (Fator de Robustez de Perda de Pacote) e lewb (Fator de deterioração do equipamento) aumentam à medida que a taxa de bits diminui, seguindo uma ordem crescente. À medida que a taxa de bits aumenta, os valores de bpl diminuem, enquanto o lewb cai inicialmente, mas depois volta a subir. De acordo com a taxa de bits, o fator R e MOS pode aumentar. Isso indica que uma taxa mais alta nem sempre resulta em melhor qualidade. Seguindo as diretrizes da recomendação ITU-T G.113 atualizada em 2019, realizamos uma pesquisa sobre esses valores que variam no caso do lewb e bpl.

Gráfico 6.1 – ppl %0



Fonte: Do Autor (2023)

Utilizando a linguagem Python, é possível construir um gráfico que representa os dados da tabela 6.13 por meio do coeficiente de correlação de Pearson. Essa medida estatística é utilizada para avaliar o grau de associação entre duas variáveis contínuas. O coeficiente de correlação varia entre -1 e 1, onde -1 representa uma correlação negativa perfeita, 1 representa uma correlação positiva perfeita e 0 representa a ausência de correlação. Ao iniciar a análise do índice MOS, variamos a Probabilidade de Perda aleatória de pacotes a partir de 0%. O valor obtido para o coeficiente de correlação de Pearson foi de 0.62, o que

indica uma correlação positiva próxima de +1. No entanto, essa correlação não é considerada perfeita, uma vez que os pontos no gráfico que representam o MOS no eixo x estão distantes do eixo y.

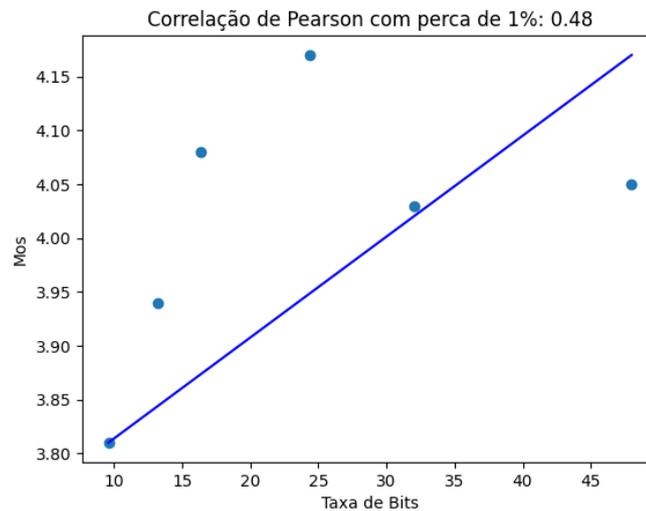
Tabela 6.14 – Cálculo do valor de R Banda Larga (WB) variando em ppl 1.0 %

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	96.5	3.81
2	13.2	11.7	17.1	104.4	3.94
3	16.4	10.8	10.3	105.3	4.08
4	24.4	11.4	7.2	108.5	4.17
5	32.0	8.7	9.3	103.5	4.03
6	48.0	9.6	10.2	104.0	4.05

Fonte: Do autor (2023)

Ao variar a taxa de bits em 1.0%, observa-se que o bpl mantém seu valor mais alto. Em contrapartida, o lewb inicialmente permanece elevado, porém diminui à medida que a taxa de bits aumenta. Embora seja comum que taxas mais altas resultem em uma diminuição do fator R e MOS, é importante ressaltar que em alguns casos a taxa de bits mais alta não causa essa diminuição, embora seja mais comum ocorrer em taxas mais baixas, que apresentam MOS mais baixo.

Gráfico 6.2 – ppl %1



Fonte: Do Autor (2023)

Ao representar os dados da tabela 6.14 em um gráfico utilizando o mesmo procedimento anterior, foi obtido um coeficiente de correlação de Pearson de 0.48. Essa correlação é considerada positiva, pois está mais próxima de +1. No entanto, não é uma correlação positiva perfeita, pois os valores do índice MOS estão distantes dos valores da taxa de bits no eixo y.

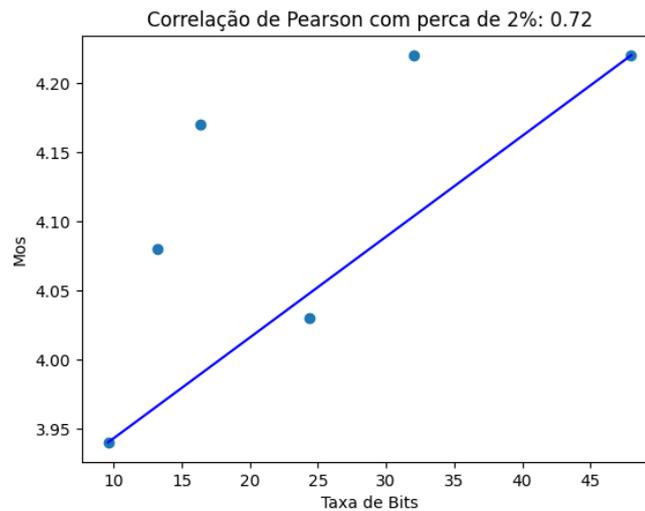
Tabela 6.15 – Cálculo do valor de R Banda Larga (WB) variando em ppl 2%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	100.4	3.94
2	13.2	11.7	17.1	105.3	4.08
3	16.4	10.8	10.3	105.5	4.17
4	24.4	11.4	7.2	103.5	4.03
5	32.0	8.7	9.3	110.7	4.22
6	48.0	9.6	10.2	110.6	4.22

Fonte: Do autor (2023)

Ao variar o Ppl (*Packet-loss Probability*) em 2.0%, pode-se observar que o parâmetro Bpl (*Packet-loss Robustness Factor*) diminui à medida que a taxa de bits aumenta, mas depois volta a aumentar. Por outro lado, o Lewb (*Equipment Impairment Factor*) mantém as mesmas características do Bpl. Já o fator R (*Reliability*) e o MOS (*Mean Opinion Score*) aumentam à medida que o Ppl aumenta, juntamente com a taxa de bits.

Gráfico 6.3 – ppl%2



Fonte: Do Autor (2023)

Aplicando o mesmo critério de avaliação dos modos anteriores, o coeficiente de correlação de Pearson agora é de 0.72, o que indica uma correlação positiva próxima de perfeita. No entanto, os valores de MOS no eixo x ainda estão distantes dos valores de taxa de bits no eixo y.

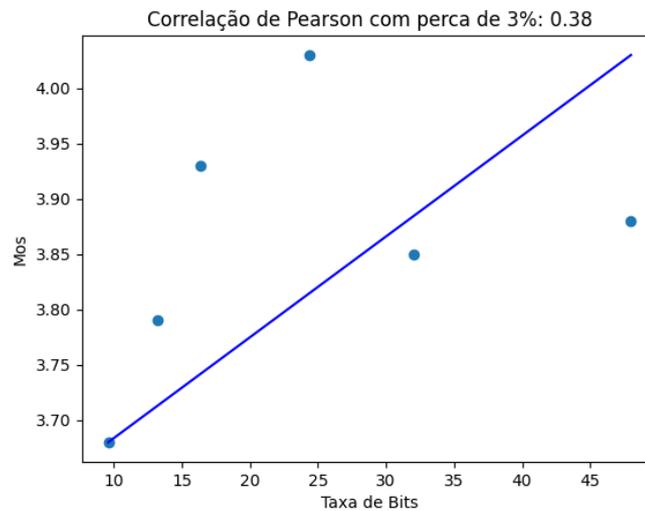
Tabela 6.16 – Cálculo do valor de R Banda Larga (WB) variando ppl 3%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	92.6	3.68
2	13.2	11.7	17.1	95.8	3.79
3	16.4	10.8	10.3	105.1	3.93
4	24.4	11.4	7.2	103.4	4.03
5	32.0	8.7	9.3	97.6	3.85
6	48.0	9.6	10.2	98.5	3.88

Fonte: Do autor (2023)

Ao calcular o valor de R Banda Larga (WB) com uma variação de 3,0% do Ppl (*Packet-loss Probability*), as características permanecem as mesmas das avaliações anteriores. No entanto, o Bpl (*Packet-loss Robustness Factor*) inicialmente aumenta em 24,4 bits, mas depois diminui em 32 bits e, em seguida, volta a subir em 48 bits. Da mesma forma, o Lewb (*Equipment Impairment Factor*) segue as mesmas características do Bpl nessa avaliação. O fator R e o MOS (*Mean Opinion Score*) aumentam em conjunto com o PPL e a taxa de bits.

