



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Avaliação da Qualidade de Voz do Serviço VoIP em Sistemas HSDPA

Autor
Leonardo Ramon Nunes de Sousa

FORTALEZA – CEARÁ
SETEMBRO 2007



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Avaliação da Qualidade de Voz do Serviço VoIP em Sistemas HSDPA

Autor

Leonardo Ramon Nunes de Sousa

Orientador

Prof. Dr. Francisco Rodrigo Porto Cavalcanti

Dissertação apresentada à Coordenação do Programa de Pós-graduação em Engenharia de Teleinformática da Universidade Federal do Ceará como parte dos requisitos para obtenção do grau de Mestre em Engenharia de Teleinformática.

FORTALEZA – CEARÁ

SETEMBRO 2007

Resumo

Nos últimos anos, observa-se o surgimento e a rápida disseminação do serviço VoIP, integrando-se ao mercado atual junto à telefonia convencional e às redes celulares. Por ser uma alternativa tecnológica que contribui para minimizar custos, assiste-se a uma preferência crescente por fazer trafegar a voz através das redes IP. O HSDPA, como sistema celular, permite a transmissão de dados em alta velocidade, aumenta a largura de banda da rede e abre novas possibilidades de serviços multimídia, como o VoIP que utiliza a transmissão em banda larga para telefones móveis. Exige-se, porém, um considerável esforço de análise deste serviço, pois o atraso inerente a esse sistema é um desafio para a garantia de qualidade de voz. Estes fatos justificam, conseqüentemente, um esforço de análise que se detenha sobre a qualidade de voz no VoIP sobre o HSDPA.

Para avaliar a qualidade de voz, neste estudo aplica-se o método MOS, que faz corresponder valores numéricos a categorias como medidas de qualidade e inteligibilidade da voz transmitida, obtendo-se esses dados de forma objetiva e subjetiva. O processo de avaliação dividiu-se em etapas de acordo com cada metodologia, seguindo recomendações técnicas e através de simulações computacionais dinâmicas. Na avaliação objetiva, utilizou-se o algoritmo PESQ para obtenção do conceito MOS, enquanto que na avaliação subjetiva, arquivos de voz com certo percentual de erro foram colocados em um endereço na Internet para escuta e atribuição de nota MOS, baseada na percepção do usuário ouvinte.

Os resultados mostraram que os dois métodos de avaliação obtiveram conceitos de qualidade satisfatórios, que o QoS nas simulações é estável e positivo, que uma boa qualidade para os arquivos de voz e a provável satisfação dos usuários do serviço VoIP sobre o sistema celular HSDPA é garantida para uma taxa de 2% de FER. Finalmente, mostra-se que a metodologia objetiva garante a obtenção de

notas MOS aproximadas da subjetiva, evitando o árduo trabalho de fazerem-se ouvir diversos arquivos de voz por uma quantidade significativa de usuários para ser válida estatisticamente.

Palavras-Chave: Avaliação da qualidade de voz, VoIP, MOS, PESQ, HSDPA, Qualidade de serviço.

Abstract

In recent years, we can observe the development and fast dissemination of VoIP services, being integrated by the present market, beside conventional telephony and cellular networks. For being a technological alternative that contributes to minimize costs, we see an increasing preference for voice transmission through IP networks. HSDPA, as a cellular system, allows high speed data transmission, increases the network width of band and creates new possibilities for multimedia services, as VoIP that transmits in wideband to mobile telephones. The delay inherent to this system, however, is a challenge for the need to assure good quality of voice transmissions, demanding a considerable effort of analysis of this service. These facts justify a study that focus on the quality of voice in VoIP over HSDPA.

To evaluate the voice quality, in this study, we applied MOS method that makes numerical values correspond to categories like quality and intelligibility of transmitted voice, getting these data through objective and subjective methodologies. The evaluation process was divided in fases according to the characteristics of each methodology and to technical recommendations, and was done through dynamic computational simulations. For objective evaluation process, algorithm PESQ was employed to obtain MOS concepts, whereas, for subjective evaluation, voice files with a percentage of error have been placed in Internet for listening and for the attribution of MOS concepts based in the perception of the listener.

The results of this research show that both evaluation methods got satisfactory concepts of quality, that QoS is steady and positive in the simulations, that a good quality for the voice files and the probable satisfaction of the users of the VoIP service on cellular system HSDPA is guaranteed for 2% FER rate. Finally, it shows that MOS concepts produced by objective methodology were close enough to those

given by subjective evaluation to dispense with the arduous work of making diverse voice files to be heard and subjectively evaluated by a statistically valid amount of users.

Index Terms: Evaluate the voice quality, VoIP, MOS, PESQ, HSDPA, Quality of service.

Agradecimentos

Primeiramente , à Deus e à Virgem Maria. Aos meus pais Antônio Domingos e Maria das Mercês, aos irmãos César, Ghaleno e Marília, às cunhadas Soraya e Lana, aos sobrinhos Maria Luiza, João Pedro, Mariah e Ian. À minha família em geral, por tudo que vivemos, pelas palavras e gestos de incentivo e apoio. Ao orientador e professor Dr. Francisco Rodrigo Porto Cavalcanti, pela disponibilidade e pela oportunidade de pesquisa e estudo. Ao Professor Dr. João César Moura Mota, pela ajuda junto à pós-graduação e pela troca de idéias. Aos amigos Janayna Passarinho, Marcelo Damasceno, Marcone Carvalho, Adriana Piauilino, Juliana Pinheiro, Livia Parente, Paulo Chaves, Valery Nicolas, Charles Barbosa, Patrick Quelemes, Fabrício Silva, Leonardo Horn e a todos os amigos membros do Movimento dos Focolares, meu muito obrigado. Aos professores da UFPI por me incentivar a fazer o Mestrado, em particular a Francisco Vieira, Magno Alves e Francisco Araújo. Aos amigos e colegas do Grupo de Pesquisa em Telecomunicações Sem Fio (GTEL), por me facilitarem o acesso à estrutura física desse laboratório e o uso do cluster computacional, pelos momentos profissionais, de ajuda e de descontração, em particular a Tarcísio Maciel, Yuri Carvalho, Marcone Carvalho, Emanuel Rodrigues, Martin Obando, Leonardo Sampaio, Emanuela Brandão, Carlos Héracles, Vicente Ângelo, Júlio Pimentel, André Braga, Raimundo Neto, Fabiano Chaves, Fábio Freitas, Rafael Marques, Cibelly Araújo, Elvis Miguel, Guilherme Barreto, Charles Cavalcante, Alexandra Medeiros, Darlan Moreira, Alex Silva, Jean Maciel, Aurélio Carvalho, Daniel Ney, Ana Livia, Williane Ferreira, Vera e Sérgio. Meus agradecimentos também à Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) pelo suporte financeiro. Agradeço a todos os professores, alunos e colegas do Departamento de Engenharia de TeleInformática da Universidade Federal do Ceará que estiveram ligados à minha formação.

Sumário

Lista de Figuras	x
Lista de Tabelas	xi
Lista de Acrônimos	xii
1 Introdução	1
1.1 Evolução das Redes de Comunicações Móveis	2
1.2 Introdução ao Tráfego de Voz sobre o Protocolo <i>Internet</i>	3
1.3 Exposição do Problema, Motivação e Objetivos	5
1.4 Revisão da Literatura	6
1.5 Metodologia Empregada	10
1.6 Contribuições e Produção Científica	11
1.7 Organização da Dissertação de Mestrado	12
1.8 Sumário	12
2 Voz sobre o Protocolo Internet	15
2.1 Codificadores de Voz	16
2.1.1 Codificador <i>Adaptive Multi-Rate</i> (AMR)	17
2.2 Protocolos para o Tráfego de VoIP	17
2.2.1 Protocolo de Tempo Real (RTP)	18
2.2.2 Protocolo de Controle de Tempo Real (RTCP)	21
2.2.3 Protocolo de Datagrama de Usuário (UDP)	22
2.2.4 Protocolo Internet (IP)	23
2.2.5 Compressão de Cabeçalho (HC)	23
2.3 Qualidade de Serviço para Voz	24
2.3.1 Avaliação da Qualidade de Voz pelos Métodos Subjetivo e Objetivo	26
2.3.2 Perda de Quadros	28
2.3.3 Atraso (<i>Delay</i>)	28
2.3.4 Variação do Atraso (<i>Delay Jitter</i>)	28
2.4 Sumário	29

3	Rede de Acesso de Rádio do UMTS	31
3.1	<i>Universal Mobile Telecommunications System</i> (UMTS)	32
3.1.1	Arquitetura do UMTS	33
3.1.2	Classes de Serviço do UMTS	35
3.2	<i>High Speed Downlink Packet Access</i> (HSDPA)	37
3.2.1	Arquitetura do HSDPA	42
3.2.2	Camada Física do HSDPA	44
3.2.3	Considerações Finais sobre o HSDPA	49
3.3	Sumário	49
4	Avaliação da Qualidade de Voz do Serviço VoIP em Sistemas HSDPA	51
4.1	Considerações sobre a oferta do serviço VoIP sobre HSDPA	52
4.2	Modelagem para Simulação	53
4.2.1	Pilha de Protocolos no UTRAN-SIM	54
4.2.2	Rede HSDPA no UTRAN-SIM	56
4.2.3	Parâmetros de Simulação	57
4.2.4	Arquivos de Voz	58
4.2.5	Estrutura das Avaliações Objetiva e Subjetiva	60
4.3	Resultados de Desempenho	62
4.3.1	Resultados da Avaliação Objetiva	63
4.3.2	Resultados da Avaliação Subjetiva	75
4.3.3	Comparação dos Resultados entre os Métodos Objetivo e Subjetivo	79
4.4	Sumário	81
5	Conclusões e Perspectivas	83
5.1	Conclusões	83
5.2	Perspectivas	85
	Apêndice A Página na <i>Internet</i> utilizada para Avaliação Subjetiva	87
	Apêndice B Produção Científica	91
	Referências Bibliográficas	108

Lista de Figuras

2.1	Transporte da voz no Protocolo IP.	16
2.2	Cabeçalho de um pacote RTP	19
2.3	Cabeçalho UDP.	22
2.4	Modelo para Qualidade de Serviço (QoS).	25
3.1	Arquitetura básica de uma rede UMTS [28].	34
3.2	Esquema de Modulação usado com o HSDPA [28].	40
3.3	Arquitetura de Protocolo no HSDPA [80].	43
3.4	Códigos alocados para o HS-PDSCH.	45
3.5	Compartilhamento do HS-PDSCH por multiplexação de códigos e de tempo.	46
3.6	Relação de tempo entre os canais HS-PDSCH e HS-PSCCH.	48
4.1	Fluxo da Informação na Pilha de Protocolos.	55
4.2	Sinal da Voz Digitalizada do Arquivo Homem 1. O eixo das abscissas indica a duração dos trechos de voz em segundos e o eixo das ordenadas reflete a amplitude do sinal, expressa em decibéis (dB), retratando a variação da frequência da voz para baixo e para cima.	59
4.3	Sinal da Voz Digitalizada de Homem 2	59
4.4	Sinal da Voz Digitalizada de Mulher 1	60
4.5	Sinal da Voz Digitalizada de Mulher 2	60
4.6	Módulos da Estrutura de Avaliação.	61
4.7	Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Homem 1, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.	64
4.8	Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Homem 2, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.	65
4.9	Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Mulher 1, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.	66

4.10	Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Mulher 2, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.	67
4.11	Gráfico das taxas de FER para o Arquivo Homem 1 em cada tempo entre chegadas de chamadas, revelando-se a frequência de erros em cada tempo.	68
4.12	Gráfico do arquivo Homem 1 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos, cuja legenda mostra que Pontos significam as taxas de quadros perdidos (FER), que Média significa o gráfico que une os pontos médios de nota MOS para cada valor de FER, que Máximo trata dos gráficos para os valores máximos e que Mínimo nomeia o gráfico dos valores mínimos. Todos eles com seus respectivos tempos de chegada.	69
4.13	Gráfico do arquivo Homem 2 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.	70
4.14	Gráfico do arquivo Mulher 1 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.	70
4.15	Gráfico do arquivo Mulher 2 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.	71
4.16	Arquivos das Vozes Avaliadas expressos variando apenas o tempo de chegada entre os usuários.	72
4.17	Arquivos de voz representando a média geral a partir da Figura 4.16.	73
4.18	Sinais Digitalizados dos Segmentos do Arquivo Homem 1 no formato de extensão <i>wav</i> , que é capaz de ser escutado por qualquer software de voz.	75
4.19	Segmentos do Arquivo Homem 2.	76
4.20	Segmentos do Arquivo Mulher 1.	76
4.21	Segmentos do Arquivo Mulher 2.	77
4.22	Histogramas que mostram as Distribuições das Notas Atribuídas a cada Segmento de Arquivo pela Avaliação Subjetiva.	78
4.23	Média das Notas Subjetivas para um Intervalo de Confiança de 95% seguida dos Limites Superior e Inferior [90].	79
4.24	Histograma da Média das Notas MOS dadas pelas Avaliações Subjetiva e Objetiva.	81
A.1	Parte Superior da Página na Internet.	88
A.2	Parte Inferior da Página na Internet.	89

Lista de Tabelas

2.1	Escala de notas da avaliação subjetiva MOS.	27
3.1	Classes de Serviço do UMTS.	36
3.2	Taxas de bits máximas para 15 códigos paralelos.	46
3.3	Comparação entre HS-PDSCH e PDSCH.	47
4.1	Parâmetros de Simulação.	57
4.2	Valores MOS(PESQ) para 2% de FER das Médias de cada Arquivo de Voz na Avaliação Objetiva, onde percebe-se que todos os resultados indicam uma percepção subjetiva equivalente entre "Aceitável" e "Bom".	74
4.3	Valores MOS das Médias de cada Arquivo de Voz na Avaliação Subjetiva para 2% de FER.	79
4.4	Valores MOS(PESQ) e MOS das Médias de cada Arquivo de Voz nas Avaliações Objetiva e Subjetiva seguido da diferença em termos percentuais de nota MOS entre o valor subjetivo e o objetivo para cada Arquivo.	80

Lista de Acrônimos

16-QAM	<i>16 Quadrature Amplitude Modulation</i>
1G	<i>1rd Cellular Generation</i>
2G	<i>2rd Cellular Generation</i>
3G	<i>3rd Cellular Generation</i>
3GPP	<i>3rd Generation Partnership Project</i>
64-QAM	<i>64 Quadrature Amplitude Modulation</i>
8PSK	<i>8 Phase Shift Keying</i>
ABC	<i>Always Best Connected</i>
ACK	<i>Acknowledge</i>
AICH	<i>Acquisition Indicator Channel</i>
AM	<i>Acknowledge Mode</i>
AMC	<i>Modulação e Codificação Adaptativas</i>
AMPS	<i>Advanced Mobile Phone Service</i>
AMR	<i>Adaptive Multi-Rate</i>
APP	<i>Application-specific RTCP packet</i>
ARIB	<i>Association of Radio Industries and Business</i>
ARQ	<i>Automatic Repeat Request</i>
AS	<i>Active Set</i>
ATM	<i>Asynchronous Transfer Mode</i>
BCS	<i>Best Channel Scheduling</i>
CC	<i>Contributing Source Count</i>
CDMA	<i>Code Division Multiple Access</i>
CN	<i>Core Network</i>
C RTP	<i>Compressed Real-Time Protocol</i>
CSRC	<i>Contributing Source Identifiers</i>

CQI	<i>Channel Quality Indicator</i>
CRC	<i>Cyclic Redundancy Check</i>
DPCH	<i>Downlink Dedicated Physical Channel</i>
DS-CDMA	<i>Direct-Sequence CDMA</i>
EDGE	<i>Enhanced Data Rates for GSM Evolution</i>
EM	<i>Estação Móvel</i>
ERB	<i>Estação Rádio-Base</i>
ETSI	<i>European Telecommunications Standards Institute</i>
FCSS	<i>Fast Cell Site Selection</i>
FEC	<i>Forward Error Correction</i>
FDDI	<i>Fiber Distributed Data Interface</i>
FER	<i>Frame Error Rate</i>
GERAN	<i>GSM/EDGE Radio Access Network</i>
GPRS	<i>General Packet Radio Service</i>
GSM	<i>Global System for Mobile Communications</i>
GTEL	<i>Wireless Telecommunications Research Group</i>
HARQ	<i>Hybrid Automatic Repeat Request</i>
HC	<i>Header Compression</i>
HSDPA	<i>High Speed Downlink Packet Access</i>
HS-DSCH	<i>High Speed Downlink Shared Channel</i>
HS-DPCCH	<i>High Speed Dedicated Physical Control Channel</i>
HS-PDSCH	<i>High Speed Physical Downlink Shared Channel</i>
HS-PSCCH	<i>High Speed Physical Shared Control Channel</i>
IETF	<i>Internet Engineering Task Force</i>
iLBC	<i>internet Low Bitrate Codec</i>
IMT	<i>International Mobile Telecommunications</i>
IS-54	<i>US Interim Standard 54</i>
IS-95	<i>US Interim Standard 95</i>
IP	<i>Internet Protocol</i>
ITU	<i>International Telecommunication Union</i>
ITU-T	<i>International Telecommunications Union-Telecommunication</i>

MAC	<i>Medium Access Control</i>
MAC-d	<i>Medium Access Control - Dedicated Transport Channels</i>
MAC-hs	<i>Medium Access Control - High Speed Shared Transport Channel</i>
ME	<i>Mobile Equipment</i>
MOS	<i>Mean Opinion Score</i>
MSC	<i>Mobile Switching Center</i>
NACK	<i>Not Acknowledge</i>
PDU	<i>Packet Data Unit</i>
PESQ	<i>Perceptual Evaluation of Speech Quality</i>
PF	<i>Proportional Fair</i>
PSTN	<i>Public Switched Telephone Network</i>
PT	<i>Payload Type</i>
QoS	<i>Quality of Service</i>
QPSK	<i>Quadrature Phase Shift Keying</i>
RAN	<i>Radio Access Network</i>
RLC	<i>Radio Link Control</i>
RNC	<i>Radio Network Controller</i>
RNS	<i>Radio Network Subsystem</i>
ROCCO	<i>Robust Checksum-Based Header Compression</i>
ROHC	<i>Robust Header Compression</i>
RR	<i>Receiver Reports</i>
RR	<i>Round Robin</i>
RTP	<i>Real-Time Protocol</i>
RTCP	<i>Real-Time Control Protocol</i>
SDES	<i>Source Description</i>
SF	<i>Fator de Espalhamento</i>
SGSN	<i>Service GPRS Support Node</i>
SMS	<i>Short Message System</i>
SR	<i>Sender Report</i>
SSRC	<i>Synchronization Source Identifier</i>
TCP	<i>Transmission Control Protocol</i>
TDMA	<i>Time Division Multiple Access</i>
TTI	<i>Transmission Time Interval</i>

TM	<i>Transparent Mode</i>
UDP	<i>User Datagram Protocol</i>
UE	<i>User Equipment</i>
UM	<i>Unacknowledged Mode</i>
UMTS	<i>Universal Mobile Telecommunications System</i>
URL	<i>Uniform Resource Locator</i>
USIM	<i>User Service Identity Module</i>
UTRAN	<i>UMTS Terrestrial Radio Access Network</i>
UTRANSIM	<i>UTRAN Simulator</i>
VoATM	<i>Voice over Asynchronous Transfer Mode</i>
VoFR	<i>Voice over Frame Relay</i>
VoIP	<i>Voice over Internet Protocol</i>
VHE	<i>Virtual Home Environment</i>
VLR	<i>Visitor Location Register</i>
WCDMA	<i>WideBand Code Division Multiple Access</i>
WRR	<i>Weighted Round-Robin</i>

Introdução

Uma vez que a telefonia é uma tecnologia eficiente, atingindo grande parte da humanidade, sua importância continua a ser vital para as telecomunicações. Nos dias atuais, entretanto, as redes de comutação de pacotes vêm mostrando um grande crescimento no número de usuários, evidenciando a importância da Internet, que se baseia no IP (do inglês *Internet Protocol*).

A estrutura da Internet é cada vez mais utilizada para tráfego de mídias contínuas, como voz e vídeo. As redes IP, porém, estão sujeitas a problemas que podem comprometer a qualidade final da voz transmitida, como o atraso e mesmo a perda de pacotes durante o transporte, causados seja por congestionamentos, seja por erros durante a transmissão ou no momento do uso do canal de rádio para tráfego, em redes sem fio. Esses problemas tornam-se mais graves quando se trata de transmitir mídias especialmente sensíveis ao atraso, como é o caso da voz. Tais problemas ocorrem porque o protocolo Internet oferece um serviço sem conexão, não alocando um caminho fixo para um determinado pacote de dados trafegar na rede e, portanto, não garantindo requisitos temporais para os pacotes transmitidos.

Apesar dessas limitações, o aparecimento das primeiras aplicações que usavam o serviço de voz sobre o protocolo Internet (VoIP, do inglês *Voice over IP*) fez com que a Internet realmente passasse a ser utilizada como a forma mais econômica de transportar voz ou vídeo. De fato, o VoIP está contribuindo para minimizar custos, provendo alternativas às formas anteriores de transmissão de voz. Por ser uma tecnologia relativamente barata, possibilitando considerável economia para as redes telefônicas particulares de corporações de um certo porte, assiste-se a uma preferência crescente por fazer trafegar a voz através das redes IP. Estes fatos justificam um esforço de análise que se detenha sobre a qualidade e as limitações

atuais desse meio de transmissão da voz.

Nesta introdução, apresenta-se, inicialmente, na seção 1.1, um breve histórico da evolução dos sistemas de comunicações celulares, apenas com a finalidade de contextualizar o tráfego de voz sobre IP no conjunto dos sistemas móveis de comunicações. A seção 1.2 consiste numa introdução ao tráfego de voz sobre IP. Uma exposição da problemática que motiva este trabalho e de seus objetivos encontra-se na seção 1.3. Em seguida, na seção 1.4, apresenta-se uma revisão da literatura atual sobre esta temática com o objetivo de mostrar quais fontes contribuíram, mais diretamente, com este estudo. A metodologia empregada na pesquisa está exposta na seção 1.5. Na seção 1.6, expõem-se as possíveis contribuições deste estudo juntamente com a produção científica gerada e publicada durante o processo de pesquisa. Finalmente, na seção 1.7, descreve-se o conjunto desta Dissertação de Mestrado, por meio do conteúdo de cada um de seus capítulos.

1.1 Evolução das Redes de Comunicações Móveis

As redes de comunicações móveis celulares passaram por alguns estágios de evolução desde o seu aparecimento [1, 2]. A chamada primeira geração dessas redes utilizava esquemas de modulação analógica, provendo apenas o serviço de voz, além de ter capacidade e qualidade de serviço limitadas. Dava, porém, o passo inicial para prover serviços de voz com mobilidade em larga escala. Como exemplo de sistemas de primeira geração, temos o Sistema Avançado de Telefonia Móvel (AMPS, do inglês *Advanced Mobile Phone Service*).

Uma segunda geração surge com o aparecimento dos sistemas celulares digitais, apresentando uma capacidade de serviço mais elevada. Os sistemas 2G apresentavam, ainda, as características de cobertura mais ampla, uso de encriptação da informação transmitida, melhor qualidade de voz e a adição de serviços de dados, baseados em chaveamento de circuito e com baixa taxa de transmissão, como por exemplo, o serviço de mensagem simples (SMS, do inglês *Short Message System*). Essa geração conquistou um aumento significativo de usuários no mundo todo, popularizando a tecnologia celular. O GSM (do inglês *Global System for Mobile Communications*), o IS-136 TDMA (do inglês *Time Division Multiple Access*) e o IS-95 CDMA (do inglês *Code Division Multiple Access*) são exemplos dos sistemas dessa geração.

A procura por novos serviços, tais como a transmissão de dados e a integração das redes celulares com a Internet, foi surgindo já no sistema 2G, porém os serviços

por chaveamento de circuitos eram ineficientes para o tráfego de dados de natureza não contínua, fazendo surgir a geração 2.5G, com serviços de dados por chaveamento de pacotes. O GPRS (do inglês *General Packet Radio Service*) e CDMA2000 1x são os principais representantes dessa geração, permitindo o desejado acesso à Internet pela rede móvel, porém ainda com taxas limitadas.

Para prover os novos serviços exigidos e a compatibilidade entre os sistemas atuais, garantir a provisão de serviços multimídia e o acesso irrestrito à Internet, aumentar as taxas de transmissão, a capacidade, a cobertura e o nível de qualidade de serviço (QoS, do inglês *Quality of Service*), criaram-se os sistemas considerados de terceira geração. Os principais exemplos dessa linha de sistemas são o EDGE (do inglês *Enhanced Data Rates for GSM Evolution*), o WCDMA (do inglês *WideBand CDMA*), o HSDPA (do inglês *High Speed Downlink Packet Access*) e o CDMA2000 3x.

O HSDPA, especificamente, é um sistema 3G que transmite pacotes de dados operando dentro do UMTS¹/W-CDMA, no enlace direto (*downlink*), permitindo a transmissão de dados em alta velocidade, melhorando a eficiência espectral do sistema e do código, além de ser uma tecnologia eficaz que aumenta a largura de banda da rede. O HSDPA, por essas características, abre novas possibilidades de serviços multimídia, como o VoIP, que utilizam a transmissão em banda larga em telefones móveis.

Uma quarta geração de sistemas está surgindo atualmente, caracterizada pela integração e pela cooperação entre diversas redes de acesso por rádio, como GERAN (do inglês *GSM/EDGE Radio Access Network*) e UTRAN (do inglês *UMTS Terrestrial Radio Access Network*). Características típicas dessa mais recente geração de sistemas são a adaptação da rede ao melhor cenário e a garantia da melhor conexão (ABC, do inglês *Always Best Connected*).

1.2 Introdução ao Tráfego de Voz sobre o Protocolo *Internet*

A tecnologia Voz sobre IP (VoIP) foi criada através de pesquisas feitas por empresas privadas, que tinham como principal objetivo fabricar hardware para telefonia e que experimentaram, no primeiro momento, fazer trafegar a voz utilizando uma rede de dados que possuía um protocolo proprietário. A comunicação fazia-se através de dois computadores clientes. Com a evolução das pesquisas, esses métodos

¹Sistema Universal de Telecomunicações Móveis (UMTS, do inglês *Universal Mobile Telecommunication System*)

separaram-se em VoATM (voz sobre ATM, do inglês *Asynchronous Transfer Mode*) e em VoFR (voz sobre *Frame Relay*), evoluindo, em seguida para VoIP, em resposta ao grande crescimento da Internet, passando-se a utilizar uma rede com protocolo TCP/IP (do inglês *Transmission Control Protocol/ Internet Protocol*) que tornava possível fazer trafegar a voz, ainda que essa tecnologia não tivesse um protocolo proprietário [3].

O VoIP é uma tecnologia que permite a digitalização de voz e seu empacotamento na forma de dados IP para a transmissão por uma rede que utilize os protocolos TCP/IP, tornando possível a realização de chamadas telefônicas com qualidade pela Internet. Antes que essa tecnologia se tornasse amplamente viável, porém, foi necessário investir na qualidade de serviço (QoS), optando-se, dentre as soluções possíveis, pelo aumento da largura da banda, que permite o aumento da velocidade de transmissão e de recepção dos dados.

Cumprindo com essas etapas, a telefonia veio a tornar-se uma das aplicações que mais crescem na Internet. Esse crescimento está ligado ao ganho de benefícios para o usuário em comparação com a rede de telefonia tradicional como a redução de custos, a possibilidade de complementação da comunicação de voz por meio de recursos multimídia e um menor consumo de largura de banda.

O tráfego de voz sobre o Protocolo Internet vem provocando uma revolução pois, além de surgir como alternativa à telefonia tradicional, permite a utilização de uma mesma rede para telefonia e tráfego de dados, abrindo a possibilidade de comunicação de voz pela Internet a partir tanto de aparelhos telefônicos comuns quanto de computadores.

Desta forma, uma simples plataforma de serviço e transporte de dados baseada em IP oferecerá também aos usuários móveis uma grande variedade de serviços interativos e em tempo real. Usar o protocolo Internet como provedor de serviços implica em que estes sejam serviços baseados totalmente em IP, tendo como principal vantagem a sua flexibilidade por não haver mais dependência de cada aplicação a diferentes tipos de rede, de modo que se alarga e facilita-se o desenvolvimento, por qualquer criador capacitado, de novas aplicações utilizáveis por um amplo espectro de usuários [4, 5].

O QoS é a medida de qualidade de serviço na rede. A qualidade de serviço nas redes IP depende de um conjunto de requisitos para um bom desempenho das aplicações em VoIP, do começo ao fim das transmissões. É necessário um reconhecimento e uma adaptação dos mecanismos, parâmetros, algoritmos e

protocolos utilizados pelos transmissores e receptores para obter-se um QoS com bom resultado em cada transmissão na rede. Exigem-se determinados parâmetros como requisitos de QoS, referentes a atrasos, vazão, perdas de pacotes, eco e largura de banda, dentre outros. Solicitando o QoS, as aplicações definem quais parâmetros devem ser garantidos para que possam ser executadas com qualidade.

A qualidade de uma conexão VoIP depende diretamente da performance, da disponibilidade e dos recursos da rede de dados utilizada. Por isto, o QoS tem o objetivo de priorizar o tráfego interativo, sensível a retardamentos, em detrimento do tráfego referente à transferência de arquivos, que não é sensível a esse fator. O serviço que usa o VoIP sobre o sistema HSDPA, tem exigido um considerável esforço de análise, pois o atraso inerente a esses sistemas é um desafio para a garantia de qualidade de voz.

Para avaliar a qualidade de voz transmitida em redes, são aplicados alguns métodos, entre eles o MOS (do inglês *Mean Opinion Score*) que faz corresponderem valores numéricos em categorias como medidas de qualidade e inteligibilidade da voz transmitida, podendo-se obter tais referências de forma objetiva e subjetiva.

1.3 Exposição do Problema, Motivação e Objetivos

A natureza do serviço de voz e de VoIP torna-o intolerante para com o atraso na transmissão dos pacotes. Por isto mesmo, o desempenho dos serviços é medido mais precisamente em termos de taxa de perdas de pacotes e exige-se a pesquisa de medidas que evitem essas perdas.

Em sistemas móveis de comunicação, muitos erros de transmissão e recepção podem ocorrer devido à própria natureza dessas redes, pois trabalham com alta frequência de pacotes descartados e alta variação de suas características, em determinados momentos. No sistema HSDPA, que cumpre com o requisito de prover qualidade de serviço, pacotes de voz podem obedecer com maior rigor ao requerimento de atraso evitando-se, em parte, as perdas, embora ainda permaneçam inúmeras formas ou causas de descarte de pacotes.

Nesta dissertação, o foco está na obtenção, objetivamente e subjetivamente, de conceitos MOS como medidas da inteligibilidade da voz transmitida no serviço VoIP sobre HSDPA. Como resultado secundário, pode-se verificar a relação entre algumas métricas de saída do simulador e suas relações com a quantidade de pacotes perdidos e com seu conseqüente impacto na qualidade final da voz trafegada.

O estudo de medição da qualidade de serviço será feito através dos métodos

objetivo e subjetivo para o tráfego de voz sobre o protocolo Internet em um sistema celular, neste caso o sistema HSDPA.

De posse das informações assim colhidas, este trabalho cumprirá com seu Objetivo Geral:

- ▶ Avaliar a qualidade final da voz transmitida no serviço VoIP sobre o sistema HSDPA, através dos métodos objetivo MOS (PESQ, do inglês *Perceptual Evaluation of Speech Quality*) e subjetivo MOS.

Cumprirá, também, com seus Objetivos Específicos:

- ▶ Comparar os resultados entre os métodos objetivo e subjetivo, averiguando se ambos chegam a valores aproximados em relação a notas MOS;
- ▶ Verificar se a metodologia objetiva garante resultados numéricos de MOS próximos aos que são obtidos através da metodologia subjetiva, com variação para mais ou para menos;
- ▶ Determinar qual é a taxa de quadros perdidos (FER, do inglês *Frame Error Rate*), que pode ser generalizada de forma a garantir-se uma boa qualidade de voz.

Como Metas Intermediárias deste estudo, têm-se:

- ▶ A obtenção de módulos de codificação e decodificação de arquivos de voz para o serviço VoIP;
- ▶ Uma melhor compreensão dos protocolos necessários ao VoIP e das características relevantes do sistema HSDPA;
- ▶ O desenvolvimento de uma ferramenta de simulação de VoIP sobre o sistema HSDPA.

1.4 Revisão da Literatura

Muitos trabalhos científicos são dedicados ao serviço VoIP. Alguns deles forneceram o embasamento teórico para o estudo aqui realizado e ofereceram exemplos úteis para o desenvolvimento deste Trabalho de Dissertação de Mestrado, que serão explicitados nos parágrafos seguintes.

