SÉRGIO COSTA MARTINS DE ALENCAR

**ESTATÍSTICA MULTIVARIADA APLICADA NO
CORRELACIONAMENTO DA QUALIDADE DE SERVIÇOS EM
CHAMADAS VOIP E A QUALIDADE DA FALA AFERIDA PELA
RECOMENDAÇÃO ITU-T G.107**

São Paulo
2011

SÉRGIO COSTA MARTINS DE ALENCAR

**ESTATÍSTICA MULTIVARIADA APLICADA NO
CORRELACIONAMENTO DA QUALIDADE DE SERVIÇOS EM
CHAMADAS VOIP E A QUALIDADE DA FALA AFERIDA PELA
RECOMENDAÇÃO ITU-T G.107**

Dissertação apresentada à Escola
Politécnica da Universidade de São Paulo
para obtenção do título de Mestre em
Engenharia

Área de Concentração:
Sistemas Digitais

Orientador:
Prof. Dr. Moacyr Martucci Junior

São Paulo
2011

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 06 de Dezembro de 2011.

Assinatura do autor

Assinatura do orientador

FICHA CATALOGRÁFICA

Alencar, Sérgio Costa Martins de

Estatística multivariada aplicada no correlacionamento da qualidade de serviços em chamadas VOIP e a qualidade da fala aferida pela recomendação ITU-T G.107 / S.C.M. de Alencar -- ed.rev. -- São Paulo, 2011.

102 p.

Dissertação (Mestrado Interinstitucional) - Departamento de Engenharia de Computação e Sistemas Digitais. Escola Politécnica da Universidade de São Paulo / Universidade Estadual do Amazonas.

1. Telecomunicações 2. Telefonia 3. Protocolos de comunicação 4. PABX 5. Voz I. Universidade de São Paulo. Escola Politécnica II. Universidade Estadual do Amazonas III. t.

DEDICATÓRIA

Dedico este trabalho ao meu pai (in memoriam), pessoa de imenso caráter e terno coração, herói de minha infância, sábio amigo e conselheiro. Sua obstinação frente às inúmeras adversidades ao longo de nossas vidas e seu verdadeiro sacerdócio no exercício da paternidade permitiu minha formação acadêmica e profissional e construíram a pessoa que sou hoje.

AGRADECIMENTOS

Agradeço primeiramente a Deus por guiar-me ao longo desta jornada e fortalecer-me diante das adversidades.

A minha esposa Maria pela imensa paciência e inabalável fé. Aos meus filhos Olivia e Matheus pela compreensão frente aos diversos momentos onde foi necessário abdicar do convívio familiar para a consecução deste trabalho e aos demais familiares pelo suporte.

Ao meu orientador, Prof. Dr. Moacyr Martucci Junior, pela paciência, boa vontade e principalmente por nunca ter desistido de mim, mesmo nos momentos mais sombrios. Sem sua confiança e determinação, não haveria a possibilidade de execução deste trabalho.

Aos coordenadores do programa de Mestrado Interinstitucional e em especial ao Prof. Dr. Marco Túlio pelo entusiasmo e disposição em ajudar.

Ao Prof. Dr. André Riyuiti Hirakawa e ao Prof. Dr. Carlos Cugnasca, pelo apoio, críticas e sugestões no processo de qualificação.

Ao amigo e colega de mestrado Marcelo Queiroz Leite, que inúmeras vezes com suas palavras de incentivo injetou ânimo e fortaleceu meu espírito frente aos obstáculos que se interpunham.

Ao amigo Edson de Souza, responsável pela infraestrutura do LAA, onde passei a maior parte do tempo durante a etapa de qualificação, por sua boa vontade em sempre estar presente auxiliando nos problemas relacionados ao laboratório.

Ao INSTITUTO CERTI AMAZÔNIA, por promover a capacitação e aperfeiçoamento de seus colaboradores e a TRÓPICO SISTEMAS E TELECOMUNICAÇÕES DA AMAZÔNIA LTDA por prover os meios tecnológicos para que este trabalho fosse concluído.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela criação e regulamentação do programa de Mestrado Interinstitucional, Minter.

A Fundação de Amparo e Pesquisas do Estado do Amazonas – FAPEAM, pela concessão da Bolsa de Mestrado ao autor.

A Superintendência da Zona Franca de Manaus – SUFRAMA, pelo apoio financeiro prestado ao programa Minter.

A Escola Politécnica da USP e a Universidade do Estado do Amazonas.

RESUMO

Vivemos atualmente uma era de convergência de tecnologias, motivada por questões tanto econômicas como de caráter operacional, na qual os serviços de dados, voz e vídeo estão migrando rapidamente para uma plataforma IP. Particularmente considerando o paradigma da telefonia IP neste processo de convergência, ocorrem desafios tecnológicos, pois temos de um lado os usuários finais que já possuem uma referência sobre a qualidade da fala, fruto das décadas de uso do sistema telefônico tradicional, e na outra extremidade as operadoras de telecomunicações que em sua última milha dependem de redes estatísticas, sem mecanismos adequados para a garantia de QoS. Assim, se torna vital para o sucesso da operação a devida identificação das relações entre os diversos componentes existentes entre terminais e sua contribuição para a qualidade de fala, percebida pelo assinante, de forma a entregar um serviço com qualidade similar ao Sistema Telefônico Fixo Comutado.

Neste contexto, este trabalho busca identificar por meio de técnicas de estatística multivariada uma correlação entre métricas objetivas de Qualidade de Serviços aplicáveis em redes IP e a qualidade subjetiva da fala predita pelo algoritmo Modelo-E definido na recomendação ITU-T G.107. Um processo para coleta e análise estatística de informações foi desenvolvido para atingir o objetivo proposto. Para sua validação um ambiente de testes foi criado, dados de operação foram coletados e ferramentas computacionais foram aplicadas para o tratamento analítico e estatístico. Os resultados obtidos pelo processo foram então aplicados em campo durante as etapas de testes e homologação de um PABX-IP-IMS desenvolvido para o mercado corporativo.

A correlação entre os diversos fatores envolvidos, suas métricas e como todo este sistema impacta na qualidade relativa, percebida pelo usuário final permitirá aos provedores de serviços avaliarem quais as melhores estratégias a serem empregadas em seus segmentos de rede de forma a garantir a excelência no nível de serviço oferecido ao consumidor final.

Palavras-chave: Estatística multivariada, qualidade da fala, qualidade de serviços, VoIP.

ABSTRACT

We live now in an convergence of technologies era, driven by economic and operational issues, where the data services, voice and video are quickly moving to an IP platform. Particularly considering the paradigm of IP telephony in the process of convergence, there are technological challenges. We have subscribers who already have a reference about the quality of speech, derived from decades of using the traditional phone system. At the other end telecom operators that rely on statistical networks, with no possibility to guarantee QoS. So it becomes vital to the operation's success the proper identification of the relationships between the various components between the terminals and their contribution to the speech quality perceived by the subscriber in order to deliver a quality service close to the PSTN.

In this context, this study sought to identify a correlation between objective metrics for Quality of Service applicable to IP networks and subjective quality of speech predicted by the algorithm "Model-E" defined in ITU-T G.107 through multivariate statistical techniques. A process of collecting and analyzing statistical information was developed to achieve the proposed objective. To validate, a test environment was created, operation data were collected and computational tools were applied to the analytical and statistical treatment. The results obtained by the process were then applied in the field during the stages of testing and approval of an IMS-IP-PBX designed for the corporate market.

The correlation between the various factors involved, their metrics and how the whole system impacts on the quality perceived by end users will enable service providers to assess what the best strategies to use in their network segments to ensure an adequate level of service offered to consumers.

Keywords: Multivariate analysis, speech quality, quality of services, VoIP.

Que força é esta, eu não sei; tudo o que sei é que existe, e está disponível apenas quando alguém está em um estado em que sabe exatamente o que quer, e está totalmente determinado a não desistir até conseguir.

Alexander Graham Bell

LISTA DE FIGURAS

Figura 1 – Evolução tecnológica da telefonia.....	10
Figura 2 – Infraestrutura VoIP.....	11
Figura 3 – A grande figura de um provedor VoIP.....	16
Figura 4 – Representação do método para estabelecimento de correlação.	40
Figura 5 – Representação gráfica do comportamento de R_{xy}	44
Figura 6 – Ambiente experimental.....	49
Figura 7 – Aplicação Java para administração do PABX-IP.....	50
Figura 8 – Interface gráfica do softphone ZOIPER 2.37.....	51
Figura 9 – SIP em modo console.....	51
Figura 10 – Página principal do VQManager.....	52
Figura 11 – Experimento para validação do método de correlação.....	54
Figura 12 – Ambiente de testes (laboratório).....	56
Figura 13 – Ambiente de testes (campo).....	58
Figura 14 – Interface gráfica do Software R para computação estatística.	59
Figura 15 – Captura de tela do R: Sumario.....	60
Figura 16 – Captura de tela do R: Transformação Z	61
Figura 17 – Captura de tela do R: Histograma do MOS_Z.....	61
Figura 18 – Histogramas das variáveis normais padronizadas	62
Figura 19 – Captura de tela do R: Gráfico QQ-Plot para MOS_Z.	63
Figura 20 – Gráfico QQPlot para todas as variáveis.	63
Figura 21 – Captura de tela do R: Gráfico de dispersão para 02 variáveis.	64
Figura 22 – Gráfico de dispersão MOS_Z x Delay_Z.....	65
Figura 23 – Gráficos de dispersão	66
Figura 24 - Captura de tela do R: Dendrograma.	67
Figura 25 – Dendrograma da amostra utilizando o índice de Jaccard.	67
Figura 26 – Captura de tela do R: Cálculo dos coeficientes de Pearson	68
Figura 27 – Captura de tela do R: Teste de significância.	69

LISTA DE EQUAÇÕES

Equação 01 – Fator R	24
Equação 02 – Média Aritimética	28
Equação 03 – Variância.....	28
Equação 04 – Desvio padrão	29
Equação 05 – Curva Normal	29
Equação 06 – Transformação Z	33
Equação 07 – Coeficiente de correlação linear de Pearson.....	42
Equação 08 – Índice de similaridade de Jaccard para conjuntos.....	45
Equação 09 – Índice de similaridade de Jaccard para amostra	45
Equação 10 – Índice t-student.....	46

LISTA DE TABELAS

Tabela 1 – Interpretação do coeficiente R_{xy}	43
Tabela 2 – Valores de t-student em níveis de 5% e 1% ($\alpha=0,05$ a $0,01$) de probabilidade.....	47
Tabela 3 – Exemplo de Arquivo .CSV do SIPP.	56
Tabela 4 – Exemplo de arquivo .CSV gerado pelo VQManager.....	57
Tabela 5 – Sumário do conjunto de dados após importação.....	60
Tabela 6 – Agrupamentos identificados no dendrograma.....	68
Tabela 7 – Coeficientes de Correlação Linear de Pearson	69
Tabela 8 – Resultado do teste t-student para verificação da significância de R_{xy}	70

LISTA DE ABREVIATURAS E SIGLAS

AAA – Authentication, Authorization and Accounting

ACR – Absolute Category Rate

AED – Análise Exploratória de Dados

CODEC – Codificador e decodificador

CPA – Controle por Programa Armazenado

DID – Direct Inward Dialing

DiffServ –Differentiated Services

ETSI – European Telecommunications Standards Institute

FIFO – First in, first out

GSM – Global System for Mobile Communications

HTTP – Hypertext Transfer Protocol

IBGE – Instituto Brasileiro de Geografia e Estatística

IETF – Internet Engineering Task Force

IntServ – Integrated Services

IP – Internet Protocol

ISDN-N – Integrated Services Digital Network – Narrow-band

ITSP – Internet Telephony Service Provider

ITU-T – International Telecommunication Union - Telecommunication
Standardization Sector

MOS – Mean Opinion Score

MPLS – Multi Protocol Label Switching

MTBF – Mean Time Between Failures

MTTR – Mean Time to Repair

NAT – Network Address Translation

PAT – Port Address Translation

PNUD – Pesquisa Nacional por Amostra de Domicílio

PQ – Priority Queuing

QoS – Quality of Services

RFC – Request For Comment

RTCP – Real Time Control Protocol

RTP – Real Time Transport Protocol

SDP – Session Description Protocol

SGBD – Sistema Gerenciador de Banco de Dados

SIP – Session Initialization Protocol

SLA – Service Level Agreement

SMTP – Simple Network Management Protocol

STFC – Serviço Telefônico Fixo Comutado

TCP – Transmission Control Protocol

TDM – Time Domain Multiplexing

TIPHON – Telecommunications and Internet Protocol Harmonization Over Networks

UA – User Agent

UAC – User Agent Client

UAS – User Agent Server

UDP – User Datagram Protocol

UML – Unified Modeling Language

UTF-8 – 8-bit Unicode Transformation Format

VOIP – Voice Over Internet Protocol

WFQ – Weighted Fair Queuing

SUMÁRIO

1 INTRODUÇÃO	1
1.1 JUSTIFICATIVA.....	2
1.2 OBJETIVOS	5
1.3 METODOLOGIA	5
1.4 ORGANIZAÇÃO DO TRABALHO	7
2 FUNDAMENTAÇÃO TEÓRICA	8
2.1 TRANSMISSÃO DA VOZ HUMANA.....	8
2.1.1 SERVIÇO TELEFÔNICO FIXO COMUTADO (STFC)	8
2.1.2 VOZ SOBRE IP.....	10
2.1.2.1 PROTOCOLO DE INICIAÇÃO DE SESSÃO	12
2.1.2.2 PROTOCOLO DE TRANSPORTE DE TEMPO REAL.....	14
2.1.2.3 INFRAESTRUTURA DE UMA OPERADORA VOIP	15
2.2 QUALIDADE NA TRANSMISSÃO DE VOZ	20
2.2.1 CONCEITO DE QUALIDADE	20
2.2.2 QUALIDADE DA FALA.....	22
2.2.3 QUALIDADE DE SERVIÇOS.....	25
2.3 TÉCNICAS ESTATÍSTICAS	27
2.3.1 DADOS MULTIVARIADOS	30
2.3.2 ESTATÍSTICA MULTIVARIADA.....	30
2.3.3 ANÁLISE DE AGRUPAMENTO HIERÁRQUICO (HCA)	31
2.3.4 VARIÁVEL NORMAL PADRONIZADA	33
3 PROCESSO PROPOSTO	34
3.1 VISÃO GERAL	34
3.1.1 PRINCIPAIS PROBLEMAS EM CHAMADAS VOIP.....	35
3.1.2 SOLUÇÃO DE PROBLEMAS EM SISTEMAS VOIP	36
3.1.3 QUALIDADE DA FALA COMO CRITÉRIO PARA GERENCIAMENTO DA REDE	38
3.2 CORRELAÇÃO ENTRE MÉTRICAS DE QoS E QUALIDADE DA FALA.....	39
3.3 APLICABILIDADE E LIMITAÇÕES	47
4 ESTUDO DE CASO	48
4.1 AMBIENTE EXPERIMENTAL.....	48
4.2 DESCRIÇÃO DO EXPERIMENTO.....	53

4.3 EXECUÇÃO DOS TESTES	55
4.3.1 TESTES DE LABORATÓRIO.....	55
4.3.2 TESTES DE CAMPO	57
4.3.3 ANÁLISE EXPLORATÓRIA DE DADOS (AED)	58
4.3.4 RESULTADOS	70
5 CONCLUSÕES	73
5.1 CONTRIBUIÇÕES	73
5.2 TRABALHOS FUTUROS	75
6 REFERÊNCIAS.....	76
ANEXO A - SCRIPTS EM XML PARA PARAMETRIZAÇÃO DO SIPP COMO UAC.....	81
ANEXO B - SCRIPTS EM XML PARA PARAMETRIZAÇÃO DO SIPP COMO UAS.....	84

1 INTRODUÇÃO

O acesso às redes IP vem progressivamente tornando-se comum no dia-a-dia do consumidor. Considerando particularmente o Brasil, o número de usuários de internet atinge a marca de 57 milhões de brasileiros ou cerca de 32% da população acima de 10 anos de idade segundo os dados oficiais disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), referentes à Pesquisa Nacional por Amostra de Domicílios do ano de 2008 (PNAD, 2009). Boa parte deste contingente passou a fazer uso de outras formas de comunicação disponibilizadas pela conectividade à grande rede, dentre elas a transmissão de voz sobre o protocolo IP (VoIP). Serviços com tal capacidade surgiram e morreram ao logo desta última década, fornecendo diversos níveis de qualidade, quase sempre sem ameaçar as redes legadas do Serviço Telefônico Fixo Comutado (STFC). No entanto, com o advento da possibilidade de entroncamento entre sistemas VoIP e redes do STFC, permitindo chamadas com “qualidade suficiente” baseadas em tarifação diferenciada, muitas vezes com custo local para chamadas de longa distância, ocorre uma divisão no mercado de telecomunicações. Esses novos provedores de serviços VoIP passaram a incomodar os grandes jogadores, que se viram compelidos a também iniciar operações dentro desse novo mercado, tentando compartilhar parte dessa receita, perdida para os novos serviços emergentes.

Com a estabilização na prestação desse serviço o novo horizonte, principalmente para os operadores das redes legadas, passa a ser tentar garantir a mesma qualidade da fala já consagrada pelas redes do STFC para esse serviço de voz sobre IP. Aqui se estabeleceu um ponto nevrálgico, pois nesse novo serviço, o operador não detém controle total do enlace. A última milha normalmente é uma conexão internet disponibilizada por algum provedor de serviços de dados local do qual o usuário final é assinante e representa na maioria das vezes uma conexão de qualidade questionável. A questão final neste embate é como garantir um nível de satisfação adequado para o assinante em suas chamadas de voz sem poder gerir o canal, uma vez que o assinante faz uso em sua última milha de redes estatísticas de terceiros. Alguns grandes operadores como a inglesa Orange (ORANGE, 2009) resolveram esta equação fornecendo a conexão para seus usuários em *combos*

domésticos ou por meio de *hot-spots* administrados pela própria empresa. No entanto, nem todas as operadoras têm o interesse ou a capacidade de alterar sua operação de forma a atender aos novos tempos.

