

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Comunicações

Detecção de Falhas em Sistemas de Comunicações

Autor: Gean Davis Breda

Orientador: Prof. Dr. Leonardo de Souza Mendes

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica. Área de Concentração: **Telecomunicações e Telemática.**

Banca Examinadora:

Leonardo de Souza Mendes, Ph.D	DECOM/FEEC/UNICAMP
Luis Geraldo Pedroso Meloni, Dr.	DECOM/FEEC/UNICAMP
Maurício Ferreira Magalhães, Dr.	DCA/DECOM/UNICAMP
Jonatas Manzolli, Dr.	NICS/UNICAMP
Marcelo Eduardo Pellenz, Dr.	PPGIA/PUC-Curitiba
Mário Lemes Proença Jr., Dr.	DC/UEL

Campinas, SP
Fevereiro / 2008

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

B743d Breda, Gean Davis
Detecção de falhas em sistemas de comunicação / Gean
Davis Breda. --Campinas, SP: [s.n.], 2008.

Orientador: Leonardo de Souza Mendes.
Tese (doutorado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas de comunicação. 2. Telefonia. 3.
Localização de falhas (Engenharia). I. Mendes, Leonardo
de Souza. II. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação. III.
Título.

Título em Inglês: Fault detection in communication systems.

Palavras-chave em Inglês: Communications systems, Fault detection, Telephony.

Área de concentração: Telecomunicações e Telemática.

Titulação: Doutor em Engenharia Elétrica

Banca examinadora: Luis Geraldo Pedroso Meloni, Maurício Ferreira Magalhães,
Jonatas Manzolli, Marcelo Eduardo Pellenz e Mário Lemes
Proença.

Data da defesa: 27/02/2008

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Gean Davis Breda

Data da Defesa: 27 de fevereiro de 2008

Título da Tese: "Detecção de Falhas em Sistemas de Comunicações"

Prof. Dr. Leonardo de Souza Mendes (Presidente): Leud.

Prof. Dr. Mario Lemes Proença Junior: Mario Lemes Proença Jr.

Prof. Dr. Marcelo Eduardo Pellenz: Marcelo Pellenz

Prof. Dr. Jônatas Manzolli: Jônatas Manzolli

Prof. Dr. Mauricio Ferreira Magalhães: Mauricio Ferreira Magalhães

Prof. Dr. Luis Geraldo Pedroso Meloni: Luis Geraldo Pedroso Meloni

Resumo

O presente trabalho concentra-se na área de detecção de falhas em sistemas de comunicações utilizando como base de informações arquivos de bilhetes contidos nas centrais. Estes arquivos, na grande maioria dos casos, são utilizados pelas operadoras de telecomunicação apenas para fins de cobrança dos usuários. A tese apresenta algoritmos para detecção de falhas e conseqüente geração de alarmes. Os algoritmos são analisados e classificados conforme o seu desempenho. O objetivo central da detecção de falhas com uma maior eficiência é diminuir as deficiências na qualidade de serviço do sistema, reduzindo as perdas econômicas e melhorando a satisfação dos clientes. Indicadores de mercado apontam que de 3% a 15% do faturamento das empresas de Telecom são perdidos por falhas na rede. Os algoritmos buscam maneiras eficientes de detectar falhas nos elementos de um sistema de comunicação.

Palavras-chave: Sistemas de Comunicação, Telefonia, Localização de falhas.

Abstract

The present work is focused in the area of fault detection in communications systems using billing tickets archives as base of information. In most of the cases, these archives are used by the telecommunication carriers only for billing purposes. The thesis presents algorithms for fault detection and consequence alarms generation. The algorithms are analyzed and classified according to their performance. The main objective of the fault detection with a bigger efficiency is to diminish the deficiencies in system's *QoS*, reducing economic losses and improving customers satisfaction. Market indicators show that from 3% to 15% of Telecommunication companies gains are lost for system failures. The algorithms search efficient ways to detect faults in the elements of a communication system.

Key-words: Communications Systems, Fault Detection, Telephony.

Agradecimentos

Percorri diversos caminhos no decorrer da vida. Durante muito tempo sempre tive a sensação de andar em círculos, sem respostas para as mesmas perguntas. Essa busca constante me conduziu a uma descoberta muito importante: somente no amor da minha esposa e filho é que pude reencontrar a essência da vida. Só consegui chegar até aqui graças a vocês.

Aos meus pais e irmãos pelo apoio e carinho constantes. A toda minha grande família.

Agradeço ao Leonardo que sempre teve paciência e dedicação necessárias.

Aos amigos Maurício, Antônio, Marcelo e Clézio, grandes companheiros desde Santa Maria.

A todos do LaRCom e da IgnisCom.

A Tânia, Paulo Zamboni e Nelson Schuch, pelo imenso carinho que sempre recebi, minha gratidão é eterna...

Em especial ao Guilherme pelo aprendizado que tive nos anos de Promon e que me ajudaram imensamente neste trabalho.

Ao Carlos, Paulo, Vitor, André, Roberto, Daniel, Priscila,..., e todos os outros, pela amizade e pelo trabalho conjunto que criou tanto conhecimento.

A Alexandre Alves, Vice-Presidente da empresa Transit Telecom, por ter colaborado cedendo bilhetes de tarifação.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, pelo suporte financeiro que permitiu a realização deste trabalho.

“Olha bem e deliberadamente para todos os caminhos. Tenta-os tantas vezes quantas forem necessárias. Depois faz a ti mesmo a pergunta: Será que esse caminho tem coração? Todos os caminhos são o mesmo: nenhum leva a parte alguma. No decurso da minha vida posso dizer que atravessei longos, longos caminhos e não cheguei a lugar algum. A pergunta agora tem um significado... Será que esse caminho tem coração? Se tem, o caminho é bom; se não tem, o caminho não tem importância. Nenhum dos caminhos leva a parte alguma; mas um tem coração, o outro não tem. Um proporciona uma viagem com alegria; na medida em que se o sequires, serás uno com o universo. Outro levar-te-á à fadiga. Um faz de ti um homem forte; o outro, um homem fraco.”

(Livro: Os Ensinamentos de Don Juan, Carlos Castañeda)

Sumário

<i>Resumo</i>	<i>ii</i>
<i>Abstract</i>	<i>ii</i>
<i>Agradecimentos</i>	<i>iii</i>
<i>Sumário</i>	<i>v</i>
<i>Índice de Figuras</i>	<i>ix</i>
<i>Índice de Tabelas</i>	<i>xii</i>
<i>Lista de Acrônimos</i>	<i>xiv</i>
<i>Trabalhos Publicados pelo Autor</i>	<i>xvii</i>
<i>Patentes Solicitadas pelo Autor</i>	<i>xviii</i>
1. Introdução	1
1.1 - Motivações para realização do trabalho	1
1.2 - Objetivos	2
1.3 - Escopo e Organização da Tese	4
2. TMN - Telecommunications Management Network	7
2.1 - Estrutura Funcional	9
2.2 - Áreas Funcionais	10
2.3 - Escopo	11
2.4 - Padronização Internacional	12
3. Bilhetes de Tarifação	15
3.1 - O que é um bilhete?	16
3.2 - Tipos de Sistemas/Bilhetes	18
3.2.1 - CDR	18
3.2.2 - IPDR	18

Sumário

3.3 - Leitura dos Bilhetes	19
3.3.1 - Formato dos Bilhetes	19
3.4 - Ambiente de Desenvolvimento	22
3.4.1 - Aplicativo para Leitura dos Bilhetes	22
3.5 - Classificação dos Bilhetes	24
3.6 - Mapeamento da Tecnologia	26
3.7 - Conclusão	27
4. Algoritmos para Processamento dos Dados e Detecção de Falhas	29
4.1 - Algoritmos para Detecção de Falhas	30
4.1.1 - Algoritmo de Tempo Real	32
4.1.2 - Algoritmo de Espaço Amostral	55
4.1.3 - Algoritmo para Chamadas OK	71
4.1.4 - Algoritmo para Detecção de Ocupação	75
4.1.5 - Processo de Poisson Homogêneo	82
4.2 - Conclusão	95
5. Algoritmos de Redes Neurais para Detecção de Falhas	97
5.1 - Redes Neurais e Reconhecimento de Padrões	98
5.1.1 - Reconhecimento de Voz	99
5.1.2 - Reconhecimento de Caracteres	99
5.1.3 - Reconhecimento de Face	99
5.1.4 - Reconhecimento de Impressões Digitais	99
5.1.5 - Detecção de Células Cancerosas	99
5.1.6 - Geração de Música	100
5.2 - Back-Propagation	100
5.2.1 - Variações sobre <i>Back-Propagation</i>	104
5.3 - Treinamento baseado no Algoritmo Back-Propagation	106
5.3.1 - Padrões	106
5.3.2 - Características	106
5.3.3 - Seleção de Características	107
5.4 - Conclusão	122
6. Curva de Abandono	124

6.1 - Curvas de Abandono para o Sistema de Telefonia Brasileiro	126
6.1.1 - Rio Grande do Sul e Santa Catarina – Telefonia Fixa	126
6.1.2 - São Paulo – Telefonia Fixa	128
6.1.3 - Considerações Gerais	129
6.2 - Conclusão	131
7. Conclusões Finais e Outras Considerações	132
7.1 - Sugestões para Trabalhos Futuros	133
7.2 - Resumo das Conclusões	134
8. Referências Bibliográficas	135
Apêndice A. Bilhetes de Tarifação	141
A.1 - CDR da Primeira Tecnologia estudada	141
A.2 - CDR da Segunda Tecnologia estudada	143
A.3 - Subeventos	145
Apêndice B. – Algoritmos para Detecção	152
B.1 - Probabilidade	152
B.1.1 - Axiomas da Probabilidade:	154
B.1.2 - Probabilidade Condicional	156
B.1.3 - Eventos Independentes	156
B.1.4 - Teorema de Bayes	156
B.1.5 - Variáveis Aleatórias	157
B.1.6 - Distribuição Uniforme Discreta	158
B.1.7 - Distribuição de Bernoulli	158
B.1.8 - Distribuição Binomial	159
B.1.9 - Distribuição Geométrica ou Binomial Negativa	159
B.1.10 - Distribuição de Poisson	159
B.1.11 - Distribuição Normal	160
B.1.12 - Variáveis Aleatórias Contínuas	161
B.1.13 - Variáveis Aleatórias Discretas	163
B.1.14 - Desigualdade de Tchebichev	164
B.1.15 - Lei dos Grandes Números	164
B.1.16 - Teoria da Renovação/Eventos Recorrentes	165
B.2 - Resultados	167

<i>Apêndice C. Sinalização entre Centrais</i>	180
C.1 - Sinalização MFC	180
C.1.1 - Sinalização Acústica	180
C.1.2 - Sinalização de Linha	180
C.1.3 - Sinalização de Registro	181
C.2 - Sinalização por Canal Comum	184
C.2.1 - “Causes” Válidos	185
<i>Apêndice D. Tempos de Ocupação do Sistema</i>	187
D.1 - Tempo de Conversação de uma Chamada OK	187
D.2 - Perfis de Ocupação do Sistema	188

Índice de Figuras

Figura 2.1 – Hierarquia em Camadas	9
Figura 3.1 – Fluxograma completo	15
Figura 3.2 – Dados sem processamento	19
Figura 3.3 – Campos Fixos do Bilhete de Tarifação.....	20
Figura 3.4 – Bilhete completo	21
Figura 3.5 – Parte fixa do bilhete	21
Figura 3.6 – Sub-registros do Bilhete.....	22
Figura 4.1 – Fluxograma	42
Figura 4.2 – Nível de qualidade de 1,7%	42
Figura 4.3 – Probabilidades.....	43
Figura 4.4 – Respostas para diferentes níveis de qualidade	46
Figura 4.5 – Desvio-Padrão em função do Nível de Qualidade	47
Figura 4.6 – Nível de Qualidade de 2%	49
Figura 4.7 – Nível de Qualidade de 7%	50
Figura 4.8 – Nível de Qualidade de 12%	51
Figura 4.9 – Nível de Qualidade de 17%	52
Figura 4.10 – Nível de Qualidade de 22%	53
Figura 4.11 – Nível de Qualidade de 27%	54
Figura 4.12 – Nível de qualidade de 1,7% (Janela 400).....	59
Figura 4.13 – Nível de qualidade de 1,7% (Janela 100).....	59
Figura 4.14 – Nível de qualidade de 1,7% (Janela 50).....	60
Figura 4.15 – Respostas para diferentes níveis de qualidade (Janela=400)	62
Figura 4.16 – Respostas para diferentes níveis de qualidade (Janela=100)	63
Figura 4.17 – Respostas para diferentes níveis de qualidade (Janela=50)	64
Figura 4.18 – Média para as diferentes janelas adotadas	66
Figura 4.19 – Desvio-Padrão para as diferentes janelas adotadas.....	66
Figura 4.20 – Janela de 50 elementos e nível de qualidade de 2%	68
Figura 4.21 – Janela de 100 elementos e nível de qualidade de 2%	69

Índice de Figuras

Figura 4.22 – Janela de 400 elementos e nível de qualidade de 2%	70
Figura 4.23 – Tempo de conversação para chamadas OK	72
Figura 4.24 – Ocupação dos circuitos	78
Figura 4.25 – Perfil de ocupação de uma central em uma grande capital (3 milhões de habitantes). 80	
Figura 4.26 – Tempo de conversação para chamada OK.....	81
Figura 4.27 – Distribuição entre ocorrências	82
Figura 4.28 – Histograma para $\lambda = 0,32$ com os intervalos de tempo sem ocorrência de chamadas 88	
Figura 4.29 – Comparação entre a Distribuição Exponencial e as chamadas de uma ERB.....	89
Figura 4.30 – Histograma para $\lambda = 0,68$ com os intervalos de tempo sem ocorrência de chamadas 90	
Figura 4.31 – Comparação entre a Distribuição Exponencial e as chamadas de uma ERB.....	91
Figura 4.32 – Interligação entre a CCC e a Plataforma.....	92
Figura 5.1 – Rede de duas camadas com alimentação direta	101
Figura 5.2 – <i>Back-Propagation</i>	103
Figura 5.3 – Comportamento da convergência da Rede Neural com 200 neurônios	108
Figura 5.4 – Comportamento da convergência da Rede Neural com 300 neurônios	109
Figura 5.5 – Comportamento da convergência da Rede Neural com 100 neurônios	110
Figura 5.6 – Comportamento da convergência da Rede Neural com 200 neurônios e $\eta=0,35$	111
Figura 5.7 – Exemplo da não-convergência de uma Rede Neural com 200 neurônios e $\eta=0,15$	112
Figura 5.8 – Rede Neural com 200 neurônios e $\eta=0,25$ e TAV=1.05	114
Figura 5.9 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.15	114
Figura 5.10 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.35	115
Figura 5.11 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.50	115
Figura 5.12 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=2,00	116
Figura 5.13 – Rede Neural com 200 neurônios e $\eta=0,25$ e $\alpha=0,1$	117
Figura 5.14 – Rede Neural com 200 neurônios e $\eta=0,25$ e $\alpha=0,2$	118
Figura 5.15 – Rede Neural com 200 neurônios e $\eta=0,25$ e TAV=1,05	121
Figura 5.16 – Comparação	122
Figura 6.1 – Taxa de AB (Abandono) Teórica.....	124
Figura 6.2 – Curva de Abandono para Rio Grande do Sul e Santa Catarina	127
Figura 6.3 – Curva de Abandono para SP – Telefonia Fixa.....	128
Figura 6.4 – Agrupamento das Curvas.....	130

Índice de Figuras

Figura 6.5 - Junção das Curvas	131
Figura 8.1 – Distribuição Normal com σ_x fixo e μ_x variável	161
Figura 8.2 – Distribuição Normal com σ_x variável e μ_x fixo	161
Figura 8.3 – Exemplo de troca de sinalização.....	183
Figura 8.4 – Sinalização por Canal Comum.....	185
Figura 8.5 – Tempo de conversação para chamada OK.....	187
Figura 8.6 – Perfis normalizados.....	189

Índice de Tabelas

Tabela 3.1 – Formato do bilhete da parte fixa.....	23
Tabela 3.2 – Formato do sub-registros das partes variáveis.....	23
Tabela 3.3 – Rótulos para classificar os bilhetes de telefonia fixa.....	25
Tabela 3.4 – Rótulos para classificar os bilhetes de telefonia celular.....	25
Tabela 4.1 – Quantidade de eventos sucessivos em função de p e N	36
Tabela 4.2 – Médias dos níveis de qualidade.....	37
Tabela 4.3 – Eventos sucessivos	38
Tabela 4.4 – Alarmes falso-positivos em função do valor de N	39
Tabela 4.5 - Níveis de qualidade.....	39
Tabela 4.6 – Eventos sucessivos	40
Tabela 4.7 – Alarmes falso-positivos em função de N	40
Tabela 4.8 – Quantidade de eventos.....	40
Tabela 4.9 – Distribuição Cumulativa de Probabilidades	44
Tabela 4.10 – Média e Desvio-Padrão	47
Tabela 4.11 – Tempos de Detecção em função do Nível de Qualidade e da Degradação	48
Tabela 4.12 – Distribuição Binomial	57
Tabela 4.13 – Média e Desvio-Padrão	60
Tabela 4.14 – Distribuição Cumulativa de Probabilidades	61
Tabela 4.15 – Média e Desvio-Padrão	65
Tabela 4.16 – Tempos de Detecção em função do Tamanho da Janela, do Nível de Qualidade e da Degradação.....	67
Tabela 4.17 – Histograma das chamadas OK em função tempo conversação	73
Tabela 4.18 – Valores para a Distribuição Binomial $c/ p=3,29%$	74
Tabela 4.19 - Valores para a Distribuição Binomial $c/ p=0,36%$	74
Tabela 4.20 – Tabela de probabilidades.....	85
Tabela 4.21 – Quantidade de eventos sucessivos em função de p e N	86
Tabela 4.22 – Probabilidade para $\lambda = 0,013$	93
Tabela 4.23 – Circuitos que apresentam problemas.....	94

Índice de Tabelas

Tabela 4.24 – Síntese dos Algoritmos.....	96
Tabela 5.1 – Tempo de convergência para 500 interações.....	119
Tabela 5.2 – Tempo de convergência para chegar ao erro de 0,0099	119
Tabela 6.1 - Taxa de AB (Abandono) Teórica.....	125
Tabela 6.2 - Relação estados/cidades	126
Tabela 6.3 - Valores de Abandono para RS e SC	127
Tabela 6.4 – Valores de Abandono para SP – Telefonia Fixa	128
Tabela 6.5 – Agrupamento das Tabelas	129
Tabela 8.1 – Classificação dos Bilhetes (eventos)	146
Tabela 8.2 – Eventos de Sistema.....	153
Tabela 8.3 – Possíveis resultados do evento DI	154
Tabela 8.4 – Montagem dos possíveis resultados	155
Tabela 8.5 – Função distribuição de probabilidade.....	157
Tabela 8.6 – Sinais para frente	182
Tabela 8.7 – Sinais para trás.....	182
Tabela 8.8 – Tráfego por Horário.....	190

Lista de Acrônimos

AB - Abandono

AMA - *Automatic Message Accounting*

AMPS - *Advanced Mobile Phone System*

ANATEL - Agência Nacional de Telecomunicações

ANSI - *American National Standards Institute*

AQL - *Acceptable Quality Level*

CCC - Central de Controle e Comutação

CCS - Cem Chamadas Segundos

CCITT - *Comité Consultatif International Téléphonique et Télégraphique*

CDMA - *Code Division Multiple Access*

CFC - *Call Final Class*

CIC - *Circuit Identification Code*

CNS - Código Nacional de Sinalização

CO - Congestionamento

CPA - Central por Programa Armazenado

CRS - Código Regional de Sinalização

CSP - Código de Seleção de Prestadora

DI - Discagem Incorreta

DP - Desligamento Prematuro

DPC - *Destination Point Code*

DT - Defeito Técnico

DTMF - *Dual Tone Multi Frequency*

EBCDIC - *Extended Binary Coded Decimal Interchange Code*

EBCH - *Equate Busy Call Hour*

EQM - Erro Quadrático Médio

ERB - Estação Rádio Base

FCAPS - *Fault, Configuration, Accounting, Performance, Security*

HMM - Hora de Maior Movimento

Lista de Acrônimos

HO - Handoff

INAP - Intelligent Network Application Part

IP - Internet Protocol

IS - Intersystem Handoff

ISUP - ISDN User Part

ITU - International Telecommunications Union

LO - Linha Ocupada

MFC - Multi Freqüencial Compelida

MLP - Multi-Layer Perceptron

MTP - Message Transfer Part

NR - Não Responde

OCRs - Optical Character Recognizers

OMAP - Operation&Maintenance Application Part

OPC - Origination Point Code

OSI - Open Systems Interconnection

PCM - Pulse Code Modulation

PGTO - Page Time Out

PSTN - Public Switch Telephone Network

PS - Ponto de Sinalização

PTS - Ponto de Transferência de Mensagens de Sinalização

QoS - Quality of Service

QRF - Queda de Rádio Freqüência

RDSI – Rede Digital de Serviços Integrados

RF - Rádio Freqüência

RNA - Redes Neurais Artificiais

RTPC - Rede de Telefonia Pública Comutada

SCC - Sinalização por Canal Comum

SCCP - Signalling Connection Control Part

SCM - Serviço de Comunicação Multimídia

SMC - Serviço Móvel Celular

SLS - Selection Long Set

Lista de Acrônimos

SP - *Signalling Point*

STFC - Sistema de Telefonia Fixa Comutada

STP - *Signalling Transfer Point*

TAV - Taxa de Aprendizado Variável

TCAP - *Transaction Capability*

TDMA - *Time Division Multiple Access*

TLDN - *Temporary Telephone Directory Number*

TMN - *Telecommunications Management Network*

TUP - *Transfer User Part*

UP - *User part*

VB - *Visual Basic*

VoIP - *Voice over Internet Protocol*

Trabalhos Publicados pelo Autor

Publicações em Revistas

1. **BREDA, Gean Davis;** MENDES, Leonardo de Souza; Bottoli, Maurício Luis. “*Failure Detection in Communication Systems*”, In: *IJCSNS International Journal of Computer Science and Network Security*, Journal ISSN: 1738-7906, VOL.7, No.10, October 2007, p.188-193.

Publicações em Congressos

2. **BREDA, Gean Davis;** MENDES, Leonardo de Souza. “*Failures Detection in Voice Communication Systems*”, In: *GLOBECOM 2006, San Francisco/CA, USA, December 2006*.
3. **BREDA, Gean Davis;** Mendes, Leonardo de Souza. “*QoS Monitoring and Failure Detection*”, In: *ITS 2006 – International Telecommunications Symposium, Fortaleza/CE, Brasil, September 2006*.
4. **BREDA, Gean Davis;** Mendes, Leonardo de Souza. “*QoS Monitoring and Fault Detection*”, In: *WINSYS 2006 - International Conference on Wireless Information Networks and Systems, Setúbal, Portugal, August de 2006*.
5. ZARPELÃO, Bruno Bogaz; MENDES, Leonardo de Souza; BOTTOLI, Maurício; **BREDA, Gean Davis;** PROENÇA JR. Mário Lemes. "Correlação de Objetos SNMP na Detecção de Anomalias em Servidores de Rede", In: XXII Simpósio Brasileiro de Telecomunicações, 2005, Campinas. Anais do SBrT'05, 2005. p. 1007-1012.
6. BOTTOLI, Mauricio Luis; **BREDA, Gean Davis;** ALBERTI, Antônio Marcos; SAKURAY, Fábio; PROENÇA Jr, Mário Lemes; Araújo, Cláudio; MENDES, Leonardo de Souza. “*Modeling of Erbium-Doped Fiber Amplifiers to Use in Optical System Simulations*”, In: *16th Modelling and Simulation IASTED - MS 2005*.
7. ALBERTI, Antônio Marcos; BOTTOLI, Maurício Luis; **BREDA, Gean Davis;** MENDES, Leonardo de Souza. “*Modeling and Simulation of ATM Traffic Management*”, In: *Proceeding of the 37th Annual Simulation Symposium, Arlington, EUA, April, 2004*.
8. ALBERTI, A. M.; BOTTOLI, M. L.; **BREDA, G. D.;** MENDES, L. S.; “Adaptação de Sequências de Tráfego Real MPEG-4 para a Transmissão em Redes ATM”, In: 3ª Conferência

Trabalhos Publicados pelo Autor

Iberoamericana em Sistemas, Cibernética e Informática, CISC2004, Orlando, Flórida, USA, 2004.

As publicações a seguir são internas à Unicamp, fruto de um trabalho desenvolvido junto a SEFAZ (Secretaria da Fazenda do Estado de São Paulo). O projeto solicitado pela SEFAZ, Monitoramento/medição de vazão de combustíveis líquidos em postos de revenda, tinha o objetivo de criar um sistema para evitar/diminuir a sonegação e também melhorar o controle sobre a qualidade do combustível. Esse projeto estava sob a tutela da Secretaria da Fazenda do Estado de SP com interesse direto do então governador Sr. Geraldo Alckmin. Uma das sugestões feitas foi a adoção de bilhetes que caracterizassem os abastecimentos. As publicações são as seguintes:

9. **BREDA, Gean Davis**; MENDES, Leonardo de Souza; BOTTOLI, Maurício Luis; Araújo, Cláudio Roberto; Costa, Max Henrique Machado. “Produto 1 - Projeto de Controle de Vazão de Combustíveis – SEFAZ”, Sumário técnico contendo descritivo técnico do sistema de medição, Unicamp, Dezembro 2005.
10. **BREDA, Gean Davis**; MENDES, Leonardo de Souza; BOTTOLI, Maurício Luis; Araújo, Cláudio Roberto; Costa, Max Henrique Machado. “Produto 3 - Transmissão de Dados dos Sistemas de Medição de Vazão de Combustíveis Líquidos em Postos de Revenda”, Relatório técnico contendo descritivo técnico do sistema de medição, Unicamp, Dezembro 2005.

Patentes Solicitadas pelo Autor

Foi dada entrada a uma solicitação de Patente junto ao INPI, nº PI 0705505-6, no dia 22/10/2007, relativa ao dispositivo denominado MMC (Multi-Mídia Center). Uma das funcionalidades do dispositivo é prover Voz sobre *IP*. Uma das contribuições foi sugerir um bilhete de tarifação específico.

Título: MMC (Multi-Mídia Center)

Objetivo: Dispositivo que realiza a interface entre uma rede banda larga e o usuário, sendo responsável por gerenciar toda a comunicação entre o dispositivo e a rede.

Solicitantes: Leonardo de Souza Mendes, **Gean Davis Breda** e Maurício Luis Bottoli.

Capítulo 1

Introdução

Essa tese concentra-se na área de detecção de falhas em sistemas de comunicações utilizando bilhetes de tarifação. A tese procura abordar desde os aspectos iniciais da utilização de bilhetes até os algoritmos para detecção de falhas, e conseqüente geração de alarmes, analisando seu desempenho e classificando-os conforme suas características. Estes algoritmos trabalham sobre as informações contidas em bilhetes de tarifação *CDR* [1], para redes convencionais (STFC – Sistema de Telefonia Fixa Comutada), e *IPDR* [2]-[5], para redes IP.

1.1 - Motivações para realização do trabalho

Boa parte da minha experiência profissional está baseada em sistemas de comunicação. Durante muitos anos trabalhei com suporte, implantação e gerenciamento de sistemas de telefonia celular e fixa. Com o passar dos anos acabei trabalhando também em projetos que buscavam a recuperação de receita em operadoras de telefonia. Esta experiência trouxe inúmeras informações a respeito da utilização de bilhetes de tarifação. No decorrer deste período fiquei convencido e impressionado pela potencialidade da utilização dos bilhetes com os mais diversos fins num sistema de telefonia. Neste trabalho procuro dar minha interpretação a tudo que aprendi e também contribuir sugerindo novas abordagens. Sendo assim, as principais motivações e objetivos que me levaram a fazer este trabalho foram:

- ✓ Contribuir para a estruturação de uma cultura de gerência baseada em bilhetes de tarifação/chamada;

Introdução

- ✓ Compilar/reunir as informações e algoritmos, buscando sempre que possível embasar o trabalho com referências bibliográficas.
- ✓ Possibilidade de inovações (publicações);
- ✓ Inúmeras possibilidades de utilização dos bilhetes;
- ✓ Entrada em operação de inúmeras empresas espelho que necessitam de soluções simples de baixo custo e alto impacto;
- ✓ Perdas não identificadas de bilhões de Reais em falhas nos sistemas de comunicações.

1.2 - Objetivos

Na análise das informações geradas por uma empresa de telecomunicações é dada ênfase a uma visão técnica e imediata que leva a empresa a desprezar importantes informações coletadas e armazenadas. Um exemplo claro disso são os bilhetes de tarifação que contêm uma vasta gama de informações sobre a rede e que são unicamente utilizados para tarifação, geração das contas telefônicas. Cada bilhete de tarifação é um registro de dados onde são armazenadas as informações relativas a uma ligação. A cada tentativa de chamada num sistema telefônico, um bilhete é gerado; nele são encontradas as informações que possibilitam a geração da conta de cobrança para os usuários. Estes bilhetes possuem uma completa gama de informações que refletem toda a história da chamada, tanto em termos de temporizações, como dos lugares por onde passou e aos quais se destinou. Alguns exemplos de informações contidas num bilhete de um sistema de telefonia celular: nome das centrais (*switches*), números dos *point codes*, número dos troncos de voz, número dos *time slots*, número das ERBs (Estações Rádio Base), número dos canais de RF (Rádio Frequência), componentes da central, número do telefone chamado e chamador, número serial do telefone, número discado, número do telefone transferido, características do telefone, tempo de início e fim da conversação, duração da ligação, tempo da sinalização, informações sobre a sinalização, tipo de resposta à ligação telefônica, o que aconteceu com a ligação, etc.

A adoção de bilhetes de tarifação permite tratar o problema da detecção de falhas de uma forma bastante completa. A quantidade de informações contidas nesses bilhetes dificilmente pode ser encontrada em outro ponto da rede, como será descrito nos próximos capítulos, de onde decorre sua importância na análise do comportamento da rede. Outra característica destes bilhetes é a sua

Introdução

confiabilidade, isto não poderia ser diferente, pois a maior parte da receita de uma operadora depende disso, o que faz com que o trabalho aqui proposto esteja embasado em informações que refletem com precisão o estado da rede. A abordagem através de bilhetes permite também localizar problemas que dificilmente seriam identificados utilizando outros modelos. Um exemplo é a detecção de problemas em centrais de telefonia em outras cidades, estados, países, bastando para isso que existam chamadas para esses locais.

O objetivo central da detecção de falhas com uma maior eficiência é diminuir as deficiências na *QoS (Quality Of Service)* [6] e, conseqüentemente, as perdas de receita nos sistemas de comunicação. Dados de empresas de consultoria e institutos internacionais de pesquisa de mercado indicam que, no ano de 1999, a perda de faturamento das prestadoras de serviços de Telecomunicações no mundo chegou a US\$ 15 bilhões [7]. Ainda segundo esse estudo, boa parte das operadoras (33%, conforme um dos levantamentos) nem sabem que estão perdendo receitas. Há um número expressivo de empresas que nem sequer conseguem identificar a origem da evasão. No Brasil, esse número deve ficar na casa de R\$ 2,5 bilhões por ano, algo em torno de 3% a 15% do faturamento. Ainda segundo esta pesquisa já é possível reduzir as perdas a, no máximo, 2% da receita da operadora. E, em 70% dos casos, isso pode ser feito com soluções de baixo custo, alto impacto e rápida implementação.

Sendo assim, a utilização dos bilhetes em conjunto com os algoritmos pode ajudar a diminuir as perdas, bem como a diminuir as reclamações associadas a uma deficiente qualidade de serviço (*QoS*) [8]. O objetivo dos algoritmos é detectar e corrigir falhas no sistema de telecomunicação, diminuindo a perda de receita e melhorando a qualidade do serviço ofertado.

Em termos gerais os algoritmos são os seguintes:

- Algoritmo sobre Redes Neurais utiliza a rede do tipo *Back-Propagation* para detectar falhas. A rede neural é inicialmente treinada e após este treinamento está apta para procurar por anormalidades nos sistemas de comunicações;
- Algoritmo de Tempo Real que consiste em analisar as saídas das chamadas esperando por uma sequência de eventos. É como se estivéssemos “jogando moedas aleatoriamente” e esperando por uma determinada sequência para confirmarmos um alarme;
- Algoritmo de Espaço Amostral no qual esperamos pelo preenchimento da janela de dados para então procurarmos por uma anomalia utilizando a distribuição binomial;

- Algoritmo para Chamadas Ok procura por anormalidades dentro de um tipo específico de chamada, ou melhor, chamadas que a princípio foram concluídas com sucesso. Este algoritmo engloba os dois mecanismos citados anteriormente: Tempo Real e Chamadas OK;
- Algoritmo para Detecção de Ocupação que trabalha sobre o tempo de ocorrência dos eventos para gerar o alarme. O algoritmo se baseia na distribuição de Poisson.

1.3 - Escopo e Organização da Tese

Este trabalho procura seguir um encadeamento lógico do assunto, definindo nos capítulos iniciais o que é um bilhete de tarifação, como é sua leitura e classificação, para depois entrar nos algoritmos de detecção. A parte final é dedicada a um estudo inicial para estimar as perdas geradas pelos problemas detectados.

Em síntese, a tese é composta de oito capítulos mais Bibliografia e Anexos:

No Capítulo 1, Introdução, é apresentada uma visão geral sobre o problema estudado.

No Capítulo 2, *TMN – Telecommunications Management Network*, falamos a respeito de uma arquitetura organizada que possibilita a integração e interoperabilidade entre vários tipos de sistemas de operação e os equipamentos de telecomunicações, utilizando modelos genéricos de rede para a gerência, modelos genéricos de informações com interfaces e protocolos padronizadas. Procuramos enquadrar o trabalho dentro desta metodologia.

No Capítulo 3, Bilhetes de Tarifação, estes são apresentados e conceituados. Primeiramente, são definidas as similaridades entre os bilhetes de tarifação da telefonia convencional e os da telefonia de Voz sobre IP. Logo após são introduzidos os conteúdos dos bilhetes de dois fabricantes distintos. Em seguida é feita a leitura dos bilhetes a partir dos dados no seu formato original. É apresentado o aplicativo de leitura e posteriormente são comentados os problemas encontrados na leitura.

Posteriormente, são definidas as possíveis classificações que cada bilhete pode assumir, dependendo do que acontece com a ligação. O resultado da classificação são os eventos. A partir dos eventos é possível procurar por determinadas falhas no sistema.

No Capítulo 4, Algoritmos para Processamento de Dados e Detecção de Falhas, são apresentados os algoritmos utilizados na detecção de falhas. Para cada algoritmo é apresentado seu modelo teórico e os resultados e conclusões pertinentes ao seu funcionamento e aplicação.

Introdução

No Capítulo 5, Algoritmos de Redes Neurais para Detecção de Falhas, apresenta-se a abordagem adotada em relação a Redes Neurais. Procura-se mostrar os diversos experimentos realizados com o objetivo de escolher o melhor modelo e, para isso, são comparados seus resultados.

No Capítulo 6, Curva de Abandono, é demonstrado o comportamento do usuário em função da retentiva nas ligações telefônicas. Com esse objetivo é apresentada a curva de abandono levantada para o sistema brasileiro de telefonia.

No Capítulo 7, Conclusões Finais e outras Considerações, a tese é finalizada descrevendo as contribuições e apontando direções para os trabalhos futuros.

Capítulo 2

TMN - Telecommunications Management

Network

Os sistemas de telecomunicações têm evoluído continuamente o que nos leva a um ambiente composto por redes e equipamentos heterogêneos. Esta heterogeneidade resulta na complexidade de se gerenciar o ambiente. As soluções proprietárias de cada fabricante tornam a gerência de redes ineficiente, envolvendo problemas como, por exemplo, a presença de múltiplas interfaces para diferentes sistemas, sistemas não interoperáveis, insuficiência de informações coletadas, bases de dados específicas e isoladas com informações redundantes e inconsistentes. A gerência integrada de rede exige que os fabricantes forneçam seus produtos dentro dos padrões estabelecidos por recomendações com interfaces padronizadas que possibilitam a interação com sistemas de gerenciamento de outras empresas.

Em 1985, a *ITU (International Telecommunications Union)* [9] iniciou estudos sobre a padronização da gerência das redes de telecomunicações, criando um conceito básico denominado Rede de Gerência de Telecomunicações (*TMN – Telecommunications Management Network*).

A *TMN* foi desenvolvida com o propósito de gerenciar redes, serviços e equipamentos heterogêneos, operando sobre os mais diversos fabricantes e tecnologias que já possuem alguma funcionalidade de gerência.

Em 1988 os estudos dos grupos de trabalho do *ITU-T* resultaram nas recomendações M.3000 [10] que definem os princípios básicos da *TMN*. Estas recomendações fornecem uma estrutura organizada que permite interconectar sistemas de suporte a operação e diversos tipos de equipamentos de telecomunicações possibilitando a interoperabilidade entre sistemas de gerência.

Veja o conjunto de recomendações da série M.3000:

- M.3010 – *Principles for a TMN* [11];
- M.3020 – *TMN Interface Specification Methodology* [12];
- M.3100 – *Generic Network Information Model* [13];
- M.3180 – *Catalogue of TMN Management Information* [14];
- M.3200 – *TMN Management Services* [15];
- M.3300 – *TMN Management Capabilities Presented at the “F” Interface* [16];
- M.3400 – *TMN Management Functions* [17].

O princípio por de trás da *TMN* é proporcionar uma arquitetura organizada, que possibilite a integração e interoperabilidade entre diferentes tipos de sistemas e equipamentos de telecomunicações, utilizando modelos genéricos de rede para a gerência, e de informações com interfaces e protocolos padronizadas de forma a criar um conjunto de padrões para administradores e fabricantes.

A *TMN* parte do princípio que as redes e os serviços de telecomunicações são um conjunto de sistemas cooperativos e gerencia-os de forma harmônica e integrada. A *TMN* interage com a rede de telecomunicações em vários pontos, através de interfaces padronizadas, podendo utilizar parte da rede de telecomunicações para realizar suas funções.

Algumas redes e serviços que podem ser gerenciadas pela *TMN* são:

- Redes públicas e privadas, incluindo a RDSI, redes de telefonia móvel, redes privadas de voz e redes inteligentes;
- Elementos de transmissão (multiplexadores, roteadores, cross-connects, equipamentos SDH);
- Sistemas de transmissão analógica e digital baseados em cabos coaxiais, fibra óptica, rádio e enlace de satélite;
- Mainframes, processadores *front-end*, controladoras remotas, servidores de arquivos, etc;
- Redes locais, geográficas e metropolitanas (*LAN*, *MAN* e *WAN*);
- Redes de comutação de circuito e pacotes;
- A própria *TMN*;
- Terminais e sistemas de sinalização incluindo Pontos de Transferência de Sinalização (*STP*) e bases de dados em tempo real;
- Serviços de suporte e teleserviços;

- Sistemas de infra-estrutura e suporte, como módulos de teste, sistemas de energia, unidades de ar condicionado, sistemas de alarme, etc.

2.1 - Estrutura Funcional

A estrutura funcional é a subdivisão da gerência em níveis ou camadas funcionais que restringem as atividades de gerência ao escopo de cada camada, veja a hierarquia na Figura 2.1.

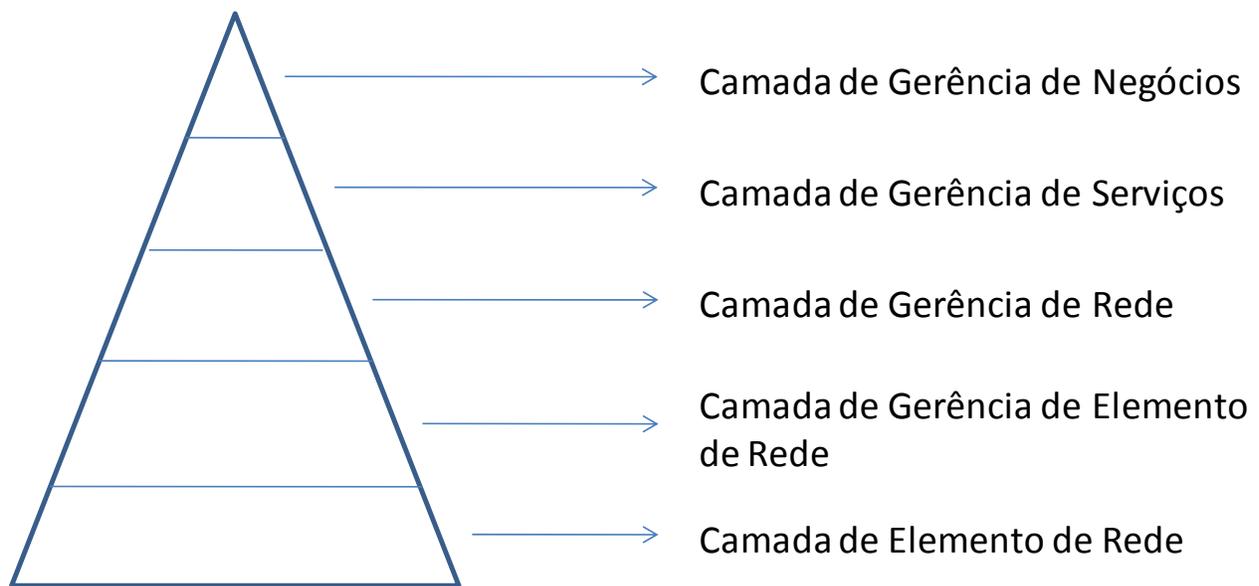


Figura 2.1 – Hierarquia em Camadas

A seguir estão detalhadas as camadas:

- Camada de Elemento de Rede - Corresponde às entidades de telecomunicações (software ou hardware) que precisam ser efetivamente monitorados e/ou controlados. Estes equipamentos devem possuir agentes para que possam fornecer as informações necessárias ao sistema de gerência, como coleta de dados de desempenho, monitoração de alarmes, coleta de dados de tráfego, etc.
- Camada de Gerência do Elemento da Rede - É responsável pelo gerenciamento dos elementos de rede, coordenando e controlando os mesmos, possibilitando manutenção preventiva e obtenção de dados relativos ao funcionamento dos elementos. Geralmente fica restrita a equipamentos de mesma tecnologia e fabricante ou na mesma região de operação.

- Camada de Gerência de Rede - Faz o gerenciamento da rede suportada pela camada de elemento de rede, fornecendo uma visão integrada para a operadora de telecomunicações no que se refere à conectividade e rotas.
- Camada de Gerência de Serviço - Esta camada relaciona os aspectos de interface com os clientes, e realiza funções como provisionamento de serviços, abertura e fechamento de contas, resolução de reclamações dos clientes (inclusive relacionados a tarifação), relatórios de falhas e manutenção de dados sobre qualidade de serviço (*QoS*).
- Camada de Gerência de Negócio - É responsável pela gerência do empreendimento, envolvendo todos os aspectos de prestação de serviços, possibilitando entre outras funções o gerenciamento de administração, organização e manutenção.

2.2 - Áreas Funcionais

De forma a se englobar toda a funcionalidade necessária ao gerenciamento de uma rede de telecomunicações (planejamento, instalação, operação, manutenção e provisionamento), identificaram-se cinco áreas funcionais:

- Gerência de Desempenho - Provê funções para avaliar e relatar o comportamento dos equipamentos de telecomunicações e a eficiência da rede. Principais funções: monitoração de desempenho, medidas de tráfego. Esta funcionalidade pode ser considerada a mais importante em um sistema de gerenciamento, pois é por meio dela que a operadora realiza o tratamento preventivo de sua rede. Esta funcionalidade garante os níveis de *SLA (Service Level Agreement)* da rede contratados pelos clientes.
- Gerência de Falhas - Possibilita a detecção, isolamento e a correção de uma operação anormal da rede. Principais funções: supervisão de alarmes, testes, relatórios de problemas. Esta funcionalidade permite à operadora identificar com precisão qual a falha em sua rede que está causando ou poderá causar a parada de determinado serviço. Permite também identificar a provável causa desta falha.
- Gerência de Configuração - Fornece funções para a modificação dos recursos físicos e lógicos da rede. Principais funções: configuração de recursos, gerenciamento da ordem de serviço, informações de recursos. Esta funcionalidade garante a agilidade que a operadora necessita para configurar os equipamentos e serviços disponíveis aos seus clientes.

- Gerência de Contabilização - Fornece funções que possibilitam determinar o uso da rede. Desta forma, a gerência de contabilidade permite com que a operadora possa tarifar estes serviços e controlar as receitas provenientes destes serviços.
- Gerência de Segurança - Fornece funções que permitem assegurar a integridade da rede. Principais funções: segurança de acesso, alarmes de segurança e segurança de dados. Controla todo tipo de intrusão seja física ou lógica. Também preserva os dados da rede da operadora. Essas áreas funcionais são muitas vezes referenciadas pela sigla *FCAPS (Fault, Configuration, Accounting, Performance, Security)*.

2.3 - Escopo

O escopo do trabalho sugerido nesta tese baseia-se na *TMN*. Como será verificado posteriormente, o trabalho está enquadrado na Camada de Elemento de Rede. As áreas funcionais correlatas são: Gerência de Desempenho e Gerência de Falhas. O gerenciamento de desempenho envolve as funções relacionadas com a avaliação, relato do comportamento dos equipamentos e com a eficiência da rede. A utilização de bilhetes propicia justamente o acompanhamento do comportamento dos diversos equipamentos ou elementos do sistema, sejam eles físicos ou lógicos. Podemos ver a seguir alguns elementos que podem ser monitorados. No Capítulo 4 daremos mais detalhes em relação aos elementos do sistema.

- Circuitos de Entrada (elemento físico);
- Circuitos de Saída (elemento físico);
- Transceptor MFC (elemento físico);
- Rotas de Entrada (elemento físico);
- Rotas de Saída (elemento físico);
- BSC (elemento físico);
- BTS origem (elemento físico);
- BTS destino (elemento físico);
- ERB origem (elemento físico);
- ERB destino (elemento físico), etc.

É possível monitorar os mais diferentes comportamentos de cada elemento, conforme é mostrado no Anexo A, item A.3 - Subeventos, Tabela 8.1.

Em relação à Gerência de falhas os algoritmos trabalham na monitoração dos elementos e conseqüente detecção de anormalidades. Com as informações geradas na detecção é possível trabalhar na análise do problema e que de imediato pode conduzir a localização do motivo da falha, possibilitando a remoção do problema.

2.4 - Padronização Internacional

Estes são alguns dos organismos internacionais, envolvidos nas atividades de padronização nos sistemas de telecomunicações:

- *ISO (International Standardization Organization)* - A *ISO* é o organismo internacional de padronização gerador dos padrões mundiais que estabelecem regras que normalizam as principais atividades de desenvolvimento. Todos os órgãos geradores de padrões devem ser filiados a *ISO*.
- *ITU (International Telecommunications Union)* - O *ITU-T* é um órgão da ONU (Organização das Nações Unidas) cuja função é emitir recomendações relativas à telefonia e telegrafia. Está dividido em grupos de estudos (SG), grupos de trabalho (WP) e subgrupos de trabalho (SWP).
- *ETSI (European Telecommunications Standards Institute)* - O Instituto *ETSI* é um gerador de padrões na área de telecomunicações na Europa.
- *ANSI (American National Standards Institute)* - O *ANSI* é uma organização não governamental, composta de fabricantes e usuários que gera padrões nos Estados Unidos. É membro da *ISO* e possui grupos na área de gerência *TMN*, transmissão e comutação.
- *EURESCOM (European Institute for Research and Strategic Studies in Telecommunication)* - O Instituto *EUROSCOM* possui projetos na área de *TMN* com o objetivo de resolver problemas de gerência de rede e serviços na Europa.
- *TINA-C (Telecommunications Information Networking Architecture Consortium)* - Consórcio formado em 1993 por operadoras de telecomunicações e vendedores de computadores. Define uma arquitetura baseada em processamento distribuído e busca definir uma arquitetura para serviços para as redes de telecomunicação de faixa larga.
- *NMF (Network Management Forum agora chamado TMF – Telecommunication Management Forum)* - Consórcio fundado em 1988 pela AT&T, British Telecom,

Telecom Canadá, Hewlett Packard, Nortel, STC e Unisys. Possui centenas de participantes e tem como missão promover, acelerar e tornar possível a implementação de sistemas de gerência baseados nas padronizações. Produz especificações e acordos, orientados a negócios, com nível de detalhe suficiente para que desenvolvedores possam produzir produtos e serviços, que possam ser automatizados e integrados, garantindo o inter funcionamento na cadeia de produção.

Capítulo 3

Bilhetes de Tarifação

Na Figura 3.1 é apresentado o fluxograma que contém todos os passos seguidos, desde a geração do bilhete, que acontece quando existe uma tentativa de chamada, até a obtenção dos resultados após a aplicação dos algoritmos.

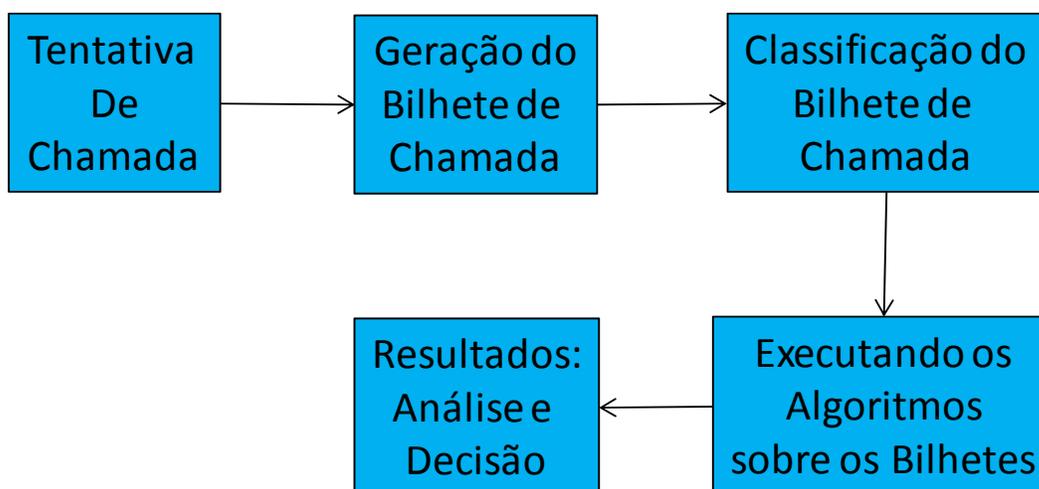


Figura 3.1 – Fluxograma completo

Conforme podemos ver através desta figura, após sua geração o bilhete precisa ser classificado. Esta classificação, que denominaremos evento, é a representação do que aconteceu com a ligação telefônica. É como colocar um rótulo para cada tipo de terminação possível. A classificação é necessária para que se possa identificar o comportamento sobre os diversos elementos da rede de telecomunicações pelos quais as chamadas passam.

Os algoritmos são empregados para monitorar os diferentes elementos de um sistema de comunicação. Para cada elemento são monitorados todos os possíveis eventos. A detecção de uma falha é feita sobre a análise dos eventos, detectando anormalidades no comportamento do elemento.

Estas anormalidades são detectadas comparando-se os valores atuais com os valores-padrão dos eventos. Os algoritmos buscam justamente isso, maneiras eficientes de fazer essa comparação, sempre buscando maximizar a taxa de acertos.

3.1 - O que é um bilhete?

Um bilhete de tarifação é um registro de dados onde são armazenadas as informações relativas a uma ligação telefônica. O bilhete contém as mais diversas informações sobre a chamada telefônica, tais como: nome das centrais (*switches*), números dos *point codes*, número dos troncos de voz, número dos *time slots*, número das ERBs, número dos canais de RF, componentes da central, número do telefone chamado e chamador, número serial do telefone, número discado, número do telefone transferido, características do telefone, tempo de início e fim da conversação, duração da ligação, tempo da sinalização, informações sobre a sinalização, tipo de resposta à ligação telefônica, o que aconteceu com a ligação, etc. Não existe norma que regulamente a criação do bilhete *CDR*, cada fabricante [18][10]-[22](Ericson, Nortel, Motorola, Lucent, Alcatel, Trópico, ZTE) adota seu padrão.

Existem poucas publicações disponíveis a respeito do *CDR*. Essa dificuldade pode ser entendida, visto que esses dados são, muitas vezes, tratados com sigilo, tanto por parte dos fornecedores de equipamentos, que procuram restringir as informações sobre as estratégias de criação e conteúdo dos bilhetes, bem como por parte das operadoras, que tratam essas informações com discrição, pois estão diretamente relacionadas à receita da empresa.

Foram encontrados trabalhos que utilizam o *CDR* para monitorar o tráfego telefônico [23][24]. Em outros trabalhos, o *CDR* é utilizado para tarefas mais complexas como, por exemplo, analisar aspectos econômicos e sociais dos usuários do sistema [25]. Esse tipo de análise pode ser feita, uma vez que cada ligação telefônica possui um correspondente bilhete de tarifação, tornando possível analisar o comportamento dos usuários do sistema, pois o bilhete possui várias informações relativas ao cliente e à ligação telefônica.

Existem também trabalhos que utilizam o *CDR* em detecções de fraudes [26][27]. Nesses, os *CDRs* são utilizados para traçar um perfil do usuário, utilizando-se das informações contidas nos bilhetes. Em função desse perfil, faz-se uma detecção do desvio no comportamento.

Encontramos também publicações que utilizam *CDRs* em *Data Mining* [28][29]. Outros utilizam a base de dados de *CDR* conjuntamente com o Modelo Bayesiano [30]-[34] para inferência

sobre o comportamento de uma rede de telecomunicações. Uma Rede Bayesiana é caracterizada por um grafo, composto por vários nós, em que cada nó da rede irá representar uma variável aleatória, ou seja, um atributo da base de dados.

Em relação ao *IPDR* o campo de pesquisa está em aberto, havendo oportunidades para trabalhar nas mais diversas áreas. A tecnologia *VoIP* (*Voice over Internet Protocol*) é recente no Brasil. Por exemplo, não existe normatização específica dessa tecnologia pela ANATEL (Agência Nacional de Telecomunicações) [35]. Isso acontece porque a regulamentação dos serviços de voz não especifica a tecnologia a ser utilizada, e sim o tipo de serviço a ser prestado pelos provedores que são divididos em dois grupos:

- STFC (Sistema de Telefonia Fixo Comutado) - prestado como serviço público de voz. É a rede tradicional de comutação de circuitos, com acessos analógicos, projetada basicamente para o serviço de telefonia e que provê suporte restrito para comunicação de dados em faixa de voz. Aspectos favoráveis: É um serviço de telecomunicações, possui uma regulamentação estruturada, possui requisitos de qualidade, universalização e continuidade, qualquer suporte tecnológico ao serviço (IP), possui modelo de Longa Distância (CSP), possui numeração, possui interconexão. Aspectos desfavoráveis: requer investimentos altos, requer o envio de relatórios para ANATEL, lidera as reclamações dos consumidores nos órgãos de defesa do consumidor [35].
- SCM (Serviço de Comunicação Multimídia) - prestado como parte dos serviços multimídia. O Serviço de Comunicação Multimídia é um serviço fixo de telecomunicações de interesse coletivo, que possibilita a oferta de capacidade de transmissão, emissão e recepção de informações multimídia a assinantes dentro de uma área de prestação de serviço, utilizando quaisquer meios. Aspectos favoráveis: Trabalha com dados multimídia, é um serviço de telecomunicações, possui uma regulamentação mais livre, requer investimento de baixo custo, não é obrigado a seguir metas de qualidade, não têm custo de interconexões. Aspectos desfavoráveis: Requisitos de qualidade ainda não regulados, mercado competitivo, sem um modelo de longa distância (CSP), não possui numeração, interconexão (sem direito de receber tarifa de interconexão) [35].

Essas licenças têm públicos distintos e não devem ser confundidas entre si. Cada tipo de licença oferece ao provedor um público específico e é dentro desse contexto que ele deve operar seus serviços, seja usando a tecnologia de telefonia convencional (STFC), seja usando a tecnologia

VoIP. Existem, entretanto, requisitos adicionais que diferenciam os serviços e as responsabilidades dos operadores. Para o STFC os requisitos de numeração, cobertura, interconexão e qualidade de serviços são bastante rígidos, e a obtenção da licença é mais complexa. Para o SCM os requisitos são voltados aos serviços multimídia, permitindo a interconexão e uso de numeração. A tecnologia *VoIP* aos poucos ganha espaço dentro do contexto nacional [36], abrindo diversas oportunidades em relação à gerência de rede.

3.2 - Tipos de Sistemas/Bilhetes

O intuito das explicações a seguir é demonstrar que não existem diferenças lógicas significativas em relação aos tipos de bilhetes. O resultado do trabalho poderá ser aplicado para vários sistemas, sejam eles de uma rede IP, de uma rede de telefonia celular, de uma rede de telefonia fixa, etc.

3.2.1 - *CDR*

No Apêndice A, itens A.1 e A.2, são apresentados dois exemplos de *CDRs* provenientes de sistemas reais, onde podem ser analisadas as similaridades com a estrutura de um *IPDR*. Essas estruturas de *CDRs* foram extraídas de centrais de comutação celular de dois fabricantes distintos.

Nos dois exemplos citados, é possível ver que os bilhetes de tarifação seguem os cinco conceitos que serão apresentando a seguir: “Quem?”, “Quando?”, “Onde?”, “Por quê?”, “O Quê?”. Todos os campos descritos para as duas tecnologias informam basicamente a identificação do assinante A e B, suas características, os componentes que estão envolvidos na chamada, tanto na central A como na B, a indicação do tempo de começo, fim, duração, as características, tipo da ligação, o que aconteceu com a ligação telefônica e por onde ela passou.