Em um desses exemplos [6], propõe-se um procedimento para medir a qualidade de serviço, colocando como resultado medidas subjetivas baseadas em MOS. Esses resultados foram obtidos a partir de arquivos de voz gravados em língua finlandesa, de acordo com [7], nas taxas 7.95 e 12.2 Kbps do codificador Adaptativo de Taxas Variáveis (AMR, do inglês *Adaptive Multi-Rate*), utilizando 8 arquivos de voz (4 masculinas e 4 femininas), com duração de 8 segundos cada um, para avaliação subjetiva por um grupo de aproximadamente 28 pessoas. Considerou-se como nota final a média das notas dadas por todos os avaliadores, que foram escolhidos entre não-especialistas em avaliar vozes, com o intuito de refletir mais fielmente a opinião de usuários comuns de telefone. Esses ouvintes utilizaram fones de ouvido de boa qualidade em um ambiente de laboratório, propício para avaliação, sem interferências ou influências externas que pudessem comprometer notas dadas. Os resultados de medidas subjetivas foram expressos em termos de "pouco atraso" e "médio atraso" na transmissão da voz, sendo que a taxa AMR de 12.2 kbps obteve uma nota MOS um pouco melhor do que a de 7.95 kbps, além de revelar uma correlação entre a Nota MOS e a taxa de perda de pacotes (FER). Verificou-se que com o aumento da FER há uma diminuição da MOS e que os arquivos transmitidos com pouco atraso obtiveram notas maiores para a MOS.

Outra abordagem importante é feita em [8], no qual o autor compara o desempenho de alguns codificadores de voz padrões da ITU-T (do inglês *International Telecommunications Union-Telecommunication*)/IETF (do inglês *Internet Engineering Task Force*) mais utilizados onde ocorre uma transmissão com perdas de pacotes como no serviço de voz sobre redes de pacotes, particularmente o serviço de voz sobre IP. Cada um dos codificadores foi testado com os seus respectivos decodificadores utilizando o algoritmo de reconstrução de quadros perdidos contidos nas recomendações através de simulações em que blocos de dados que representavam quadros de sinais de voz foram apagados simulando efeitos das perdas de pacotes em uma rede IP. Nesta abordagem, verificou-se que há um impacto na qualidade do sinal final decodificado por um algoritmo objetivo especificado na recomendação ITU-T P.832 quando usa-se cada um desses diferentes decodificadores. Este algoritmo objetivo apresenta uma alta correlação com avaliações subjetivas MOS de acordo com alguns testes de outras fontes e ele também será descrito nesta Dissertação de Mestrado nos próximos capítulos. Outra investigação realizada ainda por este autor foi uma análise para avaliar o fator que mais afetaria na qualidade do sinal de voz decodificado após sua transmissão, chegando à conclusão que os principais fatores

foram a baixa qualidade da reconstrução de quadros perdidos e a propagação dos efeitos das perdas para quadros recebidos corretamente.

Em [9,10], são propostos métodos objetivos e subjetivos para avaliar a qualidade da voz em redes de terceira geração, que também usam o AMR como codificador padrão.

Em [9], especificamente, é utilizado um novo método para avaliar a qualidade de voz em redes IP, baseado em dados tanto objetivos quanto subjetivos, para aplicações VoIP, método no qual a parte subjetiva apresenta uma medição do MOS baseado em medidas de tráfego na Internet, além de testes MOS baseados em diferentes condições da rede. Os resultados foram comparados com medidas de métodos objetivos, constatando-se alta correlação entre o método MOS tradicional e o MOS baseado no tráfego da rede Internet. Evidenciou-se que o aumento da taxa de perda de pacotes corresponde a uma diminuição do valor do MOS. Além dessas investigações, esse estudo propõe um novo método de avaliação através da medida da qualidade de voz intrusiva e não intrusiva, levando em consideração modelos matemáticos, baseados no algoritmo PESQ do ITU-T, e no modelo conhecido como E-model, que trata de medir transmissões degenerativas predizendo em seguida que qualidade de voz irá resultar dessas transmissões. Os resultados deste novo método mostraram que os métodos objetivos PESQ e E-model podem prever muito bem os escores do método subjetivo MOS. O mesmo artigo também destaca que métodos subjetivos são primordiais para a validação dos resultados colhidos com métodos objetivos.

Já o artigo [10], apresenta uma ferramenta objetiva de caracterização da qualidade do serviço de voz em redes WCDMA, destacando-se que as medidas de qualidade de voz objetivas são ótimas alternativas para as avaliações subjetivas. Os resultados desta fonte também mostraram que a qualidade de voz é altamente degradada quando se aumenta a taxa de erro de bits.

Os trabalhos [11–13] investigaram o impacto do uso do codificador AMR para o serviço VoIP ao nível de aplicação, partindo todos eles do fato de que as perdas de pacotes, muito freqüentes em VoIP, acarretam a degradação da qualidade da voz. Concluíram que a adaptação da taxa de geração de bits do AMR, dependendo das condições da rede e das taxas do sinal de ruído do canal sem fio ou da manipulação eficiente, apesar de não desejável, da largura de banda destinada ao tráfego de voz, pode ajudar a reduzir essas perdas, anunciando-se como técnica promissora para o incremento da qualidade de voz.

Os autores de [14, 15], mediram o desempenho e a eficiência do VoIP em redes celulares 3G como HSDPA, HSDPA/EUL, WCDMA, CDMA 2000, 1X-EVDO e 1X-EVDOV, evidenciando a importância do requerimento de atraso e da qualidade de serviço, que são destacados nesta Dissertação também. O primeiro desses trabalhos avalia a capacidade da rede com o VoIP, levando em consideração diferentes requerimentos de atraso associados ao canal de acesso de rádio, além de investigar outras métricas ligadas ao canal de rádio e sua melhoria. A abordagem foi feita utilizando-se uma ferramenta de simulação conhecida como Opnet, que captura processos dinâmicos em uma rede de rádio. Simularam-se atrasos variáveis, com o objetivo de avaliar o impacto das perdas de pacotes no desempenho da qualidade do VoIP. Seus resultados mostraram que, conforme as condições da rede e da minimização de uso da largura de banda por tal serviço, o HSDPA suporta no máximo 4 usuários transmitindo simultaneamente. Neste Trabalho, utilizou-se também de um simulador, mas com requerimento de atraso determinado e resultados maiores em termos de capacidade. Essa metodologia oferece um melhor conhecimento do VoIP atuando em canais de redes de pacotes e pode servir como base para o desenvolvimento do núcleo da rede e da parte de rádio. Aplicaram-se também algoritmos de compressão de cabeçalho do serviço VoIP, assim como esta Dissertação de Mestrado, demonstrando-se sua necessidade para manter alta a capacidade da rede. No segundo estudo aqui referido, examinam-se também a compressão de cabeçalho no provimento do serviço VoIP, além de melhorias na qualidade e no atraso. Seus resultados mostraram que, em células sem cobertura do HSDPA, o serviço VoIP pode ser transportado usando-se portadoras de rádio chaveadas por pacotes do WCDMA.

Nos artigos [16–18], o serviço VoIP é avaliado especificamente junto ao sistema HSDPA, abordando-se o requerimento de atraso exigido em tais redes, principalmente para esse serviço de tempo real, assim como neste Trabalho. No primeiro destes estudos, os resultados a partir de um simulador de rede quase estático para o canal de downlink mostraram que a capacidade no provimento do serviço VoIP sobre o HSDPA varia entre 72 e 104 usuários, dependendo da taxa de atraso que pode ser 80 ou 150 ms. Nesta Dissertação, fixou-se o valor 80 ms. O artigo seguinte concluiu que o desempenho no sistema do VoIP é altamente dependente do atraso e que os canais de tráfego do HSDPA devem ser reestruturados para suportar melhor e com boa capacidade as aplicações que usem VoIP e multimídia. O último artigo, por sua vez, utiliza técnicas avançadas para alcançar altas taxas de desempenho para os

usuários, e um cenário com tráfego simultâneo de voz e de vídeo, cujos resultados mostraram que o serviço VoIP pode ser suportado juntamente com o tráfego de outros dados, mantendo-se um bom nível de qualidade de serviço para todos eles. Nesta Dissertação, porém, há somente a presença do tráfego de voz.

Os trabalhos aqui mencionados serviram de base para as investigações feitas na pesquisa apresentada nesta Dissertação de Mestrado, que avalia, através dos métodos objetivo e subjetivo, a qualidade da voz do serviço VoIP no sistema HSDPA, usando-se o AMR como codificador de voz.

1.5 Metodologia Empregada

Primeiramente, fez-se um estudo das condições necessárias para a implantação do serviço VoIP em um sistema 3G de rede celular. Em seguida, fez-se a modelagem e a construção de blocos de códigos individuais que refletem os processos de codificação e decodificação de arquivos de voz, prosseguindo-se com a integração desses blocos com um simulador dinâmico já existente, que reflete as principais funcionalidades do VoIP no sistema HSDPA. O núcleo principal deste trabalho refere-se à avaliação da qualidade da voz usando o protocolo Internet no sistema celular de terceira geração HSDPA, servindo-se das metodologias objetiva e subjetiva. Assim, os processos de geração da voz, seu transporte desde o transmissor até o receptor, pelas pilhas de protocolos e pela interface de rádio, são partes integrantes do processo, embora o maior interesse da pesquisa esteja voltado para a inteligibilidade da voz recebida pelo usuário final. A metodologia de investigação adotada desenvolveu-se nas seguintes Etapas:

- ▶ **E1** - Gravação, de acordo com a recomendação P.800 do ITU-T, de 4 arquivos de voz na língua portuguesa, com extensão.wav, em períodos médios de 60 segundos, contendo períodos de fala com duração média de 3 segundos e com períodos médios também de 3 segundos de silêncio entre um período de voz e o próximo, sendo 2 arquivos com gravações de vozes do sexo masculino e outros 2 com vozes do sexo feminino.
- ▶ **E2** - Esta etapa concentra-se no módulo de codificação, onde acontece a concatenação, a codificação dos arquivos de acordo com as recomendações do codificador AMR e a geração de quadros de fala que servirão como produto final.
- ▶ **E3** - O produto do estágio anterior servirá de ponto de partida para esta

etapa, na qual se faz a simulação computacional para o tráfego de VoIP no HSDPA, gerando arquivos de voz com um certo padrão de erro resultante da transmissão pela rede móvel.

- ▶ **E4** - Nesta etapa, que tem como ponto de partida o pacote de voz gerado pela E3, encontra-se o módulo de decodificação, em que os quadros de fala são decodificados e remontados.
- ▶ **E5** - Sobre os arquivos remontados na etapa anterior procede-se à aplicação das metodologias de avaliação objetiva e subjetiva.

No caso da avaliação objetiva, na Etapa E5, aplica-se um algoritmo nos arquivos remontados na Etapa E-4, com o intuito de obter uma nota para a qualidade da voz nesses arquivos. Para a avaliação subjetiva, no entanto, os arquivos foram colocados em uma página na Internet, disponíveis para escuta e atribuição de notas pelos ouvintes, segundo uma escala de qualidade para voz também ali publicada.

Além dessa medição de qualidade, esta pesquisa colheu também outras métricas úteis para a análise de desempenho do serviço VoIP, através da rede de acesso do HSDPA tais como medidas de atraso, de variação do atraso e taxa de quadros perdidos.

1.6 Contribuições e Produção Científica

As contribuições específicas deste Trabalho para o conhecimento de seu principal objeto são:

- ▶ A análise da qualidade de voz do VoIP sobre o sistema 3G HSDPA, de forma objetiva e subjetiva, a partir da relação entre a métrica de taxa de quadros perdidos (FER), fornecida pela ferramenta de simulação, e a nota de qualidade dada em ambas as metodologias de avaliação empregadas;
- ▶ A criação de um simulador do provimento do serviço VoIP no sistema HSDPA para avaliação objetiva e subjetiva da qualidade de voz com módulos de codificação e decodificação.

Com contribuições da pesquisa que gerou esta Dissertação, foi produzido um artigo científico, apresentado em congresso internacional, aqui anexado como Apêndice B:

- ▶ **Quality of Service Evaluation of VoIP over HSDPA**, Leonardo Ramon N. Sousa, Marccone L. Carvalho, Emanuel B. Rodrigues, Leonardo Sampaio e Francisco R. P. Cavalcanti, VI International Telecommunications Symposium - ITS2006, Fortaleza-CE, Brazil, Setembro, 2006.

1.7 Organização da Dissertação de Mestrado

O texto foi organizado de modo a apresentar de forma seqüencial as técnicas utilizadas para avaliar de forma objetiva e subjetiva a qualidade da voz sobre o protocolo Internet:

- ▶ **Capítulo 2 - Conceitos de Voz sobre o Protocolo Internet:** apresenta-se e descreve-se o codificador de voz para sistemas celulares e os principais protocolos para o provimento do serviço VoIP, e as principais métricas necessárias para a garantia da qualidade de serviço. Também se descrevem, resumidamente, os métodos subjetivo e objetivo para a avaliação da qualidade de voz.
- ▶ **Capítulo 3 - Rede de Acesso de Rádio do UMTS:** abordam-se as características do UMTS, sua arquitetura, suas classes de serviço, nas quais se insere o VoIP, e descreve-se o sistema HSDPA com suas funcionalidades e aspectos mais relevantes para a inserção do serviço de voz sobre IP.
- ▶ **Capítulo 4 - Avaliação da Qualidade do Serviço VoIP em Sistemas HSDPA:** é apresentada a contextualização do serviço VoIP no sistema HSDPA, abordando-se aspectos de simulação, explicitando-se o procedimento pelo qual foi modelado o sistema em termos de estrutura da rede celular, dos mecanismos e protocolos implementados. Além disso, também são apresentados alguns parâmetros comuns de simulação, seguidos dos resultados gerados pelos métodos objetivo e subjetivo de avaliação.
- ▶ **Capítulo 5 - Conclusões e Perspectivas.**

1.8 Sumário

Este capítulo inicial apresentou uma breve introdução aos sistemas de comunicações móveis e ao tráfego da voz através da Internet, além de evidenciar

a importância de medir-se a qualidade desse tráfego. Fez-se uma breve contextualização deste estudo, indicando-se outros trabalhos nos quais se baseou e que permitiram um aprofundamento sobre os sistemas, metodologias e técnicas estudadas. Apresentaram-se também a motivação, os objetivos, a metodologia e as contribuições geradas ao longo da preparação desta Dissertação de Mestrado.

Capítulo 2

Voz sobre o Protocolo Internet

O transporte de voz sobre redes IP exige o uso de protocolos apropriados para garantir a eficiência de tal serviço, dentre os quais serão abordados, neste Trabalho, o protocolo de tempo real (RTP, do inglês *Real-Time Protocol*), o protocolo de datagrama de usuário (UDP, do inglês *User Datagram Protocol/ Internet Protocol*) e o protocolo Internet (IP). Antes de usá-los, porém, é necessário contar também com um eficiente mecanismo de codificação de voz, capaz de transformar um sinal analógico em um dado digital. Para os fins desta pesquisa será utilizado o codificador AMR. Uma vez ocorrida a sinalização e a codificação, os protocolos RTP, UDP e IP são utilizados para transportar os pacotes de voz. Na Figura 2.1, os quadros de voz são gerados no nível da aplicação, codificados e encapsulados no campo de carga útil (chamado em inglês de *payload*) de um segmento RTP. Um segmento RTP é encapsulado dentro de um segmento UDP, que é então entregue para a camada IP.

Outras métricas de qualidade de serviço como atraso, variação do atraso (em inglês *jitter*) e taxa de apagamento de quadros são também analisadas em virtude da influência que exercem sobre o serviço VoIP. Metodologias de avaliação da qualidade de serviço são igualmente necessárias para obter-se uma idéia da percepção que terá o usuário final.

Assim, este capítulo abordará a importância de um bom codificador de voz. O codificador escolhido na seção 2.1 deste Trabalho e os principais protocolos usados na transmissão de VoIP serão descritos na seção 2.2. Em seguida, na seção 2.3, serão discutidos os principais parâmetros para qualidade do serviço de voz. Finalmente, na seção 2.3.1, faz-se um breve detalhamento dos métodos de avaliação objetivo e subjetivo, aqui utilizados.

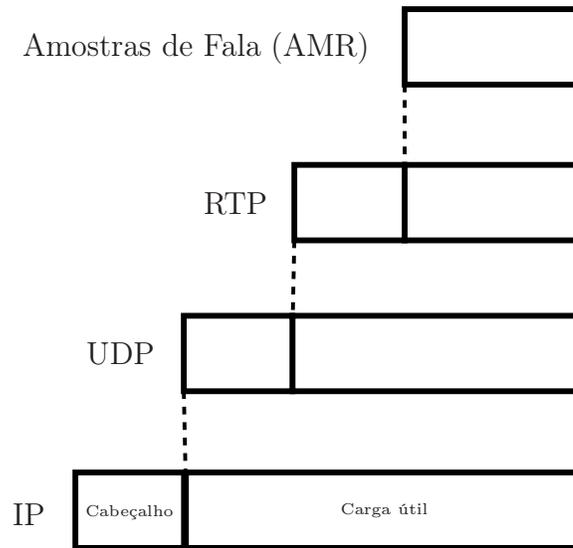


Figura 2.1: Transporte da voz no Protocolo IP.

2.1 Codificadores de Voz

Os codificadores e decodificadores de voz, também chamados de codecs, são dispositivos desenvolvidos para sistemas celulares com altas taxas de erro de bits [19]. Os codecs oferecem resposta à necessidade de que algumas técnicas operem em tempo real, devido às características do próprio serviço de comunicação interativa. Nesse caso, o uso da largura de banda é um fator muito importante, já que tem efeito na qualidade da voz reconstruída. Esses dispositivos permitem reduzir a largura de banda para a transmissão de dados utilizando técnicas de compressão. A compressão de sinais é baseada em técnicas de processamento que eliminam quase toda informação redundante ou não perceptível. Os codecs têm o objetivo de transformar sinais analógicos em dados digitais. Alguns dos codificadores mais conhecidos são G.711, G.722, G.723.1, G.726, G.729, AMR e iLBC (do inglês *internet Low Bitrate Codec*) [20–22]. O que diferencia esses codificadores entre si são os algoritmos usados por cada um deles, suas taxas de geração de bits, a média de atraso e a qualidade de voz. Todos eles, porém, são recomendados pelo Setor de Padronização de Telecomunicações da União Internacional de Telecomunicações (ITU-T).

Para transmitir voz em redes de pacotes, é necessário o uso de codificadores que trabalhem a baixas taxas de bits como, por exemplo, o codificador AMR. Antes que se possa enviar a voz na forma de dados digitais, é feito o empacotamento dos

bits gerados pelo codificador. Dependendo das características desse codificador, a quantidade de quadros de voz em cada pacote pode variar. Diferentes fenômenos como atrasos, perdas e erros sobre os dados transmitidos através das redes refletir-se-ão na qualidade do sinal reproduzido depois de decodificado. Por isso, a capacidade de decodificação com boa qualidade, por parte dos decodificadores, é uma característica fundamental para atenuar tais fenômenos através de mecanismos adequados como, por exemplo, a reconstrução de quadros perdidos.

2.1.1 Codificador *Adaptive Multi-Rate* (AMR)

O codificador de voz AMR foi escolhido como o codificador obrigatório para serviços de fala conversacional em sistemas celulares de terceira geração pelo 3GPP (do inglês *Third Generation Partnership Project*). Trata-se de um codec com 8 taxas para codificação da voz entre 4.75 e 12.2 kbps. Se a taxa de dados é 12.2 kbps, por exemplo, o codificador AMR gera pacotes de 244 bits, representando quadros de voz de 20 ms [23–25]. A possibilidade de usar vários valores de taxa permite uma melhor adequação às condições da rede. Esse procedimento tem impacto direto na amplitude da cobertura, na capacidade do sistema celular e na qualidade final percebida pelo usuário (inteligibilidade da voz).

O AMR codifica e decodifica amostras de fala digital para otimizar o uso da potência e da largura de banda disponível [26], e também suporta uma baixa taxa de codificação de ruído de conforto, o que reduz a transmissão durante os períodos de silêncio.

Devido à sua robustez e à sua flexibilidade, esse codificador é bem adequado para aplicações de voz sobre IP (VoIP). O IETF aprovou o formato do RTP payload para o AMR com o objetivo de usá-lo com o protocolo RTP [27].

Existe, também, uma relação direta entre a taxa de codificação e o valor da MOS. Quanto menor a taxa, menor o MOS [28]. O projeto desta pesquisa não se propôs como objetivo o estudo da variação das taxas de dados e por isso, para efeito desta investigação, fixou-se a taxa em 12.2 kbps.

2.2 Protocolos para o Tráfego de VoIP

Os principais protocolos que possibilitam o tráfego da voz sobre IP são descritos a seguir:

2.2.1 Protocolo de Tempo Real (RTP)

O Protocolo de Tempo Real (RTP) é um protocolo de transporte fim-a-fim [29], mas que também pode ser visto como parte da camada de aplicação, pois o desenvolvedor pode optar por integrá-lo na aplicação [30]. O RTP possibilita o transporte de tráfego multimídia sobre o IP, como voz e vídeo. Além disso, ele permite certo nível de tolerância para com a perda de pacotes introduzidos pela rede IP e para com a variação do atraso entre o tempo esperado e o tempo de chegada, que em inglês é chamado de *jitter* [31].

O protocolo RTP não oferece um mecanismo de confiabilidade, não provê qualquer mecanismo para assegurar entrega sincronizada ou outro tipo de qualidade de serviço (QoS) e de garantia de resultado. Tais garantias devem ser asseguradas por serviços das camadas inferiores. Pelo fato do protocolo RTP não prover um mecanismo de controle de congestionamento, e de não incluir a reserva de recursos ao longo do trajeto de uma conexão [29], o transmissor poderia encher o receptor com muitos dados em tempo excessivamente rápido. Para prevenir esse problema, as aplicações enviam os dados a uma taxa fixa. O RTP serve-se de um protocolo de sinalização para estabelecer uma conexão, negociando com o interlocutor os formatos de mídia que serão utilizados. O transporte de mídia exige considerações a respeito de QoS tanto no projeto da plataforma de serviço quanto na rede de suporte. Uma das principais questões a ser avaliadas, na utilização de serviços IP para o transporte de mídia, é se o RTP pode realmente ser visto e configurado como um serviço de transporte não orientado à conexão, uma vez que a QoS é de maior importância. O roteamento baseado em QoS provê um mecanismo para fixar os trajetos de mídia fim-a-fim, removendo a incerteza do roteamento de pacotes IP, resultando em uma QoS determinística do ponto de vista da rede.

O cabeçalho do pacote RTP é pequeno, como mostra a Figura 2.2, contendo apenas 12 octetos [29]. Somente os campos até o SSRC fazem parte da identificação básica. As principais características do RTP, presentes em seu cabeçalho são:

- ▶ Informação do tipo de dado carregado (PT, do inglês *Payload Type*): se é áudio ou vídeo;
- ▶ Ele permite o cálculo da variação do atraso (através do Timestamp);
- ▶ Estimação do número de pacotes perdidos e a restauração de sua seqüência (do inglês *Sequence Number*).

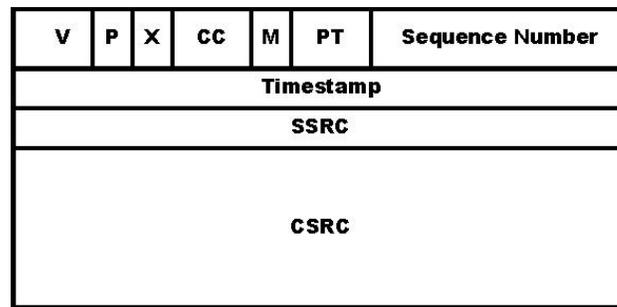


Figura 2.2: Cabeçalho de um pacote RTP

Assim, os campos do cabeçalho de um pacote RTP de acordo com a Figura 2.2 são os seguintes:

- ▶ *Version* (V, 2 bits) - identifica a versão do RTP;
- ▶ *Padding* (P, 1 bit) - determina se o conteúdo real de dados do pacote será ou não menor que o tamanho do pacote;
- ▶ *Extension* (X, 1 bit) - identifica se o cabeçalho fixo é seguido por uma extensão de cabeçalho de tamanho variável;
- ▶ *Contributing Source Count* (CC, 4 bits) - contém o número de identificadores Contributing Source que vêm após o cabeçalho fixo;
- ▶ *Marker* (M, 1 bit) - sua interpretação é definida por um *profile*, permitindo que eventos significativos, como fronteiras de quadro, sejam marcados no fluxo de pacotes;
- ▶ *Payload Type* (PT, 7 bits) - identifica o formato do *payload* do RTP, determinando sua interpretação pela aplicação;
- ▶ *Sequence Number* (16 bits) - esse campo pode ser incrementado de 1 para cada pacote RTP enviado e pode ser utilizado pelo receptor para determinar perda de pacotes ou restaurar sua seqüência;
- ▶ *Timestamp* (32 bits) - definido pelo instante de amostragem do primeiro octeto do pacote de dados do RTP;
- ▶ *Synchronization Source Identifier* (SSRC, 32 bits) - identifica a fonte de sincronismo, isto é, o transmissor;

- ▶ *Contributing Source Identifiers* (CSRC list, 0 a 15 itens, 32 bits cada) - a lista CSRC identifica as fontes contribuintes para o *payload* contido no pacote, sendo usado para fazer uma identificação correta de fonte quando os *payloads* são tocados no ponto terminal.

Para o serviço VoIP, o RTP é implementado junto com UDP/IP que permite utilizarem-se os serviços de multiplexação e de checagem de cabeçalho [32]. Normalmente o RTP é usado em cima do UDP, que fornece a porta e a detecção de erro. O RTP pode, quando usa o UDP, ser transportado por pacotes IP para atingir vários destinos.

O uso do IP como protocolo de rede apresentou o problema de atraso na entrega da mídia, em razão do fato de que roteadores e outros elementos de interconexão implementam um tratamento de armazenamento e encaminhamento (do inglês *store-and-forward*). Assim, entre outros problemas no uso do IP para transporte em tempo real, há uma grande sobrecarga de protocolo em cada pacote, que inflaciona a banda passante necessária para uma conexão.

Soluções IP usando cabeçalhos não comprimidos sofrem com a sobrecarga das camadas de rede e transporte, já antes que um único octeto de dado de usuário seja colocado no campo de dados. Logo, a compressão de cabeçalho é altamente recomendada neste caso.

O protocolo RTP pode medir o desempenho de rede como atraso, variação do atraso (*jitter*) e taxa de apagamento de quadros que podem impactar a qualidade do VoIP, sendo, por isto, os principais parâmetros de desempenho a ser considerados.

Com o objetivo de amenizar a variação do atraso, a camada de aplicação situada no destino deve incluir um reproduzidor, ou *buffer* em inglês, com considerável capacidade de armazenamento de pacotes, implementado no RTP do receptor com um tamanho variável que depende da máxima variação do atraso para uma dada aplicação [33]. Esse reproduzidor armazena a voz que chega para que possa ser reproduzida com um tempo constante na aplicação destino [34]. Se os pacotes chegam tarde demais, são descartados sem a necessidade de reprodução.

A funcionalidade do RTP é acrescida com o RTCP (do inglês *Real-Time Control Protocol*), que provê monitoramento fim-a-fim da entrega de dados e da qualidade de serviço. O RTCP faz parte da especificação do RTP [29].

2.2.2 Protocolo de Controle de Tempo Real (RTCP)

O Protocolo de Controle de Tempo Real (RTCP) provê uma resposta com a informação sobre a qualidade dos dados entregues (a quantidade de variação do atraso, a perda média de pacotes) e sobre os participantes da sessão [35]. O RTCP, associado com o RTP, monitora a sessão, propondo os tipos de pacotes mais adequados para troca de informações de controle entre o transmissor e o receptor em uma sessão RTP [29]. O transmissor envia um relatório (SR, do inglês *Sender Report*) para informar quantos pacotes são enviados e o receptor responde através de outro relatório (RR, do inglês *Receiver Reports*) indicando a taxa de pacotes perdidos. Assim, o transmissor RTP pode usar esses relatórios para fazer o controle de fluxo pela variação da taxa de transmissão [30], se ele tiver algum mecanismo que faça isso. Todos os participantes das sessões enviam pacotes RTCP, que podem ser de cinco tipos:

- ▶ Os transmissores enviam SR - pacotes que contêm informações de transmissão e recepção para transmissores ativos em uma conexão;
- ▶ Os receptores enviam RR - pacotes que contêm informações de recepção para ouvintes que não sejam ativos;
- ▶ *Source Description* (SDES) - pacotes cujo envio permite que se faça a associação entre o valor SSRC do pacote RTP e a real identificação do usuário;
- ▶ *BYE* - pacotes que informam a todos os participantes da chamada que o usuário está saindo da sessão;
- ▶ *Application-specific RTCP packet* (APP) - pacotes APP transportam informações entre diferentes aplicações que estejam ativas nos pontos terminais.

Um grande problema a ser resolvido, neste ponto, é a limitação da largura de banda, pois o tráfego cresce de acordo com a quantidade de participantes. Como o número de participantes é conhecido por todos, cada um deles pode controlar a taxa com que envia relatórios RTCP e limitar a largura de banda da sessão usada pelo RTCP a um valor razoável, que deve ser compartilhada por todos os participantes.

O RTP e o RTCP não têm qualquer influência sobre o comportamento da rede IP, pois eles não controlam a qualidade de serviço. A rede pode perder, inserir atraso

ou perder a seqüência de um pacote RTP da mesma maneira que o faz com qualquer pacote IP.

O RTCP não é implementado na pilha de protocolos aqui utilizados pois, nesta abordagem, não existem informações de controle sobre os participantes da sessão. A inserção do tópico RCTP nesta Dissertação teve a intenção não apenas de mencionar suas características junto ao protocolo RTP, mas também de mostrar que este pode ser um nicho ou tema de pesquisa a ser explorado futuramente.

2.2.3 Protocolo de Datagrama de Usuário (UDP)

O Protocolo de Datagrama de Usuário (UDP) é usado pelo VoIP como um protocolo da camada de transporte sobre o IP, previsto de forma a evitar o aumento do atraso causado por retransmissões de pacotes perdidos [35].

O UDP é um protocolo de transporte orientado a datagrama. Ele não provê confiabilidade na entrega dos datagramas já que, por exemplo, não inclui confirmação da entrega de um pacote de dados (PDU, do inglês *Packet Data Unit*). Ele assume o protocolo Internet como seu protocolo de rede [36], uma vez que sua modelagem supõe que cada UDP PDU gere apenas um IP PDU, sem fragmentação do IP.

A principal tarefa do UDP é enviar datagramas da aplicação para a camada de rede, sem garantir, porém, que o destino será alcançado [37, 38]. O tamanho do cabeçalho do UDP é padronizado em 8 bytes, como na Figura 2.3.

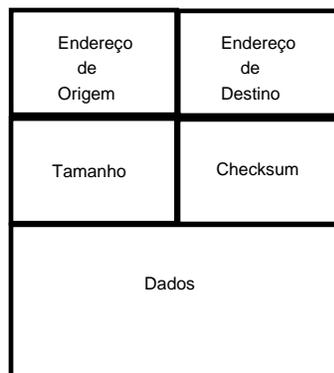


Figura 2.3: Cabeçalho UDP.

Seguem-se as definições dos campos da Figura 2.3:

- Campos que contêm os números da porta fonte e da porta destino do protocolo UDP. O uso da porta fonte é opcional e, quando é usada, especifica-se como a

porta à qual se poderia enviar uma resposta; caso contrário, ela contém apenas zeros;

- ▶ **Tamanho** - campo que contém um contador de bytes no datagrama UDP. Seu valor mínimo é de 8 bytes, indicando somente o comprimento do cabeçalho do UDP sem o tamanho do campo de carga útil;
- ▶ *Checksum* - campo opcional útil para o monitoramento da confiabilidade quanto a erros na transmissão, detectando, por exemplo, se há bits invertidos no segmento transmitido. Um valor de zero indica que não há erro computado;
- ▶ O campo de carga útil ou *payload* - caracterizado pelos datagramas das camadas superiores.

2.2.4 Protocolo Internet (IP)

O Protocolo Internet teve origem, em 1970, no desenvolvimento da rede própria do Departamento de Segurança dos Estados Unidos, mais tarde interligada à outras redes, formando um conjunto que passou a ser conhecido como Internet. Desenvolvimentos subsequentes incluíram o protocolo IP no sistema operacional UNIX, com o qual um grande número de universidades passou a constituir suas próprias redes e a ligá-las à Internet.

Esse protocolo é não-orientado à conexão, o que significa que não há nenhuma verificação de erro na transferência. O IP fornece um serviço de datagramas que é usado também por outros protocolos, como o RTP, o TCP e o UDP [37, 39, 40]. Trata-se de um protocolo que fornece um serviço de transferência de dados, de encaminhamento e de endereçamento independente da implementação da camada inferior, assegurando a comunicação através das redes pelo fato de que, em seu âmbito, cada nó da rede possui um endereço universal [37, 41].

Uma transferência de dados entre quaisquer dois nós da rede, usando o protocolo IP, pode sempre encontrar caminho para prosseguir, ainda que os dados tenham de ser transportados pelas mais diversas tecnologias, tais como Ethernet, Token-Ring, ATM, X.25 e FDDI (do inglês *Fiber Distributed Data Interface*).

2.2.5 Compressão de Cabeçalho (HC)

Os grandes cabeçalhos dos protocolos usados em voz sobre IP, em redes sem fio, constituem o maior problema para o tráfego de voz pela Internet. Para superar este problema, pode-se usar a compressão de cabeçalho (HC, do inglês *header*

compression). Para serviços multimídia, como VoIP, o uso de HC proporciona uma considerável economia de largura de banda, uma vez que, sem esse recurso, os cabeçalhos para RTP/UDP/IP fazem uso ineficiente do espectro de rádio e consomem muita largura de banda [42, 43].

O pacote multimídia consiste em 12 bytes de cabeçalho RTP, 8 bytes de cabeçalho UDP e 20 bytes do IPv4 ou 40 bytes do IPv6, totalizando 40 ou 60 bytes. O algoritmo HC comprime o cabeçalho RTP/UDP/IP para 2 octetos, no melhor caso [44].