A questão do canal de comunicação é equacionada pela relação entre os grandes operadores de telecomunicações e parceiros para provimento de acesso a rede internet, servidores de mídia, centrais de envio de mensagens de texto, entre outros serviços. Nessa parceria, a qualidade fim-a-fim percebida pelo assinante dependerá do correto gerenciamento dos indicadores subjetivos, determinados por métodos bem estabelecidos como o *Mean Opinion Score* (MOS) definido pelas recomendações da International Telecommunication Union (ITU-T), e o controle dos indicadores objetivos, representados pelas métricas de rede determinantes da Qualidade de Serviços (*Quality of Services* - QoS). O conhecimento desta relação é ponto-chave para o sucesso do serviço e permitirá a elaboração de soluções automatizadas para a sintonia dos parâmetros ao longo de redes que suportem gestão dinâmica de QoS.

1.1 JUSTIFICATIVA

Segundo Colcher (2005), o processo de convergência de tecnologias de telecomunicações e do próprio processamento de informações tem promovido profundas mudanças na sociedade contemporânea, permeando setores tão diversos como indústria, comércio, saúde e educação, alcançando em última instância a vida do cidadão comum. Este processo nos conduziu a um novo modelo social hoje conhecido como sociedade da informação. Neste novo paradigma, o acesso a informação torna-se cada vez mais veloz e ubíquo, encurtando as fronteiras e aproximando países e culturas. Em particular, a convergência entre as tecnologias empregadas em sistemas telefônicos e em redes computacionais trouxe facilidades para a implementação, operação e gerenciamento das plantas de telecomunicações ao permitir o uso de métricas, técnicas e processos já consagrados em redes de dados, sendo a grande mola impulsionadora para o crescimento acelerado do mercado de telecomunicações e serviços multimídia ao longo da última década.

Dentre os serviços multimídia convergentes, o VoIP surge como a aplicação-chave ao trazer consigo a redução dos custos tanto de operação quanto das tarifas ofertadas aos assinantes. Para Wallingford (2005), o VoIP é uma tecnologia tão revolucionária e ameaçadora para as grandes operadoras quanto o advento do computador pessoal foi frente ao mercado de grandes computadores dominado pela empresa IBM na década de oitenta. Tais operadoras de telefonia convencional passaram a demandar então de seus fornecedores de tecnologia soluções VoIP que rapidamente as colocassem no mercado (GONÇALVES, 2006). No entanto, garantir a qualidade fim-a-fim nas redes legadas de telefonia em uma rede de pacotes, na qual apenas parte da infra-estrutura pertence à operadora, tem se mostrado um desafio instigante, particularmente considerando que o protocolo IP, base das inter-redes atuais, não possui originalmente mecanismos para tratamento de QoS, o que em termos práticos representa uma incapacidade para a garantia de alocação de recursos da rede. Esta característica inerente às redes IP é uma barreira natural para aplicações interativas sensíveis a atrasos como transmissão de voz e vídeo, e que deve ser contornada.

Esforços têm sido alocados na busca e compreensão de métricas adequadas como o estudo de redes de acesso ou redes de última milha e seu suporte aos serviços VoIP (MEHMOOD; JADOON; SHEIKH, 2005), a análise de padrões para a medida da qualidade da fala (CARVALHO; MOTA; QUEIROZ, 2003), a análise do efeito do modelo de distribuição estatística referente ao atraso de pacotes (FUJIMOTO; ATA; MURATA, 2001), a análise do efeito da perda de pacotes e da duração dos quadros de áudio usados pelos diversos Codecs (DING; GOUBRAN, 2003) e a avaliação da importância do atraso, variação de atraso e perda de pacotes sobre os fluxos de mídia incluindo voz (BAI; ITO, 2006).

Questões pertinentes a implementação da camada de software tais como QoS, detecção de falhas, diagnósticos e desempenho também foram pesquisadas (SHERBURNE; FITZGERALD, 2004) e propostas de solução foram avaliadas como métodos para a troca dinâmica do codec utilizado em função da condição de rede (WÄLTERMANN et al., 2008).

No entanto, estas abordagens concentram atenção apenas no provisionamento dos clientes ou dos serviços da rede segundo os aspectos técnicos

da infraestrutura como, por exemplo, o estabelecimento de valores de limiar para as principais métricas de QoS ou a definição de políticas para tratamento de situações extremas como o descarte de pacotes baseado na marcação de prioridade. Esta visão técnica e objetiva da degradação da rede não necessariamente reflete a “sensação” de qualidade percebida pelo usuário final assinante dos serviços, e motiva a busca por novas abordagens que levem em conta a qualidade subjetiva do serviço como métrica principal para a medida da qualidade da fala em chamadas VoIP.

A observação do problema por este viés ainda possui poucos trabalhos relacionados como a identificação de correlações entre indicadores de qualidade subjetiva e parâmetros de rede e a proposição de um controle dinâmico de QoS em redes heterogêneas para uso em operadoras VoIP (GRAH; RADCLIFFE, 2008, 2009). Outra abordagem propõe o desenvolvimento de um *framework* composto por diversos modelos objetivos, utilizados para estimar a qualidade subjetiva baseado nas características físicas do sistema VoIP, aplicáveis não somente ao planejamento, mas também à avaliação de desempenho e ao gerenciamento da qualidade em chamadas VoIP (TAKAHASHI; KURASHIMA; YOSHINO, 2006). No Brasil, destacam-se os trabalhos relacionados ao monitoramento baseado em análise estatística para avaliação da qualidade da fala em VoIP (LIMA et al., 2005, 2007), a avaliação da qualidade de fala em redes wireless (PASSITO et al., 2005) e o estudo da influência de algoritmos de compensação de variação de atraso sobre a qualidade da fala (VALLE; AGUIAR; MOTA, 2009).

A compreensão da correlação entre a qualidade subjetiva percebida pelo usuário final e os diversos fatores tecnológicos envolvidos, suas métricas e como todo este sistema se relaciona, permitirá aos provedores de serviços estabelecerem meios de gerenciar a sua infraestrutura guiados pela experiência do assinante, o que deve se traduzir em um melhor aproveitamento dos recursos nas etapas de provisionamento, operação e manutenção, reduzindo a necessidade de superdimensionamento da planta e acelerando a resposta às situações de falta.

1.2 OBJETIVOS

O objetivo geral deste trabalho é, por meio da utilização de técnicas de estatística multivariada aplicadas ao reconhecimento de padrões, realizar a identificação de uma correlação entre a qualidade subjetiva da fala predita pelo algoritmo Modelo-E definido na recomendação ITU-T G.107 métricas de Qualidade de Serviços aplicáveis à infraestrutura de rede de operadoras VoIP.

Os objetivos específicos do trabalho são:

- Definição de método para a medida da qualidade subjetiva da fala baseado na recomendação ITU-T G.107;
- Definição das métricas de Qualidade de Serviços com maior aderência aos propósitos do trabalho e que permita meios ótimos para coleta (*data logging*) e posterior análise;
- Para representação do conceito, utilização de ferramenta computacional que permita verificar os requisitos propostos no projeto, demonstrando a viabilidade deste trabalho;
- Realização de estudo de caso com experimentos em ambiente controlado contendo estrutura que emule o ambiente operacional de provedores VoIP e seus assinantes, para validação da proposta.

1.3 METODOLOGIA

Para o desenvolvimento deste trabalho foram feitas revisões de literatura, especificação, customização e implementação de sistemas computacionais bem como a elaboração de um ambiente de testes VoIP para coleta de dados conforme detalhado a seguir:

- Revisão dos trabalhos relacionados e identificação das diversas abordagens empregadas;

- Elicitação das principais métricas de qualidade de serviços aplicáveis e avaliação de ferramentas para sua monitoração e coleta de dados;
- Formulação de hipóteses sobre o comportamento das variáveis;
- Estudo de ferramentas para a implementação do ambiente experimental emulado de operadora VoIP (GONÇALVES, 2006), definição da infraestrutura tecnológica necessária e de soluções para a integração e coleta de informações;
- Estudo de ferramenta para medição do MOS segundo o Modelo-E em chamadas de VoIP;
- Construção de processo para análise dos dados de operação por meio de estatística multivariada, incluindo tratamento de variáveis, gráficos e testes paramétricos de aceitação.
- Desenvolvimento de um protótipo para validação da arquitetura proposta por este trabalho.

Os serviços de análise e integração de experimentos foram baseados nas ferramentas definidas pelo Projeto Opensips (OPENSIPS, 2010) e o dimensionamento e provisionamento dos diversos serviços necessários ao estudo de caso foram realizados conforme descrito pela obra “Telefonia IP com SIP” (GONÇALVES, 2006).

A instrumentação necessária para a coleta dos dados de operação do ambiente experimental foi baseada em ferramentas gratuitas e de código aberto sempre que possível. *Sniffers* e analisadores de protocolo foram utilizados e artefatos de software foram desenvolvidos e/ou customizados para atender necessidades específicas, não cobertas pelas ferramentas anteriores.

Os testes de operação geraram um volume expressivo de dados de tal forma que estratégias de Data Warehouse (MACHADO, 2004) foram empregadas, o que permitiu a recuperação e análise de dados de diversas fontes, com caráter histórico, orientados a assuntos.

Para as correlações propostas entre MOS e QoS, assumiu-se para a pesquisa a hipótese que os dados são de caráter multivariado e, portanto técnicas

de estatística multivariada para o reconhecimento e descoberta de padrões foram também empregadas.

1.4 ORGANIZAÇÃO DO TRABALHO

A organização deste trabalho é descrita a seguir.

No Capítulo 2 é apresentada a fundamentação teórica de maior relevância ao trabalho. Questões como evolução do STFC, conceitos de qualidade, QoS e tecnologia VoIP são revisados. Os métodos para a medição da Qualidade da Fala são apresentados, com ênfase no Modelo-E, e as técnicas de estatística multivariada aplicáveis são abordadas. Por fim é feita uma revisão dos principais trabalhos correlatos e é indicado em qual contexto a proposta ora apresentada se encaixa.

No Capítulo 3 são apresentados os principais problemas existentes no serviço VoIP segundo a perspectiva do assinante e as formas atuais de mitigar tais situações. Em seguida, é apresentada a proposta de um método para estabelecer a correlação entre a métrica de QoS e o indicador fornecido pelo algoritmo Modelo-E. Suas limitações e potencialidades são discutidas e sua aplicabilidade é definida.

No Capítulo 4 é apresentado um estudo de caso baseado em um ambiente experimental que emula a estrutura de uma operadora VoIP, a rede de acesso e seus clientes. Meios e ferramentas para a coleta de dados gerados pelas chamadas de voz são introduzidos e limitações são discutidas. Este experimento foi utilizado como protótipo para a validação da proposta deste trabalho.

No Capítulo 5 são apresentadas as conclusões, ressaltando os principais resultados obtidos com o trabalho, possíveis melhorias e propostas para continuidade dessa pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta uma revisão sobre os tópicos envolvidos no desenvolvimento deste trabalho, focando nos conceitos sobre telefonia IP, qualidade da fala e estatística multivariada.

2.1 TRANSMISSÃO DA VOZ HUMANA

2.1.1 SERVIÇO TELEFÔNICO FIXO COMUTADO (STFC)

O telefone como patenteado por Alexander Graham Bell em 1876 era um dispositivo completo, vendido aos pares (TANEMBAUM, 2003). Alguém que desejasse realizar chamadas de voz entre residências deveria comprar um par de aparelhos e arcar também com a conexão entre eles, instalando a fiação necessária entre os imóveis. Se eventualmente houvesse a necessidade de falar também com n outras pessoas, ele deveria obrigatoriamente possuir um novo par de telefones e um fio interligando seus aparelhos à n residências. Em um ano após o início da venda do novo produto, as cidades norte-americanas estavam tomadas por emaranhados de fios e cabos, instalados de forma irregular sobre casas e árvores, compondo um cenário caótico e desorganizado. Ficou rapidamente evidente que o método ponto-a-ponto de conexão empregado não teria sucesso e mais uma vez Bell deu um novo passo à frente ao acrescentar uma central de comutação ao sistema e criar a *Bell Telephone Company*, em 1878, na cidade de New Haven, Connecticut, USA. Com o novo serviço oferecido por sua companhia, era agora necessária apenas uma conexão, um par de fios entre a casa do assinante do serviço e a central de comutação. Neste novo modelo, um assinante chamava a central de comutação e informava para quem desejava estabelecer uma chamada, sendo então o circuito fechado por meio de uma interligação entre as duas terminações dos assinantes, tarefa realizada por um operador humano. Em pouco tempo, as centrais da *Bell*

System se espalharam por diversas cidades dos Estados Unidos da América e uma nova necessidade passou a existir: a interconexão entre centrais de comutação para permitir que assinantes de cidades diferentes pudessem estabelecer chamadas. Já em 1890, três componentes do serviço telefônico fixo comutado eram claramente identificados: as centrais de comutação; as conexões entre as centrais de comutação e os assinantes; e as conexões de longa distância existentes entre as diversas centrais de comutação.

Muitas melhorias tecnológicas foram realizadas sobre estes três componentes, mas a idéia central, baseada no modelo idealizado por Graham Bell, permanece como base do serviço telefônico fixo comutado mundial a mais de 100 anos. As principais melhorias ocorrerem na central de comutação, hoje representada pelas modernas centrais telefônicas espalhadas ao longo da rede das diversas operadoras de telefonia fixa. Até 1950, a rede telefônica era totalmente baseada em tecnologia analógica. Um marco para as centrais telefônicas ocorreu na década de 60, quando por meio de um processo evolutivo, iniciou-se uma migração da era analógica para a era digital. Essa mudança ocorrida nos núcleos de processamento das centrais se deu por meio da troca de componentes eletromecânicos por processadores digitais, estendendo-se também a outros elementos funcionais das centrais, originando as centrais digitais com Controle por Programa Armazenado ou CPAs (COLCHER, 2005). Por fim, seguiu-se então a substituição da comutação de circuitos pela comutação de mensagens e finalmente desta pela comutação de pacotes, garantindo suporte IP no núcleo das operadoras. No entanto, estas alterações em sua grande maioria não se estendem ao usuário final, que ainda recebe na última milha um par de fios.

Traçando-se uma árvore representativa da evolução tecnológica sofrida pelo STFC, obtém-se um panorama como o exposto na figura 1, que cobre o período deste o nascimento da telefonia, passando pela era da comutação de circuitos – analógica e digital, estendendo-se até o limiar da comutação de pacotes.

Figura 1-Evolução tecnológica da telefonia.

Nesse cenário, a fronteira está em estender o suporte à comutação de pacotes já existente hoje no núcleo das operadoras, ao assinante, levando telefonia sobre protocolo IP até o usuário final. Esta meta é desejada pelo mercado, permitindo a convergência de tecnologias e garantindo facilidade na operação, baseada em uma só tecnologia em contraponto ao atual estado onde é necessária a coexistência de redes distintas para prover o serviço.

2.1.2 VOZ SOBRE IP

VoIP é o acrônimo de Voz sobre Protocolo IP (*Voice Over IP*) e denota o conjunto de tecnologias empregadas para permitir a transmissão de voz sobre redes de pacotes como a internet (COLCHER, 2005). Permite a transmissão *online* interativa de amostras de voz colocadas em pacotes de dados sobre redes IP que empregam *Transmission Control Protocol* (TCP), *Real-Time Transport Protocol* (RTP), *User Datagram Protocol* (UDP) e *Internet Protocol* (IP). Nos sistemas de VoIP, os sinais analógicos de voz são digitalizados e transmitidos como uma cadeia de pacotes sobre uma rede de dados. As redes IP permitem que cada pacote possa encontrar o caminho mais eficiente para chegar a um determinado destino num dado instante utilizando desta maneira os recursos de uma determinada rede de uma forma eficaz.

Nesta tecnologia, é possível a realização de chamadas telefônicas usando uma conexão banda larga de internet ao invés de uma linha telefônica regular (analógica ou digital). Alguns serviços usando VoIP podem somente permitir a chamada de outros usuários usando o mesmo serviço, mas outros podem permitir chamadas para qualquer número telefônico, incluindo chamadas locais, de longa distância, para telefones móveis e chamadas internacionais.

O conceito de VoIP tomou forma em meados da década de 1990, quando surgiu o primeiro software comercial para a transmissão de voz sobre IP, o *Internet Phone* da *VocalTec Communications*, que permitia a troca de pacotes IP com amostras de voz entre computadores pessoais. Ao longo da década a tecnologia e o interesse pela nova forma de transmitir voz sobre a internet evoluíram sobremaneira, suscitando a definição de padrões para o VoIP por parte das entidades normatizadoras do setor de telecomunicações.

Dois padrões principais disputam a hegemonia da telefonia IP: o H.323 da *International Telecommunications Union* (ITU), presente em muitos equipamentos e softwares VoIP; e o SIP, proposto pela *Internet Engineering Task Force* (IETF) que, apesar do curto tempo do processo de padronização, tem mobilizado muitos fabricantes da área da telefonia e dados, por causa da sua flexibilidade, aderência a padrões genuinamente Internet e de arquitetura aberta.

Para a operação do VoIP e o fornecimento dos serviços de voz e mídia segundo o padrão definido pelo IETF, algumas entidades são definidas conforme descrito na figura 2.

Figura 2 – Infraestrutura VoIP.

- UAC (*user agent client*) – cliente ou terminal que inicia a sinalização SIP;
- UAS (*user agent server*) – servidor que responde a sinalização SIP de um UAC;
- UA (*user agent*) – terminal de rede SIP (telefones SIP ou *gateways* para outras redes). Contém UAC e UAS;
- Servidor Proxy – Recebe pedidos de conexão de um UA e transfere-os para outro servidor proxy se a estação em particular não está sob sua administração;
- Servidor de Redirecionamento – Recebe pedidos de conexão e envia-os de volta ao emissor incluindo os dados de destino ao invés de enviá-los diretamente à parte chamada;
- Servidor de localização – recebe pedidos de registro de um UA e atualiza a base de dados de terminais com eles.