3.2.2 - *IPDR*

Um bilhete *IPDR* deve ser capaz de caracterizar qualquer tipo de ligação telefônica dentro de um sistema de telefonia baseado em redes IP. Existem cinco atributos básicos [37] que um bilhete *IPDR* deve conter, ou seja, ele deve ser capaz de fornecer resposta às questões:

- “Quem?” - Identificação dos participantes, origem e destino;
- “Quando?” - As temporizações referentes ao início e final das transações, tempo de originação, tempo necessário para completar a ligação, da terminação (*End Time*), tempo

da requisição de conexão, tempo da conexão, duração da conexão, etc. Os tempos devem estar num formato *ISO 8601*, facilitando a troca de informações;

- “Onde?” - O progresso da transação, ou seja, para onde a transação foi enviada e os elementos envolvidos. Quais foram os elementos das centrais que foram ocupados;
- “Por quê?” - Os finais de seleção, ou seja, o que aconteceu com a transação, se ela foi completada com sucesso, se caiu antes de atingir seu destino, se o assinante chamado não respondeu, se a rede estava congestionada, etc.;
- “O Quê?” - O tipo da ligação que está sendo feita, ou seja, livre de tarifação, tarifação normal, pré-paga, tarifação na origem ou destino, etc.

3.3 - Leitura dos Bilhetes

Os bilhetes de tarifação são criados num formato proprietário de cada fabricante. Seus dados são geralmente dispostos de maneira contínua num arquivo de dados. O primeiro passo é extrair as informações de maneira ordenada, ou seja, dispor num formato de tabela, onde cada coluna tenha o seu valor correspondente, e cada linha corresponda a um bilhete. Esse tipo de tabela possibilita que sejam feitas análises sobre os dados. A Figura 3.2 mostra um exemplo de como as informações estão no seu estado original.

```
XXKK00412c170072424c0072446c0000022c00e003040100000001000007010YY91  
111111cdf111100001700e00304040072435c0072446c0000011c170579009405ff  
fff55YY9811111206ffffffffffffffffffffffffffff9111111207ffffffffffffffffff  
fffffffff0yy911111120700ZZ
```

Figura 3.2 – Dados sem processamento

A seguir é explicado o modelo utilizado na leitura dos *CDRs* baseando-se no bilhete apresentado no Apêndice A, item A.2 - *CDR* da Segunda Tecnologia estudada.

3.3.1 - Formato dos Bilhetes

O bilhete dessa tecnologia é formado por uma parte fixa e por partes variáveis, chamadas sub-registros, que serão montadas conforme a aplicação do bilhete. Todos os bilhetes usam o mesmo conjunto de sub-registros, porém nem todos os sub-registros podem estar presentes em todos os tipos de bilhetes.

A seguir explicaremos o mecanismo de formação dos bilhetes. Esse bilhete possui um tamanho fixo de 30 bytes, sem contar os sub-registros.

Bilhetes de Tarifação

O início do bilhete, dentro de um arquivo, é sempre marcado pela seqüência “XX” e “00”. A Figura 3.3 exibe o tamanho e nome de apenas alguns campos para a parte fixa do bilhete.

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Início	Tipo	Data			Switch	Horário início				Horário fim			
		mês/dia				hora/minuto/segundo				hora/minuto/segundo			
XX	00	0DDDDH			HH	DDDDDDDH				DDDDDDDH			

Figura 3.3 – Campos Fixos do Bilhete de Tarifação

Conforme vemos, cada coluna contém três informações:

- A primeira informa o número do byte ao qual a informação se refere;
- A segunda fornece o nome do campo e, correlacionando com a linha superior, informa a quantidade de bytes que determinado campo terá. Exemplo: O campo DATA possui três bytes (Byte 3, 4 e 5);
- A terceira informa o formato de cada byte.

O último byte do bilhete da seqüência fixa traz a informação da quantidade de sub-registros que existirá no bilhete. Na seqüência, o próximo byte faz parte do sub-registro, e cabe a ele fazer a identificação deste. O significado de cada campo foi explicado no final no item A.2 - CDR da Segunda Tecnologia estudada.

Existem cerca de 50 tipos de sub-registros possíveis. É importante esclarecer que na formação de um bilhete dificilmente serão encontrados todos os sub-registros existentes. Usualmente a quantidade fica em torno de 7 a 8, ou seja, a parte fixa e mais 7 a 8 sub-registros. Assim como a parte fixa do bilhete, cada sub-registro é definido segundo uma estrutura.

Suponha que um bilhete de tarifação seja gerado e que seja composto, além da parte fixa, por quatro sub-registros, assim configurados:

- Trinta bytes referentes à parte fixa do bilhete, em que o primeiro e segundo bytes representam o início do bilhete e o tipo de ligação telefônica “XXKK”, e o último byte contém as informações relativas à quantidade de sub-registros, que no exemplo apresentado é igual a quatro;
- Dezoito bytes relativos ao primeiro sub-registro, sendo o primeiro byte relativo à identificação do sub-registro;
- Dezoito bytes relativos ao segundo sub-registro, sendo o primeiro byte relativo à identificação do sub-registro;

Bilhetes de Tarifação

- Oito bytes relativos ao terceiro sub-registro, sendo o primeiro byte relativo à identificação do sub-registro;
- Dezesete bytes relativos ao quarto sub-registro, sendo o primeiro byte relativo à identificação do sub-registro;

O bilhete será montado de forma seqüencial, ou seja, primeiro a parte fixa e depois os sub-registros. Toda vez que terminarmos de ler um sub-registro, o próximo byte será a identificação do próximo sub-registro, e assim por diante, até terminar a quantidade de sub-registros relativos ao bilhete.

3.3.1.1 - Exemplo de um Bilhete

Na Figura 3.4 é apresentado um exemplo de bilhete. Este bilhete é composto pela parte fixa e pelos sub-registros.

```
XXKK00412c170072424c0072446c0000022c00e003040100000001000005010YY91  
111111cdf111100001700e00304040072435c0072446c0000011c170579009405ff  
fff55YY9811111206ffffffffffffffffffffffff9111111207ffffffffffffffff  
ffffffff0yy911111120700ZZ
```

Figura 3.4 – Bilhete completo

A Figura 3.5 detalha a parte fixa do bilhete.

Parte Fixa - Bilhete de tarifação

```
XXKK 00412c 17 0072424c 0072446c 0000022c 00e00304 01 000000 01 00  
00 05
```

Figura 3.5 – Parte fixa do bilhete

A Figura 3.6 ilustra os sub-registros de um bilhete. Os itens (A), (B), (C), (D), (E) representam os sub-registros, e o item (F) apresenta os caracteres do final do bilhete.

A)Sub-registro – Identificação do Telefone Originador

AA 0YY91111111c df111111 00 00 17 00e0 0304

B)Sub-registro – Identificador da Terminação

BB 0072435c 0072446c 0000011c 17 0579 0094

C)Sub-registro – Número de Identificação

CC fffff55yy98111112

D)Sub-registro – Dígitos discados

DD ffffffffffffffffffffffffffffff91111112

E)Sub-registro – Número chamado

EE ffffffffffffffffffffffffffffff0YY91111112

F)Representa o final do bilhete

00ZZ

Figura 3.6 – Sub-registros do Bilhete

3.4 - Ambiente de Desenvolvimento

O aplicativo de leitura dos bilhetes é uma das contribuições da tese e foi desenvolvido utilizando a linguagem VB 6 [38], orientada a banco de dados na ferramenta MS-Access 2000 [39]. A ferramenta MS-Access propicia inúmeras funções já pré-programadas, que facilitam o trabalho com banco de dados: Criação de tabelas, concatenação de tabelas, elaboração de consultas, elaboração de consultas multi-tabelas, criação de filtros, ordenação, utilização de critérios “OR” e “AND”, adição/remoção de campos, criação de índices, estabelecimento de relacionamentos, criação de vínculos nos relacionamentos, modificação das propriedades da tabela, criação de formulários, criação de relatórios, etc.

Associando esse ambiente à capacidade de programação, é possível ter um ambiente propício para explorar o potencial das informações contidas nos bilhetes.

3.4.1 - Aplicativo para Leitura dos Bilhetes

O aplicativo foi desenvolvido de forma a atender diferentes tecnologias. Foram criadas duas tabelas: uma com a estrutura da parte fixa do bilhete, Tabela 3.1, e a outra para a estrutura da parte variável, Tabela 3.2. Caso o bilhete de determinado fabricante não possua partes variáveis, a tabela relativa a essas partes não será utilizada.

Bilhetes de Tarifação

Campo	Tipo	Nome campo	Tamanho	Forma	Usar	Ler
1	R_00	INICIO	2	A	SIM	SIM
2	R_00	TIPOBILHETE	2	A	SIM	SIM
3	R_00	DATA	6	D	SIM	SIM
4	R_00	...	2	A	SIM	SIM
5	R_00	...	8	T	SIM	SIM
6	R_00	...	8	T	SIM	SIM
7	R_00	...	8	X	SIM	SIM
8	R_00	...	4	I	SIM	SIM
9	R_00	...	4	I	SIM	SIM
10	R_00	...	2	A	SIM	SIM
11	R_00	...	6	A	SIM	SIM
12	R_00	...	2	A	SIM	SIM
13	R_00	...	2	A	SIM	SIM
14	R_00	...	2	A	NÃO	SIM
15	R_00	...	2	A	NAO	SIM

Tabela 3.1 – Formato do bilhete da parte fixa

Campo	Tipo	Nome campo	Tamanho	Forma	Usar	Ler
1	S_01	...	2	A	NAO	SIM
2	S_01	...	12	F	SIM	SIM
3	S_01	...	8	A	SIM	SIM
4	S_01	...	2	A	SIM	SIM
5	S_01	...	2	A	SIM	SIM
6	S_01	...	2	A	SIM	SIM
7	S_01	...	4	I	SIM	SIM
8	S_01	...	4	I	SIM	SIM
1	S_02	...	2	A	NAO	SIM
2	S_02	...	12	F	SIM	SIM
3	S_02	...	8	A	SIM	SIM
4	S_02	...	2	A	SIM	SIM
5	S_02	...	2	A	SIM	SIM
6	S_02	...	2	A	NAO	SIM
7	S_02	...	4	I	SIM	SIM
8	S_02	...	4	I	SIM	SIM
1	S_03	...	2	A	NAO	SIM
2	S_03	...	2	A	SIM	SIM
3	S_03	...	12	F	SIM	SIM
1	S_04	...	2	A	NAO	SIM
2	S_04	...	8	T	SIM	SIM
3	S_04	...	8	T	SIM	SIM
4	S_04	...	8	X	SIM	SIM
5	S_04	...	2	A	SIM	SIM
6	S_04	...	4	I	SIM	SIM
7	S_04	...	4	I	SIM	SIM
8	S_04	...	2	Y	SIM	SIM
...

Tabela 3.2 – Formato do sub-registros das partes variáveis

Na Tabela 3.2 não estão listados todos os sub-registros possíveis, devido ao tamanho que esta tabela assumiria. As tabelas são compostas basicamente das seguintes colunas:

- Campo – Nada mais é do que um número seqüencial que indica a ordem a ser seguida na leitura do bilhete;
- Tipo – Define o tipo de registro ou sub-registro;

- Nome do campo – É puramente ilustrativo, não tem significado no programa, apenas para controle interno;
- Tamanho – Coluna que informam o tamanho do campo a ser lido. O valor presente nesse campo deve ser multiplicado por 4 bits para termos o valor total, ou seja, quando tivermos o valor dois, teremos oito bits que equivalem a um byte.
- Forma – Indica como o valor extraído deverá ser lido/gravado;
- Usar – Indica se o campo deverá ser gravado ou não;
- Ler – Indica se o campo do bilhete original deverá ser lido ou não.

O que o programa faz com essas tabelas é encontrar o início do bilhete e o tipo dele e seguir campo a campo até o último, chamado “partesvariáveis”. Uma vez definida a quantidade de sub-registros, o programa começa a ler um a um, conforme Tabela 3.2. Após o último, teremos os caracteres de término do bilhete, definidos anteriormente como “00ZZ”.

3.5 - Classificação dos Bilhetes

Uma vez que os bilhetes tenham sido lidos, o próximo passo é classificá-los. Classificar nada mais é do que colocar um rótulo em cada bilhete dizendo à qual classe, evento, de ligação ele pertence. Na verdade, o que criamos neste item é uma taxonomia para os bilhetes. Taxonomia (do Grego verbo $\tau\alpha\sigma\sigma\epsilon\acute{\iota}\nu$ ou tassein = "para classificar" e νόμος ou nomos = lei, ciência, administrar, cf "economia"), foi uma vez, a ciência de classificar organismos vivos (alpha taxonomy), mas mais tarde a palavra foi aplicada em um sentido mais abrangente, podendo aplicar-se a uma das duas, classificação de coisas ou aos princípios subjacentes da classificação.

A taxonomia criada varia de sistema para sistema, ou seja, para telefonia celular existem eventos que são particularidades deste sistema e assim por diante. Exemplos: para um sistema baseado em *VoIP* não há como identificar problemas numa determinada central à frente, pois numa rede *IP* os pacotes de dados podem seguir caminhos distintos para alcançar o destino. Na telefonia fixa não haverá ERBs, Queda de RF, etc.

Apesar disso tudo, sempre existem eventos comuns entre diferentes sistemas, tais como: OK, CO, DP, LO, DI, NR. É possível afirmar que os eventos da telefonia celular compreendem boa parte dos eventos da telefonia fixa mais os eventos específicos.

Veja na Tabela 3.3 os eventos para um sistema de telefonia fixa¹.

Evento	Descrição
OK	Indica as chamadas completadas sem erro.
NR	Indica as chamadas que o assinante-destino não respondeu.
LO	Indica as chamadas que encontraram o assinante de destino na condição de ocupado.
CO0	Indica as chamadas não completadas devido ao esgotamento da temporização.
CO1	Indica as chamadas não completadas por congestionamento ou defeito na central-origem.
CO2	Indica as chamadas não completadas por congestionamento ou defeito na central-destino.
CO3	Indica as chamadas não completadas devido à falha na troca de sinalização MFC.
OU	Indica as chamadas não completadas devido a outros fatores.
DI	Indica as chamadas não completadas devido a erro de discagem.
DP	Indica que houve desligamento prematuro na chamada pelo assinante-origem.

Tabela 3.3 – Rótulos para classificar os bilhetes de telefonia fixa

Na Tabela 3.4 são apresentados os eventos criados especialmente para a telefonia celular. Estes eventos representam mais uma contribuição do trabalho.

Evento	Descrição
IS	Indica as chamadas não completadas devido ao assinante roamer não possui perfil.
RF1	Indica as chamadas não completadas devido à queda de RF antes da definição da rota.
RF2	Indica as cham. não completadas devido à queda de RF antes do atendimento.
RF3	Indica as chamadas onde houve queda de RF depois do atendimento c/tempo Conversação > 3 s.
PTO	Indica as chamadas não completadas devido a estouro de temporização no paging
POK	Indica as chamadas OK com tempo de conversação <= 3 segundos.
DT	Indica as chamadas não completadas devido a defeito técnico em algum componente da central.
HO	Indica as chamadas não completadas devido a um erro no handoff.
IHO	Indica as chamadas não completadas devido a um erro no InterSystem handoff.
HI	Indica as chamadas não completadas devido a um erro no handoff interno da célula.
PP	Indica as chamadas perdidas na Plataforma de Pré-pago.
FC	Indica as chamadas perdidas devido a falta de crédito.

Tabela 3.4 – Rótulos para classificar os bilhetes de telefonia celular

Por exemplo: se uma ligação telefônica foi concluída com sucesso, o assinante A conseguindo conversar com o assinante B e com terminação por uma das partes, essa ligação será classificada como “OK”. Essa classificação é necessária, para que se possa identificar o comportamento dos diversos componentes, elementos, do sistema pelos quais as chamadas estejam passando.

A localização de problemas nos elementos se dá quando da identificação de anormalidades em um ou mais eventos de um elemento. A denominação eventos [40] nada mais é do que o resultado da classificação feita nos bilhetes.

¹ Tabela retirada do Livro Telefonia Digital, Marcelo Sampaio Alencar, 4ª Edição, 2002, Editora Érica LTDA, página 213.

Para entender melhor determinados eventos é interessante ter um entendimento dos protocolos de sinalização envolvidos. O final de seleção destes protocolos indica o que aconteceu com as chamadas que saíram da central e foram encaminhadas para frente. O escopo desse trabalho não irá entrar em detalhes a respeito do funcionamento de todos esses padrões. No Apêndice B são explicados dois protocolos presentes na troca de mensagens entre centrais. São eles: Sinalização MFC (Multi Freqüencial Compelida) e Sinalização por Canal Comum nº 7.

3.6 - Mapeamento da Tecnologia

Como foi dito anteriormente, existem determinados protocolos que regulamentam a comunicação em pontos do sistema. Entretanto, internamente a uma central telefônica cada fabricante trabalha da maneira que melhor lhe convier. Como em todo bilhete existe uma “explicação” do que aconteceu com a ligação telefônica, é natural esperar que o conteúdo dos “causes”, descritos no Apêndice B, seja de alguma forma traduzido e gravado no bilhete. Essa relação existe e é feita através do campo Final de Seleção, ver Apêndice A item A.2, que indica o que aconteceu com a ligação telefônica. Esse campo é o principal na classificação, porém existem outros que também são necessários. São eles:

- Direção - Indicador do tipo da ligação:
 - fixo-móvel;
 - móvel-fixo;
 - móvel-móvel;
 - móvel-indeterminado;
 - fixo-indeterminado;
 - fixo-fixo.
- Estouro de Temporização - Se a ligação telefônica foi perdida devido a um estouro de temporização. Esse campo é bastante importante, pois é ele que indica o que estava acontecendo com a ligação antes de haver o estouro na temporização. Através dele é possível identificar se determinado evento está acontecendo com demasiada freqüência, levando a estouros na temporização do sistema;
- Duração da Ligação – Tempo de conversação da ligação. Essa informação é necessária, pois existe na norma da telefonia celular uma regra que diz que as ligações telefônicas com menos de 3 segundos não podem ser tarifadas. Isso se deve ao fato de que a ligação

passa por um canal de Rádio Frequência, podendo existir problemas no estabelecimento desse canal. Sendo assim, se a ligação durar menos que 3 segundos, ela não será cobrada. Se determinado elemento estiver com os níveis de chamadas com menos de 3 segundos elevados, pode estar com problemas na parte de RF. A norma da Anatel que regulamenta o Serviço Móvel Pessoal [41] define o seguinte: “chamadas faturáveis: somente são faturáveis as ligações com duração superior a três segundos.”

A classificação dos bilhetes foi criada baseando-se no significado dos campos anteriormente citados. O resultado da classificação foi a criação de 278 subeventos, que são apresentados no Apêndice A, item A.3. Este resultado corresponde a outra contribuição do trabalho.

O programa classifica os bilhetes, conforme as máscaras contidas neste apêndice. Quando um campo possuir o valor “x”, isso indica que esse campo, no bilhete original, pode possuir qualquer valor, ou seja, a classificação independe do conteúdo do campo em questão.

O programa foi montado de maneira a detectar quando bilhetes não estão sendo classificados devido à não-existência de uma máscara apropriada. Se isso ocorrer, uma máscara apropriada deve ser criada, e o programa, rodado novamente. Outra característica do programa é mostrar o grau de utilização de cada subevento. Se determinadas máscaras não estão sendo utilizadas, essas poderão ser eliminadas.

3.7 - Conclusão

Anteriormente citamos que os bilhetes utilizados nesse trabalho são de sistemas reais. Para a utilização desses bilhetes tivemos a preocupação de descaracterizar toda e qualquer informação contida que fosse sigilosa ou mesmo estratégica, exemplo: números dos telefones origem/destino, número de série, sistema de origem/destino, etc. Sendo assim, todos os bilhetes utilizados passaram por um processo de mascaramento, em que o conteúdo de alguns campos foi substituído. A substituição do conteúdo não prejudicou o desenvolvimento do trabalho, pois a quantidade de informação, entropia, foi mantida.

Problemas encontrados na formação dos bilhetes trouxeram alguns transtornos. Foi necessária a criação de uma função que, além de testar a coerência dos dados, sempre esteja buscando os caracteres de final e início de novo bilhete. Se o bilhete que estiver sendo lido possuir qualquer anomalia, será colocado numa tabela chamada “*CDR-defs*”. Já os bilhetes perfeitos ficam em outra tabela, chamada *CDR*. Essas tabelas possuem todos os campos possíveis do bilhete. Os que

Bilhetes de Tarifação

não fizerem parte de determinado bilhete que estiver sendo lido, simplesmente, ficam em branco, sem informação.

Foi desenvolvido um aplicativo para a leitura dos bilhetes que pode ser utilizado para diferentes fabricantes. Para a tecnologia apresentada anteriormente, o aplicativo também identifica sub-registros novos, não cadastrados. Se um novo sub-registro for identificado, ele deverá ser cadastrado, e o programa, rodado novamente. O programa leva em torno de meia hora para ler, processar e classificar 250 mil bilhetes, num computador com processador Athlon 1.7 GHz, 512 Mbytes de RAM.

Capítulo 4

Algoritmos para Processamento dos Dados e Detecção de Falhas

Neste capítulo são apresentados algoritmos para a detecção de falhas na rede e conseqüente gerência de *QoS* do sistema. Nosso objetivo principal é testar o desempenho destes algoritmos, tendo em vista que alguns já são conhecidos.

As abordagens adotadas consistem em algoritmos desenvolvidos sobre o modelo Estocástico e de Redes Neurais. O desempenho de ambos é tratado no decorrer do trabalho, guardadas algumas particularidades. Neste capítulo são apresentados quatro algoritmos.

No Apêndice B, B.1 - Probabilidade, é apresentada a teoria de probabilidades e processos estocásticos dando ênfase a confiabilidade. Esse apêndice possui os fundamentos teóricos que são à base dos algoritmos.

O termo confiabilidade pode ser definido como um número que estabelece a probabilidade que um equipamento ou sistema tem de executar sua função, satisfatoriamente, por um período de tempo determinado. O conceito de falha adotado neste trabalho considera que um equipamento não executa sua função como deveria, ou seja, qualquer equipamento, dispositivo ou sistema falham, quando seu desempenho não atende à função para a qual foi projetado.

4.1 - Algoritmos para Detecção de Falhas

Os algoritmos podem ser empregados para monitorar os mais diferentes elementos, tanto lógicos como físicos (classificação proveniente da *TMN*, conforme explicado no Capítulo 2, ítem Garantia de Configuração), dentro de um sistema de comunicação. A seguir são apresentados os elementos de um sistema de telefonia celular:

- Circuitos de Entrada (elemento físico) – circuitos de entrada por onde será estabelecida uma ligação telefônica;
- Circuitos de Saída (elemento físico) – circuitos de saída por onde será estabelecida uma ligação telefônica;
- Transceptor MFC (elemento físico) – responsável pela sinalização MFC (Multi Frecuencial Compelida);
- Troncos de entrada (elemento físico) – geralmente formado por um E1 (30 canais) de entrada;
- Troncos de saída (elemento físico) – geralmente formado por um E1 (30 canais) de saída;
- Rotas de Entrada (elemento físico) – conjunto de circuitos de entrada que formam uma rota;
- Rotas de Saída (elemento físico) – conjunto de circuitos de saída que formam uma rota;
- BSC (elemento físico) – Base Station Controle é um módulo de controle das BTS CDMA;
- BTS origem (elemento físico) – Base Transceiver Station é de onde são originadas as ligações na tecnologia CDMA;
- BTS destino (elemento físico) – Base Transceiver Station é de onde são recebidas as ligações na tecnologia CDMA;
- ERB origem (elemento físico) – estação rádio base de onde são originadas as ligações na tecnologia AMPS/TDMA;
- ERB destino (elemento físico) – estação rádio base onde são terminadas as ligações na tecnologia AMPS/TDMA;
- Canais de RF origem (elemento físico) – canal de rádio frequência pelo qual a ligação foi originada;

Algoritmos para Processamento dos Dados

- Canais de RF destino (elemento físico) – canal de rádio frequência pelo qual a ligação foi recebida;
- Assinante Origem (elemento físico) – número do telefone que originou a chamada;
- Assinante Destino (elemento físico) – número do telefone que recebeu a chamada;
- Direções das Chamadas de Origem (elemento lógico) – faixa de telefones pode ser um bairro, uma região, uma cidade, um estado, um país da qual as chamadas são originadas;
- Direções das Chamadas de Destino (elemento lógico) - faixa de telefones pode ser um bairro, uma região, uma cidade, um estado, um país para os quais as chamadas são destinadas;
- Tabelas de encaminhamento das chamadas de entrada (elemento lógico) – tabelas de encaminhamento da central de telefonia para as chamadas entrantes;
- Tabelas de encaminhamento das chamadas de saída (elemento lógico) – tabelas de encaminhamento da central de telefonia para as chamadas saíntes;
- Tabelas de ERB/BTS (elemento lógico) – tabelas das ERBs/BTS;
- Tabelas de Periféricos (elemento lógico) – tabelas dos elementos periféricos;
- Tabelas dos módulos de processamento (elemento lógico) – tabelas dos módulos de processamento, tais como: Módulo central, MS (*Message Switch*), ENET (matriz de comutação);
- Software de controle da central (elemento lógico) – software que controla a central de telefonia;
- Sinalização telefônica – sinalização que controla a troca de mensagens entre os elementos da rede.

Dentro de cada um desses elementos são monitorados os eventos, independentemente um do outro, ou seja, de forma paralela. A classificação dos eventos está apresentada na Tabela 8.2, Apêndice B, item B.1 - Probabilidade. Para cada elemento poderão ser monitorados todos os eventos pertinentes. A detecção de um problema num elemento será sempre feita sobre um ou mais eventos, dependendo de quais são os ofensores. Existe uma particularidade em relação ao evento OK, ou seja, ligações estabelecidas e com conversação. Para esse caso não faz sentido gerar alarmes, à medida que o percentual do evento cresça, pois o crescimento indica a melhora do sistema e não deterioração. Para o evento OK será descrito um algoritmo específico.

Todos os algoritmos apresentados trabalham com a garantia de que alarmes falso-positivos são gerados dentro de uma condição. Um alarme falso-positivo é aquele que, apesar do alarme ser gerado, não possui uma causa real, ou seja, na verdade a falha não existe. Esse valor pode ser definido como um critério de segurança e pode assumir uma gama de valores, por exemplo, um alarme falso-positivo a cada mil, ou então, a cada milhão de bilhetes. Para cada algoritmo essa questão é analisada. Os algoritmos são os seguintes:

- Algoritmo de Tempo Real que consiste em analisar as saídas das chamadas esperando por uma sequência de eventos. É como se estivéssemos “jogando moedas aleatoriamente” e esperando por uma determinada sequência para confirmarmos um alarme;
- Algoritmo de Espaço Amostral no qual esperamos pelo preenchimento da janela de dados para então procurarmos por uma anomalia utilizando a distribuição binomial;
- Algoritmo para Chamadas Ok procura por anormalidades dentro de um tipo específico de chamada, ou melhor, chamadas que a princípio foram concluídas com sucesso. Este algoritmo engloba os dois mecanismos citados anteriormente: Tempo Real e Chamadas OK;
- Algoritmo para Detecção de Ocupação que trabalha sobre o tempo de ocorrência dos eventos para gerar o alarme. O algoritmo se baseia na distribuição de Poisson.

4.1.1 - Algoritmo de Tempo Real

O algoritmo de Tempo Real é baseado na Teoria de Eventos Recorrentes ou Teoria da Renovação aplicada a um experimento de Bernoulli [42]-[44] e foi embasada no trabalho de Nunes [48][49], conforme veremos mais adiante. A designação de Tempo Real é única e exclusivamente devida ao fato de que este algoritmo toma decisão para cada novo CDR que é gerado, não necessitando esperar um determinado tempo ou uma determinada quantidade de bilhetes para chegar a uma conclusão sobre um alarme. A esperança e o desvio padrão são renovados sempre que um novo bilhete é gerado.

A teoria da renovação é extensamente utilizada em Confiabilidade de Sistemas [45][46] que define que para sistemas reparáveis, o tempo de operação não é contínuo. Em outras palavras, o ciclo de vida do sistema pode ser descrito por uma sequência de estados “*up*” e “*down*”. O sistema opera até a falha, então é reparado e retorna ao estágio original de operação. O sistema irá falhar novamente após um determinado tempo de operação aleatório, e será reparado novamente, e este

processo de falha e reparo irá se repetir. Isto é denominado como processo de renovação e é definido como uma seqüência de variáveis aleatórias independentes e não negativas.

4.1.1.1 - Definições

Seja X uma variável aleatória que representa um evento num experimento de Bernoulli. O espaço amostral da variável aleatória pode assumir dois valores:

$$X = \begin{cases} 1 \\ 0 \end{cases}, \text{ em que o valor } 1 \text{ é quando acontece o evento, e } 0 \text{ é quando não acontece. Seja}$$

“ p ” a probabilidade de $X = 1$.

Seja Y uma variável aleatória, relacionada com a quantidade de eventos N , até se formar, pela primeira vez, uma seqüência de “ r ” uns (1’s) seguidos.

Pela Teoria da Renovação² $N = E(Y) =$ esperança de Y .

$$E(Y) = \mu = N = \frac{1 - p^r}{p^r \cdot (1 - p)}; \quad (4.1)$$

Isolando a variável r :

$$N \cdot p^r \cdot (1 - p) = 1 - p^r; \quad (4.2)$$

$$N \cdot p^r \cdot (1 - p) + p^r = 1; \quad (4.3)$$

$$p^r \cdot (N \cdot (1 - p) + 1) = 1; \quad (4.4)$$

$$p^r = \frac{1}{N \cdot (1 - p) + 1}; \quad (4.5)$$

$$\ln(p^r) = \ln\left[\frac{1}{N \cdot (1 - p) + 1}\right]; \quad (4.6)$$

$$r \cdot \ln(p) = \ln(1) - \ln[N \cdot (1 - p) + 1]; \quad (4.7)$$

$$r = \frac{0 - \ln[N \cdot (1 - p) + 1]}{\ln p}; \quad (4.8)$$

² Equação retirada do livro: W. Feller, “An introduction to Probability Theory and Its Applications – Volume 1”, Chapter XIII, p. 303-341, John Wiley & Sons, Inc, 1968.

$$r = \frac{-\ln[N \cdot (1-p) + 1]}{\ln p}; \quad (4.9)$$

Como r deve ser inteiro, o valor resultante é sempre arredondado para cima, pois o arredondamento para baixo não garantiria a margem de segurança estipulada:

$$r = \left\lceil \frac{-\ln[N \cdot (1-p) + 1]}{\ln p} \right\rceil; \quad (4.10)$$

A função “*ceiling*” [47] de x denotada por $\lceil x \rceil$ de um número real é o próximo número inteiro maior ou igual a x .

A equação dá como resultado a quantidade r de insucessos sucessivos que devem acontecer, para que um alarme seja gerado. A quantidade está diretamente relacionada com a probabilidade p e o valor N . O valor N é determinado de tal forma, que garanta que alarmes falso-positivos somente sejam gerados dentro de um limite considerado aceitável. A maneira de analisar o valor do N é a seguinte: Suponha, por exemplo, que um evento tenha uma média de $p=1\%$ de acontecimento. Qual a quantidade de insucessos sucessivos, para que o evento seja alarmado falso-positivamente, uma vez em um milhão de bilhetes, 1.000.000? Se esses valores forem aplicados na equação 4.10, o resultado obtido é que a quantidade r de insucessos sucessivos deve ser de 3, ou seja, se houver quatro eventos consecutivos e um alarme for gerado, então a probabilidade de que esse alarme seja falso-positivo é de uma em um milhão.

Uma conclusão é que quanto maior a exigência em relação à geração do alarme, maior será a quantidade de insucessos sucessivos que deverão acontecer, para que o alarme seja gerado.

A equação 4.10 é muito semelhante à equação $r = \left\lceil \frac{-\ln[N \cdot (1-p) + 1]}{\ln p} \right\rceil + 1$ que foi utilizada por Nunes [48][49] na detecção de problemas em sistemas de telefonia utilizando bilhetes de tarifação. O termo “+1” diverge da equação deduzida do livro, possivelmente este termo foi adicionado para compensar o arredondamento para baixo, mas não temos como comprovar isso, tendo em vista que não possuímos referências bibliográficas suficientes de como foi feita a dedução da equação. Pesquisando pelas origens da equação acima citada encontramos o trabalho de Feller³ que nos propiciou a dedução da equação 4.10.

³ Equação retirada do livro: W. Feller, “An introduction to Probability Theory and Its Applications – Volume 1”, Chapter XIII, p. 303-341, John Wiley & Sons, Inc, 1968.

Na Tabela 4.1 é apresentada a relação entre o valor de p (*AQL-Acceptable Quality Level*) e a quantidade de alarmes sucessivos em função do valor de N fixo. O *AQL* é o nível de qualidade considerado aceitável pela operadora ou pelo órgão governamental, no caso brasileiro, ANATEL (Agência Nacional de Telecomunicações). Este parâmetro é um valor subjetivo. Para um melhor entendimento, podemos dizer também que o *AQL* estabelece um “nível aceitável de falhas”. Os algoritmos pesquisarão sobre os elementos, tendo como referência o nível de *AQL* estabelecido como aceitável, procurando falhas que estejam comprometendo a *QoS* do serviço.

Probabilidade	N=1000	N=10000	N=100000	N=1000000
0,1%	1	2	2	2
0,2%	2	2	2	3
0,3%	2	2	2	3
0,4%	2	2	3	3
0,5%	2	2	3	3
0,6%	2	2	3	3
0,7%	2	2	3	3
0,8%	2	2	3	3
0,9%	2	2	3	3
1%	2	2	3	3
2%	2	3	3	4
3%	2	3	4	4
5%	3	4	4	5
7%	3	4	5	6
9%	3	4	5	6
11%	4	5	6	7
13%	4	5	6	7
15%	4	5	6	8
17%	4	6	7	8
19%	5	6	7	9
21%	5	6	8	9
23%	5	7	8	10
25%	5	7	9	10
27%	6	7	9	11
29%	6	8	10	11
31%	6	8	10	12
33%	6	8	11	13
35%	7	9	11	13
37%	7	9	12	14
39%	7	10	12	15
41%	8	10	13	15
43%	8	11	13	16
45%	8	11	14	17
47%	9	12	15	18
49%	9	12	16	19
51%	10	13	17	20
53%	10	14	17	21
55%	11	15	18	22
57%	11	15	19	24
59%	12	16	21	25
61%	13	17	22	27
63%	13	18	23	28
65%	14	19	25	30
67%	15	21	26	32
69%	16	22	28	35
71%	17	24	31	37
73%	18	26	33	40
75%	20	28	36	44
77%	21	30	39	48
79%	23	33	43	52
81%	25	36	47	58
83%	28	40	53	65
85%	31	46	60	74
87%	36	52	69	85
89%	41	61	80	100
91%	48	73	97	121
93%	59	91	123	154
95%	77	122	167	211
97%	113	188	263	339
99%	239	460	688	917

Tabela 4.1 – Quantidade de eventos sucessivos em função de p e N

Simulando agora um evento qualquer, com diferentes $AQLs$, sejam eles: 5%, 9%, 21%, 41%, 61% e 81%, através da Tabela 4.1 e do índice de alarmes falso-positivos que será de um em um milhão, é possível encontrar que para ser gerado um alarme necessita-se da ocorrência de 5, 6, 9, 15, 27 e 58 insucessos sucessivos, respectivamente.

4.1.1.2 - Análise Geral

O algoritmo foi aplicado sobre condições de tráfego real, em diversas situações, tais como: rotas para estações rádio base, canais de RF, entroncamentos entre centrais, controladores específicos de periféricos, etc. Ficaria extremamente difícil e desgastante, tanto para quem lê quanto para quem escreve, descrever todos os resultados obtidos nos testes. Sendo assim são apresentados os resultados de dois casos: ERB (Estação Rádio Base) de um sistema celular com alta densidade de tráfego e o entrocamento que interliga duas centrais.

São adotados como *AQLs* os valores dos níveis percentuais dos eventos. Isso propicia trabalhar em regiões mais próximas da operação real do elemento, melhorando assim as condições para localização de falhas. O primeiro estudo levantou os níveis de qualidade da estação. As médias dos eventos estão representadas na Tabela 4.2.

EVENTO	Qdade Chamadas	%
CO0	92	5,2%
CO1	4	0,2%
CO2	2	0,1%
DP1	87	4,9%
DP2	249	14,0%
DT	31	1,7%
LO	209	11,8%
NR	21	1,2%
OK1	903	50,8%
OK2	1	0,1%
OU	12	0,7%
POK	103	5,8%
QRF2	18	1,0%
QRF3	45	2,5%
	1777	100,0%

Tabela 4.2 – Médias dos níveis de qualidade

Também foi estudado o comportamento dos eventos sucessivos que estão retratados na Tabela 4.3. O significado dessa tabela pode ser entendido da seguinte forma: Tome, por exemplo, o caso do evento qualquer, por exemplo, CO0, cuja média de ocorrência nessa ERB é de 5,2%. A tabela retrata que em 96,6% das aparições do evento CO0 ele ocorre isolado, ou seja, não são ocorrências sucessivas. Já em 2% das ocorrências o evento aparece duas vezes consecutivas, um par sucessivo, e em 1% das vezes aparece em 3 eventos consecutivos ou sucessivos. O total de ligações telefônicas nessa ERB é de cerca de 2.000 dentro de uma hora de observação. É possível observar que as ocorrências sucessivas ficam bastante próximas dos valores estabelecidos para o alarme ser gerado para N=1.000, Tabela 4.1.

Algoritmos para Processamento dos Dados

Tendo em vista que essa estação não possui problemas, os alarmes gerados serão falso-positivos e podem ser visualizados na Tabela 4.4, campos preenchidos com a palavra “sim”. O aparecimento desses alarmes pode ser entendido, tendo-se em mente que, ao se trabalhar com esta ERB, o universo de ligações foi de cerca de 2.000 chamadas ou bilhetes. Para um N com valor baixo (1.000 à 10.000) sempre existirá a possibilidade de que alarmes falso-positivos sejam gerados, lembrando-se sempre da definição do N. No decorrer do trabalho, será adotado um valor para N igual a 1.000.000, para evitar que isso aconteça.

CO0	5,2%	NR	1,2%
1	96,6%	1	100,0%
2	2,3%		
3	1,1%	OK1	50,8%
		1	45,9%
CO1	0,2%	2	27,2%
1	100,0%	3	13,0%
		4	8,0%
CO2	0,1%	5	3,0%
1	100,0%	6	1,4%
		7	1,1%
DP1	4,9%	8	0,2%
1	92,5%	9	0,2%
2	6,3%		
3	1,3%	OK2	0,1%
		1	100,0%
DP2	14,0%		
1	85,3%	OU	0,7%
2	12,8%	1	100,0%
3	0,9%		
4	0,5%	POK	5,8%
5	0,5%	1	97,0%
		2	3,0%
DT	1,7%		
1	93,1%	QRF2	1,0%
2	6,9%	1	94,1%
		2	5,9%
LO	11,8%		
1	88,1%	QRF3	2,5%
2	10,8%	1	97,7%
3	1,1%	2	2,3%

Tabela 4.3 – Eventos sucessivos

EVENTO	%	N=1000	N=10000	N=100000	N=1000000	N=10000000
CO0	5,2%	sim	não	não	não	não
CO1	0,2%	não	não	não	não	não
CO2	0,1%	sim	não	não	não	não
DP1	4,9%	sim	não	não	não	não
DP2	14,0%	sim	sim	não	não	não
DT	1,7%	sim	não	não	não	não
LO	11,8%	não	não	não	não	não
NR	1,2%	não	não	não	não	não
OK1	50,8%					
OK2	0,1%					
OU	0,7%	não	não	não	não	não
POK	5,8%	não	não	não	não	não
QRF2	1,0%	sim	não	não	não	não
QRF3	2,5%	sim	não	não	não	não

Tabela 4.4 – Alarmes falso-positivos em função do valor de N

Analisando uma rota que interliga duas centrais foi constatado que o total de ligações durante o período de observação (17:00 as 18:00 horas) foi de 43.266. As informações de desempenho podem ser vistas na Tabela 4.5.

EVENTO	Qdade Chamadas	%
CO0	38	0,1%
DP1	5494	12,7%
DP2	7773	18,0%
NR	387	0,9%
OK1	27405	63,3%
OU	43	0,1%
PGTO	166	0,4%
POK	1470	3,4%
QRF3	490	1,1%
	43266	100,0%

Tabela 4.5 - Níveis de qualidade

Os eventos sucessivos estão na Tabela 4.6. Cruzando esses dados com a Tabela 4.1, foi encontrado que, quando N=1.000 ou 10.000, são gerados vários alarmes falso-positivos. Isso pode ser melhor visualizado na Tabela 4.7. Observe que, em todos os eventos, quando N=1.000, são gerados alarmes. Para N=10.000, em apenas dois casos, não são gerados alarmes. Nesses dois eventos as suas ocorrências relativas são bastante baixas, na ordem de 0,1 ponto percentual, o que pode mascarar o comportamento.

Mais uma vez foi comprovado que, para se trabalhar com uma elevada garantia de que não serão gerados alarmes falso-positivos, deve-se utilizar o valor de N que fique na casa dos milhões.

CO0	Sucessivos	0,1%	OU	Sucessivos	0,1%
1	38	100,0%	1	43	100,0%
DP1	Sucessivos	12,7%	PGTO	Sucessivos	0,4%
1	3923	84,5%	1	160	98,2%
2	612	13,2%	2	3	1,8%
3	95	2,0%			
4	13	0,3%	POK	Sucessivos	3,4%
5	2	0,0%	1	1377	96,8%
			2	42	3,0%
			3	3	0,2%
DP2	Sucessivos	18,0%	QRF3	Sucessivos	1,1%
1	4984	79,9%	1	458	96,6%
2	1015	16,3%	2	16	3,4%
3	203	3,3%			
4	31	0,5%	NR	Sucessivos	0,9%
5	4	0,1%	1	381	99,2%
6	1	0,0%	2	3	0,8%

Tabela 4.6 – Eventos sucessivos

EVENTO	%	N=1000	N=10000	N=100000	N=1000000	N=10000000
CO0	0,1%	sim	não	não	não	não
DP1	12,7%	sim	sim	não	não	não
DP2	18,0%	sim	sim	não	não	não
NR	0,9%	sim	sim	não	não	não
OK1	63,3%					
OU	0,1%	sim	não	não	não	não
PGTO	0,4%	sim	sim	não	não	não
POK	3,4%	sim	sim	não	não	não
QRF3	1,1%	sim	sim	não	não	não

Tabela 4.7 – Alarmes falso-positivos em função de N

4.1.1.3 - Detecção de falhas

O próximo passo é utilizar o algoritmo de tempo real para detectar falhas existentes. Para isso começaremos a simular falhas nos rádios ou transceptores da BTS. A falha a ser induzida será o evento “DT” (Defeito Técnico). Será adotado como *AQL* o valor atual da frequência relativa do evento DT apresentado na Tabela 4.2, ou seja, 1,7 ponto percentual. Calculando a quantidade de eventos sucessivos, para que um alarme seja gerado, são obtidos os seguintes valores, Tabela 4.8:

Probabilidade	N=1000	N=10000	N=100000	N=1000000
1,7%	2	3	3	4

Tabela 4.8 – Quantidade de eventos

A quantidade de eventos sucessivos deverá ser de, no mínimo, quatro, para que um alarme seja gerado, adotando o valor de $N=1.000.000$.

Tendo em vista que estamos trabalhando sobre uma BTS sem falhas, ou seja, “saudável”, é necessário encontrar uma maneira de gerar falhas para testar o comportamento do algoritmo. A maneira pela qual geramos as falhas é baseada na criação aleatória de problemas nos canais de voz de uma BTS, veja Figura 4.1. A detecção desse tipo de problema é complexa, pois toda vez que existir uma ordem qualquer na degradação dos elementos, esses serão de mais fácil e rápida detecção em comparação com um evento aleatório. Uma ordem qualquer pressupõe uma entropia menor, ou então, uma quantidade de informação maior do que um evento puramente aleatório.

A Figura 4.2 mostra o resultado de aproximadamente 5.000 experimentos com degradação aleatória dos elementos, canais de voz. Para cada um destes experimentos foi feita a degradação da *QoS* da BTS e para cada nova degradação foi testado se o algoritmo detectava ou não a falha. Para cada degradação, ou seja, colocação de problemas nos canais de voz, as chamadas relativas a aqueles canais apresentavam falha no estabelecimento. À medida que mais canais apresentavam problemas maior era a degradação do nível de *QoS* da BTS, e a para cada degradação era testado o comportamento do algoritmo. A Figura 4.2 representa um histograma de como o algoritmo se comportou em relação a detecção, ou seja, em que momento da degradação da *QoS* o algoritmo detectou a falha.

É possível observar que existe uma convergência para uma Distribuição Normal, aproximação na linha em vermelho. Os valores para média e desvio-padrão são, respectivamente, $\mu = 15$ e $\sigma = 7,3$.

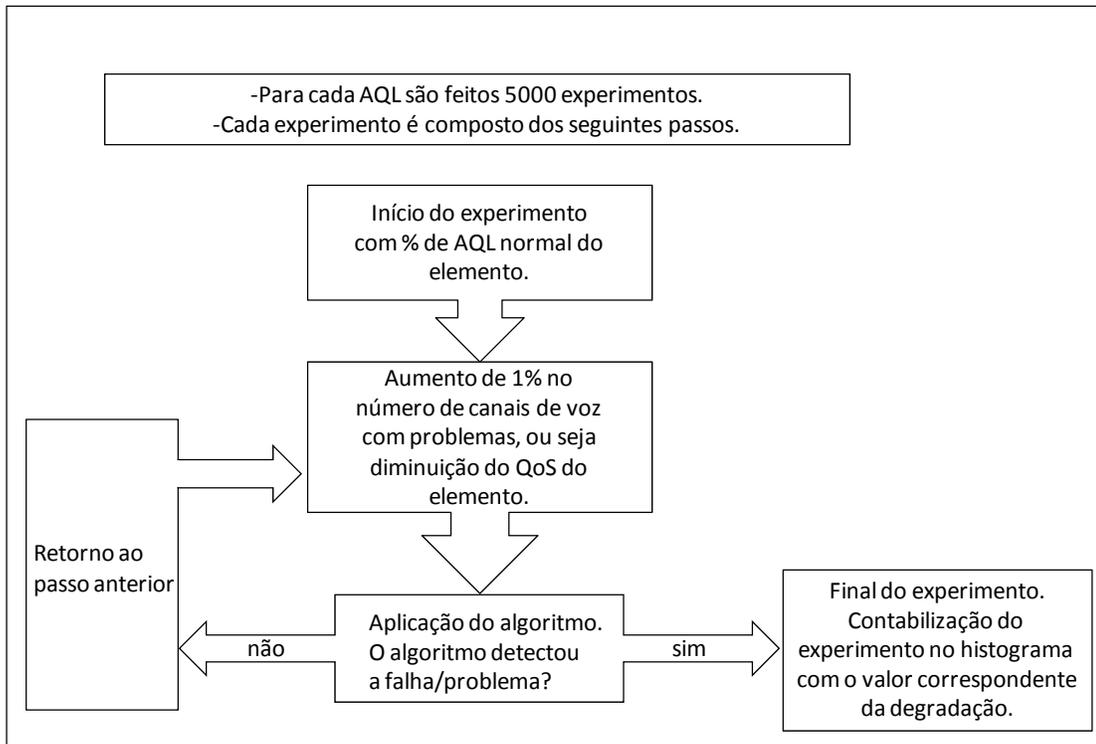


Figura 4.1 – Fluxograma

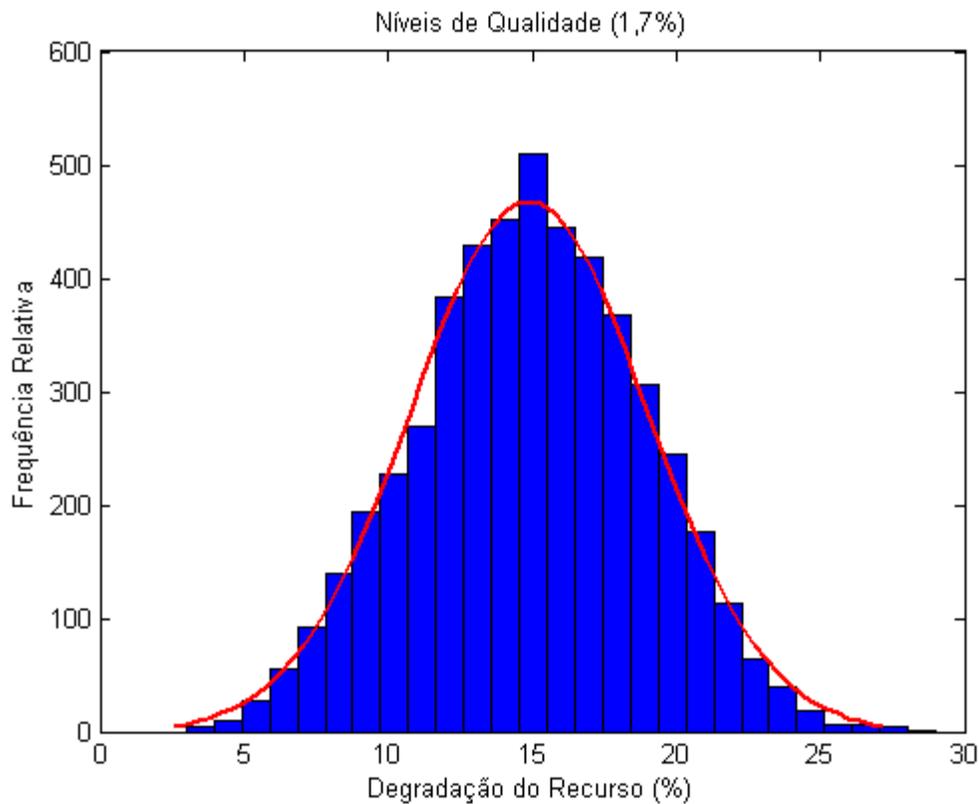


Figura 4.2 – Nível de qualidade de 1,7%

O gráfico mostra que, se um evento, com $p=AQL=1,7\%$, seja numa ERB, rota, etc, começar a degradar sua qualidade aleatoriamente, essa degradação será observada, conforme o comportamento do gráfico acima. Para conhecer as probabilidades de o evento ser detectado, quando a degradação for, por exemplo, 10%, basta fazer a integral dessa Distribuição Normal até o ponto em questão (Distribuição Cumulativa de Probabilidades). Os valores deverão ser normalizados para base um, representando 100%. Veja na Figura 4.3. Toda vez que os valores são aproximados por uma distribuição, seja ela qual for, sempre haverá um erro associado. Nesse exemplo o erro pode ser facilmente observado, tendo em vista que na Figura 4.3 a distribuição começa em -10% . A Tabela 4.9 mostra a probabilidade de o alarme ser detectado em todos os valores possíveis de degradação do elemento. São calculados os valores reais, bem como os da aproximação pela Distribuição Normal.

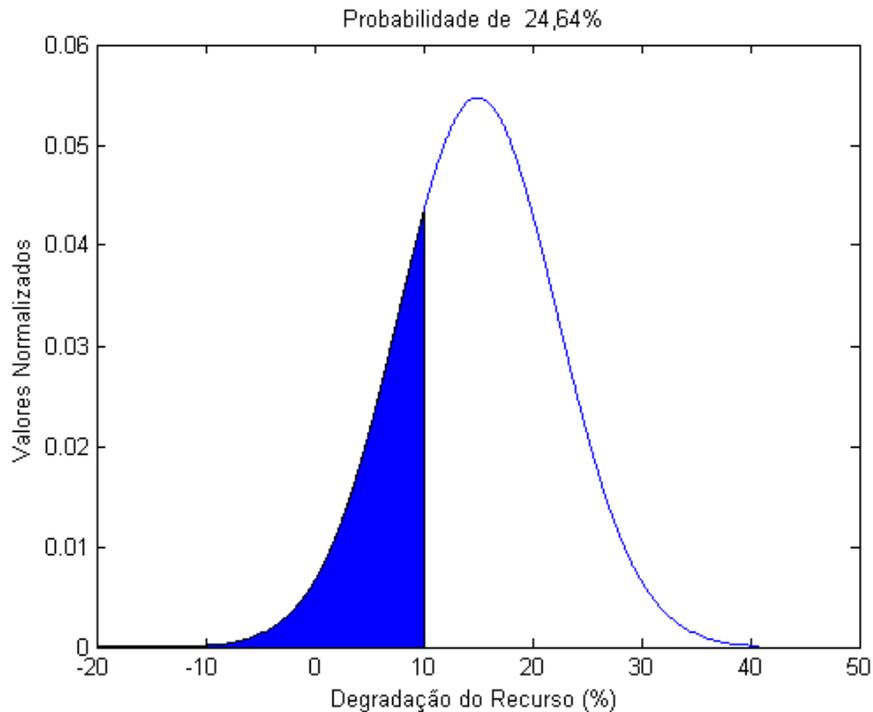


Figura 4.3 – Probabilidades

Degradação	Real	Dist. Normal
1%	0,0%	2,7%
2%	0,1%	3,7%
3%	0,3%	5,0%
4%	0,9%	6,6%
5%	2,0%	8,5%
6%	3,8%	10,9%
7%	6,6%	13,6%
8%	10,5%	16,9%
9%	15,0%	20,5%
10%	20,4%	24,6%
11%	28,1%	29,2%
12%	36,6%	34,0%
13%	45,7%	39,2%
14%	55,8%	44,5%
15%	64,7%	50,0%
16%	73,1%	55,5%
17%	80,4%	60,8%
18%	86,5%	66,0%
19%	91,4%	70,8%
20%	94,9%	75,4%
21%	97,2%	79,5%
22%	98,5%	83,2%
23%	99,3%	86,4%
24%	99,6%	89,2%
25%	99,8%	91,5%
26%	99,9%	93,4%
27%	100,0%	95,0%
28%	100,0%	96,3%
29%	100,0%	97,3%
30%	100,0%	98,0%
31%	100,0%	98,6%
32%	100,0%	99,0%
33%	100,0%	99,3%
34%	100,0%	99,5%
35%	100,0%	99,7%
36%	100,0%	99,8%
37%	100,0%	99,9%
38%	100,0%	99,9%
39%	100,0%	100,0%

Tabela 4.9 – Distribuição Cumulativa de Probabilidades

Na Figura 4.4 são apresentados gráficos similares ao da Figura 4.2, porém para diferentes AQLs: 1%, 2%, 5%, 7%, 12%, 17%, 22%, 27%, 32%, 37%, 42%, 47%, 52%, 57%, 62%, 67%, 72%, 77%, 82%, 86%, 88%, 90% e 92%. Os gráficos são descritos em função da degradação da qualidade do elemento, representada no eixo y . São feitos cerca de 5.000 experimentos para cada nível de qualidade. Os resultados são distribuições que se assemelham muito a Distribuições Normais. É possível observar, na Figura 4.5, que o desvio-padrão é menor nos níveis de qualidade na região de 1% a 5%, o que é representado por um estreitamento da distribuição. Isso também pode ser observado, com maior ênfase, nas regiões superiores. Esse comportamento pode ser entendido, uma

Algoritmos para Processamento dos Dados

vez que, ao trabalhar nas bordas, tanto inferior quanto superior, os valores para a faixa de degradação possível são estreitados, fazendo com que a variância da distribuição seja menor, o que torna a forma da figura mais estreita.

Na Figura 4.5 está a representação dos diferentes níveis de qualidade pela sua correspondente aproximação por Distribuições Normais com os respectivos desvios-padrão. A aproximação pela Distribuição Normal, nesse caso, tem o único objetivo de obter valores médios para efeito de comparação do comportamento.