Os principais exemplos de HC são o CRTP (do inglês *Compressed RTP*) [45], o ROCCO (do inglês *Robust Checksum-Based Header Compression*) [4, 42, 45] e o ROHC (do inglês *Robust Header Compression*) [46, 47]. Este Trabalho também não tem o intuito de aprofundar cada um dos exemplos de algoritmos de compressão, pois isto escaparia de seu foco específico.

2.3 Qualidade de Serviço para Voz

O crescimento da Internet trouxe uma tendência a integrar voz e dados na infra-estrutura de redes de pacotes, tornando a Qualidade de Serviço (QoS) na rede um ponto ainda mais importante para um bom desempenho das aplicações em VoIP, do começo ao fim das transmissões, já que a qualidade de uma conexão VoIP depende diretamente da performance, da disponibilidade e dos recursos da rede de dados utilizada. Como o protocolo IP não previa mecanismo algum de QoS, nem dava garantia de alocação dos recursos da rede, foi necessário desenvolverem-se protocolos que garantissem essa qualidade fim-a-fim [48], além da exigência do controle de determinados parâmetros que são requisitos de QoS como atraso, a variação do atraso (*jitter*), vazão, latência, largura de banda, perda de pacotes ou de seqüência, eco e indisponibilidade.

Outras formas de prover qualidade de serviço baseiam-se na negociação de serviços de rede para uso de múltiplos serviços, na utilização de mecanismos adaptativos que tentam reduzir as perdas e os atrasos de pacotes e no envio de correções de erro, fornecendo mecanismos de redundância para ultrapassar a perda de pacotes multimídia.

O principal objetivo da garantia de QoS é priorizar o tráfego interativo sensível ao atraso em detrimento do tráfego de transferência de arquivos [49, 50]. A qualidade de serviço deve ser fim-a-fim, tratada independentemente em cada uma das partes que passam ou repassam o pacote de dados, ou seja, o tráfego deve ser tratado tanto na rede de origem, quanto nos roteadores, com alguma política de controle de

descarte de pacotes, e na rede destino [51], como mostra a Figura 2.4.



Figura 2.4: Modelo para Qualidade de Serviço (QoS).

Voz sobre IP é um serviço que requer uma boa coordenação entre o transporte pela rede e os pontos finais responsáveis pelo processamento de voz e por sua qualidade de serviço [52]. Para obter-se uma melhor eficiência nas conexões com VoIP, é preciso fazer um melhor uso da banda, recorrendo-se aos protocolos associados às aplicações de voz, como os protocolos de compressão de dados, por meio dos quais se pode obter uma maior economia de banda e, conseqüentemente, uma melhor qualidade do sinal de voz [53].

As principais técnicas e mecanismos necessários para a obtenção de uma boa qualidade de serviço no VoIP [54] são:

- ▶ Necessidade de equilíbrio entre a taxa de amostragem e a utilização de banda;
- ▶ Controle de congestionamento e uso de algoritmos para a priorização de pacotes, em conjunto com outras técnicas para aumentar a velocidade do fluxo de pacotes de voz;
- ▶ O tamanho do pacote tem de ser fixo, de modo que a perda de alguns poucos pacotes não prejudique uma conversação;
- ▶ Emprego de tecnologias de correção de erro e buferização para compensar o atraso e a perda de pacotes;
- ▶ Uso de algoritmos de compressão de voz na supressão de silêncio - pois os períodos de silêncio utilizam, desnecessariamente, a banda de transmissão - permitindo maior disponibilidade de banda para outros pacotes úteis de voz ou de dados;
- ▶ Eliminação do eco;
- ▶ Compressão de voz;

- Uso dos protocolos adequados para transmissão da voz: RTP, UDP e IP.

Para adicionar recursos de QoS, serão considerados aqui dois modelos de classes de serviços para tráfego Internet, desenvolvidos pela IETF. O primeiro refere-se aos serviços diferenciados (DiffServ, do inglês *Differentiated Services*) [55,56], oferecendo um tratamento diferenciado, com preferência estatística por determinados tipos de fluxo. Este é o modelo no qual o VoIP se baseia. Já o segundo desses modelos refere-se aos serviços integrados (IntServ, do inglês *Integrated Services*) [56,57], que fornecem uma garantia absoluta de alocação dos recursos da rede, de modo que os recursos requeridos sejam reservados para cada fluxo antes mesmo do envio do dado atual.

Os parâmetros objetivos de qualidade de serviço mais discutidos nesta Dissertação são atraso, variação do atraso (*jitter*) e taxa de apagamento de quadro [58], que podem ter um grande impacto na qualidade final da voz e que serão descritos a seguir.

Os parâmetros subjetivos de qualidade, por sua vez, baseiam-se na percepção do usuário ouvinte e na sua correspondência com um conceito de qualidade conhecido como MOS, que também será explicado a seguir. Esses parâmetros são importantes para a avaliação do serviço porque, embora existam e usem-se fórmulas analíticas objetivas para determinar a qualidade de um sistema de comunicação de voz, identificando-se distorções entre o sinal recebido e o sinal original, nem sempre estas fórmulas são úteis se o usuário não considerar o resultado aceitável. A importância da avaliação subjetiva decorre do princípio geral de que a qualidade do serviço deve atender às expectativas do cliente, pois um usuário de telefonia necessariamente deseja compreender o significado associado ao áudio reproduzido (entendimento das palavras: inteligibilidade) e distinguir entre diferentes fontes (identificar a pessoa que fala: discernibilidade). Levando em consideração essa necessidade, optou-se, nesta pesquisa, pela utilização do método subjetivo MOS para avaliar essas características.

2.3.1 Avaliação da Qualidade de Voz pelos Métodos Subjetivo e Objetivo

A avaliação da qualidade de voz nas redes de telecomunicações é de grande importância para que se conheça a percepção do usuário final. Para avaliar a qualidade de voz usada em uma transmissão são aplicados os métodos subjetivo e objetivo, cujas estruturas e resultados serão expostos no Capítulo 4.

Como já se mencionou neste texto, o método subjetivo utiliza-se de uma classificação absoluta, dividida em categorias, chamada MOS, como principal forma

de avaliação para obter resultados reproduzíveis e confiáveis [59]. Os ouvintes das transmissões de teste devem atribuir, de maneira direta e absoluta, uma nota para cada sinal de voz avaliado, sem compará-lo com outro sinal de voz. Ou seja, este método se baseia na análise subjetiva feita por um grupo de avaliadores que atribuem, cada um deles, uma pontuação à qualidade da fala ouvida. A qualidade e a inteligibilidade da voz são medidas de uma forma geral, de acordo com uma escala de notas previamente fornecida (cf. Tabela 2.1 e Apêndice A).

Tabela 2.1: Escala de notas da avaliação subjetiva MOS.

Conceito	Nota
EXCELENTE	5
BOM	4
ACEITÁVEL	3
INSATISFATÓRIO	2
RUIM	1

Já o método objetivo, que utiliza recursos computacionais para obter a qualidade da voz que está sendo transmitida em um sistema de comunicação, permite estimar um MOS apenas aproximado, mas oferece a vantagem de poder ser aplicado e fornecer resultados com menor custo e menos investimento de trabalho para sua execução. Existem vários métodos de avaliação objetiva atualmente propostos pelo ITU-T [7, 60].

Entre eles, o mais utilizado é o PESQ, por ser o que melhor se relaciona com o MOS subjetivo [20, 61]. O PESQ baseia-se na medição da distorção entre as características do vetor do sinal original e do sinal degradado em consequência de fenômenos ocorridos durante a transmissão [62]. Além das perturbações apresentadas no sinal, na avaliação considera-se também seu atraso. Este método usa algoritmos classificados como perceptuais, pois utilizam técnicas de processamento de sinais para comparar um trecho de voz gravado com um trecho de voz degradado e construir, a partir das diferenças, um modelo de avaliação da qualidade da voz. Outros métodos utilizam os parâmetros objetivos de atraso, variação do atraso e taxa de perda de quadros como outras medidas de qualidade de serviço, parâmetros esses que, a seguir, serão descritos mais detalhadamente.

2.3.2 Perda de Quadros

Uma das causas da perda de pacotes é a variação no canal de rádio. Outra causa situa-se no momento anterior à transmissão, nos casos em que o pacote é descartado pelo fato de ser muito grande o atraso inserido pelas retransmissões. Outra causa de perdas surge quando o pacote atrasado é estimado como perdido, que acontece quando o atraso de um pacote é maior do que um determinado limitante. Todas essas perdas contabilizam uma métrica conhecida como taxa de apagamento de quadro ou *frame* (FER). Como as redes IP não garantem que os pacotes serão entregues em um tempo usual, sempre haverá pacotes descartados sob algumas configurações de carga da rede e durante períodos de congestionamento.

Embora perda de pacotes, por qualquer motivo, seja indesejável, algumas perdas podem ser toleráveis desde que possam ser usadas algumas técnicas de correção [24, 31]. Em geral, a qualidade não é afetada por muito tempo se a quantidade de pacotes perdidos for menor que 5% em relação ao número total de chamadas [14, 31, 63].

2.3.3 Atraso (*Delay*)

O atraso, que é referenciado também pelo termo em inglês *delay*, é uma das medidas mais importantes de QoS para o fornecimento de serviços VoIP. Define-se como a medida do tempo transcorrido desde a fala do transmissor até a sua chegada ao ouvido do receptor. Sobreposição de fala e eco são defeitos causados por um alto atraso em uma rede para tráfego de voz [14].

De acordo com o E-model do ITU [64], o atraso da voz entre o momento da emissão pela boca do falante até a percepção pelo ouvido do receptor deve ser menor do que 250-300 ms para alcançar uma taxa de qualidade de voz aceitável, sendo influenciado pelo codificador de voz, pela transmissão da rede, pelo roteador de origem, pelos dispositivos que compõem a rede de acesso de rádio e pelo processamento do terminal móvel. Mas para o 3GPP [65, 66], o atraso limite para transmissão de áudio é de 400 ms [16, 21].

2.3.4 Variação do Atraso (*Delay Jitter*)

Essa variação do atraso, mais conhecido pelos termos *delay jitter*, em inglês, ou simplesmente *jitter*, é a medida da diferença entre o tempo de chegada esperado e o tempo de chegada efetivo. Uma vez que pacotes de voz sejam gerados a uma taxa constante, por exemplo, a cada 20 ms, dever-se-ia esperar que todo pacote chegue

ao seu destino a cada 20 ms exatamente [14]. Entretanto, as características de funcionalidade das camadas e/ou subcamadas de uma rede causam uma quantidade imprevisível de atrasos ocorridos entre esses pacotes.

De forma a compensar a variação do atraso, o receptor deve implementar um reproduzidor ou um *buffer*, como é mais conhecido, cujo tamanho seja proporcional para uma suficiente quantidade de pacotes do fluxo contínuo de uma aplicação. O *buffer* é usado para remover essa variação do atraso entre pacotes, mas também implica em uma maior sobrecarga ou *overhead* no processo de comunicação.

A variação do atraso pode ser calculado de acordo com a equação a seguir:

$$J = \sqrt{\frac{\sum_{i=1}^L \tau_i^2}{L} - \tau_m^2} \quad (2.1)$$

onde τ_i é o atraso do pacote i , L é o número total de pacotes transmitidos e τ_m é o atraso médio do pacote [67].

2.4 Sumário

Este capítulo apresentou os principais protocolos usados na transmissão de VoIP sobre redes de pacotes, além do codificador utilizado como padrão pelo 3GPP, necessários para as simulações descritas nos capítulos seguintes. Abordou-se graficamente o processo de transporte do pacote de voz com seus cabeçalhos e carga útil, explicitando-se a dependência entre cada camada da pilha de protocolos para o tráfego de VoIP, além de explicitar a necessidade de garantir-se uma qualidade para a voz, com suas métricas mais importantes diretamente descritas. Outro ponto abordado foi a avaliação subjetiva e a avaliação objetiva, com suas interligações e diferenças, estabelecendo-se sua importância no decorrer deste Trabalho de Dissertação de Mestrado.

Capítulo 3

Rede de Acesso de Rádio do UMTS

A terceira geração de redes de telefonia móvel tem sido, atualmente, objeto de intensa pesquisa e discussão. A International Telecommunication Union (ITU) denomina as redes de terceira geração pela expressão *International Mobile Telecommunications 2000* (IMT-2000), correspondendo ao que, na Europa, designa-se como UMTS. Essa nova tecnologia pretende promover uma grande variedade de serviços, especialmente relacionados à multimídia e a alta taxa de transmissão. Nesse contexto, o WCDMA emerge como a principal solução para a interface aérea de terceira geração.

UMTS aparece para atender à crescente demanda de novas aplicações no campo das comunicações móveis em geral e para Internet. UMTS é visto como o sucessor do GSM e marca a mudança para a terceira geração de redes móveis aumentando a taxa de transmissão a até 2 Mbps por usuário móvel e permitindo padrão global de *roaming*.

HSDPA é um sub-sistema do padrão UMTS/WCDMA, proposto recentemente para transmissão de dados no enlace de descida/direto (*downlink*). O HSDPA foi concebido de forma a obter-se alta vazão na transferência de pacotes de dados, com o auxílio das técnicas de adaptação de enlace, funcionalidades que serão explicadas a seguir. HSDPA permite ao WCDMA suportar taxas de até 10Mbps para serviços de melhor esforço, utilizando a mesma banda de 5MHz, aumentando, com isso, a taxa de transmissão do WCDMA em até cinco vezes, com melhorias em termos de maior capacidade e menores atrasos.

Este capítulo tem como principal enfoque a descrição das características mais relevantes dos sistemas de terceira geração, discorrendo sobre UMTS em geral, seguido do sistema HSDPA, que foi utilizado como o sistema base para o VoIP no

trabalho que gerou esta Dissertação de Mestrado. Aqui, portanto, far-se-á apenas uma abordagem geral, sem aprofundar-se o assunto com detalhes que não interessem ao provimento do serviço VoIP.

3.1 *Universal Mobile Telecommunications System* (UMTS)

UMTS é o termo adotado para designar o padrão de terceira geração (3G) estabelecido como evolução para operadoras de GSM e que utiliza como interface de rádio o WCDMA ou o EDGE. Trata-se de uma tecnologia de dados de alta velocidade, parte dos padrões sem fio da família IMT-2000 da ITU, sendo atualmente a tecnologia mais popular do mundo e disponível em mais de 680 redes de vários países e territórios [68]. Até o ano 2000, o desenvolvimento de padrões para o GSM foi conduzido pelo ETSI (do inglês *European Telecommunications Standards Institute*). A partir dessa data, a responsabilidade passou a ser do 3GPP, fruto do esforço conjunto de várias organizações de standards, ao redor do mundo, para definir um sistema celular global de 3G UMTS.

UMTS oferece tele-serviços (como voz ou mensagem de texto) e serviços *bearer*, que definem a capacidade de transferência de informação entre os pontos de acesso. Tanto os serviços orientados à conexão quanto os serviços sem conexão são oferecidos para comunicação ponto-a-ponto e ponto a multi-pontos. Serviços *bearer* possuem diferentes parâmetros de QoS para atrasos máximos na transferência, para variação no atraso e para taxa de erro de bit.

Do ponto de vista do usuário, as aplicações atuais são chamadas de tele-serviços. Um tele-serviço pode fazer uso de vários tipos de serviços, podendo estes serem criados por cada operadora de modo independente das demais, com exceção dos tele-serviços que serão padronizados pela ETSI, como voz, fax e mensagem de texto (SMS).

UMTS também possui o conceito de Ambiente Virtual de Casa (VHE, do inglês *Virtual Home Environment*) que é um conceito de ambiente de serviço pessoal portátil através de uma rede e entre terminais. Isto significa que os usuários apresentar-se-ão sempre com as mesmas características pessoais contando sempre com uma interface que lhes garantirá o acesso imediato a serviços personalizados, seja qual for a rede ou terminal em que entrem em contato com a rede.

O WCDMA (*Wideband CMDA*), tecnologia de rádio utilizada pelo UMTS, é uma técnica de múltiplo acesso por divisão de código com espalhamento direto (DS-CDMA, do inglês *Direct-Sequence CDMA*), por meio da qual os vários usuários

são distribuídos dentro de uma largura de banda de 5 MHz, que é padrão para os serviços 3G [28, 69].

O WCDMA é uma das principais tecnologias para implementação dos sistemas celulares de 3G. Baseia-se nas técnicas de acesso e nas especificações propostas pela ETSI, segundo as quais diferentes usuários podem transmitir a diferentes taxas de dados, permitindo-se que essas taxas variem no tempo.

O WCDMA usa espalhamento direto com *chip rate* de 4096 Mcps. Uma vez que possui uma rede assíncrona, diferentes códigos (longos) são usados para as células e separação dos usuários, numa largura de banda nominal de 5 MHz [70]. Essa largura de banda permite que se alcancem as taxas de transmissão de dados esperadas na 3G. Além disso, esta banda larga proporciona mais *multipaths* que uma banda estreita, aumentando assim a diversidade e melhorando a performance dos serviços.

As principais diferenças entre WCDMA e CDMA (IS-95) referem-se à largura de banda (1,25 MHz) e ao *chip rate* (1,2288 Mcps), diferença esta que proporciona muito mais altas taxas de transmissão de dados com o uso do WCDMA.

Pelas razões acima expostas, a tecnologia WCDMA foi selecionada para fazer a interface aérea UTRAN. WCDMA UMTS é um sistema *Direct Sequence* CDMA em que os dados do usuário são multiplexados com bits pseudo-randômicos dos códigos de espalhamento do WCDMA. Em UMTS, os códigos que são usados para sincronização e espalhamento são somados ao canal. O WCDMA tem dois modos básicos de operação: Frequency Division Duplex (FDD) e Time Division Duplex (TDD), baseados, respectivamente, na divisão de frequência e na divisão do tempo de utilização do meio entre os usuários que transmitirão dados.

Nas subseções a seguir, serão descritas a Arquitetura e as Classes de Serviço do UMTS, informações necessárias para a melhor compreensão do Sistema HSDPA e do serviço VoIP focalizados nesta Dissertação.

3.1.1 Arquitetura do UMTS

Uma rede UMTS consiste no núcleo da rede (CN, do inglês *Core Network*), no UTRAN - Rede de Acesso de Rádio Terrestre do UMTS, no equipamento do usuário (UE, do inglês *User Equipment*) - e nas interfaces entre essas entidades chamadas de Uu, Cu, Iu, Iub e Iur, que serão descritas a seguir. A arquitetura do UMTS pode ser representada, simplificadamente, pela Figura 3.1.

A principal função do CN é prover chaveamento, roteamento e tráfego do usuário, além de possuir a base de dados e desempenhar as funções de gerenciamento da rede.

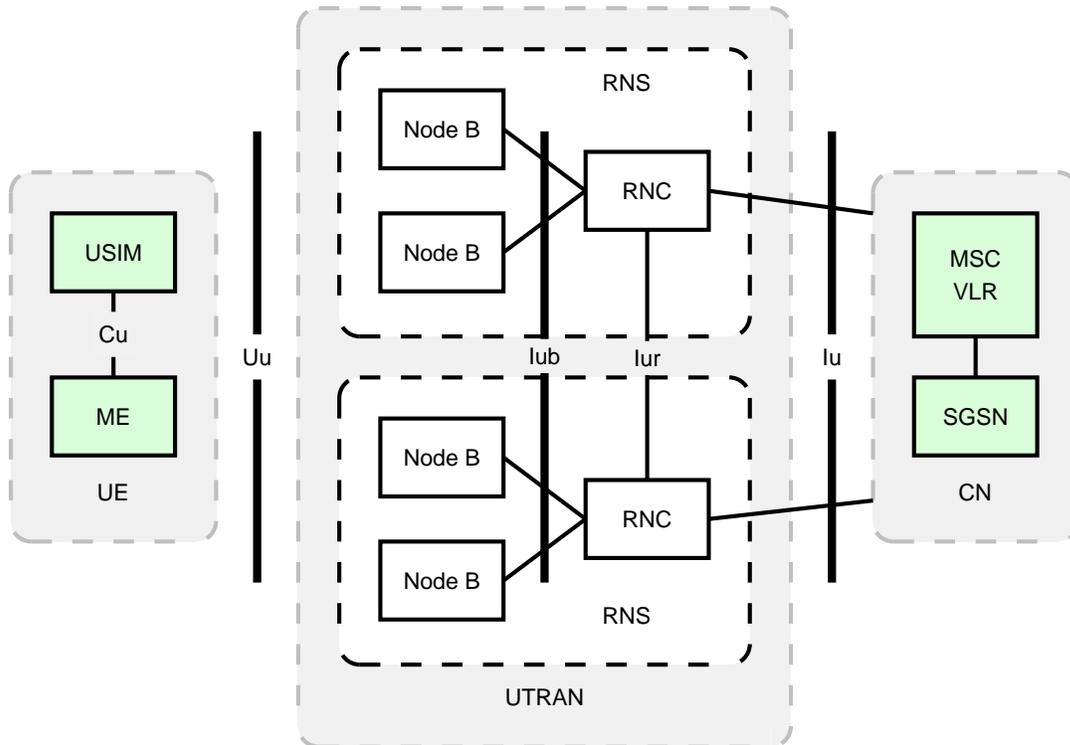


Figura 3.1: Arquitetura básica de uma rede UMTS [28].

O chaveamento pode ser dividido em chaveamento por circuito ou por pacote.

Há elementos chaveados por circuito, dentre eles a MSC (do inglês *Mobile Switching Center*) e o VLR (do inglês *Visitor Location Register*), e elementos chaveados por pacote como, por exemplo, o SGSN (do inglês *Service GPRS Support Node*).

O UTRAN tem a função de prover a interface aérea para o UE, tendo como tecnologia de múltiplo acesso o WCDMA. É composto pelas seguintes entidades:

- ▶ A RNS (do inglês *Radio Network Subsystem*) gerencia os recursos de rádio necessários à conexão do UE à UTRAN. Cada RNS é composta por uma RNC;
- ▶ A RNC (do inglês *Radio Network Controller*) é o nó lógico da RNS responsável pelo controle do uso e integridade dos recursos de rádio;
- ▶ *Node B* é a entidade responsável pela transmissão e recepção dos dados em uma célula.

O UE é, simplesmente, o equipamento móvel do usuário. Esse equipamento

móvel é composto pelo aparelho propriamente dito, representado por ME (do inglês *Mobile Equipment*) e pelo USIM (do inglês *User Service Identity Module*).

Ainda na Figura 3.1, constam as interfaces a serem definidas, a seguir:

- ▶ Uu - Interface entre o UTRAN e o UE, de fato a interface de rádio WCDMA pela qual o equipamento móvel do usuário acessa a parte física do sistema;
- ▶ Cu - Interface entre o ME e o USIM do UE que segue um formato padrão para pequenos cartões;
- ▶ Iu - Interface entre o UTRAN e a CN, que oferece aos operadores UMTS a possibilidade de adquirirem o UTRAN e a CN de diferentes fabricantes;
- ▶ Iub - Interface entre os Nodes B e a RNC do UTRAN, que atua quando, no sistema HSDPA, requer-se um mecanismo de controle de fluxo, para assegurar que os *buffers* sejam usados corretamente no Node B e que não haja perda de dados devidos a um excesso de dados nos *buffers*;
- ▶ Iur - Interface entre duas RNCs do UTRAN, permitindo a troca entre estações rádio-base ou *soft handover*, como é mais conhecido em inglês.

A UMTS reutiliza investimentos anteriores, especialmente, a infra-estrutura de rede de pacotes de dados implementada para a GPRS. Dependendo do fabricante, a atualização pode ser facilmente executada adicionando-se cartões de canais e software UMTS à infra-estrutura de rádio de GSM/GPRS/EDGE já existente, que continua a prover serviços para clientes utilizando essas novas tecnologias. Essa arquitetura em módulos reduz o custo da implementação UMTS, permitindo que as operadoras ofereçam preços mais competitivos para seus serviços 3G, se comparados com o que seriam os preços no caso de uma atualização que envolvesse a substituição dos principais elementos de infra-estrutura.

3.1.2 Classes de Serviço do UMTS

Uma das grandes características do WCDMA (UMTS) é a capacidade de negociar atributos, como atraso de transferência, *throughput* e taxa de dados errados durante o curso de uma conexão ativa. O UMTS inicialmente torna disponíveis serviços tanto no modo chaveamento de circuitos quanto serviços chaveados por pacotes. A tendência, contudo, é de que todos esses serviços sejam alocados para o modo de chaveamento por pacotes.

O QoS determina o desempenho do sistema e, conseqüentemente, o grau de satisfação do usuário. Pode ser avaliado pelas seguintes características:

- ▶ Máximo atraso de transferência;
- ▶ Variação do atraso de informação;
- ▶ Taxa de bits errados;
- ▶ Taxa de transmissão de dados.

Ao definirem-se as classes de serviços, deve-se levar em conta as restrições e limitações da interface aérea que são classificadas em serviços de tempo real e serviços de tempo não-real.

Os serviços móveis são agrupados em quatro categorias, dependendo de suas características segundo os padrões do UMTS [58, 71], e diferenciadas basicamente pela sua sensibilidade ao atraso de transferência, considerando-se que os usuários destes serviços podem se comunicar com redes fixas e com outros móveis. Essas classes são descritas na Tabela 3.1.

Tabela 3.1: Classes de Serviço do UMTS.

Classe	Características	Exemplo
Conversacional	Preserva a variação do atraso; Padrão conversacional; Atraso de transmissão restrito; Não preserva conteúdo.	Serviço de voz.
<i>Streaming</i>	Preserva a relação de tempo; Não preserva conteúdo.	Stream de vídeo.
Interativo	Padrão de requisição e resposta; Preserva informação do usuário.	Serviço WWW.
<i>Background</i>	O destinatário não espera receber a informação dentro de um certo tempo; Preserva informação do usuário.	Correio eletrônico.

Tais categorias de Serviço do UMTS são conhecidas como classes de QoS, atualmente já utilizadas para definir mais precisamente algumas classes de aplicações em sistemas 3G que atendem aos mais diversos tipos de usuários, como o comércio eletrônico, as comunicações através de vídeo-conferências, as informações

necessárias para as identificação dos terminais e para serviços de localização, notícias, entretenimento.

A classe de QoS cujas características devem ser necessariamente obedecidas para o serviço VoIP é a conversacional, cujo principal requisito é a restrição do atraso na transmissão. É importante considerá-la nesta Dissertação de Mestrado que tem como objetivo principal avaliar a qualidade de voz no serviço VoIP.

3.2 High Speed Downlink Packet Access (HSDPA)

À medida que aumenta o uso de serviços de pacotes de dados e que novos serviços são introduzidos, os usuários necessitam taxas de dados cada vez maiores junto a uma grande melhoria da Qualidade de Serviço (QoS) e os operadores necessitam maior capacidade em seus sistemas. Assim, para lidar com taxas cada vez mais altas, aprimorou-se o WCDMA, que foi introduzido a partir de uma especificação (do inglês *Release*) de número 5 do 3GPP, relativo à norma do WCDMA, e então passou a ser conhecido como HSDPA. Esse aprimoramento utiliza algumas técnicas já aplicadas em outras redes, como adaptação de enlace (do inglês *link adaptation*) e combinação de retransmissão rápida da camada física (do inglês *fast physical layer retransmission combining*).

Mudanças de arquitetura foram, portanto, necessárias para permitir essa retransmissão rápida e trazer o controle da adaptação de enlace para mais próximo da interface aérea. Tais mudanças também permitiram um processo suave de atualização para o HSDPA e asseguraram sua operação em ambientes onde nem todas as células suportam o HSDPA, podendo-se, então, comportar a crescente demanda por serviços de Internet, como o próprio VoIP.

O HSDPA foi desenvolvido, com base no conceito de transmissão por um canal compartilhado, para aumentar a capacidade de transmissão de dados no canal direto (*downlink*), em sistemas 3G, através da combinação de diversas técnicas.

No HSDPA, o fator de espalhamento variável e o controle rápido de potência, duas das principais características do WCDMA, são desativadas e substituídas por modulação e codificação adaptativa, operações extensivas de códigos múltiplos (no HSDPA pode-se usar até 15 códigos paralelos) e uma estratégia de retransmissão rápida e espectralmente eficiente. O uso de codificações mais robustas, sistema de retransmissão híbrida e operações com códigos múltiplos tornam desnecessário o uso de fator de espalhamento variável.

No WCDMA, se um usuário estiver perto da Estação Rádio-Base (ERB) ou se

a qualidade do canal for boa, ele poderá ter sua potência reduzida. No entanto, ela não pode ser reduzida além de um valor mínimo e, de qualquer forma, mesmo que esteja abaixo de um dado limiar, o impacto dessa redução na capacidade do sistema será muito pequeno. Além disso, o usuário não se beneficia dessa boa qualidade do canal para obter um ganho em eficiência espectral. Com o HSDPA, esse mesmo usuário poderá aproveitar a maior disponibilidade de potência, já que esse modo de transmissão é capaz de transformar o excesso de potência em maior vazão, sem custo adicional. A introdução do uso do HSDPA, portanto, não só acrescenta capacidade ao sistema, como também faz crescer a eficiência espectral para o usuário, oferecendo-lhe benefícios diretos.

O canal de transporte carregando os dados do usuário com operação HSDPA é denotado por HS-DSCH (do inglês *High Speed Downlink Shared Channel*).

Todos os canais HS-DSCH do HSDPA são multiplexados por código utilizando um fator de espalhamento (SF) de 16. Isto significa que as informações de todos os canais HS-DSCH são transportadas em um único canal, de forma que os sinais transmitidos são espalhados sobre uma largura de faixa muito maior do que aquela na qual a informação está contida.

Um máximo de 15 códigos paralelos são alocados ao canal HS-DSCH, sendo que esses códigos podem ser todos designados para um único usuário, durante um certo intervalo de tempo de transmissão (TTI, do inglês *Transmission Time Interval*), ou podem ser divididos entre vários usuários.

As técnicas empregadas para garantir essas altas taxas de dados, mantendo-se a compatibilidade com os equipamentos já existentes são:

- ▶ Transmissão por canal compartilhado;
- ▶ Modulação de maior ordem;
- ▶ Modulação e Codificação Adaptativas (AMC);
- ▶ Intervalo de tempo de transmissão pequeno (TTI);
- ▶ Rápida adaptação de enlace;
- ▶ Escalonamento rápido;
- ▶ Rápida retransmissão (HARQ rápido);

- ▶ Seleção da melhor célula para transmissão de dados (FCSS, do inglês *Fast Cell Site Selection*).

Com relação à transmissão por canal compartilhado, os códigos do canal HS-DSCH e a potência de transmissão são vistos como recursos comuns que são dinamicamente compartilhados entre os usuários, no domínio do tempo e do código, implicando em um uso dos códigos disponíveis e da potência mais eficiente do que antes se obtinha no WCDMA. Como já se disse aqui, o código compartilhado no qual o HS-DSCH é mapeado consiste em 15 códigos, seu fator de espalhamento (SF) é fixado em 16 e o número efetivamente empregado depende da capacidade do sistema e da quantidade de códigos suportada pelo terminal e pelas configurações da operadora.

Chama-se modulação o processo pelo qual os símbolos a serem transmitidos são multiplicados pelo sinal de outra onda, chamada de portadora, modificada para representar a mensagem a ser enviada, e cujas propriedades são mais convenientes aos meios de transmissão. Inicialmente foram propostas as modulações QPSK (do inglês *Quadrature Phase Shift Keying*), 8-PSK (do inglês *8 Phase Shift Keying*), 16-QAM (do inglês *16 Quadrature Amplitude Modulation*) e 64-QAM (do inglês *64 Quadrature Amplitude Modulation*). Entretanto, as modulações 8-PSK e 64-QAM foram descartadas por razões de performance e complexidade. Foi constatado também que a modulação 8-PSK quase não acrescentava ganho de capacidade, enquanto que o 64-QAM era praticamente inviável devido a seus altos requisitos de qualidade, ficando como viáveis as QPSK e 16-QAM.

Na solução atualmente utilizada para a modulação, o canal HS-DSCH pode usar tanto a QPSK quanto a 16QAM para prover altas taxas de transmissão (cf. Figura 3.2). A segunda delas, porém, possui o dobro da capacidade de transmissão de taxa de bits por símbolo do que tem a primeira, possibilitando um uso mais eficiente da banda, exigindo uma melhor condição do canal, além de permitir que o HSDPA atinja taxas de até 10 Mbps de pico com 15 códigos paralelos e fator de espalhamento igual a 16.

Essas modulações de alta ordem são muito mais sensíveis a ruído e interferências. Por isso, elas devem ser utilizadas em conjunto com a técnica de Modulação e Codificação Adaptativas (AMC), que escolhe o esquema de modulação e codificação, chamado de modo de transmissão, mais adequado de acordo com as condições instantâneas do canal.

Nos serviços por pacotes de dados, chegam freqüentemente ao receptor pacotes

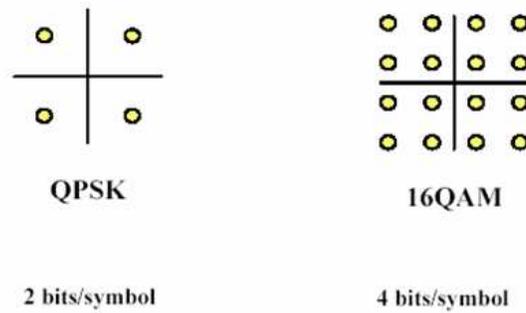


Figura 3.2: Esquema de Modulação usado com o HSDPA [28].

com erros. Por isso, o uso de esquemas de Retransmissão Automática Híbrida (HARQ) uma combinação dos protocolos FEC (do inglês *Forward Error Correction*) com protocolos ARQ (do inglês *Automatic Repeat Request*) [72] torna o sistema mais robusto frente a possíveis erros na adaptação do enlace.