2.1.2.1 PROTOCOLO DE INICIAÇÃO DE SESSÃO

O protocolo SIP, está descrito principalmente em duas *Request for Comments* (RFC) do IETF: RC 2543 e RFC 3261. O SIP é um protocolo de controle da camada de aplicação (sinalização) empregado para a criação, modificação e encerramento de sessões com um ou mais participantes. Estas sessões incluem chamadas telefônicas, distribuição multimídia (áudio e vídeo) e conferências multimídia (GONÇALVES, 2006).

O protocolo SIP é bastante similar ao protocolo de transferência de hipertexto (HTTP) ou o protocolo de gerenciamento de rede simples (SNMP). As mensagens consistem em cabeçalhos e corpo da mensagem, sendo este último, para chamadas telefônicas, definido em *Session Description Protocol* (SDP). O SIP é baseado em texto, usa a codificação *8-bit Unicode Transformation Format* (UTF-8) e a porta 5060

tanto para UDP como TCP, podendo usar outros protocolos de transporte.

Neste protocolo, mensagens SIP de INVITE, usadas para criar sessões, transportam descrições da sessão que permitem aos participantes concordar sobre um conjunto de tipos de mídia compatíveis. O SIP faz uso dos servidores proxy para ajudar a rotear requisições para a localização corrente do usuário, autenticar e autorizar usuários para determinados serviços, implementar políticas de roteamento de chamadas e fornecer funcionalidades aos usuários. O SIP também fornece uma função de registro que permite aos usuários informar sua localização atual para uso dos servidores proxy.

O SIP oferece todas as características potencialmente comuns em telefonia tais como transferência de chamadas ou de mídia, conferência e retenção de chamada, entre outras, podendo ser considerado como o protocolo habilitador para a telefonia e os serviços de voz sobre IP.

Suas principais funcionalidades são:

- Tradução de nomes e localização de usuários: assegurando que as chamadas alcancem o destinatário onde quer que ele esteja localizado e transportando qualquer mapeamento das informações descritivas para as informações de localização, bem como assegurando que detalhes da natureza da chamada (sessão) sejam suportados;
- Negociação de funcionalidades: permite ao grupo envolvido em uma chamada (uma conferência, por exemplo) concordarem sobre as funcionalidades suportadas, reconhecendo que nem todos os participantes podem suportar o mesmo nível de funcionalidades;
- Gerenciamento dos participantes da chamada: durante uma chamada um participante pode trazer outros usuários para a chamada ou cancelar conexões de outros usuários. Em adição, usuários podem ser transferidos ou colocados em espera;
- Mudanças na funcionalidade da chamada: um usuário pode ser capaz de mudar as características da chamada durante o curso da mesma;
- Negociação de mídia: os mecanismos inerentes do SIP permitem a negociação da mídia usada em uma chamada, habilitando a seleção

dos codificadores/decodificadores (codecs) apropriados para o estabelecimento de uma chamada entre os vários dispositivos. Desta forma, dispositivos menos avançados podem participar de uma chamada, por exemplo.

2.1.2.2 PROTOCOLO DE TRANSPORTE DE TEMPO REAL

À primeira vista, a definição de “tempo real” associada a um protocolo que é comumente usado sobre redes IP, não determinísticas, parece errada. Mas é preciso recordar que existem dois níveis de criticidade associado à definição de “tempo real”: a mais difundida está associada ao *hard realtime*, significando que um determinado sistema ou aplicação devem ser projetados, implementados e testados de modo a garantir que as respostas ocorram dentro de um intervalo de tempo bem definido, para todas as situações em que o sistema possa se encontrar. No entanto, a finalidade e objetivo do Protocolo de Transporte de Tempo Real (RTP) são aderentes a definição do segundo nível de criticidade, o *soft realtime*, que define os sistemas que possuem respostas para uma instrução ou comando em um tempo não necessariamente exato, e sim dentro de um tempo médio de resposta. Nestes sistemas o tempo de resposta é importante, mas os prejuízos não são tão graves se ocasionalmente um prazo for perdido.

Dentro deste espírito, o RTP é um protocolo de redes utilizado em aplicações interativas sensíveis a atrasos como o VoIP. Ele define como deve ser feita a fragmentação do fluxo de dados-áudio, adicionando a cada fragmento informação de seqüência e de tempo de entrega. O controle é realizado pelo *Real Time Control Protocol* (RTCP). Ambos utilizam o UDP como protocolo de transporte, o qual não oferece qualquer garantia de que os pacotes serão entregues em um determinado intervalo. Os protocolos RTP/RTCP são definidos pela RFC 3550 do IETF (COLCHER, 2005).

Após o protocolo de sinalização ter identificado o endereço IP de destino e as capacidades de mídia suportadas no destino, o UAC usando RTP transfere mídia

diretamente para o UAS e vive-versa. Diferente do SIP, que usa um modelo requisição-resposta, RTP é um protocolo *one-way*, significando que em uma conversação entre UAC e UAS, devem ocorrer dois fluxos de mídia (GONÇALVES, 2006).

2.1.2.3 INFRAESTRUTURA DE UMA OPERADORA VOIP

Os serviços oferecidos por uma operadora VoIP não são considerados simples. Estes serviços envolvem um núcleo básico responsável pelo controle do registro, sinalização e mídia; estratégias para permitir o fluxo de sinalização e mídia entre assinantes por meio de redes com políticas de segurança e restrições diversas; e interconexão com redes de telefonia fixa. Associados ao núcleo encontram-se serviços auxiliares que não tem relação direta com a questão de transmissão de voz e mídia, mas são essenciais para o negócio da operadora: Autenticação, políticas de acesso e contabilização; portais de acesso para assinantes e monitoração dos serviços (GONÇALVES, 2006).

Todas estas funcionalidades costumam ser providas por pacotes de software distintos que interagem para garantir a operação. Em soluções livres e de código aberto, sua configuração e manutenção tende a ser um transtorno em função da diversidade de implementações, tecnologias e estruturas usadas por cada mantenedor de pacote. De maneira geral, a infraestrutura requerida para garantir as funcionalidades desejadas pelos assinantes deve possuir os blocos funcionais indicados na figura 3.

Figura 3 - A grande figura de um provedor VoIP

SIP PROXY

Responsável pelo registro dos usuários e por manter a base de dados de localização (mapeamento dos endereços SIP para endereços IP). Toda a sinalização e o encaminhamento de chamadas são gerenciados pelo SIP Proxy. No entanto, o SIP Proxy nunca gerencia a mídia (áudio ou vídeo em RTP). Todos os pacotes RTP são enviados diretamente entre os usuários e os *gateways* ou indiretamente por meio de servidores de mídia.

Existem dois tipos de servidores SIP Proxy:

- *Stateful Proxy Server* - mantém o estado das transações e permite dividir a chamada (*fork*) para múltiplos servidores na tentativa de localizar o usuário. Dessa maneira eles criam árvores de busca, possuem maior confiabilidade, capacidade de computar o gasto do cliente e utilizam protocolo TCP;
- *Stateless Proxy Server* - não armazenam o estado da transação, apenas enviando adiante as requisições e as respostas. Possuem maior velocidade, porem menos confiabilidade e incapacidade de computar gastos do cliente.

SERVIDOR DE MÍDIA

Em um ambiente VoIP, um Servidor de Mídia é a entidade que processa e gera fluxos de mídia (áudio, usualmente baseado em RTP). Ele tipicamente é controlado por meio de mensagens SIP enviadas por um *SoftSwitch*, SIP Proxy ou mesmo um Servidor de Aplicação.

Servidores de mídia tem como principais funções:

- Decodificação de tons DTMF;
- Gerenciamento de múltiplos fluxos de mídia em uma conferência;
- Execução de *prompts* e anúncios;
- Processamento de scripts VoiceXML;
- Reconhecimento de fala;
- *Text to Speech*;
- Gravação/reprodução de áudio;
- Suporte envio/recebimento de fax.

Estas funções são empregadas, juntamente com a intervenção de algum servidor de aplicação adicional, para a criação de serviços de mídia como:

- Gerenciamento de correio de voz;
- Salas de conferência;
- Unidade de Resposta Audível (URA);
- Filas de atendimento.

ENTRONCAMENTO SIP, TRAVESSIA DE NAT E PROXY DE MÍDIA

Entroncamento SIP (*SIP Trunking*) é o meio pelo qual um provedor VoIP se conecta a um *Internet Telephony Service Provider* (ITSP), de forma a ganhar acesso à rede pública de telefonia. Na figura 3, ele está presente no *Gateway STFC*. Desta forma, o Provedor VoIP consegue, além de permitir chamadas entre os ramais

virtuais de seu domínio, também realizar e receber chamadas para telefones convencionais, tanto locais, como de longa distância ou internacionais. Um ITSP recebe de um lado requisições de um Provedor VoIP ou outros tipos de assinantes por meio de mensagens SIP, e do outro lado despacha/recebe chamadas da rede pública telefônica.

Network Address Translation (NAT) ou mascaramento, definido pela IETF RFC 1631, é um processo para a tradução dos endereços IP e portas TCP/UDP de uma rede local privada para endereços válidos da rede pública. O NAT não é um protocolo nem um padrão. O NAT é apenas uma série de tarefas que um roteador (ou equipamento equivalente) deve realizar para converter endereços IPs entre redes distintas. Inicialmente o NAT foi criado com o objetivo de suportar o crescimento das redes de computadores na Internet e a escassez de endereços IP válidos. Neste contexto, o NAT buscou oferecer um método de criação de redes privativas, com endereços não roteáveis, que poderiam passar por um sistema que faria a tradução para um endereço válido e roteável, capaz de ser reconhecido em toda a Internet.

Um roteador NAT mapeia seu endereço interno para um endereço externo mantendo uma tabela de mapeamento. Algumas vezes, o NAT também é chamado de *Port Address Translation* (PAT). O PAT mantém um mapeamento de pares “ip:porta” permitindo que um único endereço externo seja usado para vários endereços internos.

Existem quatro tipos de NAT. Para um dado endereço interno, os três primeiros tipos de NAT mantêm um mapeamento do seu endereço interno que é independente do endereço de destino sendo visto. O quarto tipo de NAT irá alocar um novo mapeamento para cada endereço de destino de forma independente. A menos que haja uma tabela de mapeamento estático, o mapeamento aberto quando o primeiro pacote é enviado de um cliente por meio do NAT pode ser válido apenas por certo tempo, tipicamente alguns minutos, a menos que os pacotes continuem a ser enviados e recebidos em uma única porta IP.

Os tipos de NAT são:

- Cone Completo: O mapeamento é bem estabelecido e qualquer pacote vindo da Internet pública que queira alcançar um cliente atrás do NAT,

precisa apenas saber o esquema de mapeamento de forma a mandar pacotes para ele;

- Cone restrito: No caso de cone restrito, o par “ip:porta” externo só é aberto uma vez que o computador interno envie dados para um endereço IP de destino específico;
- Cone restrito por porta: Um NAT do tipo “restrito por porta” é quase idêntico ao “Cone restrito”, mas neste caso o NAT irá bloquear todos os pacotes a menos que o cliente tenha previamente enviado um pacote para o IP e porta que está enviando para o NAT;
- Simétrico: O NAT simétrico é diferente dos três primeiros. Um mapeamento do “ip:porta” para um NAT público é dependente do endereço de destino para o qual o pacote é enviado.

Por seu turno, um Proxy de Mídia, também referenciado como *Media Relay Server* ou *RTP Proxy*, tem como principal propósito tornar a comunicação entre os SIP UAs atrás de NAT possível. Existem muitos casos onde a comunicação direta fim-a-fim não é possível e o fluxo RTP tem que ser encaminhado por meio de outro *host*, que atua como elemento de ligação entre os UAs. Um RTP Proxy pode ser configurado e usado para esta tarefa.

SERVIÇOS AUXILIARES

Portal de provisionamento do usuário – tipicamente baseado em soluções web, este espaço permite a interação do assinante com a operadora, onde são disponibilizados ao mesmo um conjunto de serviços como criação de contas SIP, compra de crédito para contas pré-pagas, compra de DIDs, configuração de parâmetros da conta como senha do correio de voz, siga-me, não perturbe, entre outros.

Autenticação, políticas de acesso e contabilização – Tipicamente chamado de *Authentication, Authorization and Accounting* (AAA), esta funcionalidade está diretamente ligada à lógica de negócio da operadora, sendo responsável pela

administração dos diversos planos de discagem e políticas de tarifação ofertadas por essa ao mercado consumidor. Claramente, é dentro deste conjunto de serviços que a operadora gera o seu *billing*, transformando a transmissão de voz sobre rede de pacotes disponibilizada pelo núcleo do serviço em um negócio de fato.

Monitoração – Este serviço é de longe o mais deficitário dentro do conjunto necessário à operação VoIP. Baseada hoje em dia na monitoração da disponibilidade dos serviços ofertados pelos diversos pacotes de software (*watchdog*) e apenas no acompanhamento de métricas de desempenho de rede, não conseguindo responder dinamicamente à situações de degradação do serviço oferecido que ocasionam perda de qualidade fim-a-fim, traduzida pelo assinante como dificuldades para o estabelecimento, manutenção e encerramento de chamadas de voz. A falta de um método para a gerência da infraestrutura guiada pela experiência do assinante é hoje a principal deficiência desta área.

2.2 QUALIDADE NA TRANSMISSÃO DE VOZ

2.2.1 CONCEITO DE QUALIDADE

O conceito de qualidade é inerentemente subjetivo e está relacionado à avaliação pessoal de cada indivíduo. Esta avaliação sofre influências de elementos como sua cultura, os modelos mentais pré-existentes, o tipo de produto ou serviço prestado, além das necessidades e expectativas particulares. No entanto, independentemente destas influências, um conceito de qualidade é lugar comum para todas as culturas e povos: a qualidade é aquela característica do produto ou serviço que traz satisfação ao usuário. Logo, a qualidade está relacionada às características do produto que vão ao encontro das necessidades dos clientes e, dessa forma, proporcionam a satisfação em relação ao produto (JURAN, 1990; GARVIN, 1993).

A qualidade final entregue a um produto deve ser pautada em três princípios:

planejamento, controle e melhoria.

A melhoria da qualidade está relacionada com o processo e os meios para elevar o desempenho da qualidade.

O planejamento da qualidade engloba o desenvolvimento do produto e dos processos requeridos para atender as necessidades dos clientes. De forma geral, esta etapa envolve alguns passos:

- Estabelecimento de metas de qualidade;
- Identificação dos clientes;
- Determinação das necessidades dos clientes;
- Desenvolvimento das características do produto que respondam às necessidades dos clientes;
- Desenvolvimento dos processos capazes de produzir essas características de produto;
- Estabelecimento dos controles de processo.

Já o controle da qualidade engloba como passos:

- Avaliação do atual desempenho da qualidade;
- Comparação do desempenho atual e das metas de qualidade;
- Atuação sobre as divergências.

A falta de qualidade nos produtos e serviços também foi avaliada, resultando em uma lista de custos da qualidade:

- Custos de falhas: *Scrap*, re-trabalhos, ações corretivas, reivindicações de garantia, reclamações de cliente e perdas de processo;
- Custos de avaliação: Inspeção, auditorias de conformidade e investigações;
- Custos de prevenção: Treinamento, auditorias preventivas e processo de implementação de melhorias.

Operadoras de telefonia IP prestam serviços de transmissão de voz sobre redes de pacotes. Mesmo VoIP sendo um serviço e não um produto acabado, temos que os conceitos sobre qualidade definidos são aplicáveis neste contexto uma vez que o assinante, usuário e consumidor nesta relação, possui uma expectativa em relação ao serviço contratado e em sua utilização espera encontrar satisfação. A baixa qualidade do serviço prestado também gera custos de falhas, causando uma experiência insatisfatória para o assinante, baixa confiança na solução e em última instância perda de receita para a operadora em função do baixo uso ou mesmo desistência do serviço.

O planejamento da qualidade, neste contexto, deve incluir a definição de metas de qualidade, definidas como valores mínimos de MOS para cada classe de cliente. Deve incluir ainda a identificação dos clientes, aqui entendido como tipos de enlace (conexão via cabo, via rádio/satélite ou outros meios físicos) e tipos de serviço prestado (assinante residencial, corporativo, governo, entre outros), bem como de suas necessidades.

2.2.2 QUALIDADE DA FALA

A preocupação com a qualidade do serviço de voz prestado ao usuário final não é um fato recente e nasceu junto com os serviços de telefonia e as companhias telefônicas. Esta “qualidade” para o STFC leva em conta o aspecto qualitativo, focado na experiência do usuário e permanece em uso hoje em dia para suportar indiretamente o time de operações e rede no provisionamento da infraestrutura de comunicação e para promover melhorias no sistema.

Esta avaliação do nível de serviço prestado busca observar o que se convencionou chamar de “qualidade da fala”. Durante as sessões de avaliação, realizadas por pessoas em ambientes e condições controladas, busca-se avaliar aspectos como inteligibilidade, fidelidade do áudio e conforto.