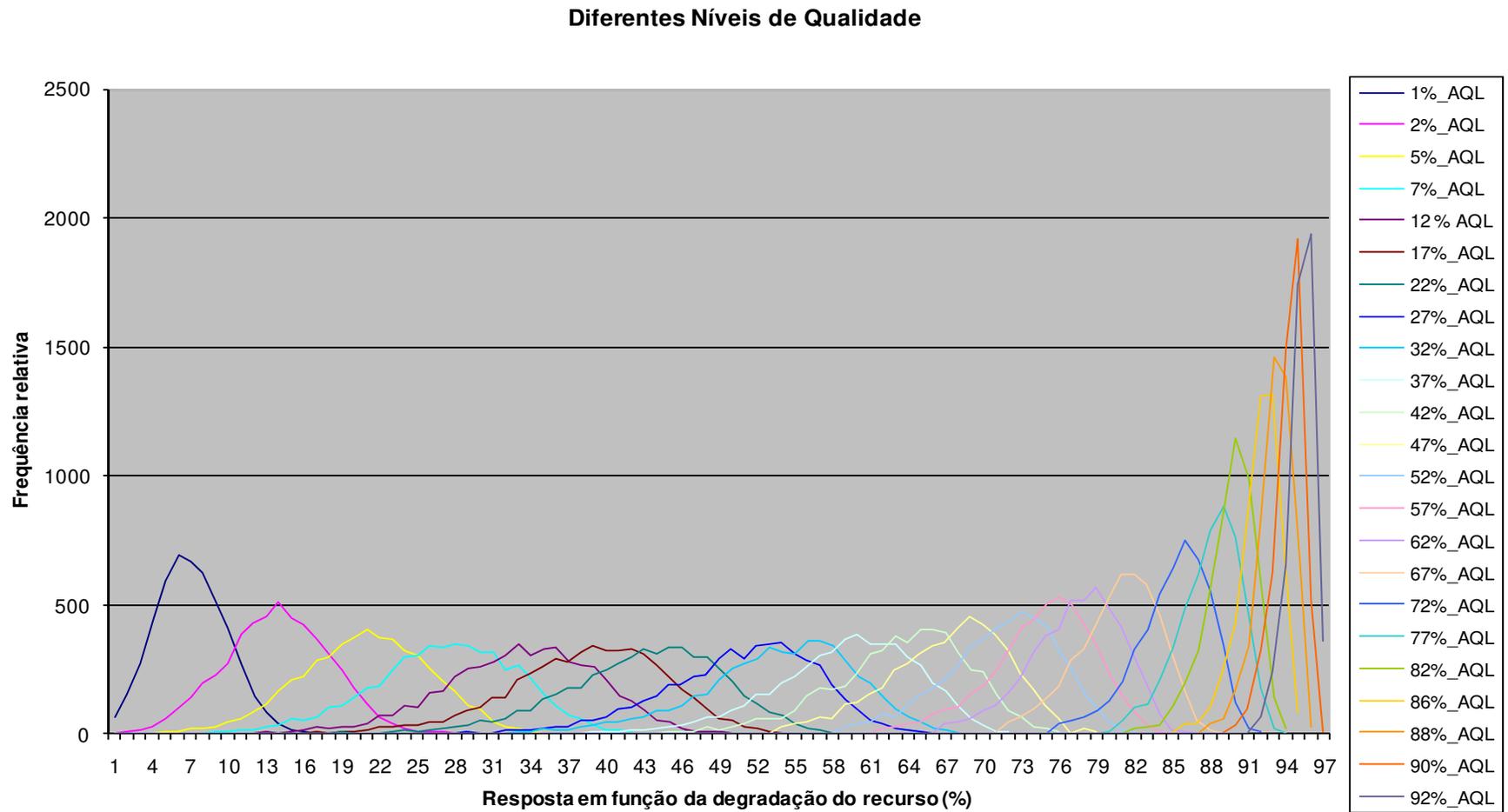


Figura 4.4 – Respostas para diferentes níveis de qualidade

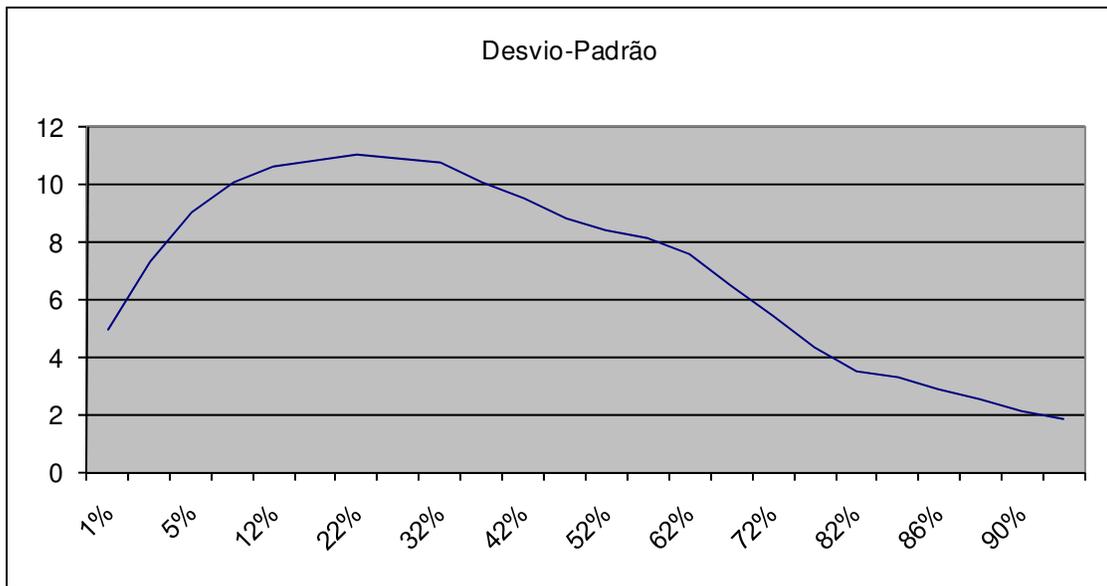


Figura 4.5 – Desvio-Padrão em função do Nível de Qualidade

%_AQL	Média	Desvio-Padrão
1%	8,03%	4,894
2%	14,80%	7,300
5%	21,86%	8,971
7%	28,22%	10,040
12%	34,14%	10,610
17%	39,48%	10,820
22%	44,00%	11,010
27%	51,63%	10,860
32%	54,62%	10,760
37%	60,06%	10,060
42%	64,46%	9,469
47%	68,06%	8,788
52%	72,42%	8,411
57%	75,91%	8,125
62%	78,83%	7,573
67%	82,49%	6,436
72%	86,00%	5,395
77%	88,92%	4,276
82%	90,57%	3,494
84%	91,79%	3,251
86%	92,92%	2,877
88%	94,09%	2,505
90%	95,26%	2,097
92%	96,27%	1,812

Tabela 4.10 – Média e Desvio-Padrão

4.1.1.4 - Tempo de detecção

Foi também elaborado um estudo em relação ao tempo necessário para a detecção dos problemas. Foram fixados níveis de qualidade, *AQLs*, usualmente encontrados nas redes de telecomunicações: 2%, 7%, 12%, 17%, 22% e 27%. O estudo foi feito degradando-se o elemento e vendo a resposta do algoritmo em relação ao tempo. Para cada nível de qualidade e para cada degradação foram feitos três mil experimentos. O tempo está no formato hh:mm (hora e minuto).

A seguir, Figura 4.6 até a Figura 4.11, são apresentados alguns resultados. Os resultados completos estão descritos no Apêndice B, item B.2- Resultados.

Nestas figuras são mostrados os níveis de qualidade de 2%, 7%, 12%, 17%, 22% e 27% em função da degradação do elemento, representado nas figuras pelas letras A), B), C) e D). A degradação está representada em ordem crescente, observe na Tabela 4.11:

		Degradação			
Figura 4.6	Nível de qualidade 2%	A) 5%	B) 14,8%	C) 44,5%	D) 84,1%
Figura 4.7	Nível de qualidade 7%	A) 14,8%	B) 44,5%	C) 74,2%	D) 94%
Figura 4.8	Nível de qualidade 12%	A) 19,8%	B) 44,5%	C) 74,2%	D) 94%
Figura 4.9	Nível de qualidade 17%	A) 24,7%	B) 54,4%	C) 74,2%	D) 94%
Figura 4.10	Nível de qualidade 22%	A) 34,6%	B) 64,3%	C) 74,2%	D) 94%
Figura 4.11	Nível de qualidade 27%	A) 44,5%	B) 74,2%	C) 84,1%	D) 94%

Tabela 4.11 – Tempos de Detecção em função do Nível de Qualidade e da Degradação

É possível concluir que a medida com que a degradação aumenta o tempo de detecção do alarme se torna menor, bem como a quantidade de vezes que o problema é detectado é maior, representado pelos valores no eixo “Frequência”. Outra conclusão é que à medida que o nível de qualidade cresce a degradação necessária para detecção do problema deve ser necessariamente maior.

- Nível de qualidade de 2%

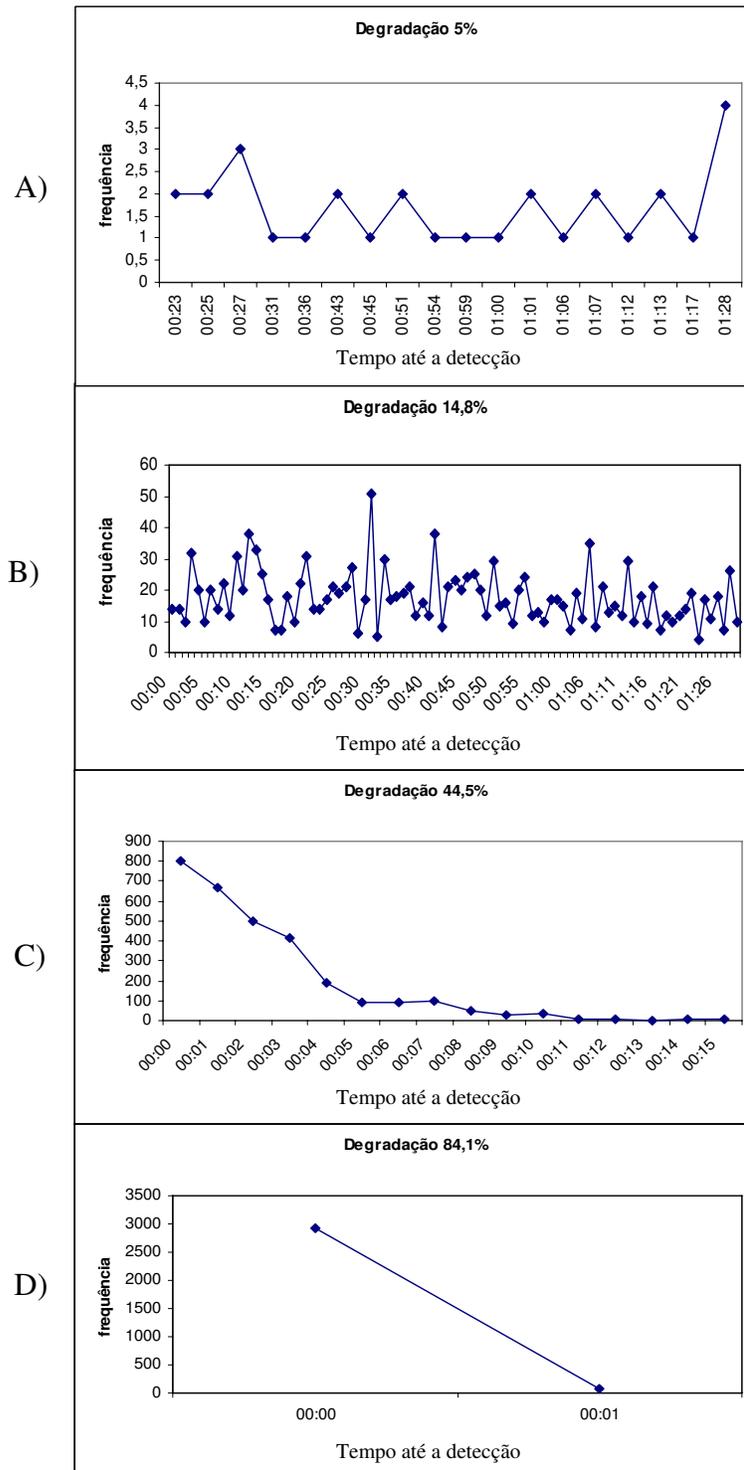


Figura 4.6 – Nível de Qualidade de 2%

- Nível de qualidade de 7%

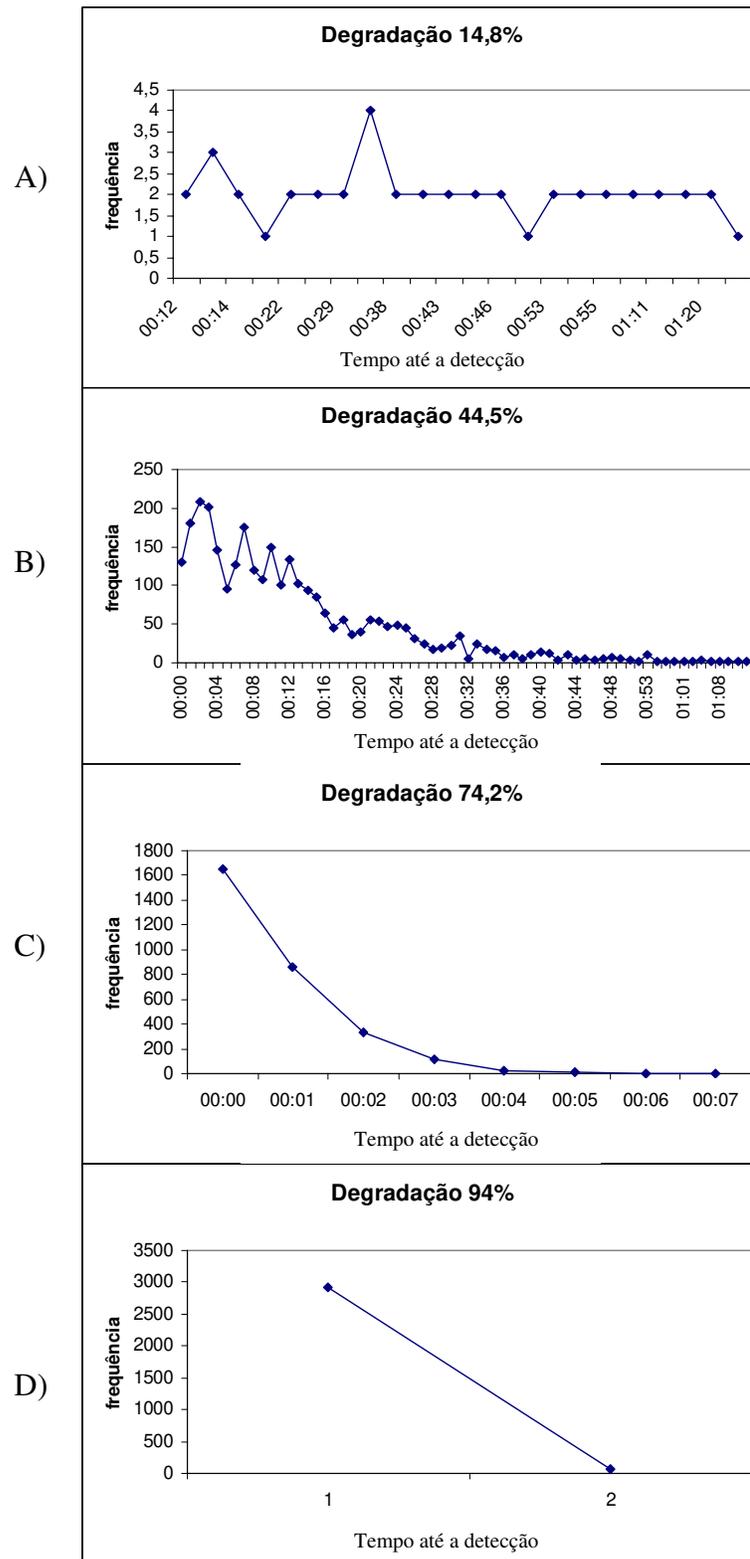


Figura 4.7 – Nível de Qualidade de 7%

- **Nível de qualidade de 12%**

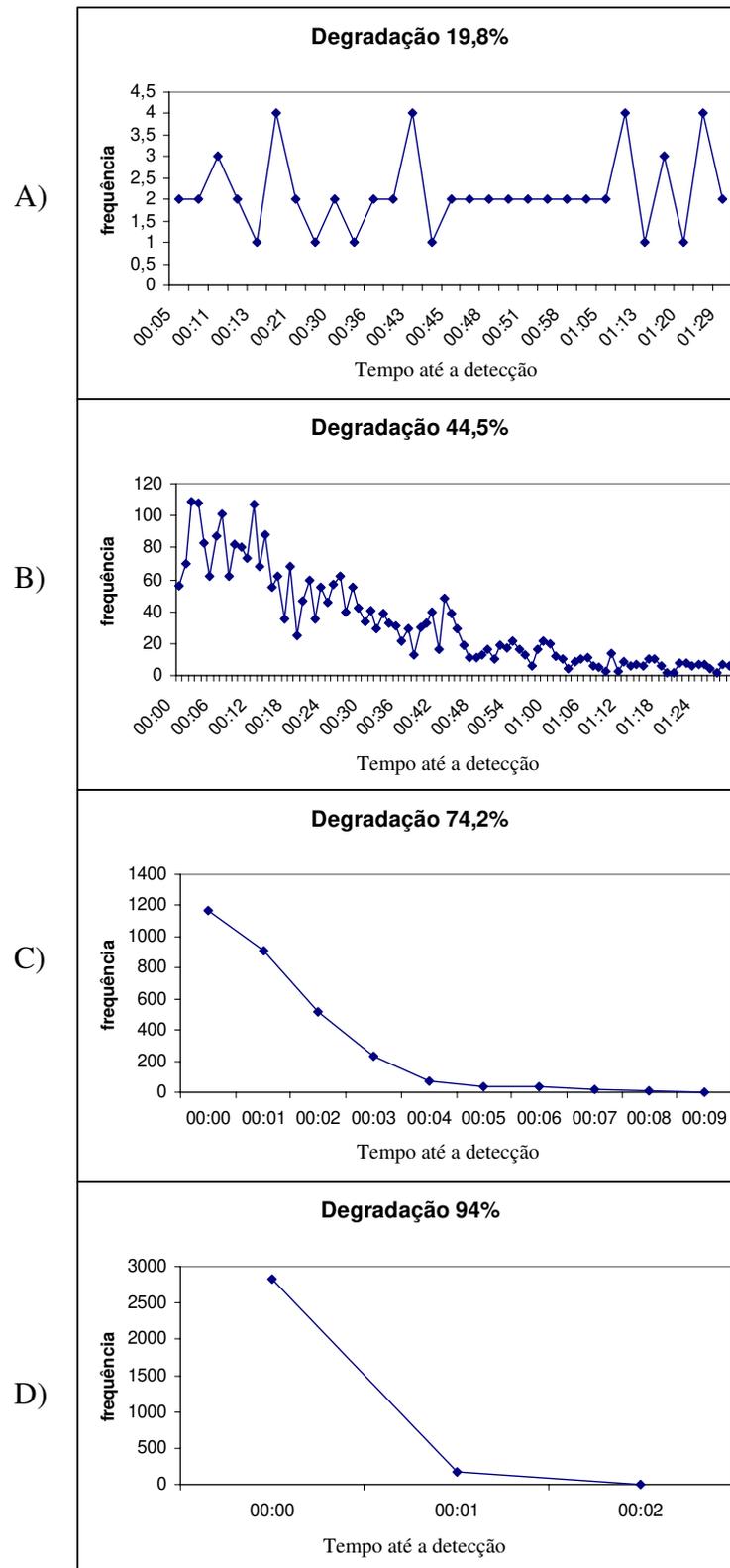


Figura 4.8 – Nível de Qualidade de 12%

- **Nível de qualidade de 17%**

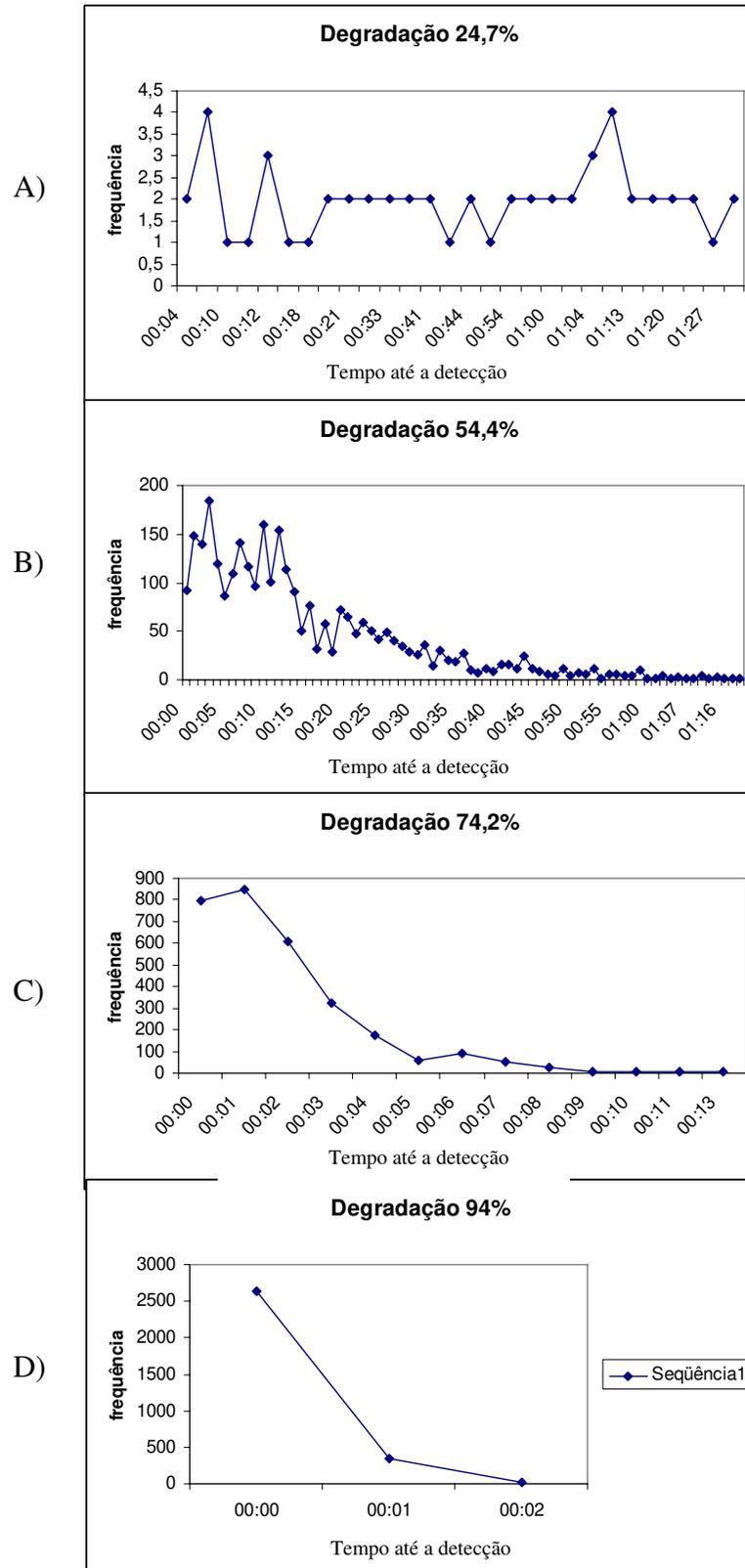


Figura 4.9 – Nível de Qualidade de 17%

- Nível de qualidade de 22%

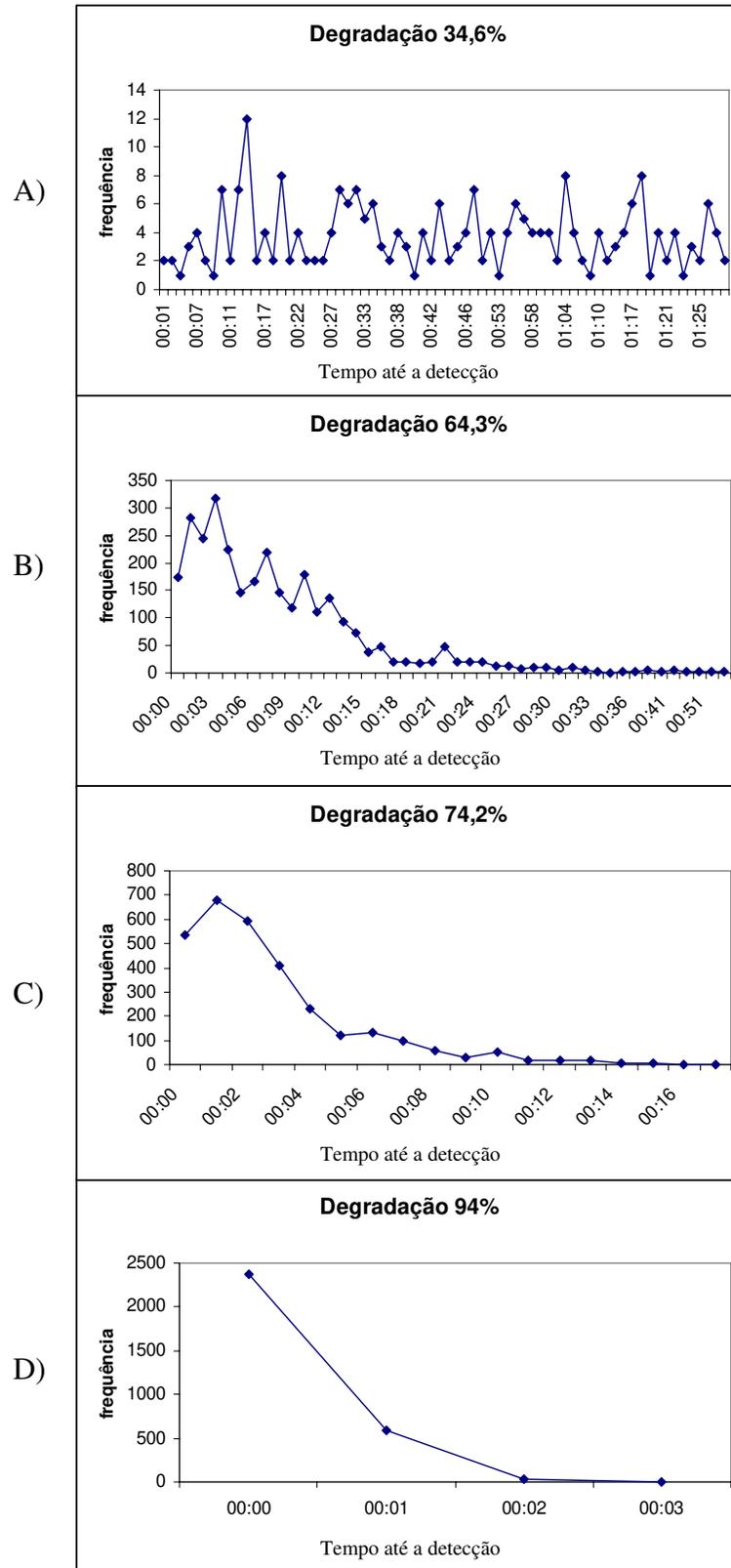


Figura 4.10 – Nível de Qualidade de 22%

- **Nível de qualidade de 27%**

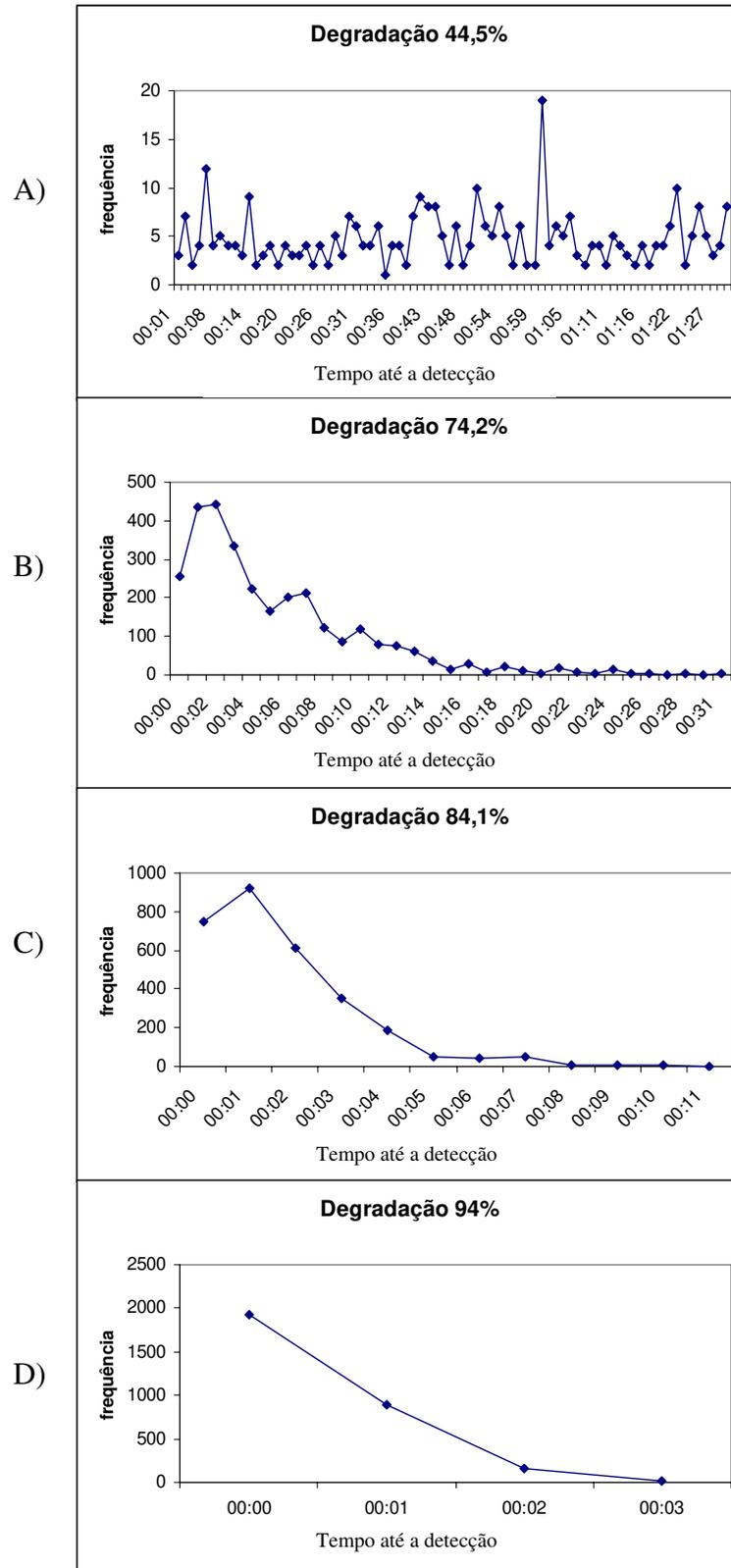


Figura 4.11 – Nível de Qualidade de 27%

É importante salientar que nem sempre são gerados alarmes quando a degradação é baixa. Dessa maneira os gráficos que representam esses valores possuem menos elementos para análise, ficando, mesmo assim, bem caracterizado o comportamento.

Observando os gráficos apresentados, é possível concluir que o comportamento do algoritmo de eventos sucessivos, em relação ao tempo de detecção, possui uma variação em função do *AQL* que está sendo utilizado e também em relação ao nível de degradação do evento no elemento. A conclusão é de que esse algoritmo é extremamente eficiente, quando houver uma grande degradação no elemento. Nesses casos a detecção do alarme é feita em poucos segundos, dependendo, logicamente, da taxa de geração de chamadas ($\lambda = \text{chamadas} / \text{segundo}$) ou geração de bilhetes no elemento. No caso estudado a detecção foi em torno de 12 segundos, com $\lambda = 0,33$ chamadas/segundo.

Por outro lado, quando a degradação não é grave, o algoritmo poderá detectar um alarme num instante qualquer. Não há como garantir em que momento a degradação na qualidade do elemento será detectada e nem se será detectada, conforme as distribuições na Figura 4.4.

4.1.2 - Algoritmo de Espaço Amostral

Esse algoritmo utiliza as seguintes distribuições de probabilidade, Binomial e Normal. O algoritmo adotado consiste na escolha de janelas de dados. O tamanho da janela define o tamanho do espaço amostral. É sobre essa quantidade de informações que são feitas as análises e tiradas as conclusões. O algoritmo não é de tempo real, pois sempre existe a necessidade de se esperar pelo preenchimento da janela para a análise ser feita.

O princípio da utilização da distribuição da Distribuição Binomial na detecção de falhas foi utilizado por Nunes [48][49]. Nossa contribuição maior neste item será testar o desempenho deste algoritmo comparando-o com o anterior.

4.1.2.1 - Distribuição Binomial

A Distribuição Binomial é baseada no modelo de um experimento com duas possíveis saídas, sucesso ou insucesso. A distribuição é utilizada da seguinte maneira: escolhe-se um evento que se queira monitorar, por exemplo, CO0. Todos os casos em que acontecer esse evento, no elemento que está sendo analisado, serão considerados como sucesso (1), e todos os eventos restantes para essa análise serão considerados como insucesso (0).

Seja X uma variável aleatória que representa um evento num experimento de Bernoulli. O espaço amostral da variável aleatória pode assumir dois valores:

$X = \begin{cases} 1 \\ 0 \end{cases}$, em que o valor 1 é quando acontece o evento, e 0 é quando não acontece. Seja

“ p ” a probabilidade de $X = 1$.

$$f_X(x; n, p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}; \quad (4.11)$$

$$f_X(x; n, p) = \frac{(n)!}{(n-x)!(x)!} \cdot p^x \cdot (1-p)^{n-x}; \quad (4.12)$$

em que n é a quantidade de elementos que estão sendo analisados.

Como no algoritmo anterior, as equações acima devem ser utilizadas para garantir que alarmes falso-positivos sejam gerados dentro de uma margem de segurança. Nesse exemplo a maneira de analisar é um pouco diferente. A margem de segurança será a mesma adotada anteriormente, ou seja, que o evento seja alarmado falso-positivamente uma vez a cada um milhão de ocorrências.

Para auxiliar o entendimento de como utilizar a margem de segurança, é necessário utilizar a Tabela 4.12. Na primeira coluna são apresentados alguns dos possíveis elementos (β) e na segunda, a Distribuição Binomial de Probabilidade de $P(X = \beta)$, que representa a probabilidade de que, num conjunto com n elementos, haja β componentes. No exemplo da Tabela 4.12 foi utilizado $n = 400$, $p = 0.017$, ou seja, 1.7%.

Na terceira coluna são apresentados os valores para a função Distribuição Cumulativa de Probabilidades, $\sum_{\beta=1}^n P(X = \beta)$, ou, então, $P(X \leq \beta)$. Essa função representa a probabilidade de que, num conjunto de n elementos, haja presente 1 ou 2 ou 3 ou, ..., β componentes.

Na quarta coluna são apresentados os valores para a função $1 - \sum_{\beta=1}^n P(X = \beta)$, ou $P(X > \beta)$. Essa função representa a probabilidade de que, num conjunto de n elementos, haja presente $\beta + 1$ ou $\beta + 2$ ou, ..., n componentes.

Beta	P(X=Beta)	P(X<=Beta)	P(X>Beta)
0	0,001050527	0,001050527	0,998949473
1	0,007267123	0,008317650	0,991682350
2	0,025072684	0,033390334	0,966609666
3	0,057525189	0,090915523	0,909084477
4	0,098737919	0,189653442	0,810346558
5	0,135239811	0,324893254	0,675106746
6	0,153973437	0,478866691	0,521133309
7	0,149878518	0,628745209	0,371254791
8	0,127331940	0,756077150	0,243922850
9	0,095912745	0,851989894	0,148010106
10	0,064855749	0,916845643	0,083154357
11	0,039766357	0,956612000	0,043388000
12	0,022293567	0,978905567	0,021094433
13	0,011507033	0,990412600	0,009587400
14	0,005501000	0,995913600	0,004086400
15	0,002448122	0,998361722	0,001638278
16	0,001018754	0,999380476	0,000619524
17	0,000397967	0,999778443	0,000221557
18	0,000146443	0,999924886	0,000075114
19	0,000050918	0,999975804	0,000024196
20	0,000016775	0,999992579	0,000007421
21	0,000005250	0,999997829	0,000002171
22	0,000001564	0,999999393	0,000000607
23	0,000000445	0,999999838	0,000000162
24	0,000000121	0,999999958	0,000000042
25	0,000000031	0,999999990	0,000000010
26	0,000000008	0,999999998	0,000000002
27	0,000000002	0,999999999	0,000000001
28	0,000000000	1,000000000	0,000000000
29	0,000000000	1,000000000	0,000000000

Tabela 4.12 – Distribuição Binomial

A maneira de aplicar a margem de segurança nessa distribuição é um pouco diferente. Existe um elemento a partir do qual a condição $1 - \sum_{\beta=1}^n P(X = \beta) < 10^{-6}$, ou $P(X > \beta) < 10^{-6}$, é satisfeita em uma Distribuição Binomial com n elementos e probabilidade p de ocorrência do evento num elemento. Esse elemento estabelece o limite, além do qual há uma certeza de 99,9999% de que o alarme que esteja acontecendo seja verdadeiro. Conceitualmente pode ser interpretado da mesma maneira que os eventos sucessivos, ou seja, em um milhão de vezes em que os bilhetes forem analisados, em janelas com tamanho de 400 elementos, e cuja probabilidade do evento no elemento é $p = 0,017$, existe a possibilidade de ser gerado um alarme falso-positivo. Tendo em vista que o universo para análise é muito menor, essa condição fornece a margem de segurança proposta para

garantir que o alarme gerado seja verdadeiro. Na Tabela 4.12, por exemplo, o valor de β , que satisfaz essa condição, é 22.

4.1.2.1.1 - Detecção de falhas

O próximo passo é utilizar o algoritmo para detectar alarmes existentes. Foi utilizada a mesma metodologia adotada anteriormente, ou seja, começar a simular falhas nos canais de voz de uma BTS. Para isso é inicializado um gerador de números aleatórios baseado numa semente cujo valor é função do relógio da CPU. Vale também a afirmação de que para esse tipo de problema a detecção é mais complexa.

A falha a ser induzida será o evento “DT” (Defeito Técnico). Será adotado como nível de qualidade, *AQL*, o valor atual da frequência relativa do evento DT apresentado na Tabela 4.2, ou seja, 1,7 ponto percentual. O valor de N a ser utilizado será 1.000.000. A quantidade de elementos para um alarme ser gerado, janelas de 50, 100 e 400 deverá ser no mínimo de:

%_AQL	Janela=400	Janela=100	Janela=50
1,7%	22	11	8

As Figuras 4.12 à 4.14 mostram que, se um evento, com nível de qualidade de 1,7%, seja numa ERB, rota, etc, começar a degradar seu *QoS*, essa degradação será observada conforme o comportamento dos gráficos ilustrados nas figuras.

Esses gráficos são o resultado de aproximadamente 5.000 experimentos, cada um com degradação aleatória dos elementos, canais de voz, de uma BTS. A média e o desvio-padrão são extraídos utilizando-se a Distribuição Normal. Esses valores serão utilizados posteriormente nas comparações com os outros algoritmos.

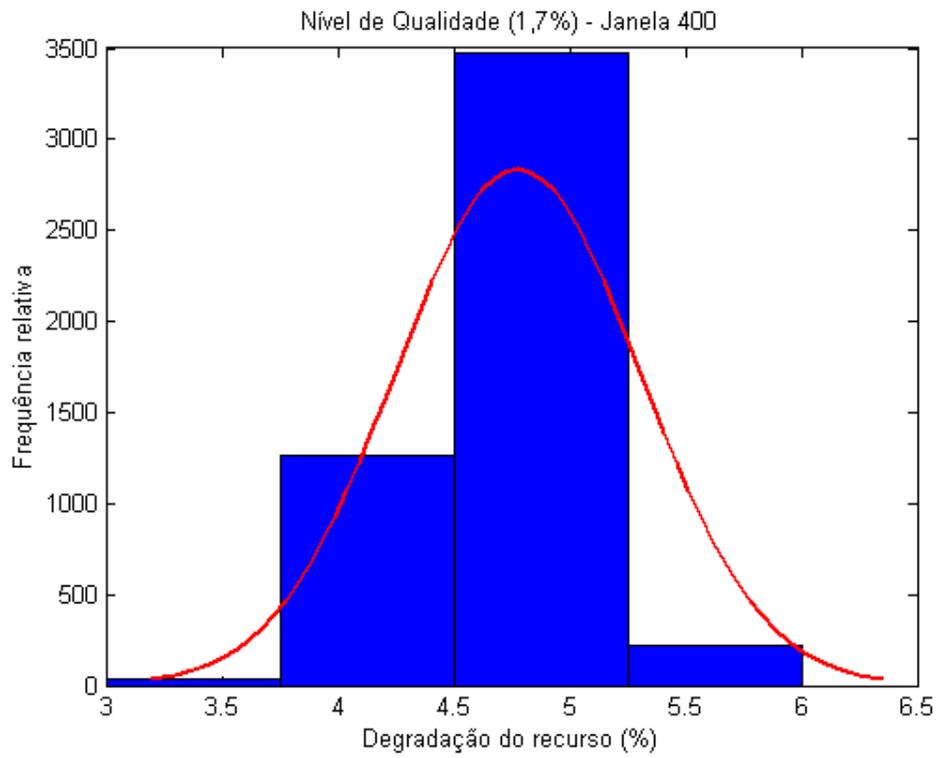


Figura 4.12 – Nível de qualidade de 1,7% (Janela 400)

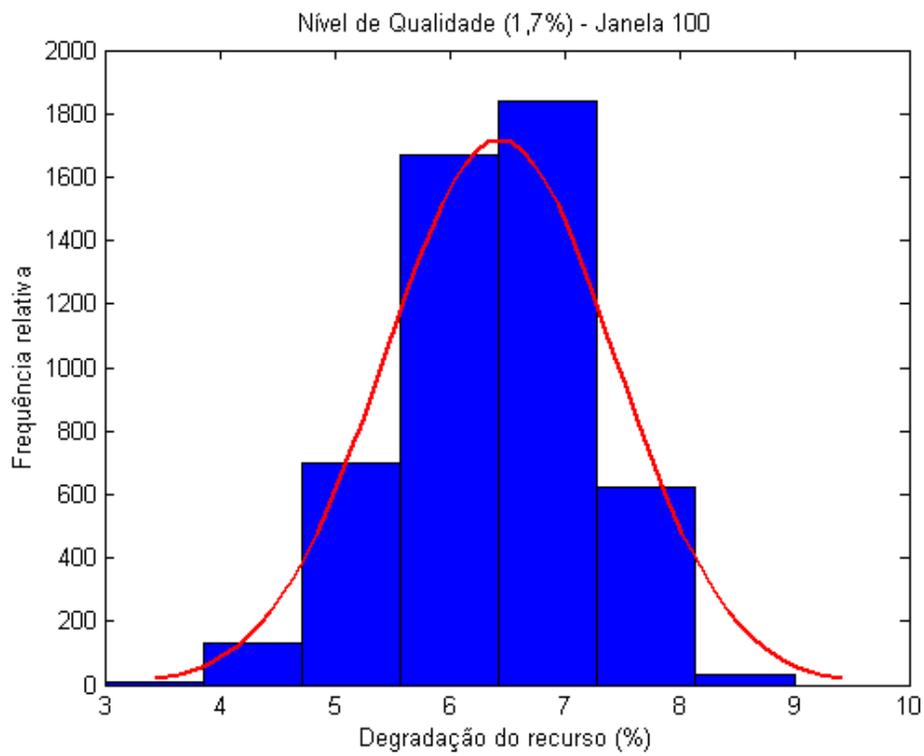


Figura 4.13 – Nível de qualidade de 1,7% (Janela 100)

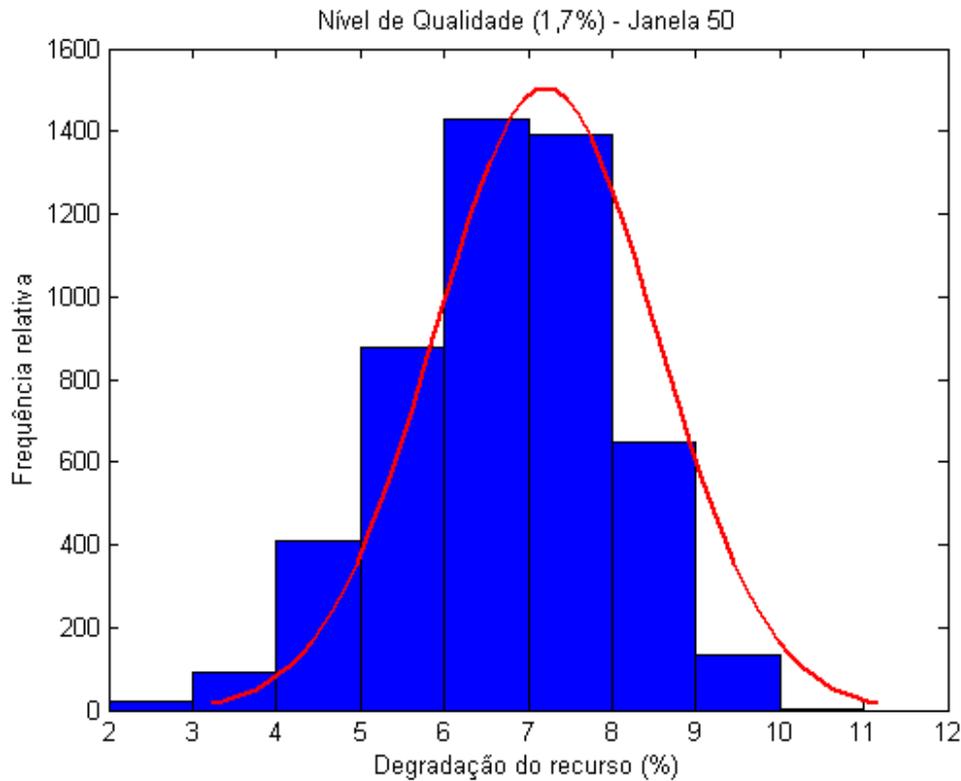


Figura 4.14 – Nível de qualidade de 1,7% (Janela 50)

Os valores para média e desvio-padrão são respectivamente, Tabela 4.13:

Janela=400		Janela=100		Janela=50	
Média	Desvio-Padrão	Média	Desvio-Padrão	Média	Desvio-Padrão
4,70%	0,939	6,40%	1,768	7,20%	2,353

Tabela 4.13 – Média e Desvio-Padrão

Uma conclusão é que quanto maior for a janela de análise, menor será a média da distribuição, bem como do desvio-padrão. Isso significa que a degradação será sentida ainda quando estiver no começo. A desvantagem de uma janela grande é a demora na detecção, ou seja, é necessário esperar que a janela termine para gerar ou não o alarme.

Para saber as probabilidades de o evento ser detectado, quando a degradação for, por exemplo, 5%, basta fazer a integral dessa Distribuição Normal até o ponto em questão (Distribuição Cumulativa de Probabilidades). Os valores deverão ser normalizados para base um, representando 100%. A Tabela 4.14 mostra a probabilidade de o alarme ser detectado em todos os valores possíveis de degradação do elemento. São calculados os valores reais, bem como os da aproximação pela Distribuição Normal:

Degradação(%)	Janela=400		Janela=100		Janela=50	
	Real	Dist.Normal	Real	Dist.Normal	Real	Dist.Normal
3%	0,8%	3,5%	0,2%	2,7%	0,3%	3,7%
4%	25,5%	22,8%	3,1%	8,7%	2,3%	8,7%
5%	94,5%	62,5%	16,4%	21,4%	10,3%	17,5%
6%	100,0%	91,7%	49,6%	41,1%	29,1%	30,5%
7%	100,0%	99,3%	85,9%	63,3%	57,8%	46,6%
8%	100,0%	100,0%	99,2%	81,7%	83,8%	63,3%
9%	100,0%	100,0%	100,0%	92,9%	97,0%	77,8%
10%	100,0%	100,0%	100,0%	97,9%	99,9%	88,3%
11%	100,0%	100,0%	100,0%	99,5%	100,0%	94,7%
12%	100,0%	100,0%	100,0%	99,9%	100,0%	97,9%
13%	100,0%	100,0%	100,0%	100,0%	100,0%	99,3%
14%	100,0%	100,0%	100,0%	100,0%	100,0%	99,8%
15%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
16%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Tabela 4.14 – Distribuição Cumulativa de Probabilidades

As Figura 4.15, 4.16 e 4.17 representam os gráficos para diferentes níveis de qualidades: 1%, 2%, 5%, 7%, 12%, 17%, 22%, 27%, 32%, 37%, 42%, 47%, 52%, 57%, 62%, 67%, 72%, 77% e 82%. Os gráficos são apresentados em função da degradação da qualidade do elemento, representada no eixo y . Foram feitos cerca de 5.000 experimentos para cada nível de qualidade. Os resultados são distribuições que se assemelham a Distribuições Normais.

Da mesma forma que nos eventos sucessivos, o desvio-padrão, nos níveis de qualidade na região de 1% a 5%, possui valores inferiores, o que representa um estreitamento da distribuição. Isso também pode ser observado, com maior ênfase, nas regiões superiores.

Nessas figuras fica também evidente o que foi dito anteriormente: para janelas maiores a média da degradação para cada nível de AQL é menor do que para janelas menores. Isso é válido também para o desvio-padrão.

Diferentes Níveis de Qualidade (Janela=400)

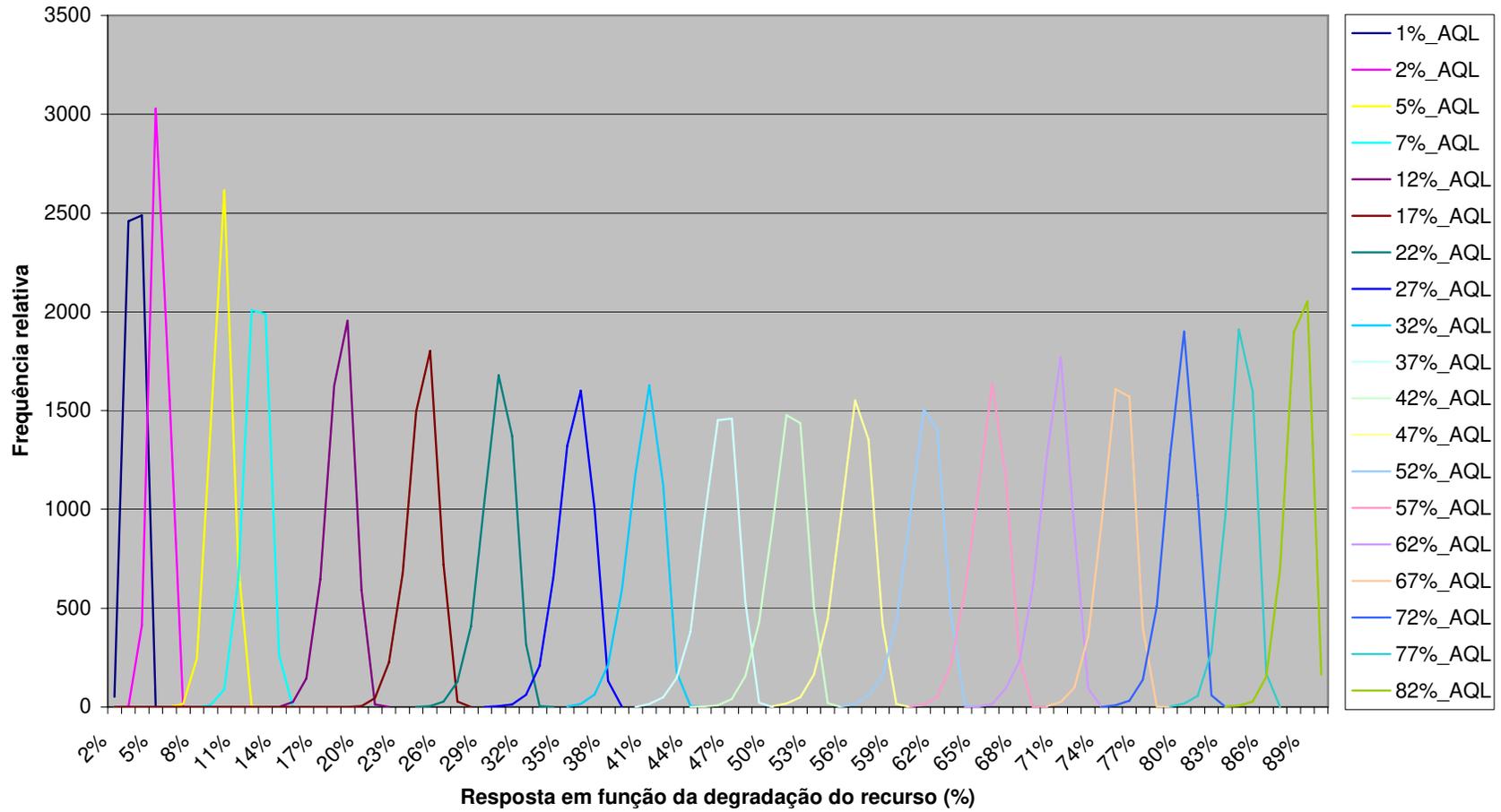


Figura 4.15 – Respostas para diferentes níveis de qualidade (Janela=400)

Diferentes Níveis de Qualidade (Janela=100)

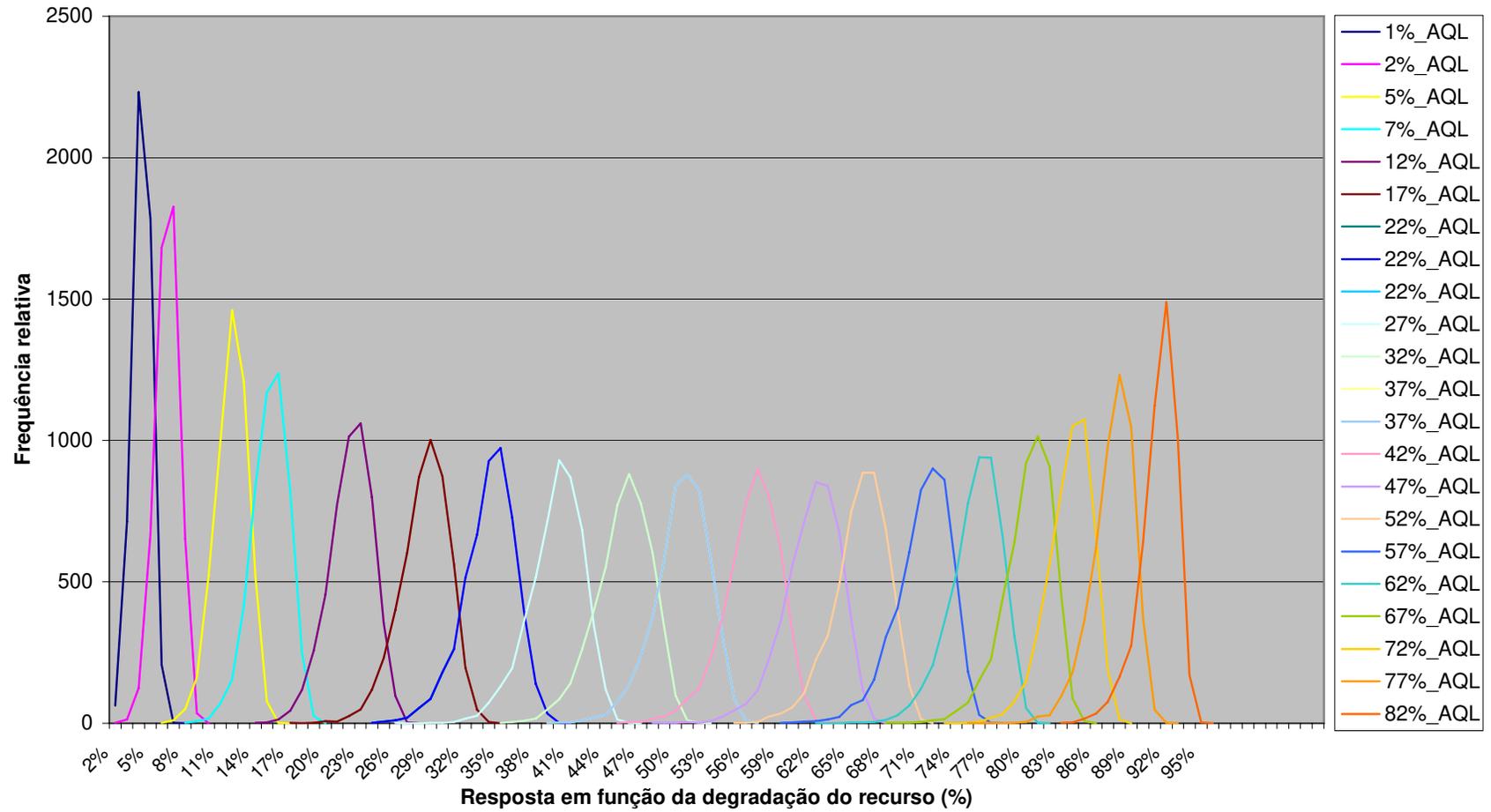


Figura 4.16 – Respostas para diferentes níveis de qualidade (Janela=100)

Diferentes Níveis de Qualidade (Janela=50)

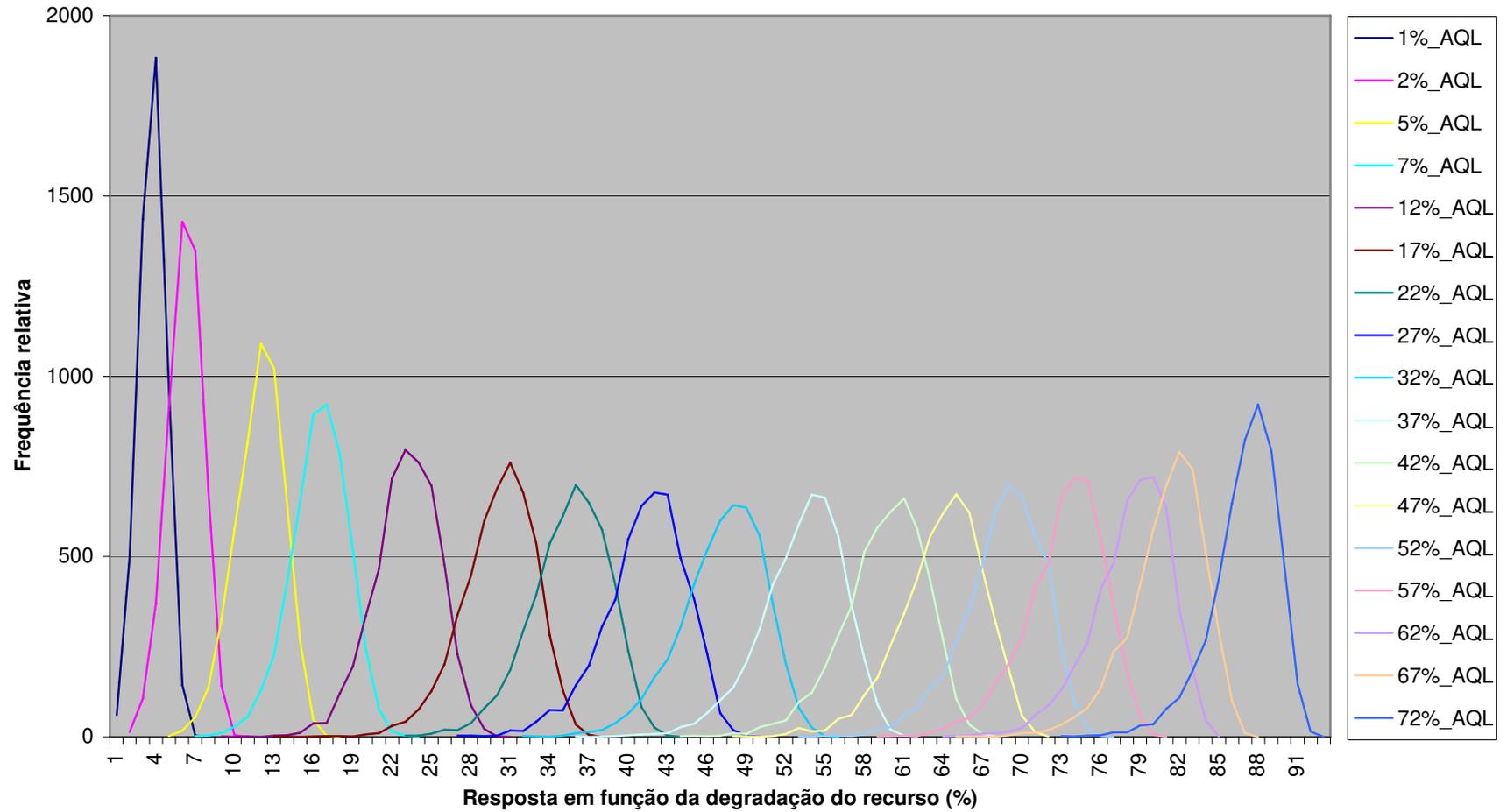


Figura 4.17 – Respostas para diferentes níveis de qualidade (Janela=50)

Algoritmos para Processamento dos Dados

Na Tabela 4.15 está a representação dos diferentes níveis de qualidade, *AQLs*, e sua correspondente aproximação por Distribuições Normais, com os respectivos desvios-padrão. A aproximação pela Normal, nesse caso, tem como único objetivo obter valores médios para efeito de comparação do comportamento.