O uso das técnicas AMC e HARQ, além de possibilitar operações com códigos múltiplos (pode-se usar até 15 códigos paralelos no HSDPA), torna desnecessário o uso de duas das principais características do WCDMA: o fator de espalhamento variável e o controle rápido de potência.

Para explorar a adaptação de enlace e o escalonamento dependente de canal, os recursos compartilhados de código são alocados dinamicamente a cada 2 ms, ou seja 500 vezes por segundo. Esse pequeno intervalo de tempo de transmissão (TTI) reduz o tempo de viagem de ida e volta, além de melhorar o rastreamento das variações do canal.

A rápida adaptação de enlace caracteriza-se pela veloz adaptação dos parâmetros de transmissão para as condições do canal, a cada instante, e pela habilitação da modulação de ordem maior, quando o canal permitir, já que as condições do canal variam significativamente segundo diferentes posições da célula para o *downlink* e segundo o tempo de transmissão.

Com o objetivo de transmitir com canal mais favorável para o usuário, em cada instante, a técnica de escalonamento rápido determina para qual usuário o canal de transmissão compartilhado deverá ser direcionado. Esse processo determinando o desempenho global do HSDPA, pois para cada TTI, o escalonador decide qual usuário HS-DSCH deverá transmitir e, em colaboração com a adaptação de enlace, decide-se qual modulação e quantos códigos devem ser usados, resultando na taxa de transmissão de dados que o usuário efetivamente receberá. O ganho obtido nas transmissões para o usuário sempre com as melhores condições de canal é conhecido

pela expressão diversidade multiusuário.

Dentre os principais algoritmos de escalonamento de pacotes para serviços de melhor esforço [73–79], listam-se:

- ▶ *Round-Robin* (RR): os pacotes são servidos de forma cíclica e uniforme, de modo que nenhum parâmetro é considerado ao distribuir o acesso dos pacotes ao canal. Cada usuário requer sua própria fila e os pacotes são servidos um por vez, desconsiderando a qualidade do canal, de modo que todos os usuários têm acesso ao canal ciclicamente;
- ▶ *Wiegthed Round-Robin* (WRR): trabalha sobre o mesmo princípio geral que rege o Round-Robin, porém esse algoritmo é aplicado quando os usuários têm requerimentos diferentes, de modo a que cada usuário passe a ser servido de acordo com seus requerimentos de taxas. Cada tipo de classe é definido com um peso, e a implementação oferecida por esse algoritmo consiste em ponderar qual quantidade de pacotes de cada classe deve ser transmitida por vez;
- ▶ *Proportional Fair* (PF): escalona pacotes de usuários de acordo com a condição instantânea de canal e um histórico de suas transmissões anteriores. Quando um usuário é ignorado por muito tempo, o algoritmo leva-o em consideração por ter sido ignorado por muito tempo, selecionando-o para transmitir;
- ▶ *Max SNR ou Best Channel Scheduling* (BCS): prioriza a transmissão de pacotes dos usuários com a melhor condição de canal, porém pode chegar a ignorar completamente o usuário com péssima condição de canal, deixando-o, nesse caso, sem nenhuma possibilidade de transmitir;

Outra característica importante na garantia de altas taxas de dados no HSDPA, é o fato de que um usuário pode pedir rapidamente a retransmissão (HARQ rápido) de um dado que chegou corrompido, até um número máximo permitido, combinando a informação da transmissão original com a da transmissão subsequente antes de decodificar a mensagem. Destaca-se aqui a abordagem chamada de *soft combining* que, além de aumentar a capacidade, provê maior robustez. Uma não confirmação de recebimento (NACK) é enviada quando estiver faltando um dado e uma confirmação (ACK) é enviada quando o dado é recebido corretamente. Existem também os processos paralelos (do inglês *stop-and-wait process*), cuja funcionalidade é a de fazer com que um pacote não precise esperar muito tempo pela sua confirmação, possibilitando que outro seja transmitido em seguida.

Outra técnica disponível é a FCSS - Seleção da melhor célula para transmissão de dados, que reconhece quando a estação móvel (EM) ou o equipamento do usuário (UE) está funcionando dentro da rede, criando a possibilidade de que mais de um Node B ou ERB venha a se comunicar com ele. Neste caso, cada EM construirá uma lista chamada *Active Set*, contendo todas as estações de Node B que poderão ser usadas para estabelecer a comunicação. Assim, FCSS permite que um equipamento móvel possa selecionar, no seu *Active Set*, a célula com as melhores características para transmitir seus pacotes de dados, obtendo-se, como vantagem deste sistema, altas taxas de dados, atingidas pela escolha sempre do melhor *Node B* para a transmissão.

Além de todas as técnicas descritas até aqui, o HSDPA ainda utiliza um esquema de Modulação e Codificação Adaptativa (AMC) que procede de modo a alocar taxas de dados mais altas para usuários com o canal em condições mais favoráveis. Baseando-se na qualidade do canal, monitorada pelo equipamento do usuário, são otimizadas as seguintes características:

- ▶ Modulação: usa-se a QPSK ou a 16QAM;
- ▶ Taxa de código efetiva do canal (code rate): é variável, podendo ser 1 ou de 1/4 a 3/4;
- ▶ Número de códigos empregados: de 1 a 15;
- ▶ Potência transmitida por código;
- ▶ A máxima taxa de dados de pico (14,4 Mbit/s) seria atingida teoricamente com modulação 16QAM, taxa de código igual a 1 e o emprego de 15 códigos simultâneos.

3.2.1 Arquitetura do HSDPA

Algumas atualizações dos protocolos de sistemas anteriores foram necessárias para suporte das novas capacidades do HSDPA [28], adicionando-lhes novas funcionalidades com o intuito de manter a compatibilidade com as novas tecnologias.

A pilha de protocolos presentes nos componentes da rede HSDPA é composta pela camada física e pela camada de enlace, subdividindo-se esta última em duas subcamadas chamadas de MAC (do inglês *Medium Access Control*) e RLC (do inglês *Radio Link Control*) como se pode ver na Figura 3.3 [80, 81].

A camada física é responsável por transmitir, através da interface de rádio, o dado ou a voz, no caso do serviço VoIP, por exemplo. A subcamada MAC é responsável pela otimização do uso dessa interface de rádio, através do seu compartilhamento dinâmico entre os usuários e através do seu escalonamento do fluxo de dados ou de voz de acordo com o requerimento de QoS, no qual o fluxo de áudio tem prioridade. Este é o caso, por exemplo, do serviço VoIP que tem alto requerimento de QoS. A RLC é responsável pelos controles de erro e de fluxo de tráfego, tendo modos específicos de seleção apropriados ao tipo de dado trafegado. No caso do VoIP, usa-se o modo não confirmado (UM, do inglês *Unacknowledged Mode*), por não exigir confirmação nem garantia de entrega, em detrimento das alternativas modo confirmado (AM, do inglês *Acknowledge Mode*) que tem baixa tolerância a erros, exigindo confirmação na entrega, e modo transparente TM, do inglês *Transparent Mode*) usado em tráfego tolerante a atrasos [58, 82, 83].

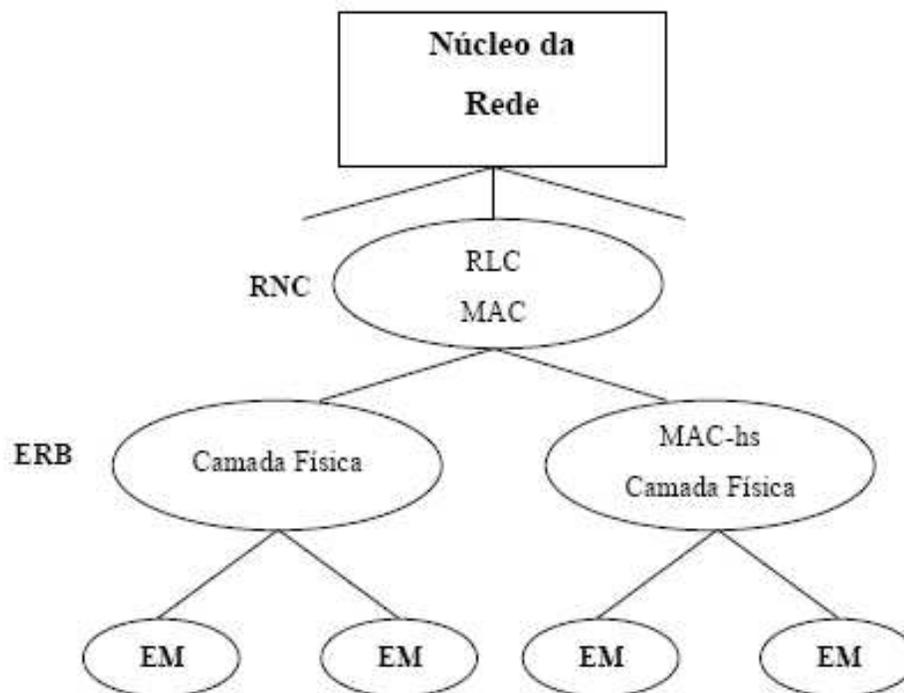


Figura 3.3: Arquitetura de Protocolo no HSDPA [80].

Na arquitetura do HSDPA há uma divisão da MAC em MAC-d (do inglês *Medium Access Control - Dedicated Transport Channels*) e MAC-hs (do inglês *Medium Access Control - High Speed Shared Transport Channel*). A MAC-d é destinada a suportar

canais de transporte dedicados, enquanto MAC-hs suporta o canal compartilhado HS-DSCH, além de prover um tempo menor de retransmissão, que era de 10ms e passa a ser de 2ms (chamado TTI).

Como se observa ainda na Figura 3.3, conectadas ao núcleo da rede estão as RNCs, cada uma delas capaz de controlar uma ou várias ERBs, mais comumente conhecidas como Node B, que, por sua vez, se comunicam com as estações móveis (EMs). As principais funções da RNC são o escalonamento, a seleção do formato do quadro (parâmetros de codificação, modulação, etc.) e as retransmissões.

Técnicas como a modulação, a codificação adaptativa, H-ARQ e o escalonamento rápido dos usuários exigem uma rápida adaptação às condições instantâneas do canal. Sendo assim, a rápida adaptação de enlace deveria ocorrer mais próxima da interface aérea, criando-se para isso uma nova entidade na ERB, a MAC-hs. No entanto, as atuais entidades da RNC, como RLC e MAC, ainda permanecem presentes para realizar uma série de funções, como a cifragem de dados, *soft handoff* etc. Dessa maneira, essas novas implementações na ERB não devem ser vistas como uma substituição às funções do RNC, mas como um complemento que possibilita um canal altamente confiável, com pequenos atrasos e com altas taxas de transmissão.

3.2.2 Camada Física do HSDPA

O HSDPA opera em conjunto com um canal dedicado (DPCH) que carrega serviços com exigências de atraso pequeno, como a voz, por exemplo. Assim, foram propostos três novos canais, além dos já existentes no WCDMA:

- ▶ HS-PDSCH (do inglês *High Speed Physical Downlink Shared Channel*)
- ▶ HS-PSCCH (do inglês *High Speed Physical Shared Control Channel*);
- ▶ Uplink HS-DPCCH (do inglês *Uplink High Speed Dedicated Physical Control Channel*)

O HS-PDSCH transporta os dados do usuário no enlace de descida com modulações de alta ordem e com taxas de pico atingindo até 10 Mbps.

O HS-PSCCH transporta as informações de controle da camada física, necessárias para uma correta decodificação dos dados do HS-PDSCH e também para possibilitar uma eventual combinação de dados enviados pelo HS-PDSCH, em caso de retransmissões.

O HS-DPCCH transporta informações de controle necessárias no enlace de subida, como mensagens de confirmação *Ack* e não-confirmação *Nack*, além de informações relativas à realimentação sobre a qualidade do enlace de descida.

HS-PDSCH

O HSDPA introduziu um novo canal físico, o HS-PDSCH, destinado ao transporte de pacotes de dados com serviço de melhor esforço, suportando todas as técnicas empregadas as altas taxas de dados. O HS-PDSCH será usado principalmente para pacotes de dados com QoS de melhor esforço e suportará todas as tecnologias de adaptação de enlace. Da mesma forma que o PDSCH, o HS-PDSCH corresponde a um recurso comum, compartilhado entre vários usuários primeiramente no domínio do tempo. A Figura 3.4 ilustra a parte da árvore de códigos que é separada para o HS-PDSCH, onde 12 códigos de fator de espalhamento 16 (SF=16) são separados para o HS-PDSCH e apenas 4 códigos são atribuídos a outros canais. Na prática, porém, esse número de códigos paralelos pode chegar a 15, pois o fator de espalhamento deste canal será fixo (igual a 16), não se utilizando fator de espalhamento variável como no WCDMA.

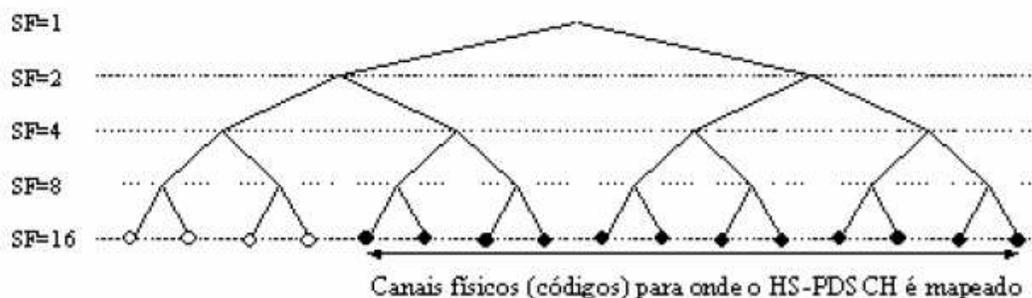


Figura 3.4: Códigos alocados para o HS-PDSCH.

A alocação de recursos de códigos compartilhados para diferentes usuários é feita para um intervalo de tempo de transmissão (TTI) do HS-PDSCH. Para reduzir os atrasos na transmissão e facilitar o processo de adaptação ao canal variante no tempo, esse intervalo de tempo de transmissão do HSDPA, que era de 10 ms no WCDMA, é fixado em 2 ms.

Inicialmente, a possibilidade de compartilharem-se os recursos de códigos do HS-PDSCH dá-se apenas no domínio do tempo, sendo que apenas um usuário transmite em um certo TTI. Poderia haver, porém, mais de um usuário transmitindo em um mesmo TTI, sendo eles separados apenas pelos códigos de canalização. Esta

possibilidade está ilustrada pela Figura 3.5, onde existem 3 usuários diferentes, havendo apenas um usuário transmitindo em alguns TTIs, enquanto em outros TTIs eles são mais de um. Caso um deles não possa preencher toda a capacidade do canal, os outros usuários poderão ser agendados para usar este mesmo TTI.



Figura 3.5: Compartilhamento do HS-PDSCH por multiplexação de códigos e de tempo.

Apesar da possibilidade de se ter até 15 códigos paralelos para um único usuário, isso depende muito da capacidade do terminal, pois existem terminais que suportam apenas 5, 10 ou 15 códigos paralelos, no máximo.

Com relação à codificação de canal, o WCDMA usa tanto códigos convolucionais quanto códigos turbo, enquanto que no HSDPA são usados apenas códigos turbo, nas taxas $\frac{1}{4}$, $\frac{1}{2}$ e $\frac{3}{4}$. As taxas teóricas de bits estão ilustradas na Tabela 3.2.

Tabela 3.2: Taxas de bits máximas para 15 códigos paralelos.

Modo de Transmissão	Modulação	Taxa de Código	Vazão Máxima(Mbps)
1	QPSK	1/4	1,8
2	QPSK	1/2	3,6
3	QPSK	3/4	5,3
4	16-QAM	1/2	7,2
5	16-QAM	3/4	10,7

O HS-PDSCH também suporta o H-ARQ, pelo qual é possível mandar retransmissões idênticas, técnica conhecida como combinação Chase, ou modificar os parâmetros e incrementar a redundância, técnica conhecida como redundância incremental.

A Tabela 3.3 mostra as principais diferenças entre o HS-PDSCH do HSDPA e o PDSCH do WCDMA.

Tabela 3.3: Comparação entre HS-PDSCH e PDSCH.

Característica	PDSCH	HS-PDSCH
Fator de Espalhamento Variável	Sim	Não
Controle de Potência Rápido	Sim	Não
Modulação e Codificação Adaptativa	Não	Sim
Operação de Códigos Múltiplos	Sim	Sim, estendida
H-ARQ	Não	Sim

HS-PSCCH

No HS-PDSCH, além da transmissão de dados, é preciso enviar ao usuário também a sinalização de controle, incluindo-se os códigos de canalização que ele deve desespalhar, o modo de transmissão utilizado e os parâmetros relacionados com H-ARQ. Este tipo de informação é relevante apenas para o usuário que está recebendo dados através do HS-PDSCH e, por isto, pode ser transmitida em um canal de controle compartilhado.

É o HS-PSCCH que carrega as informações necessárias para a demodulação do HS-PDSCH. O UTRAN tem de alocar um certo número de HS-PSCCHs, correspondente ao número máximo de usuários que serão multiplexados por códigos. Se nenhuma informação estiver sendo transmitida no HS-PDSCH, não há necessidade de transmitir um HS-PSCCH. Do ponto de vista da rede, pode haver vários HS-PSCCHs alocados, mas cada usuário só precisará considerar no máximo quatro HS-PSCCHs por vez.

Cada bloco HS-PSCCH tem a duração de três janelas de tempo e é dividido em duas partes, como está descrito na Figura 3.6 que mostra a relação de tempo entre os canais HS-PDSCH e HS-PSCCH. A primeira parte do bloco, correspondente à primeira janela de tempo, carrega todas as informações de tempo crítico que são necessárias para começar o processo de demodulação no tempo correto e evitar filas. Por sua vez, a segunda parte, que corresponde às outras duas janelas de tempo, contém parâmetros como o CRC (do inglês *Cyclic Redundancy Check*), para checar a validade da informação do HS-PSCCH, e informações do processo H-ARQ.

Ainda sobre a Figura 3.6, percebe-se que o terminal tem uma única janela de

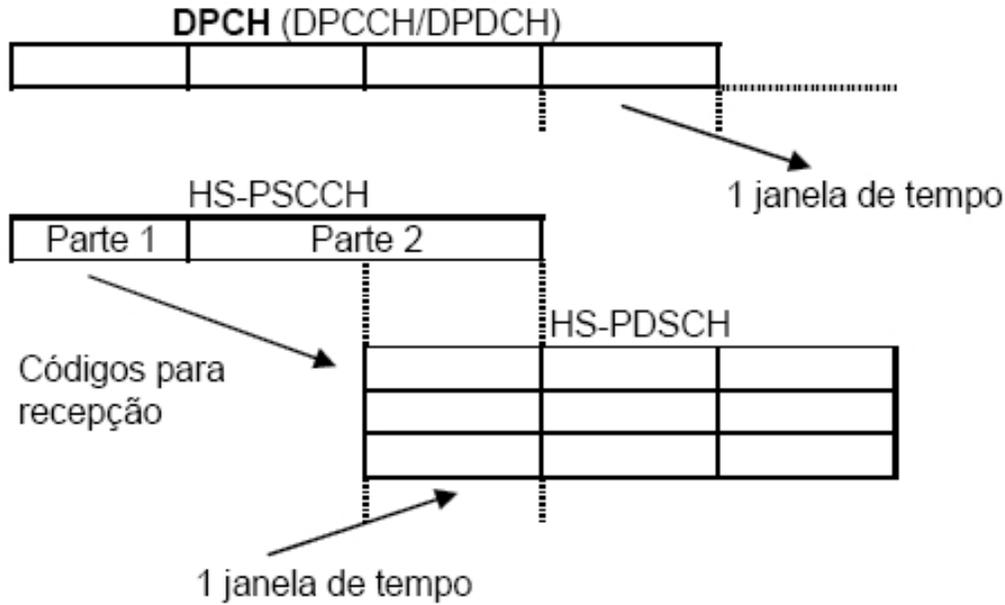


Figura 3.6: Relação de tempo entre os canais HS-PDSCH e HS-PSCCH.

tempo para determinar quais códigos do HS-PDSCH ele deve desespalhar. Quando, no HS-PDSCH, existirem dados para o mesmo usuário em TTIs consecutivos, o mesmo HS-PSCCH deve ser usado. Vê-se também que o tempo do canal dedicado DPCH não está atrelado ao tempo de nenhum dos dois canais HSDPA.

O HS-PSCCH usa fator de espalhamento igual a 128, acomodando 40 bits por janela de tempo, uma vez que não há símbolo piloto ou bits de controle de potência nesse canal. O HS-PSCCH usa código convolucional de taxa $\frac{1}{2}$ e as duas partes são codificadas separadamente, já que as informações de tempo crítico precisam estar disponíveis imediatamente depois da primeira janela de tempo.

Uplink HS-DPCCH ou HS-DPCCH do Enlace de Subida

Além da sinalização no enlace de subida, através do HS-PDSCH, precisa-se também de sinalização de controle neste enlace. O enlace de subida deve enviar as informações do H-ARQ, como as mensagens ACK e NACK, assim como a informação sobre a qualidade do canal, usada pela ERB para fazer a adaptação do enlace. O canal HS-DPCCH do enlace de subida é dividido em duas partes:

- ▶ Transmissão de mensagens ACK/NACK;
- ▶ Indicador da qualidade do canal no enlace de descida.

A informação de realimentação sobre a qualidade do canal consiste em 5 bits. Um dos estados é reservado para o estado 'sem transmissão', quando o canal estiver muito ruim. Os outros estados representam qual modo de transmissão o usuário pode receber no momento. Esses estados, então, vão desde QPSK com um único código até 16-QAM com 15 códigos paralelos, como se pode ver na Tabela 3.2.

3.2.3 Considerações Finais sobre o HSDPA

Conforme ao que já se disse em trecho anterior deste texto, o HSDPA oferece taxas de transmissão mais altas do que as do WCDMA, aumentando a capacidade do sistema de vários modos, como os que detalhamos a seguir:

- ▶ A utilização de canal compartilhado resulta em um uso mais eficiente dos recursos disponíveis no WCDMA;
- ▶ O uso de um TTI mais curto reduz o tempo da viagem de ida e volta, aumentando a capacidade de rastreamento das rápidas variações do canal;
- ▶ A adaptação de enlace maximiza o uso do canal, permitindo que a estação base opere perto da potência máxima da célula;
- ▶ O escalonamento rápido prioriza usuários com as condições de canal mais favoráveis;
- ▶ A rápida retransmissão aumenta ainda mais a capacidade do sistema;
- ▶ A modulação 16QAM permite taxas de bits maiores;
- ▶ O ganho combinado de capacidade varia entre duas a três vezes a do WCDMA, dependendo do cenário;

3.3 Sumário

Neste capítulo da Dissertação foram introduzidas as características do UMTS, sua arquitetura, suas classes de serviço, entre elas o VoIP, além da descrição do sistema HSDPA com suas características, arquitetura, camada física e aspectos mais relevantes para a inserção do serviço de voz sobre o protocolo Internet.

Avaliação da Qualidade de Voz do Serviço VoIP em Sistemas HSDPA

O serviço de VoIP, atualmente, consolida-se cada vez mais como alternativa à telefonia tradicional. Seu uso em sistemas celulares como o HSDPA vem sendo objeto de várias análises, sobretudo em busca de resolver o desafio da superação do problema do atraso, inerente à esses sistemas, para garantir a qualidade de voz [84, 85].

Este Trabalho tem a intenção de contribuir com essas análises e, para isso, procedeu-se a uma avaliação da qualidade de voz do VoIP sobre HSDPA, utilizando-se os métodos objetivo e subjetivo, com a finalidade de obter um conceito da qualidade percebida pelo usuário final. O método objetivo é referenciado pelo algoritmo PESQ e o método subjetivo pelo escore de opinião MOS. Trabalhando-se com ambos os métodos, pode-se comparar as duas metodologias e chegar a uma análise mais precisa da qualidade final de voz.

Este capítulo aborda processos de simulação e de avaliação da qualidade de voz do serviço VoIP no sistema HSDPA, introduzindo-se, na seção 4.1, algumas considerações a respeito do serviço de voz sobre IP em HSDPA. Na seção 4.2, levando-se em conta a contextualização do serviço VoIP no HSDPA, faz-se uma breve descrição da modelagem do simulador, apresentando-se uma visão geral das principais características de modelagem tanto da pilha de protocolos (na seção 4.2.1) quanto da rede HSDPA através da seção 4.2.2. Em seguida, listam-se os parâmetros de simulação utilizados (na seção 4.2.3) e apresenta-se uma visão geral dos arquivos de voz (na seção 4.2.4). Na seção 4.2.5, é descrita a metodologia de todo o processo de avaliação realizado na pesquisa que produziu esta Dissertação e, finalmente, são

apresentados os seus resultados na seção 4.3.

4.1 Considerações sobre a oferta do serviço VoIP sobre HSDPA

O atual desenvolvimento em tecnologias de rede de rádio para celular está capacitando essas redes a proverem serviços baseados no protocolo Internet. Redes 3G WCDMA, *release 5* (HSDPA), por exemplo, provêm serviços IP para redes sem fio, além de habilitar os serviços de VoIP [84]. O provimento desses serviços emergentes, de tempo real, podem ser de especial interesse para operadores de redes móveis por uma série de razões. Primeiramente, uma vez que a largura de banda para fluxos individuais em circuitos chaveados por pacotes não é reservada com antecedência, os efeitos da multiplexação trazem uma significativa economia de capacidade. Além disso, a convergência em direção aos circuitos chaveados por pacotes e a tecnologia IP podem convencer operadores móveis a migrarem para soluções que são verdadeiramente provedoras de serviços totalmente baseados no protocolo Internet, simplificando-se a interconexão e o gerenciamento de rede.

Serviços de tempo real são geralmente transportados sobre canais dedicados por conta da sensibilidade ao atraso enquanto que dados são geralmente transportados por canais compartilhados por conta da sua natureza de tráfego em rajada. Embora os canais compartilhados dos padrões 3G serem originalmente desenvolvidos com pacotes de dados em mente, modificações recentes para os padrões foram feitas para aumentar o suporte à serviços de tempo real e, em particular, serviços de baixa largura de banda tal como o VoIP [86].

As aplicações de voz, pelo fato de serem de tempo real, são intolerantes ao atraso, à perda de pacotes, à seqüência fora de ordem dos pacotes e à variação do atraso. Todos esses problemas degradam a qualidade da voz transmitida para o receptor. Infelizmente, redes sem fio potencializam os problemas que prevalecem nas redes com fio: uma alta freqüência de pacotes perdidos, alto tempo de latência e uma grande variação do atraso. Como possibilidade de mitigar a perda de pacotes e a conseqüente degradação na qualidade final do serviço VoIP sobre HSDPA, pode-se investigar a relação entre o aumento da quantidade de usuários no sistema e a quantidade de pacotes perdidos. Outra forma de investigação útil ao mesmo escopo, seria a análise de impacto do acúmulo de pacotes de voz no RTP do transmissor, medindo-se a quantidade de pacotes perdidos e a qualidade final pois, se o RTP tem a possibilidade de armazenar até quatro pacotes, antes de passar para as

camadas seguintes, esses pacotes podem chegar muito atrasados e ser descartados, comprometendo a qualidade.

O desempenho do VoIP é medido mais precisamente em termos de taxa de perda de pacotes [73]. No HSDPA, as principais causas desse tipo de perdas são: descarte de pacote ao nível da MAC, devido ao atraso inserido pelo enfileiramento (espera pelo escalonamento), erro devido à interface aérea (canal rádio-móvel) e pacotes descartados no buffer do receptor devido à ultrapassagem do valor do requerimento de atraso (número elevado de retransmissões). Ocorre ainda que as funcionalidades de retransmissão junto com processos paralelos, na MAC do HSDPA, podem provocar não só a variação do atraso como a entrega de pacotes fora de ordem. Todos estes eventos afetam de forma direta a qualidade da voz transmitida.

O sistema HSDPA provê um ganho de desempenho para o tráfego que é geralmente limitado por requerimentos de atraso [14,16]. Entretanto, portadoras de acesso de rádio conversacional podem ser mapeadas para um canal compartilhado de *downlink* de alta velocidade (HS-DSCH), mas seu desempenho dependerá de como se gerencia a eficiência dos recursos de rádio em diferentes cargas de sistema. Um dos objetivos na operação das redes emergentes de alta velocidade de dados, como UMTS-HSDPA, é o de prover qualidade de serviço, em particular conhecendo os requerimentos de taxa de dados e atraso de pacote dos usuários de voz em tempo real.

Uma vez que a finalidade principal deste estudo é de contribuir para esse conhecimento, neste capítulo serão apresentadas considerações a respeito de toda a modelagem necessária para o provimento do serviço VoIP sobre o sistema HSDPA e serão expostos os procedimentos e resultados desta pesquisa de avaliação da qualidade final da voz transmitida.

4.2 Modelagem para Simulação

Pelo fato de que as redes de comunicações modernas têm um grau elevado de complexidade, as simulações computacionais surgem como alternativa para emular o seu funcionamento em determinados aspectos, obtendo-se uma caracterização de seu desempenho em um cenário próximo da realidade.

Com o objetivo de investigar as características relativas ao sistema HSDPA - UTRAN, como parte de um conjunto integrado de pesquisas, o Grupo de Pesquisa em Telecomunicações Sem Fio (GTEL), da Universidade Federal do Ceará (UFC), criou um simulador dinâmico ao nível sistêmico, chamado UTRANSIM, em cuja

elaboração o autor desta Dissertação teve participação direta no desenvolvimento das classes no nível de aplicação e nos processos de medição da qualidade objetiva e subjetiva da voz. Sua validação é feita comparando-se os resultados obtidos com os resultados de outras ferramentas de simulação semelhantes e com os resultados reportados na literatura pertinente, entre os quais podem ser encontrados, especificamente em [73], a descrição mais detalhada e os resultados de validação.

O simulador UTRANSIM foi construído em C++, usando a estrutura de orientação a objetos, dado o seu melhor desempenho computacional comparado a outras linguagens de programação. Ele modela as principais características do sistema HSDPA, as funcionalidades necessárias para o tráfego do serviço VoIP e a para a medição da sua qualidade.

O processo de simulação de provimento do serviço VoIP, desenvolvido neste trabalho, desde o transmissor até a chegada ao receptor será descrito como um usuário que fala em seu terminal móvel, deixando-se implícitos os processos de digitalização, codificação e geração do quadro de voz. Depois de gerado, um quadro de voz ou *frame* (expressão em inglês, mais conhecida) é segmentado em pacotes RTP, que depois são segmentados em pacotes UDP/IP e em seguida em pacotes das camadas RLC e MAC do HSDPA. Esses pacotes são armazenados no transmissor, em um *buffer* da subcamada MAC-hs, com um tempo de descarte de pacotes. A cada passo de transmissão dos pacotes, o sistema verifica se o tempo de descarte foi ultrapassado e, em caso afirmativo, descarta-os antes da transmissão; se o usuário for escalonado, entretanto, o pacote será transmitido pelo canal sem fio. Obtendo sucesso na transmissão ou nas retransmissões pelo canal sem fio, o pacote faz o processo inverso pela pilha de protocolos e se não tiver ultrapassado o máximo atraso permitido para a chegada no receptor - pode ser ouvida e avaliada a qualidade da voz final.

4.2.1 Pilha de Protocolos no UTRANSIM

A Figura 4.1 mostra o fluxo da informação através da pilha de protocolos usados pelo serviço VoIP. O codificador gera um quadro de voz a cada 20 ms, com uma taxa de 12.2 kbps, correspondendo a um pacote de 256 bits. Esse quadro de voz é entregue às camadas RTP/UDP/IP, que tratam de inserir neste ponto seus cabeçalhos. Se existe algum algoritmo de compressão de cabeçalho ou HC, ele comprime os cabeçalhos RTP/UDP/IP para o menor tamanho possível, com a finalidade de economizar largura de banda no momento de transmitir pelo canal sem

figo. A camada de enlace abrange os protocolos da interface de rádio da RAN. Neste Trabalho escolheu-se a camada de enlace do sistema WCDMA/HSDPA, composta pelas subcamadas RLC e MAC-hs. No receptor, o quadro de voz é recebido pela camada de enlace e encaminhado para as camadas superiores até chegar ao decodificador. Nas simulações feitas aqui, o fluxo de informações é modelado em detalhes, uma vez que todos os pacotes são gerados e transmitidos através das camadas de protocolos.