Visando a uniformização dos procedimentos para a realização destas avaliações, entidades normalizadoras do mercado de telecomunicações envidaram esforços para a elaboração de recomendações a serem seguidas pelo setor, sendo

a recomendação ITU-T P.800 (1996) do International Telecommunication Union sediada nos EUA, aceita como padrão “de fato” pelo mercado. Esta recomendação fundamentalmente descreve métodos e procedimentos para a condução de avaliações subjetivas da qualidade da transmissão, onde estas avaliações são realizadas com grupos de pessoas. Dentre os vários métodos recomendados, destaca-se o *Absolute Category Rate* (ACR), do tipo somente audição (*listening-only*), onde grupos formados por pessoas retiradas do conjunto de usuários do sistema telefônico convencional passam a ouvir, por meio do enlace sob teste e em condições controladas, conjuntos de frases pré-gravadas no idioma nativo – entre 2 e 5 frases por conjunto, com duração de 2-3 segundos cada. É solicitado ao avaliador que pontue sua experiência em uma escala pré-definida de 5 níveis chamada MOS (*Mean Opinion Score*), para cada frase ouvida. Ao final dos testes, uma média destes *scores*, referente a cada grupo, é obtida, resultando em um índice numérico final de caráter qualitativo. O MOS propriamente pontua a qualidade da fala, mas outras duas escalas, MOS_{LE} (*Listening Effort*) e MOS_{LP} (*Loudness Preference*) também são usadas para identificar o esforço necessário para compreender o que foi ouvido e o nível do áudio percebido pelo avaliador, respectivamente. Como apresentado, imediatamente se observa que a determinação do MOS é uma tarefa dispendiosa e demorada. Os procedimentos ali definidos são eficazes na qualificação da experiência do usuário mas foram idealizados como ferramenta auxiliar no provisionamento da rede e dos serviços e não para a determinação da qualidade em tempo de execução de chamadas, permanecendo a lacuna no tocante a formas de identificação do nível de serviço prestado durante a operação. Para atender a esta demanda do setor, novo direcionamento foi dado visando concentrar esforços na busca de métodos automatizados para a predição da qualidade subjetiva de chamadas.

A Recomendação ITU-T G.107 (2000) responde a esta demanda e trás a especificação de um algoritmo chamado “modelo-E”, que permite determinar dinamicamente um fator de qualidade denominado “Fator R”, um índice numérico escalar, que representa uma predição da qualidade “Boca à Orelha” (*Mouth-to-Ear*). Este algoritmo realiza esta predição da qualidade percebida pelo usuário com grande acurácia em sistemas STFC, pois o modelo leva em conta os principais fatores de degradação da transmissão, fatores estes que refletem os efeitos das

redes e equipamentos baseados em *Time Domain Multiplexing* (TDM).

O fator R é determinado baseado na eq.(1) a seguir:

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

Na eq.(1), “Ro” representa a relação sinal/ruído básica, “Is” representa a degradação simultânea ocorrida no processo de captura da fala – variação na intensidade da voz e perdas de quantização, “Id” representa a participação do atraso de transmissão, “Ie” representa a degradação devido aos equipamentos – perdas de pacotes, variação de atraso e codecs. O termo “A” é um fator de compensação quando a tecnologia apresenta algum tipo de vantagem adicional para o acesso do usuário e “R” é a saída do algoritmo, podendo assumir um escalar entre 0 e 100.

Estratégia similar foi adotada pelo *European Telecommunications Standards Institute* (ETSI) em sua Especificação Técnica ETSI TS-101-329-5 (2003). O ETSI, sediado na França, possui um projeto iniciado no ano de 1997 denominado *Telecommunications and Internet Protocol Harmonization Over Networks* (TIPHON), que visa permitir a interoperabilidade entre STFC, redes digitais de serviços integrados de banda estreita (ISDN-N), *Global System for Mobile Communications* (GSM) e redes baseadas em IP. A TS-101-329-5 faz parte deste projeto, tendo como objetivo cobrir os aspectos da qualidade fim-a-fim. Neste documento, um novo algoritmo chamado “VQMon”, um modelo-E estendido, é proposto em seu Anexo “E”. A contribuição dada pelo ETSI no modelo contempla a modificação da parcela “Ie” contida na equação da recomendação ITU-T G.107, que passa a considerar um período de perdas isoladas intercalado a um período de perdas em rajadas, sendo a modelagem desta abordagem realizada segundo cadeias de Markov. Esta alteração torna a algoritmo mais aderente ao comportamento real desse tipo de fluxo de mídia sobre redes heterogêneas, melhorando consideravelmente a predição.

Embora os dois algoritmos para predição tenham sido elaborados por entidades distintas, o mesmo mantém meios para realizar um mapeamento entre os valores obtidos por seus respectivos fatores e a escala MOS de cinco níveis.

2.2.3 QUALIDADE DE SERVIÇOS

Da perspectiva do usuário final, qualidade de serviços (QoS) se refere à sua percepção da qualidade recebida da rede para um serviço ou aplicação específica como, por exemplo, voz, dados ou vídeo. Da perspectiva da rede, o termo QoS se refere à capacidade de fornecer a qualidade esperada pelas aplicações e em última instância, pelo usuário final (WALLINGFORD, 2005). Nesta pesquisa a visão de QoS na perspectiva da rede é preferida. Nesse contexto, essa capacidade de fornecer a qualidade esperada, por sua vez, pode ser dividida em dois tipos: primeiro, para fornecer QoS, uma rede deve ser capaz de diferenciar classes de tráfego para que o usuário final possa tratar uma ou mais classes de forma diferente de outras; segundo, uma vez que a rede diferencia as classes de tráfego, deve ser capaz de tratar essas classes também de forma distinta, fornecendo garantias de recursos e diferenciação dos serviços dentro da rede. De qualquer forma, a qualidade de serviços (QoS) nas redes IP é então um aspecto operacional fundamental para o desempenho fim-a-fim das novas aplicações e deve ser suportada pela rede, seus componentes e equipamentos utilizados.

Para esta finalidade, existem essencialmente duas formas de oferecer garantias de QoS. A primeira procura oferecer recursos suficientes para o pico esperado, com uma margem de segurança substancial. Esta abordagem é chamada de “IntServ” ou serviços integrados. É simples e eficaz, mas na prática é assumida como dispendiosa, e tende a ser ineficaz se o valor de pico aumentar além do previsto, pois reservar recursos gasta tempo. O segundo método é obrigar os elementos de rede a reservar os recursos, e apenas aceitar as reservas se os roteadores conseguirem servi-las com confiabilidade. Esta abordagem chama-se “DiffServ” ou serviços diferenciados.

Estas garantias oferecidas pela rede às aplicações e serviços são expressas em termos de um *Service Level Agreement* (SLA).

Na especificação dos SLAs são definidos parâmetros de qualidade de serviço sendo os mais comumente utilizados:

- Vazão (Banda): refere-se à capacidade de transmissão de informações de um determinado meio físico por unidade de tempo, normalmente

expressa em bits por segundo. Algumas aplicações necessitam de quantidade específica de banda para funcionar de acordo com o desempenho esperado. O projeto de rede deve dimensionar vazão suficiente para o funcionamento adequado das aplicações;

- **Atraso (Latência):** refere-se ao tempo necessário para que um pacote chegue ao destino. O atraso total é a soma dos atrasos fixos ou constantes, associados aos atrasos variáveis. Atrasos constantes são originados pelas características tecnológicas da solução adotada. Por exemplo, tempos gastos com codificação e serialização em uma aplicação VoIP compreendem atrasos constantes e dependem da velocidade dos enlaces, capacidade de processamento do codificador e complexidade do codec utilizado. Os atrasos variáveis são resultado de condições “ambientais” isto é, estão relacionados a um estado da rede. A causa mais importante desse tipo de atraso é o congestionamento. Em situações de contenção, os pacotes têm que aguardar nos *buffers* para que sejam despachados para a interface;
- **Variação de atraso (*jitter*):** o *jitter* pode ser definido como a diferença entre os atrasos na recepção de pacotes consecutivos. Esse tipo de atraso afeta os fluxos de tempo real e, se não é controlado, causa problemas de degradação do desempenho para aplicações sensíveis à alteração na cadência de recepção de fluxos de pacotes. Para reduzir o *jitter*, são utilizados *buffers* de de-*jitter* que recebem os pacotes em tempos variados e despacham paulatinamente, mantendo as diferenças aproximadamente constantes. O uso desses *buffers* provoca aumento do atraso total, o que requer cuidados no dimensionamento;
- **Taxa de perdas de pacotes:** mede a relação entre o número de pacotes entregues corretamente ao destino e o total de pacotes transmitidos. É expressa em porcentagem de pacotes perdidos e normalmente está relacionada à disponibilidade do serviço. Em uma rede altamente disponível, a taxa de perda de pacotes, em períodos sem congestionamento, deve ser zero. Em casos de congestionamento, os

mecanismos de QoS devem entrar em operação justamente para decidir quais pacotes, preferencialmente, devem ser descartados.

Outros parâmetros adicionais podem ser considerados como disponibilidade, Tempo Médio entre Falhas (*Mean Time Between Failures* – MTBF), Tempo Médio para Reparo (*Mean Time to Repair* – MTTR) e taxa de erros. Alguns desses parâmetros afetam a rede como um todo ou afetam todos os serviços demandados pelos usuários e não um serviço ou aplicação em particular. Uma vez que a rede transporta uma mistura de tráfegos com diferentes requisitos de desempenho, um determinado índice ou parâmetro importante para um serviço ou aplicação em particular pode não ser tão relevante para outros tipos de serviço ou aplicação e vice-versa. Os mecanismos de QoS implementados na rede devem considerar os diversos conflitos quanto aos requisitos de desempenho, buscando uma solução otimizada que possa atender a todos os serviços.

2.3 TÉCNICAS ESTATÍSTICAS

A estatística é uma parte da matemática aplicada que fornece métodos para coleta, organização, descrição, análise e interpretação de dados e para a utilização dos mesmos na tomada de decisões (FONSECA, 1996). No entanto, sua maior valia está em permitir se chegar a conclusões sobre uma população, partindo da observação de partes dessa população (amostras).

População – É um conjunto completo de elementos que apresentam um determinado conjunto de características ou parâmetros.

Amostra – É um subconjunto da população.

É de interesse para este trabalho as técnicas para medidas de posição que são as estatísticas que representam uma série de dados orientando-nos quanto à

posição da distribuição em relação ao eixo horizontal do gráfico da curva de frequência. As medidas de posições mais importantes são as medidas de tendência central ou promédias (verifica-se uma tendência dos dados observados a se agruparem em torno dos valores centrais). As medidas de tendência central mais utilizadas são: média aritmética, moda e mediana.

Média aritmética - É igual ao quociente entre a soma dos valores do conjunto e o número total dos valores e pode ser calculada conforme a eq.(2).

$$\bar{X} = \frac{\sum X_i}{n} \quad (2)$$

Onde “Xi” são os valores da variável e “n” o número de valores.

Moda (Mo) - É o valor que ocorre com maior frequência em uma série de valores.

Mediana (Md) - A mediana de um conjunto de valores, dispostos segundo uma ordem (crescente ou decrescente), é o valor situado de tal forma no conjunto que o separa em dois subconjuntos de mesmo número de elementos.

Variância - A variância é a soma dos quadrados dividida pelo número de observações do conjunto menos uma. A variância é representada por σ^2 , sendo calculada pela fórmula representada pela eq.(3).

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (3)$$

Desvio padrão - O desvio padrão é uma das mais utilizadas medidas de variação de um grupo de dados. A vantagem que apresenta sobre a variância é de permitir uma interpretação direta da variação do conjunto de dados, pois o desvio padrão é expresso na mesma unidade que a variável. É representado por “ σ ” e calculado usando-se a eq.(4).

$$\rho = \sqrt{\rho^2} \quad (4)$$

Outro conceito importante é a normalidade de uma variável. Uma distribuição de probabilidades é uma função matemática usada para descrever o padrão de variação de uma variável contínua. A função matemática que representa a distribuição normal envolve dois parâmetros - média e variância/desvio padrão. A curva que a descreve tem forma de “sino” e sua principal propriedade é a simetria em torno da média. A curva normal é também conhecida como curva de Gauss devido à suposição de que Gauss foi o primeiro a fazer uso de suas propriedades para aplicações práticas. Esta distribuição é importante porque os métodos paramétricos baseiam-se na suposição de que os dados observados na amostra são provenientes de uma população com distribuição de forma teórica conhecida. A suposição de que os dados seguem uma distribuição normal é assumida para a maioria dos métodos estatísticos e aceitá-la simplifica a abordagem do problema.

Denotamos $N(\mu, \sigma)$ à curva Normal com média μ e desvio padrão σ , onde a média refere-se ao centro da distribuição e o desvio padrão ao espalhamento (ou achatamento) da curva. A eq.(5) representa matematicamente a curva normal ou gaussiana.

$$f(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}. \quad (5)$$

2.3.1 DADOS MULTIVARIADOS

Um conjunto de dados multivariados é aquele onde cada caso ou observação de um fenômeno é descrito por um conjunto de várias variáveis, sendo representado tipicamente por um padrão $(n+1)$ -dimensional, onde $n > 2$ é o número de variáveis necessárias para descrever o fenômeno ou observação e a variável $(n+1)$ descreve a classe à qual este determinado padrão pertence.

Em Estatística, porém, faz-se diferenciação se um fenômeno é baseado em uma, duas ou muitas variáveis (mono, bi ou multivariado), pois as técnicas estatísticas utilizadas para cada um desses três casos variam muito e são tratadas separadamente.

2.3.2 ESTATÍSTICA MULTIVARIADA

A extração de informações úteis a partir de experimentos requer a análise de grande número de variáveis. De maneira geral, um pequeno número destas variáveis contém as informações mais relevantes, enquanto a maioria das variáveis adiciona pouco ou nada à interpretação dos resultados em termos práticos. A decisão sobre quais variáveis são importantes é feita, geralmente, com base na experiência do profissional.

No entanto, em Estatística, existem técnicas utilizadas para reconhecimento e descoberta de padrões em situações onde os fenômenos são descritos em função de uma grande variedade de dados, conhecidas como Análise Exploratória de Dados (AED), Estatística Exploratória ou Estatística Multivariada (CORRAR; PAULO; DIAS, 2007).

A estatística multivariada guarda relação com o conceito de Mineração de Dados (*Data Mining*), por meio da qual se busca a identificação de relacionamentos sistemáticos entre variáveis quando não existem expectativas *a priori* acerca da natureza destes relacionamentos. Nesta busca, muitas variáveis diferentes são

consideradas e comparadas. Isto é realizado utilizando-se uma grande variedade de técnicas e modelos matemáticos com o objetivo de se encontrar padrões nestes dados, dos quais se pode destacar:

- Análise de Correspondências (*Correspondence Analysis*);
- Análise de Agrupamentos Hierárquicos (*Cluster Analysis*);
- Análise Fatorial (*Factor Analysis*);
- Análise de Discriminantes (*Discriminant Function Analysis*);
- Escalonamento Multidimensional (*Multidimensional Scaling*);
- Análise Log-Linear (*Log-linear Analysis*);
- Correlação Canônica (*Canonical Correlation*);
- Regressão Parcialmente Linear e Não-Linear (*Stepwise Linear and Nonlinear Regression*);
- Análise de Séries Temporais (*Time Series Analysis*);
- Árvores de Classificação/Decisão (*Classification Trees*).

2.3.3 ANÁLISE DE AGRUPAMENTO HIERÁRQUICO (HCA)

A análise de agrupamento hierárquico consiste no tratamento matemático de cada amostra como um ponto no espaço multidimensional descrito pelas variáveis escolhidas (MOITA NETO; MOITA, 1998). Nesta técnica, no entanto, também é possível tratar cada variável em particular como um ponto neste espaço multidimensional descrito pelas amostras. Desta forma, o uso da análise de agrupamento hierárquico permite reunir amostras ou variáveis de acordo com o interesse em cada situação. Quando uma determinada amostra é tomada como um ponto no espaço das variáveis, é possível calcular a distância deste ponto a todos os outros pontos, constituindo-se assim uma matriz que descreve a proximidade entre todas as amostras estudadas. Baseada nesta matriz de proximidade entre as amostras se constrói um diagrama de similaridade denominado dendrograma.

Dentre as diversas formas para o cálculo da distância entre dois pontos, a mais utilizada é a distância euclidiana. Nesta, para duas variáveis em particular, aplica-se o Teorema de Pitágoras ($a^2 = b^2 + c^2$) sendo a distância entre os pontos a hipotenusa (a), que por sua vez é igual à raiz quadrada da soma dos quadrados dos comprimentos dos catetos (b) e (c).

Tomando um conjunto de pontos e suas respectivas n-ésimas distâncias, pode-se aglomerar matematicamente estes pontos no espaço multidimensional para formar os agrupamentos hierárquicos que usam as informações da matriz de proximidade para criar um dendrograma de similaridade.

A interpretação de um dendrograma de similaridade entre amostras fundamenta-se na intuição, pois duas amostras próximas devem ter também valores semelhantes para as variáveis medidas sendo, portanto, próximas matematicamente no espaço multidimensional. Donde se conclui que quanto maior a proximidade entre as medidas relativas às amostras, maior a similaridade entre elas.

O dendrograma hierarquiza esta similaridade. Quando se constrói dendrogramas de variáveis, a similaridade entre estas aponta forte correlação.

Não obstante, a aplicação da análise de agrupamento hierárquico requer cuidados. Quando temos variáveis com ordens de grandeza diferentes, deve-se preceder a aplicação do método por um tratamento prévio dos dados. O pré-tratamento mais comumente empregado é a transformação Z, que transforma as medidas de cada variável de tal modo que o conjunto de dados tenha média 0 (zero) e variância 1 (um). Esta passa a ser a variável normal padronizada "Z". A finalidade deste procedimento é igualar a importância estatística de todas as variáveis utilizadas.

De forma resumida, a técnica de agrupamento hierárquico interliga as amostras por suas associações, produzindo um dendrograma onde as amostras semelhantes, segundo as variáveis escolhidas, são agrupadas entre si. A suposição básica de sua interpretação é esta: quanto menor a distância entre os pontos, maior a semelhança entre as amostras. Os dendrogramas são especialmente úteis na visualização de semelhanças entre amostras ou objetos representados por pontos em espaço com dimensão maior do que três, onde a representação de gráficos convencionais não é possível.

2.3.4 VARIÁVEL NORMAL PADRONIZADA

É obtida por meio de uma transformação linear da variável normal “X” conforme eq.(6), obtendo-se assim uma nova variável “Z” que pertence a escala relativa de valores na qual a média é tomada como ponto de referência e o desvio padrão como medida do afastamento da média.