Gráfico 6.4 – ppl%3



Fonte: Do Autor (2023)

Com base nos critérios de avaliação anteriores, o coeficiente de correlação de Pearson agora é de 0.38, o que indica uma correlação positiva, porém fraca. Essa fraqueza da correlação ocorre devido à distância entre os valores do MOS no eixo x e os valores da taxa de bits no eixo y.

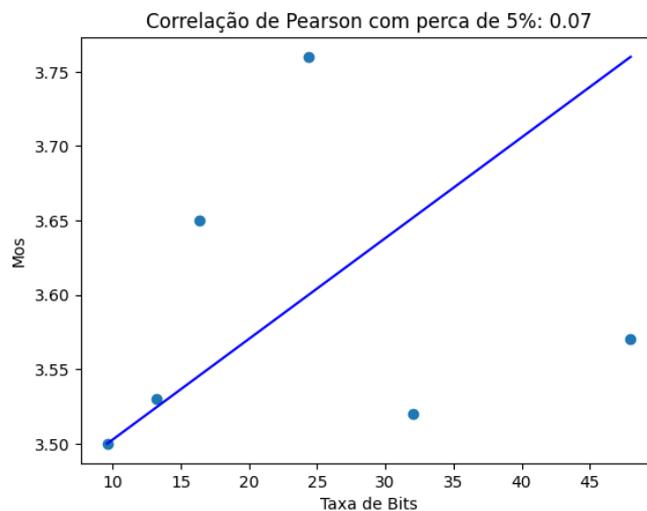
Tabela 6.17 – Cálculo do valor de R Banda Larga (WB) variando ppl 5%

EVS	Taxa de bits (kpbs)	bpl	lew	R	MOS
1	9.6	13.0	22.7	87.8	3.50
2	13.2	11.7	17.1	88.4	3.53
3	16.4	10.8	10.3	91.7	3.65
4	24.4	11.4	7.2	94.9	3.76
5	32.0	8.7	9.3	88.3	3.52
6	48.0	9.6	10.2	89.6	3.57

Fonte: Do autor (2023)

Ao variar o percentual do Ppl para 5,0%, nota-se que as características são mantidas conforme as avaliações anteriores. No entanto, é perceptível que, à medida que o percentual do Ppl aumenta, os fatores R e MOS tendem a diminuir. Inicialmente, tanto o fator R quanto o MOS apresentam queda, mas nas taxas de 16,4% e 24,4% ocorre um aumento temporário, seguido por uma queda contínua nas taxas seguintes de 32,0% e 48,0%.

Gráfico 6.5 – ppl%5



Fonte: Do Autor (2023)

De acordo com a mesma avaliação dos modos anteriores, o coeficiente de correlação de Pearson apresenta um valor de 0.07, o que indica uma correlação fraca próxima de zero. Ao observar o gráfico, podemos notar que apenas dois pontos, representados pelos valores do MOS, estão próximos um do outro, enquanto os demais valores de MOS no eixo x e os valores de taxa de bits no eixo y estão distantes entre si.

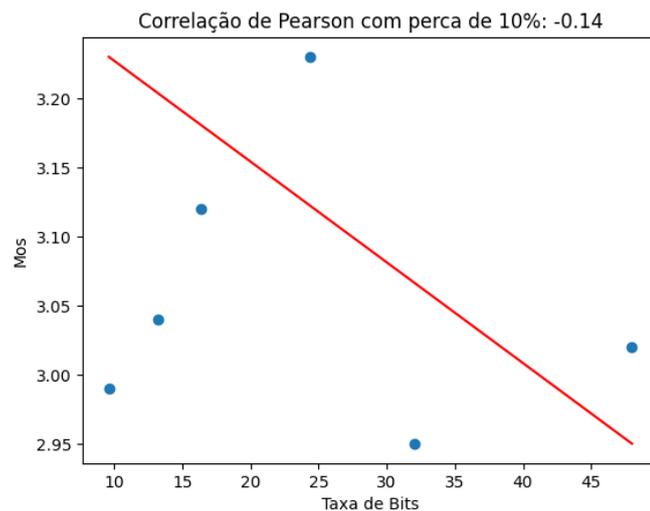
Tabela 6.18 – Cálculo do valor de R Banda Larga (WB) variando ppl 10%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	74.7	2.99
2	13.2	11.7	17.1	75.8	3.04
3	16.4	10.8	10.3	77.8	3.12
4	24.4	11.4	7.2	80.6	3.23
5	32.0	8.7	9.3	73.7	2.95
6	48.0	9.6	10.2	75.4	3.02

Fonte: Do autor (2023)

Ao variar o Ppl para 10.0%, nesta avaliação, a situação permanece a mesma em relação à avaliação anterior. No entanto, é esperado que tanto o fator R quanto o MOS continuem apresentando valores em queda.

Gráfico 6.6 – ppl%10



Fonte: Do Autor (2023)

Considerando os critérios de avaliação dos modos anteriores, o coeficiente de correlação de Pearson atual é -0.14. Isso implica em uma mudança na direção e na cor da linha do gráfico, uma vez que a correlação de Pearson agora é negativa, embora fraca. Essa fraqueza é atribuída à distância entre os valores do MOS em relação aos pontos no eixo x e à taxa de bits no eixo y.

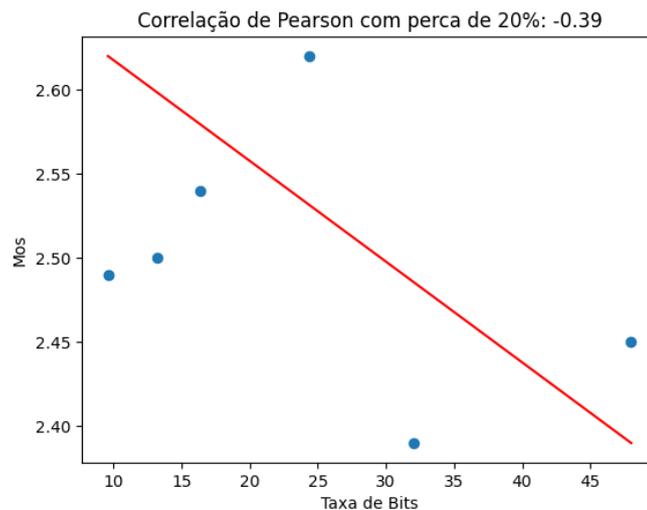
Tabela 6.19 – Cálculo do valor de R Banda Larga (WB) variando ppl 20%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	62.3	2.49
2	13.2	11.7	17.1	62.6	2.50
3	16.4	10.8	10.3	63.5	2.54
4	24.4	11.4	7.2	65.7	2.62
5	32.0	8.7	9.3	59.8	2.39
6	48.0	9.6	10.2	61.3	2.45

Fonte: Do autor (2023)

Ao variar o percentual de Ppl em 20.0%, é importante observar que a forma de avaliação permanece a mesma da avaliação anterior. No entanto, é provável que tanto o fator R quanto o MOS (*Mean Opinion Score*) continuem diminuindo à medida que o percentual de Ppl aumenta. Essas métricas são utilizadas para avaliar a qualidade de transmissão de voz em telecomunicações. Conforme a recomendação ITU-T G.109, a qualidade da transmissão é determinada pelas métricas atribuídas. Valores entre 80 e 89 são considerados de alta qualidade, enquanto valores entre 70 e 79 são classificados como médios. Quando os valores estão entre 60 e 69, a qualidade é considerada baixa, e valores entre 50 e 59 indicam uma qualidade ruim. Qualquer avaliação abaixo de 50 é considerada insatisfatória em termos de qualidade. Assim, ao elevar o percentual de Ppl para 20.0%, é esperado que a qualidade da transmissão sofra impacto negativo, com uma tendência de queda nos valores de R e MOS. De acordo com os critérios estabelecidos pela ITU-T G.109, as pontuações das métricas têm diminuído, indicando uma deterioração progressiva da qualidade.