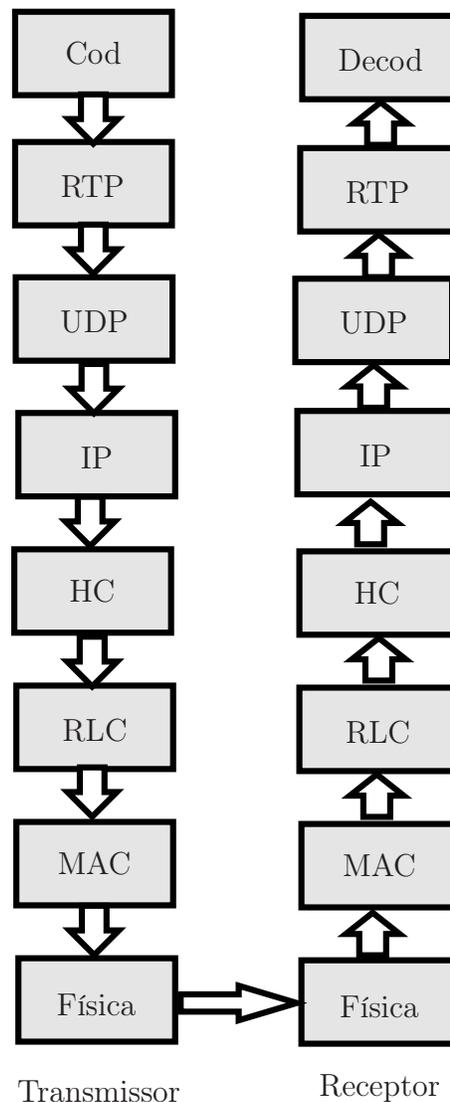


Figura 4.1: Fluxo da Informação na Pilha de Protocolos.

A modelagem do tráfego de voz ON/OFF do usuário é implementada atualmente com a leitura dos períodos de fala de acordo com um detector de atividade de voz,

que suporta no máximo 256 bits a cada 20 ms de períodos de atividade. Durante os períodos de silêncio, o transmissor é considerado como desligado.

O RTCP, é responsável pelo monitoramento de uma sessão através da troca de informações de controle entre o transmissor e o receptor, não é modelado nessas simulações, pois nesta abordagem não existem informações de controle sobre os participantes da sessão, como foi já se mencionou no Capítulo 2.

O requerimento de atraso do VoIP é caracterizado na camada RTP, tendo neste projeto um valor de 80 ms. Isto significa que se o pacote de voz chega ao receptor com um atraso maior que 80 ms é considerado como pacote perdido, incrementando a métrica de FER, que é computada no simulador. Note-se, porém, que o *buffer* do RTP receptor - normalmente usado no combate ao problema de variação do atraso (*delay jitter*) e da entrega de pacotes fora de ordem - não foi implementado na atual versão do simulador. A razão disso é a de que aqui se considera que um pacote, gerado no transmissor, percorrendo o canal de comunicação e chegando livre de erros ao receptor será imediatamente usado pela camada de aplicação.

Embora a segmentação do RTP e UDP esteja implementada no simulador, ela não será utilizada no cenário de simulação porque se assumiu que o quadro de voz gerado no codificador é encapsulado em um simples segmento RTP e, em seguida, em um só segmento UDP.

Finalmente, o simulador usa a técnica de compressão de cabeçalho (HC), reduzindo um segmento RTP/UDP/IP de 40 bytes para 4 bytes, para isso utilizando o algoritmo de compressão de cabeçalho ROHC, assumindo-se a exigência de diminuição do tamanho do cabeçalho a ser transmitido em redes sem fio, do grande consumo de largura de banda e do uso ineficiente do espectro de rádio.

4.2.2 Rede HSDPA no UTRANSIM

Adota-se o modo UM, uma vez que o VoIP é um serviço de tempo real com restrito requerimento de atraso, sem tolerância para com o atraso causado pelas retransmissões na camada RLC, que são permitidas pelo modo confirmado (AM).

Com relação à camada MAC, foram modeladas algumas funcionalidades da subcamada MAC-hs, de uma forma simplificada. No caso do pacote de voz ser recebido com erro no receptor, é permitido um número máximo de 5 retransmissões no nível da MAC, reenviando-se o pacote até que seja recebido corretamente ou até que se alcance o número máximo de retransmissões, considerando-se que as retransmissões adicionam atraso aos fluxos. Além disso, foram adotados 6 processos

paralelos, de tal forma que os pacotes de voz possam ser levados simultaneamente por processos que aguardam sua vez de usar o canal e transmitir na MAC.

Adotou-se nas simulações o algoritmo de escalonamento de pacotes RR, com o qual cada usuário tem a mesma probabilidade de acessar o canal compartilhado, por escalonamento cíclico, proporcionando-se uma distribuição igual de recursos entre todos os usuários na fila.

4.2.3 Parâmetros de Simulação

Os principais parâmetros utilizados no UTRANSIM, através de simulações, e seus respectivos valores padrões estão resumidos na Tabela 4.1.

Tabela 4.1: Parâmetros de Simulação.

Parâmetros	Valores Padrões	Valores Utilizados
Taxa da voz em atividade	Entre 4.75 e 12.2 kbps	12.2 kbps
Taxa da voz em silêncio	0, 1 ou 6 kbps	0 kbps
Duração dos quadros de voz AMR	20 ms	20 ms
Duração média das chamadas de voz	60 ou 120 s	60 s
Duração média dos períodos ON/OFF	3 s	3 s
Tamanho dos pacotes de voz gerados pelo AMR	256 bits	256 bits
Tamanho do cabeçalho RTP	96 bits	96 bits
Tamanho do cabeçalho UDP	64 bits	64 bits
Tamanho do cabeçalho IPv4	160 ou 320 bits	160 bits
Modo RLC	Modo Confirmado (AM), Não-Confirmado (UM) ou Transparente (TM)	Modo Não-Confirmado (UM)
Número máximo de retransmissões na MAC-hs	Variável	5
Número de processos paralelos na MAC-hs	Variável	6
Algoritmo de escalonamento de pacote	<i>Round-Robin (RR), Wiegthed Round-Robin (WRR), Proportional Fair (PF), Max SNR ou Best Channel Scheduling (BCS)</i>	<i>Round-Robin</i>
Requerimento de Atraso	Entre 80 e 150 ms	80 ms

Justifica-se a escolha e o impacto de cada um dos parâmetros definidos na Tabela

4.1 - taxa de atividade de voz de 12.2 kbps, a conseqüente duração de quadros de voz de 20 ms e o tamanho de 256 bits dos pacotes de voz gerados - pelo fato de que estes são os padrões do codificador AMR no serviço VoIP, definidos pelo 3GPP. Não foi utilizada taxa para o período de silêncio por não se haver modelado nenhum procedimento de tratamento de ruído durante a transmissão. A duração média das chamadas de voz foi escolhida com base na bibliografia relativa ao tema e pelo fato de um trecho menor de chamada de voz ser mais fácil para avaliar-se de forma subjetiva. A duração média de atividade e inatividade de voz é de 3 segundos, porque é o que definem as normas. Pela mesma razão, o tamanho dos cabeçalhos do RTP/UDP/IP é de 320 unidades de bits. O modo UM da camada RLC do HSDPA também foi utilizado por não necessitar de confirmação de entrega, característica determinante para o serviço VoIP que é sensível ao atraso. A quantidade de retransmissões e o número de processos paralelos foram definidos em 5 e 6, respectivamente, por não serem valores altos demais para iniciar uma nova transmissão e para a espera de uma confirmação, permitindo a transmissão de outro pacote logo em seguida. O algoritmo de escalonamento de pacote *Round-Robin (RR)* foi escolhido por transmitir os pacotes de maneira cíclica e uniforme. Todos os valores aqui utilizados geraram resultados e conclusões que serão descritos nas próximas seções.

4.2.4 Arquivos de Voz

Foram escolhidas, para este estudo, quatro vozes diferentes de adultos, sendo duas masculinas e duas femininas. Esta escolha foi baseada em [61,87] e também no fato de que os adultos emitem mais claramente a voz, cujo timbre presta-se melhor à digitalização. As vozes foram gravadas em português com o *software* SoundMax Digital Audio versão 5.1 da Microsoft, usando-se um microfone de marca Clone (*Multimedia System*) com o formato PCM e atributos 8 KHz, 16 bits, Mono. O conteúdo das falas é composto de frases aleatórias, sem ligação entre si, pois o objetivo era avaliar a inteligibilidade geral de toda a fala durante a escuta de tais frases.

Os sinais de voz foram digitalizados pelo software editor de voz digital *GoldWave* versão 5.10 de teste, disponível em [88]. Nele, os quatro arquivos de extensão .wav com duração média de 60 segundos cada um foram divididos em segmentos para a avaliação subjetiva. As Figuras 4.2 e 4.3 mostram os arquivos de voz por completo do sexo masculino gravados com os seus sinais digitalizados para transporte pelo simulador UTRANSIM. Em seguida, nas Figuras 4.4 e 4.5 aparecem

os sinais digitalizados femininos também completos e de mesma duração, onde há uma distribuição mais igualitária entre os períodos de atividade de voz (ON) e inatividade (OFF).

Sobre a Figura 4.2, especificamente, verifica-se a existência de um período sem sinal ou também chamado período OFF com uma duração maior que os outros OFF, podendo influenciar tanto na avaliação da qualidade objetiva quanto na subjetiva, além de ser um período pouco maior a 5 segundos sem atividade de voz, pois há uma barra numérica na parte de baixo indicadora de referência em segundos. Já nas figuras de Homem 2, Mulher 1 e Mulher 2 seus períodos de atividade e inatividade de sinal de voz são intercalados de uma forma mais constante sem haver um comportamento que chame atenção de uma forma tão diferenciada.

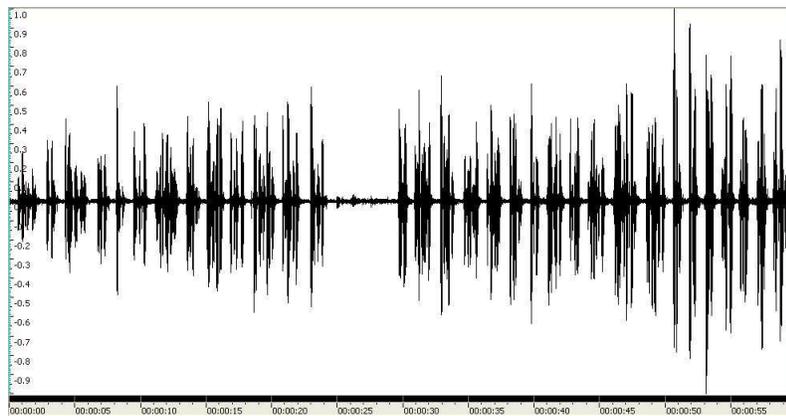


Figura 4.2: Sinal da Voz Digitalizada do Arquivo Homem 1. O eixo das abscissas indica a duração dos trechos de voz em segundos e o eixo das ordenadas reflete a amplitude do sinal, expressa em decibéis (dB), retratando a variação da frequência da voz para baixo e para cima.

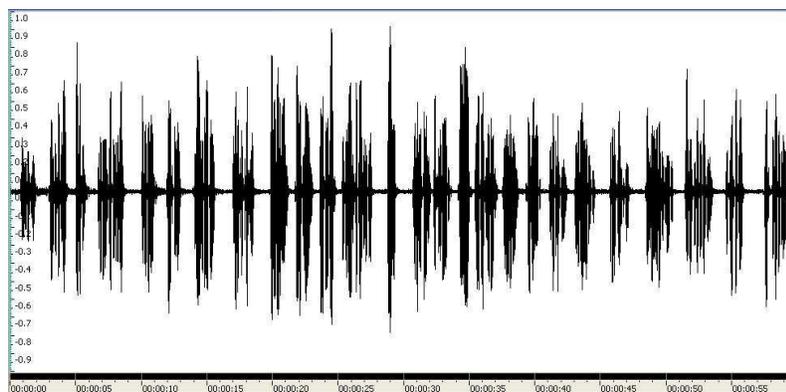


Figura 4.3: Sinal da Voz Digitalizada de Homem 2

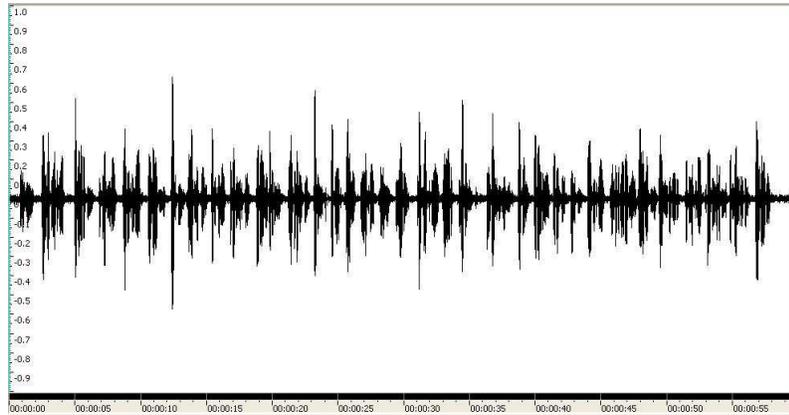


Figura 4.4: Sinal da Voz Digitalizada de Mulher 1

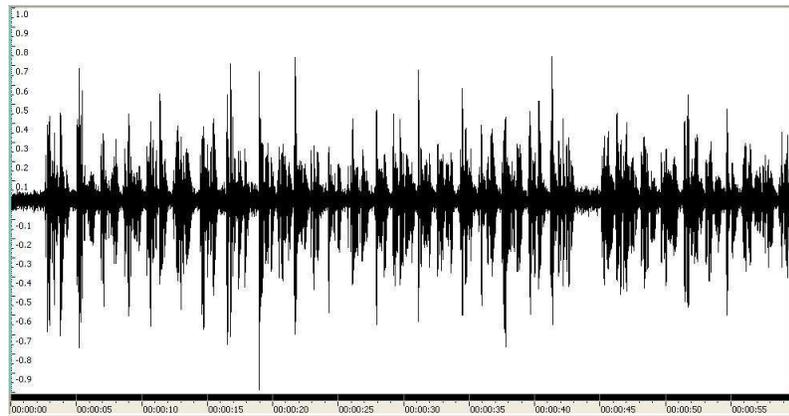


Figura 4.5: Sinal da Voz Digitalizada de Mulher 2

4.2.5 Estrutura das Avaliações Objetiva e Subjetiva

A estrutura desta Dissertação de Mestrado pode ser melhor visualizado através da Figura 4.6, que apresenta esquematicamente a seqüência dos passos dados para obterem-se vozes disponíveis para a avaliação tanto objetiva quanto subjetiva.

Uma vez selecionadas e digitalizadas as vozes para utilização neste Trabalho, (quatro arquivos, dois de vozes masculinas e dois de vozes femininas) retiraram-se seus cabeçalhos, mantendo-se somente a informação de voz a ser codificada, empacotada e transmitida pelo simulador UTRANSIM. Transformou-se, através de um algoritmo, esta informação em arquivo binário para indicar atividade e inatividade de voz, representando-se um detector de atividade de voz. Cada bit indicativo de atividade de voz, que representa um trecho de voz, foi transmitido pelo UTRANSIM, podendo ser ou não perdido durante a transmissão. Caso fosse perdido, um marcador de perdas naquele trecho era colocado em um arquivo de saída chamado padrão de erro, que será usado no momento da decodificação do

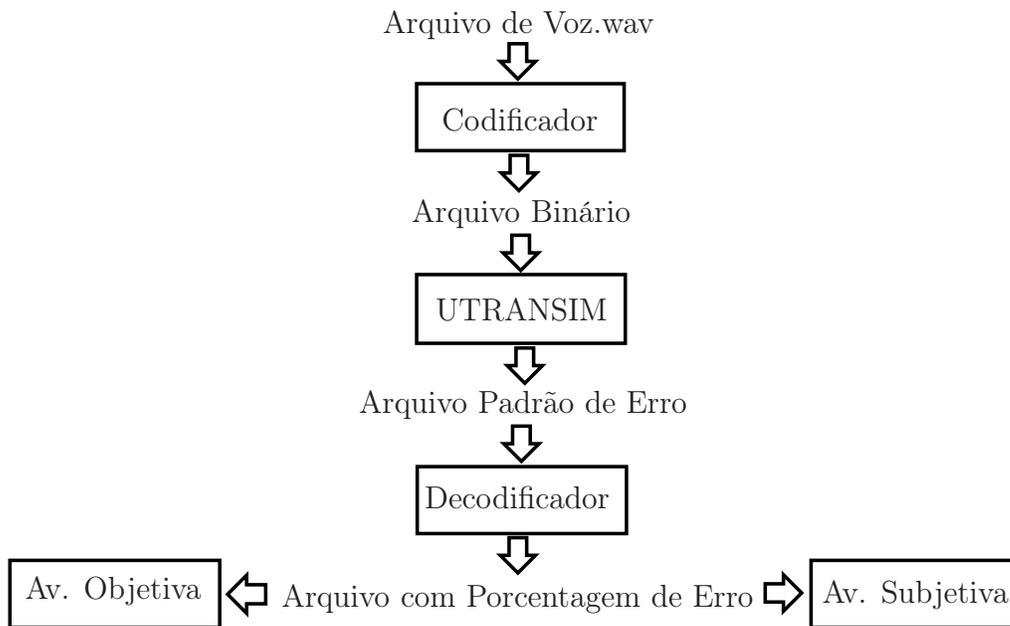


Figura 4.6: Módulos da Estrutura de Avaliação.

arquivo para indicar onde houve a perda de um trecho de voz. No arquivo padrão de erro colocam-se também os trechos correspondentes aos períodos *OFF*, contendo os trechos correspondentes à atividade, inatividade e perdas, para a decodificação do arquivo original.

Antes da transmissão pelo UTRANSIM, porém, fez-se com que cada um dos quatro arquivos de voz passasse por um ciclo de cerca de 2500 simulações de transmissão, obtendo-se, portanto, a mesma quantidade de arquivos padrões de erros relativos a cada arquivo de voz.

Ao final da simulação, obteve-se também uma lista com o nome de todos os arquivos padrões de erros gerados. Através desta lista, os arquivos foram separados para o processo de decodificação relativo a cada um dos quatro arquivos de voz. Na decodificação, utilizou-se um algoritmo onde cada arquivo padrão de erro é aplicado ao seu respectivo arquivo de voz original, gerando arquivos de voz com certa porcentagem de erro, que pode variar de 0 a 100%.

Até este ponto, o processo de codificação e simulação é o mesmo para as avaliações objetiva e subjetiva. Na avaliação objetiva, porém, depois da geração dos arquivos com porcentagens de erro o processo diferencia-se do subjetivo pois, de posse desses arquivos, o algoritmo PESQ é aplicado com o intuito de obter as notas MOS, que consistem nas notas de avaliação objetiva para cada um dos arquivos codificados, transmitidos e decodificados. Esta Dissertação de Mestrado não tem o

objetivo de avaliar nem analisar de forma profunda esse algoritmo PESQ de avaliação da qualidade de voz, pois ele foi obtido através de [61] e disponível pela ITU-T para utilização. O interesse no uso de PESQ, neste Trabalho, foi somente de obter sua nota final para cada arquivo a fim de compará-la com os resultados obtidos pela metodologia subjetiva. Vale lembrar, entretanto, que o algoritmo MOS (PESQ) opera comparando os sinais originais com os sinais decodificados, dando-lhes uma nota numérica indicadora da qualidade final, de acordo com a recomendação da ITU-T para este tipo de avaliação.

Na metodologia subjetiva, entre os arquivos de voz aos quais já haviam sido aplicados os padrões desejados de erros com certo valor de perda, escolheram-se novamente apenas quatro deles, cada um representando uma das vozes para escuta e avaliação. Nessa metodologia trabalha-se com um número reduzido de arquivos em virtude da impossibilidade de obter conceitos subjetivos para o grande número de arquivos de voz utilizados na avaliação objetiva. Ou seja, de cada voz (duas masculinas, duas femininas) foi escolhido um arquivo padrão de erro com uma porcentagem de 2% de FER, por ser esta a taxa considerada suficiente para a satisfação do usuário com boa qualidade do serviço de voz. Em seguida, adicionando-se o cabeçalho de cada um, foram digitalizados e decodificados. Na etapa seguinte, esses arquivos foram fragmentados e colocados em uma página da Internet para avaliação dos internautas, através da URL (do inglês *Uniform Resource Locator*) <http://www.gtcl.ufc.br/~ramon>, que pode ser visualizada através de figuras e descrições, no Apêndice A nesta Dissertação de Mestrado. Desta forma, ao final de 3 semanas de exposição dos arquivos de voz na Internet, colheram-se as notas de 54 avaliadores, cuja média produziu uma só nota MOS subjetiva para cada um dos quatro arquivos de voz selecionados.

Como etapa final, fez-se uma comparação das médias das notas subjetivas e objetivas dos mesmos arquivos de voz e com as mesmas taxas de FER, chegando-se a conclusões bem expressivas e que estão expostas na próxima seção.

4.3 Resultados de Desempenho

Os resultados obtidos através das campanhas de simulações com a utilização do UTRANSIM e através da disponibilização dos arquivos numa URL, já decodificados com padrões de erros, estão dispostos a seguir.

4.3.1 Resultados da Avaliação Objetiva

Nesta seção, os gráficos mostram os dados da parte objetiva e os percentuais de taxa de quadros com erros ou perdidos (FER) com relação à nota objetiva MOS, dada pelo algoritmo PESQ, aqui referido como MOS (PESQ).

Os dois arquivos de voz de homem e os dois de vozes de mulher são avaliados em diferentes cenários, com tempos de chegada distintos na região celular simulada através do UTRANSIM, causando uma sobrecarga de usuários no caso de menor tempo de chegada. Como a taxa de FER para satisfação de um usuário com relação ao serviço de voz deve ser 2%, de acordo com [73] e [89], este Trabalho também se limitou a essa taxa percentual como referência de qualidade aceitável para voz.

Os arquivos serão, daqui em diante, referidos segundo a nomenclatura Homem 1, Homem 2, Mulher 1 e Mulher 2, para descrever os comportamentos adotados, com variação nos tempos de chegada, entre cada usuário que nasce na célula, nos valores 0.5, 1.0, 1.5 e 2.0 segundos, por exercerem influência mais direta na carga do sistema, sendo que quanto menor for o tempo de chegada entre os usuários, mais sobrecarregado ficará o sistema e vice-versa.

As Figuras 4.7, 4.8, 4.9 e 4.10 revelam todos os percentuais de taxa de perda de quadros (FER), em todas as 2500 simulações, em cada tempo de chegada entre os usuários, para cada um dos arquivos de voz transmitidos pelo simulador. Percebe-se, ainda, que a maioria dos pontos se concentram em valores menores que 10% de FER, além de um comportamento dos percentuais relacionados mostrando que com o aumento da taxa de FER, há uma diminuição no valor do MOS (PESQ), atingindo valores entre 4 e 4.5 de MOS (PESQ) para as taxas de FER menores e valores próximos a 1 de MOS (PESQ) para os maiores percentuais de quadros perdidos. É preciso ressaltar também que quanto maior o valor de MOS (PESQ), melhor é a qualidade da voz, ou seja, existe uma menor quantidade de quadros perdidos no arquivo de voz que foi transmitido e aplicado o algoritmo PESQ.

A Figura 4.11 mostra a distribuição das taxas percentuais de FER em forma de histogramas do arquivo Homem 1 para cada tempo entre chegadas de chamadas, como pode-se perceber nos tempos 0.5 segundos, 1.0 segundo, 1.5 segundos e 2.0 segundos através das partes (a), (b), (c) e (d), respectivamente. Observa-se, entretanto, que há presença de taxas de FER bastante próximas nos menores tempos entre chegadas de chamadas, revelando a frequência de ocorrência de erros em cada tempo, mostrando-se intensa no tempo 0.5 segundos seguida de valores de FER mais alternados à medida que aumenta-se o tempo até chegar ao 2.0 segundos. Este

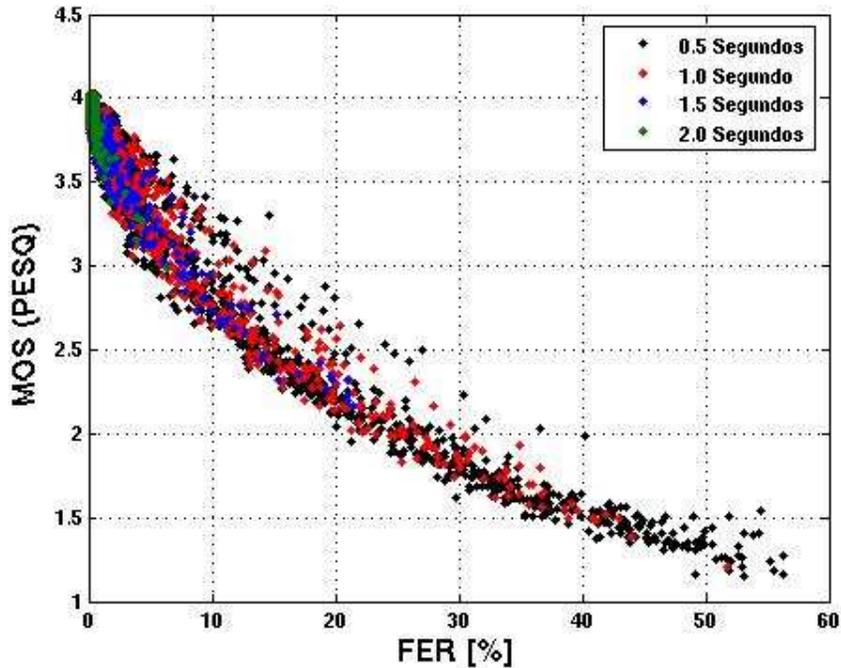


Figura 4.7: Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Homem 1, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.

comportamento também é similar nos arquivos Homem 2, Mulher 1 e Mulher 2.

Outra informação importante a ser ressaltada, também, é que para cada tempo de chegada, podem existir diferentes notas MOS (PESQ) para arquivos com um mesmo percentual de FER. Isto se deve ao fato de que o algoritmo pode ser influenciado por trechos seguidos com muitos quadros perdidos de um arquivo, dando diferentes notas de MOS (PESQ) para arquivos com a mesma FER. Sendo assim, há valores máximos, medianos e mínimos de MOS (PESQ) para uma mesma taxa de FER; unindo-se todos esses pontos máximos da distribuição obtém-se um gráfico que revela o comportamento dos valores máximos, e assim também se podem produzir os gráficos de valores médios e mínimos.

Separando cada um dos quatro arquivos em figuras distintas, obtém-se gráficos com as distribuições de perdas, valores mínimos, médios e máximos para todos os tempos de chegada, como mostram as Figuras 4.12, 4.13, 4.14 e 4.15. Em todas essas figuras, que mostram uma análise completa da distribuição de perdas para cada arquivo nos quatro diferentes tempos de chegada, percebe-se que não houve taxas de FER maiores que 60%. As legendas dessas representações gráficas indicam, para cada tempo de chegada, os percentuais de taxa de FER, chamadas de Pontos, os

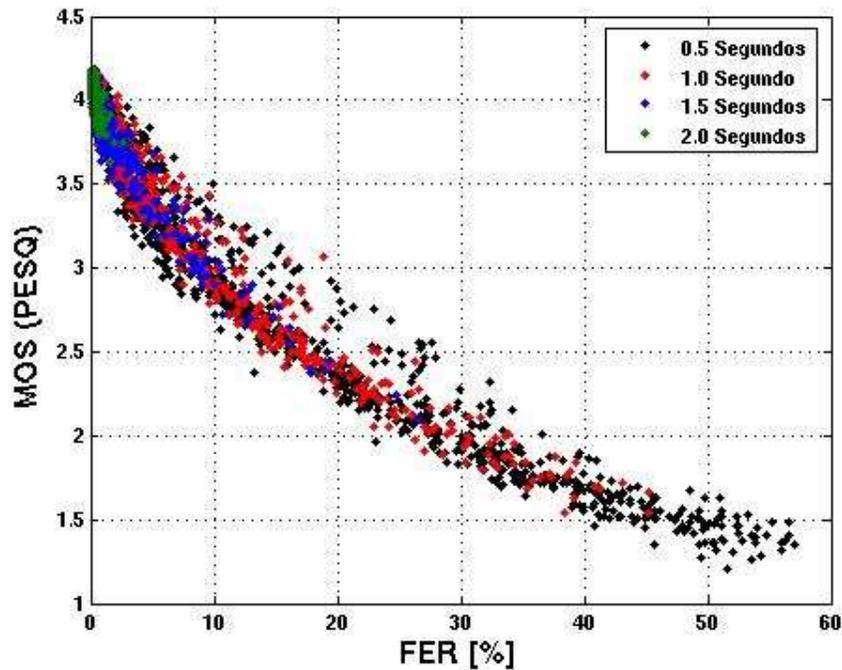


Figura 4.8: Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Homem 2, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.

gráficos dos pontos medianos, máximos e mínimos como Média, Máximo e Mínimo, respectivamente.

Por existirem pontos máximos, médios e mínimos nas distribuições, neste Trabalho, os gráficos dos valores medianos servirão como referencial comparativo para as notas médias da avaliação subjetiva, que será descrita na próxima Subseção. Assim, se forem selecionados somente os gráficos dos valores médios para os quatro diferentes tempos de chegada, têm-se a Figura 4.16, que mostra comportamentos similares dos gráficos, destacando-se apenas as cinco primeiras porcentagens de FER, já que a taxa a ser analisada para uma qualidade de voz e sua correspondente nota MOS (PESQ) foi de 2%.

Verifica-se, ainda, na Figura 4.16, que quase todos os gráficos dos valores médios para o tempo de chegada de 2.0 segundos não conseguem alcançar o valor de 5% de FER, pelo fato da maioria dos percentuais concentrarem-se abaixo desse número somente para este tempo de chegada entre os usuários. Observa-se, portanto, que para esse maior tempo de chegada entre os usuários, não se registram altos percentuais de taxas para quadros perdidos, pelo fato de que uma menor carga na região celular permite que todos os usuários sejam servidos e transmitam seus

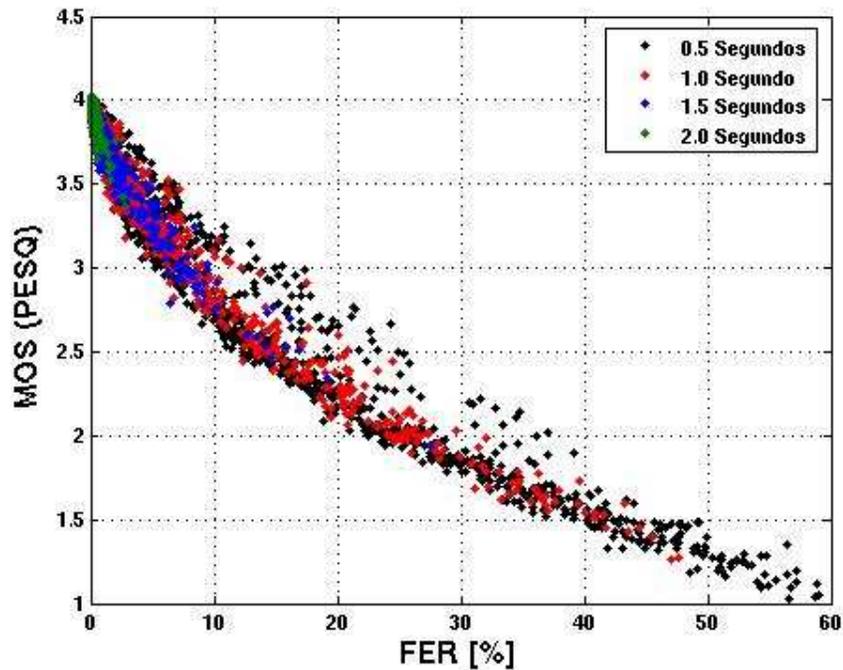


Figura 4.9: Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Mulher 1, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.

arquivos de voz em um tempo mais flexível, sem competir entre si por recursos da rede. Diminui-se assim a quantidade de perdas de quadros que puxaria para baixo sua nota MOS (PESQ). Outro fato importante a ser observado é que os gráficos dos diferentes tempos de chegada mostram notas MOS (PESQ) para 2% de FER muito próximas umas das outras, comportamento que se repete, na maioria das vezes, também para os demais percentuais.