O conjunto de valores transformados passa a apresentar média 0 (zero) e desvio padrão igual a 1 (um), ou seja $N(0, 1)$.

$$Z = \frac{X - \mu}{\sigma} \quad (06)$$

Onde:

Z = Valor da variável normal padronizada Z

X = Valor de x

μ = média de x

σ = desvio padrão de x

3 PROCESSO PROPOSTO

Neste capítulo são discutidas as deficiências atuais para a transmissão de voz sobre redes de pacotes para então apresentar os meios atualmente empregados para mitigar os problemas oriundos de tais deficiências e sua dissociação com a experiência do usuário final. Posteriormente é apresentada a proposição de um processo para estabelecer a correlação entre as métricas utilizadas atualmente, centradas em QoS, e indicadores associados a qualidade da fala percebida subjetivamente pelo assinante. Suas limitações e potencialidades são discutidas e sua aplicabilidade é definida.

3.1 VISÃO GERAL

Sob a ótica do assinante, o principal atrativo dos serviços de voz sobre IP tem sido o custo baixo para a realização de chamadas a longa distância e não a qualidade da fala. Graças aos entroncamentos entre as operadoras VoIP e as operadoras do STFC, é possível para um assinante de serviços VoIP realizar chamadas telefônicas entre seu *softphone*, telefone IP o ATA e telefones fixos ou móveis de outros estados ou mesmo países a custo de centavos por minuto (MEGGELEN; SMITH; MADSEN, 2005). Neste cenário, o assinante tem tolerado a instabilidade na qualidade da chamada em razão do custo-benefício.

Tanto operadoras exclusivamente VoIP quanto operadoras do sistema legado adentrando neste mercado sabem que a comparação com a qualidade de chamadas do STFC é automaticamente feita pelo assinante, que em última instância é quem detém o poder de decisão sobre o serviço.

Sabe-se hoje que a qualidade da fala percebida pelo usuário sofre influência de três parâmetros principais, relacionados ao QoS: o atraso total fim-a-fim, a variação de atraso fim-a-fim e a perda de pacotes ao longo do enlace. (VIANNA et al., 2006). Tais parâmetros são relacionados com a infraestrutura de rede de dados e

sua tratativa de forma a tentar estabilizar a transmissão tem utilizado técnicas próprias para redes de computadores. No entanto, o tráfego de voz sobre rede de pacotes tem caráter urgente, sendo definido como de “tempo real” no sentido de ser sensível a variações da condição de rede. A monitoração destas métricas e seu controle, por si, não são suficientes para garantir a qualidade da fala, pois não são guiadas pela experiência do usuário (LIPOVAC, 2009).

O processo proposto neste trabalho objetiva criar meios para que o gerenciamento dos parâmetros de rede importantes para a transmissão de voz sobre redes de pacotes sejam guiados diretamente pela experiência do usuário definida pela sua percepção subjetiva da qualidade da fala.

3.1.1 PRINCIPAIS PROBLEMAS EM CHAMADAS VOIP

Do ponto de vista do assinante, os principais problemas relativos à qualidade da fala estão relacionados à “metalização” da voz (voz robotizada) e cortes na transmissão (picotes). A metalização da voz quase sempre ocorre devido a problemas de sub-dimensionamento de largura de banda disponibilizada por algum dos elementos componentes da nuvem entre os terminais (BRUSCATO et al., 2004). A interrupção momentânea no fluxo de voz, por sua vez, está relacionada com a incapacidade do codec utilizado de reconstruir o áudio em função da perda de pacotes sofrida e também com o atraso total no enlace.

Estes comportamentos são naturais e ocorrem devido às características da transmissão da tecnologia VoIP. A mídia trafega pela rede de dados por meio de pacotes RTP sobre protocolo UDP. Pacotes RTP são muito pequenos e possuem carga útil de 20 a 150 bytes, mais um cabeçalho RTP/UDP/IP de 40 bytes. A transmissão de pacotes RTP não é eficiente, pois existe uma alta relação entre o tamanho da carga útil e o tamanho do cabeçalho. Adicionalmente, quando estes pacotes RTP trafegam por linhas com tráfego pesado, dividindo banda com grandes pacotes de dados (1000-1500 bytes ou maiores), eles precisam ser enfileirados em *buffers* de roteadores ao longo do percurso, tendo que esperar sua vez para acessar

o canal físico de dados por um tempo variável. Este atraso derivado desta e de outras questões não deve ser superior a 250 milissegundos, idealmente (ITU-T G.114, 2003). Atrasos entre 100 e 250 milissegundos fornecem uma interatividade aceitável. Acima desse valor é possível realizar a comunicação com um pouco de treino mas a conversa não é natural.

Outra questão importante é o fato de usar-se UDP como protocolo de transmissão. Roteadores são parametrizados para avaliar fluxos não responsivos como o UDP dando-lhe baixa prioridade, o que aumenta a probabilidade de pacotes deste tipo de tráfego serem descartados (CISCO, 2010).

3.1.2 SOLUÇÃO DE PROBLEMAS EM SISTEMAS VOIP

A maneira mais simples de garantir a qualidade do serviço para o VoIP é manter o enlace com a maior largura de banda possível de forma a atender os requisitos de aplicação (STEINBERG; SOUTHWICK, 2010). Claro que está também é a forma mais dispendiosa de resolver o problema o que é intolerável do ponto de vista de uma corporação, pois envolve superdimensionar a infraestrutura. Outros mecanismos foram desenvolvidos para evitar congestionamentos por meio do controle e priorização de tráfego ou reserva de recursos. A implementação desses mecanismos pode ocorrer na rede de pacotes ou nos equipamentos terminais e procura minimizar:

- Filas em congestionamento;
- Atrasos fim-a-fim ou latência;
- Variação do atraso fim-a-fim;
- Perda de pacotes ou seqüência.

Destes, as filas são de longe o principal entrave, pois impactam diretamente sobre a variação do atraso. Por isso foram desenvolvidos diversos tratamentos de filas por meio de algum algoritmo de priorização ou classificação de tráfego nos

roteadores. Além dos tratamentos de fila, o IETF desenvolveu algumas soluções como os serviços diferenciados (DiffServ) e a de serviços integrados (IntServ). Outra tecnologia que também foi definida pela IETF é o *Multi Protocol Label Switching* (MPLS).

Destas várias técnicas, destacam-se:

- *First in, first out* (FIFO): os pacotes que chegam são colocados na fila de transmissão, sem prioridade e sem classificação. A transmissão ocorre na ordem em que foram recebidos. Se não houver espaço suficiente na memória, a política de descarte de pacotes determinará se o pacote será descartado ou se outros pacotes serão eliminados da fila;
- *Weighted Fair Queuing* (WFQ): é um método de escalonamento em que cada fila é tratada de forma imparcial por varreduras cíclicas. Cada fila recebe uma fatia imparcial da largura de banda, cada fluxo possui sua própria fila, e não pode afetar outros fluxos. A técnica FQ ainda oferece mais largura de banda para fontes que usam pacotes grandes, porque o critério de imparcialidade é baseado na taxa de pacotes;
- *Priority Queuing* (PQ): os pacotes são classificados na entrada em níveis de prioridade. Cada classificação forma uma fila e pode ser atribuída a um tipo de tráfego ou a uma interface. A transmissão dos pacotes é feita seguindo essa prioridade. A transmissão começa pelos pacotes de maior prioridade até que todos estejam transmitidos. Só então começa a transmissão da próxima classe. Por isso, em enlaces de menor velocidade podem ocorrer atrasos e aumento de *jitter* ou até mesmo o descarte de todos os pacotes da classe de menor nível de prioridade;
- Fragmentação e *Interleaving*: a fragmentação e *interleaving*, ou LFI (*Link Fragmentation and Interleaving*) é uma técnica bastante usada em conexões de baixa velocidade. O objetivo é evitar que o tempo de serialização de grandes datagramas cause uma indesejável variação no atraso do tráfego sensível ao *jitter*. A fragmentação divide os pacotes de dados que excederem determinado tamanho em pacotes menores. Os pacotes de voz são intercalados entre os fragmentos do pacote de dados. Quando um pacote de dados é fragmentado, o transmissor insere um cabeçalho em cada fragmento

o que permite ao receptor remontar o pacote original. Esse cabeçalho representa um acréscimo de dados que somados ao tempo de processamento causam certo atraso. Por isso, é importante estabelecer um tamanho adequado do fragmento para que a fragmentação não se torne ineficiente.

3.1.3 QUALIDADE DA FALA COMO CRITÉRIO PARA GERENCIAMENTO DA REDE

O gerenciamento tecnicista das métricas de QoS e a mitigação dos diversos efeitos causados pelos problemas de infraestrutura e pelo próprio comportamento das redes IP tentam resolver a questão de forma abrangente, englobando todas as prováveis causas conhecidas, tratando do efeito e não da causa. Esta abordagem é mais onerosa, pois precisa contemplar solução para todas as situações, atuando tanto sobre a rede de dados em si como sobre equipamentos e softwares associados, de forma a garantir uma região de conforto.

Outras abordagens para monitoração já conseguem calcular em tempo de execução o fator R e o MOS e utilizar tais parâmetros para a sintonia de equipamentos terminais e alguns elementos-chave de rede (CONWAY, 2005). Esta direção, no entanto, apenas reconhece a existência do problema ao indicar valores de qualidade de fala fora dos limiares definidos, que são utilizados para orientar na modificação de diversos parâmetros de rede, mas não apontam a causa.

Portanto, a descoberta de uma relação causa-efeito entre a percepção do usuário indicada pelo MOS e as métricas de QoS associadas à infraestrutura, em uma perspectiva *top-down*, permitirá atuar diretamente sobre a causa da degradação, otimizando recursos e esforços.

3.2 CORRELAÇÃO ENTRE MÉTRICAS DE QOS E QUALIDADE DA FALA

No início da pesquisa foram definidas hipóteses para o comportamento da população e premissas para os ensaios experimentais no ambiente de testes, de forma a limitar o escopo do estudo.

As hipóteses levantadas foram:

- Hipótese 1: Os dados que compõem a população são multivariados uma vez que o evento sob investigação, a qualidade da fala, precisa de um conjunto de três ou mais variáveis para ser adequadamente caracterizado;
- Hipótese 2: O conjunto de dados que compõem a população é bem comportado e obedece a distribuição estatística normal, pois esta é a condição para o uso da maioria dos métodos estatísticos para análise de correlações.

As premissas para os ensaios experimentais foram:

- SIP como protocolo para inicialização, controle e encerramento de chamadas VoIP;
- RTP/RTCP como protocolo para transporte e gerenciamento do tráfego de mídia de tempo real;
- UDP como protocolo da camada de transporte;
- RTP Proxy exclusivamente como solução para travessia de NAT do fluxo de mídia, em detrimento de STUN;
- Codec G.711 (u-law ou a-law) como padrão nos terminais;
- Utilização de *softphones* para o provisionamento inicial;
- Utilização de geradores de tráfego SIP/RTP como ferramenta para artificialmente gerar carga de chamadas sobre o sistema em laboratório.

Dado o panorama traçado na seção 3.1 e os balizadores definidos anteriormente, a pesquisa objetivou identificar as correlações entre os indicadores de qualidade da fala e os indicadores de qualidade de serviços, estabelecendo um método sistematizado para sua obtenção, conforme apresentado na figura 4.

Figura 4 – Representação do método para estabelecimento de correlação.

COLETA DE DADOS

A pesquisa utiliza dados numéricos obtidos a partir de relatórios de operação (*logs*) referente às chamadas entre terminais, armazenados na forma de registros de dados (*data records*).

Estes registros contém, para cada chamada observada:

- As leituras para as métricas de qualidade de serviço impactantes sobre o problema – atraso, variação de atraso e perda de pacotes;
- As leituras para os indicadores de qualidade de fala calculados segundo o modelo-E definido pela G.107 – Fator R e o correspondente na escala MOS.

Estes dados têm caráter qualitativo para o evento sob observação, pois representam as métricas com impacto sobre o desempenho do serviço.

TRANSFORMAÇÃO Z E ANÁLISE EXPLORATÓRIA DE DADOS (AED)

As variáveis de interesse - MOS, atraso, variação de atraso e perda de pacotes - possuem ordens de grandeza diferentes. Neste caso particular as variáveis relacionadas a tempo são expressas em milissegundos e seus valores habituais variam entre 0 e 1000 milissegundos para o atraso e 0 a 500 milissegundos para a variação de atraso. A perda de pacotes é expressa de forma percentual, com valores possíveis entre 0% e 100%. O fator R pode assumir valores entre 0 e 100 e o MOS valores entre 1 e 5.

Quando os dados da amostra estão em escalas diferentes, deve-se proceder a um tratamento prévio por meio da transformação de dados para evitar que variáveis com valores numéricos mais altos sejam mais importantes que as variáveis com valores numéricos mais baixos durante a avaliação.

As transformações diretas dos dados mais comumente utilizadas são logarítmica, raiz quadrada dos dados, raiz quadrada dos (dados + 1, ou mais 1/2), raiz cúbica dos dados, transformação angular, transformação hiperbólica de primeiro

grau (ou o inverso dos dados), hiperbólica de segundo grau, transformação percentual, e a transformação em valores de z.

Utilizamos a transformação Z por ser mais simples e indicada para conjuntos de dados que já possuem comportamento normal, conforme hipótese 2 formulada. Esta transformação modifica as medidas de cada variável de tal modo que o conjunto de dados tenha média 0 (zero) e variância 1 (um). A finalidade deste procedimento é equalizar a importância estatística de todas as variáveis utilizadas, forçando-as a assumirem um comportamento normal padronizado.

Vencida a etapa de padronização, os dados relativos às variáveis já transformados foram submetidos às técnicas de análise exploratória de dados (AED). O que se desejou neste ponto é realizar um exame da relação entre variáveis.

Inicialmente uma análise da amostra do conjunto de dados de cada variável é realizada por meio de histogramas e gráficos de dispersão para verificar-se o comportamento da distribuição estatística.

Em seguida, foi realizada a busca por relações entre duas ou mais variáveis, isto é, saber se as alterações sofridas por uma das variáveis são acompanhadas por alterações nas outras. Em particular se busca determinar a existência da correlação e o grau de relação entre as variáveis.

Fez-se a medida de correlação por meio da determinação do coeficiente de correlação linear de Pearson (RODRIGUES, 2008), sobre uma amostra dos dados da população. Foram calculados coeficientes para todos os pares de variáveis possíveis, sendo seus resultados tabulados.

O estimador de correlação r_{xy} , é definido pela eq.(7):

$$R_{xy} = \frac{Cov(x,y)}{\sqrt{\rho_x^2 \cdot \rho_y^2}} \quad (7)$$

Onde:

$Cov(x,y)$ – Covariância

ρ_x, ρ_y – Variância

Este coeficiente de correlação r_{xy} é um número puro que varia de (-1) a $(+1)$ e mede o grau da correlação e a direção dessa correlação (se positiva ou negativa). O valor 0 (zero) significa que não há relação linear, o valor 1 (um) indica uma relação linear perfeita e o valor -1 (menos um) também indica uma relação linear perfeita mas inversa. Neste caso, em termos práticos significa que quando o valor numérico de uma das variáveis aumenta o valor numérico da outra diminui.

Quanto mais próximo estiver de (-1) ou $(+1)$, mais forte é a associação linear entre as duas variáveis. A interpretação do resultado para diversos intervalos é mostrada na tabela 1.

Tabela 1 – Interpretação do coeficiente R_{xy}

Intervalo	Interpretação
$r_{xy} = -1$	Correlação perfeita negativa
$-1 < r_{xy} < 0$	Correlação negativa
$r_{xy} = 0$	Correlação nula
$0 < r_{xy} < 1$	Correlação positiva
$r_{xy} = 1$	Correlação perfeita positiva
$0,2 < r_{xy} < 0,4$	Correlação fraca*
$0,4 < r_{xy} < 0,7$	Correlação moderada*
$0,7 < r_{xy} < 0,9$	Correlação forte*

*Possui o mesmo significado para os casos positivo e negativo.

Os dados de toda a população, relativo às variáveis em estudo, foram também plotados em um diagrama de dispersão, aos pares, e seu comportamento observado. Assim como o coeficiente de correlação de Pearson, os diagramas de dispersão, quando plotados para duas ou mais variáveis, permitem observar seus relacionamentos.

A maneira como os dados estão distribuídos na forma de pontos no plano possui relação com coeficiente de correlação e seu sinal, caracterizando também graficamente o tipo de correlação, conforme apresentado na figura 5.

Figura 5 – Representação gráfica do comportamento de R_{xy} .

Ao final, um terceiro procedimento ainda foi adotado para a identificação de correlações: um método de análise de agrupamento hierárquico. Este método será aplicado sobre uma amostra da população e irá calcular as distâncias entre os conjuntos de dados de cada variável, estabelecendo índices de similaridade.

Existem vários métodos baseados em diversos índices de similaridade, cada qual com aplicabilidade restrita em função do tipo de evento investigado. São exemplos os métodos e/ou índices de Sorensen, Morisita-Horn e Jaccard.

Para a análise realizada, utilizou-se o método/índice de Jaccard. Também chamado de coeficiente de similaridade de Jaccard, mede a similaridade entre conjuntos e é formalmente definido pelo tamanho da intersecção dividido pelo tamanho da união dos conjuntos de amostras A e B (KREBS, 1989), conforme eq.(8).

$$J_{AB} = \frac{|A \cap B|}{|A \cup B|} \quad (8)$$

Para uma população finita de indivíduos, o índice de Jaccard é calculado conforme eq.(9).

$$Ij = \frac{a_{xy}}{a_{xy}+b_{xz}+c_{yz}} \quad (9)$$

Considerando x e y como as variáveis para os quais se deseja calcular o índice e z como qualquer outra variável diferente de x e y, temos que (a) é o conjunto de vezes em que (x, y) apresentam alguma relação (proximidade), (b) é o conjunto de vezes em que (x, z) apresentam alguma relação e (c) o conjunto de vezes em que (y, z) apresentam alguma relação.

O valor calculado pela equação varia entre 0 (zero) e 1 (um). O valor zero indica total similaridade entre elementos enquanto um indica ausência completa ou dissimilaridade. O coeficiente de Jaccard aplica o mesmo peso a todas as variáveis e é uma medida da assimetria sobre as informações armazenadas.