%_AQL	Janela=400		Janela=100		Janela=50	
	Média	Des.Padrão	Média	Des.Padrão	Média	Des.Padrão
1%	3,5%	0,923	4,3%	1,424	4,7%	1,810
2%	5,2%	1,044	6,5%	1,777	7,2%	2,330
5%	9,8%	1,364	11,9%	2,466	12,9%	3,220
7%	12,3%	1,497	15,2%	2,820	17,1%	3,919
12%	18,4%	1,769	22,2%	3,333	24,1%	4,445
17%	24,4%	1,970	28,5%	3,752	31,1%	4,960
22%	29,9%	2,114	34,0%	3,962	36,5%	5,281
27%	35,5%	2,214	39,7%	4,128	42,2%	5,520
32%	40,6%	2,266	45,5%	4,274	48,2%	5,645
37%	46,0%	2,346	50,5%	4,148	54,2%	5,702
42%	51,0%	2,334	56,5%	4,303	60,5%	5,722
47%	55,9%	2,343	61,7%	4,307	64,9%	5,719
52%	61,0%	2,358	65,8%	4,138	69,6%	5,449
57%	65,7%	2,265	71,2%	4,205	74,2%	5,341
62%	70,5%	2,216	75,7%	3,943	79,3%	5,262
67%	75,1%	2,094	80,2%	3,784	82,0%	5,074
72%	79,6%	1,957	83,7%	3,407	87,9%	4,504
77%	84,0%	1,803	87,4%	3,163		
82%	88,3%	1,610	91,4%	2,813		

Tabela 4.15 – Média e Desvio-Padrão

Nas Figuras 4.18 e 4.19 são ilustradas as representações gráficas dos valores da média e desvio-padrão, em função das janelas de 50, 100 e 400 elementos. Essas figuras servem para mostrar, com mais clareza, as diferenças em função do tamanho das janelas, comentado anteriormente.

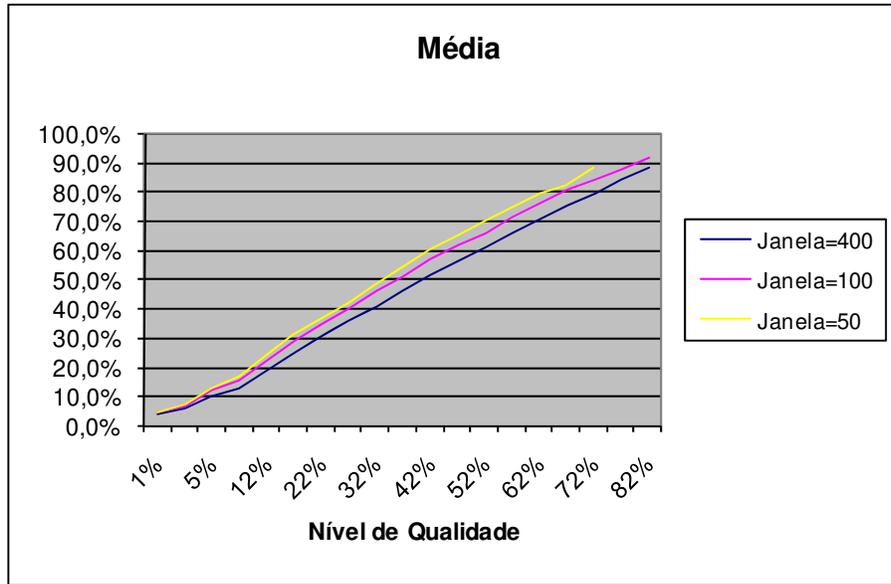


Figura 4.18 – Média para as diferentes janelas adotadas

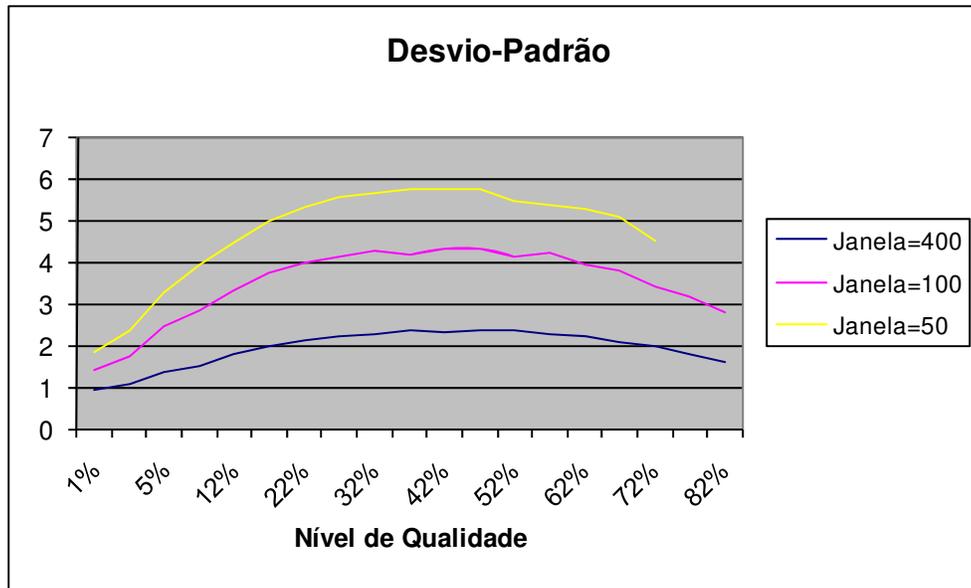


Figura 4.19 – Desvio-Padrão para as diferentes janelas adotadas

4.1.2.1.2 - Tempo de detecção

Foi feito um estudo sobre o tempo para detecção dos alarmes. Será descrito apenas um nível de qualidade, *AQL*, usualmente encontrado nas redes de telecomunicações: 2%. Com esse valor é possível comparar o desempenho entre os algoritmos.

Algoritmos para Processamento dos Dados

Foram analisadas as janelas com tamanho de 50, 100 e 400 elementos. O estudo foi feito degradando-se o elemento e vendo a resposta do algoritmo, em relação ao tempo necessário para a detecção.

A análise foi elaborada sobre aproximadamente dois mil e trezentos experimentos. É importante salientar que nem sempre são gerados alarmes, quando a degradação é baixa. Dessa maneira os gráficos que representam esses valores possuem menos elementos para análise.

O tempo está no formato hh:mm (hora e minuto).

As Figuras 4.20, 4.21 e 4.22 mostram o comportamento do algoritmo de Espaços Amostrais – Distribuição Binomial, em relação ao tempo de detecção, veja Tabela 4.16.

			Degradação			
Figura 4.20	Janela 50 elementos	Nível de qualidade 2%	A) 5,9%	B) 8,9%	C) 11,9%	D) 23,8% - 29,8% e 35,7%
Figura 4.21	Janela 100 elementos	Nível de qualidade 2%	A) 5,9%	B) 8,9%	C) 11,9%	D) 17,8% e 20,8%
Figura 4.22	Janela 400 elementos	Nível de qualidade 2%	A) 5,9%	B) 8,9%		

Tabela 4.16 – Tempos de Detecção em função do Tamanho da Janela, do Nível de Qualidade e da Degradação

- Janela de 50 elementos – Nível de qualidade de 2%

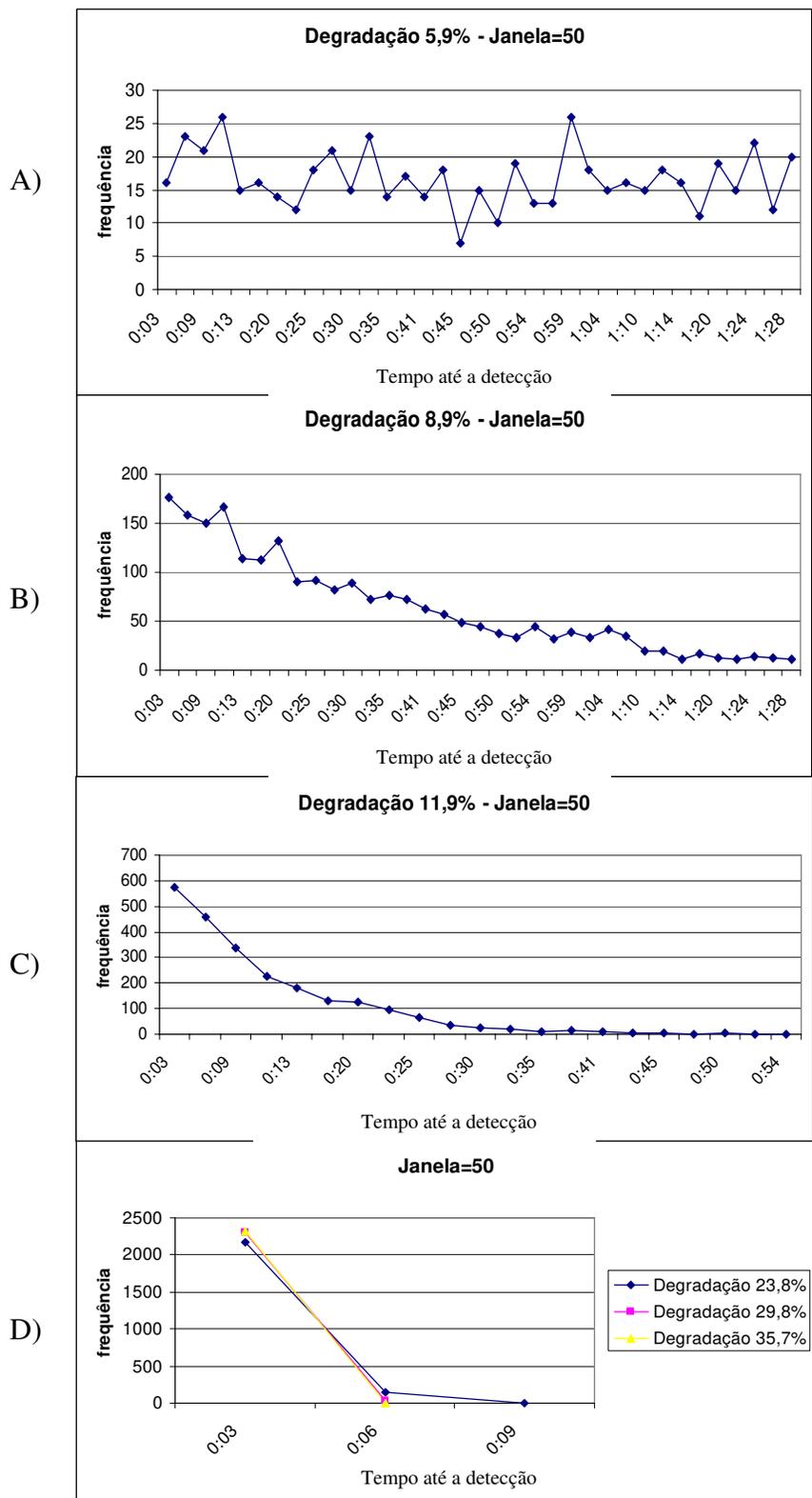


Figura 4.20 – Janela de 50 elementos e nível de qualidade de 2%

- **Janela de 100 elementos – Nível de qualidade de 2%**

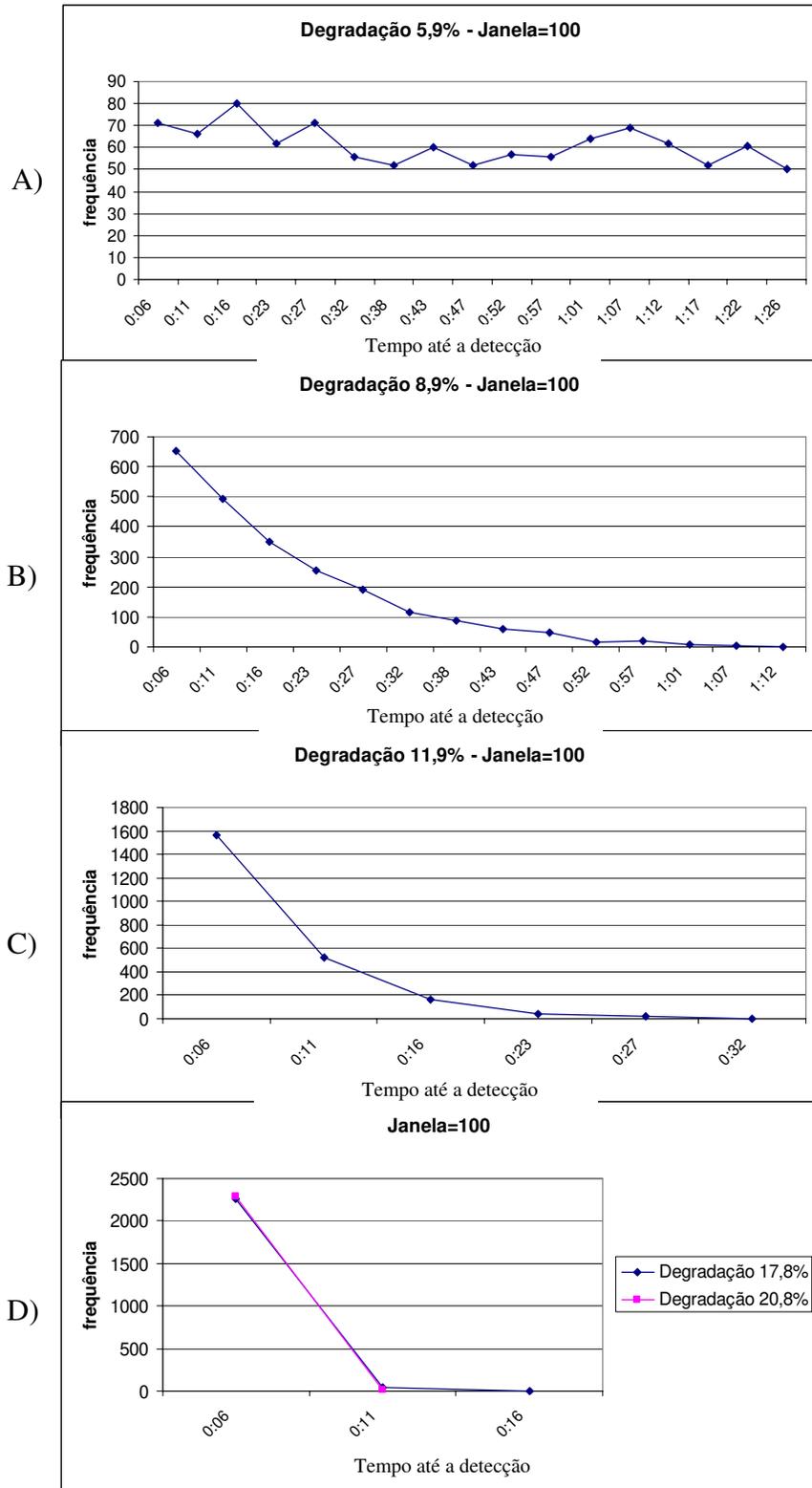


Figura 4.21 – Janela de 100 elementos e nível de qualidade de 2%

- **Janela de 400 elementos – Nível de qualidade de 2%**

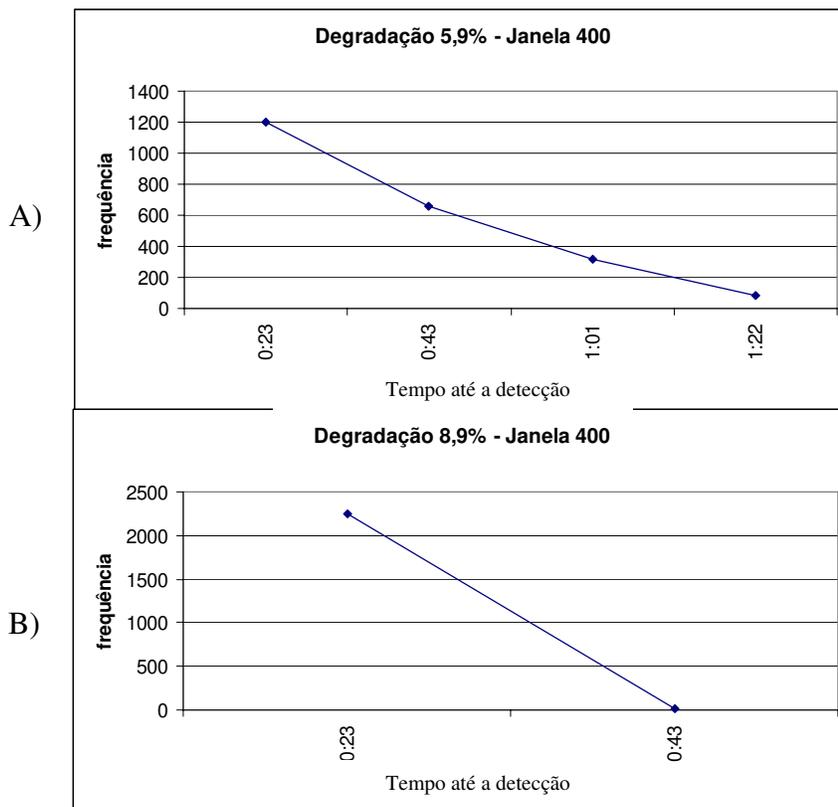


Figura 4.22 – Janela de 400 elementos e nível de qualidade de 2%

Para degradações superiores a 8,9% todos os alarmes são detectados no tempo de 23 minutos.

A análise mostra que quanto maior é o tamanho da janela de dados, menor é a degradação necessária para detectar o alarme, porém maior é o tempo necessário, ou seja, será necessário mais tempo, para que o alarme seja detectado. Também pode ser observado que o algoritmo gera alarmes para pequenas degradações na qualidade do elemento.

A conclusão é que esse algoritmo é bastante eficiente na detecção de pequenas degradações na qualidade. O algoritmo pode ser implementado, utilizando uma variada quantidade de tamanhos para as janelas. Aqui são apresentados os tamanhos 50, 100 e 400, porém podem ser utilizados outros valores. Isso faz com que haja uma complementaridade na detecção, abrangendo os mais diferentes comportamentos.

4.1.3 - Algoritmo para Chamadas OK

O algoritmo para chamadas OK não deve trabalhar sobre o crescimento do percentual do evento, e sim o contrário. O crescimento do percentual indica uma melhora no sistema e não degradação. Os mesmos raciocínios feitos até agora são válidos, porém com o sentido complementar, ou seja, é possível aplicar tanto o algoritmo de Tempo Real, quanto o de Espaços Amostrais.

Entretanto é importante acrescentar mais informações em relação às ligações OK. Nem toda ligação estabelecida e com conversação atinge o seu objetivo, que é satisfazer em 100% o desejo dos assinantes A e B de conversarem, ou seja, para determinadas chamadas OK o comportamento está mascarado pelo sistema. É o caso de algumas ligações com tempo curto ou longo de conversação. Esses comportamentos acontecem particularmente em juntores chamados “anormais”. Eles são “anormais”, tanto no aspecto de cortar a conversação alguns segundos depois que ela é iniciada, quanto no sentido de reter o juntor com o status de ocupado, mesmo após a conversação entre os usuários ter sido terminada. Nesse segundo sentido, ele é “anormal”, pois fica retendo o juntor e impossibilitando que novas ligações trafeguem por ele.

Esses comportamentos podem ser detectados. Para isso é necessário recorrer à Figura 4.23, que é uma reprodução da Figura 8.5, definida no Apêndice D, onde é tratada a função densidade de probabilidade para as ligações OK, em função do tempo de conversação. A figura ilustra o comportamento em função do tempo de conversação das chamadas. O eixo vertical apresenta a probabilidade das chamadas, enquanto o eixo horizontal apresenta o tempo em segundos cursado pelas chamadas. A figura foi obtida com 64888 bilhetes Ok de uma central localizada na cidade de Blumenau-SC. A equação da exponencial negativa aproximada é $p_T(t) = 0,023e^{-0,023t}$, onde o tempo médio de conversação é $t_m = \frac{1}{\mu} = \frac{1}{0,023} = 43,48$ segundos.

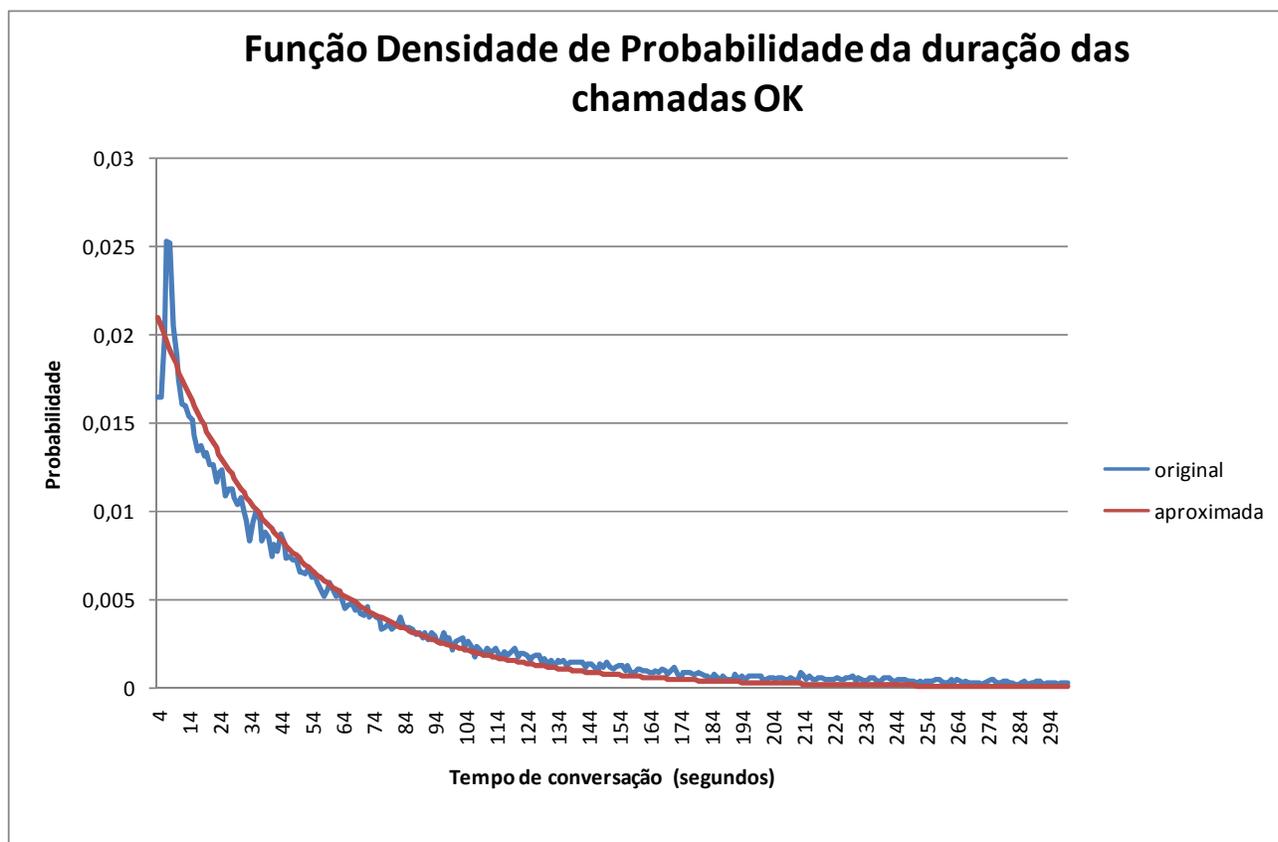


Figura 4.23 – Tempo de conversação para chamadas OK

Premissa adotada: capturar os juntores com comportamento “anormal”, com tempo igual ou inferior a 5 e superior a 1.000 segundos. Vale lembrar que o espaço amostral em relação a esse algoritmo se restringe às chamadas OK. Analisando a Tabela 4.17, é possível encontrar que a probabilidade de $P(X \leq 5)$ é de 0,0313, ou 3,13%. Já para o caso de $P(X > 1000)$ o valor da probabilidade é de 0,0036, ou 0,36% .

Beta (segundos)	P(X=Beta)	P(X<=Beta)	P(X>Beta)
4	0,016520774	0,016520774	0,983479226
5	0,016428307	0,032949081	0,967050919
6	0,02001911	0,052968191	0,947031809
7	0,025305141	0,078273333	0,921726667
8	0,025212674	0,103486007	0,896513993
9	0,02054309	0,124029096	0,875970904
10	0,018847861	0,142876957	0,857123043
20	0,01329984	0,290870423	0,709129577
30	0,010417951	0,406854888	0,593145112
40	0,008537788	0,499984589	0,500015411
50	0,006595981	0,576208236	0,423791764
60	0,005964123	0,637298114	0,362701886
70	0,004176427	0,685843299	0,314156701
80	0,00329799	0,724386635	0,275613365
90	0,002881889	0,75816792	0,24183208
100	0,002142153	0,785969671	0,214030329
200	0,000724325	0,916856738	0,083143262
300	0,00026199	0,958497719	0,041502281
400	9,2467E-05	0,975249661	0,024750339
501	4,62335E-05	0,984311429	0,015688571
1000	1,54112E-05	0,996347553	0,003652447
2016	1,54112E-05	0,999244853	0,000755147
3026	1,54112E-05	0,999568487	0,000431513
5355	1,54112E-05	0,999707188	0,000292812

Tabela 4.17 – Histograma das chamadas OK em função tempo conversação

4.1.3.1 - Caso $P(X \leq 5)$

Analisando quais os valores necessários para a utilização dos algoritmos de Tempo Real e Espaços Amostrais:

4.1.3.1.1 - Tempo Real

Utilizando a fórmula $r = \left\lceil \frac{-\ln[N \cdot (1-p) + 1]}{\ln p} \right\rceil$, lembrando que $N=1.000.000$ e $p=3,13\%$,

chegamos que r deve ser no mínimo igual a quatro elementos.

Dessa maneira, ao monitorar o universo de chamadas OK no elemento junto, toda vez que chegarem quatro bilhetes consecutivos, com temporização igual ou inferior a cinco segundos, um alarme poderá ser gerado sobre uma anomalia.

4.1.3.1.2 - Espaços Amostrais

Para esse caso é necessário encontrar a quantidade de bilhetes OK, com temporização igual ou inferior a cinco segundos, para janelas com tamanhos de 50, 100 e 400 elementos, e margem de segurança igual ao caso anterior, 1.000.000 e $p=3,29\%$.

Calculando através da Distribuição Binomial, chega-se aos seguintes valores, Tabela 4.18:

Probabilidade	Janela=400	Janela=100	Janela=50
3,29%	33	15	10

Tabela 4.18 – Valores para a Distribuição Binomial c/ $p=3,29\%$

4.1.3.2 - Caso $P(X > 1000)$

Analisando quais os valores necessários para a utilização dos algoritmos de Tempo Real e Espaços Amostrais:

4.1.3.2.1 - Tempo Real

Utilizando a fórmula $r = \left\lceil \frac{-\ln[N \cdot (1-p) + 1]}{\ln p} \right\rceil$, lembrando que $N=1.000.000$ e $p=0,36\%$,

chega-se que r deve ser no mínimo igual a três elementos.

Dessa maneira, ao monitorar o universo de chamadas OK no elemento juntor, toda vez que chegarem três bilhetes consecutivos com temporizações superiores a mil segundos, um alarme poderá ser gerado sobre uma anomalia.

4.1.3.2.2 - Espaços Amostrais

Para esse caso é necessário encontrar a quantidade de bilhetes OK, com temporização superior a mil segundos, para janelas com tamanhos de 50, 100 e 400 elementos, e margem de segurança igual ao caso anterior, 1.000.000 e $p=0,36\%$.

Calculando através da Distribuição Binomial são obtidos os seguintes valores, Tabela 4.19:

Probabilidade	Janela=400	Janela=100	Janela=50
0,36%	10	6	5

Tabela 4.19 - Valores para a Distribuição Binomial c/ $p=0,36\%$

Para ambos os casos, $P(X \leq 5)$ e $P(X > 1000)$, não serão feitos os cálculos do desempenho e temporização que já foram amplamente cobertos.

Esclarecendo um ponto importante: todas as ligações com conversação igual ou inferior a três segundos não são consideradas chamadas OK num sistema de telefonia celular, conforme deliberação da ANATEL [50]. Esse tipo de chamada é classificada no decorrer do trabalho como POK (Parcialmente OK).

Outra observação pertinente é que esse estudo pode ser replicado em relação aos tempos de outros eventos, conforme descrito no Apêndice A. Para os outros eventos faz sentido pensar no tempo de estabelecimento e não no tempo de conversação.

4.1.4 - Algoritmo para Detecção de Ocupação

O algoritmo para Detecção de Ocupação é complementar aos demais modelos. O objetivo desse quarto algoritmo é preencher uma lacuna existente sobre o monitoramento dos elementos, em caso de bloqueio de tráfego. Um exemplo disso é a queda de uma rota que interliga centrais ou mesmo, uma central com uma estação rádio base. A queda desse tipo de elemento pode ser tanto física, quanto lógica. A consequência é uma só: interrupção do tráfego de voz/dados nos circuitos. Outro exemplo é quando, por uma falha qualquer, um elemento começa a bloquear o tráfego das ligações, indisponibilizando o elemento.

O algoritmo se baseia na taxa de chamadas λ , que circulam por cada elemento para estimar a presença ou não de um comportamento anômalo, no menor intervalo de tempo possível e com a maior garantia possível de que o alarme seja verdadeiro. O algoritmo pode ser empregado para monitorar os mais diferentes elementos, tanto lógicos como físicos, dentro de um sistema de comunicação:

- Circuitos de Entrada (elemento físico) – circuitos de entrada por onde será estabelecida uma ligação telefônica;
- Circuitos de Saída (elemento físico) – circuitos de saída por onde será estabelecida uma ligação telefônica;
- Transceptor MFC (elemento físico) – responsável pela sinalização MFC (Multi Freqüencial Compelida);
- Troncos de entrada (elemento físico) – geralmente formado por um E1 (30 canais) de entrada;
- Troncos de saída (elemento físico) – geralmente formado por um E1 (30 canais) de saída;

- Rotas de Entrada (elemento físico) – conjunto de circuitos de entrada que formam uma rota;
- Rotas de Saída (elemento físico) – conjunto de circuitos de saída que formam uma rota;
- BSC (elemento físico) – Base Station Controle é um módulo de controle das BTS CDMA;
- BTS origem (elemento físico) – Base Transceiver Station é de onde são originadas as ligações na tecnologia CDMA;
- BTS destino (elemento físico) – Base Transceiver Station é de onde são recebidas as ligações na tecnologia CDMA;
- ERB origem (elemento físico) – estação rádio base de onde são originadas as ligações na tecnologia AMPS/TDMA;
- ERB destino (elemento físico) – estação rádio base onde são terminadas as ligações na tecnologia AMPS/TDMA;
- Canais de RF origem (elemento físico) – canal de rádio frequência pelo qual a ligação foi originada;
- Canais de RF destino (elemento físico) – canal de rádio frequência pelo qual a ligação foi recebida;
- Assinante Origem (elemento físico) – número do telefone que originou a chamada;
- Assinante Destino (elemento físico) – número do telefone que recebeu a chamada;
- Direções das Chamadas de Origem (elemento lógico) – faixa de telefones pode ser um bairro, uma região, uma cidade, um estado, um país da qual as chamadas são originadas;
- Direções das Chamadas de Destino (elemento lógico) - faixa de telefones pode ser um bairro, uma região, uma cidade, um estado, um país para os quais as chamadas são destinadas;

Esse algoritmo trabalha com o mesmo princípio apresentado anteriormente, garantindo que alarmes falso-positivos sejam gerados dentro de um limite pré-estabelecido.

4.1.4.1 - O Problema

Um dos problemas que levaram a desenvolver um algoritmo para detecção de elementos com bloqueio de tráfego foi a observação de uma plataforma de valor agregado de uma operadora. Foi observado que determinados juntores, circuitos de um tronco, de um momento para o outro,

interrompiam o tráfego. A análise foi sendo refinada, até ficar claro que o problema estava na plataforma, embora não se possa afirmar a causa raiz. O interesse nesse caso e em todos os outros possíveis casos similares não é encontrar a causa raiz do problema e sim criar mecanismos para detecção da falha.

A seguir são apresentadas algumas definições necessárias ao entendimento do trabalho.

4.1.4.2 - Tráfego

O tráfego telefônico é não-estacionário, o que tende a levar o modelo a ser complexo. No entanto, durante a hora de maior movimento (HMM), pode-se supor que a intensidade de tráfego permaneça aproximadamente constante. Dessa forma, a quantidade de ligações simultâneas existentes varia estatisticamente em torno de um valor médio, denominado intensidade de tráfego. Nos estudos da teoria de Tráfego, isso corresponde à condição de que o tráfego na central de comutação se encontra em estado de equilíbrio estatístico ou em estado estacionário [54][55].

Processos aleatórios podem ser classificados de uma maneira preliminar em duas categorias: processos aleatórios estacionários e processos aleatórios não-estacionários. Por seu turno, processos aleatórios estacionários podem ser classificados como ergódicos ou não ergódicos. Os processos aleatórios não estacionários podem ser categorizados em termos de propriedades estatísticas não-estacionárias que fogem do escopo do presente trabalho.

Quando um fenômeno qualquer é estudado em termos de um processo aleatório, as propriedades de tal fenômeno podem hipoteticamente ser descritas, em qualquer instante de tempo, por meio de grandezas estatísticas calculadas sobre uma coleção de realizações do fenômeno. De uma maneira geral, pode-se dizer que um dado processo é estacionário quando as grandezas estatísticas calculadas, tendo como base uma coleção de realizações do fenômeno associado, não variam, quando se tomam diferentes instantes de tempo. Ou, em outras palavras, quando as grandezas estatísticas calculadas nas realizações são invariantes com o tempo. Dessa forma, em processos aleatórios estacionários é possível calcular-se as propriedades estatísticas do fenômeno associado por meio de cálculos operados sobre coleções de realizações em instantes quaisquer de tempo.

Processos ergódicos são processos estacionários que apresentam uma propriedade adicional em relação aos processos estacionários: o cálculo das propriedades estatísticas pode ser feito por meio de médias temporais operado sobre uma única realização do fenômeno em estudo.

4.1.4.2.1 - Como medir o tráfego?

Observando a ocupação de um juntor, ao longo do tempo, encontramos dois estados possíveis: livre ou ocupado. Nos instantes em que o circuito, juntor, está livre, esse poderá ser tomado por qualquer chamada que tenha acesso a ele e vice-versa.

Imaginemos agora um sistema constituído por cinco circuitos, cujo comportamento é observado em um intervalo de uma hora. Os circuitos estão representados na parte superior do gráfico, Figura 4.24.

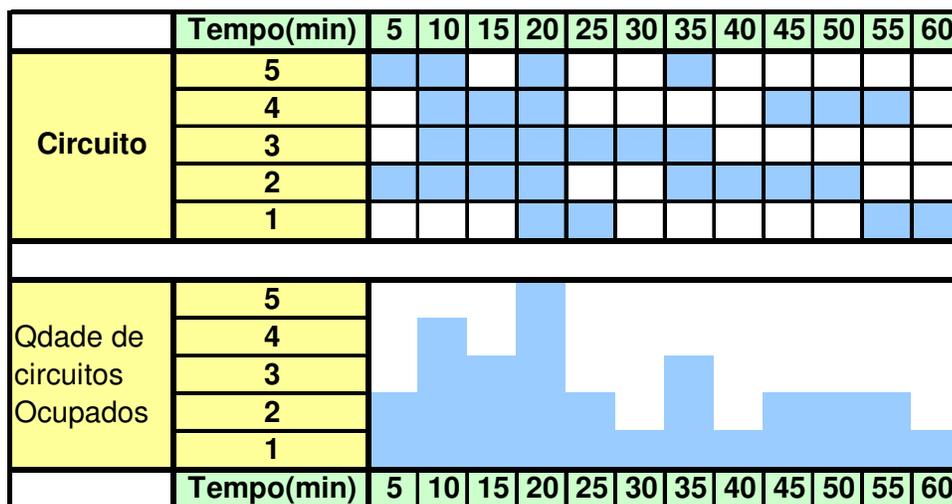


Figura 4.24 – Ocupação dos circuitos

Na parte inferior do gráfico, Figura 4.24, é apresentado um registro acumulado do número de ocupações simultâneas.

É possível observar que todos os circuitos estiveram ocupados no intervalo de tempo entre $t = 20$ e $t = 25$ minutos. Qualquer chamada que aparecer nesse intervalo será recusada. Outras definições:

- Tempo de ocupação - É o intervalo de tempo em que uma chamada está ocupando um circuito;
- Volume de tráfego - É a soma dos tempos de ocupação dos circuitos ou órgãos de um sistema, ou seja:

$$V = \sum_{j=1}^{j=n} t_j ; \tag{4.13}$$

em que n é o número total de ocupações, e t_j é o tempo de ocupação de cada chamada j sobre cada circuito. Para o caso apresentado anteriormente, o volume de tráfego é:

$$\begin{aligned} n &= 10 \\ V &= t_1 + t_2 + t_3 + \dots + t_n && ; (4.14) \\ V &= 10 + 10 + 20 + 20 + 30 + 15 + 15 + 10 + 5 + 5 = 140 \text{ min} \end{aligned}$$

- Taxa de tráfego (λ) - É o número de ocupações ou de chamadas num grupo de circuitos ou órgãos durante um período de observação. É possível observar, na Figura 4.24, que houve 10 ligações num período de observação de 60 minutos. Portanto, a taxa de tráfego é de:

$$\lambda = \frac{10}{60} = 0,166 \text{ chamadas/minuto}; \quad (4.15)$$

- Intensidade de Tráfego – É o quociente entre o volume de tráfego e o período de observação:

$$A = \frac{V}{T} = \frac{140}{60} = 2,33; \quad (4.16)$$

A esse valor acrescentamos a palavra Erlang.

Existem outras medidas de Intensidade de Tráfego, como por exemplo: CCS, que equivale a cem chamadas de 1 segundo de duração. A relação entre as unidades é a seguinte: $36 \cdot CCS = 1 \cdot Erlang$; outra medida é o EBCH, que equivale a 120 chamadas de 1 segundo de duração ou a uma chamada de 120 segundos de duração. A relação entre as unidades é a seguinte: $30 \cdot EBCH = 1 \cdot Erlang$.

Como já foi dito, o tráfego está sujeito a alterações no decorrer do dia. No Apêndice A, item D.2 - Perfis de Ocupação do Sistema, está a representação de como o tráfego varia durante o dia. A Figura 4.25 representa a quantidade de tráfego no decorrer de um dia, para uma central de controle e comutação. O eixo “y” representa a ocupação em Erlangs, e o eixo “x”, o horário durante o dia. O valor em Erlangs permite ter uma noção do tamanho da central.

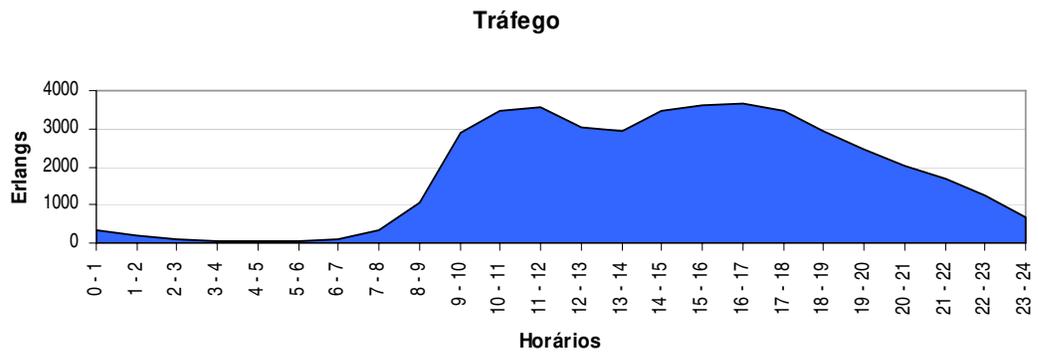


Figura 4.25 – Perfil de ocupação de uma central em uma grande capital (3 milhões de habitantes)

A partir do tráfego é possível encontrar a quantidade de chamadas por período. Basta que para isso se tenha o tempo médio de uma ligação telefônica. A Figura 4.26 é semelhante à curva de tempo de duração de chamadas OK anteriormente, Figura 4.23. A equação exponencial negativa aproximada é $p_T(t) = 0,017e^{-0,017t}$, onde o tempo médio de conversação é $t_m = \frac{1}{\mu} = \frac{1}{0,017} = 58,82$ segundos. A Função Densidade de Probabilidade foi criada em função de 74.721 chamadas de uma central localizada na cidade de São Paulo – SP.

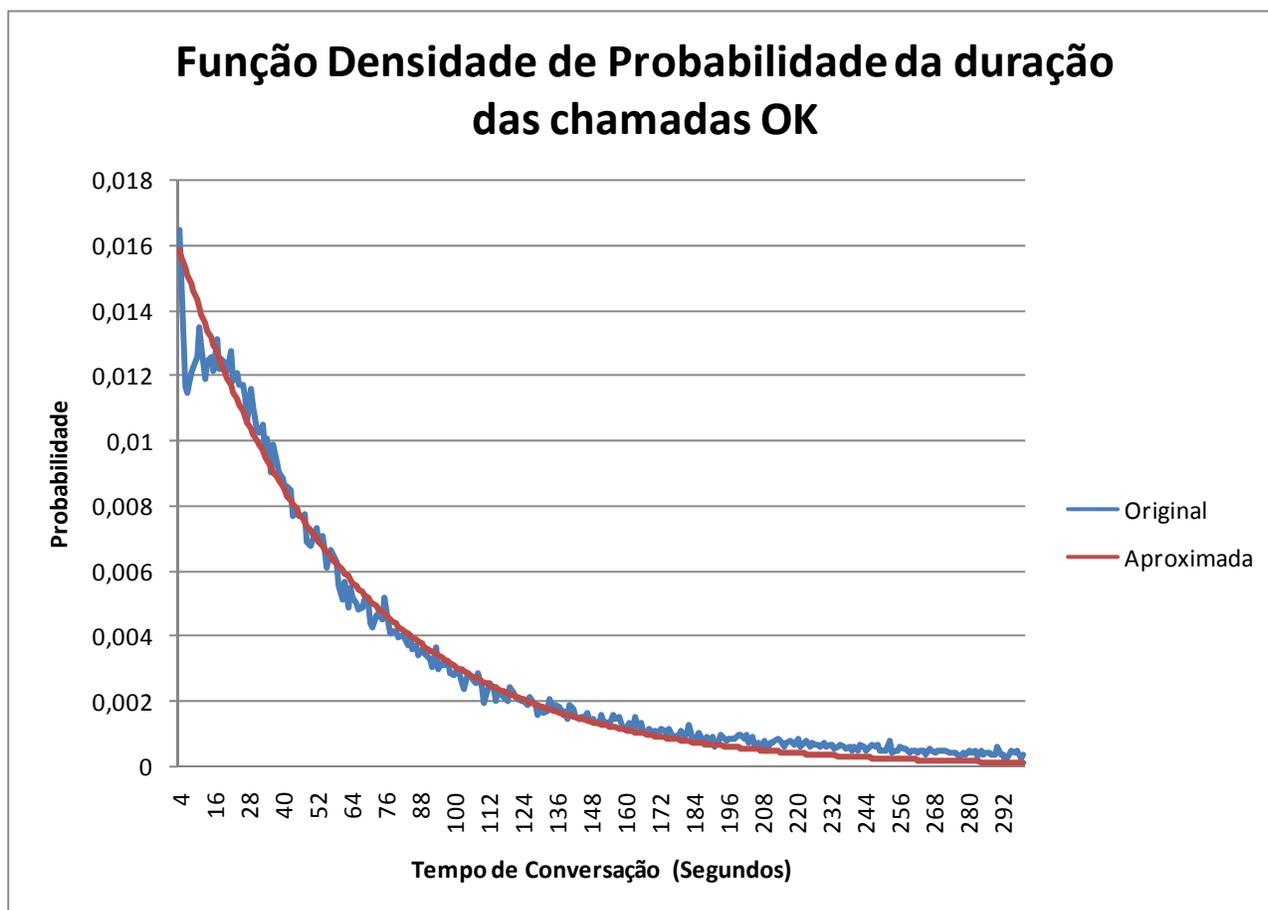


Figura 4.26 – Tempo de conversação para chamada OK

De posse do tempo médio de conversação, é possível encontrar a quantidade de chamadas por período (λ), da seguinte forma:

$$\frac{\text{erlangs} * \text{período}}{\text{tempo médio}}; \tag{4.17}$$

$$\frac{3.500 * 60}{0,98} \cong 214.286;$$

ou seja, algo em torno de 214.286 chamadas/hora.

Alguns esclarecimentos:

- 3.500 erlangs é o tráfego apresentado na Figura 4.25 no HMM;
- 0,98 representa o tempo médio de conversação expresso em minutos, o equivalente a 58,82 segundos;
- 60 representa o período de observação em minutos, o equivalente a uma hora.

A taxa de chamadas λ é utilizada pelo algoritmo descrito no próximo item.

4.1.5 - Processo de Poisson Homogêneo

Toda a Teoria de Tráfego, desde Erlang até a Teoria de Tentativas Repetidas, assume que o processo de chegadas de chamadas é Poisson Homogêneo [56][57]. Por essa razão esse é o primeiro fundamento no desenvolvimento do algoritmo. A distribuição de Poisson pode ser deduzida da Distribuição Binomial, utilizando a teoria de Limites [58], ou seja:

$$Bernoulli = \Pr\{Y = x\} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (4.18)$$

$$\lim_{n \rightarrow +\infty} \Pr\{Y = x\} = \lim_{n \rightarrow +\infty} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad (4.19)$$

Nesse caso, a probabilidade de $P_n(\Delta t)$, em que x chamadas cheguem ao juntor, no intervalo de tempo de comprimento t , satisfaz a distribuição de Poisson, conforme a equação:

$$P_n(t) = \frac{(\lambda \cdot t)^x}{x!} e^{-\lambda \cdot t} \quad (4.20)$$

O dual da distribuição de Poisson é uma distribuição exponencial. Suponha as ocorrências de chamadas obedecendo a uma distribuição de Poisson, como ilustra a Figura 4.27. Seja T a variável aleatória que representa o tempo entre as ocorrências sucessivas. Quer-se calcular a distribuição da variável aleatória T .

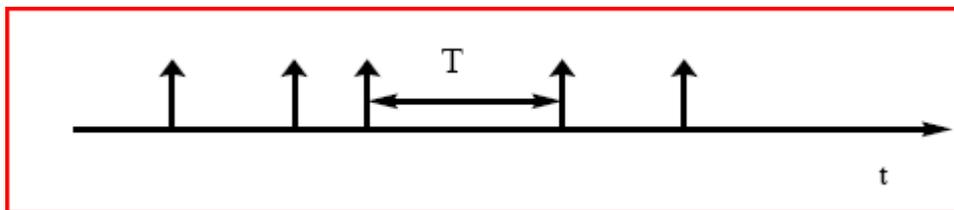


Figura 4.27 – Distribuição entre ocorrências

Seja $F_T(t) = \Pr\{T < t\}$ a função distribuição de probabilidade da variável aleatória T . A derivada dessa distribuição é a função densidade de probabilidade e será denotada por $p_T(t)$.

A função complementar $\Pr\{T > t\}$ representa a probabilidade de não haver ocorrência num intervalo de duração t .

Portanto,

$$\Pr\{T > t\} = \Pr\{\text{zero_chamadas_até_t}\} \quad (4.21)$$

$$\Pr\{T > t\} = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t} \quad (4.22)$$

$$F_T(t) = 1 - \Pr\{T > t\} = 1 - e^{-\lambda t}; t \geq 0 \quad (4.23)$$

Derivando, obtém-se

$$p_T(t) = \lambda e^{-\lambda t}; t \geq 0 \quad (4.24)$$

A distribuição do tempo entre as chamadas é uma exponencial negativa. A média e a variância da exponencial negativa são $E\{T\} = \frac{1}{\lambda}$ e $\sigma^2 = \frac{1}{\lambda^2}$.

Pode-se caracterizar o processo de chegada como poissoniana de média λ , ou, alternativamente, como exponencial negativa de média $\frac{1}{\lambda}$, dependendo se está contando o número de chamadas ou se está observando o tempo entre as chamadas.

O processo de conversação pode ser caracterizado de maneira análoga ao processo de chegada. Basta, neste caso, olhar o tronco de saída e verificar os tempos entre as partidas que podem ser modelados como exponencial negativa.

Assim,

$$F_T(t) = 1 - e^{-\mu t}; t \geq 0 \quad (4.25)$$

e

$$p_T(t) = \mu e^{-\mu t}; t \geq 0 \quad (4.26)$$

onde μ é a taxa média de término de conversação, e $t_m = \frac{1}{\mu}$ é o tempo médio de conversação.

4.1.5.1 - Probabilidade de t segundos sem ocupação

Inicialmente será analisado o intervalo de tempo, $t = 15$ segundos e $n = 0$, isto é:

$$P_0(10) = (15 \cdot \lambda)^0 \cdot \frac{e^{-15 \cdot \lambda}}{0!} = e^{-15 \cdot \lambda}; \quad (4.27)$$

O termo $P_0(15)$ é equivalente à probabilidade da primeira chamada aparecer após quinze segundos.

4.1.5.2 - Intervalos sucessivos de t segundos sem ocupação

- Adotando o mesmo princípio utilizado no Algoritmo de Tempo Real que consiste em analisar as saídas das chamadas esperando por uma sequência de eventos. É como se estivéssemos “jogando moedas aleatoriamente” e esperando por uma determinada sequência para confirmarmos um alarme;
- Algoritmo de Espaço Amostral no qual esperamos pelo preenchimento da janela de dados para então procurarmos por uma anomalia utilizando a distribuição binomial;
- Algoritmo para Chamadas Ok procura por anormalidades dentro de um tipo específico de chamada, ou melhor, chamadas que a princípio foram concluídas com sucesso. Este algoritmo engloba os dois mecanismos citados anteriormente: Tempo Real e Chamadas OK;
- Algoritmo para Detecção de Ocupação que trabalha sobre o tempo de ocorrência dos eventos para gerar o alarme. O algoritmo se baseia na distribuição de Poisson.

Algoritmo de Tempo Real, baseado na Teoria da Renovação [42]-[44], temos que:

$$r = \left\lceil \frac{-\ln[N \cdot (1-p) + 1]}{\ln p} \right\rceil \quad (4.28)$$

A equação dá como resultado a quantidade r de insucessos sucessivos que devem acontecer, para que um alarme seja gerado. A quantidade está diretamente relacionada com a probabilidade p e o valor N . O valor N é determinado da mesma forma que foi apresentado anteriormente.

A aplicação desse modelo ao presente caso é feita definindo o evento e a probabilidade p . O evento é “intervalo de t segundos sem receber chamadas” e é definido pela fórmula: $p = e^{-t \cdot \lambda}$.

Uma conclusão é que quanto maior a exigência em relação à geração do alarme, maior será a quantidade de insucessos sucessivos que deverão acontecer, para que o alarme seja gerado.

A Tabela 4.20 apresenta a probabilidade p de um elemento não receber chamadas num intervalo de t segundos, em função da taxa de chegada de chamadas λ . O valor de probabilidade é função da taxa de chamadas no elemento e da janela de tempo escolhida para monitorar o elemento.

		Lambda (chamadas/segundo)											
		0,01	0,1	1	2	4	6	8	10	15	20	30	40
tempo(seg)	1	0,990049834	0,904837418	0,367879441	0,135335283	0,018316	0,002479	0,000335	4,53999E-05	3,05902E-07	2,06115E-09	9,35762E-14	4,24835E-18
	2	0,980198673	0,818730753	0,135335283	0,018315639	0,000335	6,14E-06	1,13E-07	2,06115E-09	9,35762E-14	4,24835E-18	8,75651E-27	1,80485E-35
	3	0,970445534	0,740818221	0,049787068	0,002478752	6,14E-06	1,52E-08	3,78E-11	9,35762E-14	2,86252E-20	8,75651E-27	8,19401E-40	7,66765E-53
	4	0,960789439	0,670320046	0,018315639	0,000335463	1,13E-07	3,78E-11	1,27E-14	4,24835E-18	8,75651E-27	1,80485E-35	7,66765E-53	3,25749E-70
	5	0,951229425	0,60653066	0,006737947	4,53999E-05	2,06E-09	9,36E-14	4,25E-18	1,92875E-22	2,67864E-33	3,72008E-44	7,1751E-66	1,3839E-87
	6	0,941764534	0,548811636	0,002478752	6,14421E-06	3,78E-11	2,32E-16	1,43E-21	8,75651E-27	8,19401E-40	7,66765E-53	6,71418E-79	5,8793E-105
	7	0,93239382	0,496585304	0,000911882	8,31529E-07	6,91E-13	5,75E-19	4,78E-25	3,97545E-31	2,50657E-46	1,58042E-61	6,28288E-92	2,4977E-122
	8	0,923116346	0,449328964	0,000335463	1,12535E-07	1,27E-14	1,43E-21	1,6E-28	1,80485E-35	7,66765E-53	3,25749E-70	5,8793E-105	1,0611E-139
	9	0,913931185	0,40656966	0,00012341	1,523E-08	2,32E-16	3,53E-24	5,38E-32	8,19401E-40	2,34555E-59	6,71418E-79	5,5016E-118	4,508E-157
	10	0,904837418	0,367879441	4,53999E-05	2,06115E-09	4,25E-18	8,76E-27	1,8E-35	3,72008E-44	7,1751E-66	1,3839E-87	5,1482E-131	1,9152E-174
	11	0,895834135	0,332871084	1,67017E-05	2,78947E-10	7,78E-20	2,17E-29	6,05E-39	1,68891E-48	2,19488E-72	2,85242E-96	4,8175E-144	8,1363E-192
	12	0,886920437	0,301194212	6,14421E-06	3,77513E-11	1,43E-21	5,38E-32	2,03E-42	7,66765E-53	6,71418E-79	5,8793E-105	4,508E-157	3,4566E-209
	13	0,878095431	0,272531793	2,26033E-06	5,10909E-12	2,61E-23	1,33E-34	6,81E-46	3,48111E-57	2,05388E-85	1,2118E-113	4,2184E-170	1,4685E-226
	14	0,869358235	0,246596964	8,31529E-07	6,9144E-13	4,78E-25	3,31E-37	2,29E-49	1,58042E-61	6,28288E-92	2,4977E-122	3,9475E-183	6,2386E-244
	15	0,860707976	0,22313016	3,05902E-07	9,35762E-14	8,76E-27	8,19E-40	7,67E-53	7,1751E-66	1,92195E-98	5,1482E-131	3,6939E-196	2,6504E-261
	20	0,818730753	0,135335283	2,06115E-09	4,24835E-18	1,8E-35	7,67E-53	3,26E-70	1,3839E-87	5,1482E-131	1,9152E-174	2,6504E-261	0
	25	0,778800783	0,082084999	1,38879E-11	1,92875E-22	3,72E-44	7,18E-66	1,38E-87	2,6692E-109	1,379E-163	7,1246E-218	0	0
	30	0,740818221	0,049787068	9,35762E-14	8,75651E-27	7,67E-53	6,71E-79	5,9E-105	5,1482E-131	3,6939E-196	2,6504E-261	0	0
	35	0,70468809	0,030197383	6,30512E-16	3,97545E-31	1,58E-61	6,28E-92	2,5E-122	9,9296E-153	9,8946E-229	9,8597E-305	0	0
	40	0,670320046	0,018315639	4,24835E-18	1,80485E-35	3,26E-70	5,9E-105	1,1E-139	1,9152E-174	2,6504E-261	0	0	0
50	0,60653066	0,006737947	1,92875E-22	3,72008E-44	1,38E-87	5,1E-131	1,9E-174	7,1246E-218	0	0	0	0	
75	0,472366553	0,000553084	2,67864E-33	7,1751E-66	5,1E-131	3,7E-196	2,7E-261	0	0	0	0	0	
100	0,367879441	4,53999E-05	3,72008E-44	1,3839E-87	1,9E-174	2,7E-261	0	0	0	0	0	0	
200	0,135335283	2,06115E-09	1,3839E-87	1,9152E-174	0	0	0	0	0	0	0	0	

Tabela 4.20 – Tabela de probabilidades

A Tabela 4.21 é uma reprodução da Tabela 4.1 e também é válida para esse caso.