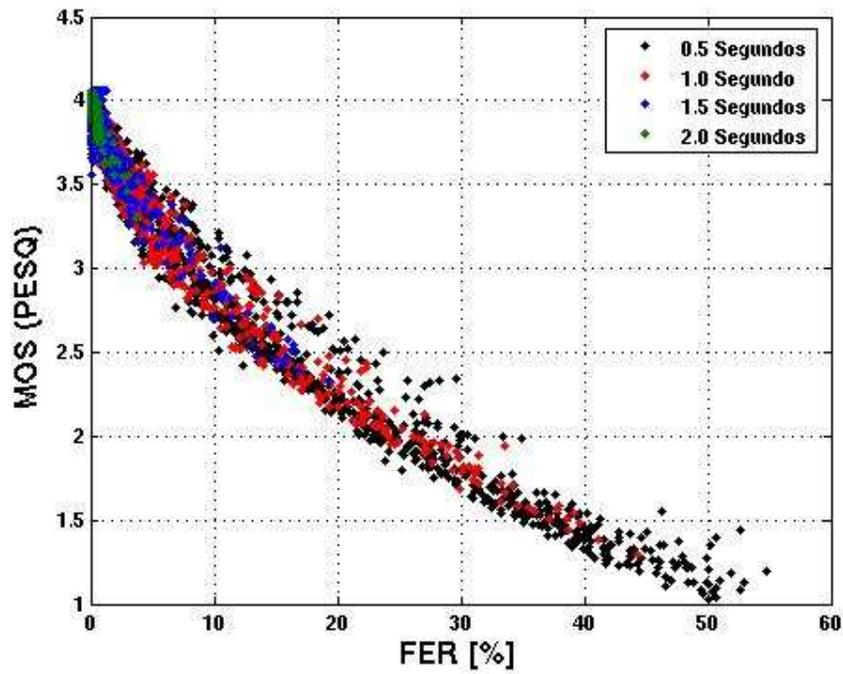
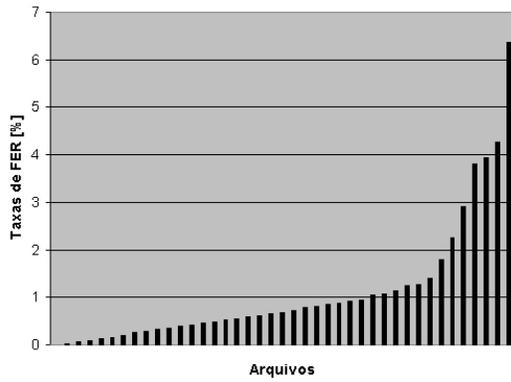
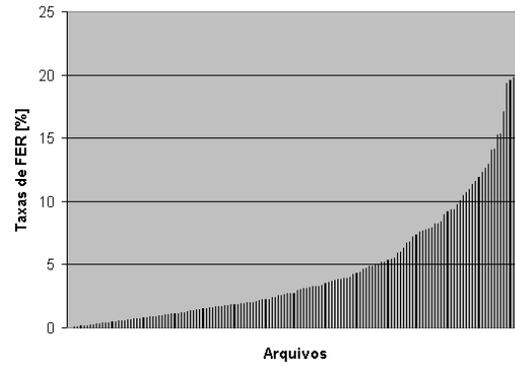


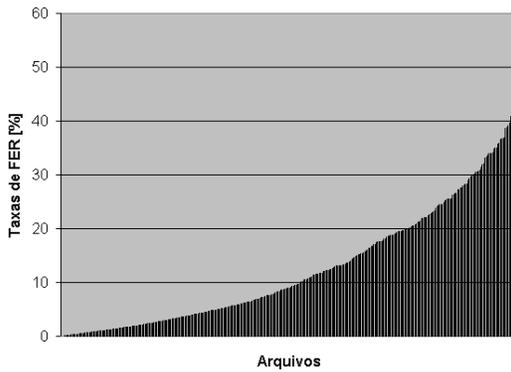
Figura 4.10: Distribuições dos percentuais das taxas de quadros perdidos variando o tempo de chegada entre os usuários para o arquivo Mulher 2, sendo que foram feitas aproximadamente 2500 simulações para este arquivo.



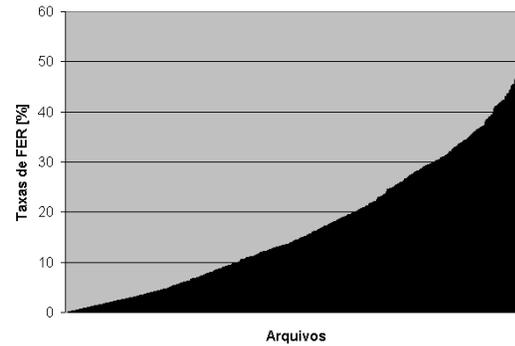
(a) Tempo de Chegada = 2.0 segundos



(b) Tempo de Chegada = 1.5 segundos



(c) Tempo de Chegada = 1.0 segundo



(d) Tempo de Chegada = 0.5 segundos

Figura 4.11: Gráfico das taxas de FER para o Arquivo Homem 1 em cada tempo entre chegadas de chamadas, revelando-se a frequência de erros em cada tempo.

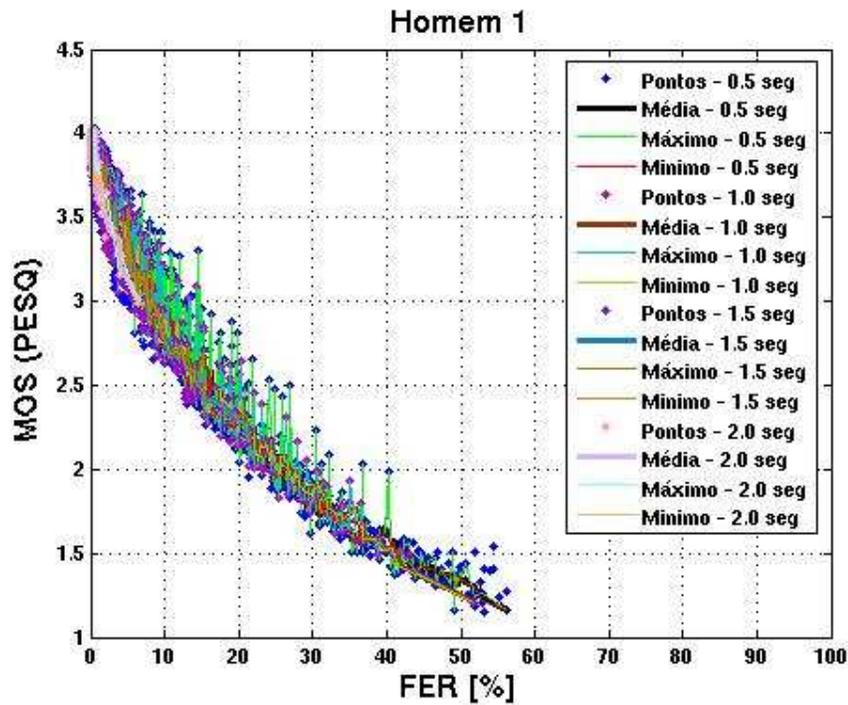


Figura 4.12: Gráfico do arquivo Homem 1 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos, cuja legenda mostra que Pontos significam as taxas de quadros perdidos (FER), que Média significa o gráfico que une os pontos médios de nota MOS para cada valor de FER, que Máximo trata dos gráficos para os valores máximos e que Mínimo nomeia o gráfico dos valores mínimos. Todos eles com seus respectivos tempos de chegada.

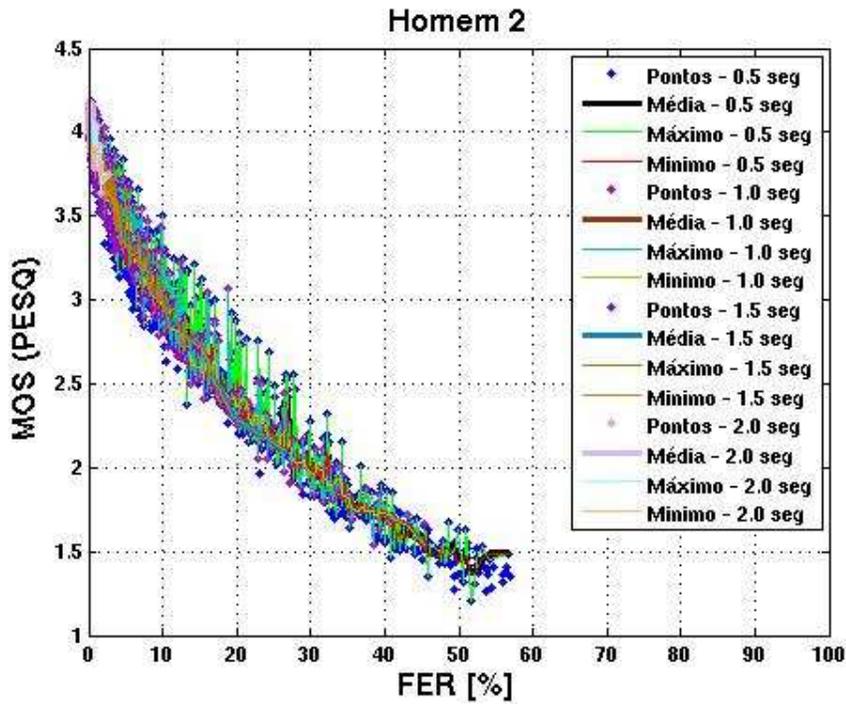


Figura 4.13: Gráfico do arquivo Homem 2 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.

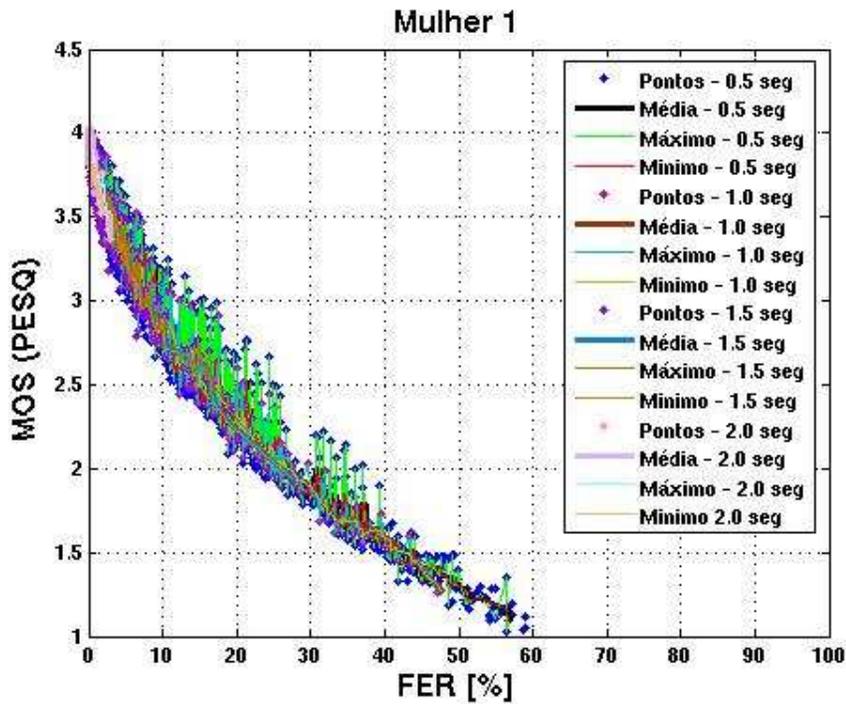


Figura 4.14: Gráfico do arquivo Mulher 1 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.

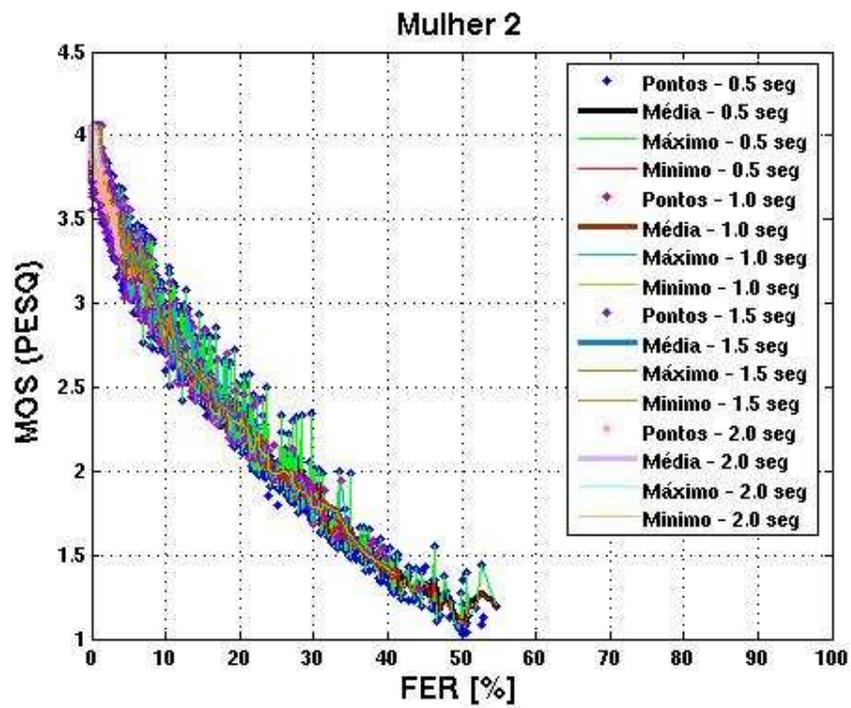
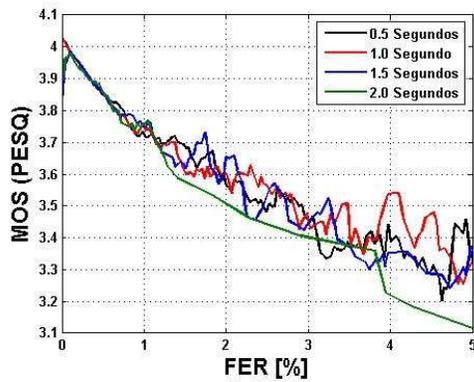
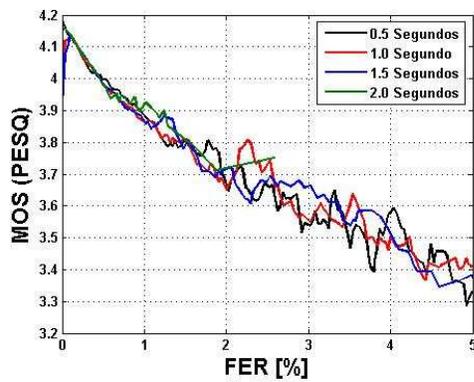


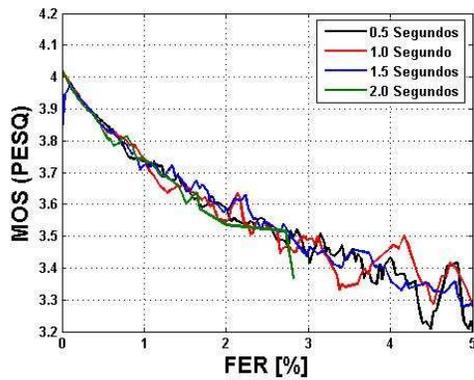
Figura 4.15: Gráfico do arquivo Mulher 2 para 2500 simulações com os tempos de chegada entre usuários de 0.5, 1.0, 1.5 e 2.0 segundos.



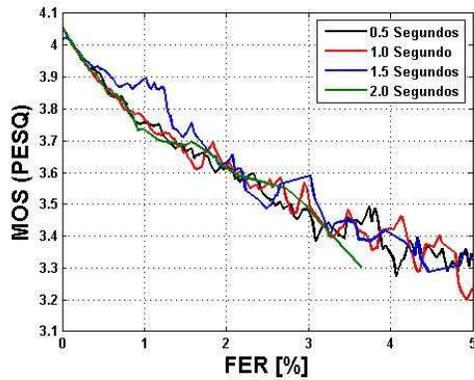
(a) Homem 1



(b) Homem 2



(c) Mulher 1



(d) Mulher 2

Figura 4.16: Arquivos das Vozes Avaliadas expressos variando apenas o tempo de chegada entre os usuários.

Analisando os percentuais médios dados pelos gráficos dos valores médios, para cada tempo de chegada entre os usuários, apresentados na Figura 4.16, e calculado-se uma média entre eles, têm-se o gráfico comportamental relativo aos quatro tempos de chegada para cada arquivo de voz, expresso através da Figura 4.17. Esse gráfico geral das médias trata de evidenciar a correspondência entre a nota MOS (PESQ) e a respectiva FER.

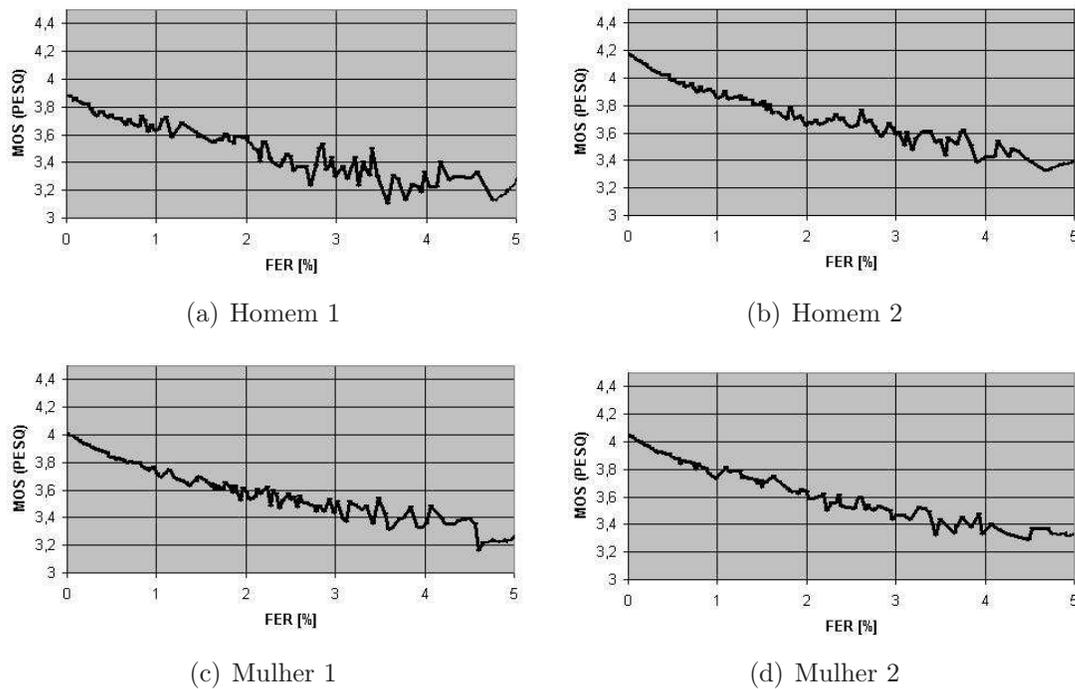


Figura 4.17: Arquivos de voz representando a média geral a partir da Figura 4.16.

A partir dos gráficos das médias dadas para cada um dos arquivos de voz, na Figura 4.17, encontram-se os valores de nota MOS (PESQ) para 2% de FER que estão destacados na Tabela 4.2, podendo-se perceber que os arquivos Homem 2 e Mulher 1 obtiveram as maiores notas objetivas para este valor de FER.

Tabela 4.2: Valores MOS(PESQ) para 2% de FER das Médias de cada Arquivo de Voz na Avaliação Objetiva, onde percebe-se que todos os resultados indicam uma percepção subjetiva equivalente entre "Aceitável" e "Bom".

Arquivo	MOS(PESQ)
Homem 1	3.5581
Homem 2	3.6761
Mulher 1	3.5883
Mulher 2	3.6392

4.3.2 Resultados da Avaliação Subjetiva

Na avaliação feita com uma metodologia subjetiva, foram usados os mesmos quatro arquivos de voz decodificados: Homem 1, Homem 2, Mulher 1 e Mulher 2 com o quais se fez a avaliação objetiva. Para o procedimento subjetivo, cada um deles foi separado em três segmentos de, em média, 20 segundos, agora com extensão .wav, formato capaz de ser ouvido através de qualquer aplicativo que suporte mídia sonora. A Figura 4.18 representa os três segmentos digitalizados do arquivo Homem 1. Essa separação em segmentos, que foi também utilizada ao tornarem-se os arquivos disponíveis na URL, teve como finalidade evitar a má influência que poderia exercer um pequeno trecho, eventualmente muito ruim, no momento da atribuição da nota final à qualidade de recepção de todo o arquivo de voz. Em seguida, mostram-se os segmentos dos arquivos Homem 2, Mulher 1, Mulher 2, respectivamente nas Figuras 4.19, 4.20 e 4.21.

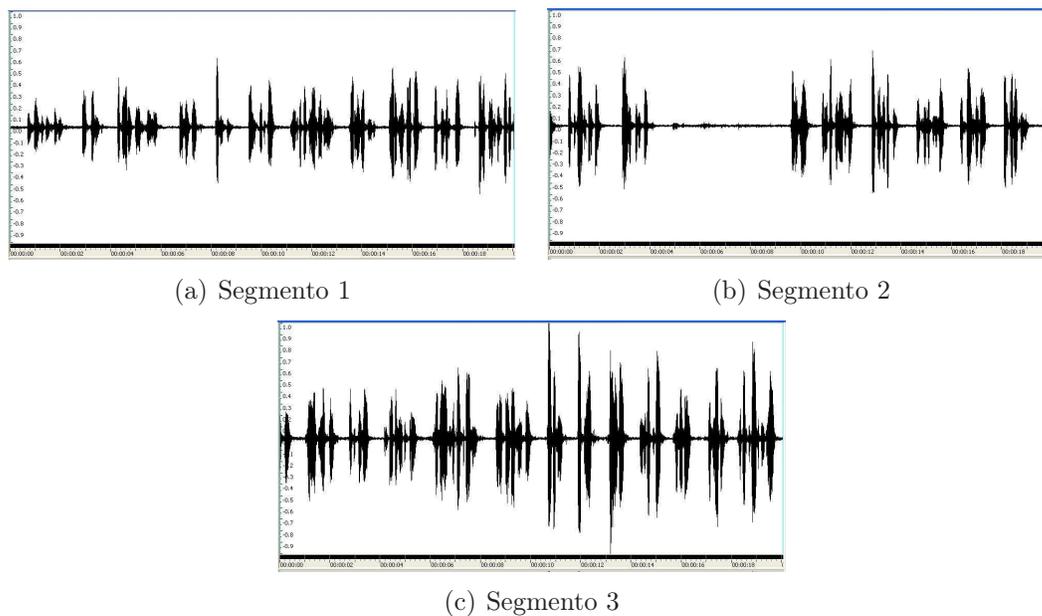


Figura 4.18: Sinais Digitalizados dos Segmentos do Arquivo Homem 1 no formato de extensão *wav*, que é capaz de ser escutado por qualquer software de voz.

Como já se disse, os arquivos de voz sob exame nesta pesquisa foram submetidos à taxa de 2% para a FER - por ser este o valor de satisfação para uma boa qualidade de voz de um usuário do serviço VoIP - tornando-se então comparáveis as notas MOS obtidas por arquivos de voz com igual taxa de FER através da metodologia subjetiva.

Na avaliação subjetiva foram coletadas, ao todo, 162 notas, oriundas de 54 usuários, tendo, cada um deles, atribuído uma nota a cada um dos três segmentos

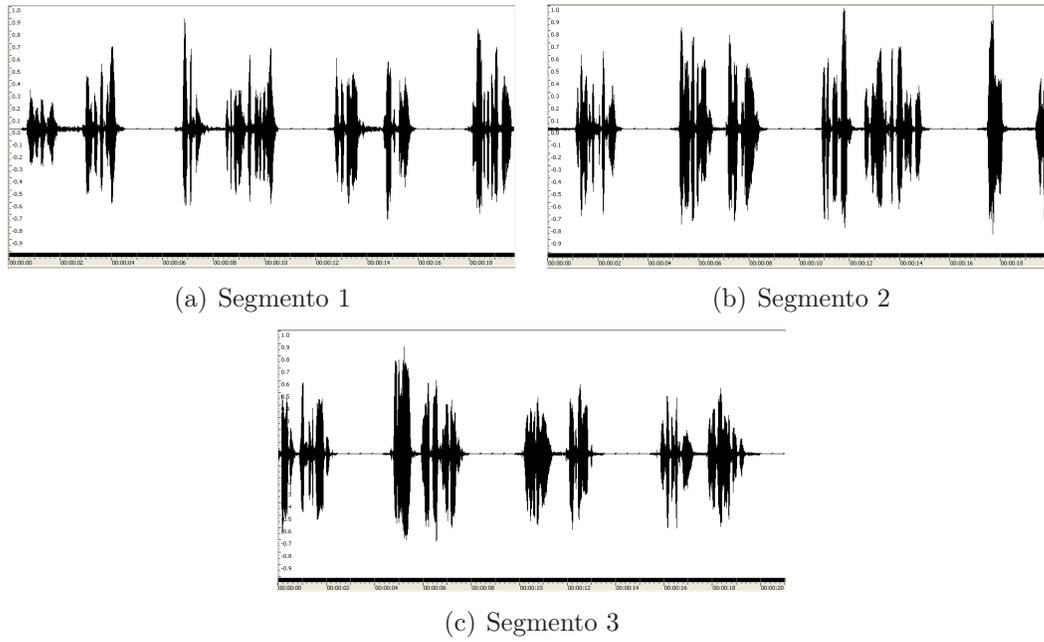


Figura 4.19: Segmentos do Arquivo Homem 2.

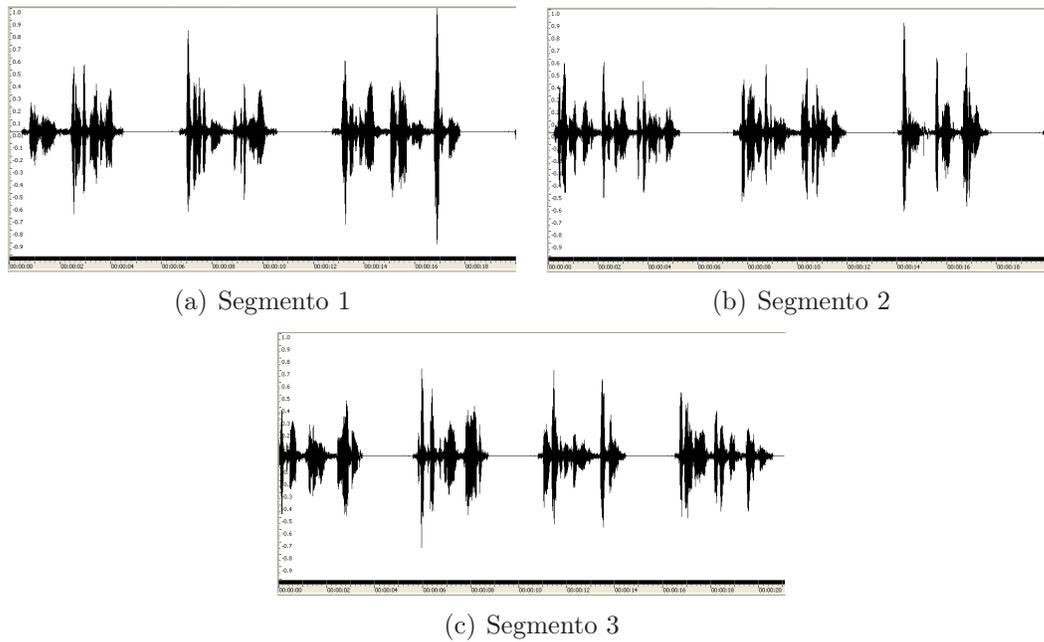


Figura 4.20: Segmentos do Arquivo Mulher 1.

por arquivo.

Neste Trabalho, quanto à avaliação subjetiva, tomou-se como referência inicial o artigo científico [90], que trabalhou com a opinião de apenas 24 usuários e o estudo [6], que contou com a participação de 28 pessoas para a avaliação de voz.

As notas finais obtidas por esse método podem ser visualizadas, na Figura 4.22,

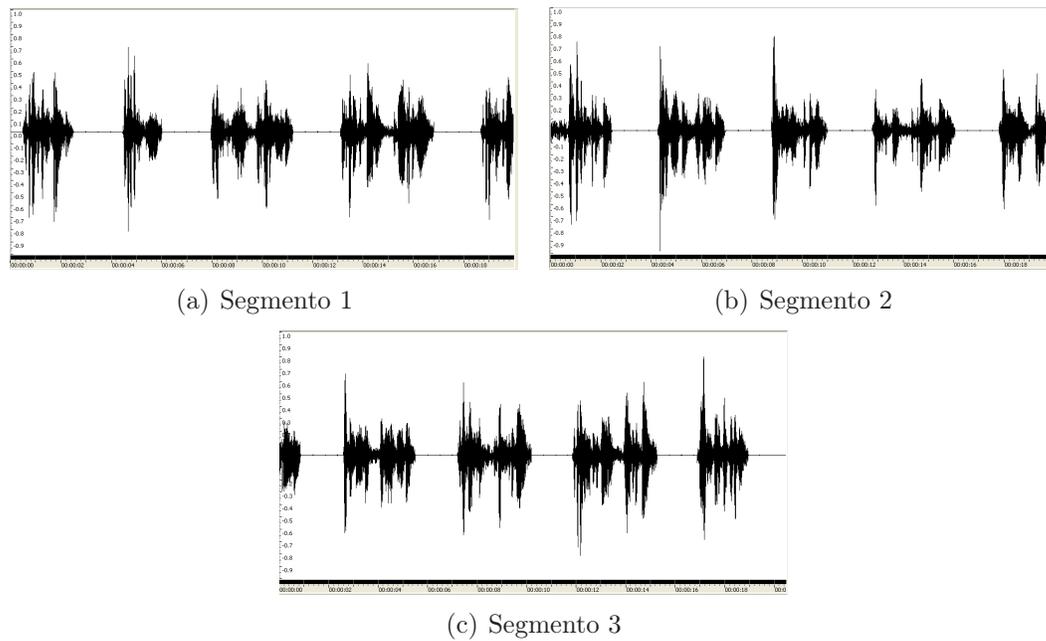


Figura 4.21: Segmentos do Arquivo Mulher 2.

de forma mais precisa através de sua distribuição nos histogramas gerados pelo aplicativo Microsoft Excel, a partir de todas as notas atribuídas pelos usuários avaliadores. A nota final foi calculada como a média das notas e tomada como referência para o método subjetivo nesta Dissertação. Como a divisão dos arquivos de voz em três segmentos, permitia que o ouvinte desse notas influenciadas por um trecho bom ou ruim de algum dos segmento, ao fazer-se uma média, essa influência se ameniza para um valor aceitável, como é definido em [7,91].

A Figura 4.22 mostra também uma grande quantidade de notas MOS máximas, como pode-se perceber nos histogramas de Homem 1, Homem 2, Mulher 1 e Mulher 2 através das partes (a), (b), (c) e (d), respectivamente. Observa-se, entretanto, que há presença de todos os valores possíveis de nota MOS, desde a mínima até a máxima, devido às diferentes interpretações dos usuários para cada um dos arquivos avaliados em seus respectivos segmentos. Percebe-se também que no arquivo Homem 1 houve quase a mesma quantidade de notas 4 e 5 para MOS, correspondentes a boa e ótima qualidade de entendimento, enquanto no arquivo Homem 2 a nota máxima (5, ótima) aparece em quantidade bem mais significativa do que todas as outras. Quanto aos arquivos Mulher 1 e Mulher 2, o comportamento na distribuição das notas foi similar, predominando o valor 4, que significa uma boa qualidade para a inteligibilidade de voz.

As notas médias de MOS da avaliação subjetiva para 2% de FER são

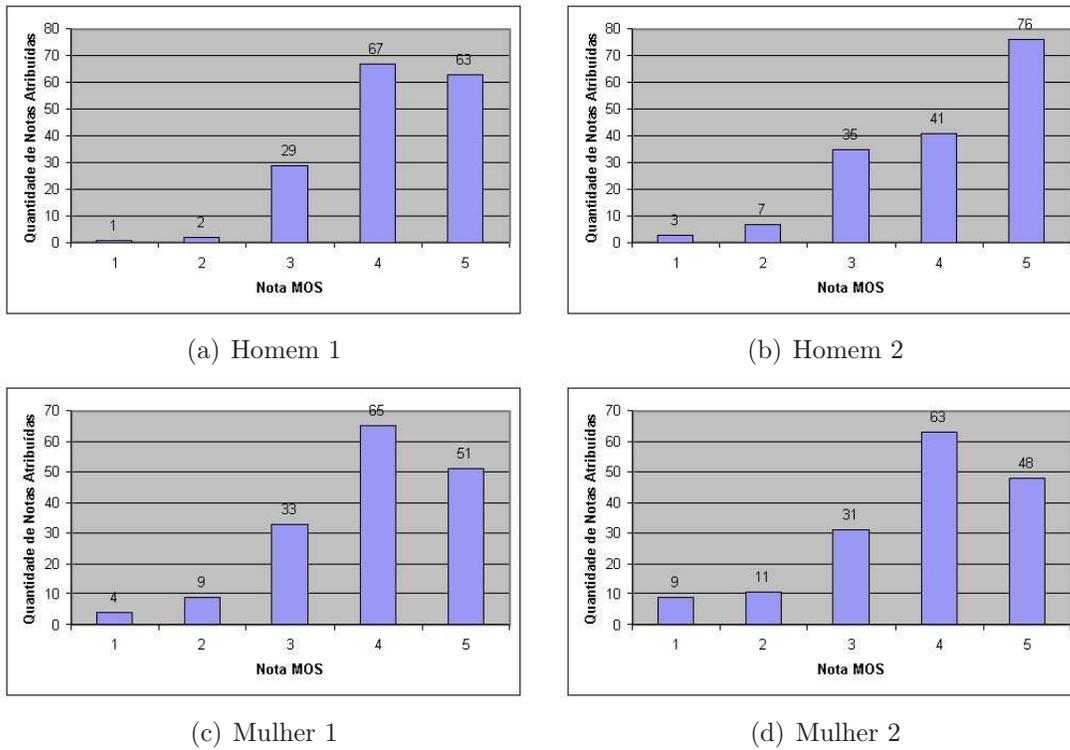


Figura 4.22: Histogramas que mostram as Distribuições das Notas Atribuídas a cada Segmento de Arquivo pela Avaliação Subjetiva.

mostradas na Tabela 4.3, de forma numérica e detalhada, podendo-se notar que as vozes de homens obtiveram as melhores notas MOS para a avaliação subjetiva. Provavelmente contribuiu para isso o fato de as vozes femininas terem sofrido mais com um fenômeno conhecido como reverberação, que é ocasionado pela reflexão do som, e mesmo múltiplas reflexões, que prolongam a sensação auditiva, o que as torna menos claras para os usuários, levando-os a atribuir-lhes notas menores e mais variadas [6, 8].

Outro dado verificado é que o arquivo Homem 1 recebeu as melhores notas MOS, apesar de incluir um período seguido de inatividade maior que 5 segundos, que pode ser visualizado pelo segmento 2 da Figura 4.18, já mostrada.