Os índices obtidos para todos os conjuntos de relações foram então plotados na forma de um dendograma. Esta ferramenta permite visualizar graficamente a proximidade entre as amostras.

Ao final da execução destes três processos de AED o conjunto de resultados gerados – tabelas e gráficos - foi avaliado e então foram elicitadas as relações entre variáveis com maior similaridade.

As relações obtidas, no entanto, ainda não podem ser consideradas válidas, pois o coeficiente de correlação R_{xy} é apenas uma estimativa do coeficiente de correlação populacional ρ_{xy} , uma vez que seu valor é calculado com base em uma amostra finita composta por um conjunto de “n” pares de dados. Estes pontos da amostra podem apresentar uma correlação e ainda assim ocorrer da população não seguir este comportamento, o que representaria um problema de inferência. Este impasse foi resolvido aplicando-se um teste paramétrico de hipótese para verificar se o valor de R_{xy} é coerente com o tamanho da amostra n, a um nível de significância α , demonstrando que realmente existe correlação linear entre as

variáveis.

Os testes paramétricos podem ser aplicados sobre amostras extraídas de populações com distribuição normal (hipótese 2) e variâncias iguais ou muito próximas (população após transformação Z). Estes testes são robustos, podendo ser aplicados mesmo quando ocorram pequenos desvios de normalidade ou da variância entre as amostras.

Como as amostras são pareadas e independentes, foi escolhido o teste paramétrico “t-student” ou “t”. Na pesquisa os dados são tratados de forma diferente, ou seja, possuem o mesmo tamanho de amostra n, mas não pertencem ao mesmo domínio nem há comparação entre os estados antes e depois. Nestas condições o teste t é aplicável, salientando que a variável em análise tem que apresentar os dados em distribuição normal ou aproximadamente normal (novamente hipótese 2).

O resultado obtido pelo teste foi então comparado com valores de referência, contidos na tabela 2, para então realizar a extrapolação do resultado para toda a população com nível de significância α (GOMES, 1990). A inferência é realizada com base em novas hipóteses acerca da validade dos resultados:

- $h_0: \rho_{xy} = 0$ (Não existe correlação entre X e Y na população)
- $h_1: \rho_{xy} \neq 0$ (Existe correlação entre X e Y na população)

A tabela 2 traz os valores de significância de 5% ($\alpha=0,05$) ou 1% ($\alpha=0,01$), obtidos de acordo com o grau de liberdade do teste. O valor t-student com $(n - 2)$ graus de liberdade é calculado pela eq.(10):

$$t_c = \frac{r_{xy} \cdot \sqrt{n - 2}}{\sqrt{1 - r_{xy}^2}} \quad (10)$$

Tabela 2- Valores de t-student em níveis de 5% e 1% ($\alpha=0,05$ a $0,01$) de probabilidade

Grau de Liberdade	5% ($\alpha = 0,05$)	1% ($\alpha = 0,01$)
1	12,71	63,66
2	4,30	9,92
3	3,18	5,84
4	2,78	4,60
5	2,57	4,03
6	2,45	3,71
7	2,36	3,50
8	2,31	3,36
9	2,26	3,25
10	2,23	3,17
50	2,01	2,68
100	1,98	2,63
Inf.	1,96	2,58

Fonte: GOMES (1990)

Ao final, as correlações que demonstraram significância foram então rotuladas como correlações válidas e assumido que representam a relação MOS – QoS.

3.3 APLICABILIDADE E LIMITAÇÕES

Embora sistemas baseados em protocolo SIP/RTP respondam hoje em dia pela absoluta maioria dos sistemas VoIP *opensource* em produção, não se pode ignorar a fatia alocada para sistemas H.323 e demais soluções proprietárias de grandes fabricantes como CISCO SYSTEMS e outras. Os resultados obtidos são portanto válidos para as condições e limitações definidas neste trabalho, cabendo avaliações sobre como adaptar o processo para realidades distintas ou mesmo sua aplicabilidade em outras condições de operação.

4 ESTUDO DE CASO

A implementação do processo proposto neste trabalho na forma de um estudo de caso foi realizada como parte das atividades de um projeto custeado pelo Fundo para o Desenvolvimento Tecnológico das Telecomunicações – FUNTTEL.

O projeto visava o desenvolvimento de um produto de software baseado na integração de soluções VoIP de código aberto associadas ao desenvolvimento de aplicações proprietárias, que forneceria serviços de roteamento de chamadas em uma rede de telefonia IP (*Proxy/Redirect*) e serviços de SIP B2BUA (*Back-To-Back User Agent*), baseados nos protocolos SIP e em RADIUS e/ou DIAMETER para a contabilização. Este ambiente, disponível por meio do projeto, foi utilizado para a execução do estudo de caso.

4.1 AMBIENTE EXPERIMENTAL

O ambiente utilizado para a realização dos experimentos, apresentado na figura 6, foi planejado de tal forma a operar conforme as necessidades do mercado corporativo, provendo um conjunto de serviços básicos voltados às necessidades típicas dessa clientela. Cenários que poderiam ser atendidos pelo ambiente experimental incluem provedores de serviços de telecomunicações e corporações que estejam realizando testes para a substituição gradual de equipamentos PBX TDM para uma solução que forneça mais recursos com menor custo.

Este ambiente contém quatro áreas funcionais distintas:

- a) PABX-IP;
- b) Equipamentos terminais;
- c) Estrutura para coleta e armazenamento dos dados de operação (*data log*);
- d) Conectividade.

Figura 6 – Ambiente experimental.

A infraestrutura do provedor VoIP propriamente dita é representada pelo “PABX-IP” (figura 7). O PABX-IP possui diversos blocos funcionais, necessários a execução de suas diversas funcionalidades. No entanto, apenas os blocos de SIP PROXY, RTP SERVER e Aplicação Java são pertinentes para o experimento.

- O SIP PROXY é o responsável pelo registro e localização de assinantes, sinalização e controle das chamadas SIP, estratégias de roteamento de chamadas para os *gateways*, suporte a SNMP, presença e contabilização de chamadas. No PABX-IP-IMS, este bloco funcional é implementado pelo pacote de software Opensips (OPENSIPS, 2010).
- O RTP proxy é o responsável por instrumentar a travessia entre redes distintas para UAs que estão em redes locais atrás de NAT. No PABX-IP-IMS, este bloco é implementado pelo RTP Proxy Versão 1.2.1 (RTPPROXY, 2010).
- A Aplicação Java é a responsável pelo gerenciamento centralizado dos serviços disponibilizados pelo SIP Proxy, SIP/PSTN Gateway e Servidor de Mídia (*Media Server*), permitindo configuração de parâmetros, reinício

de serviços e visualização de relatórios de chamadas. No PABX-IP, este bloco funcional é implementado por uma aplicação Java Enterprise desenvolvida exclusivamente para o PABX-IP-IMS, rodando sobre o container EJB Glassfish 2.0.

O PABX-IP para este experimento foi instalado em um computador dedicado locado à GoDaddy Inc. (www.godaddy.com), localizada no Texas, USA, rodando o Sistema Operacional Red Hat Enterprise Linux 4.0.

Figura 7 – Aplicação Java para administração do PABX-IP.

Os equipamentos terminais são notebooks que possuem “*headsets*” (conjunto fones de ouvido e microfone) instalados e utilizam clientes SIP baseados no *softphone* Zoiper Versão 2.37, mostrado na figura 8.

Figura 8 – Interface gráfica do softphone ZOIPER 2.37.

Embora o *softphone* tenha sido usado inúmeras vezes durante a preparação do ambiente para realizar validações e testes iniciais, ele de fato não foi utilizado diretamente para o experimento, pois era necessário realizar grande volume de chamadas. Optou-se por utilizar o pacote de software SIPP, uma ferramenta gratuita de código aberto (GPL) para teste e geração de tráfego SIP e RTP (SIPP, 2010).

```

ocadmin@vista:~/sipp
----- Scenario Screen ----- [1-4]: Change Screen --
Call-rate(length)  Port  Total-time  Total-calls  Remote-host
    10 cps(0 ms)   5061      4.01 s       40  127.0.0.1:5060(UDP)

10 new calls during 1.000 s period      16 ms scheduler resolution
0 concurrent calls (limit 30)           Peak was 1 calls, after 0 s
0 out-of-call msg (discarded)
1 open sockets

      Messages  Retrans  Timeout  Unexpected-Msg
INVITE ----->      40      0      0
  100 <-----      0      0      0
  180 <-----      40      0      0
  200 <----- E-RTD  40      0      0
  ACK ----->      40      0
    [    0 ms]
  BYE ----->      40      0      0
  200 <-----      40      0      0

----- [+-|*|/]: Adjust rate ---- [q]: Soft exit ---- [p]: Pause traffic -----

```

Figura 9 – SIPP em modo console.

O SIPP, mostrado na figura 9, é capaz de gerar requisições e respostas a partir de cenários criados de forma declarativa utilizando arquivos XML. Estes cenários são disponibilizados pelo fornecedor de software na forma de arquivos padronizados para testes utilizando requisições SIP do tipo INVITE, tanto para o UAC, que gera as requisições, quanto para o UAS, que as responde. Também é possível injetar uma massa de testes – dados de assinantes, entre outras – que é usada pelo SIPP em conjunto com o script XML de forma a simular condições de campo. No entanto, a função mais interessante do pacote é a capacidade de gerar e receber fluxos RTP por meio de arquivos PCAP (captura de pacotes).

Após a pesquisa por soluções para o monitoramento das chamadas, optou-se pelo software VQManager (VQMANAGER, 2010) para a coleta e armazenamento dos dados de operação. Sua interface gráfica é apresentada na figura 10.

Figura 10 – Página principal do VQManager.

Desenvolvido para redes VoIP, o VQManager foi escolhido por sua simplicidade de operação e configuração, sua capacidade de monitoração de QoS

em tempo real e por ser capaz de calcular durante a execução da chamada o Fator R e respectivo MOS. A versão utilizada foi a shareware, com período de avaliação de trinta dias. Ela é capaz de monitorar qualquer equipamento VoIP que suporte protocolo SIP e use RTP/RTCP como protocolos de transporte de mídia. Todas as informações são armazenadas em arquivos de log que podem ser exportados em diversos formatos, incluído arquivos tabulados .CSV.

Além da ferramenta de monitoração, ajustes nos equipamentos de rede (*switches*) são necessários. Para que os indicadores de QoS possam ser calculados pelo VQManager, é necessário configurar o espelhamento de porta (*port mirroring*) em cada um dos switches e configurar o software para passar a escutar todo o tráfego nesta porta. O VQManager trabalha realizando a captura dos pacotes VoIP (*sniffing*) de todas as chamadas que ocorrem na rede. Todo o tráfego da rede precisa ser espelhado para uma porta de tal forma a permitir que o software possa capturar os pacotes e realizar suas contabilizações.

A conectividade foi representada por conexões internet de banda larga existentes nas dependências do Instituto Certi Amazônia (link Embratel 512 Kpbs), e nas unidades Trópico de Manaus (2 Mbps), em Manaus, Amazonas e Trópico de Campinas, São Paulo (2 Mbps), além da conectividade da GoGaddy, no estado do Texas, USA – servidor PABX-IP.

4.2 DESCRIÇÃO DO EXPERIMENTO

Na figura 11 é apresentado o procedimento adotado para a realização dos experimentos bem como o arranjo de equipamentos e ferramentas de software utilizado.

Figura 11 – Experimento para validação do método de correlação.

Dois terminais são usados: o primeiro para geração da carga SIP/RTP e o segundo para recepção do fluxo de mídia. O “terminal 01”, parametrizado por meio do script de teste XML, age como UAC. Ele utiliza os dados de carga contidos no arquivo .CSV para fornecer as credenciais necessárias ao registro no PABX-IP e para realizar a chamada, enviando a mídia por meio da execução do arquivo PCAP (que contem pacotes RTP). Este fluxo de mídia é representado em (1), partindo do UAC e chegando ao “Terminal 02”, parametrizado para atuar como um UAS. Este processo é automatizado, tendo sido ajustado para uma cadência de 20 (vinte) chamadas por minuto.

Em (2), a ferramenta “VQManager” faz a monitoração durante a execução das chamadas, escutando na porta espelhada do switch o tráfego de dados. Em tempo de execução, o VQManager realiza a contabilização das chamadas, armazenando as informações de QoS e de MOS.

Em (3), o log de cada bateria de testes é exportado no formato .CSV e

importado pelo sistema gerenciador de banco de dados MySQL, sendo armazenado na tabela “Resul_testes” do banco de dados.

Os dados são importados pela ferramenta estatística “R” (4), que aplica a técnica de Análise Exploratória de Dados (AED) e outras técnicas auxiliares, permitindo a geração de dendrogramas e outros gráficos (5) que irão aparelhar o pesquisador na análise dos resultados.

4.3 EXECUÇÃO DOS TESTES

Especificamente para os interesses deste trabalho foram realizados experimentos em conjunto com as atividades de projeto, ao longo do ano de 2010. As atividades foram realizadas no período de Janeiro a Março de 2010 como parte da etapa de Testes e Homologação do Produto PABX-IP e foram organizadas conforme indicado a seguir.

4.3.1 TESTES DE LABORATÓRIO

Os testes foram realizados nas dependências do Instituto Certi Amazônia, em ambiente confinado, sobre rede local, ao longo de 15 dias. Foram realizadas 500 mil chamadas, utilizando-se o software SIPP como gerador de carga.

Figura 12 – Ambiente de testes (laboratório).

O ambiente de teste, indicado na figura 12, era composto por:

- 1 Switch gerenciável 24 portas 3Com 3300XM;
- 1 Servidor Dell Poweredge R710 (PABX-IP);
- 2 Notebooks HP Pavillion dv4000 (Terminais);
- 1 Computador Dell Desktop OptiPlex 780 (VQManager).

Os arquivos XML utilizados para a parametrização do SIPP constam dos Apêndices A e B. Um exemplo referente ao arquivo .CSV com a massa de testes é apresentado na tabela 3.

Tabela 3 - Exemplo de Arquivo .CSV do SIPP.

SEQUENTIAL		
User01	8.8.49.100	[authentication username=User01 password=1000]
User02	8.8.49.100	[authentication username=User02 password=2000]
...	...	

Fonte: SIPP

Cada sessão de testes tinha duração de 1 (uma) hora. Ao final da sessão, o arquivo .CSV gerado pelo VQManager era coletado e seus dados armazenados em no SGBD (MySQL Versão 5.1). Um exemplo do arquivo .CSV gerado pelo VQManager é apresentado na tabela 4.

Tabela 4 – Exemplo de arquivo .CSV gerado pelo VQManager

Initiator	Participant	Duration	Call Status	Protocol	Initiator MOS	Participant MOS	MOS	Delay	Jitter	Packet Loss
2000	sip:1000@voip.manaus.certi.org.br	4 mins43 secs	Completed	SIP	1,9	1,96	1,9	835	106	4,2
2000	sip:1000@voip.manaus.certi.org.br	9 mins12 secs	Completed	SIP	1,3	1,3	1,3	2463	312	12,3
2000	sip:1000@voip.manaus.certi.org.br	30 secs	Completed	SIP	3,8	3,8	3,8	276	35	1,4

Fonte: VQManager

4.3.2 TESTES DE CAMPO

Contou com a utilização de um servidor PABX-IP remoto instalado nos USA e clientes instalados nas dependências do Instituto Certi Amazônia em Manaus-AM e da Trópico Sistemas e Telecomunicações Ltda., em Manaus-AM e Campinas-SP. Foram realizadas cerca de 300 mil chamadas ao longo de 45 (quarenta e cinco) dias entre uso de *softphones* com interlocutores humanos e carga artificial utilizando-se o software SIPP.

O ambiente de teste em Manaus e Campinas, exibido na figura 13, era similar e composto por:

- 1 Switch gerenciável (no caso da Certi, o mesmo 24 portas 3Com 3300XM);
- 2 Notebooks HP Pavillion dv4000 (Terminais);
- 1 Computador Dell Desktop OptiPlex 780 (VQManager);
- 1 link banda larga para conexão internet.

O procedimento de foi similar ao descrito na seção anterior, acrescentam-se testes com operadores humanos, em chamadas via *softphone* Zoiper.

Figura 13 – Ambiente de testes (campo).

4.3.3 ANÁLISE EXPLORATÓRIA DE DADOS (AED)

O uso de softwares estatísticos para a análise de dados é rotineiro quando se faz necessária a interpretação de resultados e naturalmente foi utilizado nesta pesquisa. Para este trabalho optou-se pela ferramenta computacional R (RPROJECT, 2010), gratuita e de código fonte aberto. Como é licenciada sob GPL, pode ser modificada ou melhorada com novos procedimentos desenvolvidos pelo usuário a qualquer momento. O R é capaz de realizar análise e manipulação de dados, com testes paramétricos e não paramétricos, modelagem linear e não linear, análise de séries temporais, análise de sobrevivência, simulação e estatística espacial, além de apresentar facilidade na elaboração de diversos tipos de gráficos. Sua interface gráfica é apresentada na figura 14.

Figura 14 – Interface gráfica do Software R para computação estatística.

O software R permite conexão direta com o SGBD MySQL por meio de sua biblioteca RMySQL, sendo possível importar os dados de teste diretamente para a ferramenta estatística. Desta forma, foi importada uma amostra representativa da população, totalizando 150 mil chamadas, utilizada no processo de análise. Os dados importados pelo R são armazenados internamente como uma estrutura do tipo “*data.frames*”. Para nosso experimento, o *data.frame* criado chama-se “dvoip150k”.

Após a importação, alguns testes iniciais nos permitem um vislumbre sobre a natureza dos dados:

Sumário – obtido através da função *summary()*, o sumário apresenta algumas informações sobre a amostra, como valores mínimos e máximos, quartis, media e mediana. Para o conjunto *data.frame* dvoip150k, o resultado para o *summary()* é mostrado na figura 15 e na tabela 5.