Probabilidade	N=1000	N=10000	N=100000	N=1000000
0,1%	1	2	2	2
0,2%	2	2	2	3
0,3%	2	2	2	3
0,4%	2	2	3	3
0,5%	2	2	3	3
0,6%	2	2	3	3
0,7%	2	2	3	3
0,8%	2	2	3	3
0,9%	2	2	3	3
1%	2	2	3	3
2%	2	3	3	4
3%	2	3	4	4
5%	3	4	4	5
7%	3	4	5	6
9%	3	4	5	6
11%	4	5	6	7
13%	4	5	6	7
15%	4	5	6	8
17%	4	6	7	8
19%	5	6	7	9
21%	5	6	8	9
23%	5	7	8	10
25%	5	7	9	10
27%	6	7	9	11
29%	6	8	10	11
31%	6	8	10	12
33%	6	8	11	13
35%	7	9	11	13
37%	7	9	12	14
39%	7	10	12	15
41%	8	10	13	15
43%	8	11	13	16
45%	8	11	14	17
47%	9	12	15	18
49%	9	12	16	19
51%	10	13	17	20
53%	10	14	17	21
55%	11	15	18	22
57%	11	15	19	24
59%	12	16	21	25
61%	13	17	22	27
63%	13	18	23	28
65%	14	19	25	30
67%	15	21	26	32
69%	16	22	28	35
71%	17	24	31	37
73%	18	26	33	40
75%	20	28	36	44
77%	21	30	39	48
79%	23	33	43	52
81%	25	36	47	58
83%	28	40	53	65
85%	31	46	60	74
87%	36	52	69	85
89%	41	61	80	100
91%	48	73	97	121
93%	59	91	123	154
95%	77	122	167	211
97%	113	188	263	339
99%	239	460	688	917

Tabela 4.21 – Quantidade de eventos sucessivos em função de p e N

Dessa maneira, depois da probabilidade p e do valor N estarem definidos, é possível encontrar a quantidade de intervalos sucessivos de tempo t que devem acontecer, para que um alarme seja gerado.

4.1.5.3 - Detecção de Falhas

O próximo passo é utilizar o algoritmo para monitorar elementos com o objetivo de detectar a existência ou não de alarmes. O primeiro teste do algoritmo foi feito sobre uma ERB com $\lambda = 0,32$ chamadas/segundo. O período de observação foi de uma hora no HMM, das 17h às 18h.

Para $t = 10$ segundos foi encontrado, através da Tabela 4.20, uma probabilidade $p = 0,04 = 4\%$. Com $p = 4\%$ e $N = 1.000.000$, através da Tabela 4.21, foi encontrado que a quantidade de intervalos sucessivos é de 5 intervalos, ou seja, $5 \cdot 10 = 50$ segundos.

Aplicando o algoritmo sobre o elemento, nenhuma falha foi detectada, pois, na verdade, a ERB não apresenta problema de bloqueio de tráfego. A Figura 4.28 mostra um histograma de todos os intervalos de tempo em que a ERB ficou sem ter tráfego. O maior intervalo é de 24 segundos. Esse valor fica abaixo dos 50 segundos necessários para a detecção do alarme.

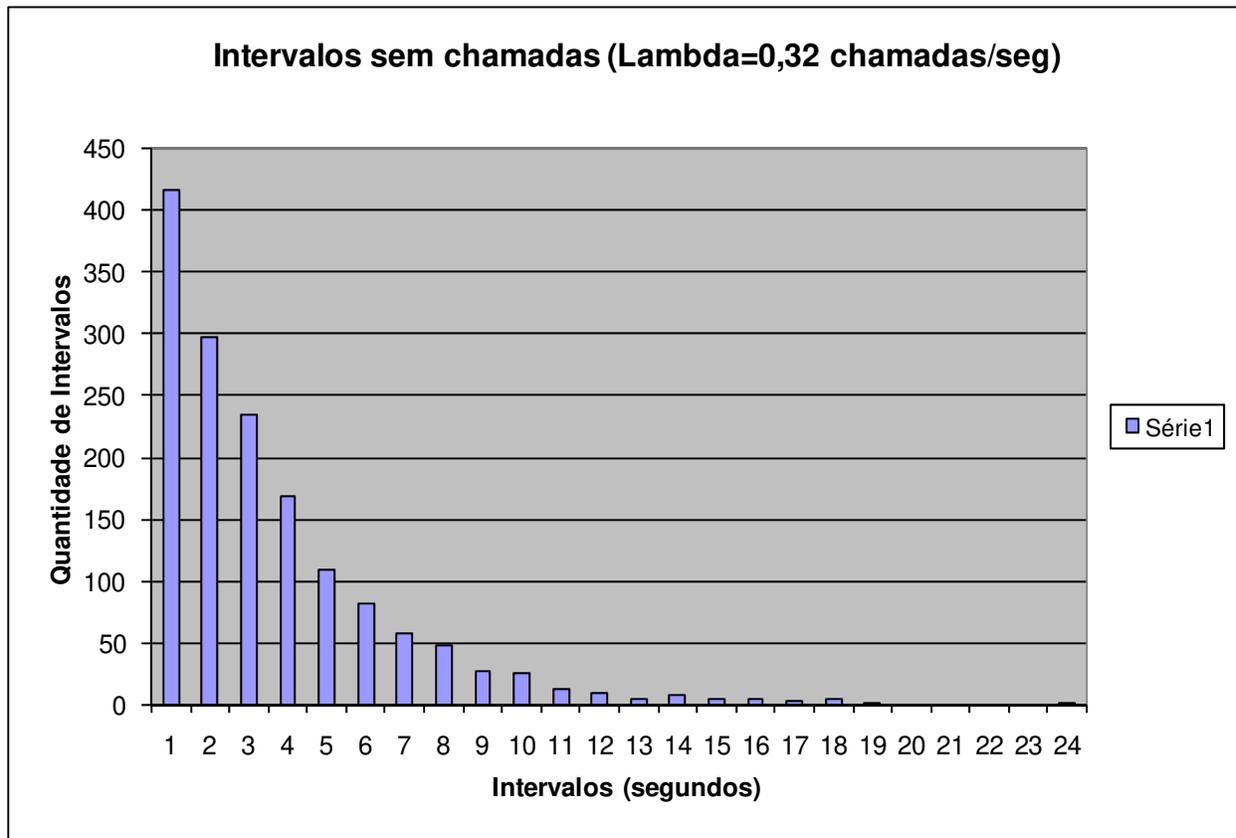


Figura 4.28 – Histograma para $\lambda = 0,32$ com os intervalos de tempo sem ocorrência de chamadas

A Figura 4.29 mostra uma comparação entre a distribuição exponencial negativa obtida pela equação $\lambda \cdot e^{-\lambda \cdot t}$, que representa o tempo entre ocorrências sucessivas, e os valores percentuais (Eixo “Taxa” do gráfico abaixo) equivalentes aos intervalos sem chamadas da ERB. Observando a figura, é possível concluir que os valores teóricos são semelhantes aos práticos.

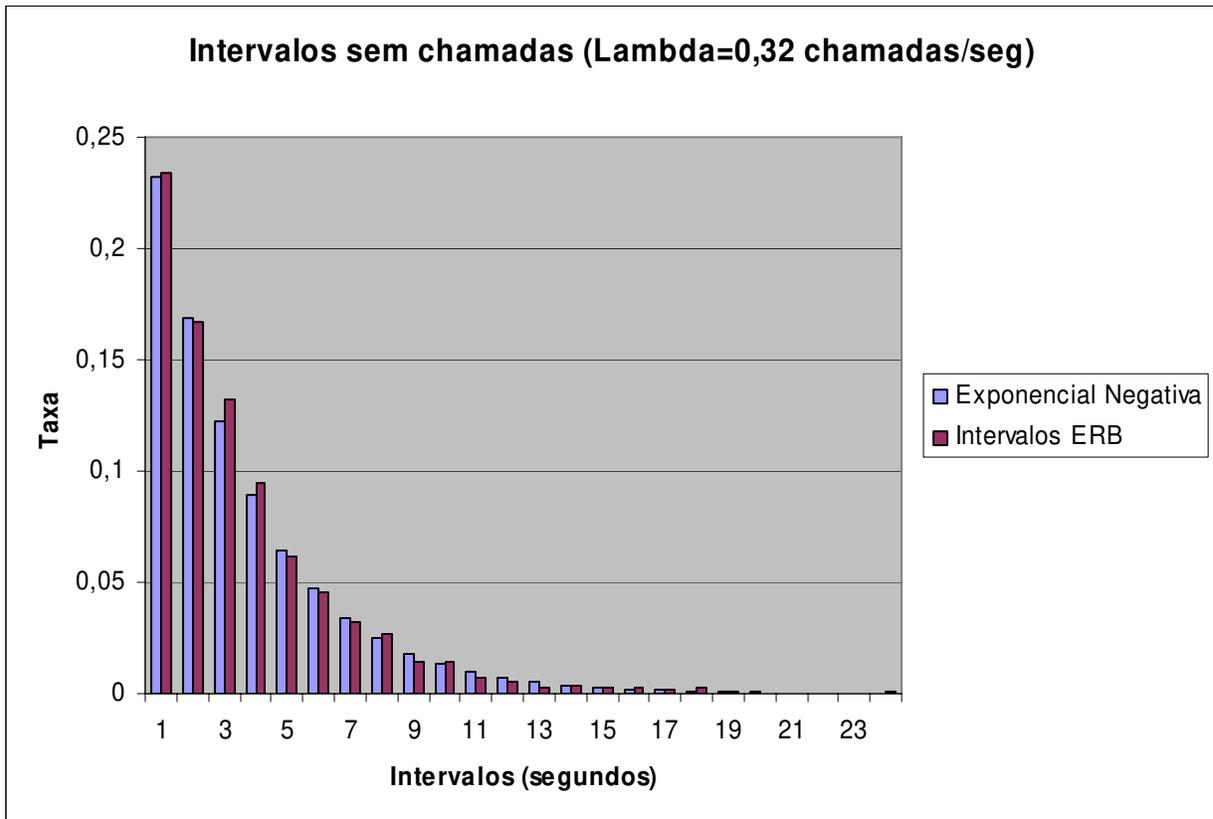


Figura 4.29 – Comparação entre a Distribuição Exponencial e as chamadas de uma ERB

Foi feito outro experimento, agora com um tronco que interconecta centrais, $\lambda = 0,68$ chamadas/segundo. Esse valor representa o dobro do λ apresentado anteriormente. Tomando $t = 5$ segundos foi encontrada uma probabilidade $p = 0,033 = 3,3\%$. Olhando na Tabela 4.21, é possível encontrar que a quantidade de intervalos consecutivos de 5 segundos com $N = 1.000.000$ é de 4 intervalos, ou seja, 20 segundos.

Entretanto, aplicando outra vez o algoritmo sobre o elemento, não foi detectada a existência de problemas de interrupção de tráfego. A Figura 4.30 mostra um histograma de todos os intervalos de tempo em que o entroncamento fica sem ter tráfego. O maior intervalo é de 13 segundos, que aconteceu duas vezes, em períodos distintos. Esse valor fica abaixo dos 20 segundos necessários para a detecção do alarme.

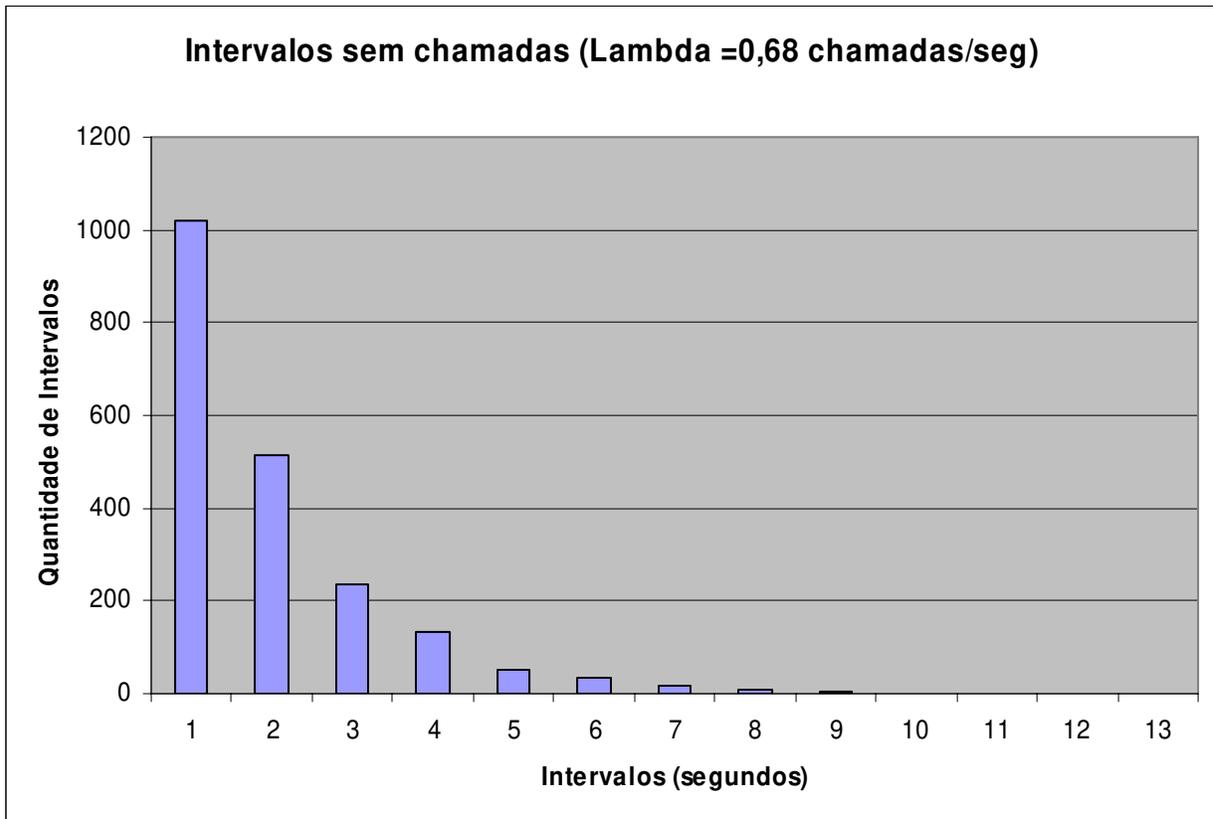


Figura 4.30 – Histograma para $\lambda = 0,68$ com os intervalos de tempo sem ocorrência de chamadas

A Figura 4.31 mostra a mesma relação apresentada anteriormente. Novamente existe a convergência entre o modelo teórico e os valores de um entroncamento real.

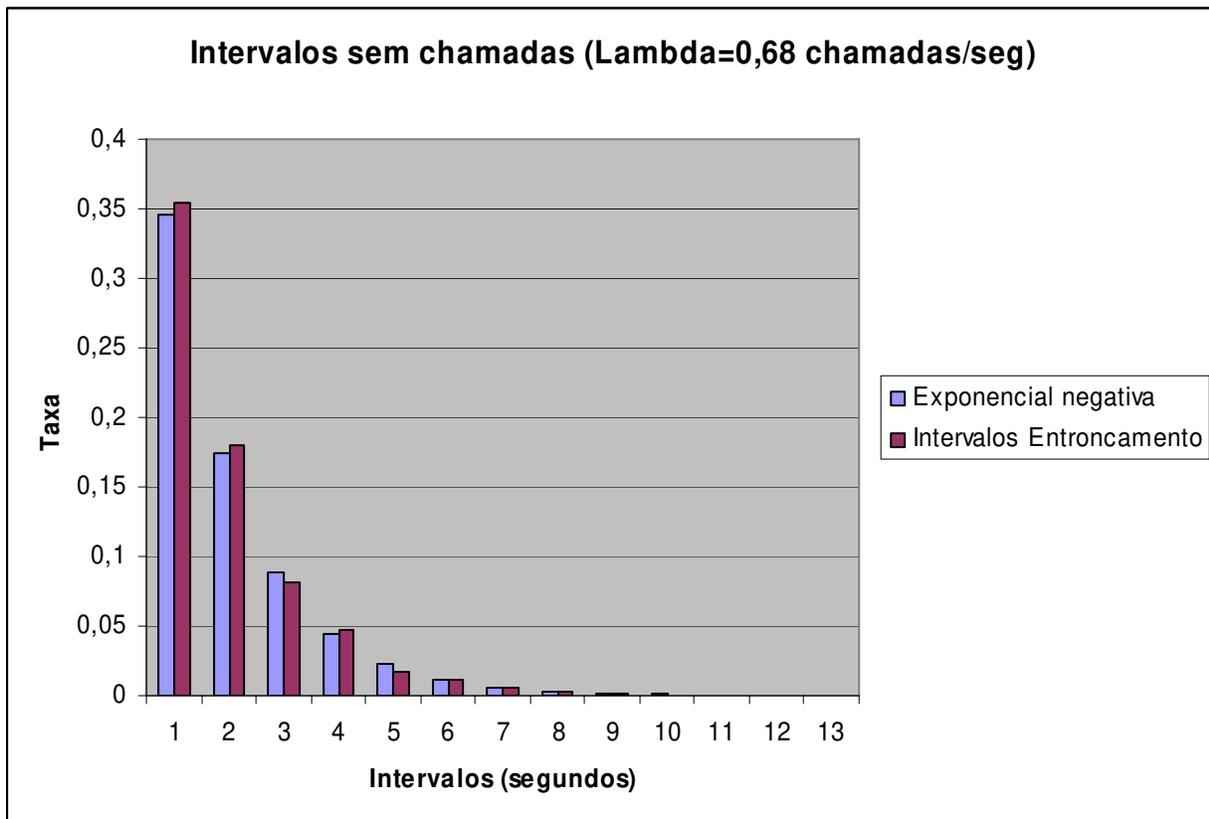


Figura 4.31 – Comparação entre a Distribuição Exponencial e as chamadas de uma ERB

A detecção de um problema veio por intermédio da análise dos circuitos relativos à plataforma de valor agregado. Essa plataforma possui um problema que bloqueia o tráfego nos juntores, circuitos. O bloqueio é aleatório, não há como criar qualquer relação de circuitos que sofram esse problema.

A plataforma é um componente externo à Central, Figura 4.32. Na origem dos sistemas celulares esta funcionalidade não existia. Com o passar dos anos, esse novo serviço foi adicionado ao sistema.

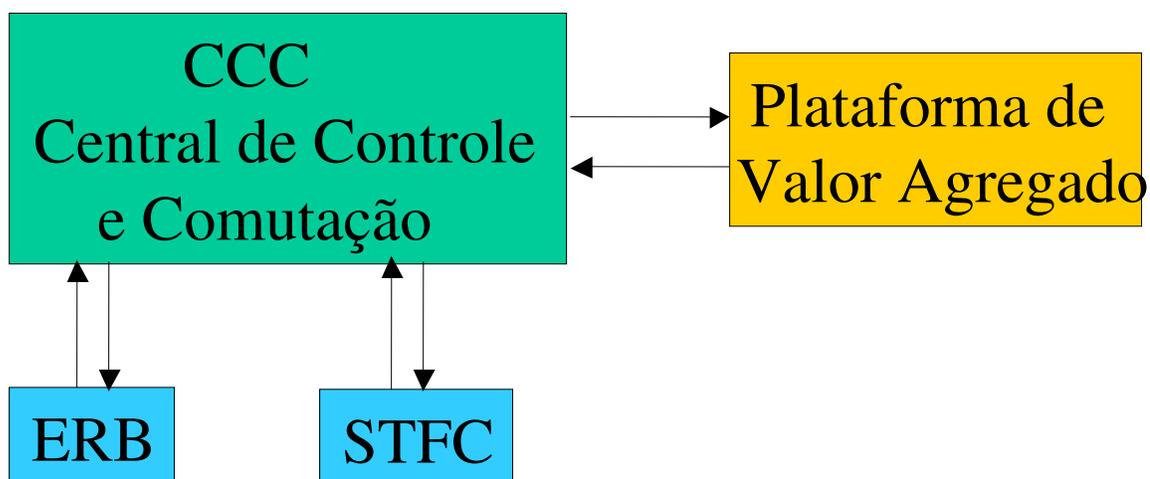


Figura 4.32 – Interligação entre a CCC e a Plataforma

Na grande maioria dos casos a plataforma e a CCC são de fabricantes diferentes. A interligação é feita por troncos baseados no padrão *PCM-30* canais, com sinalização ISUP entre eles.

As plataformas são uma tecnologia recente, que em muitos casos ainda está se consolidando tecnologicamente. As plataformas implantadas no Brasil são equipamentos importados que foram adaptados muito rapidamente aos padrões nacionais. O melhor exemplo disso é em relação ao padrão de interconexão que nos EUA é diferente do adotado no Brasil. Nos EUA a interconexão é feita por entroncamentos baseados no padrão *PCM-24* canais. Na maioria dos casos a migração foi feita em prazos curtos, sem uma bateria de testes adequada, deixando algumas falhas de software/hardware ocultas e que aparecem em situações específicas. Dessa maneira, a plataforma é um ponto potencial para falhas, em especial o entroncamento que a interliga com a CCC.

Na central que está sendo analisada, existem aproximadamente 450 juntores, que interconectam a CCC com a plataforma. Desses serão analisados aproximadamente 100 juntores, escolhidos aleatoriamente. A taxa de chamadas para os juntores é em média de $\lambda = 0,013$ chamadas/segundo.

		Lambda (chamadas/segundo)		
		0,013	0,68	1
tempo(seg)	1	0,987084135	0,506616992	0,367879441
	2	0,97433509	0,256660777	0,135335283
	3	0,961750709	0,130028711	0,049787068
	4	0,949328867	0,065874754	0,018315639
	5	0,937067463	0,03337327	0,006737947
	6	0,924964427	0,016907466	0,002478752
	7	0,913017711	0,008565609	0,000911882
	8	0,901225297	0,004339483	0,000335463
	9	0,889585193	0,002198456	0,00012341
	10	0,878095431	0,001113775	4,53999E-05
	11	0,866754069	0,000564257	1,67017E-05
	12	0,85555919	0,000285862	6,14421E-06
	13	0,844508903	0,000144823	2,26033E-06
	14	0,83360134	7,33697E-05	8,31529E-07
	15	0,822834658	3,71703E-05	3,05902E-07
	20	0,771051586	1,2405E-06	2,06115E-09
	25	0,722527354	4,13994E-08	1,38879E-11
	30	0,677056874	1,38163E-09	9,35762E-14
	35	0,634447968	4,61096E-11	6,30512E-16
	40	0,594520548	1,53883E-12	4,24835E-18
50	0,522045777	1,71391E-15	1,92875E-22	
75	0,377192354	7,09547E-23	2,67864E-33	
100	0,272531793	2,93748E-30	3,72008E-44	
200	0,074273578	8,6288E-60	1,3839E-87	

Tabela 4.22 – Probabilidade para $\lambda = 0,013$

As probabilidades relativas a $\lambda = 0,013$ estão relacionadas na primeira coluna da Tabela 4.22: para $t = 200$ segundos encontra-se uma probabilidade $p = 0,074 = 7,4\%$. Olhando na Tabela 4.21, encontra-se que a quantidade de intervalos sucessivos para $p = 7,4\%$ e para $N = 1.000.000$ é de 6 intervalos, ou seja, $6 \cdot 200 = 1200$ segundos, ou 20 minutos.

Aplicando o algoritmo sobre o elemento, foram detectados dois circuitos com problemas, os de número 82 e 86. Na Tabela 4.23 é descrita a evolução das chamadas nos dois circuitos que apresentaram o problema. Pode-se observar que para o circuito 86 não houve mais ligações após a chamada efetuada às 17:26:29, que teve seu término às 17:27:26. O término da janela de observação foi às 18:00:00. São aproximadamente 1.954 segundos sem chamadas.

Para o circuito 82 a última ligação foi efetuada às 17:35:25, com término às 17:36:10. Desse horário até o término da janela de observação são 1.430 segundos. Os dois circuitos apresentam um comportamento anormal.

Tronco	Circuito	Início Ligação	Término ligação	Segundos
1901	86	17:00:42	17:02:45	
1901	86	17:03:15	17:03:43	30
1901	86	17:04:14	17:04:20	31
1901	86	17:04:53	17:06:24	33
1901	86	17:06:57	17:07:20	33
1901	86	17:07:44	17:07:51	24
1901	86	17:08:15	17:08:43	24
1901	86	17:09:17	17:09:44	34
1901	86	17:10:19	17:10:52	35
1901	86	17:11:18	17:12:17	26
1901	86	17:12:44	17:19:04	27
1901	86	17:19:35	17:20:12	31
1901	86	17:20:46	17:21:22	34
1901	86	17:21:48	17:22:06	26
1901	86	17:22:31	17:23:45	25
1901	86	17:24:58	17:25:07	73
1901	86	17:25:30	17:25:31	23
1901	86	17:25:57	17:26:00	26
1901	86	17:26:29	17:27:26	29
		18:00:00		1954
Tronco	Circuito	Início Ligação	Término ligação	Segundos
1901	82	17:01:15	17:07:24	
1901	82	17:07:46	17:09:55	22
1901	82	17:10:28	17:10:31	33
1901	82	17:10:59	17:11:32	28
1901	82	17:11:56	17:12:19	24
1901	82	17:13:08	17:13:17	49
1901	82	17:13:40	17:14:11	23
1901	82	17:14:35	17:14:42	24
1901	82	17:15:13	17:15:20	31
1901	82	17:15:46	17:16:40	26
1901	82	17:17:03	17:17:04	23
1901	82	17:17:24	17:17:46	20
1901	82	17:18:14	17:18:32	28
1901	82	17:19:03	17:19:09	31
1901	82	17:19:39	17:19:41	30
1901	82	17:20:12	17:20:17	31
1901	82	17:20:51	17:23:26	34
1901	82	17:23:54	17:24:06	28
1901	82	17:24:36	17:26:13	30
1901	82	17:27:10	17:27:18	57
1901	82	17:27:48	17:28:41	30
1901	82	17:28:59	17:29:24	18
1901	82	17:29:49	17:29:55	25
1901	82	17:30:24	17:30:43	29
1901	82	17:31:46	17:31:47	63
1901	82	17:32:09	17:32:11	22
1901	82	17:32:39	17:32:42	28
1901	82	17:33:11	17:34:03	29
1901	82	17:34:29	17:35:00	26
1901	82	17:35:25	17:36:10	25
		18:00:00		1430

Tabela 4.23 – Circuitos que apresentam problemas

Uma conclusão que pode ser extraída, baseando-se na taxa de chamadas/segundo, $\lambda = 0,013$, é que existe um super dimensionamento do entroncamento entre a CCC e a plataforma. Dessa maneira, os circuitos que sofrem pane não chegam a interferir na taxa de bloqueio dessa canalização. Podem existir danos caso não haja manutenção por um longo período; isto causaria um acúmulo de circuitos com problemas aumentando o bloqueio no tronco.

4.2 - Conclusão

Foi constatado, no decorrer do trabalho, que os algoritmos são complementares, e podem ser implementados segundo essa filosofia de complementaridade. O algoritmo de Tempo Real destina-se a detectar problemas de grande proporção. Um problema de grande proporção é aquele que gera uma quantidade excessiva de perdas no sistema, uma degradação de grandes proporções do evento no elemento, podendo levá-lo à paralisação. Esse algoritmo possui um grande apelo, pois como o próprio nome diz, trabalha em tempo real. O tempo de detecção para esse tipo de problema é da ordem de segundos, dependendo da taxa λ de chegada de chamadas ou geração de bilhetes.

O algoritmo de Espaços Amostrais se mostrou bastante eficiente para detectar problemas de menor proporção, ou seja, é mais sensível. A desvantagem está no maior tempo para detecção.

Foi observado que quanto maior o tamanho da janela de dados, maior é a sensibilidade e menor o desvio-padrão, ou seja, menor a dispersão no comportamento. Porém quanto maior a janela, maior também é o tempo de detecção. Por outro lado, quanto menor a janela, menor a sensibilidade e maior o desvio-padrão, o que caracteriza uma maior dispersão no comportamento. A vantagem de uma janela menor é que o tempo de detecção também diminui. Não é possível dizer que esse algoritmo é baseado em tempo real, uma vez que é necessário esperar a quantidade de elementos definida, para, então, fazer a análise e concluir se existe alguma anormalidade.

O algoritmo para chamadas OK fecha uma lacuna existente nos modelos anteriores que cobriam todos os eventos com exceção do evento OK. Em relação a esse algoritmo é sugerida uma implementação complementar aos dois algoritmos anteriores.

O algoritmo para Detecção de Ocupação, como o próprio nome diz, é eficiente na detecção de elementos com bloqueio de tráfego. Uma conclusão é que o desempenho do algoritmo melhora, à medida que o tráfego, ou seja, a quantidade de chamadas por segundo é definida para cada elemento que está sendo monitorado. A adoção de um valor padrão qualquer degrada o desempenho do algoritmo, podendo ocasionar falsos alarmes positivos. Outra característica desse algoritmo é que

quando aplicado para monitorar elementos como circuitos de um tronco, canais de RF de uma ERB, ele implicitamente monitorará os elementos “maiores” como o próprio tronco (conjunto dos circuitos) e a ERB, respectivamente. O contrário não é verdade, ou seja, ao monitorar um tronco o algoritmo não estará monitorando os circuitos.

Na Tabela 4.24 é apresentada uma síntese em relação aos algoritmos com as características descritas neste capítulo.

	Utilização				Mais eficiente		Distribuição		
	Todas chamadas (eventos) exceto OK	Todos os Recursos	Chamadas (Eventos) OK	Ausência de Chamadas	Problemas de Grande Proporção	Problemas de Menor Proporção	Teoria da Renovação	Distribuição Binomial	Distribuição Poisson
Algoritmo - Tempo Real	X	X			X		X		
Algoritmo - Janelas Amostrais	X	X				X		X	
Algoritmo - Chamadas OK		X	X		X	X	X	X	
Algoritmo - Detecção da Ocupação		X		X	X				X

Tabela 4.24 – Síntese dos Algoritmos

Obs: Problemas ou falhas de grande proporção são aquelas que paralisam boa parte dos elementos de um determinado recurso, por exemplo, paralisação de 80% dos juntores de uma rota. Já os problemas de menor proporção, como o próprio nome diz, são aqueles em que apenas uma pequena parte dos elementos de um elemento são paralisados, por exemplo, 10% dos juntores de uma rota.

Capítulo 5

Algoritmos de Redes Neurais para Detecção de Falhas

O objetivo neste capítulo é utilizar Redes Neurais Artificiais (RNAs) [59] na detecção de falhas. As redes neurais são treinadas sobre padrões extraídos do sistema de telefonia. Esses padrões estão baseados nos bilhetes de tarifação, como veremos no decorrer deste capítulo. O objetivo não é comparar o algoritmo de Redes Neurais com os apresentados no Capítulo 4, e sim estudar uma nova abordagem ao problema.

Dentre as principais utilidades das RNAs está a sua utilização para reconhecimento de padrões, que no passado eram realizadas apenas por seres humanos, pois dependiam de uma inspeção visual, da audição ou de outras habilidades sensoriais que as máquinas não conseguiam identificar.

Sistemas biológicos realizam o reconhecimento de padrões via interconexão de células físicas chamadas neurônios. Essa propriedade forneceu motivação para tentar simular tal sistema computacionalmente. Surgiu com grande força um campo de estudos centrado ao redor da criação e análise de sistemas inteligentes pela formulação de sistemas computacionais que simulem a estrutura do funcionamento do cérebro humano. É bastante nova a idéia de que é possível simular o comportamento humano por meio de interações de um grande número de simples unidades de processamento. A adaptabilidade, tolerância ao erro, grande capacidade de memória e capacidade de processamento de informações em tempo real sugerem uma arquitetura alternativa competitiva. O

fato de as células nervosas serem extremamente lentas (faixa de milissegundos) e o processamento no cérebro ser relativamente rápido nos leva a crer que a base computacional do cérebro possui poucos passos em série, sendo maciçamente paralela.

Algoritmos de Aprendizado das Redes Neurais Artificiais

Algumas das RNAs de maior destaque junto com os nomes de seus inventores são: *PERCEPTRON* (Rosenblatt / 1957); *MADALINE* (Widrow / 1960-1962); *AVALANCHE* (Grossberg / 1967); *CEREBELLATION* (Marr, Albus & Pellionez / 1969); *BACK-PROPAGATION* (BPN) (Werbos, Parker, Rumelhart / 1974, 1985), mais comumente referenciada como *MULTI-LAYER PERCEPTRON* (MLP); *BRAIN STATE IN A BOX* (Anderson / 1977); *NEOCOGNITRON* (Fukushima / 1978-1984); *ADAPTIVE RESONANCE THEORY* (ART) (Carpenter, Grossberg / 1976, 1986); *SELF-ORGANISING MAP* (Kohonen / 1982); *HOPFIELD* (Hopfield / 1982); *BI-DIRECTIONAL ASSOCIATIVE MEMORY* (Kosko / 1985); *BOLTZMANN/CAUCHY MACHINE* (Hinton, Sejnowsky, Szu / 1985, 1986); *COUNTERPROPAGATION* (Hecht-Nielsen / 1986); *RADIAL BASIS FUNCTION* (Broomhead, Lowe / 1988); *PROBABILISTIC* (PNN) (Specht / 1988); *GENERAL REGRESSION NEURAL NETWORK* (GRNN) (Specht / 1991).

Existem muito mais RNAs do que essas citadas. A RNA de uso mais comum na atualidade é a *back-propagation* ou rede neural *MLP*. É uma rede simples de se compreender e geralmente se obtêm bons resultados com a sua aplicação. Entretanto, essa modalidade de RNA tem suas desvantagens, como por exemplo:

1. Só pode ser usada com treinamento supervisionado;
2. Necessita de um grande número de exemplos, para que se possa realizar um treinamento satisfatório;
3. O treinamento pode ser lento.

5.1 - Redes Neurais e Reconhecimento de Padrões

As RNAs têm sido largamente utilizadas na tarefa de reconhecer padrões presentes em conjuntos de dados. O objetivo desse processo é apresentar à rede neural um conjunto de dados conhecidos, de modo que, através de um processo de aprendizado, ela determine subconjuntos de dados, com características semelhantes (padrões), e seja capaz de reconhecer tais características, em

um novo elemento qualquer que a ela seja apresentado, de acordo com um dos padrões previamente fixados.

Uma referência clássica sobre o assunto é tratada no trabalho de Bishop [60] em 1995. Diversas técnicas e aplicações podem ser encontradas e analisadas. Para complementar a leitura básica, pode-se citar Ripley & Hjort [61] em 1995.

A seguir descrevemos algumas das aplicações possíveis de reconhecimento de padrões.

5.1.1 - Reconhecimento de Voz

Utilizada principalmente no processamento de comandos de voz. Outras aplicações compreendem a transcrição da fala contínua, ou seja, transformar a fala humana em texto digital [62]-[69].

5.1.2 - Reconhecimento de Caracteres

É o reconhecimento automático de caracteres e textos. É utilizado em sistemas bancários para leitura de cheques e em ferramentas chamadas de *OCRs (Optical Character Recognizers)* [70]-[73].

5.1.3 - Reconhecimento de Face

Trata-se do reconhecimento automático de faces humanas. O sistema extrai características da face humana (tamanho e forma da boca, olhos, nariz, rosto, sobrancelha) e utiliza essas informações para fazer a identificação do indivíduo. Costuma ser utilizado em sistemas de segurança para identificar pessoas [74]-[75].

5.1.4 - Reconhecimento de Impressões Digitais

Realiza o reconhecimento automático de impressões digitais. Parte do princípio que cada ser humano possui uma impressão digital diferente de qualquer outro indivíduo. Utilizado principalmente em sistemas para identificação pessoal [76]-[77].

5.1.5 - Detecção de Células Cancerosas

As RNAs são treinadas na detecção de células cancerosas de tecidos humanos. O treinamento é feito sobre padrões conhecidos [78]-[79].

5.1.6 - Geração de Música

A proposta é a utilização de algoritmos bio-inspirados que possuem características intrínsecas dos sistemas auto-organizados como definidores do processo de geração e estruturação dos diferentes elementos sonoros [80]-[82].

5.2 - *Back-Propagation*

O algoritmo de *Back-Propagation* [83]-[86] é um dos modelos que mais foi utilizado no reconhecimento de padrões. Esse algoritmo foi inventado, de forma independente, por Bryson e Ho(1969), Werbos(1974), Parker(1985) e Rumelhart, Hinton, e Willinas(1986). Uma versão muito próxima a esse algoritmo foi sugerida por Le Cun(1985). A base do funcionamento de uma rede *Back-Propagation* é um modelo de otimização baseada no gradiente descendente. Considere primeiro uma rede com duas camadas, tal como a ilustrada na Figura 5.1. A convenção adotada está ilustrada nessa mesma figura, onde as unidades de saída estão denotadas por O_i , as unidades escondidas por V_j , e os terminais de entrada por Y_k . Há conexões w_{jk} da entrada à camada intermediária e W_{ij} da camada intermediária até a saída. Observe que o índice i sempre se refere a uma unidade de saída. O índice j a uma unidade intermediária, e k a um terminal de entrada.

As entradas estão sempre fixadas a valores previamente definidos. Denotaremos os diferentes elementos por um subscrito μ . Assim, a entrada k é o conjunto y_k^μ . Os valores de entrada podem ser binários (0/1, ou ± 1) ou contínuos. Nós utilizamos N para denotar o número de unidades de entrada e p para o número de características de entrada ($\mu = 1, 2, \dots, p$).

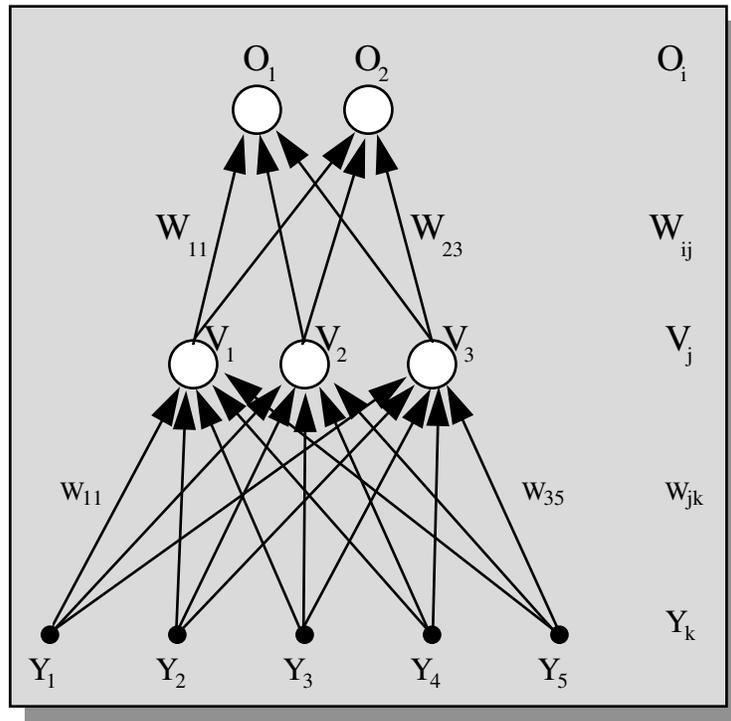


Figura 5.1 – Rede de duas camadas com alimentação direta

Dado um elemento μ , a unidade intermediária j recebe da rede de entrada

$$h_j^\mu = \sum_k w_{jk} y_k^\mu, \quad (5.1)$$

e produz a saída

$$V_j^\mu = g(h_j^\mu) = g\left(\sum_k w_{jk} y_k^\mu\right). \quad (5.2)$$

Onde $g(\)$ é a função de transferência. A unidade de saída i recebe, assim,

$$h_i^\mu = \sum_j W_{ij} V_j^\mu = \sum_j W_{ij} g\left(\sum_k w_{jk} y_k^\mu\right), \quad (5.3)$$

e produz para saída final

$$O_i^\mu = g(h_i^\mu) = g\left(\sum_j W_{ij} V_j^\mu\right) = g\left(\sum_j W_{ij} g\left(\sum_k w_{jk} y_k^\mu\right)\right). \quad (5.4)$$

A função erro quadrático médio (EQM) medida é a seguinte

$$E[W] = \frac{1}{2} \sum_{\mu} [\zeta_i^\mu - O_i^\mu]^2, \quad (5.5)$$

onde ζ_i^μ é o padrão esperado.

$$E[W] = \frac{1}{2} \sum_{\mu} \left[\zeta_i^\mu - g \left(\sum_j W_{ij} g \left(\sum_k w_{jk} y_k^\mu \right) \right) \right]^2. \quad (5.6)$$

Essa é uma função contínua diferenciável em relação a W_{ij} , w_{jk} . Logo, é possível utilizar um algoritmo de gradiente descendente para calcular e atualizar os pesos. Isso é a essência do algoritmo de *Back-propagation*, mas há grande importância na forma com que as atualizações vão ser feitas.

Para as conexões da camada intermediária à camada de saída, as regras para o gradiente descendente fornecem

$$\begin{aligned} \Delta W_{ij} &= -\eta \frac{\partial E}{\partial W_{ij}} = \eta \sum_{\mu} [\zeta_i^\mu - O_i^\mu] g'(h_i^\mu) V_j^\mu, \\ &= \eta \sum_{\mu} \delta_i^\mu V_j^\mu, \end{aligned} \quad (5.7)$$

se os erros (ou deltas) δ_i^μ forem definidos por

$$\delta_i^\mu = g'(h_i^\mu) [\zeta_i^\mu - O_i^\mu]. \quad (5.8)$$

Para as conexões da camada de entrada e camada intermediária Δw_{jk} deve-se derivar, pois eles estão mais “profundos” na rede

$$\begin{aligned} \Delta w_{jk} &= -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \sum_{\mu} \frac{\partial E}{\partial \mathcal{N}_j^\mu} \frac{\partial \mathcal{N}_j^\mu}{\partial w_{jk}} \\ &= \eta \sum_{\mu} [\zeta_i^\mu - O_i^\mu] g'(h_i^\mu) W_{ij} g'(h_j^\mu) y_k^\mu \\ &= \eta \sum_{\mu} \delta_i^\mu W_{ij} g'(h_j^\mu) y_k^\mu, \\ &= \eta \sum_{\mu} \delta_j^\mu y_k^\mu \end{aligned} \quad (5.9)$$

com

$$\delta_j^\mu = g'(h_j^\mu) \sum_i W_{ij} \delta_i^\mu . \quad (5.10)$$

A constante η é chamada de taxa de aprendizado. Observe que a equação 5.9 tem a mesma forma da equação 5.7, mas com a definição diferente de δ 's. Em geral, com um número arbitrário de camadas, a regra de atualização sempre tem a forma

$$\Delta w_{pq} = \eta \sum_{\text{elementos}} \delta_{saída} * V_{entrada} , \quad (5.11)$$

em que a saída e entrada referem-se aos dois finais p e q da conexão feita. A Figura 5.2 mostra o algoritmo *Back-Propagation* sendo utilizado para uma rede com três camadas. As linhas sólidas mostram a propagação para frente, enquanto as linhas pontilhadas mostram a propagação do erro para trás.

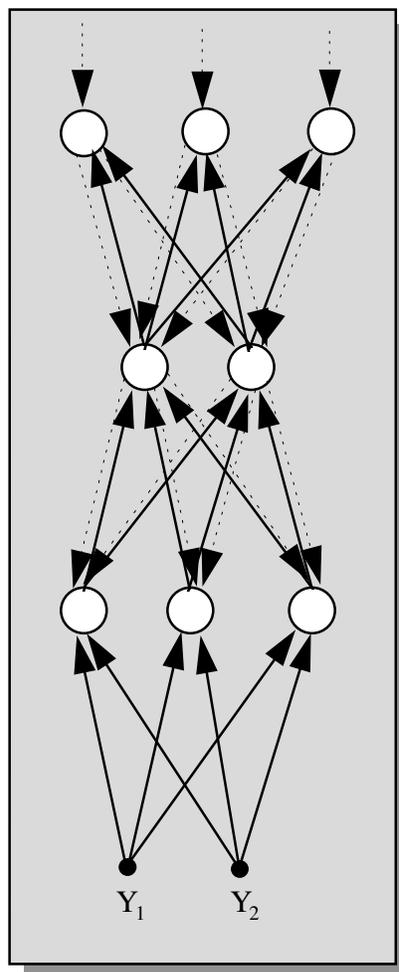


Figura 5.2 – Back-Propagation

Apesar das regras para a atualização, equações 5.7 e 5.9, estarem escritas como somas sobre todos os elementos μ , será utilizada uma filosofia um pouco diferente. Um elemento μ é apresentado à entrada da rede, e, então, todos os pesos são atualizados, antes do próximo elemento ser considerado. Isso diminui o tempo gasto em cada passo. É possível implementar esse processo sugerindo que a ordem de apresentação dos elementos à entrada da rede é aleatória. Se os pesos forem atualizados somente após todos os elementos passarem pela rede, será necessário ter um espaço adicional para armazenarmos variáveis. A eficiência relativa dos dois métodos depende do problema, mas o primeiro parece superior na maioria dos casos, especialmente para conjuntos com elementos regulares ou redundantes.

É possível utilizar a função sigmóide, equação 5.12, ou também a função tangente hiperbólica, equação 5.13, como função de transferência. A função, é claro, deve ser diferenciável, e normalmente é desejável que ela sature nos dois extremos. Para as funções abaixo existem os seguintes extremos: 0/1 ou ± 1 , respectivamente.

$$g(h) = f_{\beta}(h) = \frac{1}{1 + \exp(-2\beta h)} \quad (5.12)$$

$$g(h) = \tanh \beta h \quad (5.13)$$

O parâmetro β é geralmente estabelecido em 1 ou 1/2. As derivadas dessas funções são expressas da seguinte forma: $g'(h) = 2\beta g(1 - g(h))$, para a equação 5.12, e $g'(h) = \beta(1 - g^2)$, para a equação 5.13.

5.2.1 - Variações sobre *Back-Propagation*

O algoritmo *Back-Propagation* tem sido muito estudado, e muitas extensões e modificações foram produzidas. O método acima é relativamente lento para convergir com múltiplas camadas, e muitas variações têm proposto torná-lo mais rápido, incluindo cuidados com pontos mínimos locais e com a eficiência geral do método.

5.2.1.1 - Momento

O gradiente descendente pode ser muito lento, se η é pequeno, e pode oscilar largamente, se η é muito grande. Há um grande número de maneiras de tratar esse problema, incluindo até a substituição do gradiente descendente por algoritmos mais sofisticados de minimização. Porém uma

maneira simples é a adição de um termo chamado momento [86], que é geralmente eficiente e muito fácil de ser utilizado. A idéia é dar alguma inércia ou momento para cada conexão. Conforme o algoritmo converge para um ponto de mínimo, essa inércia se ajusta, de forma que não haja oscilação para se encontrar esse ponto. A taxa de aprendizado efetiva pode ser aumentada, sem provocar oscilações, quando o algoritmo estiver longe de pontos de mínimo. Esse esquema é apresentado a seguir

$$\Delta w_{pq}(t+1) = -\eta \frac{\partial E}{\partial w_{pq}} + \alpha \Delta w_{pq}(t). \quad (5.14)$$

O valor para o parâmetro do momento, α , precisa estar entre 0 e 1. É aconselhável utilizar o valor 0.9.

Se o algoritmo estiver trabalhando numa região plana do espaço da função, então a derivada $\partial E / \partial w_{pq}$ será praticamente sempre a mesma a cada passo, e a equação 5.14 convergirá para

$$\Delta w_{pq} \approx -\frac{\eta}{1-\alpha} * \frac{\partial E}{\partial w_{pq}}, \quad (5.15)$$

com taxa efetiva de aprendizado de $\eta/(1-\alpha)$. Se, por exemplo, chegarmos a uma região de oscilação, w_{pq} responde somente com o coeficiente η . Podemos concluir que produzimos um aumento na taxa comum, quando possível, mas, quando chegamos a uma região onde exista uma tendência a um mínimo, essa aceleração diminui gradativamente até zero.

5.2.1.2 - Parâmetros Adaptativos

Não é fácil escolher valores adequados para os parâmetros η e α para um problema particular. Além disso, os melhores valores no começo do treinamento podem não ser tão bons no final. Assim, muitos autores têm sugerido que esses parâmetros sejam ajustados automaticamente com um processo de aprendizado (exemplo, Cater, 1987; Franzini, 1987; Vogl et al., 1988; Jacobs, 1988) [86].

O procedimento usual é checar se a atualização dos pesos realizada reduziu a função erro. Se isso não aconteceu, então o processo não está convergindo, e η precisa ser diminuído. Se por outro lado, em alguns passos, o erro diminuiu, então talvez estejamos sendo muito moderados e poderíamos tentar aumentar η . Parece melhor incrementar η por uma constante, mas decrementá-lo

geometricamente para permitir um decaimento rápido, quando necessário. Teremos, então, os seguintes casos

$$\Delta\eta = \begin{cases} a_{se_}\Delta E < 0 \\ -b\eta_{se_}\Delta E > 0 \\ 0_{caso_contrário} \end{cases}, \quad (5.16)$$

em que ΔE é a média da função erro, e a e b são constantes apropriadas.

5.3 - Treinamento baseado no Algoritmo Back-Propagation

A primeira tarefa a ser executada no treinamento de uma rede neural é saber quais características extrair e como montar o vetor de entradas. Essa extração deve necessariamente caracterizar padrões, de forma que esses sejam captados pela rede neural artificial. Mesmo assim, não há garantias de que a rede neural irá conseguir assimilar esses padrões, de forma a reconhecê-los posteriormente.

5.3.1 - Padrões

Grande parte das informações que nos rodeiam manifestam-se em forma de padrões. A facilidade com que humanos classificam e descrevem padrões, geralmente, nos conduz a assumir incorretamente que essa capacidade possa ser facilmente automatizada. Um padrão pode ser tão básico quanto um conjunto de medidas ou observações (pressão sanguínea, idade, peso, altura), podendo ser representado numa notação de vetor ou matriz.

5.3.2 - Características

Características são quaisquer medidas extraídas de um padrão. Um exemplo de características de nível mais baixo são as intensidades dos sinais. Características podem ser simbólicas ou numéricas, como cores e dimensões, respectivamente. Características podem também ser resultado da aplicação de um algoritmo de extração ou operador. Adicionalmente, características podem ser de nível mais alto, por exemplo, descrições geométricas de uma região ou de um objeto em 3-D, aparecendo numa imagem. Observe que: (1) esforços significantes podem ser requeridos na extração de características; (2) as características extraídas podem conter erros ou ruídos.

Os problemas relatados de seleção e extração de características precisam ser resolvidos, conforme as necessidades do sistema de reconhecimento de padrões. A chave é escolher e extrair características que:

- Sejam computacionalmente flexíveis;
- Representem o conjunto de padrões, de forma que possam ser injetadas num sistema de reconhecimento eficaz;
- Sejam reduzidas a uma quantidade de informação administrável, sem descartar informações vitais.

5.3.3 - Seleção de Características

A maneira encontrada para identificar padrões que representam problemas, num sistema de telecomunicações, foi extrair as características relativas aos eventos:

- A taxa percentual do evento que está sendo monitorado;
- A taxa percentual do evento OK;
- A taxa percentual do restante dos eventos;
- O valor do *AQL*;

A saída foi assim modelada:

- O valor representando o estado do alarme (-1 para representar que o vetor caracteriza um estado alarmado, e 1 para representar que o vetor caracteriza um estado sem alarme). Nesse caso o limiar para a decisão é o valor 0 (Zero).

O estado do alarme é o vetor de saída. Dessa maneira, a rede neural possui quatro entradas, uma camada intermediária de neurônios e uma saída, ou seja uma rede com três camadas. Com esses valores é possível montar uma seqüência de vetores que pode ser útil no treinamento da rede neural.

A camada de entrada da rede possui uma função de transferência chamada Sigmóide, que limita os valores da sua saída entre 0 e 1. A camada intermediária que conduz a saída da rede neural possui a função de transferência do tipo Tangente hiperbólica, tendo em vista que os valores de saída variam entre 1 e -1.

A adoção dessas funções de transferência aumentou a capacidade de convergência da rede neural. Foram feitos testes alterando as funções de transferência, e os melhores resultados foram obtidos com as funções definidas acima.

5.3.3.1 - Back-Propagation

Nos primeiros experimentos foi utilizada a rede neural do tipo back-propagation, em sua versão mais básica, que significa trabalhar com gradiente descendente, sem momento e sem parâmetros adaptativos. Posteriormente, serão abordadas redes com um grau crescente de complexidade, isto é, Taxa de Aprendizado Variável (TAV) e Momento (α). Os parâmetros da rede neural ficaram assim configurados:

- Neurônios na entrada = 4;
- Quantidade de saídas = 1;
- Neurônios na camada Intermediária = 200;
- Função de transferências no neurônio de entrada = Sigmóide;
- Função de transferência nos neurônios de saída = Tangente hiperbólica;
- Quantidade de interações = 500;
- Erro quadrático médio (EQM) desejado = 0,005;
- Taxa de aprendizado (η) = 0,25;
- Vetores de entrada = 1.000 (vetores representando um *AQL*).

A Figura 5.3 mostra o comportamento da rede em relação à convergência.

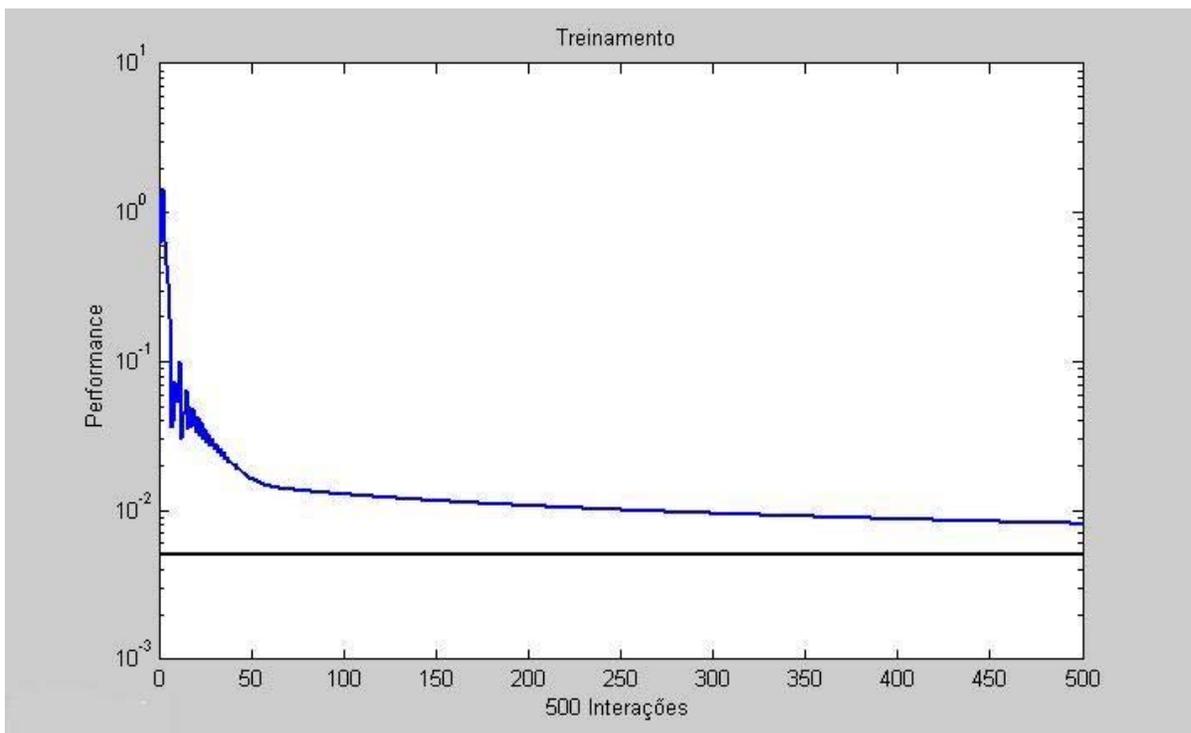


Figura 5.3 – Comportamento da convergência da Rede Neural com 200 neurônios

O valor estipulado, como meta para o EQM, foi de 0.005, porém, após 500 interações, o valor encontrado foi de 0,00867. Sempre será dada ênfase ao parâmetro relativo ao EQM, pois é ele que reflete o estado da rede neural e é através dele que é possível concluir se houve ou não convergência. Um EQM elevado significa uma não-convergência da rede neural, implicando que a rede não conseguiu aprender e, conseqüentemente, está inapta ao reconhecimento de padrões. Apesar do algoritmo não ter encontrado o EQM desejado, houve convergência da rede neural. A convergência pode ser averiguada, após testar a rede e verificar que para o EQM encontrado a rede acertou aproximadamente 99,8% dos testes executados. A grande quantidade de vetores para o treinamento dificulta a obtenção do EQM estipulado.

Alterando a quantidade de Neurônios

A seguir, fizemos algumas tentativas no sentido de reduzir o erro. A primeira delas foi aumentar o número de neurônios na camada intermediária. Na Figura 5.4 vemos o resultado do aumento do número de neurônios para 300.