A partir da média das notas subjetivas, foram calculados um valor máximo e um valor mínimo com o intuito de determinar uma margem de segurança da nota, como está descrito na Figura 4.23. A forma utilizada para encontrar os valores mínimo e máximo partiu de uma Distribuição Normal de probabilidade com intervalo de confiança de 95% e a validade de sua utilização teve como referência também o artigo científico [90], que também trabalha com valores mínimos e máximos originados a partir de uma média de avaliações subjetivas de arquivos de voz.

Tabela 4.3: Valores MOS das Médias de cada Arquivo de Voz na Avaliação Subjetiva para 2% de FER.

Arquivo	MOS
Homem 1	4.1666
Homem 2	4.1111
Mulher 1	3.9259
Mulher 2	3.8024

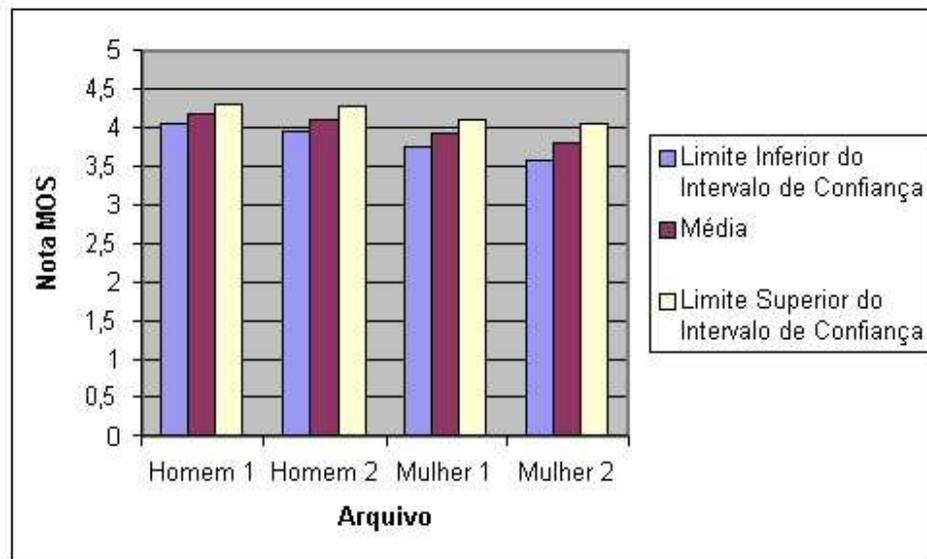


Figura 4.23: Média das Notas Subjetivas para um Intervalo de Confiança de 95% seguida dos Limites Superior e Inferior [90].

Ainda sobre a Figura 4.23, observa-se, de maneira geral, que a avaliação subjetiva para os quatro arquivos de voz disponíveis na URL criada para esta pesquisa, de acordo com o Apêndice A e com a tabela de conceito MOS, gerou conceitos acima de aceitável para a inteligibilidade dos arquivos avaliados, mesmo calculando-se valores máximos e mínimos para um intervalo de confiança.

4.3.3 Comparação dos Resultados entre os Métodos Objetivo e Subjetivo

Comparando-se os valores de nota MOS no método objetivo, mostrados na Tabela 4.2 da Seção 4.3.1, com os valores obtidos pelo método subjetivo, contidos na Tabela 4.3 da Seção 4.3.2, criou-se a Tabela 4.4 e a Figura 4.24, representando todos os arquivos e suas respectivas notas médias de MOS (PESQ) e MOS para uma taxa de 2% de FER.

Percebe-se que os dois métodos de avaliação obtiveram conceitos de qualidade

satisfatórios para uma taxa de quadros perdidos de 2%, podendo-se concluir que essa taxa garante uma boa qualidade de audição de arquivos de voz, suficiente para a satisfação do usuário do serviço VoIP sobre o sistema celular HSDPA.

Tabela 4.4: Valores MOS(PESQ) e MOS das Médias de cada Arquivo de Voz nas Avaliações Objetiva e Subjetiva seguido da diferença em termos percentuais de nota MOS entre o valor subjetivo e o objetivo para cada Arquivo.

Arquivo	MOS(PESQ)	MOS	Diferença (%)
Homem 1	3.5581	4.1666	17 %
Homem 2	3.6761	4.1111	11 %
Mulher 1	3.5883	3.9259	09 %
Mulher 2	3.6392	3.8024	05 %

Outra conclusão importante, proveniente da comparação das duas avaliações, é a de que o método objetivo, através da aplicação do algoritmo PESQ, garante a obtenção notas MOS bastante aproximadas, para menos, da MOS obtida pelo método subjetivo, podendo evitar, para o controle da qualidade das transmissões de voz, o árduo trabalho de fazerem-se ouvir diversos arquivos de voz, com diferentes taxas de FER, por uma quantidade significativa de usuários para que lhes atribuam notas subjetivas.

Os dados mostrados na Tabela 4.4 com as diferenças percentuais entre as duas metodologias também foram detectados na literatura vigente, nos quais compararam o MOS objetivo e o subjetivo, encontrando-se valores similares ao deste Trabalho.

Em [92], mostrou-se que o PESQ tem um coeficiente de correlação mais exato com o MOS subjetivo que em comparação com outros métodos objetivos. Revela-se na parte análise de resultados e testes de [9], um gráfico comparativo entre os métodos de teste objetivo e subjetivo. Para a taxa de 2% de perda de pacotes, percebe-se que o subjetivo tem melhor conceito MOS que o objetivo PESQ. Outra abordagem semelhante é feita em [93], cujos resultados mostraram que medidas objetivas são capazes de prever as subjetivas.

Os resultados dispostos nesta Dissertação são, portanto, esperados e explicáveis, já que foram anteriormente encontrados em trabalhos comparativos entre as avaliações objetiva e subjetiva.

Vale observar também, a partir da Tabela 4.4 e da Figura 4.24, comparativas das duas metodologias, que os arquivos masculinos receberam melhores notas MOS na

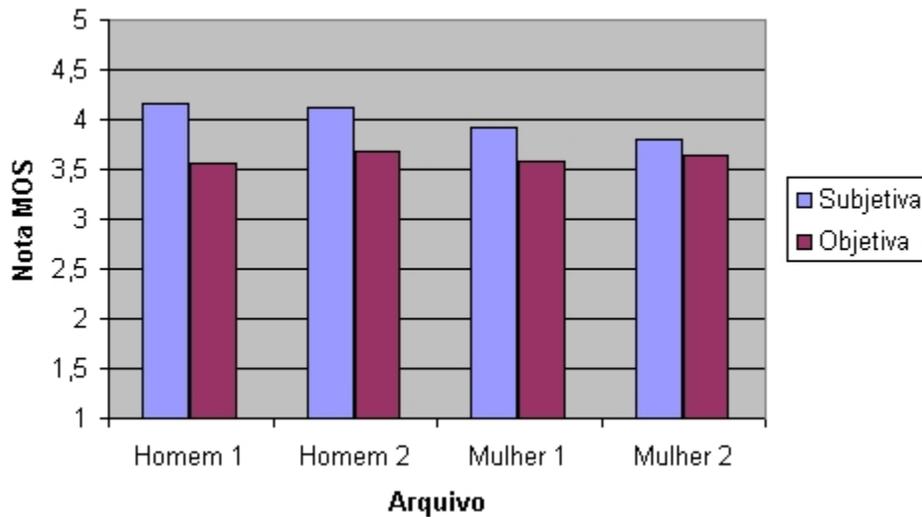


Figura 4.24: Histograma da Média das Notas MOS dadas pelas Avaliações Subjetiva e Objetiva.

avaliação subjetiva, enquanto que, na avaliação objetiva, tanto o arquivo Homem 2 quanto o arquivo Mulher 2 obtiveram os melhores valores.

Assim, os dois métodos chegam a valores aproximados para a nota MOS de um mesmo arquivo para o valor de 2% de FER e com conceitos satisfatórios de inteligibilidade para a qualidade de voz no serviço VoIP sobre o HSDPA.

Em outras palavras, conclui-se que o método objetivo, cuja aplicação demanda menor investimento, pode, por ele mesmo, proporcionar uma avaliação suficiente da qualidade da recepção de voz pelos usuários, sem necessidade de utilização concomitante do método subjetivo cujos resultados seriam reiterativos.

4.4 Sumário

Este capítulo descreveu o processo de avaliação da qualidade de voz do serviço VoIP em sistemas HSDPA, realizado na pesquisa que baseia esta Dissertação, desde as considerações sobre a oferta do serviço VoIP em HSDPA, passando pela modelagem do sistema para simulação com os parâmetros definidos, até chegar-se aos resultados obtidos através dos métodos utilizados para a avaliação. Com isso, pode-se chegar a uma série de conclusões que serão descritas no próximo capítulo.

Conclusões e Perspectivas

5.1 Conclusões

A maior contribuição deste Trabalho consiste na análise da qualidade de voz do serviço de VoIP em sistemas HSDPA, comparando os resultados obtidos pelas metodologias objetiva e subjetiva, a partir da relação que se pôde estabelecer entre a métrica de taxa de quadros perdidos (FER), fornecida por um simulador e as notas de qualidade obtidas com as duas diferentes metodologias. Para os fins desta pesquisa, criou-se, também, um simulador de provimento do serviço VoIP no sistema HSDPA que poderá ser usado em outros experimentos.

O processo de comparação dos dois métodos seguiu um só trajeto até a geração de arquivos padrões de erro, registrando os locais onde ocorriam perdas nos arquivos de voz originais. A partir desse ponto, bifurcaram-se os caminhos, obedecendo a dois métodos diferentes para a obtenção de conceitos avaliativos da qualidade da voz decodificada pelo receptor. Para a avaliação pelo método objetivo, utilizou-se o algoritmo chamado PESQ, recomendado pela ITU-T no documento P.800. Para a avaliação pelo método subjetivo, os arquivos de voz foram publicados em uma URL para que internautas pudessem, espontaneamente, avaliá-los e atribuir notas MOS para a qualidade da voz, de acordo com as recomendações de avaliação de MOS subjetivo já estabelecidas na literatura pertinente. O conjunto do processo favoreceu algumas observações a respeito de ambas as metodologias que interessam ao escopo desta pesquisa e que serão detalhadas a seguir.

Observações relativas ao Método Objetivo:

- ▶ Não houveram taxas de FER que chegassem a 100% para os tempos de chegada usados, que foram escolhidos por serem tempos críticos;

- ▶ A maioria dos percentuais de FER obtidos nas simulações concentraram-se abaixo dos 2%;
- ▶ Os gráficos dos diferentes tempos de chegada obtiveram notas MOS (PESQ) para 2% de FER muito próximas umas das outras;
- ▶ Mantendo-se até 2% de FER para cada arquivo, o comportamento dos gráficos dos diferentes tempos de chegada entre os usuários é muito semelhante;
- ▶ Os arquivos Homem 2 e Mulher 2 obtiveram melhores notas objetivas.

Observações relativas ao Método Subjetivo:

- ▶ No resultado final, apareceram todos os valores possíveis de Nota MOS, desde a nota mínima até a máxima, devido às diferentes interpretações dadas pelos usuários avaliadores, como sempre se pode esperar de qualquer tipo de avaliação subjetiva;
- ▶ Os arquivos de vozes masculinas (Homem 1 e Homem 2) obtiveram as melhores notas de MOS pelo método subjetivo, ao contrário do que constatou a Dissertação de Mestrado de [8], na qual arquivos de vozes femininas obtiveram notas melhores do que as masculinas;

Finalmente, podemos assim resumir as conclusões gerais a que chegou esta pesquisa:

- ▶ A taxa de 2% de FER garante uma boa qualidade para os arquivos de voz e a provável satisfação dos usuários do serviço VoIP sobre o sistema celular HSDPA;
- ▶ O método objetivo produz notas MOS (PESQ) aproximadas para menos das notas MOS obtidas pela metodologia subjetiva, podendo, portanto, substituí-la e evitar um dispendioso trabalho de recolhimento de notas subjetivas de uma quantidade de usuários suficientemente grande para ser estatisticamente válida;
- ▶ Os dois métodos chegam a valores aproximados para a nota MOS de um mesmo arquivo, para o valor de 2% de FER, com conceitos satisfatórios de inteligibilidade da voz transmitida pelo serviço VoIP sobre o HSDPA;

- ▶ O simulador UTRANSIM, ao prover o serviço de VoIP, fornece uma QoS muito positiva e estável, verificada pelo fato de que a maioria dos arquivos processados por ele obtiveram taxas de FER menores que 2%.

Tais conclusões permitem afirmar que tanto o objetivo geral quanto os objetivos específicos e as metas intermediárias que se propunha o autor deste Trabalho, conforme anunciados na Seção 1.3 do Capítulo 1 desta Dissertação, foram plenamente alcançadas.

5.2 Perspectivas

O Trabalho até aqui realizado sugere várias perspectivas de continuidade e aprofundamento dos conhecimentos sobre os métodos de avaliação da qualidade e transmissão de voz pelo serviço VoIP sobre HSDPA. A busca pelo mais confiável e econômico método para esse tipo de avaliação pode prosseguir experimentando-se com a utilização e análise dos efeitos de outras variáveis:

- ▶ Diferentes algoritmos de escalonamento;
- ▶ Diferentes requerimentos de atraso, como o máximo limiar permitido de 150 ms;
- ▶ Cenários mistos, com a presença de serviços *Web* concorrendo com o serviço de VoIP;
- ▶ Outros modelos que retratam codificadores como, por exemplo, o AMR-WB e o iLBC;
- ▶ Outros algoritmos de avaliação objetiva;
- ▶ Uma combinação de outros métodos subjetivos;
- ▶ Um acúmulo de pacotes de voz no RTP do transmissor;
- ▶ Outras taxas de atividade de voz;
- ▶ O mapeamento entre a nota MOS e o atraso;
- ▶ O mapeamento entre a nota MOS e a variação do atraso também;
- ▶ Validação da relação MOS(PESQ) e MOS(Subjetivo) para valores mais altos de FER;

- ▶ Verificação e validação da correspondência entre MOS(PESQ) e MOS(Subjetivo) para taxas maiores de FER;
- ▶ Segmentação dos arquivos de voz avaliados na metodologia objetiva e o uso do algoritmo PESQ em cada segmento com o objetivo de encontrar a média de notas PESQ, assim como realizou-se no método subjetivo;
- ▶ Implementação de outros algoritmos de compressão de cabeçalho;
- ▶ Implementação do RTCP e o seu impacto na qualidade de voz.

Página na *Internet* utilizada para Avaliação Subjetiva

Este apêndice descreve e mostra a página como se apresenta a URL ou sítio eletrônico *http://www.gtél.ufc.br/~ramon*, no qual ficaram disponíveis para avaliação subjetiva os quatro arquivos de voz decodificados com as devidas perdas e divididos em segmentos espalhados de forma não seqüencial.

A primeira parte, correspondente à Figura A.1, contém a descrição do que é o conceito MOS, juntamente com uma tabela que relaciona as notas numéricas que devem ser dadas a cada conceito. Em seguida, na Figura A.2, estão dispostos os segmentos de voz com os campos de atribuição de nota, além de algumas observações que são feitas com o intuito de facilitar a escuta das vozes e a atribuição das notas pelo internauta avaliador.

Os arquivos de voz aqui utilizados para a escuta são de dois homens e de duas mulheres, gravados de acordo com os critérios descritos na Seção 4.2.4 do Capítulo 4 desta Dissertação de Mestrado.

Os quatro arquivos de voz gravados têm, cada um, 60 segundos de duração em média. Porém, cada um deles foi dividido em 3 segmentos, compondo-se um total de 12 segmentos de voz disponíveis para escuta e avaliação numa URL, onde foram dispostos de forma aleatória para não cansar o ouvinte. Essa divisão teve o objetivo de facilitar o processo de escuta e atribuição de nota caso houvesse algum trecho excepcionalmente bom ou ruim que pudesse influir fortemente na nota geral, como se explicou na Seção 4.3.2 do Capítulo 4 desta Dissertação.

A URL utilizada para a coleta de notas subjetivas foi formatada de modo que, ao final da escuta e do preenchimento dos campos para as notas de todos

os segmentos, cada usuário enviase os dados para um registro armazenador das notas, calculando-se então a média para cada arquivo e dispondo-se seus resultados, como se encontram na Seção 4.3.2 do Capítulo 4.

Avaliação MOS

Mean Opinion Score (MOS) é um método subjetivo usado para determinar a qualidade de um som percebido por um ouvinte, sendo especificado na Recomendação P.800 do ITU-T.

O ouvinte escuta um arquivo de voz e segundo sua percepção, fornece um conceito a cada sinal de voz em teste de acordo com a Tabela 1.

Tabela 1: Conceito MOS

Conceito	Nota	Descrição
Excelente	5	Sinal de voz com perfeito entendimento
Bom	4	Qualidade de uma chamada telefônica de longa distância
Aceitável	3	Requer algum esforço no entendimento
Insatisfatório	2	Fala de baixa qualidade e difícil de entender
Ruim	1	Fala não clara, muito quebrada, não dá pra entender

Assim, pedimos para que você escute cada um dos seguintes arquivos de voz e forneça uma respectiva nota. Cada um dos arquivos a serem avaliados consiste de um segmento de voz diferente.

OBSERVAÇÃO:

Figura A.1: Parte Superior da Página na Internet.

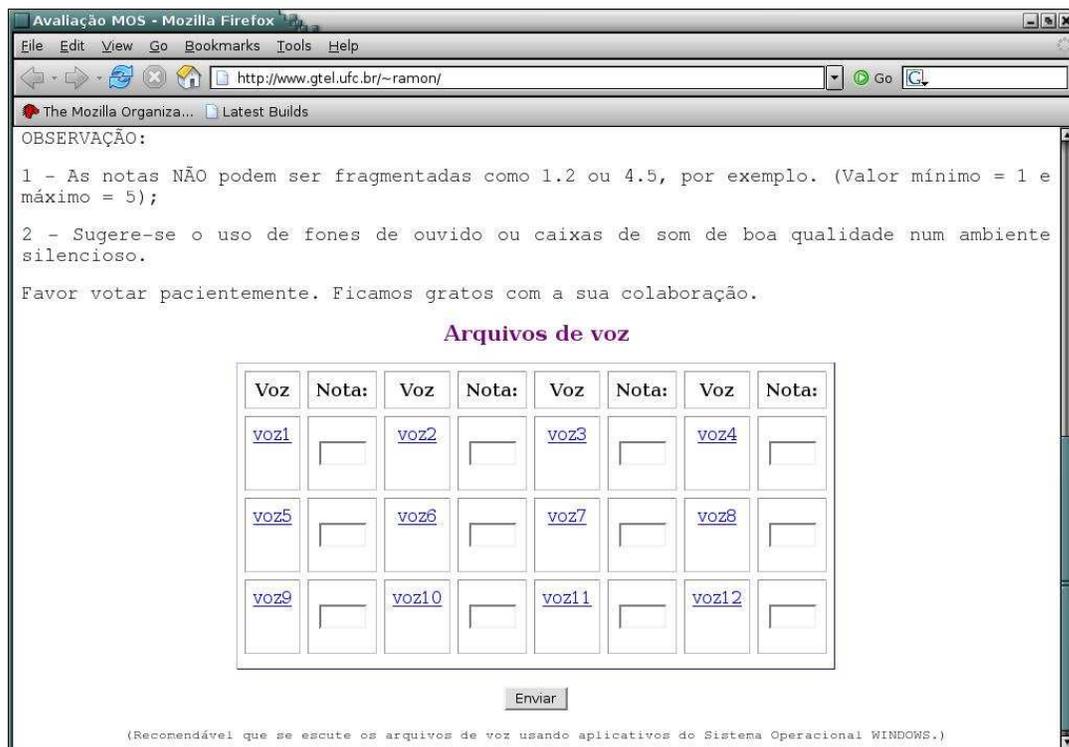


Figura A.2: Parte Inferior da Página na Internet.

Produção Científica

Como produção científica, relacionada ao tema desta Dissertação de Mestrado, publicou-se o artigo transcrito nas próximas páginas deste Apêndice, fruto do trabalho do autor desta Dissertação em colaboração com outros pesquisadores no laboratório do Grupo de Pesquisa em Telecomunicações Sem Fio (GTEL) na Universidade Federal do Ceará (UFC).

Esse artigo trata, também, da avaliação da qualidade de serviço (QoS) do VoIP em Sistemas HSDPA. Na introdução, são destacados os conceitos do serviço de Voz sobre IP e as características necessárias para o seu provimento no HSDPA. Em seguida, descreve-se a pilha de protocolos para o carregamento de voz em redes IP, no caso o codificador AMR, os protocolos RTP, UDP e IP, além do algoritmo de compressão de cabeçalho. O processo de simulação também é relatado através da sua pilha de protocolos, da modelagem do sistema HSDPA, do modelo de canal utilizado para transmissão no meio sem fio e, finalmente, das métricas de qualidade de serviço. Os resultados seguem-se logo após, com os parâmetros utilizados no processo de simulação em função da taxa de quadros com erro, do atraso e do atraso variável. Como conclusões, têm-se que, com o aumento da carga oferecida no sistema e da probabilidade de erro do canal, automaticamente incrementa-se o número de pacotes perdidos e o atraso médio de uma sessão, de modo que a qualidade do serviço VoIP no HSDPA excede ao WCDMA. A capacidade de transmissão, portanto, pode ser aumentada na medida em que se aumenta a complexidade da implementação, já que ela assume algumas simplificações. O artigo finaliza-se com a perspectiva de que novas abordagens devem ser feitas para diminuir as perdas de pacotes e os atrasos, além de alguns outros detalhes que podem ser melhor investigados.

Quality of Service Evaluation of VoIP over HSDPA

Leonardo Ramon N. Sousa, Marcene L. Carvalho, Emanuel B. Rodrigues, Leonardo Sampaio and Francisco R. P. Cavalcanti

Abstract—In this paper we investigate the characteristics of the Voice over IP service deployed over 3G mobile network HSDPA. The necessary protocols for Voice over IP may provide controlled delivery of multimedia traffic over Internet and wireless technologies. We evaluate the VoIP service using a simulator that models the HSDPA system. Results are manifested in terms of packet loss rate, delay and delay jitter, which are important performance measures to guarantee QoS. Moreover, the VoIP performance depends on the system capacity and radio access channel configurations. The understanding of how the VoIP service behaves under different system configurations is given in terms of offered traffic load and wireless channel error probability.

Index Terms—VoIP, HSDPA, Quality of Service, Performance Analysis.

I. INTRODUCTION

Internet telephony is becoming one of the fastest-growing application areas for the Internet today. This quick growth has directed a lot of interest towards Voice over IP (VoIP) and is driven by its several fundamental benefits over traditional Public Switched Telephone Network (PSTN) from the user and network operator's viewpoint: substantial cost reduction, enhancement of voice communication with multimedia, and lower bandwidth consumption.

On the other front, the current development in the cellular radio network technologies is paving the way towards IP capable radio networks. 3G WCDMA networks, e.g. 3GPP release 5 (HSDPA), will provide IP over wireless services, therefore also enabling VoIP. The provision of this emerging real-time service may be of special interest for the mobile network operators for multiple reasons. Firstly, as the bandwidth for individual flows in packet switched domain is not reserved in advance, the multiplexing effects should bring significant capacity savings. Secondly, VoIP service will be supported by the Session Initiation Protocol (SIP), which is a text-based protocol, similar to HTTP and SMTP, for initiating interactive communication sessions between users. Finally, the convergence towards packet switched and IP technology may convince mobile operators to go for solutions that are truly all-IP in order to simplify network interconnection and network management. A common, IP-based transport and service platform will offer mobile users a multitude of real-time and interactive services. IP as a service enabler implies IP all the way to the mobile terminal. The main advantage of running IP all the way over the air interface is service flexibility. That is, there are no dependencies between applications and the network and as a consequence, almost anyone can develop new applications [1, 2].

The authors are with the Wireless Telecommunications Research Group - GTEL, Department of Teleinformatics Engineering - DETI, Federal University of Ceará - UFC, Fortaleza, Brazil. URL: www.gtel.ufc.br E-mail: {ramon.marcone, emanuel, leosam, rodrigo}@gtel.ufc.br

Traditionally, real-time services (e.g. voice) are transported over dedicated channels because of their delay sensitivity while data are transported over shared channels because of its bursty nature. Although the packet data channels of the 3G standards were originally designed with packet data performance in mind, recent modifications to the standards have been made to improve the support of real-time services and, in particular, low bandwidth services such as VoIP.

Because voice applications are real-time, they are intolerant of lengthy delays, packet losses, out-of-order packets and jitter. All these problems gravely degrade the quality of the voice transmitted to the recipient. Unfortunately, wireless networks exacerbate the problems that are intrinsically prevalent in their wire line counterparts: a higher frequency of dropped packets, larger latency and more jitter [3].

HSDPA gives best performance gains for traffic that has no or limited delay requirements as well as no delay variation requirements [4, 5]. However, Conversational Radio Access Bearers (RABs) can be mapped onto an HS-DSCH but its performance will depend on how efficiently radio resources are managed in different system loads.

Providing Quality of Service (QoS), in particular meeting the data rate and packet delay constraints of real-time voice users, is one of the requirements in emerging high-speed data networks, such as HSDPA. However, it is quite a challenge to combine the HSDPA best effort transport channel with QoS guarantees demanded by the VoIP service from a Radio Resource Management (RRM) perspective.

The main objective of this work is to perform a basic provision evaluation of VoIP service on a simplified HSDPA system. The focus will be on the understanding of how the VoIP service behaves under different system configurations in terms of offered traffic load and wireless channel error probability.

This article is organized as follows. The section II speaks of the main necessary protocols for the VoIP service. The following section III describes the simulation modeling and the system HSDPA is described too with your characteristics and our modeling. After that, in the section IV the evaluation and the evidence of the benefits of the solutions proposals is presented by means of simulation results. Finally, a summary of pertinent conclusions and perspectives are shown in the section V.

II. PROTOCOL STACK DESCRIPTION

In order to carry voice on IP networks, appropriate protocols must be used. The main protocols are Real Time Protocol (RTP), User Datagram Protocol (UDP) and Internet Protocol (IP). Before using them, it is necessary also to use an efficient voice encoding and decoding mechanism in order to transform analog signal into digital data [3]. Once signaling and

encoding occur, RTP and UDP are utilized to transport the voice packets. In Fig. 1, the voice frames are generated in the application layer, encoded and encapsulated within payload of an RTP SDU. The RTP PDU is encapsulated into an UDP SDU, which is delivered to the IP layer.

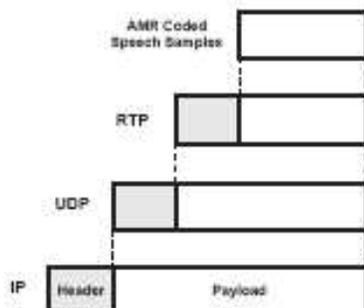


Fig. 1. Transport of speech in IP.

A. Adaptive Multi-Rate (AMR) Speech Codec

The Adaptive Multi-Rate (AMR) Speech Codec was chosen as the mandatory codec for conversational speech services within third generation (3G) cellular systems by the Third Generation Partnership Project (3GPP). It is a codec with 8 narrow-band speech encoding modes with bit rates between 4.75 and 12.2 kbps. If the data rate is 12.2 kbps, the AMR codec generates packets of 244 bits which represent voice frames of 20 ms [6].

The AMR encodes and decodes the digitally sampled speech to make optimum use of power and bandwidth available [7]. It also supports a low-rate background noise encoding mode to reduce transmission during silence periods, further reducing bandwidth and battery usage in the user equipment.

Due to robustness and flexibility of this codec, it is well suited also for voice over IP (VoIP) applications. The Internet Engineering Task Force (IETF) has approved the RTP payload format for AMR to facilitate the usage of this codec with the RTP protocol.

B. RTP

The Real Time Protocol (RTP) is an end-to-end transport protocol [8], but it can be viewed as part of the application layer, because the developer must integrate RTP into the application. It is used to transport multimedia traffic (audio and video communication) in IP networks, supporting unicast and multicast traffic. In the case of VoIP service, it is implemented together with UDP/IP to make use of its multiplexing and checksum services [3].

RTP offers no reliability mechanisms and it does not provide any mechanism to assure synchronized delivery or another type of quality of service (QoS) and guarantee of delivery. This must be assured by the services of the other layers. The RTP protocol also does not approach the aspect of resource reservation [8].

The RTP PDU consists of a header of 12 bytes, optional payload headers and the payload itself. It defines important features:

- Information on the type of carried data (Payload Type);
- It allows the synchronization of media streams carried in more than one session and calculation of jitter (Timestamp);
- Estimation of the number of lost packets and restoration of their sequence (Sequence Number).

Networks metric such as packet delay, jitter and packet loss rate can impact on VoIP quality and they can be measured by the RTP protocol. They are the main performance parameters that should be considered.

C. UDP/IP

The User Datagram Protocol (UDP) is used by VoIP as a transport layer protocol over Internet Protocol (IP) in order to avoid the delay increase caused by retransmissions of lost voice packets.

The UDP protocol is a simple datagram-oriented transport protocol. It provides no reliability on datagram delivery, i.e., there is no confirmation on the delivery of a PDU to its destination. It assumes the Internet Protocol (IP) as the network protocol [9] and it is modeled assuming that each UDP PDU will generate just one IP PDU, i.e. without IP fragmentation.

The main UDP task is to send datagrams from applications to the IP layer without guaranteeing that the destination will be reached [10]. The UDP header size is standardized in 8 bytes and 20 bytes of IPv4 or 40 bytes of IPv6.

D. Header compression (HC)

The large headers of the protocols used when voice data is sent over the Internet constitute a major problem with voice over IP over wireless. Wireless links have a high bit error probability (BEP) due to shadow fading, multi-path fading and mobility. Header compression may be used to minimize the impact of these problems in the performance of the protocols and the VoIP service. For multimedia services, HC achieves a considerable saving in bandwidth, because the headers for RTP/UDP/IP make inefficient use of the radio spectrum and consume too much bandwidth [11].

The multimedia stream packet consists of 12 bytes of the RTP header, 8 bytes of the UDP header and 20 bytes of IPv4 or 40 bytes of IPv6, yielding a total of 40 or 60 bytes. The HC algorithms compress the RTP/UDP/IP header down to 2 octets in the best case [12].

There is a degree of redundancy between header fields of consecutive packets belonging to the packet stream. The compressed headers carry changes to the context and the HC schemes must have mechanisms for repairing and detecting the context when it is out of date. Some examples of HC include Compressed RTP (CRTP) [13], Robust Checksum-Based Header Compression (ROCCO) [11, 13] and Robust Header Compression (ROHC) [14].

III. SIMULATION MODELING

The preliminary performance evaluation of the VoIP service conducted in this study has utilized a static simulator that models in a concise way the HSDPA system.

The simulation consists of several snapshots that represent the life of a VoIP session. This session is related to a user of interest, who competes with other dummy users to get access to the HSDPA shared channel (scheduling process). The user of interest has its own protocol stack composed by the AMR codec, RTP, UDP and IP layers (section III-A). The simplified modeling of some important functionalities of the HSDPA system is presented in section III-B, while the characterization of the wireless channel is shown in section III-C. And in section III-D, it is presented the QoS requirements.

A. Protocol Stack

Fig. 2 shows the information flow through the protocol stack used by a VoIP service. The voice frame of 20 ms is translated into a packet of 244 bits. This voice packet is delivered to the RTP/UDP/IP layers, which insert their respective headers. If there is any HC algorithm, it compresses the RTP/UDP/IP headers into a much lower header providing bandwidth saving in the limited radio spectrum of the wireless channel. The link layer embraces the radio interface protocols of the considered Radio Access Network. In this work, we consider the link layer of the WCDMA/HSDPA system, composed by the RLC and MAC-hs sublayers. In the receiver, the frame is received by the link layer and passed to the upper protocol layers until the decoder. The information flow is modeled in detail in the simulations since all the packets (PDUs) are generated and transmitted through the protocol layers accurately. However some functionalities and protocol layers are simplified or simply not modeled, as indicated in the following.

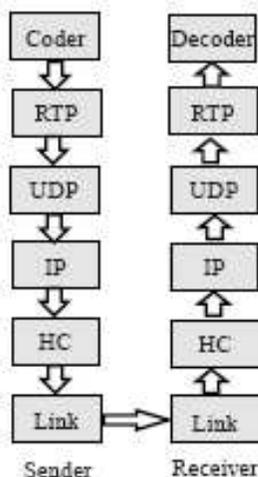


Fig. 2. Information flow in the protocol stack.

The ON/OFF voice traffic model of the user of interest is implemented in detail in the simulator, how a two-state traffic model assuming the use of a slow voice activity detector. The 12.2 kbps AMR codec functioning is also modeled, where

244 bits of encoded data is generated every 20 ms of voice activity periods. During the silence periods, the transmitter is considered to be totally turned off. Regarding the dummy users, their service class is irrelevant. Thus, we consider that each time they are scheduled they have something to transmit.

The VoIP delay budget is characterized in the RTP protocol layer. A fixed delay budget D_{lim} is considered in this work. It means that if the voice packet arrives at the receiver with delay higher than D_{lim} , it is considered as a lost packet, increasing the Frame Erasure Rate (FER) metric. This functionality is implemented in the simulator. However, the RTP buffer was not modeled in the current version of the simulator. This buffer is used to combat the problem of delay jitter (consecutive packets arrive with varying delays) and out-of-order delivery on the receiver. We consider that a packet generated in the transmitter that crosses the communication channel and arrives error-free at the receiver is used by the application layer immediately. We assume that delay jitter and out-of-order delivery are corrected at the RTP buffer.