Figura 15 – Captura de tela do R: Sumario

Tabela 5 – Sumário do conjunto de dados após importação

MOS	Delay	Jitter	PacketLoss
Min. :1.300	Min. : 0.0	Min. : 0.00	Min. : 1.00
1st Qu.:2.200	1st Qu.: 249.0	1st Qu.: 33.00	1st Qu.: 1.30
Median :3.000	Median : 348.0	Median : 48.00	Median : 2.00
Mean :2.857	Mean : 427.2	Mean : 65.86	Mean : 13.81
3rd Qu.:3.500	3rd Qu.: 548.0	3rd Qu.: 83.00	3rd Qu.: 4.00
Max :4.200	Max :1200.0	Max :253.00	Max :309753.70

Fonte: Software R

Transformação Z, histogramas e QQ-Plot - permitem avaliar isoladamente a distribuição estatística de cada variável. No entanto, antes é necessário realizar a transformação Z usando-se a função *scale()*, de forma a criar um conjunto normal padronizado de dados. Isto permite comparar os resultados para as variáveis de interesse usando a mesma escala de referência. As variáveis MOS_Z, Delay_Z, Jitter_Z e PacketLoss_Z foram criadas para receber estes dados. Usa-se a função

`summary()` para exibir os dados do *data.frame* e verificar que a transformação foi feita. O resultado para o `scale()` e `summary()` é mostrado na figura 16.

Figura 16 – Captura de tela do R: Transformação Z

Após esta etapa, os histogramas são construídos por meio das funções `hist()`, `mean()`, `sd()` e `curve()`. A figura 17 mostra a construção do histograma para a variável `MOS_Z`. O Procedimento foi o mesmo para as demais variáveis. Todos os histogramas estão reunidos na figura 18.

Figura 17 – Captura de tela do R: Histograma do MOS_Z

Figura 18 – Histogramas das variáveis normais padronizadas

Gráfico QQ-Plot – este diagrama, assim como o histograma, também fornece uma estimativa sobre o comportamento da variável. Tomando a reta normal de referência que corta a área do gráfico na diagonal, observa-se a distribuição dos dados plotados. A maior ou menor semelhança da curva gerada pela distribuição dos pontos com a curva de referência determina sua normalidade. A figura 19 mostra a curva QQ-Plot para a variável MOS_Z. O Procedimento é o mesmo para as demais variáveis e todos os gráficos estão reunidos na figura 20.

Como se pode observar na figura 20, novamente apenas a variável “PacketLoss” diverge da hipótese inicial de normalidade assumida no início da pesquisa, estando os pontos das demais variáveis distribuídos em torno da reta normal.

Figura 19 – Captura de tela do R: Gráfico QQ-Plot para MOS_Z.

Figura 20 – Gráfico QQPlot para todas as variáveis.

Feita a avaliação do comportamento de cada variável, iniciou-se a busca de correlações. Procurou-se verificar se existia relação entre duas ou mais variáveis, isto é, saber se as alterações sofridas por uma das variáveis são acompanhadas por alterações nas outras. Em particular se busca determinar a existência da correlação e o grau de relação entre as variáveis.

Gráficos de dispersão - a primeira abordagem para a avaliação conjunta de variáveis foi novamente gráfica. Os valores das variáveis referentes ao total da população foram colocados num diagrama de dispersão aos pares e seu comportamento observado. Para tal, foi utilizada a função *pairs()*, cujo resultado é apresentado na figura 21. Os diagramas de dispersão permitem por meio de simples observação obter-se uma idéia razoável de como duas variáveis se relacionam.

Figura 21 – Captura de tela do R: Gráfico de dispersão para 02 variáveis.

Na figura 22 temos o detalhe do gráfico de dispersão para as variáveis MOS_Z e Delay_Z. No quadro superior direito temos a distribuição para MOS_Z x Delay_Z e no quadro inferior esquerdo os mesmos dados agora tomados em ordem inversa, Delay_Z x MOS_Z. Comparando o comportamento das curvas formadas na figura 22 com os gráficos de referência da figura 5 da seção 3.2 pode-se afirmar que existem fortes correlações, positivas em alguns intervalos e negativas em outros,

entre as variáveis DELAY_Z e MOS_Z.

Figura 22 – Gráfico de dispersão MOS_Z x Delay_Z

O Procedimento foi o mesmo para os demais pares de variáveis. Todos os gráficos de dispersão foram agrupados na figura 23 de forma a facilitar a avaliação. Por inferência, observa-se que existem correlações entre MOS e Delay, MOS e Jitter e entre Delay e Jitter. Por outro lado, existem correlações fracas ou ausência de correlação nas relações entre PacketLoss e as demais variáveis.

Figura 23 – Gráficos de dispersão

Dendrogramas - Outra maneira de graficamente avaliar as correlações é a construção de dendrogramas. Antes, porém, algum método de agrupamento hierárquico deve ser aplicado sobre uma amostra da população de forma a calcular as distâncias entre variáveis, estabelecendo índices de similaridade. O coeficiente de Jaccard foi escolhido, conforme já detalhado em 3.2. Aplicou-se o método de agrupamento hierárquico sobre uma amostra menor, formada por 102 (cento e duas) chamadas, tomadas aleatoriamente. Utilizou-se a função *vegdist()* da biblioteca *vegan()*, informando como parâmetro o método (*method="jaccard"*). Esta biblioteca do R é na verdade direcionada aos especialistas em ecologia e ciências naturais, mas contém métodos para análise de diversidade populacional entre outros recursos de HCA sendo, portanto útil ao estudo. Após esta etapa, o dendrograma pôde ser preparado utilizando-se a função *agnes()* da biblioteca *cluster()* e depois desenhado utilizando-se a função *plot()*.

A figura 24 mostra a construção do dendrograma utilizando as funções mencionadas. Na figura 25, temos o detalhe do dendrograma, onde se observa que três agrupamentos ou “clusters” significativos foram formados. Os índices de similaridade obtidos pelo coeficiente de Jaccard e os elementos que compõem cada agrupamento são apresentados na tabela 6.

Figura 24 - Captura de tela do R: Dendrograma.

Figura 25 – Dendrograma da amostra utilizando o índice de Jaccard.

Tabela 6 – Agrupamentos identificados no dendrograma

Índice de similaridade (Jc)	Cluster	Composição
0,28	02	Variáveis MOS e Jitter (Variação de Atraso)
0,48	01	Variáveis MOS e Delay (Atraso)
0,58	03	Variáveis MOS e PacketLoss (Perda de Pacotes)

Fonte: Software R

Coeficiente de Correlação de Pearson - além dos métodos de análise gráfica, a correlação entre duas variáveis foi calculada matematicamente por meio do coeficiente de correlação linear de Pearson (R_{xy}). Para tal, utilizou-se a função `cor()`, aplicada sobre a mesma amostra utilizada na elaboração do dendrograma. Na figura 26 temos o cálculo dos coeficientes realizados no R. Os coeficientes e as interpretações destes conforme definido na seção 3.2, para cada par de variáveis avaliado, são apresentados na tabela 7.

```

RGui
Arquivo  Editar  Visualizar  Msc  Pacotes  Janelas  Ajuda

R Console
> C_mos_delay = cor(dvoip102$MOS, dvoip102$Delay, use = "everything", method = c("pearson"))
> C_mos_jitter = cor(dvoip102$MOS, dvoip102$Jitter, use = "everything", method = c("pearson"))
> C_mos_PacketLoss = cor(dvoip102$MOS, dvoip102$PacketLoss, use = "everything", method = c("pearson"))
> C_Delay_jitter = cor(dvoip102$Delay, dvoip102$Jitter, use = "everything", method = c("pearson"))
> C_Delay_PacketLoss = cor(dvoip102$Delay, dvoip102$PacketLoss, use = "everything", method = c("pearson"))
> C_Jitter_PacketLoss = cor(dvoip102$Jitter, dvoip102$PacketLoss, use = "everything", method = c("pearson"))
>
> C_mos_delay
[1] -0.5370857
> C_mos_jitter
[1] -0.7116251
> C_mos_PacketLoss
[1] -0.2032659
> C_Delay_jitter
[1] 0.7204282
> C_Delay_PacketLoss
[1] -0.123062
> C_Jitter_PacketLoss
Erro: objeto 'C_Jitter_PacketLoss' não encontrado
> C_Jitter_PacketLoss
[1] -0.07891786
>

```

Figura 26 – Captura de tela do R: Cálculo dos coeficientes de Pearson

Tabela 7 – Coeficientes de Correlação Linear de Pearson

X	Y	Rxy	Interpretação
MOS	Delay	-0,5370857	Correlação negativa moderada
MOS	Jitter	-0,7116251	<u>Correlação negativa forte</u>
MOS	Packet Loss	-0,2032659	Pode ser tomada como 0. Sem correlação.
Delay	Jitter	0,7204282	<u>Correlação positiva forte</u>
Delay	Packet Loss	-0,123062	Pode ser tomada como 0. Sem correlação
Jitter	Packet Loss	-0,07891786	Pode ser tomada como 0. Sem correlação.

Fonte: Software R

Teste de significância - as correlações MOS-Jitter e Delay-Jitter são consideradas fortes segundo o coeficiente de correlação de Pearson. No entanto, o coeficiente de correlação Rxy é apenas uma estimativa do coeficiente de correlação populacional ρ_{xy} , conforme mostrado na seção 3.2. Aplicou-se o teste paramétrico “t-student” para verificar se o valor de Rxy é coerente com o tamanho da amostra n, a um nível de significância α . Utilizou-se a função *r.test()* da biblioteca *psych()* para calcular o valor t-student do Rxy obtido para a correlação MOS-Jitter e para a correlação Delay-Jitter. A figura 27 mostra a execução dos testes no R.

```

RGui
Arquivo  Editar  Visualizar  Plots  Pacotes  Janelas  Ajuda

R Console
> library(psych)
> 
> n=102
> 
> r.test(n,C_mos_jitter)
Correlation Tests
Call:r.test(n = n, r12 = C_mos_jitter)
Test of significance of a correlation
t value -10.13 with probability < 0
and confidence interval -0.8 -0.6>
> 
> r.test(n,C_Delay_jitter)
Correlation Tests
Call:r.test(n = n, r12 = C_Delay_jitter)
Test of significance of a correlation
t value 10.39 with probability < 0
and confidence interval 0.61 0.8>
  
```

Figura 27 – Captura de tela do R: Teste de significância.

Em seguida, recorreu-se à tabela 1 contida na seção 3.2. Segundo o procedimento metodológico, deve-se confrontar o resultado dos testes e o grau de liberdade das variáveis com a tabela e verificar o enquadramento.

Para a correlação MOS-Jitter, devemos verificar os valores críticos na tabela com cem graus de liberdade ($n-2$). Desta forma, encontramos os valores 1,98 (5% ou $\alpha=0,05$) e 2,63 (1% ou $\alpha=0,01$). O valor calculado de $t=10,13$ é maior que os valores tabelados. Conclui-se pela rejeição de h_0 em nível de 1% de probabilidade ($\alpha=0,01$). Em termos práticos, conclui-se que o MOS possui uma forte correlação negativa com o Jitter com 99% de probabilidade. De maneira similar, temos que para a correlação Delay-Jitter o valor obtido foi $t=10,39$, também maior que os valores tabelados. Entende-se a rejeição de h_0 também com 1% de probabilidade o que resulta na confirmação da forte correlação positiva entre as variáveis com probabilidade de 99%. Os resultados estão resumidos na tabela 8.

Tabela 8 – Resultado do teste t-student para verificação da significância de R_{xy} .

Correlação	R_{xy}	T_c^*	α	Hipótese
MOS – Jitter	-0,7116251	10,13	1% ($\alpha=0,01$)	$\rho \neq 0$
Delay – Jitter	0,7204282	10,39	1% ($\alpha=0,01$)	$\rho \neq 0$

Fonte: Software R

4.3.4 RESULTADOS

As variáveis MOS, Atraso e Variação de Atraso apresentaram comportamento normal conforme hipótese 2 ou seja, os conjuntos de dados resultantes dos experimentos mostraram-se bem comportados e respeitando a distribuição estatística normal. As figuras 17, 18, 19 e 20 corroboram com esta conclusão, apresentando a distribuição das variáveis por meio de uma representação gráfica. Para as variáveis em questão, a representação gráfica se apresenta muito próxima a curva da distribuição estatística normal. Tal comportamento é coerente considerando o volume de dados obtidos nos experimentos e submetidos à análise estatística pois, conforme o teorema central do limite, toda soma de variáveis aleatórias independentes de média finita e variância limitada é aproximadamente normal, desde que o número de termos da soma seja suficientemente grande (KREBS,

1989), o que foi o caso.

A variável Percentual de Perda de Pacotes, por sua vez, divergiu da hipótese 2, conforme indicado pelas figuras 18 e 20. Este comportamento não era esperado, embora seja justificável considerando que a perda de pacotes não está associada apenas a parâmetros de qualidade de serviços. Outros eventos como erros de Verificação de Seqüência de Quadros (*Frame Check Sequence* – FCS) e taxa de erros de bit (*Bit Error Rate* – BER), relacionados com a variação na condição de rede, além de fragmentação de pacotes e perda de seqüência, contribuem para este indicador. No entanto tais métricas possuem distribuição estatística reconhecidamente diversa da normal.

Os diagramas de dispersão apontaram existência de correlações entre MOS x Atraso e MOS x Jitter, mas pouco acoplamento entre as demais relações. Estes comportamentos são evidenciados na figura 23. A existência de correlações entre os pares avaliados é graficamente identificada pela aglomeração dos pontos segundo um padrão ordenado, passível de ser representado por uma curva. Tomando o detalhe da figura 22 que apresenta o par MOS x Delay, e considerando a referência da figura 5 da seção 3.2, é possível observar correlações que ora se manifestam positivas e ora se manifestam negativas, conforme o intervalo de pontos considerado. As curvas formadas tendem à representações pseudo-lineares ou pseudo-exponenciais, indicando consistência entre os pontos do aglomerado e a existência de uma tendência. Observando novamente a figura 23 que contém o conjunto de todos os pares de variáveis testados, nota-se que as curvas formadas por pares como MOS x PacketLoss e Delay x PacketLoss não expressam correlações, indicando que os conjuntos são disjuntos.

Os agrupamentos identificados por meio da HCA confirmam a tendência, pois os pares MOS-Jitter e MOS-Delay apresentaram índices de Jaccard de 0,28 e 0,48 respectivamente (maior acoplamento) enquanto o par MOS-PacketLoss apresentou Índice de Jaccard de 0,58. Tais relações podem ser observados por meio da figura 25 onde os agrupamentos estão indicados. Índices de jaccard menores de 0,50 apontam grande relação entre os elementos do conjunto, apontando proximidade.

Por seu turno, os coeficientes de correlação linear de Pearson para os mesmos pares destacados pela HCA acrescentaram ao resultado a direção da tendência onde MOS-Jitter se apresentou com forte correlação negativa e MOS-

Delay com correlação moderada positiva. Este resultado analítico não representa os intervalos ora positivos e ora negativos indicados nos gráficos de dispersão, mas sim a tendência geral composta pelo intervalo de maior valor absoluto ou intervalo dominante. Portanto, permanece coerente com o já apresentado pelos gráficos de dispersão. O comportamento das variáveis de interesse, indicado pelos métodos formais aplicados, também concorda com a compreensão empírica da relação entre a qualidade de fala e a variação de atraso, pois o aumento do atraso total fim a fim e da variação de atraso na chamada causam também o aumento no picotamento do áudio e a conseqüente piora na experiência do usuário final.

O teste paramétrico t-student aplicado ao conjunto amostral permitiu extrapolar os resultados obtidos por meio do Coeficiente de Pearson para toda a população p , mostrando que as correlações apontadas pelo método são válidas dentro de um intervalo de confiança de 99%.

O processo sugerido para a obtenção das correlações necessitou de 4 (quatro) variáveis distintas para ficar adequadamente representado, sendo então aderente ao enunciado pela hipótese 1, pois este número de variáveis de interesse pertence a um espaço $(n+1)$ -dimensional com $N > 2$.

Finalizados os experimentos e realizada a análise dos dados coletados, foi possível comprovar por meio do processo proposto a existência de correlações entre a qualidade da fala aferida pelo MOS, uma medida subjetiva, e as métricas de qualidade de serviços, uma medida objetiva. Estas correlações provaram-se válidas para toda a população p com 99% de probabilidade de sucesso, sob as condições apresentadas, conforme objetivo inicial.

Grande parte dos marcos apontada inicialmente foram atingidos. Para a definição de um processo para a medida da qualidade subjetiva, foi realizada a escolha de uma implementação em particular dado que métodos baseados no modelo-E já estão bem definidos e consolidados. O estudo de caso foi realizado por meio da construção de ambiente VoIP de produção. Foi gerada carga de chamadas artificialmente, realizada a coleta de dados e aplicado o processo. No entanto, os testes em campo propostos não foram concretizados, pois não foi viabilizada nenhuma parceria para utilização do PABX-IP em operadoras de telefonia nacionais.

5 CONCLUSÕES

Neste capítulo são apresentadas as considerações finais sobre este trabalho. Primeiramente é apresentada uma retrospectiva das contribuições do projeto desenvolvido e são apresentadas opções de trabalhos futuros.

5.1 CONTRIBUIÇÕES

A observação dos histogramas gerados mostrou que à exceção do percentual de perda de pacotes, o comportamento das variáveis de interesse de fato apresenta distribuição estatística normal conforme suposição inicial. Este resultado é importante, pois cria a possibilidade de se utilizar outras técnicas estatísticas que possuem esta premissa para sua adoção.