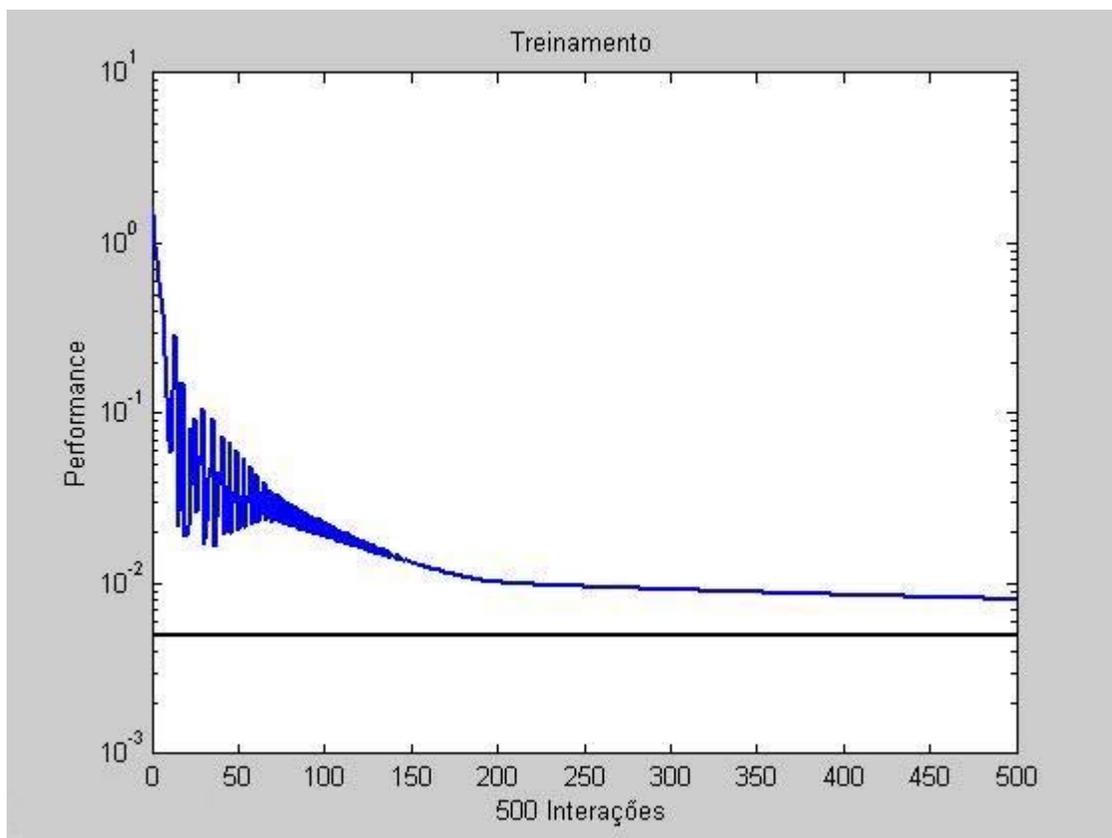


Figura 5.4 – Comportamento da convergência da Rede Neural com 300 neurônios

Basicamente, o erro encontrado foi o mesmo que o da rede anterior, 0,00808. Sempre que trabalhamos com redes neurais, devemos fazer os experimentos mais que uma vez, pois a convergência pode diferir, mesmo sem alterações nos parâmetros. Assim sendo, realizamos cinco outros experimentos. Em dois deles o valor do EQM ficou em 0,00815 e 0,00811, e nos três restantes não houve convergência. A composição dos resultados leva a concluir que 300 neurônios para a camada intermediária está acima do necessário, o que pode conduzir à rede a não-convergência.

A Figura 5.5 mostra o comportamento para uma rede com 100 neurônios na camada intermediária.

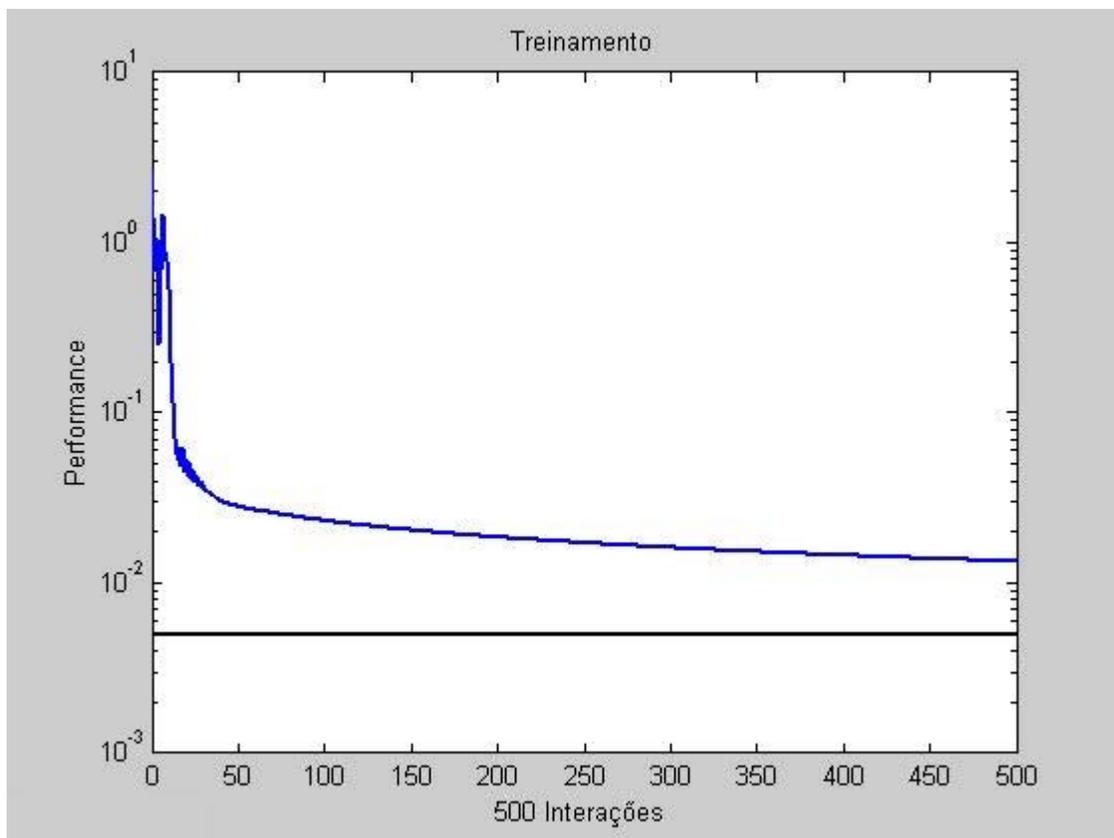


Figura 5.5 – Comportamento da convergência da Rede Neural com 100 neurônios

Com 100 neurônios na camada intermediária o resultado em relação ao EQM foi de:

- Primeiro experimento: 0,01335;
- Segundo experimento: 0,01512;
- Terceiro experimento: 0,01883.

Mesmo com os valores mais elevados para o EQM, a taxa de acerto ficou na casa dos 98,2%. Entretanto, observando a ativação dos neurônios, podemos verificar que em muitos casos o limiar entre a ativação correta e a incorreta foi muito pequeno, na casa dos centésimos. Dessa maneira, é possível concluir que 100 neurônios para a camada intermediária está abaixo do necessário. Esses experimentos mostraram que 200 neurônios é uma boa escolha para a camada intermediária.

Alterando a taxa de aprendizado (η)

Outro teste que pode ser feito é alterar a taxa de aprendizado da rede, com o objetivo de reduzir o tempo de convergência e também o EQM. Foram feitos alguns experimentos, cujos resultados estão na Figura 5.6, para uma taxa de aprendizado (η) de 0,35.

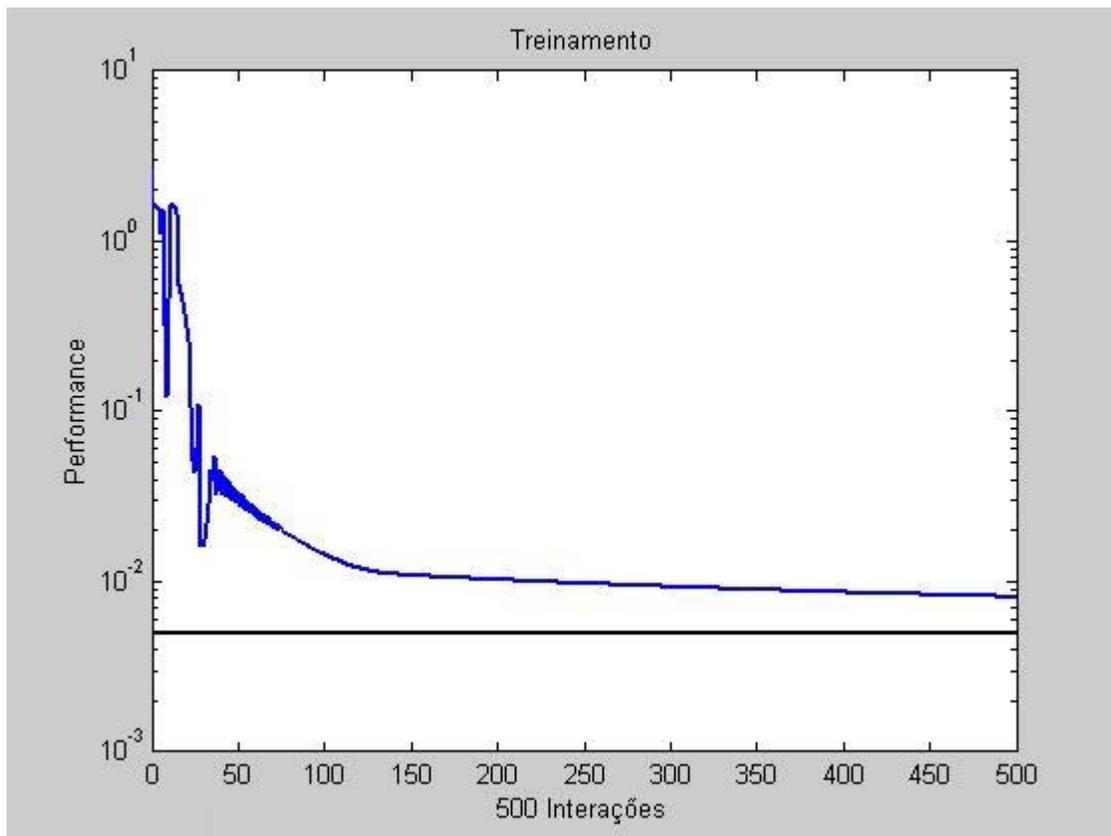


Figura 5.6 – Comportamento da convergência da Rede Neural com 200 neurônios e $\eta=0,35$

Com $\eta=0,35$ o resultado em relação ao erro foi de:

- Primeiro experimento: 0,00811;
- Segundo experimento: 0,00852;
- Terceiro experimento: 0,00902.

Com $\eta=0,15$ o resultado da rede neural não foi satisfatório, pois das cinco tentativas três não convergiram. A Figura 5.7 mostra um exemplo de uma não-convergência. Os valores para os experimentos foram os seguintes:

- Primeiro experimento: (1,6111) não convergiu;
- Segundo experimento: 0,01071;
- Terceiro experimento: 0,01104;
- Quarto experimento: (1,6111) não convergiu;
- Quinto experimento: (1,6111) não convergiu.

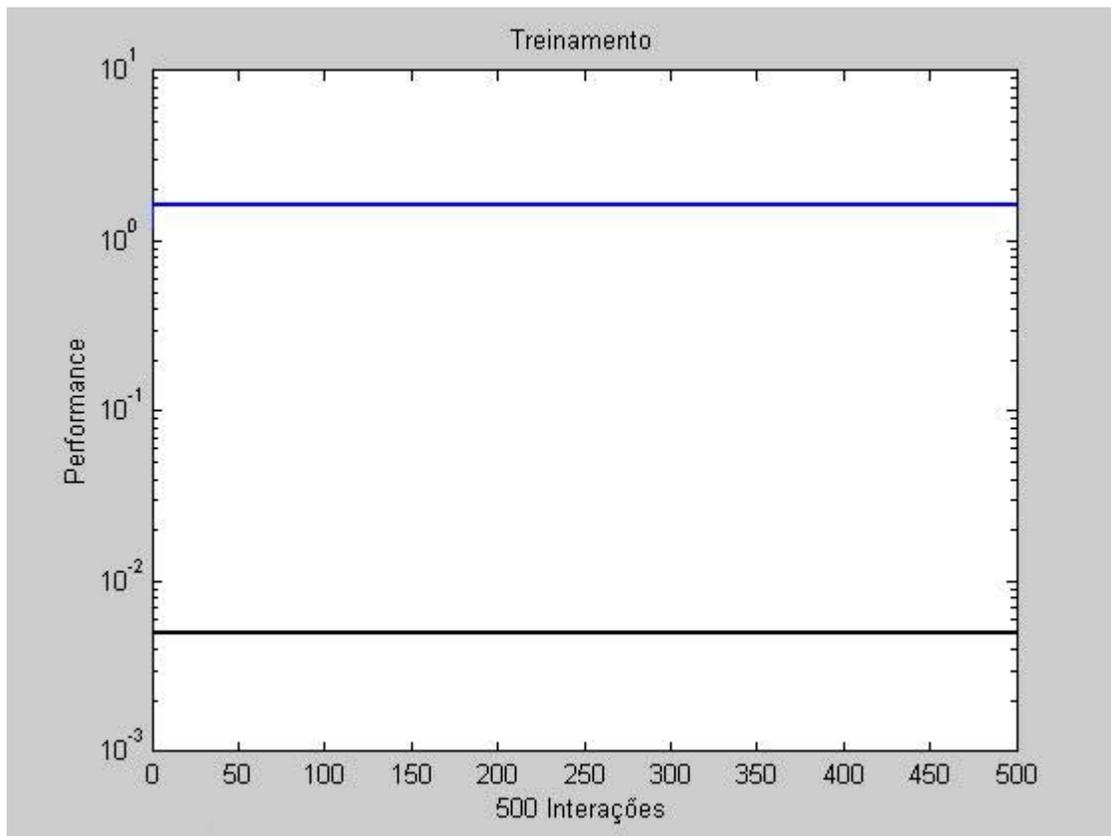


Figura 5.7 – Exemplo da não-convergência de uma Rede Neural com 200 neurônios e $\eta=0,15$

Os resultados para taxas de aprendizado de 0,35 e 0,15 foram piores que os obtidos com $\eta=0,25$. Sendo assim, o valor de $\eta=0,25$ parece ser uma boa escolha.

5.3.3.2 - Back-Propagation com Taxa de Aprendizagem variável

O algoritmo anterior mantinha a taxa de aprendizagem constante durante todo o processo. Essa é a forma mais simples possível que o parâmetro da taxa de aprendizagem pode assumir. Esse tipo de algoritmo é muito sensível à escolha apropriada da taxa de aprendizagem. Se a taxa de aprendizagem for estabelecida num valor muito elevado, o algoritmo pode oscilar e se tornar estável. Por outro lado, se a taxa for um valor muito baixo, o algoritmo pode demorar muito para convergir. Nos próximos testes será adotada uma taxa de aprendizagem variável, durante o treinamento da rede, e será verificado se isso beneficia a convergência do algoritmo. Uma taxa de aprendizagem adaptativa tenta manter o aprendizado sempre no máximo possível. Esse novo parâmetro será designado como TAV ou Taxa de Aprendizagem Variável.

Foram feitos diversos experimentos com o objetivo de testar esse novo parâmetro. Os resultados estão descritos a seguir. É possível visualizar o comportamento para diferentes TAVs, observando a Figura 5.8 até a Figura 5.12. Adotando-se valores para o TAV entre 1,05 e 1,35, é possível chegar a um desempenho similar ao apresentado no algoritmo anterior.

- **EQM=0,008417**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,05**
- **EQM=0,008644**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,05**
- **EQM=0,008965**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,05**
- **EQM=0,007690**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,15**
- **EQM=0,008781**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,15**
- **EQM=0,010398**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,15**
- **EQM=0,010367**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,35**
- **EQM=0,008351**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,35**
- **EQM=0,014920**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,35**
- **EQM=0,011025**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,50**
- **EQM=0,010367**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,50**
- **EQM=0,011664**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=1,50**
- **EQM=0,009917**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=2,00**
- **EQM=0,011296**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=2,00**
- **EQM=0,011008**, 500 Interações, 200 Neurônios, $\eta=0,25$ e **TAV=2,00**

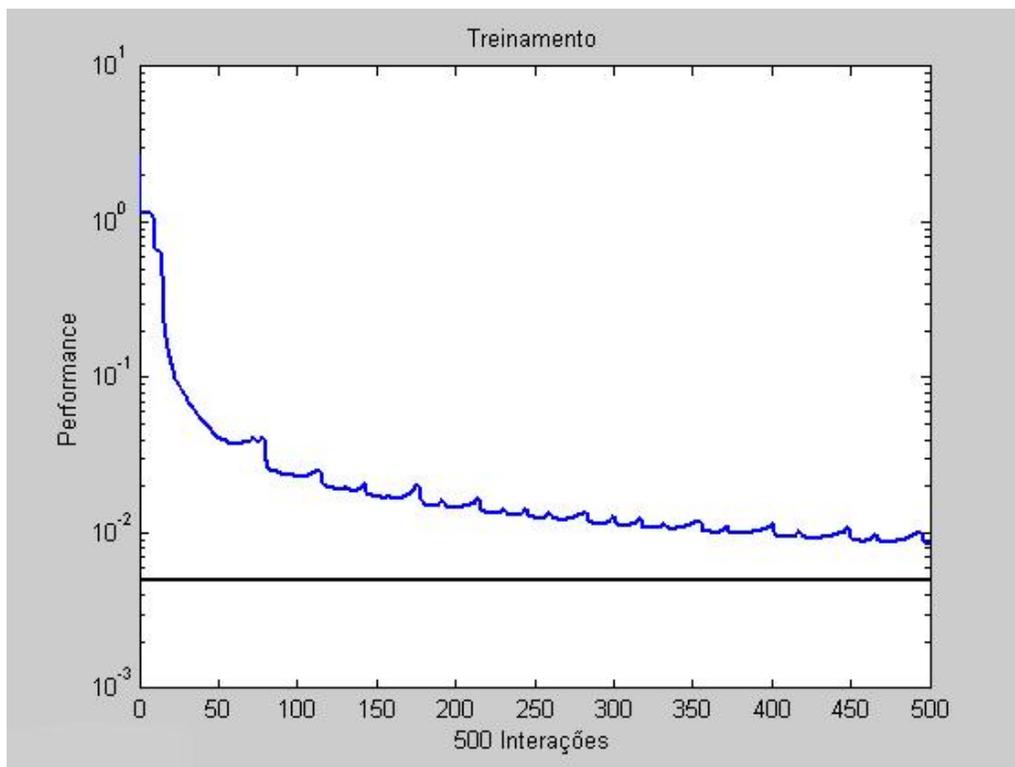


Figura 5.8 – Rede Neural com 200 neurônios e $\eta=0,25$ e TAV=1.05

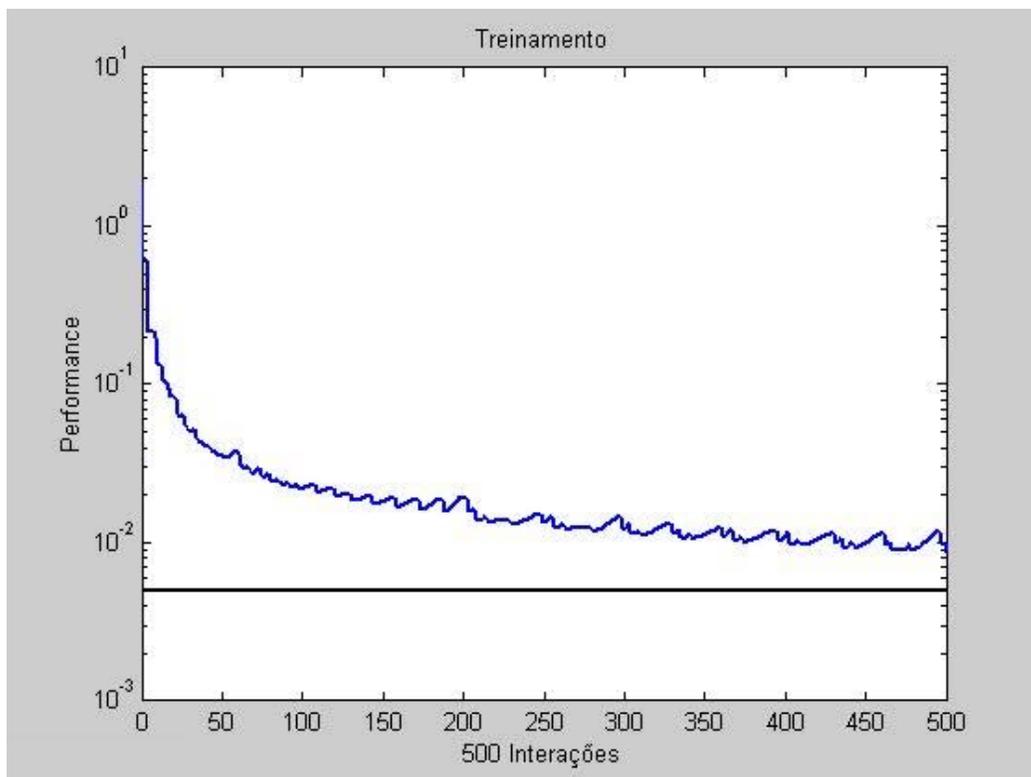


Figura 5.9 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.15

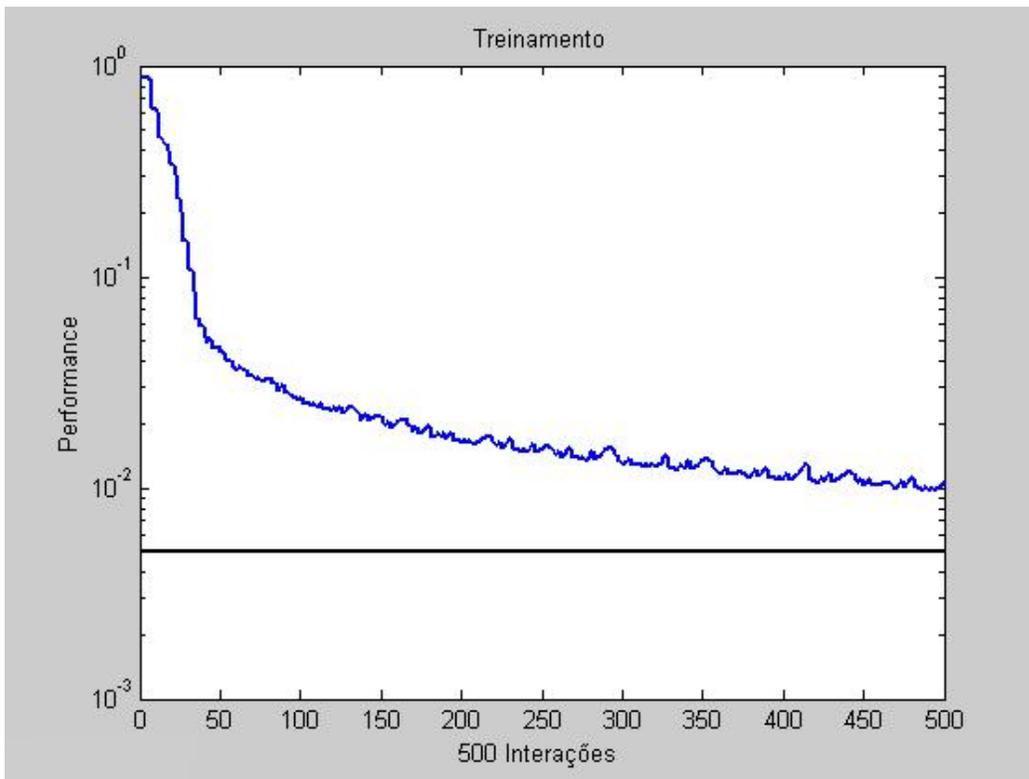


Figura 5.10 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.35

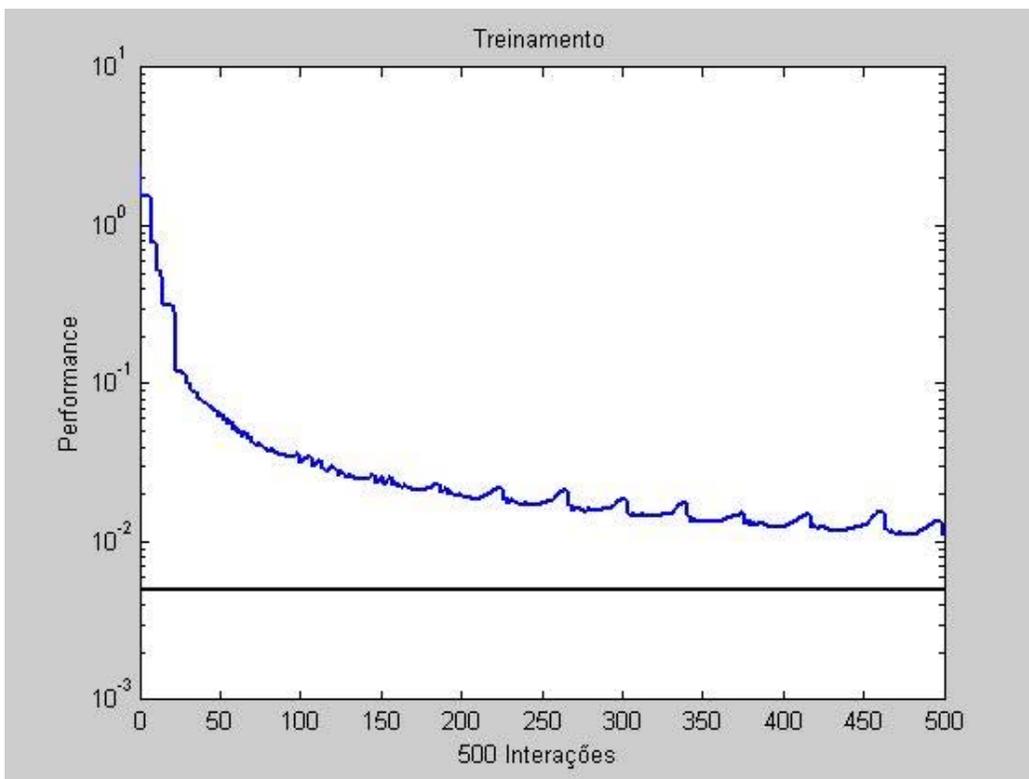


Figura 5.11 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=1.50

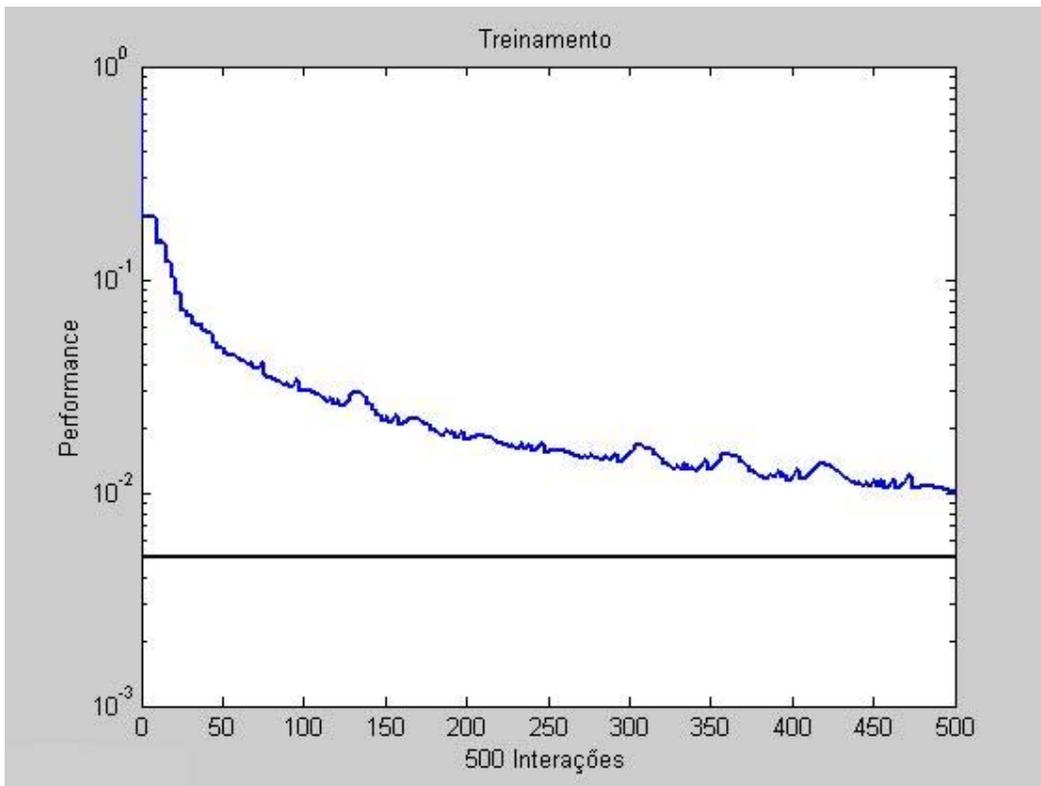


Figura 5.12 – Rede Neural com 200 neurônios e $\eta=0,25$ e Taxa de variação da TAV=2,00

Adotando-se valores para o TAV entre 1,05 e 1,35, é possível chegar a um desempenho similar ao apresentado no algoritmo anterior.

5.3.3.3 - Back-Propagation com Momento

Existe outro modelo de rede neural Back-Propagation que oferece a possibilidade de uma convergência mais rápida, chamada de gradiente descendente com momento (α). O momento permite à rede responder não somente ao gradiente local, mas também às tendências da superfície do erro, atuando como um filtro passa-baixa. O momento permite que o algoritmo ignore pequenas falhas na superfície do erro. Sem o momento, o algoritmo pode ficar preso nessas pequenas imperfeições e convergir para o mínimo da pequena falha.

Foram feitos diversos experimentos, com o objetivo de testar diferentes valores para o momento. Designaremos esse parâmetro pela letra α . Os resultados estão descritos a seguir. Na Figura 5.13 e Figura 5.14 é possível visualizar o comportamento para os diferentes momentos.

- **EQM=0,00825**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,1$
- **EQM=0,00880**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,1$

- **EQM=0,00944**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,1$
- **EQM=0,01476**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,2$
- **EQM=ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,2$
- **EQM=ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,2$
- **EQM=ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,2$
- **EQM=ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,3$
- **EQM= ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,3$
- **EQM= ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,3$
- **EQM= ñ.convergiu**, 500 Interações, 200 Neurônios, $\eta=0,25$ e $\alpha=0,3$

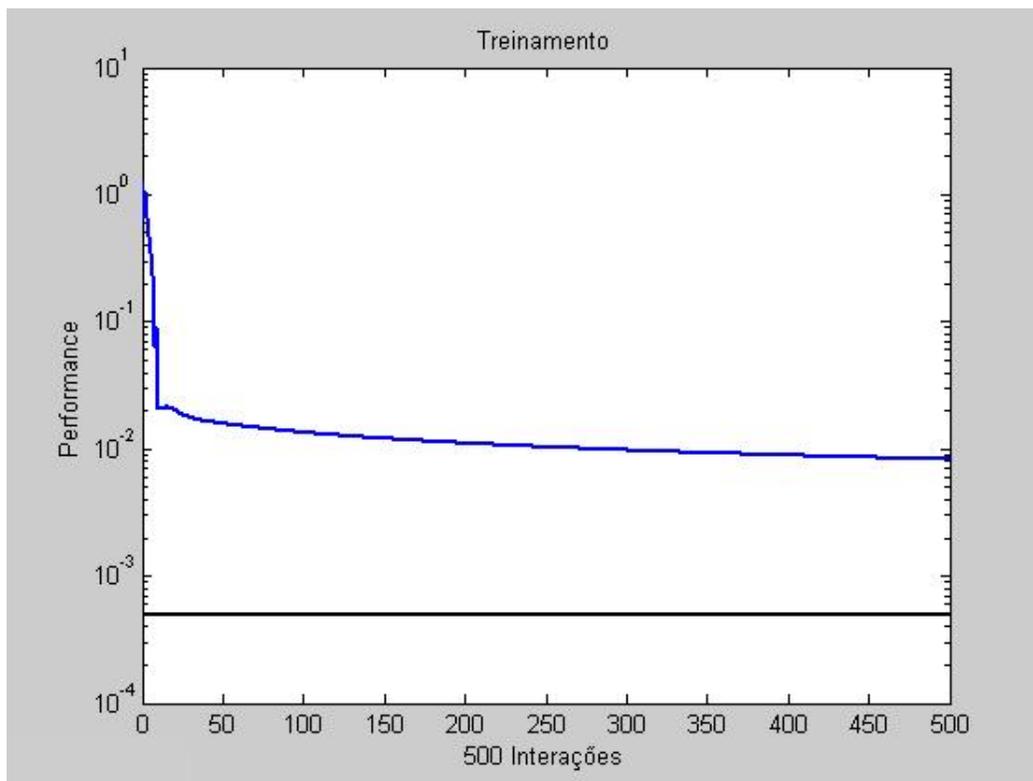


Figura 5.13 – Rede Neural com 200 neurônios e $\eta=0,25$ e $\alpha=0,1$

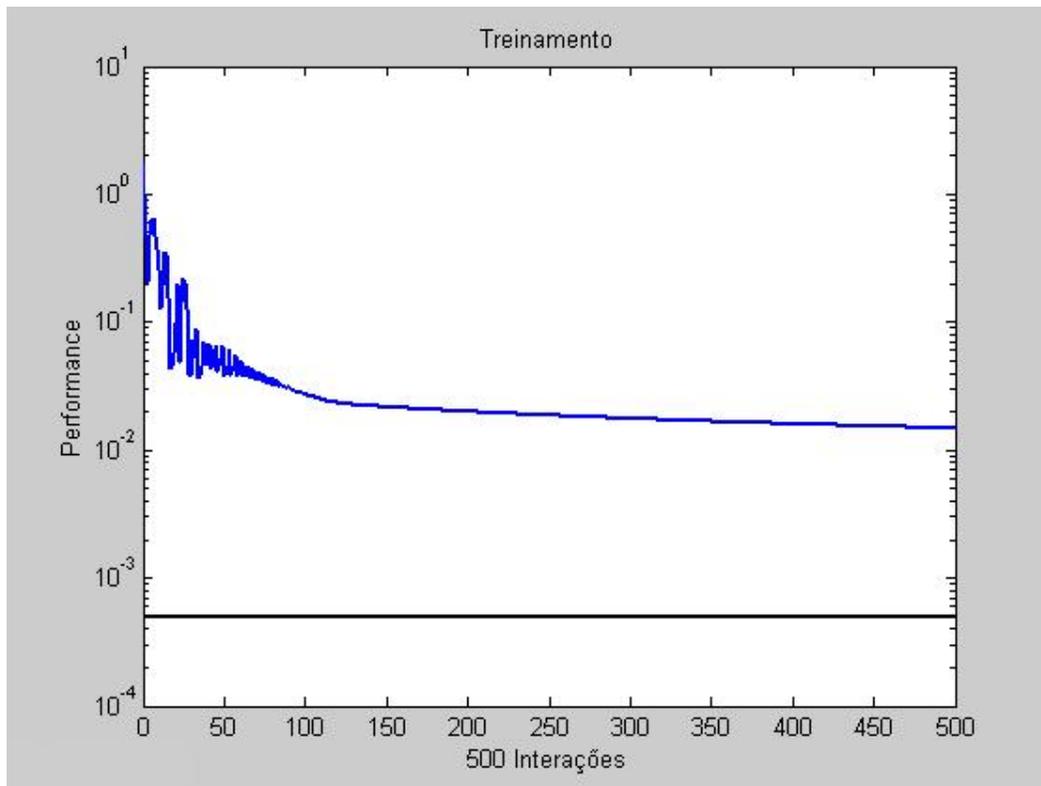


Figura 5.14 – Rede Neural com 200 neurônios e $\eta=0,25$ e $\alpha=0,2$

Adotando-se o valor $\alpha=0,1$, o desempenho se assemelha bastante aos descritos nos algoritmos anteriores.

5.3.3.4 - Análise dos tempos

A seguir, é apresentado o desempenho em relação ao tempo para os três algoritmos descritos. Os valores de tempo foram obtidos, utilizando um computador com processador Athlon XP2800+, 2.08 GHz e 1 Gbyte de RAM.

Na Tabela 5.1 pode ser visualizado o tempo de convergência para 500 interações. A tabela mostra que o algoritmo com momento é o que mais demanda, em termos de processamento, seguido pelo algoritmo descendente simples e depois pelo TAV.

Back-Propagation		
Simplex	TAV	Momento
Tempo (segundos)	Tempo (segundos)	Tempo (segundos)
118,14	114,68	122,85
117,93	113,89	123,93
118,23	112,35	121,37
117,81	112,64	120,41
118,37	114,03	122,35
117,75	101,89	121,41
117,82	114,85	122,38
117,81	115,03	123,07
117,51	114,31	121,91
118,75	113,18	120,49

Tabela 5.1 – Tempo de convergência para 500 iterações

A Tabela 5.2 exibe o tempo que cada algoritmo gastou para convergir para um valor de erro quadrático médio estipulado em 0,099. Os parâmetros utilizados na obtenção dos tempos foram os que tiveram as melhores respostas para as simulações anteriores, ou seja, 200 neurônios, $\eta=0,25$, TAV=1,15 e $\alpha=0,1$.

Back-Propagation		
Simplex	TAV	Momento
Tempo (segundos)	Tempo (segundos)	Tempo (segundos)
89,54	89,71	104,43
80,10	41,28	79,15
123,12	32,93	119,21
100,40	105,64	95,09
96,64	89,10	100,70
105,40	81,98	78,64
Média:	99,2	73,44
		96,20

Tabela 5.2 – Tempo de convergência para chegar ao erro de 0,0099

O melhor desempenho foi obtido com o algoritmo TAV, seguido pelo Momento e depois pelo algoritmo descendente simples. Isso vem comprovar a teoria a respeito de Redes Neurais do tipo Back-Propagation que coloca as redes, com taxa de aprendizado variável e momento, como sendo mais rápidas. No caso do momento, o ganho foi menor, principalmente, pelo fato de o algoritmo ser computacionalmente mais “pesado” que os anteriores.

5.3.3.5 - Aumentando o número de AQLs treinados

Uma vez determinada a versão mais adequada, relativa ao algoritmo Back-Propagation, conjuntamente com os parâmetros, foram feitos testes, no sentido de aumentar o número de AQLs,

para os quais a rede seria treinada. Para os casos anteriores o treinamento foi executado apenas sobre um nível de qualidade (12%). Foram aproximadamente 1.000 vetores, isso porque os valores para a taxa do evento em questão foram representados com até uma casa decimal.

O objetivo é treinar a rede para 12 diferentes níveis de qualidade (*AQL*): 1%, 2%, 7%, 12%, 17%, 22%, 27%, 32%, 37%, 42%, 47%, 52%. Caso fosse utilizado o mesmo nível de detalhamento anterior, existiriam aproximadamente 12.000 vetores. Optou-se, então, por não utilizar casas decimais para representar o valor percentual do evento, o que diminuiu a quantidade de vetores por *AQL*. Utilizando esse princípio, o total para representar os 12 diferentes *AQLs* ficou em aproximadamente 5.200 vetores. Uma consequência direta disso foi a necessidade de redimensionar a camada intermediária, que passou de 200 para 800 neurônios. Esse foi o melhor valor obtido por testes repetitivos com diferentes números de neurônios, sempre analisando os resultados obtidos para o EQM e a convergência ou não da rede. Entretanto, mesmo sendo esse o melhor valor para a camada intermediária, a rede neural não respondeu satisfatoriamente. O EQM ficou na casa de 0,35, muito superior aos valores obtidos anteriormente, e a taxa de acerto caiu para aproximadamente 85%. Com 1.000 neurônios na camada intermediária, o erro quadrático médio foi de aproximadamente 0,39, e a taxa de acerto ficou em 82%.

Depois de muitos testes e sucessivos insucessos, para conseguir aumentar a taxa de acerto, foi feita uma tentativa de colocar uma segunda camada intermediária de neurônios. Após essa mudança recuperou-se boa parte do desempenho da rede neural.

O melhor resultado foi obtido com a rede neural na seguinte configuração:

- Tipo de rede: Back-Propagation com TAV;
- Neurônios na entrada: 4;
- Neurônios na primeira camada intermediária: 800;
- Neurônios na segunda camada intermediária: 200;
- Número de saídas: 1;
- Função de transferências nos neurônios de entrada = Sigmóide;
- Função de transferência na camada Intermediária = Linear;
- Função de transferência na camada Intermediária-Saída = Tangente hiperbólica;
- Quantidade de interações = 400;
- Erro quadrático médio desejado = 0,005;
- Taxa de aprendizado (η) = 0,25;

- Taxa de Aprendizado Variável = 1,05;
- Vetores de entrada = 5200 (vetores representando 12 diferentes *AQLs*).

A Figura 5.15 mostra o comportamento da rede em relação à convergência. A taxa de acerto da rede foi de 99,01%.

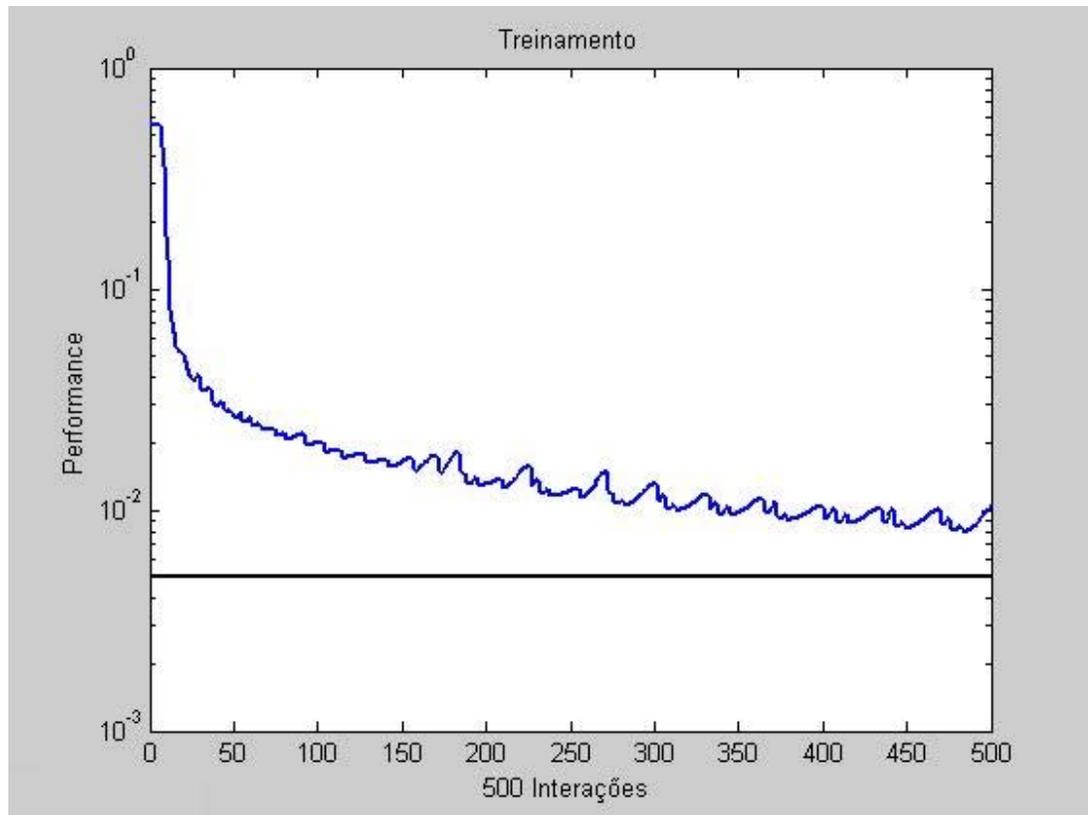


Figura 5.15 – Rede Neural com 200 neurônios e $\eta=0,25$ e TAV=1,05

Após os desenvolvimentos feitos sobre Redes Neurais, é possível concluir que os resultados obtidos, utilizando a rede neural do tipo Back-Propagation, foram extremamente positivos. A taxa de acerto ficou em 99,01%. A única observação dessa configuração final foi o tempo de treinamento necessário para a convergência que ficou em 46 minutos. Boa parte disso se deve à adição de uma segunda camada intermediária e também ao grande vetor de entrada (treinamento) que possui aproximadamente 5.200 elementos.

5.4 - Conclusão

Os testes descritos anteriormente são uma primeira tentativa de enquadrar o trabalho também sob o ponto de vista de Redes Neurais. Foi utilizado um modelo de rede conhecido como Back-Propagation.

Para a primeira parte dos treinamentos foi identificado que o melhor desempenho foi obtido com o algoritmo TAV, seguido pelo Momento e depois pelo algoritmo descendente simples. Isso vem comprovar a teoria a respeito de Redes Neurais do tipo Back-Propagation que coloca as redes, com taxa de aprendizado variável e momento, como sendo mais rápidas. No caso do momento, o ganho foi menor, principalmente, pelo fato de o algoritmo ser computacionalmente mais “pesado” que os anteriores.

Após o aumento do vetor de treinamento ocorreu um aumento significativo no tempo gasto no treinamento, devido à adição de mais uma camada intermediária de neurônios. Houve também uma piora no erro quadrático médio. Entretanto, a taxa de acerto ficou num patamar muito bom, 99,01%.

Na Figura 5.16 temos uma comparação entre os algoritmos mostrados anteriormente e o algoritmo de Redes Neurais.

	Utilização				Mais eficiente		Distribuição		
	Todas chamadas (eventos) exceto OK	Todos os Recursos	Chamadas (Eventos) OK	Ausência de Chamadas	Problemas de Grande Proporção	Problemas de Menor Proporção	Teoria da Renovação	Distribuição Binomial	Distribuição Poisson
Algoritmo - Tempo Real	X	X			X		X		
Algoritmo - Janelas Amostrais	X	X				X		X	
Algoritmo - Chamadas OK		X	X		X	X	X	X	
Algoritmo - Detecção da Ocupação		X		X	X				X
Algoritmo de Redes neurais	X	X	X	X	X	X			

Figura 5.16 – Comparação

Os próximos passos em relação a redes neurais serão os de buscar e extrair novas características, de forma a enriquecer o treinamento. Outra sugestão é utilizar outros modelos de redes neurais.

Capítulo 6

Curva de Abandono

Na década de 70, em telefonia, desenvolveu-se a Teoria das Tentativas Repetidas de Chamadas [87]-[90]. O principal resultado da teoria é um gráfico denominado de curva de abandono, que está ilustrada na Figura 6.1. Essa curva deriva de dados da rede de telefonia de Paris do final da década de 70. Os valores numéricos podem ser visualizados na Tabela 6.1.

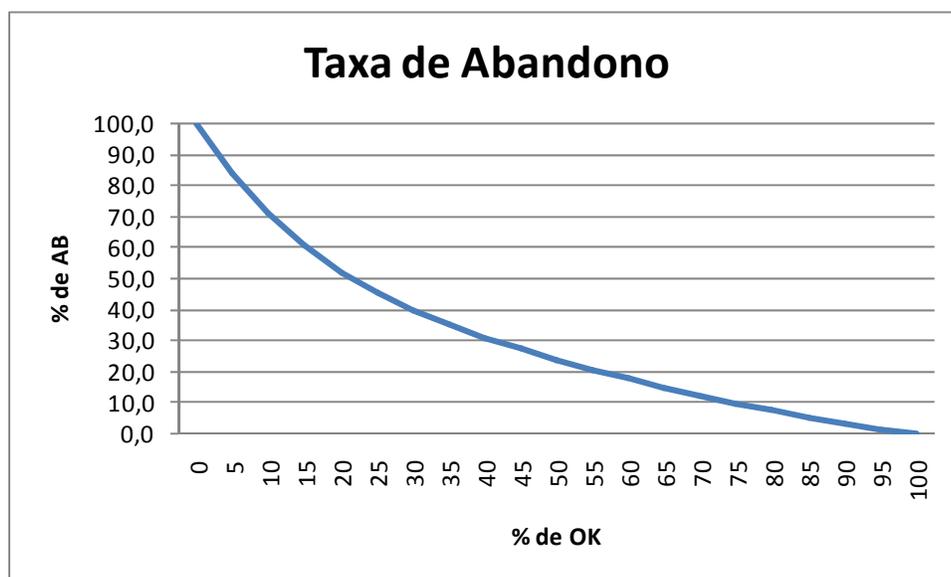


Figura 6.1 – Taxa de AB (Abandono) Teórica

Tive contato com esta curva através do trabalho de Nunes [48][49] que a utilizou de forma brilhante para cálculo de perda de receita. O Sr. Nunes deduziu uma equação a partir da qual é possível estimar as perdas nos elementos da rede que estejam apresentando problemas.

Taxa Ok (%)	Abandono (%)
0%	100,00%
5%	84,00%
10%	71,00%
15%	61,00%
20%	51,50%
25%	45,50%
30%	39,60%
35%	35,00%
40%	30,40%
45%	27,50%
50%	23,90%
55%	20,50%
60%	17,60%
65%	14,80%
70%	12,30%
75%	9,40%
80%	7,40%
85%	5,00%
90%	3,40%
95%	1,40%
100%	0,00%

Tabela 6.1 - Taxa de AB (Abandono) Teórica

Basicamente, a curva ilustra que à medida que a taxa de *OK* de um sistema de telefonia aumenta a taxa de abandono diminui e vice-versa. Essa é uma noção de certa forma empírica, porém a interpretação mais comum e equivocada é que, se um sistema, por exemplo, possui uma taxa de *OK* de 55%, a taxa de abandono seria o restante do percentual, ou seja, 45%. Entretanto, “abandono” não significa a mesma coisa que *NOK*, “Não *OK*”. O princípio que está por trás da palavra “abandono” envolve o princípio básico de uma chamada telefônica, que é o interesse de um assinante “A” em falar com outro assinante “B”. Até que esse princípio não seja satisfeito por completo, não poderemos considerar que houve sucesso na chamada. Se uma chamada for estabelecida de “A” para “B” e antes de terminar voluntariamente por uma das partes, a chamada

falhar, por qualquer motivo, poderemos ter como resultado uma retentativa, ou seja, uma nova chamada sendo estabelecida.

Dessa forma, uma taxa de 50% nas chamadas *OK* representa aproximadamente 23,90% na taxa de abandono, ou seja, de todas as chamadas que não foram completadas apenas 23,90% das pessoas não irão tentar novamente. O restante, 26,10%, irá retentar, até que a chamada seja completada.

Existem trabalhos relacionados à teoria do abandono que buscam modelar o fenômeno. Este modelamento é extremamente complexo [91], por causa das implicações matemáticas na dinâmica das filas e também pela falta de dados sobre o comportamento dos usuários. Nos trabalhos de Gans [92] são mostradas curvas relativas ao tempo entre retentativas das chamadas.

6.1 - Curvas de Abandono para o Sistema de Telefonia Brasileiro

Nosso objetivo neste capítulo foi levantar curvas de abandono para o sistema de telefonia brasileiro. O trabalho consistiu em analisar os bilhetes, verificando o abandono dos assinantes em relação à taxa *OK* do sistema. Essa tarefa exigiu tempo e esforço, pois a quantidade de bilhetes a serem analisados foi elevada. Na Tabela 6.2 são mostrados os estados/cidades, bem como o tipo de sistema e a quantidade de bilhetes analisados:

Estados/Cidades	Tipo	Quantidade Bilhetes
Rio Grande do Sul, Sta Catarina ⁴	Fixo	975.542
São Paulo - SP ⁵	Fixo	450.350

Tabela 6.2 - Relação estados/cidades

6.1.1 - Rio Grande do Sul e Santa Catarina – Telefonia Fixa

Na Figura 6.2 temos a curva de abandono para Rio Grande do Sul e Santa Catarina.

⁴ Foram analisados os bilhetes de tarifação de seis centrais telefônicas das principais cidades desses estados.

⁵ Foram analisados os bilhetes de tarifação de duas centrais telefônicas de São Paulo - SP.

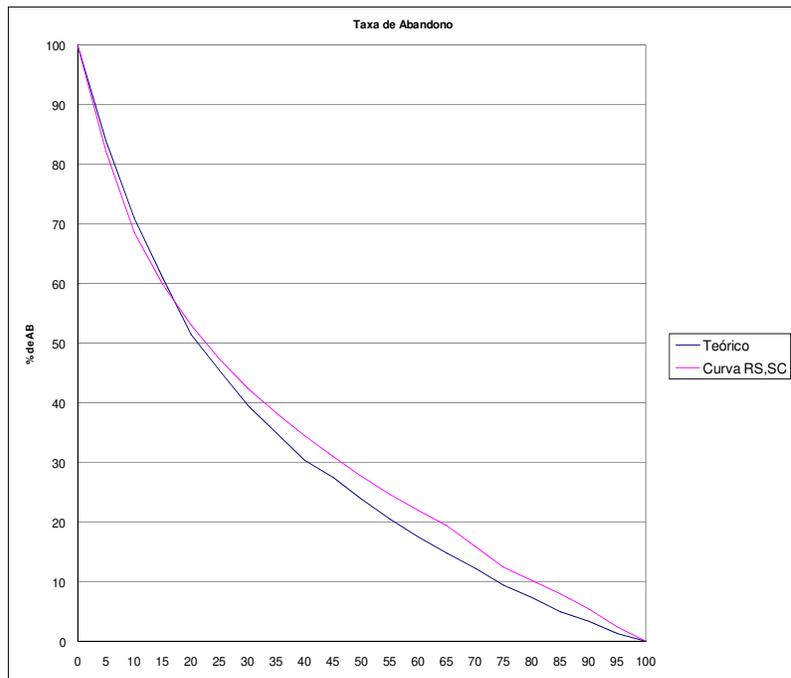


Figura 6.2 – Curva de Abandono para Rio Grande do Sul e Santa Catarina

A curva teórica diz respeito aos dados da rede de Paris, no final da década de 70, e desde então é assumida como padrão. Na Tabela 6.3 estão listados os valores.

% de Abandono		
Telefonia Fixa		
% OK	% Teórico	RS/SC
0	100,0	100,0
5	84,0	82,6
10	71,0	68,4
15	61,0	60,1
20	51,5	53,1
25	45,5	47,2
30	39,6	42,4
35	35,0	38,2
40	30,4	34,4
45	27,5	29,9
50	23,9	27,6
55	20,5	24,6
60	17,6	22,0
65	14,8	19,5
70	12,3	15,8
75	9,4	12,3
80	7,4	10,0
85	5,0	7,9
90	3,4	5,3
95	1,4	2,1
100	0,0	0,0

Tabela 6.3 - Valores de Abandono para RS e SC

6.1.2 - São Paulo – Telefonia Fixa

A curva de abandono gerada para essa cidade pode ser vista na Figura 6.3:

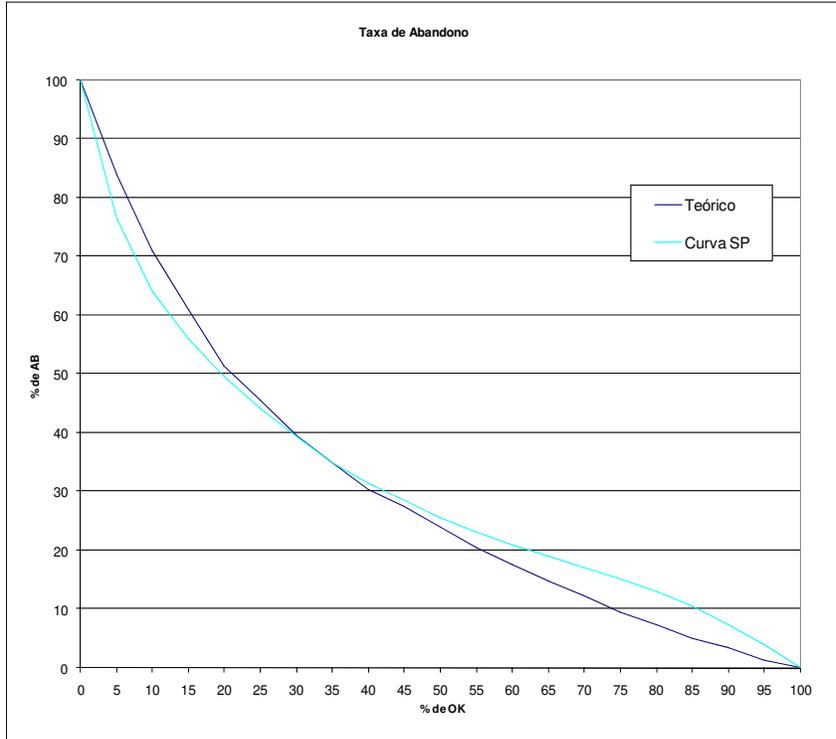


Figura 6.3 – Curva de Abandono para SP – Telefonia Fixa

Na Tabela 6.4 estão listados alguns valores em formato de tabela.

% de Abandono		
Telefonia Fixa		
% OK	% Teórico	SP
0	100,0	100,0
5	84,0	76,6
10	71,0	64,1
15	61,0	56,0
20	51,5	49,7
25	45,5	44,1
30	39,6	39,4
35	35,0	35,0
40	30,4	31,5
45	27,5	28,5
50	23,9	25,5
55	20,5	23,1
60	17,6	21,0
65	14,8	19,0
70	12,3	17,0
75	9,4	15,2
80	7,4	13,0
85	5,0	10,5
90	3,4	7,4
95	1,4	4,0
100	0,0	0,0

Tabela 6.4 – Valores de Abandono para SP – Telefonia Fixa

6.1.3 - Considerações Gerais

Na Figura 6.4 podemos visualizar todas as curvas de abandono mais a curva teórica num mesmo gráfico. De uma maneira geral, o comportamento é bastante parecido. Na Tabela 6.5 temos os valores de abandono dos sistemas analisados.

% de Abandono			
Telefonia Fixa			
% OK	% Teórico	RS/SC	SP
0	100,0	100,0	100,0
5	84,0	82,6	76,6
10	71,0	68,4	64,1
15	61,0	60,1	56,0
20	51,5	53,1	49,7
25	45,5	47,2	44,1
30	39,6	42,4	39,4
35	35,0	38,2	35,0
40	30,4	34,4	31,5
45	27,5	29,9	28,5
50	23,9	27,6	25,5
55	20,5	24,6	23,1
60	17,6	22,0	21,0
65	14,8	19,5	19,0
70	12,3	15,8	17,0
75	9,4	12,3	15,2
80	7,4	10,0	13,0
85	5,0	7,9	10,5
90	3,4	5,3	7,4
95	1,4	2,1	4,0
100	0,0	0,0	0,0

Tabela 6.5 – Agrupamento das Tabelas

Na Figura 6.5 temos a união das curvas apresentadas. Podemos ver que a curva resultante é semelhante à curva teórica, porém o comportamento não é exatamente o mesmo para toda faixa de valores de *OK*.