Gráfico 6.7 – ppl%20



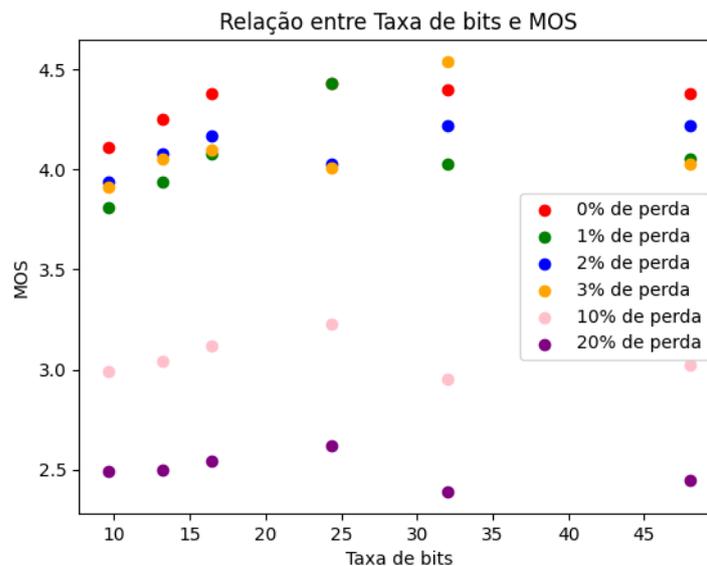
Fonte: Do Autor (2023)

Seguindo o mesmo critério de avaliação dos modos anteriores, o coeficiente de correlação de Pearson agora é de -0,39. A linha do gráfico continua mostrando uma relação inversa, e a reta que corta o

gráfico é caracterizada pela cor vermelha, como definido no código utilizado para avaliar essa correlação de Pearson. No entanto, essa correlação agora é considerada quase imperfeita, uma vez que está próxima de -1. Ainda assim, é uma correlação fraca, uma vez que os valores do MOS no eixo x do gráfico estão distantes uns dos outros e dos demais valores da taxa de bits no eixo y.

Para melhor exemplificar as taxas de perda, apresentamos um único gráfico que exibe diferentes valores de MOS e taxas de bits, baseado nos dados registrados nas planilhas anteriores.

Gráfico 6.8 – ppl em vários valores de perda



Fonte: Do Autor (2023)

Informações sobre os valores disponíveis dos fatores de comprometimento do equipamento de Super Banda Larga (SWB), Fator de deterioração do equipamento (Ie), e fatores de Robustez de Perda de Pacotes (Bpl). Esses valores devem ser usados em uma escala de classificação de transmissão estendida (escala R), conforme definido em (G., 2015)

Tabela 6.20 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 0.0%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	125.3	4.19
2	13.2	11.7	17.1	130.9	4.30
3	16.4	10.8	10.3	137.7	4.41
4	24.4	11.4	7.2	140.8	4.44
5	32.0	8.7	9.3	138.7	4.42
6	48.0	9.6	10.2	137.8	4.41

Fonte: Do autor (2023)

A partir deste momento em diante, a determinação do valor de R será para o modo Super Banda Larga (SWB), começando com 0.0% para os mesmos atributos dos testes do modo banda larga (WB) anteriores.

Tabela 6.21 – Cálculo do valor de R Super Banda Larga (SWB) variando 1.0%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	117.5	4.0
2	13.2	11.7	17.1	121.9	4.11
3	16.4	10.8	10.3	127.4	4.23
4	24.4	11.4	7.2	130.7	4.30
5	32.0	8.7	9.3	161.1	4.20
6	48.0	9.6	10.2	126.3	4.21

Fonte: Do autor (2023)

Cálculo do valor de R Super Banda Larga (SWB) iniciando ppl a 1.0% os testes são mantidos com as mesmas características com a tendencia do fator R e MOS subir a medida o ppl aumenta o percentual e a taxa de bits aumenta.

Tabela 6.22 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 2%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	110.7	3.81
2	13.2	11.7	17.1	114.1	3.91
3	16.4	10.8	10.3	118.7	4.03
4	24.4	11.4	7.2	122.2	4.12
5	32.0	8.7	9.3	115.8	3.95
6	48.0	9.6	10.2	116.8	3.98

Fonte: Do autor (2023)

Calculando o valor de R Super Banda Larga (SWB) iniciando ppl a 2.0% nota-se que mantém os padrões dos testes anteriores porém os valores de R e MOS tendem a baixar a medida que o ppl aumenta o percentual e as taxas de bits aumentam.

Tabela 6.23 – Cálculo do valor de R Super Banda Larga (SWB) variando bpl 3%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	104.8	3.63
2	13.2	11.7	17.1	107.5	3.72
3	16.4	10.8	10.3	111.2	3.83
4	24.4	11.4	7.2	114.8	3.93
5	32.0	8.7	9.3	107.2	3.71
6	48.0	9.6	10.2	108.8	3.76

Fonte: Do autor (2023)

Calculando o valor de R Super Banda Larga (SWB) iniciando ppl agora em 3.0% os testes procedem da mesma maneira seguindo o mesmo protocolo de quando ppl aumenta o percentual e taxa de bits também mantém o crescimento, dessa forma o fator R e MOS também continua caindo seus valores.

Tabela 6.24 – Cálculo do valor de R Super Banda Larga (SWB variando ppl 5%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	94.9	3.31
2	13.2	11.7	17.1	96.5	3.36
3	16.4	10.8	10.3	99.2	3.45
4	24.4	11.4	7.2	102.8	3.57
5	32.0	8.7	9.3	93.9	3.28
6	48.0	9.6	10.2	96.1	3.35

Fonte: Do autor (2023)

Cálculo do valor de R Super Banda Larga (SWB) agora iniciando ppl em 5.0% segue procedimento sem nenhuma alteração os testes tendem a ter o mesmo tipo de avaliação, o que é notável é a queda do fator R e MOS, a medida que o ppl aumenta juntamente com a taxa de bits.

Tabela 6.25 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl 7%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	87.0	3.04
2	13.2	11.7	17.1	87.9	3.07
3	16.4	10.8	10.3	89.8	3.14
4	24.4	11.4	7.2	93.3	3.26
5	32.0	8.7	9.3	84.0	2.93
6	48.0	9.6	10.2	86.4	3.02

Fonte: Do autor (2023)

Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 7.0% seguindo mesmo padrão dos testes realizados anteriormente porém segue o mesmo ritmo de queda em relação ao fator R e MOS, quanto mais o ppl aumenta menor fica os valores de R e MOS.

Tabela 6.26 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 10.0%

EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	77.8	2.71
2	13.2	11.7	17.1	78.0	2.72
3	16.4	10.8	10.3	79.2	2.76
4	24.4	11.4	7.2	82.5	2.88
5	32.0	8.7	9.3	73.1	2.54
6	48.0	9.6	10.2	75.7	2.63

Fonte: Do autor (2023)

Cálculo do valor de R Super Banda Larga (SWB) agora variando ppl a 10.0% seguindo mesmo padrão dos testes anteriores nada mudou apenas a queda do fator R e MOS tende a continuar cainda a medida que o ppl aumenta o percentual.

Tabela 6.27 – Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 20.0%

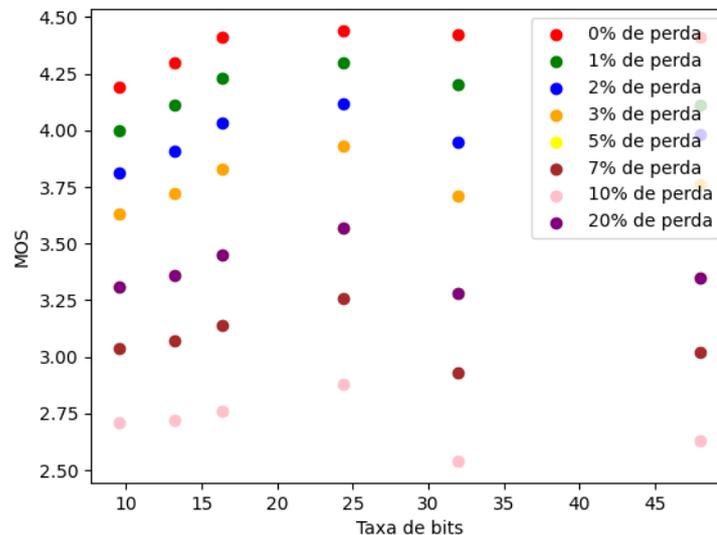
EVS	Taxa de bits (kpbs)	bpl	lewb	R	MOS
1	9.6	13.0	22.7	94.9	3.31
2	13.2	11.7	17.1	96.5	3.36
3	16.4	10.8	10.3	99.2	3.45
4	24.4	11.4	7.2	102.8	3.57
5	32.0	8.7	9.3	93.9	3.28
6	48.0	9.6	10.2	96.1	3.35

Fonte: Do autor (2023)

Cálculo do valor de R Super Banda Larga (SWB) variando ppl a 20.0% os procedimentos continuam os mesmos dos acontecimentos anteriores com o fator R e MOS caindo a medida que o ppl aumenta.