Although RTP and UDP segmentation is implemented in the simulator, they were not used in the simulated scenario because we assumed that the voice frame generated in the AMR codec was encapsulated in a single RTP SDU and afterwards on a single UDP SDU.

Finally, Header Compression algorithms were not allowed in the simulations. They were left as perspectives for further studies.

B. HSDPA Modeling

The *High Speed Downlink Packet Access* (HSDPA) is a third generation system, where it represents an evolution of the mobile communication bringing a major increase of the data traffic and high bit rate services in downlink direction. It has been included by the 3GPP in the specifications of the Release 5 as an evolution to boost the WCDMA performance for downlink packet traffic. Its concept consists of a new downlink time shared channel that supports a multi-code transmission, smaller transmission time interval (TTI), link adaptation and fast physical layer retransmission combining [15, 16].

The HSDPA technology should be able to increase the system capacity over other technologies and to meet the QoS demands of third generation mobile data services. The QoS is important here because there exists a trade-off between the QoS provided to the end user and the network capacity.

In the HSDPA architecture it is introduced a new transport channel, the High Speed Downlink Shared Channel (HS-DSCH), where a large amount of power and code resources are assigned to a single user at every TTI. The HS-DSCH is directly terminated at the Node B, allowing the acquisition of channel quality reports that enable the tracking of the instantaneous signal quality for low speed mobiles.

Here, the RLC layer of the WCDMA/HSDPA system is modeled in a simplified manner. The RLC Unacknowledged Mode (UM) is adopted since the Voice over IP is a real-time service with stringent delay requirements and cannot tolerate the delay caused by the retransmissions at the RLC level allowed in the Acknowledged Mode (AM).

Regarding the MAC layer, we model some functionalities of the MAC-hs sublayer in a simplified way. It is allowed a maximum number of 3 retransmissions at the MAC level in case the voice packet is received erroneously at the receiver. Moreover, we adopt 4 parallel H-ARQ processes so that voice packets can be processed simultaneously by parallel MAC ARQ stop-and-wait processes. These two functionalities (MAC retransmission and parallel H-ARQ processes) can provoke the occurrence of delay jitter and out-of-order packet delivery. We consider that the HSDPA is always able to allocate a high enough transmission mode (CQI) to transmit the voice packet in just one TTI.

The Packet Scheduling (PS) algorithm adopted in the simulations was the Round Robin (RR), where each user has the same probability to access the shared channel on a cyclic scheduling.

C. Channel Model

The used wireless channel model is a discrete-time first-order Markov model as proposed in [17]. The pattern of errors is described by the transition matrix as shown in equation 1, where P_{BC} is the probability of the channel to change from bad to good state, P_{CB} is the channel transition probability from good to bad state, P_{BB} is the probability of the channel to stay on the bad state, and P_{CC} is the probability of the channel to stay on the good state. The channel characterization is depicted in Fig. 3, where the transitions among states described above can be seen.

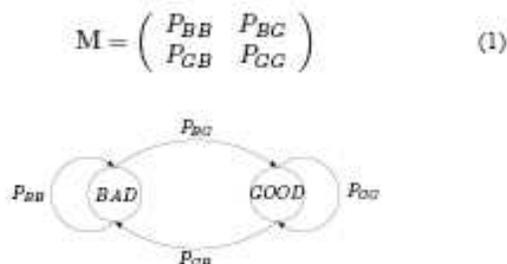


Fig. 3. Channel model

According to [17], given the matrix M , the channel properties are completely characterized. The average slot error probability presented in [17] was used as the effective error probability of the wireless channel. Its calculation is presented in the following equation:

$$\varepsilon = \frac{P_{CB}}{(P_{CB} + P_{BC})} \quad (2)$$

D. Quality of Service Metrics

1) *Frame Erasure Rate*: One of the causes of packet loss is the varying behaviour of the wireless channel. Another possibility is that the delayed packet is regarded as lost (packet delay higher than a threshold). As IP networks do not guarantee that packets will be delivered in a timely manner, packets will be dropped under heavy network traffic loads and during periods of congestion.

Although packet loss of any kind is undesirable, some loss can be tolerated since error-concealment techniques can be used. In general, the quality is not adversely affected as long as the amount of packet loss is less than 5% for the total number of calls [4].

2) *Delay*: The delay is one of the most important QoS measures in the provision of VoIP services. It is defined as the measure of time that it takes the talker's voice to reach the listener's ear. Talk overlap and echo are two drawbacks that are caused by high end-to-end delay in a voice network [4].

According to the ITU E-model [18], the one-way mouth-to-ear delay, which is contributed by the voice encoder, ethernet transmission, the ingress router, Internet, GGSN, SGSN, RNC, Node-B and mobile user processing, should be less than 250-300 ms to achieve a good to acceptable voice quality rating. As we only simulate the Node-B functionalities and the user reception of VoIP packets, we consider that the maximum acceptable latency is D_{lim} at the radio access network only [5].

3) *Delay Jitter*: Jitter is the measure of time difference between a packet's expected arrival time and its actual arrival time. Since voice packets are generated at a constant rate (every 20 ms), every packet would be expected to arrive at the destination exactly every 20 ms [4]. However, functionalities of the MAC-hs sublayer cause an unpredictable amount of delay to occur between these packets.

In order to compensate the jitter, the receiver must implement a buffer, whose size is big enough to keep a continuous flow for the application. This buffer is used to remove jitter of the packets but it also implies in more overhead in the communication process.

The delay jitter of the i^{th} packet can be calculated as shown in equation 3:

$$J = \sqrt{\frac{\sum_{i=1}^L \tau_i^2}{L} - \tau_m^2} \quad (3)$$

where τ_i is the delay of packet i , L is the total number of packets transmitted and τ_m is the mean packet delay [19].

4) *Out-of-Order Packet Delivery*: When diverse and complex networks exist between a transmitter and a receiver in a communication process, the delivery of packets out of order can occur. As example we have the Internet. In order to compensate this effect, the receiver must process these packets in a buffer, which accumulate two functions: packet reordering and elimination of jitter. An RTP entity is responsible for the reordering of voice packets and the basic process uses the sequence number from the RTP header [20].

IV. PRELIMINARY PERFORMANCE RESULTS

The simulation parameters and their values are summarized in Table I.

It should be emphasized that the values of the quality figures depicted in the following graphics should not be considered as absolute performance indicators. The reader should focus on the presented relative comparisons, since the results depend highly on the specific parameters defining the scenario and are

TABLE I
SIMULATION PARAMETERS.

Parameter	Value
Simulation runs	10000
Load	Variable
Channel error probability	Variable
AMR source data rate	12.2 kbps
AMR voice frame duration	20 ms
Voice call mean duration	120 s
Voice on/off mean duration	3 s
AMR voice packet payload size	144 bits
RTP header size	96 bits
RTP padding size	12 bits
UDP header size	64 bits
IPv4 header size	160 bits
RLC mode	Unacknowledged mode (UM)
Maximum number of MAC-hs retransmissions	3
Number of MAC-hs H-ARQ parallel processes	4
Packet scheduling algorithm	Round-Robin
Delay budget (D_{lim})	100 ms

generally only illustrative. Performance of real products and systems may differ considerably.

The next sections will show the evaluation of the VoIP service over HSDPA looking at the main QoS metrics that influence its performance: Frame Erasure Rate (FER), packet delay and packet delay jitter. We will vary two simulation parameters: channel error probability (equation 2) and offered load in terms of number of users.

A. Frame Erasure Rate (FER)

To summarize the evaluation of the FER, Figure 4 shows the 90th percentile for several loads and channel error probabilities, where it can be seen that the higher number of users being scheduled in the HSDPA shared channel, the higher number of lost frames due to delays exceeding the delay budget.

B. Packet Delay

The mean packet delay of the VoIP sessions is assessed in this section. The channel error probability influences the packet delay in the sense that VoIP packets that were received erroneously can be retransmitted up to 3 times by the MAC-hs sublayer. Naturally, these retransmissions increase the overall packet delay. And the variation of the number of users being scheduled in HSDPA also impacts the delay perceived by the voice packets. This happens because the higher the number of

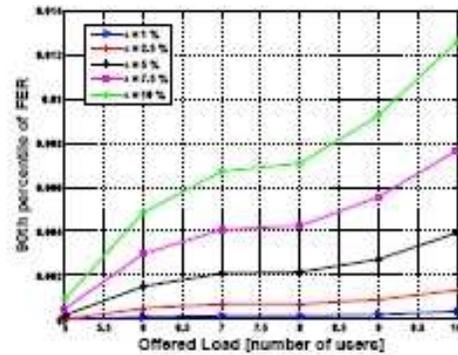


Fig. 4. 90th percentile of FER for several loads and channel error probabilities.

users in the cell, the higher the queueing delay, i.e. the time that the user has to wait to receive data.

The evaluation of the mean packet delay is showed with Figure 5, where the 90th percentile for several loads and channel error probabilities is presented. One perceives that the increase of the channel error probabilities, the delay increases due the retransmissions and when the cell serves more users, the mean delay perceived by the VoIP sessions also increases.

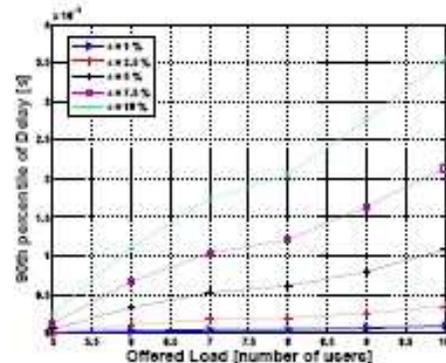


Fig. 5. 90th percentile of mean packet delay for several loads and channel error probabilities.

C. Packet Delay Jitter

Figure 6 shows the Probability Density Function (PDF) of the mean packet delay jitter considering several values of the channel error probability and a fixed offered load of 5 users per cell. It can be seen that an increase in the mentioned parameter increases the delay jitter (delay variance). This is due to the fact that the functionalities of the MAC-hs sublayer (retransmission and H-ARQ processes) can cause delay jitter in the case of non-error-free wireless channel.

V. CONCLUSIONS AND PERSPECTIVES

In this paper, we presented a method to analyze VoIP performance in HSDPA which depends on the system capacity and radio access channel. Delay, delay jitter and frame loss

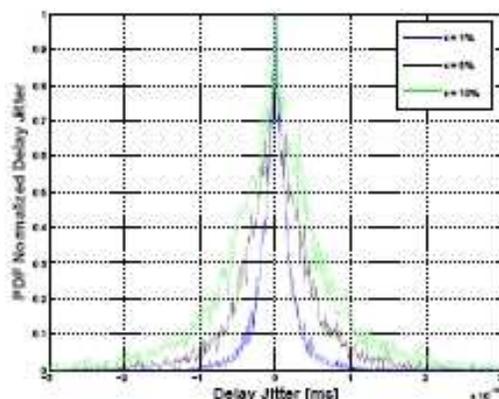


Fig. 6. PDF of the mean packet delay jitter for several channel error probabilities and fixed offered load of 5 users.

rate are the most important performance measurements of the Quality of Service (QoS). We perceived that increasing the offered load and channel error probability, automatically increases the number of lost frames and mean delay perceived by sessions. The quality of VoIP service depends on the implemented protocols and speech encoders. We hope that the capacity of the VoIP service over HSDPA exceed the circuit-switched voice service (WCDMA). The capacity can be improved through increased implementation complexity and this study should not be used for absolute capacities because it made simplifying assumptions. It can be used to understand the gains achievable by the design parameters and features.

In order to provide better QoS for VoIP and to determine specific capacities, new approaches and solutions should be investigated to mitigate the problems associated with delay, delay jitter and packet loss. More detailed simulations are needed to quantify the capacity, as well as to increase the buffer size of the receiver, for example. In the future, the study of header compression algorithms could be evaluated.

ACKNOWLEDGMENTS

We would like to thank Alex P. Silva, Julio F. Pimentel, Tarcisio F. Maciel, and Yuri C. B. Silva for several useful discussions. The authors Leonardo Ramon N. Sousa and Marcone L. Carvalho are scholarships supported by FUNCAP. Francisco R. P. Cavalcanti was partly funded by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, grant no.304477/2002-8.

REFERENCES

- [1] K. Svanbro, H. Hannu, L.-E. Jansson, and M. Degermark, "Wireless real-time ip services enabled by header compression," *Proceedings of IEEE VTC2000*, May 2000.
- [2] R. Cuny and A. Lakanemi, "Voip in 3g networks: An end-to-end quality of service analysis," *IEEE*, 2003.
- [3] P. Mehta and S. Udani, "Voice over ip," *IEEE Potentials*, vol. 20, pp. 36-40, February 2001.
- [4] H. Zheng, G. Rittenhouse, and M. Recchiona, "The performance of voice ip over 3g downlink shared packet channels under different delay budgets," *IEEE Vehicular Technology Conference - VTC*, vol. 4, pp. 2501-2505, October 2003.

- [5] B. Wang, K. I. Pedersen, T. E. Kolding, and P. E. Mogensen, "Performance of voip on hsdpa," *Vehicular Technology Conference - VTC*, July 2005.
- [6] 3GPP, "Mandatory speech codec speech processing functions; smr speech codec; error concealment of lost frames," 3rd Generation Partnership Project, Tech. Rep. TS 26.091, 1999. [Online]. Available: <http://www.3gpp.org>
- [7] M. C. Bale, "Voice and internet multimedia in umts networks," *BT Technology*, vol. 19, no. 1, pp. 48-67, January 2001.
- [8] IETF, "Rtp: A transport protocol for real-time applications," The Internet Engineering Task Force, Tech. Rep. RFC 3550, 2003.
- [9] —, "User datagram protocol," The Internet Engineering Task Force, Tech. Rep. RFC 768, 1980.
- [10] W. R. Stevens, *TCP/IP illustrated, vol. 1: the protocols*. Addison - Wesley Professional Computing Series, 1994.
- [11] G. A. Eriksson, B. Olin, K. Svanbro, and D. Turina, "The challenges of voice-over-ip-over-wireless," *Ericsson Review*, no. 1, June 2000.
- [12] IETF, "Compressing ip/udp/tp headers for low-speed serial links," The Internet Engineering Task Force, Tech. Rep. RFC 2508, 1999.
- [13] A. Cellioglou, S. Fabri, S. Worrall, A. Sadka, and A. Koudoz, "Robust header compression for real-time services in cellular networks," *3G Mobile Communication Technologies*, no. 477, 2001.
- [14] S. Rein, F. H. Fitzek, and M. Reisslein, "Voice quality evaluation for wireless transmission with robc," *International Conference on Internet and Multimedia Systems and Applications - IMSA*, pp. 461-466, August 2003.
- [15] H. Holms and A. Toskala, *WCDMA for UMTS, Radio Access for Third Generation Mobile Communications*. John Wiley & Sons Ltd, 2004.
- [16] J. Korhonen, *Introduction to 3G Mobile Communications*. Artech House, 2003.
- [17] M. Zorzi, E. R. R. Rao, and L. B. Milstein, "On the accuracy of a first-order markov model for data transmission on fading channels," *Fourth IEEE International Conference*, pp. 211 - 215, November 1995.
- [18] I. R. G.114, "One-way transmission time," ITU, Tech. Rep. G.114, 2003.
- [19] M. Kazmi, P. Godlewski, and C. Cordier, "Admission control strategy and scheduling algorithms for downlink packet transmission in wcdma," *IEEE Vehicular Technology Conference - VTC*, 2000.
- [20] A. Kos, B. Klepac, and S. Tomazic, "Techniques for performance improvement of voip applications," *IEEE Mediterranean Electrotechnical Conference - MELTECON*, 2002.

Referências Bibliográficas

- [1] J. Laiho, A. Wacker e T. Novosad, *Radio Network Planning and Optimisation for UMTS*, 1ª edição. John Willey & Sons, 2002.
- [2] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 1ª edição. Prentice Hall PTR, 1999.
- [3] TELECO, “Informação para o aprendizado contínuo em telecomunicação,” acessado em 23 de Maio de 2006. [Online]. Disponível: <http://www.teleco.com.br/voip.asp>
- [4] K. Svanbro, H. Hannu, L.-E. Jonsson e M. Degermark, “Wireless real-time ip services enabled by header compression,” *IEEE 51st Vehicular Technology Conference Proceedings - VTC - Spring*, vol. 2, pp. 1150 – 1154, Maio 2000.
- [5] R. Cuny e A. Lakaniemi, “Voip in 3g networks: An end-to-end quality of service analysis,” *IEEE Vehicular Technology Conference - VTC*, vol. 2, pp. 930–934, Abril 2003.
- [6] A. Lakaniemi, J. Rosti e V. Raisanen, “Subjective voip speech quality evaluation based on network measurements,” *IEEE International Conference on Communications - ICC*, vol. 3, pp. 748–752, Junho 2001.
- [7] I. R. P.800, “Methods for objective and subjective assessment of quality,” ITU, Relatório técnico P.860, 1996.
- [8] M. D. B. Obando, “Comparação do desempenho dos codificadores de voz g.711, g.729, g.723.1 e ilbc em transmissão com perda de pacotes,” Dissertação de mestrado, Universidade de Brasília - UnB, Dezembro 2005.

- [9] L. Sun e E. Ifeachor, “New methods for voice quality evaluation for ip networks,” *International Teletraffic Congress - ITC*, pp. 1201–1210, Setembro 2003.
- [10] H. Hosseini, B. Rohani e B. Rohani, “Objective characterization of voice service quality in wideband cdma,” *IEEE Vehicular Technology Conference - VTS 54th*, vol. 4, pp. 2708 – 2711, Outubro 2001.
- [11] J. W. Seo, S. J. Woo e K. S. Bae, “A study on the application of an amr speech codec to voip,” *IEEE Conference on Acoustics Speech and Signal Processing - ICASSP*, Maio 2001.
- [12] I. Johansson, T. Frankkila e P. Synnergren, “Bandwidth efficient amr operation for voip,” *IEEE Speech Coding Workshop*, pp. 150–152, Outubro 2002.
- [13] M. Kaindl e N. Gortz, “Amr voice transmission over mobile internet,” *Proceedings of the International Conference on Acoustics, Speech and Signal Processing - ICASSP*, pp. 2049–2052, Maio 2002.
- [14] H. Zheng, G. Rittenhouse e M. Recchione, “The performance of voice ip over 3g downlink shared packet channels under different delay budgets,” *IEEE Vehicular Technology Conference - VTC*, vol. 4, pp. 2501–2505, Outubro 2003.
- [15] S. Wanstedt, F. Rui, M. Ericson e M. Nordberg, “Providing reliable and efficient voip over cellular networks,” *Ericsson Review*, no. 1, 2005.
- [16] B. Wang, K. I. Pedersen, T. E. Kolding e P. E. Mogensen, “Performance of voip on hsdpa,” *Vehicular Technology Conference - VTC*, Junho 2005.
- [17] G. Rittenhouse e H. Zheng, “Providing voip service in umts-hsdpa with frame aggregation,” *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP*, Março 2005.
- [18] W. Xiao, A. Ghosh e D. S. L. Downing, “Voice over ip (voip) over cellular: Hrpd-a and hsdpa/hsupa,” *IEEE Vehicular Technology Conference - VTC*, Setembro 2005.
- [19] K. Svanbro, J. Wiorek e B. Olin, “Voice-over-ip-over-wireless,” *Proc. IEEE Int. Symposium on Personal, Indoor & Mobile Radio Communications - PIMRC*, pp. 24–28, 2000.

- [20] L. Sun e E. C. Ifeachor, “Voice quality prediction models and their application in voip networks,” *IEEE transactions on multimedia*, vol. 8, pp. 809–820, Agosto 2006.
- [21] A. D. Elbayoumy e S. J. Shepherd, “Qos control using an end-point cpu capability detector in a secure voip system,” *Proceedings of the 10th IEEE Symposium on Computers and Communications, ISCC 2005.*, pp. 792–797, Junho 2005.
- [22] M. Mehmood, T. Jadoon e N. Sheikh, “Assessment of voip quality over access networks,” *The First IEEE and IFIP International Conference in Central Asia on Internet*, p. 5 pp., Setembro 2005.
- [23] 3GPP, “Mandatory speech codec speech processing,” 3rd Generation Partnership Project, Relatório técnico TS 26.071 v5.0.0, 2002. [Online]. Disponível: <http://www.3gpp.org>
- [24] —, “Mandatory speech codec speech processing functions; amr speech codec; error concealment of lost frames,” 3rd Generation Partnership Project, Relatório técnico TS 26.091 v3.1.0, 1999. [Online]. Disponível: <http://www.3gpp.org>
- [25] —, “Adaptive multi-rate (amr) speech codec; interface to iu and uu,” 3rd Generation Partnership Project, Relatório técnico TS 26.102 v6.0.0, 2004. [Online]. Disponível: <http://www.3gpp.org>
- [26] M. C. Bale, “Voice and internet multimedia in umts networks,” *BT Technology*, vol. 19, Janeiro 2001.
- [27] A. Lakaniemi, P. Ojala e H. Toukomaa, “Amr and amr-wb rtp payload usage in packet switched conversational multimedia services,” *IEEE Workshop Proceedings Speech Coding*, pp. 147–149, Outubro 2002.
- [28] H. Holma e A. Toskala, *WCDMA for UMTS. Radio Access for Third Generation Mobile Communications*. John Wiley & Sons Ltd, 2004.
- [29] IETF, “Rtp: A transport protocol for real-time applications,” The Internet Engineering Task Force, Relatório técnico RFC 3550, 2003.

- [30] T. Koistinen, “Protocol overview: Rtp and rtcp,” *Nokia Telecommunications*, 2000.
- [31] A. Kos, B. Klepec e S. Tomazic, “Techniques for performance improvement of voip applications,” *IEEE Mediterranean Electrotechnical Conference - MELECON*, Fevereiro 2002.
- [32] P. Mehta e S. Udani, “Voice over ip,” *IEEE Potentials*, vol. 20, pp. 36–40, Fevereiro 2001.
- [33] J. L. Ucles e C. N. Manikopoulos, “Use of the burst voice arrival lag (bval) as a measure of performance that detects packet burst errors (pber) for voice over ip systems,” *IEEE Military Communications Conference Proceedings - MILCOM*, 1999.
- [34] H. Nyberg, C. Johansson e B. Olin, “A streaming video traffic model for the mobile access network,” *IEEE Vehicular Technology Conference - VTC*, vol. 1, 2001.
- [35] U. Varshney, A. Snow, M. McGivern e C. Howard, “Voice over ip,” vol. 45, no. 1, 2002.
- [36] IETF, “User datagram protocol,” The Internet Engineering Task Force, Relatório técnico RFC 768, 1980.
- [37] W. R. Stevens, *TCP/IP illustrated, vol. 1: the protocols*. Addison - Wesley Professional Computing Series, 1994.
- [38] A. Mena e J. Heidemann, “An empirical study of real audio traffic,” *IEEE Computer and Communication Societies - INFOCOM*, vol. 1, pp. 26–30, Março 2000.
- [39] T. Lang, “Evaluation of different tcp versions in non—wireline environments,” Tese de doutorado, The University of South Australia, Institute for Telecommunications Research, Agosto 2002.
- [40] A. de Sena Almeida, “Análise do problema do protocolo mac ieee 802.11 em redes ad hoc multihop,” Dissertação de mestrado, Universidade Federal de Pernambuco, Centro de Informática, Pós-Graduaçãoem Ciência da Computação, Fevereiro 2003.

- [41] IETF, “Internet protocol,” The Internet Engineering Task Force, Relatório técnico RFC 791, 1981.
- [42] G. A. Eriksson, B. Olin, K. Svanbro e D. Turina, “The challenges of voice-over-ip-over-wireless,” *Ericsson Review*, no. 1, 2000.
- [43] Y.-S. Kim, “Capacity of voip over hsdpa with frame bundling,” *The Institute of Electronics, Information and Communication Engineers (IEICE) Transactions Communications*, no. 12, pp. 3450–3453, Dezembro 2006.
- [44] IETF, “Compressing ip/udp/rtp headers for low-speed serial links,” The Internet Engineering Task Force, Relatório técnico RFC 2508, 1999.
- [45] A. Cellatoglu, S. Fabri, S. Worrall, A. Sadka e A. Kondo, “Robust header compression for real-time services in cellular networks,” *3G Mobile Communication Technologies*, no. 477, Março 2001.
- [46] S. Rein, F. H. Fitzek e M. Reisslein, “Voice quality evaluation for wireless transmission with rohc,” *International Conference on Internet and Multimedia Systems and Applications - IMSA*, pp. 461–466, Agosto 2003.
- [47] G. Malmgren, “Voice over ip over wireless,” *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications - PIMRC*, Setembro 2000.
- [48] R. Agrawal e N. Gupta, “To improve the voice quality over ip using channel coding,” *6th International Conference On ASIC, ASICON 2005.*, vol. 1, pp. 111–117, Outubro 2005.
- [49] T. Yletyinen, “The quality of voice over ip,” Dissertação de mestrado, Helsinki University of Technology, Faculty of Electrical and Communications Engineering, Laboratory of Telecommunications Technology, Março 1998.
- [50] B. Larish e P. Nguyen, “Quality of service analysis using the georgia tech network simulator,” *IEEE Military Communications Conference, 2005. MILCOM 2005.*, vol. 4, pp. 2649–2654, Outubro 2005.
- [51] J. Zhang, Y. Dacheng e Z. Quan, “Voice quality of voip in mobile communication systems,” *IEEE Radio and Wireless Symposium*, pp. 131–134, Janeiro 2006.

- [52] W. Jiang, K. Koguchi e H. Schulzrinne, “Qos evaluation of voip end-points,” *In Proceedings of IEEE International Conference on Communications - ICC 2003*, vol. 3, pp. 1917–1921, Maio 2003.
- [53] T.-K. Chua e D. C. Pheanis, “Effects of loss characteristics on loss-recovery techniques for voip,” *Proceedings of the International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL’06)*, p. 204, 2006.
- [54] Q. de Serviço, “Desafios para a implementação de voip com qos,” acesso em 06 de Junho de 2007. [Online]. Disponível: <http://www.dei.unicap.br/almir/seminarios/2005.1/ts04/voip/12-DesafiosImplementacao.htm>
- [55] IETF, “An architecture for differentiated services,” The Internet Engineering Task Force, Relatório técnico RFC 2475, 1998.
- [56] M. Rahmaniheris e M. Sedighi, “A new algorithm for resource manager to enhance the quality of service in voip systems,” *International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL’06)*, p. 2, 2006.
- [57] IETF, “Integrated services in the internet architecture: an overview,” The Internet Engineering Task Force, Relatório técnico RFC 1633, 1994.
- [58] R. Lloyd-Evans, *QoS in Integrated 3G Networks*, 1ª edição. Artech House Publishers, 2002.
- [59] O. Hersent, D. Gurle e J. P. Petit, *IP Telephony*, 1ª edição. Prentice Hall, 2002.
- [60] I. R. P.861, “Objective quality measurement of telephone-band (300-340 hz) speech codecs,” ITU, Relatório técnico, 1998.
- [61] I. R. P.862, “Perceptual evaluation of speech quality (pesq), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” ITU, Relatório técnico P.862, 2001.
- [62] K. Doh-Suk, “A cue for objective speech quality estimation in temporal envelope representations,” *IEEE Signals Processing Letters*, vol. 11, no. 10, pp. 849–852, Outubro 2004.

- [63] Y. Bai e M. Ito, “A study for providing better quality of service to voip users,” *20th International Conference on Advanced Information Networking and Applications (AINA'06)*, vol. 1, pp. 799–804, 2006.
- [64] I. R. G.114, “One-way transmission time,” ITU, Relatório técnico G.114, 2003.
- [65] 3GPP, “Services and service capacities,” 3rd Generation Partnership Project, Relatório técnico TS 22.105 v3.5.0, 1999. [Online]. Disponível: <http://www.3gpp.org>
- [66] A. Schieder e T. Ley, “Enhanced voice over ip support in gprs and egprs,” *IEEE Wireless. Communications and Networking Conference*, vol. 2, pp. 803–8008, Setembro 2000.
- [67] M. Kazmi, P. Godlewski e C. Cordier, “Admission control strategy and scheduling algorithms for downlink packet transmission in wcdma,” *IEEE Vehicular Technology Conference - VTC*, Outubro 2000.
- [68] U. W. H. Page, “Umts world - news and information about 3g mobile networks,” 2006, acessado em 06 de Março de 2006. [Online]. Disponível: <http://www.umtsworld.com/default.htm>
- [69] J. Korhonen, *Introduction to 3G Mobile Communications*, 2^a edição. Artech House, 2003.
- [70] R. G. de Queiroz, “Análise de requisitos para o encaminhamento eficiente de tráfego sobre w-cdma,” Dissertação de mestrado, Universidade Federal de Pernambuco, Centro de Informática, Pós-Graduação em Ciência da Computação, Fevereiro 2003.
- [71] 3GPP, “Quality of service (qos) concept and architecture,” 3rd Generation Partnership Project, Relatório técnico TS 23.107 v5.10.0, 2003. [Online]. Disponível: <http://www.3gpp.org>
- [72] J. M. C. Brito, “Controle de erro em redes atm sem fio,” Tese de doutorado, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Departamento de Comunicações, Junho 2003.
- [73] A. R. Braga, “Controle de congestionamento para voz sobre ip em hsdpa,” Dissertação de mestrado, Universidade Federal do Ceará - UFC, Maio 2006.

- [74] J. M. HOLTZMAN, “Forward link waterfilling power control,” *Vehicular Technology Conference - VTC 2000 Spring*, vol. 3, pp. 1663–1667, 2000.
- [75] R. C. ELLIOTT e et al, “Scheduling algorithms for the cdma2000 packet data evolution,” *Vehicular Technology Conference - VTC 2002 Fall*, vol. 1, pp. 304–310, 2002.
- [76] J. M. HOLTZMAN, “Asymptotic analysis of the proportional fair algorithm,” *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. F–33–F37, Setembro 2001.
- [77] H. Lundevall, B. Olin, J. Olsson, N. Wiberg, S. Wanstedt, J. Eriksson e F. Eng, “Streaming applications over hsdpa in mixed service scenarios,” *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 2, pp. 841–845, Setembro 2004.
- [78] H. J. Bang, T. Ekman e D. Gesbert, “A channel predictive prportional fair scheduling algorithm,” 2006, acessado em: Agosto de 2006. [Online]. Disponível: <http://www.eurecom.fr/gesbert/papers/Spawc2005Hans.pdf>
- [79] A. R. Braga, E. B. Rodrigues e F. R. P. Cavalcanti, “Packet scheduling for voip over hsdpa in mixed traffic scenarios,” *The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 06)*, pp. 1–5, Setembro 2006.
- [80] R. J. de Assis Correa, “Análise de desempenho de enlaces com modulação adaptativa aplicada a sistemas wcdma/hsdpa,” Dissertação de mestrado, Pontifícia Universidade Católica do Rio de Janeiro, Julho 2003.
- [81] M. C. Ramos, “Desempenho da modulação adaptativa em enlaces wcdma/hsdpa em presença de multipercursos,” Dissertação de mestrado, Pontifícia Universidade Católica do Rio de Janeiro, Agosto 2004.
- [82] 3GPP, “Rlc protocol specification,” *3rd Generation Partnership Project*, Relatório técnico TS 25.322 v3.0.0, 1999. [Online]. Disponível: <http://www.3gpp.org>
- [83] J. Li, “A performance evaluation of the radio link control protocol in 3g umts,” Dissertação de mestrado, School of Mathematics and

- Statistics, Ottawa-Carleton Institute of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada, Agosto 2003.
- [84] H. Holma, M. Kuusela, E. Malkamaki, K. Ranta-aho e C. Tao, “Voip over hspa with 3gpp release 7,” *The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 06)*, pp. 1–5, Setembro 2006.
- [85] L. R. N. Sousa, M. L. Carvalho, E. B. Rodrigues, L. Sampaio e F. R. P. Cavalcanti, “Quality of service evaluation of voip over hsdpa,” *VI International Telecommunications Symposium - ITS2006*, Setembro 2006.
- [86] P. Hosein, “Scheduling of voip traffic over a time-shared wireless packet data channel,” *IEEE International Conference on Personal Wireless Communication - ICPWC*, Janeiro 2005.
- [87] J. S. Garofolo, L. F. Lamel, W. M. Fisher e et al, *TIMIT Acoustic-Phonetic Continuous Speech Corpus*, TIMIT data set.
- [88] G. Inc., “Goldwave - audio software,” acesso em 29 de Abril de 2006. [Online]. Disponível: <http://www.goldwave.com>
- [89] 3GPP, “Performance characterization of the adaptive multi-rate (amr) speech codec,” 3rd Generation Partnership Project, Relatório técnico TR 26.975 v6.0.0 - Release 6, Dezembro 2004. [Online]. Disponível: <http://www.3gpp.org>
- [90] H. L. M. J. P Ojala, A Lakaniemi, “The adaptive multirate wideband speech codec: system characteristics, quality advances, and deployment strategies.” *IEEE Communications Magazine*, vol. 44, no. 5, pp. 59–65, Maio 2006.
- [91] I. R. P.830, “Subjective performance assessment of telephoneband and wideband digital codecs,” ITU, Relatório técnico P.830, 1996.
- [92] J. C. Magro, “Estudo da qualidade de voz em redes ip,” Dissertação de mestrado, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e Computação, Departamento de Comunicações, Julho 2005.
- [93] T. Rohdenburg, V. Hohmann e B. Kollmeier, “Objective perceptual quality measures for the evaluation of noise reduction schemes,” *In Proc. of the*

International Workshop on Acoustic Echo and Noise (IWAENC), pp. 169–172,
Setembro 2005.