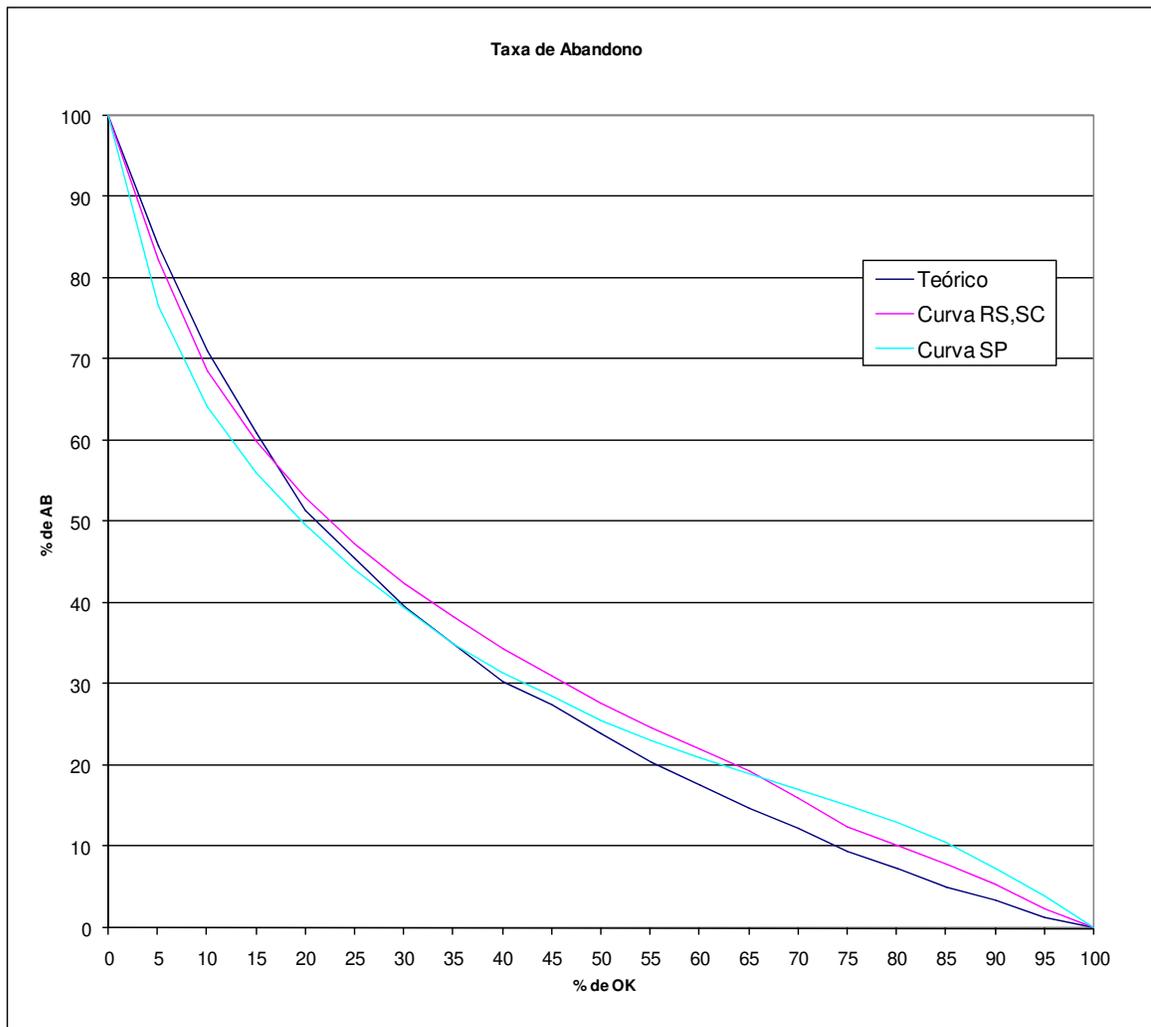


Figura 6.4 – Agrupamento das Curvas

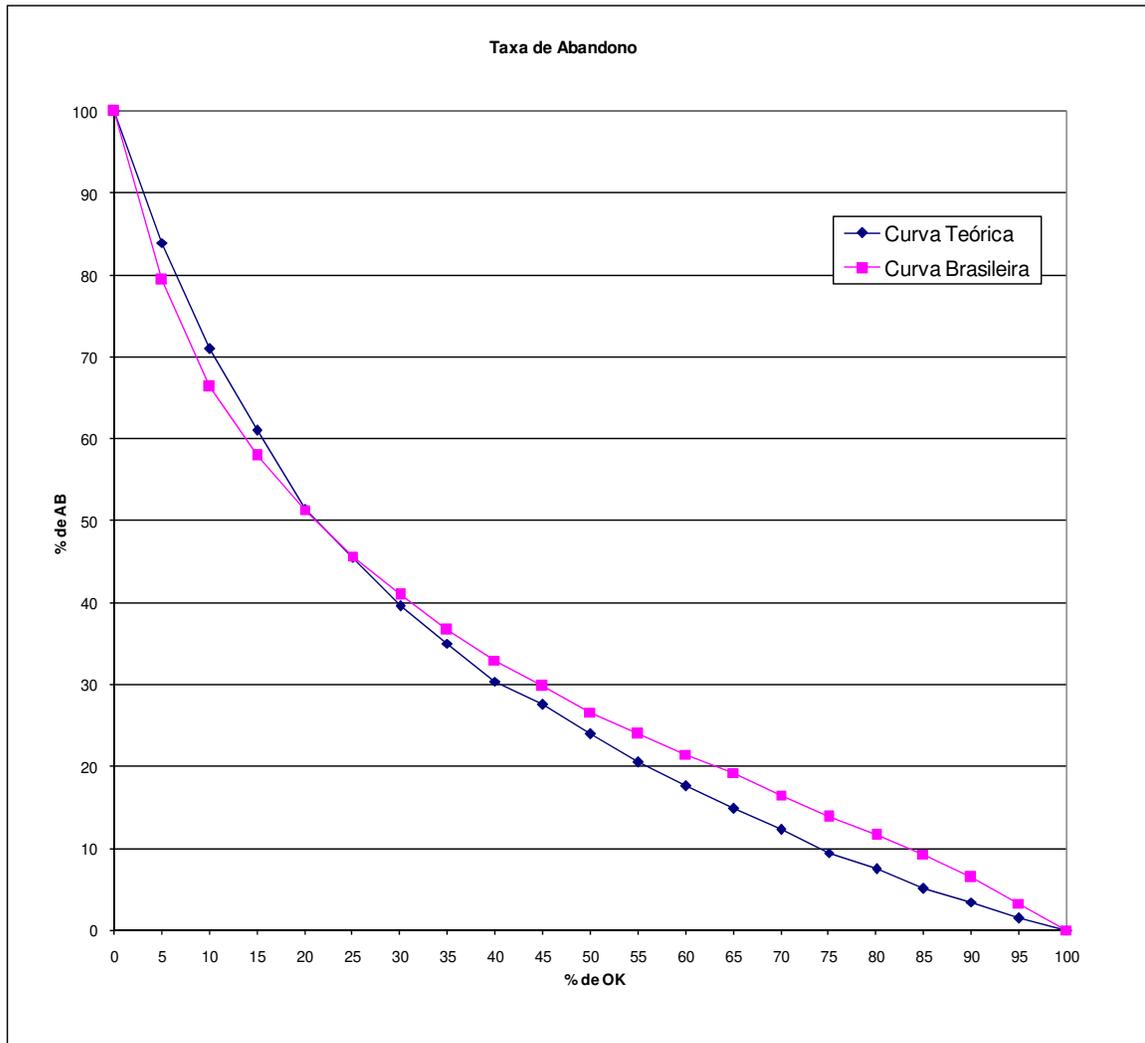


Figura 6.5 - Junção das Curvas

6.2 - Conclusão

Neste capítulo introduzimos importantes fundamentos em relação à conceituação dos problemas pelo ponto de vista da perda de receita. Levantamos as curvas para o sistema de telefonia brasileiro. Como observamos, o levantamento dessas curvas se baseou nos bilhetes de tarifação de sistemas de telefonia fixa para as diferentes regiões do país. Essas curvas são relevantes para o sistema de telefonia nacional, pois não encontramos nada similar na literatura nacional. As curvas possuem uma elevada confiabilidade, tendo em vista que foram construídas sobre um espaço amostral representativo. A curva final, que representa a união de todas as demais, pode ser utilizada como um padrão para os sistemas de telefonia nacional.

Capítulo 7

Conclusões Finais e Outras

Considerações

Nesta tese apresentamos um estudo da aplicabilidade de bilhetes de tarifação na detecção de falhas em sistema de telecomunicações, utilizando para isso algoritmos baseados em teoria de probabilidade e redes neurais.

As principais contribuições obtidas com este trabalho foram:

- Introduzir no meio científico o conceito de *CDR/IPDR* na detecção de falhas e Gerência de Redes (tomada de decisão) - Conforme pesquisamos, nos meios científicos/acadêmicos não existiam trabalhos similares sendo desenvolvidos. A introdução de bilhetes de tarifação é uma nova abordagem que contribui para termos diferentes pontos de vista sobre as tecnologias existentes e sobre novas tecnologias que estão sendo desenvolvidas. Os artigos publicados em congressos e revistas internacionais são as primeiras contribuições na introdução desse assunto nos meios acadêmicos.
- Comparação de resultados entre diferentes algoritmos - Os resultados de diferentes algoritmos são comparados com o objetivo de definir qual a abrangência e aplicação de cada um.
- Os algoritmos além de serem utilizados na detecção de falhas em sistemas de comunicações, também podem ser utilizados para outros fins. Uma evidência disso está no trabalho desenvolvido [93][94] para a Secretaria da Fazenda do Estado de São Paulo, no qual foi

sugerida a utilização dos bilhetes e algoritmos na detecção de evasão de receita/combustível em postos de revenda.

- Introdução do conceito de Redes Neurais na detecção de falhas utilizando CDRs.
- Criação e apresentação de uma Curva de Abandono para o sistema brasileiro – A tese apresenta uma curva de abandono para o sistema de telefonia brasileiro, sendo comparada a uma curva teórica original.
- Incorporação desses algoritmos nas redes de próxima geração – Os conceitos aqui desenvolvidos e explorados podem ser totalmente aplicados a novas tecnologias, tais como *VoIP (Voice over Internet Protocol)*.
- Análise do perfil de usuário através dos bilhetes – Também nesse caso os bilhetes de tarifação oferecem uma grande gama de aplicações, podendo ser utilizados desde a detecção de possíveis fraudes até a análise do perfil de consumo e otimização de planos de serviço que melhor se encaixem a cada usuário.

7.1 - Sugestões para Trabalhos Futuros

Para o futuro pretendemos dar continuidade a este trabalho tratando dos seguintes assuntos:

- Extração automática de características relativas à qualidade dos elementos de um sistema de telefonia. Atualmente as características são extraídas manualmente e o *AQL* é estabelecido da mesma forma. É importante que novos algoritmos sejam desenvolvidos para que esses parâmetros possam ser extraídos de maneira automática. Isso pode ser alcançado utilizando-se Redes Neurais ou, então, modelos estatísticos.
- Desenvolvimento de novos algoritmos utilizando outros modelos de redes neurais. Aliado a isso é importante que seja aperfeiçoada a extração das características das redes de telefonia;
- Levantamento de curvas de abandono para outras regiões do país, bem como para sistemas de telefonia celular, VoIP. Levantamento das curvas relativas aos tempos entre as tentativas de chamadas.
- Utilização do *CDR* na extração de características dos assinantes, procurando utilizar essas informações na detecção do perfil do usuário. Esse perfil do usuário pode ser útil na detecção de fraude, na prospecção de novas oportunidades de negócios, no aumento da satisfação do cliente tendo em vista que este terá um serviço otimizado ao seu perfil, etc.

7.2 - Resumo das Conclusões

A seguir apresentamos um resumo das conclusões que estão descritas com maior detalhe no final de cada capítulo:

- No capítulo 3 que trata do bilhete de tarifação, sua leitura e classificação, podemos concluir que problemas encontrados na formação do bilhete atingem cerca de 0,3% de todos os bilhetes lidos e que isto pode ocasionar perdas tanto para as operadoras quanto para os usuários no caso de cobranças indevidas. Outra conclusão foi o ótimo nível de detalhamento que se chegou ao elaborar uma classificação para os bilhetes com 278 diferentes subeventos. Uma consequência disto foi a criação dos eventos para os sistemas celulares.
- No capítulo 4 que trata do processamento dos dados e detecção das falhas apresenta como conclusão a complementariedade na aplicação dos algoritmos apresentados. Enquanto o algoritmo de Tempo Real é mais eficiente na detecção de falhas de grande proporção, o algoritmo de Espaços Amostrais é mais eficiente na detecção de problemas de menor proporção, ou seja, ambos podem ser implementados conjuntamente. O algoritmo para Chamadas OK e para Detecção de Ocupação complementam os anteriores no sentido de fechar lacunas existentes e assim oferecer uma varredura ampla sobre os problemas que possam surgir.
- No capítulo 5 é apresentado um modelo para detecção de falhas baseado em redes Neurais, a conclusão que o capítulo apresenta é a viabilidade de utilizar estas redes para detecção de falhas no sistema. Os resultados mostram que o melhor desempenho foi obtido com o algoritmo TAV, seguido pelo Momento e depois pelo algoritmo Descendente Simples.
- No capítulo 6 apresentada a Curva de Abandono para o sistema de telefonia nacional. A curva mostra que o comportamento para o sistema de telefonia nacional difere em algumas regiões do gráfico da curva inicialmente proposta na década de 70. A curva apresentada é importante, pois retrata o comportamento do usuário brasileiro em relação a retentativa nas chamadas.

Referências Bibliográficas

- [1] P. Ginzboorg, “Seven Comments on charging and billing”, in Communication of the ACM, Vol.43, No II, pp: 89-92, November 2000.
- [2] IPDR Organization, “Business Solution Requirements -Network Data Management- Usage (NDM-U)”, Version 3.5.0.1, November 2004.
- [3] Sandra Borthick, “Call accounting and billing for IP services”, Business Communications Review, pp: 28-33, August 2001.
- [4] Cotton, S., “Network Data Management – Usage (NDM-U) for IP-Based Services – Version 3.1”, Cotton Management Consulting, August 21, 2002.
- [5] Cotton, S., “Network Data Management – Usage (NDM-U) for IP-Based Services – Service Specification - Version 2.5 A.0”, Cotton Management Consulting, April 13, 2001.
- [6] *ITU-T*, “Telephone Network and ISDN *QoS*, Network Management and Traffic Engineering - *ITU-T E.800*”, International Telecommunication Union, 1994.
- [7] Sposito, R., “Em busca da qualidade para evitar a perda de receita”, Revista Telecom, v.13, nº159, Setembro de 2001.
- [8] M. C. Mozer, R. Wolniewicz, D. B. Grimes, E. Johnson, and H. Kaushansky, “Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry”, in IEEE Transactions on Neural Networks, vol.1, pp: 690-696, 2000.
- [9] [Sitio da *ITU-T*]:
<http://www.itu.int/ITU-T/>
- [10] [Sitio da *ITU-T – TMN M3000*]:
<http://www.itu.int/rec/T-REC-M.3000-200002-I/en>
- [11] [Sitio da *ITU-T – TMN M3010*]:
<http://www.itu.int/rec/T-REC-M.3010-199210-S/en>
- [12] [Sitio da *ITU-T – TMN M3020*]:
<http://www.itu.int/rec/T-REC-M.3020-199507-S/en>
- [13] [Sitio da *ITU-T – TMN M3100*]:
<http://www.itu.int/rec/T-REC-M.3100-200504-I/en>
- [14] [Sitio da *ITU-T – TMN M3180*]:
<http://www.itu.int/rec/T-REC-M.3180-199210-I/en>
- [15] [Sitio da *ITU-T – TMN M3200*]:
<http://www.itu.int/rec/T-REC-M.3200-199704-I/en>
- [16] [Sitio da *ITU-T – TMN M3300*]:
<http://www.itu.int/rec/T-REC-M.3300-199806-I/en>
- [17] [Sitio da *ITU-T – TMN M3400*]:

- <http://www.itu.int/rec/T-REC-M.3400-199704-S/en>
- [18] [Sitio da Ericson]:
<http://www.ericsson.com/>
- [19] [Sitio da Motorola]:
<http://www.motorola.com/>
- [20] [Sitio da Alcatel-Lucent]:
<http://www.alcatel-lucent.com/wps/portal>
- [21] [Sitio da Trópico]:
<http://www.tropiconet.com.br/wwwtsa.asp>
- [22] [Sitio da ZTE]:
<http://www.naukri.com/gpw/zte/index.htm>
- [23] Q. Chen, M. Hsu, U. Dayal, “A data-warehouse/OLAP framework for scalable telecommunication tandem traffic analysis”, in Proc. International Conference in Data Engineering, pp: 201-210, March 2000.
- [24] M. Sestak, A.R. Kaye, “Traffic monitoring in a network of PBXs”, in Global Telecommunications Conference, 1990, vol. 1, pp: 402-407, December 1990.
- [25] L. Yan, M. Fassino, P. Baldasare, “Predicting customer behavior via calling links”, in Proc. IEEE International Joint Conference, vol. 4, pp: 2555-2560, August 2005.
- [26] W. Dong; W. Quan-yu; Z. Shou-yi; L. Feng-xia; W. Da-zhen, “A feature extraction method for fraud detection in mobile communication networks”, in Intelligent Control and Automation, vol.1, pp: 1853-1856, June 2004.
- [27] S. Rosset, U. Murad, E. Neumann, Y. Idan, G. Pinkas, “Discovery of fraud rules for telecommunications-challenges and solutions”, in Proc. ACM SIGKDD International Conference on Knowledge discovery and data mining, pp: 409-413, August 1999.
- [28] C. Cortes, K. Fisher, D. Pregibon, A. Rogers, F. Smith, “Hancock: A language for analyzing transactional data streams”, in ACM Transactions on Programming Languages and Systems (TOPLAS), Vol. 26, Issue 2, pp: 301-338, March 2004.
- [29] S. Trisolini, M. Lenzerini, D. Nardi, “Data integration and warehousing in Telecom Italia” in proc. ACM SIGMOD international conference on Management of data SIGMOD, Vol. 28, Issue 2, pp: 538-539, June 1999.
- [30] Chien, Chen-Fu et. All, “Using Bayesian Network for Fault Location on Distribution Feeder” in IEEE Transactions on Power delivery, Vlo. 17, pp:785-793, july 2002.
- [31] G. Box, G. Tiao, “Bayesian Inference in Statistical Analysis”, Wiley Classics Library, 1992.
- [32] Fayyad, U. Piatetsky-Shapiro, G.; Smyth, P., “The KDD Process for Extracting Useful Knowledge from Volumes of Data Communication of the ACM”, vol. 39, nº 11, p. 27-34, November 1996.
- [33] Ramoni, M.; Sebastiani, P., “Discovering Bayesian Networks in Incomplete Databases”. Knowledge Media Institute, The Open University, Technical Report, nº 46, 1997.
- [34] Turban, E.; Aronson, J. E., “Decision Support Systems and Intelligent Systems”, 6ª Ed. Prentice-Hall, 2001.
- [35] [Sítios com regulamento/normas da ANATEL]:
Geral:<http://www.anatel.gov.br/home/default.asp>
SCM:http://www.anatel.gov.br/Comunicacao_Multimidia/default.asp?CodArea=33&CodPrinc=1

- [36] [Siteio com informações estatísticas sobre VoIP]:
http://www.teleco.com.br/voip_estatis.asp
- [37] IPDR, Inc, “Network Data Management – Usage (NDM-U) IP-based Services”, Version 3.1, IPDR Org(www.ipdr.org), 1999-2002.
- [38] Winemiller, E., Roff, J.T., Heyman, B., Groom, R., “Visual Basic 6 Database”, Macmillan Computer Publishing, 1998.
- [39] [Siteio com dados da ferramenta Microsoft Access]:
<http://www.microsoft.com/en/us/default.aspx>
- [40] Alencar, M.S., “Telefonia Digital”, 4ª Edição, Editora Érica, 2002.
- [41] [Siteio com a norma da Anatel relativa ao SMP]:
http://www.anatel.gov.br/Portal/documentos/biblioteca/resolucao/2002/anexo_res_316_2002.pdf
- [42] Feller, W., “An introduction to Probability Theory and Its Applications – Volume 1”, John Wiley & Sons, Inc, 1968.
- [43] Feller, W., “An introduction to Probability Theory and Its Applications – Volume 2”, John Wiley & Sons, Inc, 1968.
- [44] Cox, D.R., “Renewal Theory”, Mthuen & Company.LTD, 1970.
- [45] Hunter, J.J., “Renewal Theory in Two Dimensions: Basic Results, Advances in Applied Probability”, 6, 376-391, 1974.
- [46] Barlow, R.E., Proschan, F., “Mathematical Theory of Reliability”, Chapter 3, Published January 1996, (ISBN-13: 9780898713695 | ISBN-10: 0898713692).
- [47] [Siteio com a definição do operador matemático]:
http://en.wikipedia.org/wiki/Floor_function#Some_properties_of_the_ceiling_function
- [48] Nunes, G., “Teoria de Probabilidade”, Curso de Processos Estocásticos aplicados à Engenharia, 2001, não publicado.
- [49] Nunes, G., “Teoria da Renovação”, Curso de Processos Estocásticos aplicados à Engenharia, 2001, não publicado.
- [50] Site da ANATEL com as normas, resoluções, portarias e decretos do sistema de Telefonia Celular:
- [51] Gomes, S.M.C., “Tráfego - Teoria e Aplicações”, Mkrón McGraw-hill, 1991.
- [52] Bear, D., " Principles of Telecommunication Traffic Engineering", Peter Peregrinus LTD., England, 1976.
- [53] Flood, J.E., "Telecommunications Switching, Traffic and Networks", Prentice Hall, 1995.
- [54] Bellanu, J., “Digital Telephony”, second Edition, John Wiley & Sons Inc, 1999.
- [55] Siemens, A.G., Teoria de Tráfego Telefônico:Tabelas e Gráficos”, Edgard Blucher, Rio de janeiro, 1975.
- [56] Sokolnikoff, I.S., “Mathematics of Physics and Modern Engineering – Chap. 1 e 8 “, McGraw-Hill – New York, 1998.
- [57] Korn, G.A., Korn, T.M., “Mathematical Handbook for Scientists and Engineers – Chap. 18”, Dover Publications – New York, 1996.
- [58] Motoyama, S., Bonatti, I., “Teoria de Tráfego Telefônico”, Siteio com o material descritivo: <http://www.dt.fee.unicamp.br/~motoyama/ie670/aulas/poisson-Aulas2e3.pdf>
- [59] Haykin, S., “Redes Neurais: Princípios e Práticas”, Bookman Editor, 2ª Edição, 2007.
- [60] Bishop, C.M., “Neural Networks for Pattern Recognition”, 1995.

- [61] Ripley, B.D. & Hjort, N.L., “Pattern Recognition & Neural Networks”, 1995.
- [62] Waibel, A. & Lee, K., “Readings in Speech Recognition. Morgan Kaufmann”, 1990.
- [63] Lippmann, R.P., “Review of Neural Networks for Speech Recognition”. *Neural Computation* 1, 1-38, 1989.
- [64] Bourlard, H.A., “How Connectionist Models Could Improve Markov Models for Speech Recognition. *Advanced Neural Computers*”, pp.247-254, Amsterdam: North-Holland, 1991.
- [65] Morgan, N. & Bourlard, H., “Continuous Speech Recognition Using Multilayer Perceptrons With Hidden Markov Models. *International Conference on Acoustics, Speech, and Signal Processing*, pp.413-416, Albuquerque, NM, 1990.
- [66] Renals, S., Morgan, N., Cohen, M., Franco, H. & Boulard, H., “Improving Statistical Speech Recognition. *International Joint Conference on Neural Networks*”, Vol. 2, pp. 302-307, Baltimore, MD, 1992.
- [67] Robinson, A.J. & Fallside, F., “A Recurrent Error Propagation Speech Recognition System”. *Computer Speech and Language* 5, 259-274, 1991.
- [68] Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. & Lang, K.J., “Phoneme Recognition Using Time-Delay Neural Networks”. *IEEE Transactions on Acoustic, Speech, and Signal Processing ASSP-37*, 328-339, 1989.
- [69] Renals, S., “Radial Basis Function Network for Speech Pattern Classification. *Electronic Letters*”, 25, 437-439, 1989.
- [70] Guyon, I., “Applications of Neural Networks to Character Recognition”. *International Journal of Pattern Recognition and Artificial Intelligence* 5, 353-382, 1991.
- [71] Jackel, L.D., Boser, B., Denker, J.S., Graf, H.P., LeCun, Y., Guyon, I., Henderson, D., Howard, R.E., Hubbard W. & Solla, S.A., “Hardware Requirements for Neural Net Optical Character Recognition”. *International Joint Conference on Neural Networks*, Vol. 2, pp. 855-861, San Diego, CA, 1990.
- [72] Simard, P., Victorri, B., LeCun, Y. & Denker, J., “Tangent Prop-A Formalism for Specifying Selected Invariances in an Adaptive Network”. *Advances in Neural Information Processing Systems 4* (J.E. Moody, S.J. Hanson, and R.P. Lippmann, eds.), pp. 895-903, San Mateo, CA, Morgan Kaufmann, 1992.
- [73] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. & Jackel, L.D., “Handwritten Digit Recognition With a Back-propagation Network. *Advances in Neural Information Processing Systems*”, 2 (D.S. Touretzky, ed.), pp. 598-605, San Mateo, CA, Morgan Kaufmann, 1990.
- [74] Jiang, Q., “Principal Component Analysis and Neural Network Based Face Recognition”. URL: <http://people.cs.uchicago.edu/~qingj/ThesisHtml/>.
- [75] Lawrence, S., Giles, C.L., Tsoi, A.C. & Back, A.D., “Face Recognition: A Convolutional Neural Network Approach”. *IEEE Transactions on Neural Networks*, 8, 98-113, 1997.
- [76] Coetzel, L., Botha, E.C., “Fingerprint recognition in low quality images”, *Pattern Recognition* 26, n° 10, p.441-460, 1998.
- [77] Srinivasan, V.S., Murthy, N.N., “Detection of Singular Points in Fingerprint Images,. *Pattern recognition* 25, n° 2, p.139-153,1997.
- [78] Mitra, P., Mitra, S., Pal, S.K., “Evolutionary Modular MLP with Rough Sets and ID3 Algorithm for Staging of Cervical Cancer.”, *Neural Computing \& Applications*, pages 67-76, 10, Springer-Verlag, London, 2001.

- [79] Mitchell, A., “A Survey of Existing Digital Artificial Neural Network Implementations, Concerning Their Suitability for Integration into the POetic Tissue”. 2002.
- [80] Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 a) Application of an Artificial Immune System in a Compositional Timbre Design Technique. In C. Jacob et al. (Eds): ICARIS 2005, Lecture Notes in Computer Science 3627, pp 389-403.
- [81] Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 b) Interactive Control of Evolution Applied to Sound Synthesis. in Markov, Z., Russel, I. (eds.) Proceedings of the 18th International Florida Artificial Intelligence Research Society (FLAIRS), Clearwater Beach, Florida, EUA, pp. 51-56.
- [82] Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 c) “Topological Self-Organizing Timbre Design Methodology Using a Kohonen's Neural Network”. 10 Simpósio Brasileiro de Computação e Música, Belo Horizonte, Brazil.
- [83] Haykin, S., “Redes Neurais - Princípios e Prática”, BookMan Company Editora, São Paulo, 2002.
- [84] Schalkoff, R., “Pattern Recognition: Statical, Structural and Neural Approaches”, John Wiley & Sons, Inc., Singapore, 1992.
- [85] Freeman, J.A.,; Skapura, D.M., “Neural Networks-Algoritms, Applications, and Programming Techniques”, Addison-Wesley Publishing Company, Massachusetts, 1991.
- [86] Hertz, J., Krogh, A., Palmer, R.G., “Introduction to the Theory of Neural Computation”, Addison-Wesley Publishing Company, Massachusetts, 1991.
- [87] Le Gall, P., “Sur les problèmes du trafic téléphonique général direct et sans attente”, Ann Télécommun, vol 34, no 9-10, 1979.
- [88] Le Gall, F., Bernoussou, J., Garcia, J.M., “A state dependent one moment model grade of service and traffic evaluation in circuit networks”, presented at the 11th Int. Teletraffic Congr., Kyoto, Japan, Sept. 1985.
- [89] Le Gall, F., Bernoussou, J., Garcia, J.M., “A one moment model for telephone networks with dependence on link blocking probabilities”, presented at Symposiun IFIP WG 7-3, Perform 84, Paris December 1984.
- [90] Le Gall, F., Bernoussou, J., “An Analytical formulation for grade of service determination in telephone networks”, IEEE Trans. Commun., vol. COM-31, PP.420-424, 1983.
- [91] Mehrota, V., Fama J., “Call center simulation modeling: methods, challenges, and opportunities”, Proceedings of the 2003 Winter Simulation Conference, 2003.
- [92] Gans, N., Koole, G., Mandelbaum, A., “Commissioned Paper Telephone Call Centers: Tutorial, Review, and Research Prospects”, Manufacturing & Service Operations Management 2003; 5(2): 79–141.
- [93] Breda, G.D., Mendes, L.S., Bottoli, M.L., Araújo, C.R., Costa, M.H.M., “Produto 1 - Projeto de Controle de Vazão de Combustíveis – SEFAZ”, Sumário técnico contendo descritivo técnico do sistema de medição, Unicamp, Dezembro 2005.
- [94] Breda, G.D., MENDES, L.S., Bottoli, M.L., Araújo, C.R., Costa, M.H.M., “Produto 3 - Transmissão de Dados dos Sistemas de Mediçãõ de Vazão de Combustíveis Líquidos em Postos de Revenda”, Relatório técnico contendo descritivo técnico do sistema de medição, Unicamp, Dezembro 2005.
- [95] Montgomery, D.C., “Applied Statistics and Probability for Engineers”, John Wiley & Sons, Inc, 1994.

- [96] Lipschutz, S., “Probabilidade”, Makron Books, 1994.
- [97] Spiegel, M.S., “Probabilidade e Estatística”, MaKron Books, 1978.
- [98] Prática Telebrás SDT 210.110.703.
- [99] Promon Eletrônica, “Guia de Sinalização Telefônica – MFC”, Guia de treinamento TD-TD0-008, Promon Eletrônica Ltda, Novembro de 1993, não publicado.
- [100] Promon Eletrônica, “Sinalização por Canal Comum nº7”, Guia de Treinamento TD-TD0-017, Promon Eletrônica Ltda, Maio de 1996, não publicado.
- [101] ITU, “*Specification of Signalling System n° 7*”, Recommendation Q.700-Q.716, *ITU-Internacional Telecommunications Union*, March 1993.
- [102] ITU, “*Specification of Signalling System n° 7 - Causes*”, Recommendation Q.767, *ITU-Internacional Telecommunications Union*, March 1993.
- [103] Carvalho, F.T.A; Soares Neto, V., “Telecomunicações: tecnologia de centrais telefônicas – CPA-E, CPA-T, tráfego e sinalização”, Editora Érica, 1999.

Apêndice A. Bilhetes de Tarifação

A.1 - CDR da Primeira Tecnologia estudada

O bilhete é dividido nos seguintes campos:

- Campo que apresenta as características do assinante chamador (A);
- Campo que apresenta as características do assinante chamado (B);
- Campo que apresenta o número do assinante A, seja ele fixo ou móvel;
- Campo que apresenta a identificação do sistema no qual o assinante A está fazendo a ligação telefônica;
- Campo que identifica a área do sistema em que o assinante A está;
- Campo que indica se o assinante A é um telefone visitante, *roaming*;
- Campo que identifica que tipo de ligação telefônica é essa, tarifada, livre, etc.;
- Campo que apresenta o número serial do telefone do assinante A;
- Campo que indica o estado do assinante B, bem como seu perfil;
- Campo que indica o número que foi discado;
- Campo que indica o número que foi chamado, ou seja, número para o qual a ligação telefônica foi encaminhada;
- Campo que faz a identificação do sistema do assinante B;
- Campo que identifica a região do sistema do assinante B;
- Campo que identifica se o assinante B está fora de seu sistema, *roaming*;
- Campo que identifica o número serial do telefone do assinante B;
- Campo que identifica o tipo de ligação telefônica que está em curso, celular-celular, celular-fixo, fixo-celular, etc.;
- Campo que identifica o número para o qual deverá ser gerado o bilhete;
- Campo que indica se o assinante a ser tarifado é regular, ou se possui algum plano especial de tarifação;

- Campo que indica o que acontece com a ligação telefônica;
- Campo que indica o número serial do telefone que será tarifado;
- Campo que identifica qual foi o tipo de desconexão da ligação telefônica;
- Campo que serve para o propósito de correlacionar/juntar bilhetes fruto do *handoff* no sistema A;
- Campo que identifica a tecnologia envolvida na ligação telefônica no sistema A;
- Campo que serve para o propósito de correlacionar bilhetes oriundos de *handoff* no sistema B;
- Campo que identifica a tecnologia envolvida na ligação telefônica no sistema B;
- Campo que identifica o número do tronco de saída da central A, no início da ligação;
- Campo que identifica o número do juntor, dentro do tronco de saída da central A, no início da ligação;
- Campo que identifica o número do tronco de saída da central A, no final da ligação;
- Campo que identifica o número do juntor, dentro do tronco de saída da central A, no final da ligação;
- Campo que identifica o horário em que começou a ligação;
- Campo que identifica se ocorreram *handoffs* durante a ligação;
- Campo que identifica o número do tronco de entrada na central A, no início da ligação;
- Campo que identifica o número do juntor, dentro do tronco de entrada na central A, no início da ligação telefônica;
- Campo que identifica o número do tronco de entrada na central A, no final da ligação;
- Campo que identifica o número do juntor, dentro do tronco de entrada na central A, no final da ligação;
- Campo que identifica o horário em que terminou a ligação;
- Campo que identifica qual foi o serviço de traduções de ligações utilizado;
- Campo que apresenta em segundos o tempo que durou a ligação telefônica, após o atendimento do assinante B;
- Campo que identifica se a ligação telefônica feita foi completada com sucesso, ou, então, o motivo do insucesso;
- Campo que apresenta o tipo de resposta do assinante B.

A.2 - CDR da Segunda Tecnologia estudada

O bilhete é dividido nos seguintes campos:

- Campo que identifica os dois principais tipos de ligações;
- Campo que identifica o horário em que o assinante apertou o “*send*”, ou em que a ligação chegou à central;
- Campo que identifica o horário em que houve alguma desconexão da chamada, seja do assinante A ou B;
- Campo que identifica a diferença entre o tempo de originação e terminação;
- Campo que identifica o tempo de conversação;
- Campo que identifica o número do tronco de saída da central A, no início da ligação;
- Campo que identifica o número do juntor, dentro do tronco de saída da central A, no início da ligação;
- Campo que identifica o número do tronco de saída da central A, no final da ligação;
- Campo que identifica o número do juntor, dentro do tronco de saída da central A, no final da ligação;
- Campo com a identificação da central em que foi iniciada a ligação;
- Campo que identifica a central em que o móvel estava quando a ligação terminou;
- Campo que contém uma informação necessária quando o bilhete for referente a um *Intersystem Handoff*;
- Campo indicador do que aconteceu com a ligação telefônica;
- Campo que indica a quantidade de partes variáveis que terá o bilhete;
- Campo que identifica o sub-registro;
- Campo que indica as características que foram utilizadas durante a ligação;
- Campo indicador do tipo da ligação: fixo-móvel, móvel-fixo, móvel-móvel, móvel-indeterminado, fixo-indeterminado, fixo-fixo;
- Campo utilizado no caso de *Intersystem Handoff* para identificar as pernas do bilhete;
- Campo utilizado para identificar o bilhete;
- Campo que contém o número do assinante A, seja ele fixo ou móvel;
- Campo que contém o número serial do telefone do assinante A;

- Campo que indica se o assinante A é local ou visitante, *roaming*;
- Campo que indica a quantidade de *handoffs* que foram feitos pelo assinante A;
- Campo que indica a quantidade de *handoffs* que foram feitos pelo assinante B;
- Campo que indica se o assinante B é local ou visitante, *roaming*;
- Campo que indica a central em que estava o assinante A quando a ligação foi concluída;
- Campo que identifica o número do tronco de entrada na central A, quando a ligação foi concluída;
- Campo que identifica o número do juntor, dentro do tronco de entrada na central A, quando a ligação foi concluída;
- Campo que indica se a ligação for terminada no móvel, esse campo será preenchido com o número do telefone do assinante B;
- Campo que indica se a ligação for terminada no móvel, esse campo será preenchido com o número serial do telefone do assinante B;
- Campo que indica para quem o apelido (*alias*) do móvel será utilizado: 0 (zero) para móvel atual ou 1 (um) para transferência para outro telefone móvel;
- Campo que contém o *alias* do móvel, ou seja, o *TLDN*, número temporário atribuído aos assinantes que estão fora de sua área de serviço, *roaming*;
- Campo que indica o horário em que o assinante B atendeu;
- Campo que indica o horário em que o assinante B desligou;
- Campo que apresenta em segundos o tempo que durou a ligação, após o atendimento do assinante B;
- Campo que indica a central em que estava o assinante A, quando a ligação foi iniciada;
- Campo que identifica o número do tronco de entrada na central A, quando a ligação foi iniciada;
- Campo que identifica o número do juntor, dentro do tronco de entrada na central A, quando a ligação foi iniciada;
- Campo que indica os números que foram discados do aparelho;
- Campo que indica se houve alguma transferência, esse campo apresentará o número do telefone para o qual a ligação foi transferida.

A.3 - Subeventos

Na Tabela 8.1 estão listados os 278 subeventos fruto da classificação dos bilhetes. Os subeventos apresentados são derivados dos eventos apresentados no capítulo 2, item 3.5 - Classificação dos Bilhetes.

Tabela 8.1 – Classificação dos Bilhetes (eventos)

SUB-EVENTO	Comentários
DI.1	Central à frente acusou nº de dígitos incompletos- Chamada Móvel-Fixo
DI.2	Central à frente acusou nº de dígitos incompletos- Chamada Fixo-Indefinido
DI.3	Central à frente acusou nº de dígitos incompletos- Chamada Fixo-Fixo
DI.4	Chamada não encaminhada devido à falha nas traduções - Chamada Móvel-Móvel
DI.5	Chamada não encaminhada devido à falha nas traduções - Chamada Móvel-Indefinido
DI.6	Chamada não encaminhada devido à falha nas traduções - Chamada Fixo-Móvel
DI.7	Chamada não encaminhada devido à falha nas traduções - Chamada Fixo-Indefinido
DI.8	Classe de serviço não possui plano de tradução
DI.9	Não existe plano de rota para esse código de grupo - Chamada Móvel-Indefinido
DI.10	Não existe plano de rota para esse código de grupo - Chamada Fixo-Indefinido
DI.11	Não existe plano de rota para esse código de grupo - Chamada Móvel-Fixo
DI.12	Não existe plano de rota para esse código de grupo - Chamada Fixo-Móvel
DI.13	Chamada não encaminhada devido a nº insuficiente de dígitos - Chamada Móvel-Indefinido
DI.14	Chamada não encaminhada devido a nº insuficiente de dígitos - Chamada Fixo-Indefinido
DI.15	Sem chave de pré-tradução válida ou nº de dígitos discados não combina com o formato - cham. Móvel-Indef.
DI.16	Sem chave de pré-tradução válida ou nº de dígitos discados não combina com o formato - cham. Fixo-Indef.
DI.17	Roteamento de exceção de Chamada foi designado p/gerar esse Final de Seleção - Chamada Móvel-Indefinido
DI.18	Roteamento de exceção de Chamada foi designado p/gerar esse Final de Seleção - Chamada Móvel-Indefinido
DI.19	Chamada não encaminhada devido a quantidade de dígitos serem insuficientes
DI.20	Dialing empregado pelo assinante não está dentro de sua classe de serviço - Chamada Móvel-Indef.
DI.21	Dialing format empregado pelo assinante não está dentro de sua classe de serviço - Chamada Fixo-Indef.
DI.22	Nº discado ou entrante possui muitos dígitos
DI.23	Roteamento de exceção de Chamada foi designado p/gerar esse Final de Seleção
DI.24	Roteamento de exceção de Chamada foi designado p/gerar esse Final de Seleção
DI.25	Roteamento de exceção de Chamada foi designado p/gerar esse Final de Seleção
DI.26	Tentativa de exceder o nº máximo de pernas devido à transferência da Chamada - Chamada Fixo-Móvel
DI.27	Tentativa de exceder o nº máximo de pernas devido à transferência da Chamada - Chamada Móvel-Móvel
DI.28	Tel. móvel não habilitado a originar Cham. - Cham. Móvel-Indefinido
DI.29	Caractere inválido detectado em uma sequência de Caracteres discados - Chamada Móvel-Indefinido
DI.30	Caractere inválido detectado em uma sequência de Caracteres discados - Chamada Fixo-Indefinido
DI.31	Tabelas de traduções não fornecem um plano de encaminhamento pré-definido - Cham. Móvel-Fixo
DI.32	Tabelas de traduções não fornecem um plano de encaminhamento pré-definido - Cham. Móvel -
COO.1	Nº Tel. móvel originador está listado como negado a serviço
COO.2	Nº Tel. móvel originador está listado como negado a serviço - Estouro de temporização 01
COO.3	Nº Tel. móvel originador está listado como negado a serviço - Estouro de temporização 47
COO.4	Nº Tel. móvel originador está listado como vago (assinante válido, mas não atribuído a ninguém)
COO.5	Nº Tel. móvel originador não está cadastrado na central
COO.6	nº serial do Tel. móvel originador é inválido - Chamada Móvel-Indefinido
COO.7	Tel. móvel chamador tentando utilizar feature que não está disponível -Chamada Móvel-Indefinido
COO.8	Falha na programação remota de Feature
COO.9	Não há canais de rádio disponíveis para originação - Chamada Móvel-Fixo
COO.10	Não há canais de rádio disponíveis para originação - Chamada Móvel-Móvel
COO.11	Não há canais de rádio disponíveis para originação - Chamada Móvel-Indefinido
COO.12	Tentativa de Chamada por um Tel. de uma unidade móvel, cujo nº foi mudado ou desconectado

SUB-EVENTO	Comentários
COI.1	Não existem troncos de saída - Chamada Móvel-Fixo
COI.2	Não existem troncos de saída - Chamada Fixo-Fixo
COI.3	Não há canais de rádio disponíveis para terminação - Chamada Fixo-Móvel
DP1.1	Tel. móvel chamador desligou a Chamada antes de ela ser completada - Chamada Móvel-Fixo
DP1.2	Tel. móvel chamador desligou a Chamada antes de ela ser completada - Chamada Móvel-Móvel
DP1.3	Tel. móvel chamador desligou a Chamada antes de ela ser completada - Chamada Móvel-Indefinido
DP1.4	Tel. móvel chamador desligou a Chamada antes de ela ser completada - Chamada Fixo-Fixo
DP1.5	Tel. B desligou a Chamada antes de ela ser completada - Chamada Fixo-Móvel
DP1.6	Tel. B desligou a Chamada antes de ela ser completada - Chamada Fixo-Indefinido
DP1.7	Tel. B desligou a Chamada antes de ela ser completada - Chamada Fixo-Fixo
DP2.1	Mensagem de supervisão não foi recebida da central - Chamada Móvel-Fixo
DP2.2	Mensagem de supervisão não foi recebida da central - Chamada Móvel-Móvel
DP2.3	Mensagem de supervisão não foi recebida da central - Chamada Fixo-Fixo
DP2.4	Mensagem de supervisão não foi recebida da central - Chamada Fixo-Móvel
LO.1	Tel. móvel chamado está ocupado - Chamada Móvel-Móvel
LO.2	Tel. móvel chamado está ocupado - Chamada Fixo-Móvel
LO.3	Tel. Fixo chamado está ocupado - Chamada Móvel-Fixo
LO.4	Tel. Fixo chamado está ocupado - Chamada Fixo-Fixo
NR.1	Tel. móvel chamado falhou no responder a Cham. dentro do periodo de tempo de toque - Cham. Fixo-Móvel
NR.2	Tel. móvel chamado falhou no responder a Cham. dentro do periodo de tempo de toque - Cham. Móvel-Móvel
NR.3	Tel. móvel chamado falhou no responder a Cham. dentro do periodo de tempo de toque - Cham. Fixo-Fixo
NR.4	Tel. fixo chamado não responde no período de toque - Chamada Fixo-Móvel
NR.5	Tel. fixo chamado não responde no período de toque - Chamada Móvel-Fixo
NR.6	Tel. fixo chamado não responde no período de toque - Chamada Fixo-Fixo
OU.1	Assinante Chamado não está mais em serviço
OU.2	Tentativa de Chamada por um móvel tendo igual nº de identificação de um móvel já engajado em outra Chamada.
OU.3	Chamada de teste tipo 103
OU.4	A CCITT Signalling System nº7 send digits of na unobtained number
OU.5	Assinante chamado está listado como vago - Chamada Fixo-Móvel
OU.6	Assinante chamado está listado como vago - Chamada Móvel-Fixo
OU.7	Assinante chamado está listado como vago - Chamada Móvel-Móvel
OU.8	Assinante chamado está listado como vago - Chamada Fixo-Fixo
OU.9	Assinante Chamado com nº mudado - Chamada Móvel-Fixo
OU.10	Assinante Chamado com nº mudado - Chamada Móvel-Móvel
OU.11	Tentativa de Chamada para um Tel. móvel, cujo nº foi alterado ou modificado
OU.12	Reservado para uso futuro - Chamada Móvel-Fixo
OU.13	Reservado para uso futuro - Chamada Fixo-Fixo
OU.14	Tel. B(Destino) não habilitado a receber Chamadas - Chamada Fixo-Móvel
OU.15	Tel. B(Destino) não habilitado a receber Chamadas - Chamada Móvel-Fixo
OU.16	Tel. B(Destino) não habilitado a receber Chamadas - Chamada Móvel-Móvel
OU.17	Tel. B(Destino) não habilitado a receber Chamadas - Chamada Fixo-Fixo
OU.18	Chamando para um Tel. móvel que está listado como serviço negado
OU.19	Tel. móvel chamado está com nº serial inválido
OU.20	Não existe descrição desse evento na documentação
OU.21	Número de identificação do móvel está incompleto ou com problemas
OU.22	Área de busca inválida especificada no banco de dados do assinante - Chamada Fixo-móvel
OU.23	Área de busca inválida especificada no banco de dados do assinante - Chamada Móvel-Móvel
OU.24	Reservado para uso futuro
OU.25	Reservado para uso futuro
OU.26	Reservado para uso futuro
OU.27	Reservado para uso futuro
OU.28	Reservado para uso futuro
OU.29	Reservado para uso futuro
OU.30	Reservado para uso futuro

SUB-EVENTO	Comentários
DT1.1	Sinal de terminação recebido de Central à frente, Chamada sem conversação - Chamada Móvel-Fixo
DT1.2	Sinal de terminação recebido de Central à frente, Chamada sem conversação - Chamada Fixo-Fixo
DT1.3	Falha numa chamada de voz na central - Chamada Móvel-Fixo
DT1.4	Falha numa chamada de voz na central - Chamada Móvel-Indefinido
DT1.5	Falha numa chamada de voz na central - Chamada Fixo-Móvel
DT1.6	Falha numa chamada de voz na central - Chamada Móvel-Móvel
DT1.7	Falha numa chamada de voz na central - Chamada Fixo-Indefinido
DT1.8	Falha numa chamada de voz na central - Chamada Fixo-Fixo
DT1.9	Switch Ticketing failure - Chamada Fixo-Móvel
DT1.10	Switch Ticketing failure - Chamada Móvel-Fixo
DT1.11	Switch Ticketing failure - Chamada Móvel-Móvel
DT1.12	Switch Ticketing failure - Chamada Fixo-Fixo
DT1.13	Rota de sinalização nº7 recebeu um procedimento de erro remoto - Chamada Móvel-Fixo
DT1.14	Rota de sinalização nº7 recebeu um procedimento de erro remoto - Chamada Fixo-Fixo
DT1.15	Rota de sinalização nº7 recebeu uma mensagem de serviço incompatível - Chamada Móvel-Fixo
DT1.16	Rota de sinalização nº7 recebeu uma mensagem de serviço incompatível - Chamada Fixo-Fixo
DT1.17	Erro de protocolo entre a BSC e a central
DT1.18	Erro de protocolo entre a BSC e a central
DT1.19	Erro de protocolo entre a BSC e a central
DT1.20	Erro de protocolo entre a BSC e a central
DT1.21	Erro de protocolo entre a BSC e a central
DT1.22	Erro de protocolo entre a BSC e a central
DT1.23	Erro de protocolo entre a BSC e a central
DT1.24	Erro de protocolo entre a BSC e a central
DT1.25	Erro de protocolo entre a BSC e a central
DT1.26	Erro de protocolo entre a BSC e a central
DT1.27	Erro de protocolo entre a BSC e a central
DT1.28	Erro de protocolo entre a BSC e a central
DT1.29	Tronco BSS foi bloqueado
DT1.30	Tel. Móvel não reconhece uma Cham. ou não tem nível de Threshold adequado da célula - Cham. Móvel-Fixo
DT1.31	Tel. Móvel não reconhece uma Cham. ou não tem nível de Threshold adequado da célula - Cham. Fixo-Móvel
DT1.32	Tel. Móvel não reconhece uma Cham. ou não tem nível de Threshold adequado da célula - Cham. Móvel-Móvel
DT1.33	Tel. Móvel não reconhece uma Cham. ou não tem nível de Threshold adequado da célula - Cham. Fixo-Fixo
DT1.34	Tel. móvel reconhece a Chamada, mas não aloca canal de voz - Chamada Fixo-Móvel
DT1.35	Tel. móvel reconhece a Chamada, mas não aloca canal de voz - Chamada Móvel-Fixo
DT1.36	Tel. móvel reconhece a Chamada, mas não aloca canal de voz - Chamada Móvel-Móvel
DT1.37	Tel. móvel reconhece a Chamada, mas não aloca canal de voz - Chamada Móvel-Indefinido
DT1.38	Símbolo da sinalização MFC-R2 recebido indevidamente
DT1.39	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.40	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.41	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.42	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.43	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.44	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.45	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.46	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.47	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.48	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.49	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.50	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.51	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.52	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.53	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.54	Chamada caiu devido à falha no gerenciamento pelo Switch

SUB-EVENTO	Comentários
DT1.55	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.56	Chamada caiu devido à falha no gerenciamento pelo Switch
DT1.57	Um tronco de saída falhou durante o estabelecimento de Chamadas
DT1.58	A CCITT Signalling System enviou uma mensagem incompatível
DT1.59	Timeout entre dígitos da sinalização MFC-R2
DT1.60	Nenhum resposta para a IFAM(CCITT nº 7) Chamada móvel-Fixo
DT1.61	Nenhum resposta para a IFAM(CCITT nº 7) Chamada Fixo-Fixo
DT1.62	Terminating mobile phone reached his target, but did not return the alert acknowledge message
DT1.63	Falha de hardware do canal de voz - Chamada Fixo-Móvel
DT1.64	Falha de hardware do canal de voz - Chamada Móvel-Fixo
DT1.65	Falha de hardware do canal de voz - Chamada Móvel-Móvel
DT1.66	Falha de hardware do canal de voz - Chamada Móvel-Indefinido
DT1.67	Falha de hardware do canal de voz - Chamada Fixo-Indefinido
DT2.1	Chamada caiu após conversação devido à falha na central - Chamada Móvel-Fixo
DT2.2	Chamada caiu após conversação devido à falha na central - Chamada Móvel-Móvel
DT2.3	Chamada após conversação devido à falha na central - Chamada Móvel
DT2.4	Chamada após conversação devido à falha na central - Chamada Fixo-Móvel
DT2.5	Chamada após conversação devido à falha na central - Chamada Fixo-Indefinido
DT2.6	Chamada após conversação devido à falha na central - Chamada Móvel-Fixo
DT2.7	Terminating mobile phone reached his target, but did not return the alert acknowledge message
DT2.8	Chamada Perdida devido a erro de protocolo entre a BSC e a central
DT2.9	Cham. Perdida devido a erro de protocolo entre a BSC e a central com estouro de temporização 02(No answer charge)
HO.1	Queda de RF numa operação de Handoff
HO.2	Queda de RF numa operação de Handoff
HO.3	Queda de RF numa operação de Handoff
HO.4	Queda de RF numa operação de Handoff
RF2.1	Chamada OK atendida e Tarifável que foi perdida devido à Queda RF - Chamada Móvel-Fixo
RF2.2	Chamada OK atendida e não-tarifável que foi perdida devido à Queda RF - Chamada Móvel-Fixo
RF2.3	Chamada OK atendida e Tarifável que foi perdida devido à Queda RF - Chamada Móvel-Móvel
RF2.4	Chamada OK atendida e Tarifável que foi perdida devido à Queda RF - Chamada Móvel-Fixo
RF2.5	Chamada OK atendida e Tarifável que foi perdida devido à Queda RF.
	Transceiver na BSC detectou perda da portadora - Chamada Fixo-Móvel
RF2.6	Chamada perdida devido à Queda RF, Transceiver na BSC detectou perda da portadora - Chamada Móvel-Fixo
RF2.7	Chamada OK atendida e Tarifável que foi perdida devido à Queda RF.
	Transceiver na BSC detectou perda da portadora - Chamada Móvel-Móvel
RF2.8	Chamada perdida devido à Queda RF, Transceiver na BSC detectou perda da portadora - Chamada Móvel-Fixo
RF2.9	Chamada perdida devido à Queda RF, Transceiver na BSC detectou perda da portadora - Chamada Móvel-Indef.
RF2.10	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF2.11	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF2.12	Chamada atendida - Mobile disconnect lost
RF2.13	Chamada atendida - Mobile disconnect lost
RF2.14	Chamada atendida - Mobile disconnect lost
RF2.15	Downlink RF loss - Tel. móvel detectou perda de RF
RF1.1	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.2	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.3	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.4	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.5	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.6	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.7	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.8	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.9	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.10	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.11	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo

Bilhetes de Tarifação

SUB-EVENTO	Comentários
RF1.12	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.13	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.14	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.15	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.16	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.17	Chamada perdida devido à Queda RF - Chamada Móvel-Fixo
RF1.18	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.19	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.20	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.21	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.22	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.23	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.24	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.25	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.26	Chamada perdida devido à Queda RF - Chamada Móvel-Móvel
RF1.27	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.28	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.29	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.30	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.31	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.32	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.33	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.34	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.35	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.36	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.37	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.38	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.39	Chamada perdida devido à Queda RF - Chamada Móvel-Indefinido
RF1.40	Chamada perdida devido à Queda RF
RF1.41	Chamada perdida devido à Queda RF
RF1.42	Chamada perdida devido à Queda RF
RF1.43	Chamada perdida devido à Queda RF
RF1.44	Chamada perdida devido à Queda RF
RF1.45	Chamada perdida devido à Queda RF
RF1.46	Chamada perdida devido à Queda RF
RF1.47	Chamada perdida devido à Queda RF
RF1.48	Chamada perdida devido à Queda RF
RF1.49	Chamada perdida devido à Queda RF
RF1.50	Perda de enlace de subida - Unidade móvel de terminação(Transceptor detectou perda de portadora)
RF1.51	Mobile disconnect lost
RF1.52	Mobile disconnect lost
RF1.53	Downlink de RF foi perdido - Tel. móvel detectou perda de RF
RF1.54	Downlink de RF foi perdido - Tel. móvel detectou perda de RF
OKWAP.1	2ª perna de uma chamada pré-pago WAP OK, a primeira perna até a plataforma possui Final de Seleção 01 numa Chamada OK
OKWAP.2	Chamada pós-pago WAP OK.
POK1.1	Chamada atendida com TC<3 - Chamada Fixo-Móvel
POK1.2	Chamada atendida com TC<3 - Chamada Móvel-Fixo
POK1.3	Chamada atendida com TC<3 - Chamada Móvel-Móvel
POK1.4	Chamada atendida com TC<3 - Chamada Fixo-Fixo
POK1.5	Chamada atendida com TC<3 com estouro de temporização
POK1.6	Assinante Romeiro originando - Chamada Móvel-Fixo com TC<3
POK1.7	Assinante Romeiro originando - Chamada Móvel-Móvel com TC<3
POK1.8	Chamada atendida com estouro de temporização= 48 e TC<=3

SUB-EVENTO	Comentários
OK1.1	Chamada atendida e tarifável - Chamada Fixo-Móvel
OK1.2	Chamada atendida e tarifável - Chamada Móvel-Fixo
OK1.3	Chamada atendida e tarifável - Chamada Móvel-Móvel
OK1.4	Chamada atendida e tarifável - Chamada Fixo-Fixo
OK1.5	Chamada atendida e tarifável com estouro de temporização
OK1.6	Chamada atendida e não tarifável - Chamada Móvel-Fixo
OK1.7	Chamada atendida e não tarifável - Chamada Fixo-Fixo
OK1.8	Assinante Romeiro originando - Chamada Móvel-Fixo
OK1.9	Assinante Romeiro originando - Chamada Móvel-Móvel
OK1.10	Chamada atendida e tarifável com estouro de temporização=48
OK2.1	Sucesso na programação remota de feature
OK3.1	Mensagem SMS OK
IS.1	Falha no estabelecimento da Chamada, envolvendo Intersystem-Roaming, devido à falha na comunicação entre centrais
IS.2	Chamada falhou devido a erro no acesso a dados de Assinante Roaming - Chamada Móvel-Indefinido
PTO.1	Telefone móvel destino sofreu de-registração
PPago.1	Chamada não completada devido a tratamento Final de Seleção 8F
PPago.2	Chamada não completada devido a tratamento Final de Seleção 8F
CO2.1	Foi recebida uma mensagem CCCITT de congestionamento do terminal - Chamada Móvel-Fixo
CO2.2	Foi recebida uma mensagem CCCITT de congestionamento do terminal - Chamada Fixo-Fixo
CO2.3	Congestionamento na rede à frente - Chamada Móvel-Fixo
CO2.4	Congestionamento na rede à frente - Chamada Fixo-Fixo

Apêndice B. – Algoritmos para Detecção

B.1 - Probabilidade

O objetivo nesse item é apresentar a teoria de probabilidades, com enfoque em confiabilidade. O termo confiabilidade pode ser definido como a probabilidade de um equipamento ou sistema executar sua função, satisfatoriamente, por um período de tempo determinado. O conceito de falha adotado nesse trabalho considera que um equipamento não executa sua função como deveria, ou seja, qualquer equipamento, dispositivo, ou sistema falham, quando seu desempenho não atende à função para a qual foi projetado.