O gráfico 6.9 a seguir mostra o resultado de todas as informações mencionadas nas planilhas, em que o ppl variou de 0 a 20% em SWB.

Gráfico 6.9 – ppl em vários valores de perda (SWB)



Fonte: Do Autor (2023)

Ao explicitar vários valores de ppl em SWB em um único gráfico de dispersão e usar o coeficiente de correlação de Pearson para analisar a relação entre as variáveis. A correlação de Pearson parece válida porque as variáveis são contínuas e apresentam uma relação linear aproximada. Porém a medida que o percentual aumenta a relação torna-se uma relação não linear ou discreta. Porque os dados começam a dispersar e começam a ficar longe um outro.

6.2.2 Avaliação Codec Adaptive Multi-Rate Wideband (AMR-WB) com E-model

Avaliação da qualidade da fala para áudio banda larga E-MODEL WB para *Adaptive Multi-Rate Wideband* (AMR-WB). AMR-WB é um padrão de codificação de fala amplamente usado para comunicação de áudio de alta qualidade. Usaremos o mesmo percentual na avaliação do codec AMR-WB: 0%, 1%, 2%, 3%, 5%, 10%, 20% iniciando com: Cálculo do valor de R Banda Larga (WB) com ppl iniciando a 0.0% valores tirados de (Nguyen, et. al, 2016). O codec AMR-WB também inclui dois parâmetros adicionais, *Ie* e *Bpl*, que são usados para otimizar o desempenho do codec com base nas características do sinal de áudio.

Tabela 6.28 – Cálculo do valor de R Banda Larga (WB) variando ppl a 0.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	89.8	3.58
1	8.85	13.5	25.0	103.8	4.04
2	12.65	4.3	13.0	115.8	4.33
3	14.25	14.1	10.0	118.8	4.39
4	15.85	13.1	7.0	121.8	4.43
5	18.25	12.5	5.0	123.8	4.46
6	19.85	12.3	4.0	124.8	4.47
7	23.05	13.0	1.0	127.8	4.49
8	23.85	4.9	8.0	120.8	4.42

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 0.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta.

Tabela 6.29 – Cálculo do valor de R Banda Larga (WB) variando ppl a 1.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	85.8	3.43
1	8.85	13.5	25.0	99.0	3.90
2	12.65	4.3	13.0	100.4	3.94
3	14.25	14.1	10.0	113.2	4.28
4	15.85	13.1	7.0	115.6	4.33
5	18.25	12.5	5.0	117.2	4.36
6	19.85	12.3	4.0	118.0	4.37
7	23.05	13.0	1.0	121.1	4.42

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 1.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que

prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair.

Tabela 6.30 – Cálculo do valor de R Banda Larga (WB) variando ppl a 2.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	82.3	3.29
1	8.85	13.5	25.0	94.8	3.76
2	12.65	4.3	13.0	89.8	3.58
3	14.25	14.1	10.0	108.3	4.16
4	15.85	13.1	7.0	110.2	4.21
5	18.25	12.5	5.0	111.4	4.24
6	19.85	12.3	4.0	112.1	4.26
7	23.05	13.0	1.0	115.3	4.32
8	23.85	4.9	8.0	95.6	3.78

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 2.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair.

Tabela 6.31 – Cálculo do valor de R Banda Larga (WB) variando ppl a 3.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	79.2	3.17
1	8.85	13.5	25.0	91.1	3.63
2	12.65	4.3	13.0	82.1	3.29
3	14.25	14.1	10.0	103.9	4.05
4	15.85	13.1	7.0	105.4	4.09
5	18.25	12.5	5.0	105.8	4.10
6	19.85	12.3	4.0	107.0	4.13
7	23.05	13.0	1.0	110.2	4.21
8	23.85	4.9	8.0	87.8	3.51

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 3.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair.

Tabela 6.32 – Cálculo do valor de R Banda Larga (WB) variando ppl a 5.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	74.1	2.97
1	8.85	13.5	25.0	84.9	3.40
2	12.65	4.3	13.0	71.8	2.87
3	14.25	14.1	10.0	96.6	3.82
4	15.85	13.1	7.0	97.5	3.85
5	18.25	12.5	5.0	98.8	3.87
6	19.85	12.3	4.0	98.5	3.88
7	23.05	13.0	1.0	101.7	3.98
8	23.85	4.9	8.0	76.9	3.08

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 5.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair.

Tabela 6.33 – Cálculo do valor de R Banda Larga (WB) variando ppl a 10.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	65.3	2.61
1	8.85	13.5	25.0	74.1	2.97
2	12.65	4.3	13.0	58.5	2.33
3	14.25	14.1	10.0	83.6	3.34
4	15.85	13.1	7.0	83.8	3.35
5	18.25	12.5	5.0	83.8	3.35
6	19.85	12.3	4.0	84.0	3.36
7	23.05	13.0	1.0	87.0	3.47
8	23.85	4.9	8.0	62.5	2.49

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 10.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair. Porém vale lembrar que valores de R abaixo de 60 são considerados ruins.

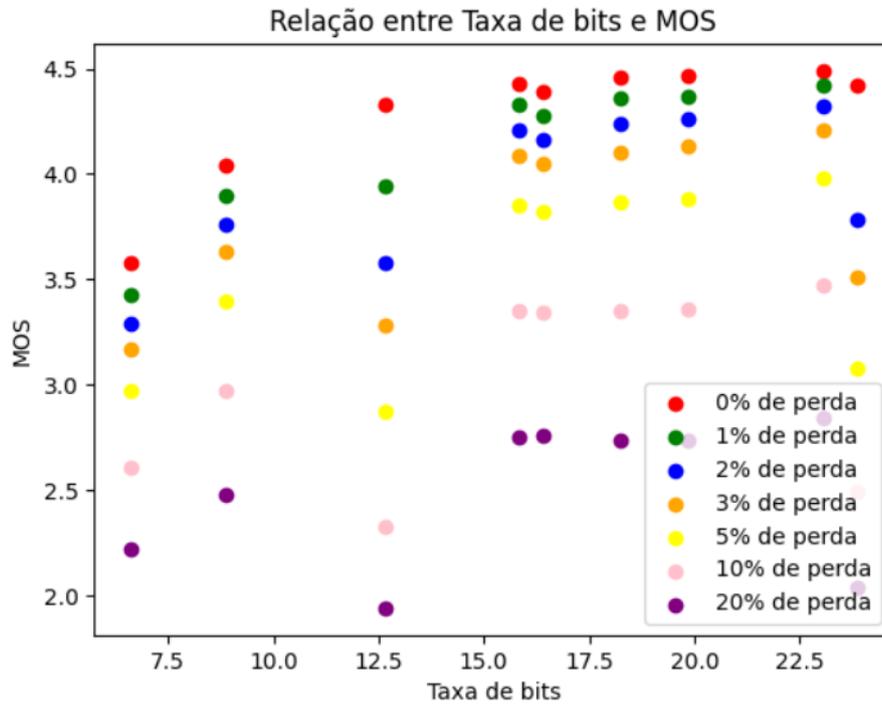
Tabela 6.34 – Cálculo do valor de R Banda Larga (WB) variando ppl a 20.0%

AMR-WB	Taxa de bits (kpbs)	bpl	lewb	R	MOS
0	6.60	12.8	39.0	55.7	2.22
1	8.85	13.5	25.0	62.1	2.48
2	12.65	4.3	13.0	48.4	1.94
3	14.25	14.1	10.0	69.0	2.76
4	15.85	13.1	7.0	68.7	2.75
5	18.25	12.5	5.0	68.5	2.74
6	19.85	12.3	4.0	68.5	2.74
7	23.05	13.0	1.0	70.9	2.84
8	23.85	4.9	8.0	51.0	2.04

Fonte: Do autor (2023)

Cálculo do valor de R e MOS Banda Larga (WB) com ppl iniciando a 20.0% com nove modos iniciando em 6.60 até 23.85 seguindo o mesmo padrão de avaliação do codec EVS anteriormente. O que prevalece é que a qualidade nem sempre é melhor com uma taxa mais alta. Mais a medida o ppl aumenta o fator R e o MOS tendem cair. Porém vale lembrar que valores de R abaixo de 60 são considerados ruins. Para uma melhor representação das taxas de perda, apresentamos a seguir a Figura 6.12 que exhibe diferentes valores de MOS e taxas de bits, com base nos dados das planilhas anteriores agora usando o codec AMR-WB antecessor do EVS variando a taxa de perda de 0 a 20% conforme foi realizado nos outros testes. Ao usar os valores de ppl em AMR-WB em um único gráfico de dispersão e através do coeficiente de correlação de Pearson para analisar a relação entre as variáveis. A correlação de Pearson parece válida até certo ponto e as variáveis são contínuas e apresentam uma relação linear aproximada. Porém a medida que o percentual aumenta a relação torna-se uma relação não linear ou discreta,