Todos os métodos empregados na análise, tanto gráficos como procedurais, convergiram para o mesmo resultado. A avaliação conjunta do comportamento das variáveis por meio dos diagramas de dispersão apontou para a existência de correlações entre MOS e as métricas de atraso e variação de atraso enquanto também mostrou pouco acoplamento entre o percentual de perda de pacotes e o restante dos indicadores. A abordagem por meio de análise de agrupamentos hierárquicos (HCA) e dendrogramas avançou na compreensão destas correlações. Os três agrupamentos identificados, referentes às correlações MOS-Jitter, MOS-Delay e MOS-PacketLoss, apresentaram índices de Jaccard de 0,28, 0,48 e 0,58 respectivamente, o que evidencia maior similaridade entre a qualidade da fala e a variação de atraso. A determinação dos coeficientes de correlação linear de Pearson para os mesmos pares destacados pela análise de agrupamentos hierárquicos manteve a coerência entre os resultados, novamente apontando o par MOS-Jitter como possuidor de uma forte correlação. Além desta confirmação sobre a correlação, trouxe adicionalmente informações sobre a natureza da mesma, mostrando-a como negativa. Isto implica dizer que uma redução no valor da primeira variável implica no aumento do valor de segunda, o que concorda com a

compreensão empírica da relação entre a qualidade de fala e a variação de atraso. É senso comum que a redução no *jitter* durante chamadas VoIP ocasiona uma melhora na qualidade do áudio, reduzindo principalmente o picotamento da voz.

Com a validação obtida após o teste paramétrico t-student, é possível afirmar como resultado principal da pesquisa que existe uma forte correlação entre o MOS e a variação de atraso, válida para toda a população ρ .

A pesquisa almejava estabelecer uma correlação entre a qualidade subjetiva da fala e uma métrica de qualidade de serviços. Um processo baseado em estatística multivariada foi proposto para alcançar o objetivo e um estudo de caso foi realizado para implementar o processo, testar as hipóteses levantadas e obter a correlação.

O método escolhido para determinação da qualidade da fala, baseado na recomendação ITU-T G.107, mostrou-se adequado para o estudo proposto dentro das condições definidas no escopo, embora não contemple a predição de fluxos em rajada.

A decisão pela escolha do conjunto de métricas de qualidade de serviços adotado frente às demais disponíveis para avaliação mostrou-se a mais acertada, pois estas possuem grande impacto sobre o desempenho do serviço avaliado, o que restou demonstrado pelos coeficientes de correlação obtidos.

Referente aos métodos estatísticos adotados para a análise observou-se após a pesquisa que não contemplaram totalmente os objetivos propostos, em parte devido à inadequação na escolha do estimador para os agrupamentos hierárquicos. Um estimador baseado em distância ou dependência poderia produzir resultados melhores sobre os conseguidos com o estimador de similaridade.

Independente das limitações apresentadas, foi possível comprovar estatisticamente a existência de forte correlação linear entre a qualidade da fala aferida pelo MOS, uma medida subjetiva, e a métrica de qualidade de serviços definida como variação de atraso fim-a-fim, uma medida objetiva.

5.2 TRABALHOS FUTUROS

O estabelecimento da correlação entre os indicadores subjetivos da qualidade da fala e as métricas objetivas de qualidade de serviços é um primeiro passo na direção da busca de um sistema automatizado para o controle de QoS baseado no MOS ou Fator R.

Algumas possibilidades para a continuidade do trabalho são a avaliação de novos métodos estatísticos para a eliciação de correlações, a escolha de novos estimadores para balizar a técnica de agrupamentos hierárquicos, a avaliação de distribuições estatísticas que descrevam adequadamente a correlação MOS-Jitter e a proposição de uma equação característica que descreva seu comportamento.

6 REFERÊNCIAS

BAI, Y.; ITO, M. R. **A Study for Providing Better Quality of Service to VoIP Users. Advanced Information Networking and Applications, 2006 (AINA 2006)**, 2006, Vienna, Austria. 20th International Conference, [S.l.], 2006. p. 799 – 804.

BRUSCATO, A. C. et al. **Análise de Serviços de Voz Sobre Redes de Comunicação de Dados**. Brasil Telecom. Brasil. 2004.

CARVALHO, L. S. G.; MOTA, E. S. ; QUEIROZ, J. M. **Análise comparativa de padrões para medida de qualidade da fala**. In: SUCESU, 2003, Salvador. SUCESU, 2003.

CISCO **CISCO Systems. Understanding Delay in Packet Voice Networks**. Disponível em http://www.cisco.com/en/US/partner/tech/tk652/technologies_white_paper09186a00800a8993.shtml. Acessado em 14 de Jul. de 2009.

COLCHER, S. et. al. **VoIP: Voz sobre IP**. Rio de Janeiro: Campus, 2005. 288 p.

CONWAY, A. **Systems and Methods for Automatic Evaluation Of subjective quality of packetized telecommunication signals while varying implementation parameters**. Verizon Laboratories Inc. USA. 2005.

CORRAR, L. J.; PAULO, E.; DIAS, J. M. F. **Análise Multivariada**. [S.l.]: Atlas, 2007. 544 p.

DING, L.; GOUBRAN, R. A. **Assessment of effects of packet loss on speech quality in VoIP. Haptic, Audio and Visual Environments and Their Applications, 2003 (HAVE 2003)**, 2003, Ottawa, Canadá. Proceedings: The 2nd IEEE International Workshop, [S.l.], 2003. p. 49 – 54.

FONSECA, J. F. **Curso De Estatística**. São Paulo: Editora Atlas, 1996. 320 p.

FUJIMOTO, K., ATA, S., MURATA, M. **Playout control for streaming applications by statistical delay analysis**. ICC – 2001.

GARVIN, D. A. **Gerenciando a qualidade - A visão estratégica e competitiva**. Rio

de Janeiro: Editora Quality Mark, 1993. 355 p.

GOMES, F.P. **Curso de Estatística Experimental**, Piracicaba: Editora Nobel, 1990. 467p.

GONÇALVES, F. E. A. **Telefonia IP com SIP**. Florianópolis: Editora V. Office, 2006. 276 p.

GRAH, M.; RADCLIFFE, P. **Dynamic QoS For Commercial VoIP In Heterogeneous Networks**. In: Software, Telecommunications and Computer Networks (SoftCOM 2008), 2008, Split - Dubrovnik, Croatia. 16th International Conference, [S.I.], 2008. p. 163 – 168.

_____. **Dynamic QoS and Network Control for Commercial VoIP Systems in Future Heterogeneous Networks**. In: The IEEE International Symposium on Multimedia 2008 (ISM 2008), 2008, Berkeley, USA. Tenth IEEE International Symposium, [S.I.], 2008. p. 356 – 363.

_____. **Dynamic network control with QoS and resource priority monitor based on active “eM” for commercial VoIP**. In: International Conference on Industrial Technology 2009 (ICIT'09), 2009, Churchill, Australia. IEEE International Conference, [S.I.], 2009. p. 1 – 6.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T P.800**: Methods for subjective determination of transmission quality. [S.I.], 1996.

_____. **ITU-T P.861**: Objective quality measurement of telephone band (300-3400 Hz) speech codecs. [S.I.], 1998.

_____. **ITU-T G.107**: The E-model, a computational model for use in transmission planning. [S.I.], 2000.

_____. **ITU-T G.114**: One-way transmission time. [S.I.], 2003.

EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE. **ETSI 101-329-5**: Quality of Service (QoS) measurement methodologies. [S.I.], 2003.

JURAN, J.M. **Juran na Liderança pela Qualidade**. São Paulo: Pioneira, 1990. 386

p.

KREBS, C. J. 1989. **Ecological Methodology**. New York: Harper-Collins, 370 p.

LIMA, A. F. M. et al. **Monitoring based on statistical analysis for evaluating quality of calls in VoIP environment**. In: Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing (WCNM). Wuhan, 2005.

LIMA A. et al. **VoIP Management: Speech quality monitoring architecture based on statistical analysis**. International Journal of Network Management. Vol. 17(4) 263-274. 2007.

LIPOVAC, V. **Testing Integrated Qos of Voip: Packets to Perceptual Voice Quality**. [S.l.]: Editora AUERBACH, 2009. 350 p.

MACHADO, F. N. R. **Tecnologia e Projeto de Data Warehouse**. [S.l.]: Editora Erica, 2004. 320 p.

MEGGELEN, J. V.; SMITH, J., MADSEN, L. **Asterisk: The Future of Telephony**. [S.l.]: Editora O'Reilly. 2005. 376 p.

MEHMOOD, M.A.; JADOON, T.M.; SHEIKH, N.M. **Assessment of VoIP quality over access networks**. In: The First IEEE and IFIP International Conference in Central Asia 2005 (ICI2005), Bishkek, Kyrgyz Republic, 2005. p. 26-29.

MOITA NETO, J. M.; MOITA, G. C. **Uma Introdução à Análise Exploratória de Dados Multivariados**. São Paulo: Editora Química Nova, 1998.

OPENSIPS **OpenSIPS Project - Former OpenSER Project**. Disponível em <http://www.opensips.org>. Acessado em 30 de Nov. de 2010.

ORANGE **Home Broadband - Broadband Deals - Orange Shop**. Disponível em <http://shop.orange.co.uk/broadband/>. Acessado em 30 de Jul. de 2009.

PASSITO, A. et al. **Evaluating Voice Speech Quality in 802.11b Networks with VPN/IPSec**. In: Proceedings of IEEE International Conference on Networks (ICON). Kuala Lumpur, Malaysia, 2005.

PNUD. **Pesquisa Nacional por Amostra de Domicílios: Acesso à Internet e posse de telefone móvel celular para uso pessoal 2008**. [S.l.], 2009.

RPROJECT **The R Project for Statistical Computing**. Disponível em <http://www.r-project.org/>. Acessado em 17 de Ago. de 2009.

RTPPROXY **Sippy RTP Proxy**. Disponível em <http://b2bua.org/wiki/RTPproxy>. Acessado em 30 de Nov. de 2010.

SHERBURNE, P. FITZGERALD, C. **You Don't Know Jack About VoIP**. ACM Queue vol. 2, no. 6 – September 2004.

SIPP. **SIPP**. Disponível em <http://sipp.sourceforge.net/>. Acessado em 30 de Nov. de 2010.

STEINBERG, M.; SOUTHWICK, P. **VoIP and Telephony**. Editora Colchester: Hill Associates, 2003. Disponível em <http://www.hill.com/resources/whitepapers.html>. Acessado em 23 de Ago. de 2009.

TAKAHASHI, A.; KURASHIMA, A.; YOSHINO, H. **Objective Assessment Methodology for Estimating Conversational Quality in VoIP**. IEEE Transactions on Audio, Speech, and Language Processing, Volume 14, Issue 6, p. 1984 – 1993. 2006.

TANEMBAUM, A. S. **Redes de Computadores**. Rio de Janeiro: Campus, 2003. 948 p.

VALLE, R. et al.. **Estudo Comparativo do Impacto de Algoritmos de Compensação de Jitter sobre a Qualidade da Fala**. In: I Escola Regional de Informática – Regional Norte 1. ERIN 2009. Brasil. 2009.

VIANNA, B. A. et al.. **aMOS: Algoritmo MOS-Adaptativo para Fontes VoIP**. In: Simpósio Brasileiro de Sistemas Multimídia e Web. WebMedia 2006. Brasil. 2006.

VQMANAGER **VQManager VoIP Quality Monitoring Software**. Disponível em <http://www.manageengine.com/products/vqmanager/>. Acessado em 02 de Ago. de 2010.

WALLINGFORD, T. **Switching to VoIP**. O'Reilly Media, 2005. 504 p.

WÄLTERMANN, M. et al. **A Technique for Seamless VoIP-Codec Switching in Next Generation Networks**. In: IEEE International Conference on Communications (ICC '08), Beijing. 2008. p. 1772 - 1776.

ANEXO A - SCRIPTS EM XML PARA PARAMETRIZAÇÃO DO SIPP COMO UAC

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<scenario name="Basic Sipstone UAC">

  <!-- In client mode (sipp placing calls), the Call-ID MUST be      -->
  <!-- generated by sipp. To do so, use [call_id] keyword.          -->
  <send retrans="1000">
    <![CDATA[

      INVITE sip:[service]@[remote_ip]:[remote_port] SIP/2.0
      Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
      From: sipp <sip:[field0]@[field1]:5060>;tag=[pid]SIPpTag00[call_number]
      To: sut <sip:[service]@[remote_ip]:[remote_port]>
      Call-ID: [call_id]
      CSeq: 1 INVITE
      Contact: sip:[field0]@[local_ip]:[local_port]
      Max-Forwards: 70
      Subject: Performance Test
      Content-Type: application/sdp
      Content-Length: [len]

      v=0
      o=user1 53655765 2353687637 IN IP[local_ip_type] [local_ip]
      s=-
      c=IN IP[media_ip_type] [media_ip]
      t=0 0
      m=audio [media_port] RTP/AVP 0
      a=rtpmap:0 PCMU/8000

    ]]>
  </send>

  <recv response="407" auth="true" rrs="true">
  </recv>

  <send retrans="5000">
    <![CDATA[

      INVITE sip:[service]@[remote_ip]:[remote_port] SIP/2.0
      Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
      From: sipp <sip:[field0]@[field1]:5060>;tag=[pid]SIPpTag00[call_number]
      To: sut <sip:[service]@[remote_ip]:[remote_port]>
      Call-ID: [call_id]
      CSeq: 1 INVITE
      Contact: sip:[field0]@[local_ip]:[local_port]
      [field2]
      Max-Forwards: 70
      Subject: Performance Test
      Content-Type: application/sdp
```

```

Content-Length: [len]

v=0
o=user1 53655765 2353687637 IN IP[local_ip_type] [local_ip]
s=-
c=IN IP[media_ip_type] [media_ip]
t=0 0
m=audio [media_port] RTP/AVP 0
a=rtpmap:0 PCMU/8000

]]>
</send>

<recv response="100"
  optional="true" rrs="true">
</recv>

<recv response="180" optional="true" rrs="true">
</recv>

<recv response="183" optional="true" rrs="true">
</recv>

<!-- By adding rrs="true" (Record Route Sets), the route sets -->
<!-- are saved and used for following messages sent. Useful to test -->
<!-- against stateful SIP proxies/B2BUAs. -->
<recv response="200" rtd="true" rrs="true">
</recv>

<!-- Packet lost can be simulated in any send/recv message by -->
<!-- by adding the 'lost = "10"'. Value can be [1-100] percent. -->
<send>
  <![CDATA[

    ACK sip:[service]@[remote_ip]:[remote_port] SIP/2.0
    Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
    From: sipp <sip:[field0]@[local_ip]:[local_port]>;tag=[pid]SIPpTag00[call_number]
    To: sut <sip:[service]@[remote_ip]:[remote_port]>[peer_tag_param]
    Call-ID: [call_id]
    CSeq: 1 ACK
    [routes]
    Contact: sip:[field0]@[local_ip]:[local_port]
    Max-Forwards: 70
    Subject: Performance Test
    Content-Length: 0

  ]]>
</send>

<!-- This delay can be customized by the -d command-line option -->
<!-- or by adding a 'milliseconds = "value"' option here. -->
<pause/>

<!-- The 'crlf' option inserts a blank line in the statistics report. -->

```

```

<send retrans="500">
  <![CDATA[

    BYE sip:[service]@[remote_ip]:[remote_port] SIP/2.0
    Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
    From: sipp <sip:[field0]@[field1]:[local_port]>;tag=[pid]SIPpTag00[call_number]
    To: sut <sip:[service]@[remote_ip]:[remote_port]>[peer_tag_param]
    Call-ID: [call_id]
    CSeq: 2 BYE
    [routes]
    Contact: sip:[field0]@[local_ip]:[local_port]
    Max-Forwards: 70
    Subject: Performance Test
    Content-Length: 0

  ]]>
</send>

<recv response="200" crlf="true">
</recv>

<!-- definition of the response time repartition table (unit is ms) -->
<ResponseTimeRepartition value="10, 20, 30, 40, 50, 100, 150, 200"/>

<!-- definition of the call length repartition table (unit is ms) -->
<CallLengthRepartition value="10, 50, 100, 500, 1000, 5000, 10000"/>

</scenario>

```

ANEXO B - SCRIPTS EM XML PARA PARAMETRIZAÇÃO DO SIPP COMO UAS

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE scenario SYSTEM "sipp.dtd">
```

```
<send>
  <![CDATA[

    SIP/2.0 180 Ringing
    [last_Via:]
    [last_From:]
    [last_To:];tag=[pid]SIPpTag01[call_number]
    [last_Call-ID:]
    [last_CSeq:]
    Contact: <sip:[local_ip]:[local_port];transport=[transport]>
    Content-Length: 0

  ]]>
</send>
```

```
<send retrans="500">
  <![CDATA[

    SIP/2.0 200 OK
    [last_Via:]
    [last_From:]
    [last_To:];tag=[pid]SIPpTag01[call_number]
    [last_Call-ID:]
    [last_CSeq:]
    Record-route: <sip:[local_ip]:[local_port];lr>
    Contact: <sip:[local_ip]:[local_port];transport=[transport]>
    Content-Type: application/sdp
    Content-Length: [len]

  ]]>
```



```

v=0
o=user1 53655765 2353687637 IN IP[local_ip_type] [local_ip]
s=-
c=IN IP[media_ip_type] [media_ip]
t=0 0
m=audio [media_port] RTP/AVP 0
a=rtpmap:0 PCMU/8000

]]>
</send>

```

```

<recv request="ACK"
  optional="true"
  rtd="true"
  crlf="true"
  rrs="true">
</recv>

```

```

<recv request="BYE" rrs="true">
</recv>

```

```

<send>
<![CDATA[

SIP/2.0 200 OK
[last_Via:]
[last_From:]
[last_To:]
[last_Call-ID:]
[last_CSeq:]
Contact: <sip:[local_ip]:[local_port];transport=[transport]>
Content-Length: 0

]]>
</send>

```

```

<!-- Keep the call open for a while in case the 200 is lost to be -->

```

```
<!-- able to retransmit it if we receive the BYE again.      -->  
<timewait milliseconds="4000"/>
```

```
<!-- definition of the response time repartition table (unit is ms)  -->  
<ResponseTimeRepartition value="10, 20, 30, 40, 50, 100, 150, 200"/>
```

```
<!-- definition of the call length repartition table (unit is ms)  -->  
<CallLengthRepartition value="10, 50, 100, 500, 1000, 5000, 10000"/>
```

```
</scenario>
```