Gráfico 6.10 – ppl em vários valores de perda (AMR-WB)



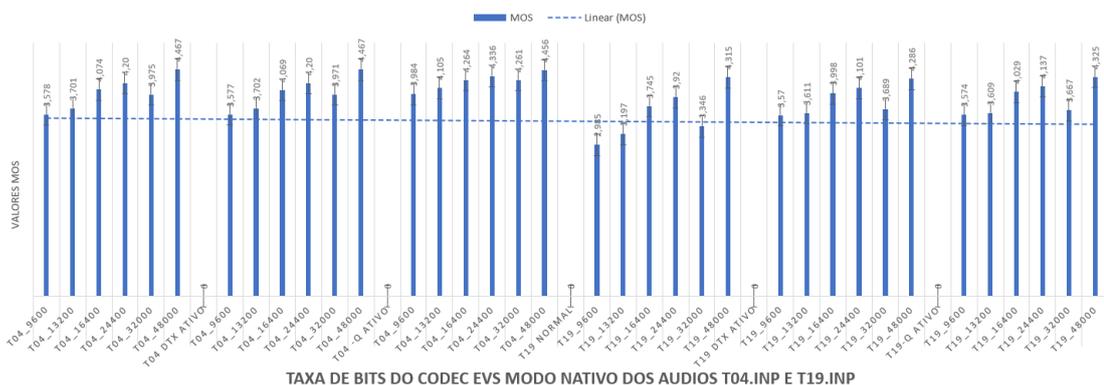
Fonte: Do Autor (2023)

6.3 Gráfico Rec. P.862 e E-model

Ao combinar as avaliações objetivas do P.862 (também conhecido como PESQ) e as previsões objetivas do E-Model, é possível obter uma visão mais completa do desempenho do codec EVS em diferentes situações. Enquanto o P.862 fornece informações sobre a qualidade da voz em condições ideais, o E-Model leva em consideração as degradações reais encontradas nas redes, permitindo uma avaliação mais realista do codec.

Através dos dois gráficos abaixo apresentamos o resultado do estudo em ambas métricas de avaliação da voz:

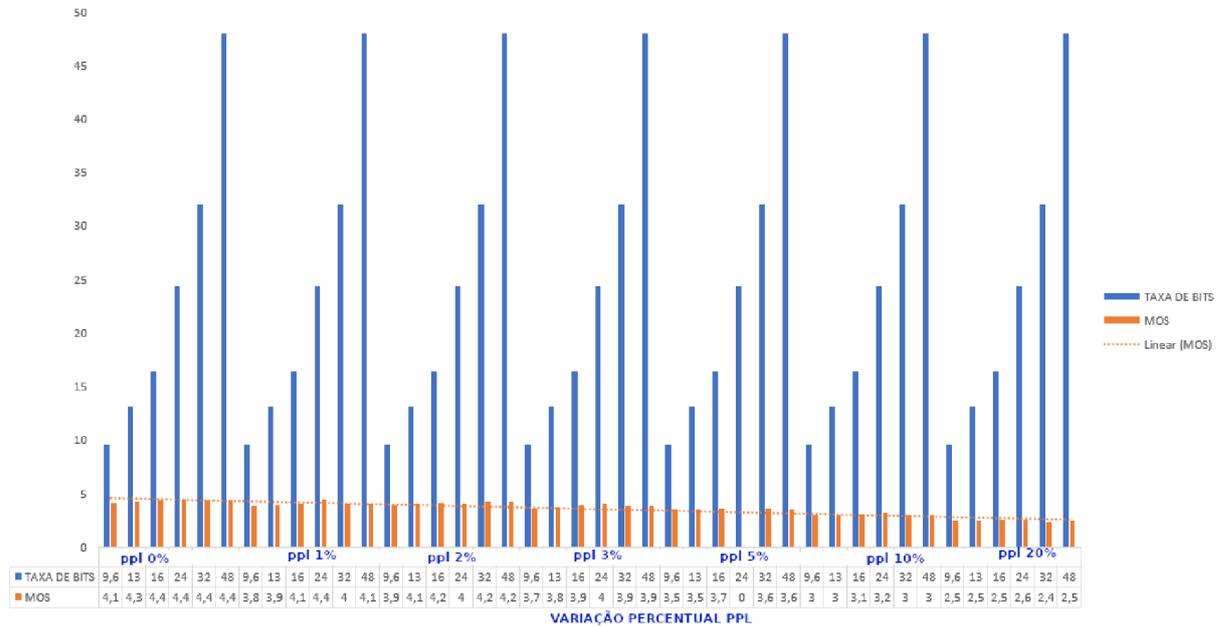
Gráfico 6.11 – MOS - P.862



Fonte: Do Autor (2023)

Em seguida temos um gráfico E-Model:

Gráfico 6.12 – MOS - Modelo E



Fonte: Do Autor (2023)

A Recomendação P.862 (PESQ) é uma métrica objetiva que se foca na qualidade perceptual da fala, enquanto o E-Model é uma métrica mais abrangente que leva em consideração vários parâmetros de rede para calcular a qualidade da chamada de voz. Ambas as métricas têm suas aplicações e são utilizadas para medir e comparar a qualidade de diferentes sistemas e redes de comunicação.

7 CONCLUSÃO

Nesse trabalho destacamos a recomendação do uso do codec EVS para serviços que buscam priorizar a qualidade de chamada e oferecer uma experiência de voz aprimorada aos usuários. O EVS é conhecido por sua eficiência no atendimento de áudio, garantindo a largura de banda necessária durante as chamadas sem comprometer a excepcional qualidade de voz. Isso é especialmente benéfico em redes com recursos limitados, como conexões de Internet lentas ou áreas com cobertura de celular fraca. Comparado ao seu antecessor, o codec AMR, o EVS oferece a melhor qualidade de áudio em taxas de bits mais altas. Portanto, para obter o máximo das melhorias do EVS, é recomendável usar taxas de bits mais elevadas.

O estudo analisou a qualidade da voz usando o codec EVS e avaliou a pontuação MOS com diferentes parâmetros. Os resultados mostram que o desempenho da transmissão de voz é beneficiado por taxas de bits mais altas e pela ativação de certos parâmetros, como o DTX (*Discontinuous Transmission*) e o parâmetro -q. O DTX permite economizar largura de banda ao interromper temporariamente a transmissão de voz durante momentos de silêncio. O parâmetro -q refere-se à qualidade da transmissão de voz, onde um valor mais alto indica melhor qualidade, gerado em taxas de bits mais altas e, simultaneamente, a qualidade da transmissão.

O estudo também utilizou o E-model para medir a qualidade da voz, levando em conta fatores como largura de banda, taxa de bits, perda de pacotes e latência da rede, para determinar a qualidade percebida pelo usuário durante as chamadas de voz.

Em conclusão, a pesquisa contínua na área da qualidade da voz em sistemas de comunicação móvel é essencial para avanços tecnológicos e melhorias contínuas nos codecs, como o EVS, proporcionando uma melhor experiência de comunicação para os usuários de dispositivos móveis em todo o mundo. Essas estruturas podem beneficiar diversos setores, incluindo produtos de aparelhos móveis, fornecedores de serviços de comunicação e estudiosos interessados em aprimorar a experiência do usuário em links de voz.

REFERÊNCIAS

- 3GPP. **Technical Specification Group Services and System Aspects**. 2011. <<https://www.3gpp.org/ftp/Specs/html-info/26273.htm>>. Accessed: 2023-8-21.
- GOUVEIA, D.; ROSA, R.; RODRÍGUEZ, D. **Avaliação da Qualidade da Voz em Serviços de Comunicação usando Deep Learning**. In: **Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**. Porto Alegre, RS, Brasil: SBC, 2018. p. 183–196. ISSN 2177-9384. Disponível em: <<https://sol.sbc.org.br/index.php/sbrc/article/view/2415>>.
- SHIGUEOKA, R. S. **ANÁLISE COMPARATIVA DE TÉCNICAS PARA OCULTAMENTO DE PERDAS DE PACOTES EM APLICAÇÕES DO TIPO VOZ SOBRE IP (VOIP)**. 2016.
- ITU-TP.800. **ITU-TP.800. Métodos para determinação subjetiva da qualidade de transmissão**. 1996. <<https://www.itu.int/rec/T-REC-P.800-199608-I>>. (Acessado em 23/07/2023).
- SOUSA, A. L. R. N. de. **Avaliação da Qualidade de Voz do Serviço VoIP em Sistemas HSDPA**. 2007. <https://repositorio.ufc.br/bitstream/riufc/16076/1/2007_dis_lrnsousa.pdf>. Accessed: 2023-4-17.
- RODRÍGUEZ, D. Z. et al. **Speech Quality Assessment in Wireless Communications With MIMO Systems Using a Parametric Model**. *IEEE Access*, v. 7, p. 35719–35730, 2019.
- ANDRADA L. & MARINHO, R. . **Metodologias de Avaliação de Qualidade de Fluxos de Voz**. 2007. <<http://www.midiacom.uff.br/~debora/fsmm/trab-2007-1/qvoz.pdf>>. Accessed: 2022-8-28.
- EMANUEL et al. **Uso de métricas objetivas para avaliação de qualidade de voz com diferentes degradações**", booktitle = "Org.br. 2016. <http://www.sbrc.org.br/sbrc2016/anais/minicursos/uso_de_metricas.pdf> Accessed: 2022-9-17.
- GOUVEIA et al. **Avaliação da Qualidade da Voz em Serviços de Comunicação usando Deep Learning**. In: **Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC 2018)**. [S.l.]: Sociedade Brasileira de Computação - SBC, 2018.
- TYMCHENKO, O. et al. **Speech quality measurement methods and models over ip-networks**. In: **2020 21st International Conference on Research and Education in Mechatronics (REM)**. [S.l.: s.n.], 2020. p. 1–6.
- RODRÍGUEZ, A. R. **Evaluación de desempeño del códec de banda ancha EVS**. Tese (Doutorado) — Universidad Central “Marta Abreu” de Las Villas, Facultad de Ingeniería Eléctrica, Departamento de Electrónica y Telecomunicaciones, jun. 2018.
- PREGO, T. de M. **ACERCADAREVERBERAC,~AOEMSINAISDEVOZ:QUANTIFICAC,~AO PERCEPTUALEAPERFEIC,~OAMENTODEALGORITMOSDE DESREVERBERAC,~AO**. 2012. <<http://www.pee.ufrj.br/index.php/pt/producao-academica/teses-de-doutorado/2012/2012053101-2012053101/file>>. Accessed: 2023-4-23.
- SOUSA, J. M. d. **Avaliação do índice de qualidade de voz em redes móveis celulares baseada na percepção dos usuários**. 2011.
- FAGHIHI, E.; BEHDADFAR, M. **Adaptive Resource Utilization and Quality Management in VoIP Networks with Quantitative Relationship between QoE and QoS Using HMM**. In: **2020 International Conference on Information Networking (ICOIN)**. [S.l.: s.n.], 2020. p. 493–498.
- RODRIGUEZ, D. Z. et al. **Assessment of Quality-of-Experience in Telecommunication Services**. v. 6, p. 241, 10 2016.
- LAGHARI et al. **Toward total quality of experience: A qoe model in a communication ecosystem**. *IEEE Communications Magazine*, p. 58–65, 2012.

SANTOS, C. J. d. Seleção de codificadores de voz (codecs) através do balanceamento entre qos e qoe auxiliado por algoritmo baseado em análise hierárquica de processos (ahp). 2019.

TARUK, M. et al. **Quality of Service Voice over Internet Protocol in Mobile Instant Messaging**. In: **2018 2nd East Indonesia Conference on Computer and Information Technology (EIconCIT)**. [S.l.: s.n.], 2018. p. 285–288.

T. ZOURZOUVILLYS. & RESCORLA. **An Introduction to Standards-Based VoIP: SIP, RTP, and Friends**. **IEEE Internet Computing**, v. 14, n. 2, p. 69–73, March 2010. ISSN 1089-7801.

TELECO. **Arquitetura de Rede VoIP**. 2019. <https://www.teleco.com.br/tutoriais/tutorialvoipconv/pagina_4.asp>.

SINIARSKI, B. et al. **OpenFlow based VoIP QoE monitoring in enterprise SDN**. In: **2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)**. [S.l.: s.n.], 2017. p. 660–663.

HARTE, et al. **TCD-VoIP, a research database of degraded speech for assessing quality in VoIP applications**. 07 2015.

SONASKAR, S & GIRIPUNJE, S. **Voice over intranet based private branch exchange system design**. In: **2011 3rd International Conference on Electronics Computer Technology**. [S.l.: s.n.], 2011. p. 287–291.

AHUJA, S.; ENSOR, J. **VoIP: What is it Good for?** **ACM Queue**, v. 2, p. 48–55, 09 2004.

CAVALCANTE, A. et al. **Análise dos efeitos dos codecs de áudio na avaliação de desvios vocais**. In: **Anais de XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais**. [S.l.]: Sociedade Brasileira de Telecomunicações, 2017.

M. R. HIKMATULLAH & S. HARYADI. **Perceptual evaluation of speech quality over the top call service**. In: **2017 3rd International Conference on Wireless and Telematics (ICWT)**. [S.l.: s.n.], 2017. p. 181–185.

COMER, D. **Interligação de Redes com TCP/IP–: Princípios, Protocolos e Arquitetura**. [S.l.]: Elsevier Brasil, 2016. v. 1.

LINS, T. **Internet das Coisas : coletando dados na IoT - Laboratório iMobilis**. 2015. <<http://www2.decom.ufop.br/imobilis/iot-coletando-dados/>>. (Acessado em 24/07/2023).

FONSECA, H. T. d. S. **Segurança, privacidade, QoS e QoE em plataformas VoIP**. Dissertação (Mestrado), 2014.

UHL, T. **QoS by VoIP Under Use Different Audio Codecs**. In: **2018 Joint Conference - Acoustics**. [S.l.: s.n.], 2018,. p. 1–4. ISSN null.

ISPBLOG. **5G e VoIP**. 2020. <<https://www.ispblog.com.br/2020/07/01/5g-e-voip/>>. (Acessado em 30/12/2021).

A. ALHAYAJNEH et al. . **Quality of Service Analysis of VoIP Services**. In: **2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)**. [S.l.: s.n.], 2018. p. 812–818.

REC. ITU-T. <https://www.itu.int/rec/T-REC-G.107>. 2013. <<https://www.itu.int/rec/T-REC-G.107>>. Accessed: 2023-2-9.

VIANNA, B. de A. **Universidade Federal Fluminense**. 2007. <<https://site.ic.uff.br/wp-content/uploads/2021/11/350.pdf>>.

- BORGES, R. S. **Projetos para Redes VoIP Homologação da tecnologia VoIP em Redes de Dados**. 2012. <https://repositorio.ufsc.br/bitstream/handle/123456789/184626/Projetos2-2012-2_RAPHAEL_BORGES.pdf?sequence=-1>. Acessado em 7 de junho de 2023.
- REC. ITU-T. **Tutorial**. 2008. <<https://www.itu.int/ITU-T/2005-2008/com12/emodelv1/tut.htm>>. Accessed: 2022-9-10.
- STASINSZKY, O. **How do you measure R-Factor in VoIP quality?** [S.l.]: VoIP-Info.org, 2022. <<https://www.voip-info.org/r-factor-in-voip/>>. Accessed: 2023-7-9.
- CHOCHOL, P. **QUALITATIVE FACTORS THAT IMPACT REAL IMPLEMENTING VOIP IN PRIVATE NETWORKS**. 07 2023.
- RIX, A. et al. Perceptual evaluation of speech quality (pesq)-a new method for speech quality assessment of telephone networks and codecs. In: **2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)**. [S.l.: s.n.], 2001. v. 2, p. 749–752 vol.2.
- GOUVEA et al. **Vista do Avaliação da Qualidade da Voz em Serviços de Comunicação usando Deep Learning**. 2018. <<https://sol.sbc.org.br/index.php/sbrc/article/view/2415/2379>>. Accessed: 2022-9-10.
- StudFiles. 2015. Disponível em: <URL:<https://studfiles.net/preview/1688917/page:5/>>.
- ITU-TP.862. **R. P.862: Avaliação perceptiva da qualidade da fala (PESQ): Um método objetivo para avaliação ponta a ponta da qualidade da fala de redes telefônicas de banda estreita e codecs de fala**. 2001. <<https://www.itu.int/rec/T-REC-P.862-200102-I/en>>. (Acessado em 24/07/2023).
- PEREIRA, C. H.; RODRIGUEZ, D. Z. . Estudo da distribuição temporal das perdas de pacotes e o seu impacto na qualidade de voz. In: **Anais do XXXIV Simpósio Brasileiro de Telecomunicações**. [S.l.]: Sociedade Brasileira de Telecomunicações, 2016.
- RODRÍGUEZ D.Z. , **Algoritmo para determinação da taxa de transmissão em uma rede IP**. 2009. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/3/3142td-30032010-153910/pt-br.php>>.
- L. Malfait et al. **P.563—The ITU-T Standard for Single-Ended Speech Quality Assessment**. **IEEE Transactions on Audio, Speech, and Language Processing**, v. 14, n. 6, p. 1924–1934, Nov 2006. ISSN 1558-7916.
- REC.G.107, I.-T. **G.701 : Vocabulary of digital transmission and multiplexing, and pulse code modulation (PCM) terms**. 1993. <<https://www.itu.int/rec/T-REC-G.701-199303-I/en>>. Accessed: 2023-4-21.
- HARFF, S. **Navegação Ciências Exatas e da Terra por Autor “Harff, Simone”**. 2008. <<https://lume.ufrgs.br/handle/10183/12236/browse?type=author&value=Harff%2C+Simone>>. Accessed: 2023-4-21.
- TECHNOLOGIES, L. V. **G.726**. 2017. Disponível em: <<https://www.vocal.com/speech-coders/g-726/>>.
- COELHO, V. T. **Monitoramento de Qualidade de Serviço em Voz Sobre IP**. 2008.
- KONG, X. et al. **Implementation of G.729 Codec Based on DaVinci Technology**. In: **2008 International Conference on MultiMedia and Information Technology**. [S.l.: s.n.], 2008. p. 11–14.
- SJOBERG, L. et al. **RFC 4867 - RTP payload format and file storage format for the adaptive multi-rate (amr) and adaptive multi-rate wideband (amr-wb) audio codecs**. 2007. Disponível em: <<https://tools.ietf.org/html/rfc4867>>.

ABICHANDANI, J. et al. **A Comparative Study of Voice Quality and Coverage for Voice over Long Term Evolution Calls Using Different Codec Mode-sets**. *IEEE Access*, v. 5, p. 10315–10322, 2017. ISSN 2169-3536.

SALAMI, R. et al. The adaptive multi-rate wideband codec: history and performance. In: **Speech Coding, 2002, IEEE Workshop Proceedings**. [S.l.: s.n.], 2002. p. 144–146.

BYUN, K. J. et al. **Real-time Implementation of AMR and AMR-WB using the Fixed-point DSP for WCDMA Systems**. In: **2006 IEEE International Symposium on Consumer Electronics**. [S.l.: s.n.], 2006. p. 1–6. ISSN 0747-668X.

BOÏL, J. . **Impact of the GSM AMR Codec on Automatic Vowel Formant Measurement in Praat and VoiceSauce**. In: **2018 41st International Conference on Telecommunications and Signal Processing (TSP)**. [S.l.: s.n.], 2018. p. 1–4.

BESSETTE, B. et al. **The Adaptive Multirate Wideband Speech Codec (AMR-WB)**. *IEEE Transactions on Speech and Audio Processing*, v. 10, n. 8, p. 620–636, Nov 2002. ISSN 1063-6676.

SJOBERG, J. et al. **RFC ft-ietf-avt-rtp-amr-bis: RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) audio codecs**. 2007. <<https://datatracker.ietf.org/doc/html/rfc4867>>. Accessed: 2023-4-21.

PEDRO, J. P. e. J. R. C. **Codec de Voz em Comunicações Móveis**. 2011. Disponível em: <http://www.img.lx.it.pt/~fp/cav/ano2010_2011/Trabalhos_MEEC_2011/Artigo%207/site/AMRF-WB.html>.

VOICEAGE. **VoiceAge - codec solutions**. 2019. <<https://voiceage.com/index.html>>. Accessed: 2023-4-22.

JÄRVINEN, K. **Enhanced Voice Services Codec for LTE**. 2014. <<https://www.3gpp.org/news-events/3gpp-news/evs-news>>. Acessado em 13/08/2023.

SCHNELL, M. e. G. B. e. G. J. e. M. L. e. J. r. K. e. L. L. e. H. N. e. N. N. e. o. B. S. P. H. E. **Conferência Internacional IEEE de 2015 sobre Acústica, Processamento de Fala e Sinais**. [S.l.: s.n.], 2015.

MITTAG, G. et al. **Quantifying Quality Degradation of the EVS Super-Wideband Speech Codec**. In: **2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)**. [S.l.: s.n.], 2018. p. 1–6.

3GPP. **Universal Mobile Telecommunications System (UMTS); LTE; EVS Codec General Overview (3GPP TS 26.441 version 12.0.0 Release 12)**. 2014. <https://www.etsi.org/deliver/etsi_ts/126400_126499/126441/12.00.00_60/ts_126441v120000p.pdf>. Accessed: 2022-10-6.

VOICEAGE. **VoiceAge EVS | EVS Standard**. 2015. <https://voiceageevs-com.translate.google/evs.aspx?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt-PT&_x_tr_pto=op,sc>. (Accessed on 12/12/2021).

COHEN, E. **Testando o EVS Codec**. 2016. <<https://www.spirent.com/blogs/evs-codec-reveals-superior-performance-over-amr-wb>>. (Accessed on 01/04/2022).

FUCHS, G. et al. **Low delay LPC and MDCT-based audio coding in the EVS codec**. In: **2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.: s.n.], 2015. p. 5723–5727. ISSN 1520-6149.

NAHUZ, S. J. et al. **Um Módulo de Monitoramento para uma Arquitetura de Transmissão de Mídia Contínua**. Universidade Federal do Maranhão, 2008.

PRADO, A. d. O. **Otimização da rede de telecomunicações das unidades de direção apoio e execução da Polícia Militar do Estado de Minas Gerais situadas na região metropolitana de Belo Horizonte com a implantação da tecnologia VoIP.** 2009.

FLYNN, K. **Enhanced voice services codec for LTE.** 2022. <https://www.3gpp.org/news-events/1639-evs_news>.

GSA, G. M. S. A. **New Benchmark Studies Evaluate Enhanced Voice Services (EVS) in the Lab and Live Network.** 2016. <<https://www.businesswire.com/news/home/20160803005477/en/New-Benchmark-Studies-Evaluate-Enhanced-Voice-Services-EVS-in-the-Lab-and-Live-Network>>. Accessed: 2022-8-19.

CHEWE, M. **VoLTE por IP Multimedia Subsystem: Implementação Programática do Serviço de Voz em LTE.** 2015.

MARTÃO, V. B. **Codificação de canal para redes 5G utilizando códigos polares.** Universidade Estadual Paulista (UNESP), 2018.

PRASAD, A. et al. **Enhanced Voice Services Based VoLTE Rate Adaptation Mechanism to Improve Quality of Experience.** In: **2019 IEEE Wireless Communications and Networking Conference (WCNC).** [S.l.: s.n.], 2019. p. 1–6.

HUMEIDA, Y. et al. **A method for comparing the performance of EVS and other voice codecs under bursty packet loss.** In: **2018 Principles, Systems and Applications of IP Telecommunications (IPTComm).** [S.l.: s.n.], 2018. p. 1–8.

3GPP. **Codec de serviços de voz aprimorado para LTE.** 2022. <https://www.3gpp.org/news-events/1639-evs_news>. (Accessed on 03/07/2022).

ETSI. **Codec for Enhanced Voice Services (EVS); General overview.** 2020. Disponível em: <https://www.etsi.org/deliver/etsi_ts/126400/126499/126441/16.00.00_60/ts_126441v160000p.pdf>.

3GPP. **3GPP Report WorkItem.** 2017. <https://www.etsi.org/deliver/etsi_ts/126400/126499/126442/17.00.00_60/ts_126442v170000p.pdf>. Accessed: 2022-10-6.

3GPP. **3GPP TS 26.443 version 17.0.0 Release 17.** 2022. <https://www.etsi.org/deliver/etsi_ts/126400/126499/126443/17.00.00_60/ts_126443v170000p.pdf>. Accessed: 2022-10-6.

3GPP. **3GPP TS 26.445 version 17.0.0 Release 17.** 2022. <https://www.etsi.org/deliver/etsi_ts/126400/126499/126445/17.00.00_60/ts_126445v170000p.pdf>. Accessed: 2022-10-6.

LUTZKY, M. S. M. **Enhanced Voice Service (EVS) Codec.** 2015. <http://pub.dega-kustik.de/DAGA_2015/data/articles/000517.pdf>. Accessed: 2023-8-14.

BRUHN, S.; EKSLER, E. A. B. V.; GIBBS, F. I. J. **Workshop at the 140th AES Convention 2016.** 2016. <<https://www.aes.org/technical/documentDownloads.cfm?docID=548>>. Accessed: 2023-8-14.

BRUHN, e. a. **Workshop at the 140th AES Convention 2016.** 2016. <<https://www.aes.org/technical/documentDownloads.cfm?docID=548>>. Accessed: 2023-4-23.

3GPP. **VoiceAge - codec Enhanced Voice Services (EVS).** 2021. <http://webcache.googleusercontent.com/search?q=cache:im2lWK_LP2YJ:www.voiceage.com/EVS.html+&cd=8&hl=pt-PT&ct=clnk&gl=br>. (Accessed on 03/07/2022).

FEASLEY, J. **Enhanced Voice Services (EVS) codec.** [S.l.]: VOCAL Technologies, 2021. <https://vocal-com.translate.google.com/speech-coders/enhanced-voice-services-codec/?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt-BR&_x_tr_pto=sc>. Accessed: 2023-7-10.

3GPP. **Codec para Serviços de Voz Aprimorados (EVS); Código ANSI C (ponto fixo) (3GPP TS 26.442 versão 14.0.0 Release 14)**. 2017. <https://www.etsi.org/deliver/etsi_ts/126400_126499/126442/14.04.00_60/ts_126442v140400p.pdf>. Accessed: 2022-9-20.

3GPP. **Código ANSI-C para o codec de voz Adaptive Multi-Rate - Wideband (AMR-WB)**. 2019. <https://www.etsi.org/deliver/etsi_ts/126100_126199/126173/15.01.00_60/ts_126173v150100p.pdf>. Accessed: 2022-9-25.

3GPP. **Enhanced voice services codec for LTE**. 2014. <<https://www.3gpp.org/news-events/3gpp-news/evs-news>>. Accessed: 2022-10-7.

VOICEAGE. **VoiceAge - Enhanced Voice Services (EVS) codec**. 2021. <<https://voiceage.com/EVS.html>>. Accessed: 2023-4-23.

Rodriguez, et. al. **Avaliação de um Algoritmo de Determinação da Taxa de Codificação baseado na qualidade do sinal de voz em uma rede IP emulada**. In: . [S.l.: s.n.], 2009.

Magro, et.al. **Estudo da Qualidade de Voz em Redes IP**. 2005. <<https://repositorio.unicamp.br/Busca/Download?codigoArquivo=468880>>. Accessed: 2023-7-10.

SOUSA, J. M. d. **Avaliação do índice de qualidade de voz em redes móveis celulares baseada na percepção dos usuários**. 2011.

RIBEIRO, D. A. et al. **Avaliação da qualidade perceptiva dos sinais de voz transmitidos sob diferentes condições de canal sem fio**. *Res. Soc. Dev.*, v. 10, n. 9, p. e27610918053, 2021.

JÚNIOR, B.; CARLOS, L. **Determinação de um modelo não intrusivo de qualidade de voz fundamentado na análise do sinal no domínio do tempo usando aprendizagem de máquina**. 2019.

BARBEDO, J. G. A. **AVALIAÇÃO OBJETIVA DE QUALIDADE DE CODECS DE VOZ NA FAIXA DE TELEFONIA**. 2001. <https://www.google.com.br/url?sa=i&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=0CAIQw7AJahcKEwjo-Keo_qmAAxUAAAAAHQAAAAAQAg&url=https%3A%2F%2Frepositorio.unicamp.br%2FBusca%2FDownload%3FcodigoArquivo%3D466722&psig=AOvVaw324SoOJDB4zzhzBBaGGinr&ust=1690371615778366&opi=89978449>. Accessed: 2023-7-25.

3GPP. **3GPP TS 26.441**. 2022. <https://www.etsi.org/deliver/etsi_ts/126400_126499/126441/17.00.00_60/ts_126441v170000p.pdf>. Accessed: 2022-9-19.

SW. **E-model**. 2020. <<https://www.itu.int/ITU-T/studygroups/com12/emodel/index.htm>>. Accessed: 2023-6-10.

JORDAN, H. **O codec 3GPP Enhanced Voice Services (EVS)**. 2016. <http://Nokia_3GPP_EVS_Codec_White_Paper_EN.pdf>. Accessed: 2023-7-26.

3GPP. **3GPP specification series: 26series**. 2022. <<https://www.3gpp.org/dynareport?code=26-series.htm>>. Acessado em 15/08/2023.

3GPP. **3GPP specification:**. 2017. <<http://www.3gpp.org/DynaReport/26441.htm>>. Accessed: 2023-4-26.

Qualcomm Technologies. **Enhanced Voice Services – EVS**. 2014. <http://www.iwaenc2014.org/files/2014_IWAENC_EVS_Varga.pdf>. Accessed: 2023-7-26.

G.192, I.-T. **G.192 : A common digital parallel interface for speech standardization activities**. 1996. <<https://www.itu.int/rec/T-REC-G.192-199603-I/en>>. Accessed: 2023-4-25.

- 3GPP. **Codec for Enhanced Voice Services (EVS) ANSI C code (fixed-point) (3GPP TS 26.442 version 14.4.0 Release 14)**. 2020. <https://www.etsi.org/deliver/etsi_ts/126400_126499/126442/14.04.00_60/ts_126442v140400p.pdf>. Accessed: 2022-9-22.
- 3GPP. **Codec for Enhanced Voice Services (EVS);Test sequences(3GPP TS 26.444 version 14.4.0 Release 14)**. 2020. <https://www.etsi.org/deliver/etsi_ts/126400_126499/126444/14.04.00_60/ts_126444v140400p.pdf> Accessed: 2022-9-22.
- 3GPP. **ITeh standards**. 2022. <<https://standards.iteh.ai/catalog/standards/etsi/2e86333b-a429-4fc9-8846-0987ae877232/etsi-ts-126-174-v17-0-0-2022-05>>. Accessed: 2022-9-25.
- CHENG, e. **MASS: Microphone Array Speech Simulator in Room Acoustic Environment for Multi-Channel Speech Coding and Enhancement**. *Applied Sciences*, v. 10, p. 1484, 02 2020.
- Rec. ITU-T. **P.800 : Methods for subjective determination of transmission quality**. 1996. <<https://www.itu.int/rec/T-REC-P.800/en>>. Accessed: 2023-5-3.
- ITU-T, G.107. **The narrowband E-model G.107 (06/2015)**. 2015. <<https://www.itu.int/ITU-T/studygroups/com12/emodel/calculnb.php>>. Accessed: 2023-6-11.
- G107, I.-T. **E-MODEL-BASED VOICE QUALITY ESTIMATION (VoIP)**. 2007. <<http://what-when-how.com/voip/e-model-based-voice-quality-estimation-voip/>>. Accessed: 2023-5-6.
- G.107, I.-T. **G.107: O modelo E: um modelo computacional para uso no planejamento de transmissão**. 2015. <<https://www.itu.int/rec/T-REC-G.107-201506-I>>. Accessed: 2023-5-9.
- BERNARDINO, D. **Correlação de Pearson: de que trata esse coeficiente?** 2019. <<https://www.questionpro.com/blog/pt-br/correlacao-de-pearson/>>. Accessed: 2023-6-4.
- OLIVEIRA, B. **Coeficientes de correlação: Para que servem e como interpreta-los?** 2019. <<https://statplace.com.br/blog/coeficientes-de-correlacao/>>. Accessed: 2023-6-11.
- ITU-T, R. **G.107: O modelo E: um modelo computacional para uso no planejamento de transmissão**. 2015. <<https://www.itu.int/rec/T-REC-G.107>>. Accessed: 2023-7-14.
- G., R. **G.113: Novo Apêndice V - Valores provisórios de planejamento para o fator de comprometimento do equipamento de banda completa e o fator de robustez de perda de pacote de banda completa**. 2019. <<https://www.itu.int/rec/T-REC-G.113-201905-I!Amd2>>. Accessed: 2023-7-14.
- G., I.-T. **G.107: O modelo E: um modelo computacional para uso no planejamento de transmissão**. 2015. <<https://www.itu.int/rec/T-REC-G.107>>. Accessed: 2023-7-14.
- Nguyen,et. al. **A new non-intrusive model for measuring VoLTE quality based on wideband E-model**. v. 5, n. 8, p. 1–8, 2016.