Probabilidade matemática, significando possibilidade, é um número variando entre zero e um que mede a expectativa de um evento ocorrer. A probabilidade zero é atribuída para um evento impossível de ocorrer. Se for certo que o evento ocorrerá, a probabilidade é um. Um experimento que pode gerar diferentes resultados, mesmo sendo sempre repetido da mesma maneira, todas as vezes, é chamado de experimento aleatório.

O conjunto de todos os possíveis resultados do experimento aleatório é chamado de espaço amostral do experimento, denotado pela letra “ S ”. Um Espaço Amostral é discreto se consiste de um número finito de resultados. Um evento E é um elemento ou conjunto de elementos do Espaço Amostral S de um experimento aleatório.

Por exemplo, toda vez que uma ligação telefônica é executada, essa pertencerá a uma das possibilidades definidas no seguinte espaço amostral S , definido na Tabela 8.2. O conjunto dos eventos apresentados nesta tabela foi introduzido no capítulo 2, item 3.5 - Classificação dos Bilhetes.

Utilizaremos dois livros como referência para os modelos probabilísticos que serão descritos neste apêndice, são eles: “Applied Statistics and Probability for Engineers”, “Probabilidade” e “Probabilidade e Estatística”, as respectivas referências são: [95]-[97].

Evento	Descrição
OK	Indica as chamadas completadas sem erro.
NR	Indica as chamadas que o assinante destino não respondeu.
LO	Indica as chamadas que encontraram o assinante de destino na condição de ocupado.
CO0	Indica as chamadas não completadas devido ao esgotamento da temporização.
CO1	Indica as chamadas não completadas por congestionamento ou defeito na central origem.
CO2	Indica as chamadas não completadas por congestionamento ou defeito na central destino.
CO3	Indica as chamadas não completadas por devido à falha na troca de sinalização MFC.
OU	Indica as chamadas não completadas devido a outros fatores.
DI	Indica as chamadas não completadas devido a erro de discagem.
DP	Indica que houve desligamento prematuro na chamada pelo assinante origem.
IS	Indica as chamadas não completadas devido ao assinante roamer não possui perfil.
RF1	Indica as chamadas não completadas devido à queda de RF antes da definição da rota.
RF2	Indica as cham.não completadas devido à queda de RF antes do atendimento.
RF3	Indica as chamadas onde houve queda de RF depois do atendimento c/tempo Conversação > 3 s.
PTO	Indica as chamadas não completadas devido a estouro de temporização no paging
POK	Indica as chamadas OK com tempo de conversação <= 3 segundos.
DT	Indica as chamadas não completadas devido a defeito técnico em algum componente da central.
HO	Indica as chamadas não completadas devido a um erro no handoff.
IHO	Indica as chamadas não completadas devido a um erro no InterSystem handoff.
HI	Indica as chamadas não completadas devido a um erro no handoff interno da célula.
PP	Indica as chamadas perdidas na Plataforma de Pré-pago.
FC	Indica as chamadas perdidas devido a falta de crédito.

Tabela 8.2 – Eventos de Sistema

Cada evento E do espaço amostral S é formado por um conjunto de elementos ou possíveis resultados. Como exemplo é possível citar o caso do evento DI, discagem incorreta, que possui os seguintes subeventos. Veja na Tabela 8.3:

Sub-evento	Significado
DI.1	Central a frente acusou nº de dígitos incompletos- Chamada Móvel-Fixo
DI.2	Central a frente acusou nº de dígitos incompletos- Chamada Fixo-Indefinido
DI.3	Central a frente acusou nº de dígitos incompletos- Chamada Fixo-Fixo
DI.4	Chamada não encaminhada devido a falha nas traduções - Chamada Móvel-Móvel
DI.5	Chamada não encaminhada devido a falha nas traduções - Chamada Móvel-Indefinido
DI.6	Chamada não encaminhada devido a falha nas traduções - Chamada Fixo-Móvel
DI.7	Chamada não encaminhada devido a falha nas traduções - Chamada Fixo-Indefinido
DI.8	Classe de serviço não possui plano de tradução
DI.9	Não existe plano de rota para este código de grupo - Chamada Móvel-Indefinido
DI.10	Não existe plano de rota para este código de grupo - Chamada Fixo-Indefinido
DI.11	Não existe plano de rota para este código de grupo - Chamada Móvel-Fixo
DI.12	Não existe plano de rota para este código de grupo - Chamada Fixo-Móvel
DI.13	Chamada não encaminhada devido a nº insuficiente de dígitos - Chamada Móvel-Indefinido
DI.14	Chamada não encaminhada devido a nº insuficiente de dígitos - Chamada Fixo-Indefinido
DI.15	Sem chave de pré-tradução válida ou nº de dígitos discados não combina com o formato - Ch. Móvel-Indefinido
DI.16	Sem chave de pré-tradução válida ou nº de dígitos discados não combina com o formato - Ch. Fixo-Indefinido
DI.17	Roteamento de exceção de Chamada, foi designado p/gerar este CFC - Chamada Móvel-Indefinido
DI.18	Roteamento de exceção de Chamada, foi designado p/gerar este CFC - Chamada Móvel-Indefinido
DI.19	Chamada não encaminhada devido a quantidade de dígitos serem insuficientes
DI.20	Formato de discagem empregado pelo assinante não está dentro de sua classe de serviço - Chamada Móvel-Indefinido
DI.21	Formato de discagem empregado pelo assinante não está dentro de sua classe de serviço - Chamada Fixo-Indefinido
DI.22	Nº discado ou entrante possui muitos dígitos
DI.23	Roteamento de exceção de Chamada, foi designado p/gerar este CFC - Chamada Móvel-Móvel
DI.24	Roteamento de exceção de Chamada, foi designado p/gerar este CFC - Chamada Móvel-Fixo
DI.25	Roteamento de exceção de Chamada, foi designado p/gerar este CFC - Chamada Móvel-Indefinido
DI.26	Tentativa de exceder o nº máximo de pernas devido a transferência da Chamada - Chamada Fixo-Móvel
DI.27	Tentativa de exceder o nº máximo de pernas devido a transferência da Chamada - Chamada Móvel-Móvel
DI.28	Tel.móvel não habilitado a originar chamada - Chamada Móvel-Indefinido
DI.29	Caracter inválido detectado em uma sequência de caracteres discados - Chamada Móvel-Indefinido
DI.30	Caracter inválido detectado em uma sequência de caracteres discados - Chamada Fixo-Indefinido
DI.31	A classe de serviço não provê um plano de roteamento - Chamada Móvel-Fixo
DI.32	A classe de serviço não provê um plano de roteamento - Chamada Móvel-indefinido

Tabela 8.3 – Possíveis resultados do evento DI

Esses são os possíveis resultados de uma ligação telefônica que não obteve sucesso devido a erro de discagem. Todos esses resultados fazem parte do evento DI.

Algumas definições:

- Para um espaço amostral discreto, a probabilidade de um evento, denotado por $P(E)$, é igual à soma das probabilidades dos elementos do evento E ;
- Dois eventos, denominados E_1 e E_2 , tal que, $E_1 \cap E_2 = \emptyset$, são ditos mutuamente exclusivos;
- Uma coleção de eventos, E_1, E_2, \dots, E_k , são ditos mutuamente exclusivos para todos os pares, quando $E_i \cap E_j = \emptyset$, então, $P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) + \dots + P(E_k)$.

B.1.1 - Axiomas da Probabilidade:

Probabilidade é um número que é designado para cada membro de uma coleção de eventos de um experimento aleatório, que satisfaz as seguintes propriedades: Se S é o espaço amostral e E é um evento qualquer do experimento aleatório, então:

$$P(S) = 1; \quad (\text{b.1})$$

$$0 \leq P(E) \leq 1; \quad (\text{b.2})$$

1. Para dois eventos E_1 e E_2 com $E_1 \cap E_2 = \emptyset$ (eventos mutuamente exclusivos)

$$P(E_1 \cup E_2) = P(E_1) + P(E_2). \quad (\text{b.3})$$

O mapeamento da tecnologia apresentada nesse projeto foi formulado, de maneira que todos os eventos sejam mutuamente exclusivos, não existe sobreposição, área comum. Isso fica mais evidente, quando observamos os resultados isoladamente. Não existe classificação dúbia. Cada possível resultado é único e exclusivo. Veja na Tabela 8.4 um exemplo em relação ao evento DI (discagem incorreta).

SUB-EVENTO	Final de Seleção	Direção da cham.	Estouro da temp.	Duração da Cham.
DI.1	0E	01	00	X
DI.2	0E	04	00	X
DI.3	0E	05	00	X
DI.4	35	02	00	X
DI.5	35	03	00	X
DI.6	35	00	00	X
DI.7	35	04	00	X
DI.8	D1	X	00	X
DI.9	65	03	00	X
DI.10	65	04	00	X
DI.11	65	01	00	X
DI.12	65	00	00	X
DI.13	36	03	00	X
DI.14	36	04	00	X
DI.15	3A	03	X	X
DI.16	3A	04	X	X
DI.17	6B	03	00	X
DI.18	6C	03	00	X
DI.19	36	00	00	X
DI.20	3C	03	00	X
DI.21	3C	04	00	X
DI.22	4E	X	X	X
DI.23	67	X	X	X
DI.24	6A	X	X	X
DI.25	6D	X	X	X
DI.26	B5	00	00	X
DI.27	B5	02	00	X
DI.28	5F	03	00	X
DI.29	4F	03	00	X
DI.30	4F	04	00	X
DI.31	D0	01	00	X
DI.32	D0	03	00	X

Tabela 8.4 – Montagem dos possíveis resultados

Todo bilhete possui uma única classificação e se encaixa em um único evento, de tal forma que as propriedades para eventos mutuamente exclusivos podem ser aplicadas.

B.1.2 - Probabilidade Condicional

A probabilidade condicional de um evento A , dado um evento B , é denotada por $P(A | B)$, isso é:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \text{ onde:} \quad (\text{b.4})$$

$$P(B) = \frac{(\text{número de possíveis resultados no evento } B)}{n \text{ (número total de possíveis resultados)}}; \quad (\text{b.5})$$

$$P(A \cap B) = \frac{(\text{número de possíveis resultados em } A \cap B)}{n \text{ (número total de possíveis resultados)}}. \quad (\text{b.6})$$

Regra da Multiplicação:

$$P(A \cap B) = P(A | B) \cdot P(B) = P(B | A) \cdot P(A) \quad (\text{b.7})$$

B.1.3 - Eventos Independentes

Dois eventos são independentes, se, e somente se, qualquer uma das seguintes regras é válida:

$$P(A | B) = P(A); \quad (\text{b.8})$$

$$P(B | A) = P(B); \quad (\text{b.9})$$

$$P(B \cap A) = P(A) \cdot P(B). \quad (\text{b.10})$$

B.1.4 - Teorema de Bayes

O teorema de Bayes é sintetizado das definições anteriores da probabilidade condicional, sendo:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad (\text{b.11})$$

porém: $P(A \cap B) = P(B | A) \cdot P(A)$, dessa maneira:

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}; \quad (\text{b.12})$$

B.1.5 - Variáveis Aleatórias

Geralmente é interessante sintetizar a saída de um experimento aleatório por um simples número. Muitas vezes o espaço amostral tem somente uma descrição dos possíveis resultados. Em alguns casos, a descrição é suficiente, porém, em outros, não é, sendo necessário associar um número com uma saída no espaço amostral. A variável que associa um número com a saída do experimento aleatório é dita uma variável aleatória. Uma variável aleatória é uma função que associa um número real para cada saída, no espaço amostral de um experimento aleatório.

Uma variável aleatória X é discreta, se o conjunto de valores que pode assumir é finito ou numerável. Um evento englobando todas as saídas para qual $X = x$ é designado como $\{X = x\}$, e a probabilidade desse evento é dada por $P(X = x)$. A função distribuição de probabilidade de uma variável aleatória X é a descrição do conjunto de todos os possíveis valores de X . Frequentemente, a função distribuição de probabilidade de uma variável aleatória é o melhor sumário de um experimento aleatório. Suponha que x seja todos os possíveis eventos que uma ligação telefônica possa ser classificada, conforme Tabela 8.5.

X	P(X=x)
OK	0,510
NR	0,050
LO	0,150
CO0	0,015
CO1	0,011
CO2	0,020
CO3	0,020
OU	0,060
DI	0,030
DP	0,030
IS	0,025
RF1	0,021
RF2	0,013
RF3	0,010
PTO	0,001
POK	0,010
DT	0,001
HO	0,001
IHO	0,001
HI	0,001
PP	0,010
FC	0,010
Total:	1,000

Tabela 8.5 – Função distribuição de probabilidade

A função Distribuição Cumulativa de uma variável aleatória discreta X é expressa por $F_X(x)$:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i). \quad (\text{b.13})$$

Para uma variável aleatória discreta X , $F_X(x)$ satisfaz as seguintes propriedades:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f_X(x_i); \quad (\text{b.14})$$

$$0 \leq F_X(x) \leq 1; \quad (\text{b.15})$$

$$\text{se } x \leq y, \text{ então } F_X(x) \leq F_X(y). \quad (\text{b.16})$$

Se X é uma variável aleatória, o valor esperado, ou média, denominado μ_X ou $E(X)$ é:

$$\mu_X = E(X) = \sum_x x \cdot f_X(x), \quad (\text{b.17})$$

Se X é uma variável aleatória, a variância denominada σ_X^2 ou $V(X)$ é dada por:

$$\sigma_X^2 = E(X - \mu_X)^2 = \sum (x - \mu_X)^2 \cdot f_X(x), \quad (\text{b.18})$$

Se X é uma variável aleatória, o desvio-padrão denominado σ_X é a raiz quadrada da variância, σ_X^2 .

Serão abordadas a seguir algumas distribuições de probabilidade, começando pela mais simples, que é a Distribuição Uniforme Discreta.

B.1.6 - Distribuição Uniforme Discreta

Essa é a distribuição mais simples, pois assume um número finito de possíveis valores, cada um com igual probabilidade. Uma variável aleatória uniforme X assume cada um dos valores x_1, x_2, \dots, x_n , com igual probabilidade $\frac{1}{n}$. Dessa maneira:

$$f_X(x; n) = \frac{1}{n}; \quad (\text{b.19})$$

B.1.7 - Distribuição de Bernoulli

Caracteriza-se pelo fato de que apenas dois resultados são possíveis: "sucesso", com probabilidade p , e "fracasso", com probabilidade $(1 - p)$.

$$f_X(x; p, n) = (1 - p)^{n-1} \cdot p; \quad (\text{b.20})$$

Logicamente os termos sucesso ou fracasso são apenas meras denominações, poderíamos utilizar “A” ou “B”, “0” ou “1”.

B.1.8 - Distribuição Binomial

A Distribuição Binomial é também baseada no modelo de um experimento com duas possíveis saídas, sucesso ou insucesso. O tipo de experimento que gera essa distribuição é definido como uma tentativa de Bernoulli, vista anteriormente. A Distribuição Binomial é baseada num experimento aleatório com n tentativas ou repetições, todas independentes umas das outras. Cada nova repetição gera um novo resultado, sucesso ou insucesso.

Se X é uma variável aleatória binomial com os valores p para sucesso e n para a quantidade de tentativas, então:

$$f_X(x; p, n) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}; \quad (\text{b.21})$$

$$\mu_X = E(X) = n \cdot p; \quad (\text{b.22})$$

$$\sigma_X^2 = V(X) = n \cdot p \cdot (1 - p). \quad (\text{b.23})$$

B.1.9 - Distribuição Geométrica ou Binomial Negativa

Imagine um experimento aleatório que é intimamente ligado a uma Distribuição Binomial. Assuma novamente que as séries são independentes, tentativas de Bernoulli com probabilidade constante p de sucesso de cada tentativa. Porém, ao invés de fixar um número de tentativas ou repetições, as tentativas são conduzidas, até que um sucesso seja obtido.

$$f_X(x; p, r) = \binom{x-1}{r-1} \cdot (1 - p)^{x-r} \cdot p^r; \quad (\text{b.24})$$

$$\mu_X = E(X) = \frac{r}{p}; \quad (\text{b.25})$$

$$\sigma_X^2 = V(X) = \frac{r \cdot (1 - p)}{p^2}. \quad (\text{b.26})$$

B.1.10 - Distribuição de Poisson

Tal como nos modelos determinísticos, nos quais algumas relações funcionais desempenham um importante papel (exemplo: linear, quadrática, exponencial, trigonométrica, etc), também verificamos que, na construção de modelos não determinísticos para fenômenos observáveis, algumas distribuições de probabilidade surgem mais frequentemente que outras. Um motivo para isso é que, da mesma maneira que no caso determinístico, alguns modelos matemáticos relativamente simples parecem ser capazes de descrever uma classe bastante grande de fenômenos. A função de Distribuição de Poisson é usada como uma aproximação da Distribuição Binomial, quando n é muito grande, tendendo ao infinito e p é muito pequeno. A Distribuição de Poisson pode ser deduzida da Distribuição Binomial, utilizando a teoria dos limites. Não iremos fazer essa demonstração, pois está amplamente coberta na literatura matemática. Após realizarmos todas as manipulações chegaremos à:

$$f_X(x; \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \text{ onde,} \quad (\text{b.27})$$

$$\sigma_x^2 = V(X) = \lambda . \quad (\text{b.28})$$

O modelo de Poisson é muito utilizado em confiabilidade, em especial no aprovisionamento de peças de reposição.

B.1.11 - Distribuição Normal

A Distribuição Normal ou de Gauss é a mais familiar dentre as distribuições de probabilidade contínuas e também uma das mais importantes em estatística. É definida como se segue:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad (\text{b.29})$$

onde μ_x e σ_x são constantes maiores que zero, dizemos que:

- μ_x = Média;
- σ_x^2 = Variância;
- σ_x = Desvio-Padrão.

O comportamento da Distribuição Normal em relação a μ_x e σ_x é apresentado na Figura 8.1 e Figura 8.2:

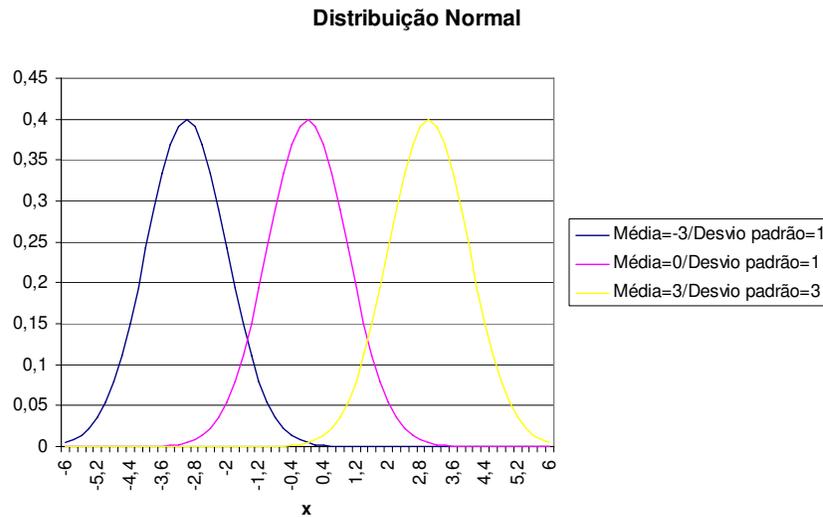


Figura 8.1 – Distribuição Normal com σ_x fixo e μ_x variável

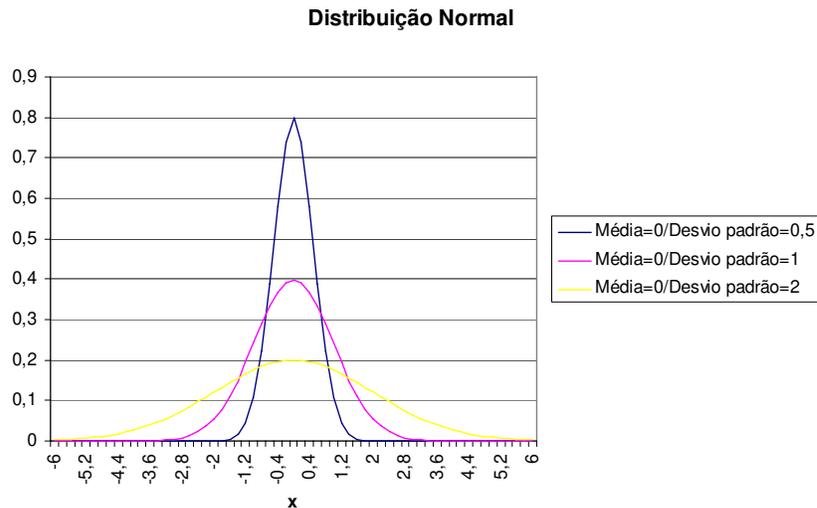


Figura 8.2 – Distribuição Normal com σ_x variável e μ_x fixo

A média (μ_x) desloca a função em relação ao eixo x , enquanto o desvio-padrão (σ_x) altera o formato da função. Quanto menor o desvio-padrão, mais aguda será a forma, em outras palavras, menor será a dispersão dos elementos em relação à média. Quanto maior o desvio-padrão, mais aberta será a curva e maior será a dispersão dos elementos em relação à média.

B.1.12 - Variáveis Aleatórias Contínuas

Se o range de uma variável X contém um intervalo, seja finito ou infinito, de números reais, então X é uma variável aleatória contínua. Em alguns exemplos, a variável aleatória é discreta, porém, devido ao range de valores possíveis ser muito grande, é mais conveniente utilizar modelos baseados em Variáveis Aleatórias Contínuas.

A função $f_X(x)$ é uma Função Densidade de Probabilidade de uma variável aleatória contínua X , se para qualquer intervalo de números reais $[x_1, x_2]$:

$$f_X(x) \geq 0; \quad (\text{b.30})$$

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1; \quad (\text{b.31})$$

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f_X(u) du. \quad (\text{b.32})$$

O ponto importante é que $f_X(x)$ é utilizado para calcular a área que significa a probabilidade que X assume dentro do intervalo $[x_1, x_2]$. Uma consequência de trabalharmos com uma variável aleatória contínua é que $P(X = x) = 0$. O resultado advém do fato de fazermos uma integral cujo intervalo é $[x, x]$, ou seja:

$$\int_x^x f_X(u) du = 0. \quad (\text{b.33})$$

Logo, se X é uma variável aleatória contínua, para qualquer x_1 e x_2 , vale a seguinte afirmação:

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2). \quad (\text{b.34})$$

A função $F_X(x)$ representa a Função Distribuição Cumulativa de probabilidade, ou seja, através dela encontramos a probabilidade que uma variável X assume dentro do intervalo $[-\infty, x]$. Calculamos a probabilidade, nesse intervalo, integrando a função $f_X(x)$ dentro desse intervalo:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(u) du, \text{ para } -\infty < x < \infty. \quad (\text{b.35})$$

Uma consequência disso é que:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f_X(u) du = 1. \quad (\text{b.36})$$

A média de X , denominada $E(X)$ ou μ_X , é:

$$E(X) = \mu_X = \int x \cdot f_X(x) dx. \quad (\text{b.37})$$

A variância de X , denominada $V(X)$ ou σ_X^2 , é:

$$V(X) = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f_X(x) dx. \quad (\text{b.38})$$

O desvio-padrão de X é $\sigma_X = [V(X)]^{1/2}$.

B.1.13 - Variáveis Aleatórias Discretas

Supondo X uma variável aleatória em S , com contradomínio infinito e numerável, diz-se que $X(S) = \{x_1, x_2, \dots\}$. Tais variáveis aleatórias são chamadas variáveis aleatórias discretas. Como no caso finito, $X(S)$ será um espaço de probabilidade, se definirmos a probabilidade de x_i como $f(x_i) = P(X = x_i)$ e chamarmos f de distribuição de X :

$$\begin{aligned} x_1 &= f(x_1); \\ x_2 &= f(x_2); \\ x_3 &= f(x_3); \\ &\dots \end{aligned} \quad (\text{b.39})$$

A média e variância de uma variável discreta são definidas de maneira similar à variável contínua. A diferença é que a integral é substituída pelo somatório, ou seja, a integral é utilizada num ambiente contínuo, e o somatório é utilizado num ambiente discreto.

A esperança, ou média, $E(X)$ e a variância $V(X)$ ou σ_X^2 são respectivamente:

$$E(X) = \mu_X = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots = \sum_{i=1}^{\infty} x_i \cdot f(x_i); \quad (\text{b.40})$$

$$V(X) = \sigma_X^2 = (x_1 - \mu)^2 \cdot f(x_1) + (x_2 - \mu)^2 \cdot f(x_2) + \dots = \sum_{i=1}^{\infty} (x_i - \mu)^2 \cdot f(x_i); \quad (\text{8.41})$$

A função Distribuição Cumulativa de uma variável aleatória discreta X , denotada por $F(x)$, é dada por:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i); \quad (\text{b.42})$$

Para uma variável aleatória discreta X , $F(x)$ satisfaz as seguintes propriedades:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i); \quad (\text{b.43})$$

$$0 \leq F(x) \leq 1; \quad (\text{b.44})$$

$$\text{Se } x \leq y, \text{ então } F(x) \leq F(y). \quad (\text{b.45})$$

B.1.14 - Desigualdade de Tchebichev

Um teorema de grande importância em probabilidade e estatística, que revela uma propriedade geral de variáveis aleatórias contínuas ou discretas, com média e variância finitas, é a desigualdade de Tchebichev.

Seja X uma variável aleatória (discreta ou contínua) com média μ_x e variância σ_x^2 , ambas finitas. Então, se ε é um número positivo,

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad (\text{b.46})$$

ou, com $\varepsilon = kv$,

$$P(|X - \mu| \geq kv) \leq \frac{1}{k^2}. \quad (\text{b.47})$$

B.1.15 - Lei dos Grandes Números

O teorema conhecido como Lei dos Grandes Números é uma consequência interessante da desigualdade de Tchebichev.

Teorema I (lei dos grandes números) - Sejam X_1, X_2, \dots variáveis aleatórias, discretas e contínuas, mutuamente independentes, cada uma com média μ_x e variância σ_x^2 , ambas finitas.

Então, se:

$$S_n = X_1 + X_2 + \dots + X_n \quad (n = 1, 2, \dots),$$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0; \quad (\text{b.48})$$

Como $\frac{S_n}{n}$ é a média aritmética de X_1, X_2, \dots, X_n , o teorema afirma que a probabilidade de a média aritmética $\frac{S_n}{n}$ diferir de seu valor esperado μ_x de mais de ε , tende a zero, quando $n \rightarrow \infty$. Um resultado mais forte, que se poderia esperar, é que $\lim_{n \rightarrow \infty} S_n/n = \mu$, mas tal resultado é, na realidade, falso. Pode-se, entretanto, demonstrar que o $\lim_{n \rightarrow \infty} S_n/n = \mu$ com

probabilidade igual a um. Esse resultado é por vezes denominado lei forte dos grandes números, e, em contraposição, o teorema I (lei dos grandes números) é designado lei fraca dos grandes números. A referência “lei dos grandes números”, sem especificação, indica a lei fraca dos grandes números.

B.1.16 - Teoria da Renovação/Eventos Recorrentes

A teoria da renovação [42]-[44] começou com o estudo de algumas probabilidades associadas à falha e substituição de componentes, tal como componentes eletrônicos. Após essa aplicação ficou claro que essa teoria poderia ser aplicada em muitos outros campos.

De uma maneira geral, um experimento padrão ε qualifica-se para a teoria da renovação, se, após cada ocorrência de ε , os experimentos começarem do zero, no sentido de que as próximas ocorrências de ε formam uma réplica do experimento. O caso mais simples acontece, quando ε representa a ocorrência de sucesso num experimento de Bernoulli. O tempo de espera até o primeiro sucesso tem uma Distribuição Geométrica. Quando o primeiro sucesso ocorre, o experimento inicia-se novamente, e o número de experimentos entre o n e $n+1$ sucessos tem a mesma Distribuição Geométrica. O tempo de espera até o sucesso n é a soma das n variáveis independentes.

Aplicação num Experimento de Bernoulli

Suponhamos que r denote número inteiro positivo e ε seja a representação da ocorrência de uma seqüência de sucessos de comprimento r , num experimento de Bernoulli. A variável u_n é a probabilidade de ε no experimento de número n , e f_n é a probabilidade de que a primeira seqüência de comprimento r ocorra no experimento de número n .

A probabilidade de que r tentativas, $n, n-1, n-2, \dots, n-r+1$, resultarem em sucesso é dado por p^r . A probabilidade de que ε ocorra na tentativa $n-k$ ($k=0,1,\dots,r-1$) e a seguinte k tentativa resulte em k sucessos é igual a $u_{n-k} \cdot p^k$. Assim essas r possibilidades são mutuamente excludentes. Teremos, então, uma relação recorrente⁶:

$$u_n + u_{n-1} \cdot p + \dots + u_{n-r+1} \cdot p^{r-1} = p^r, \quad (\text{b.49})$$

válido para $n \geq r$;

⁶ A dedução a seguir foi retirada do livro: W. Feller, “An introduction to Probability Theory and Its Applications – Volume 1”, Chapter XIII, p. 303-341, John Wiley & Sons, Inc, 1968.

$$u_1 = u_2 = u_3 = \dots = u_{r-1} = 0, u_0 = 1. \quad (\text{b.50})$$

Multiplicando a equação **b.49** por s^n , o lado esquerdo fica :

$$\{U(s) - 1\} \cdot (1 + p \cdot s + p^2 \cdot s^2 + \dots + p^{r-1} \cdot s^{r-1}); \quad (\text{b.51})$$

e no lado direito $p^r \cdot (s^r + s^{r+1} + \dots)$. As duas séries são geométricas:

$$\{U(s) - 1\} \cdot \frac{1 - (p \cdot s)^r}{1 - p \cdot s} = \frac{p^r \cdot s^r}{1 - s}, \text{ ou:} \quad (\text{b.52})$$

$$U(s) = \frac{1 - s + q \cdot p^r \cdot s^{r+1}}{(1 - s) \cdot (1 - p^r \cdot s^r)}. \quad (\text{b.53})$$

As funções gerações de $\{u_n\}$ e $\{f_n\}$ são relacionadas por:

$$U(s) = \frac{1}{1 - F(s)}. \quad (\text{b.54})$$

$$F(s) = \frac{p^r \cdot s^r \cdot (1 - p \cdot s)}{1 - s + q \cdot p^r \cdot s^{r+1}} = \frac{p^r \cdot s^r}{1 - p \cdot s(1 + p \cdot s + \dots + p^{r-1} \cdot s^{r-1})}, \quad (\text{b.55})$$

a média μ pode ser obtida diretamente da equação (1), uma vez que $\mu_n \rightarrow \mu^{-1}$.

$$\mu = \frac{1 - p^r}{q \cdot p^r}, \quad (\text{b.56})$$

a variância é obtida calculando-se a derivada de $F(s)$:

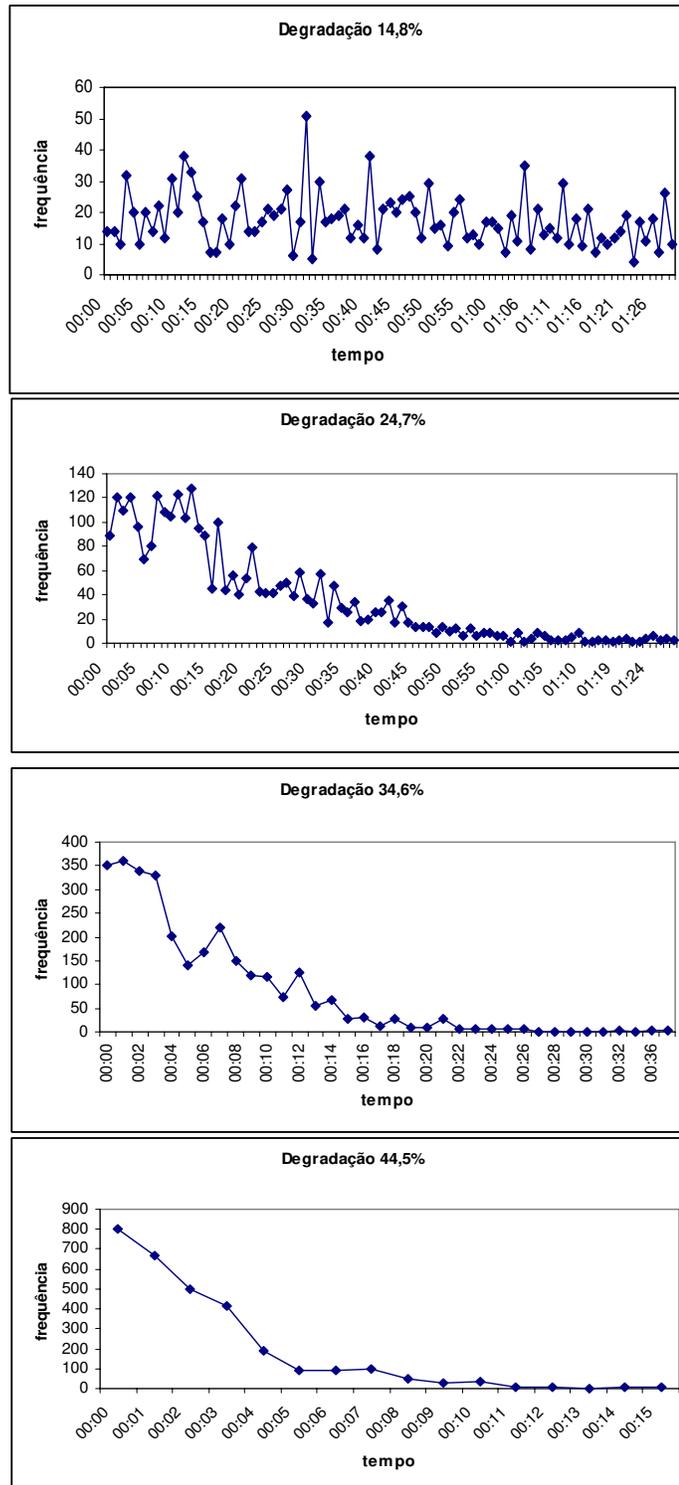
$$\sigma^2 = \frac{1}{(q \cdot p^r)^2} - \frac{2 \cdot r + 1}{q \cdot p^r} - \frac{p}{q^2}, \quad (\text{b.57})$$

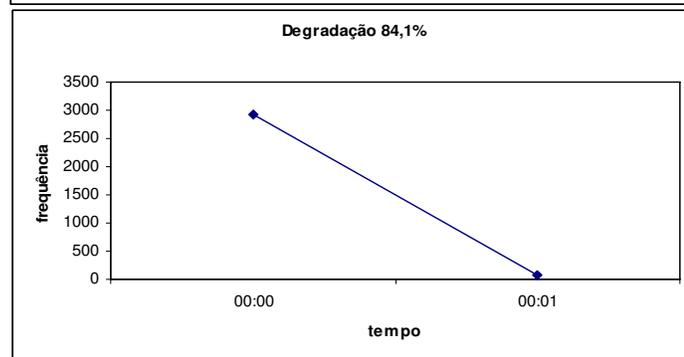
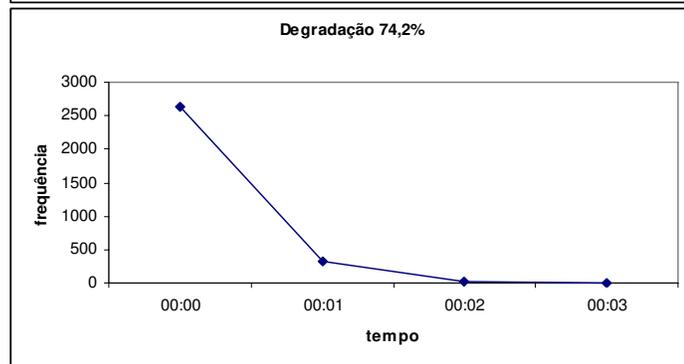
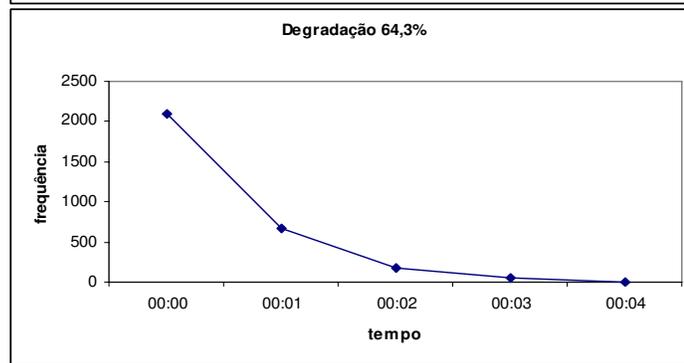
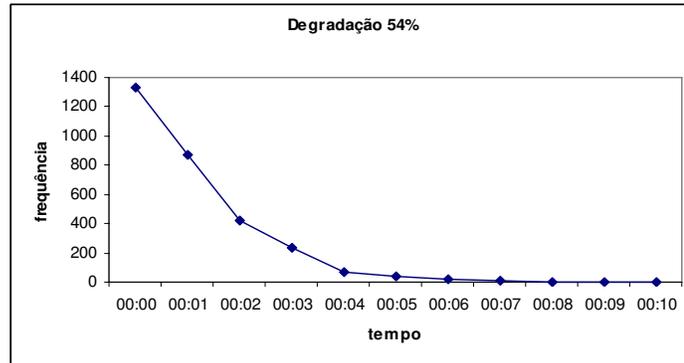
onde:

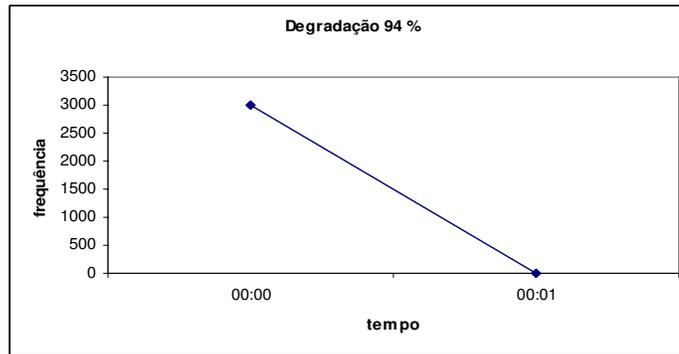
$$q = 1 - p. \quad (\text{b.58})$$

B.2 - Resultados

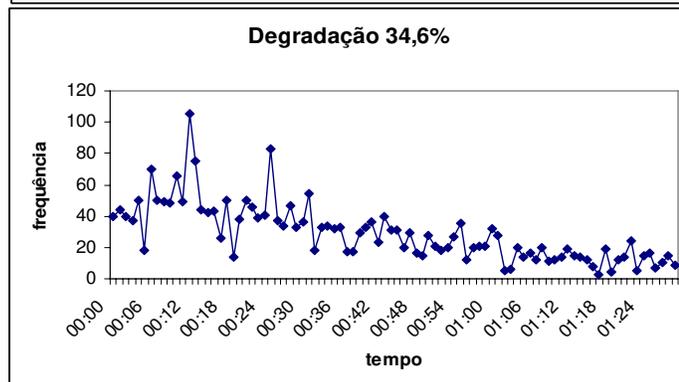
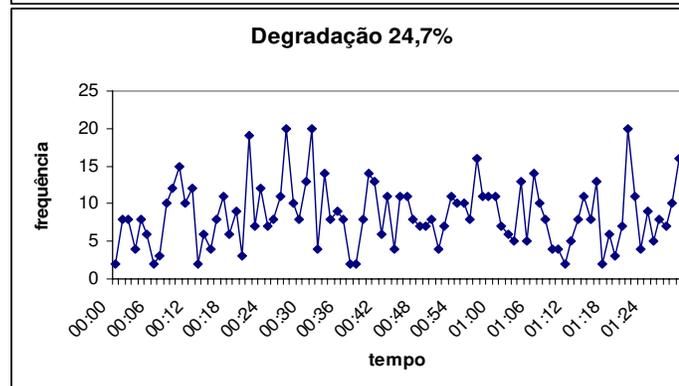
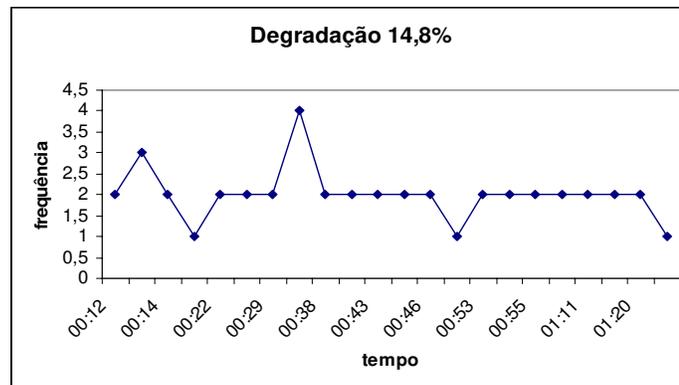
- **Nível de qualidade de 2%**

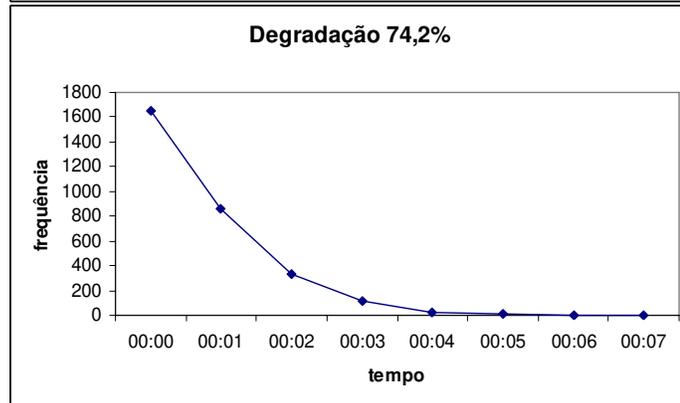
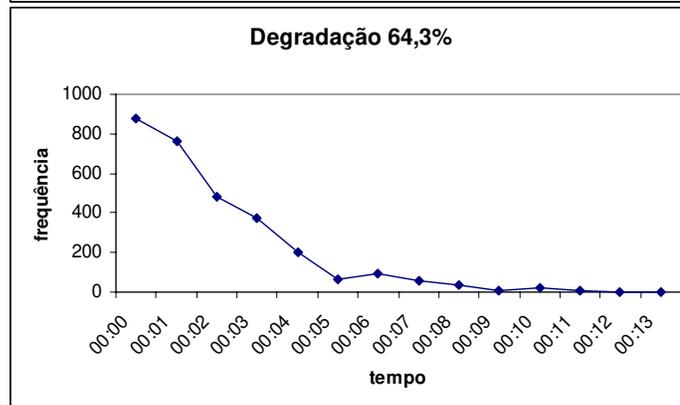
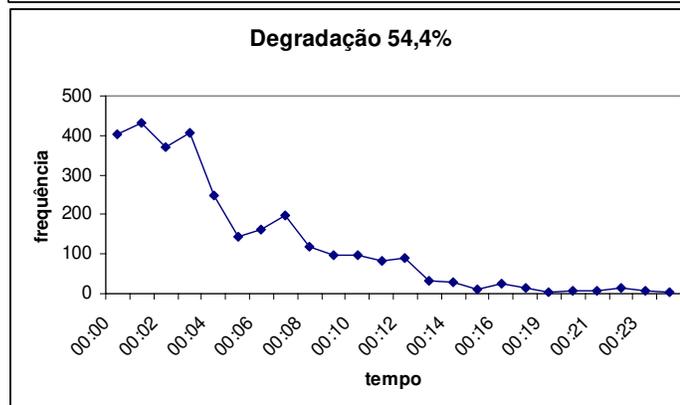
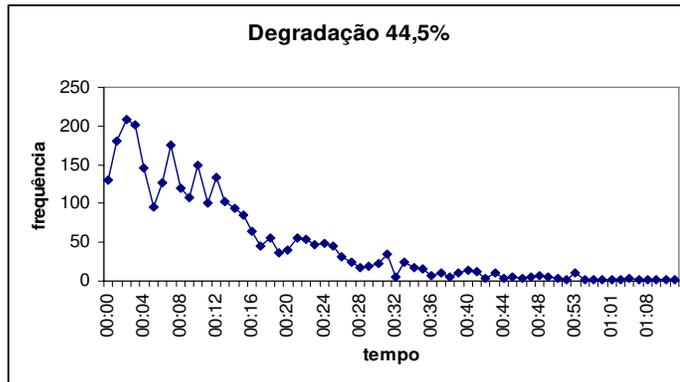


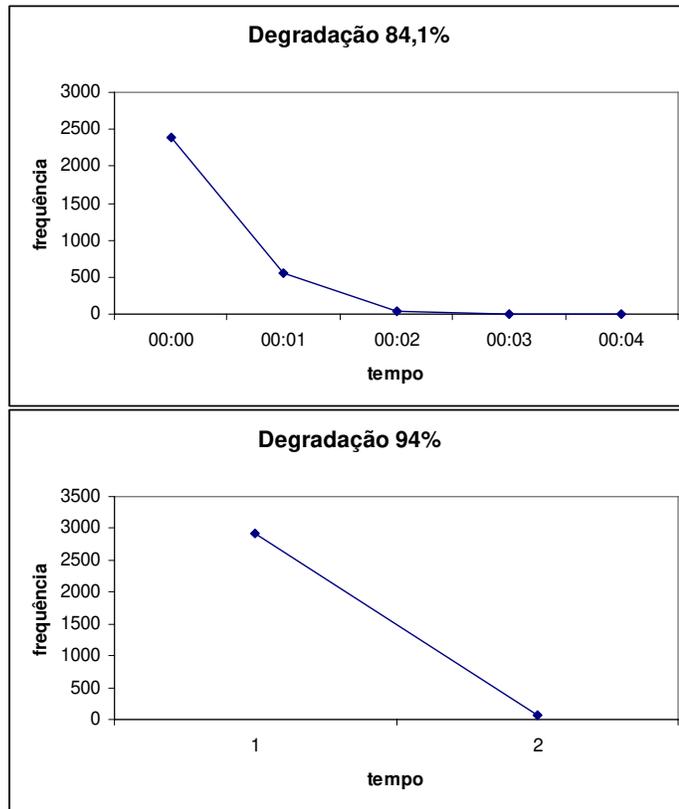




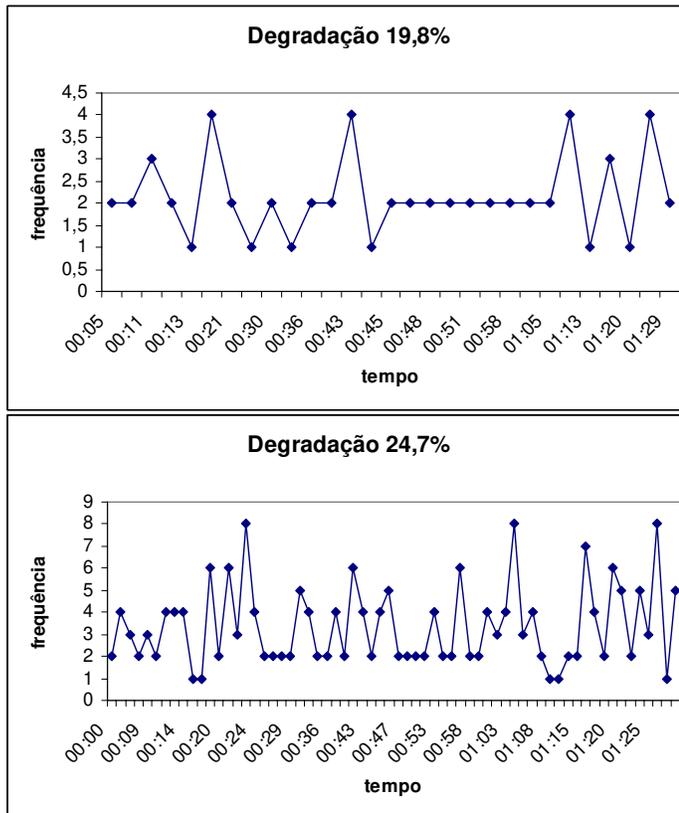
- **Nível de qualidade de 7%**

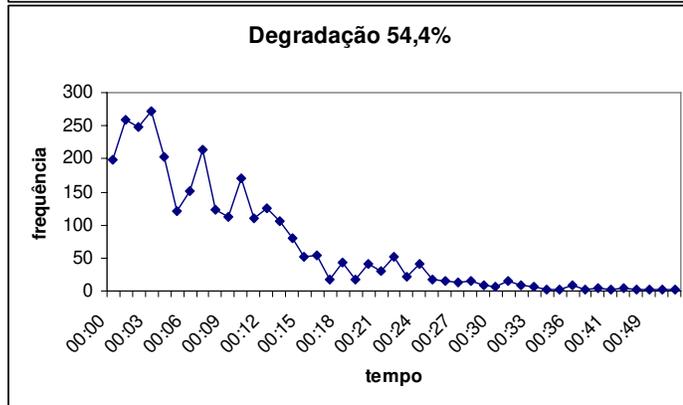
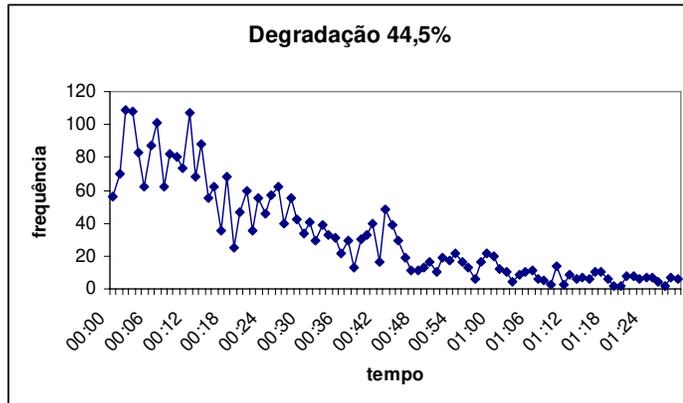
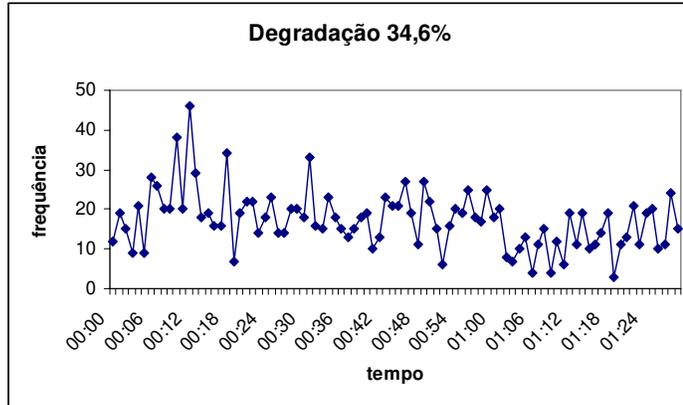


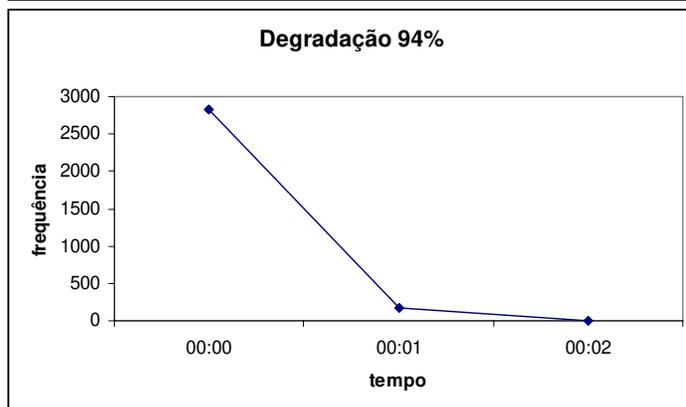
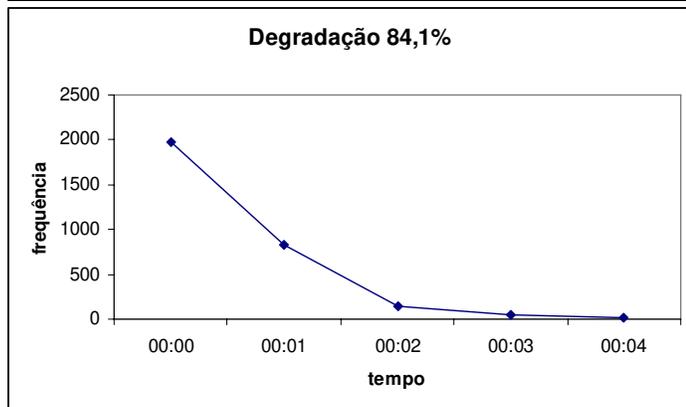
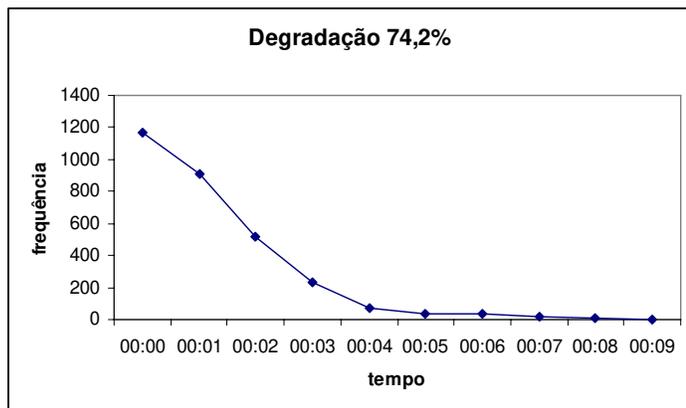
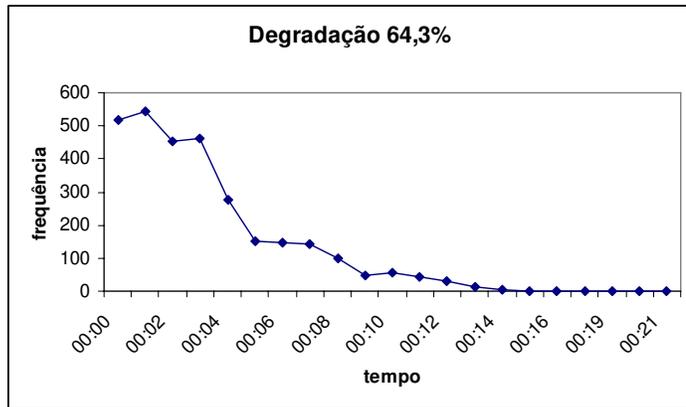




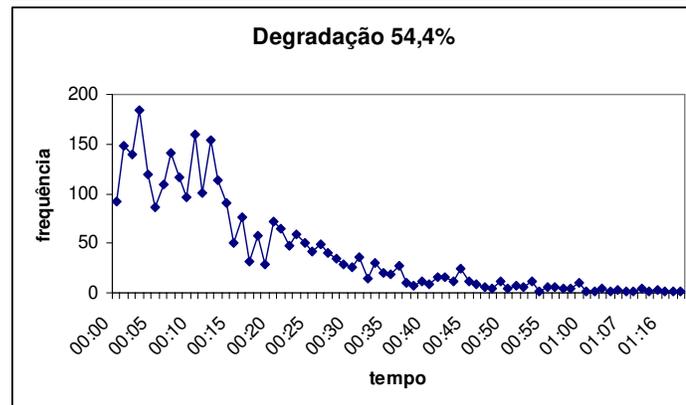
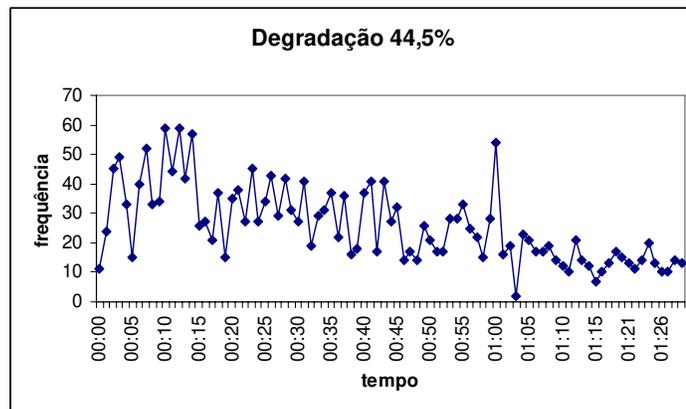
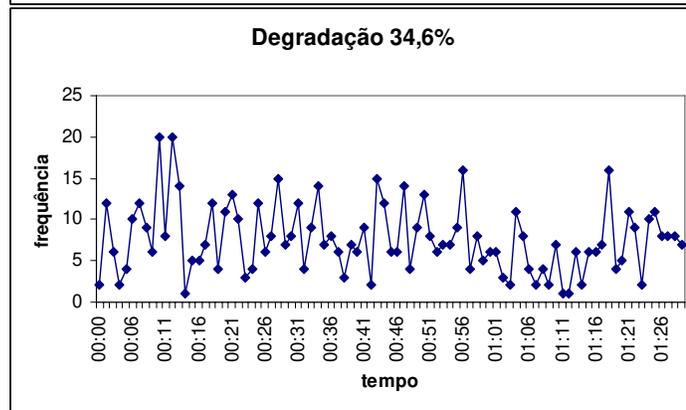
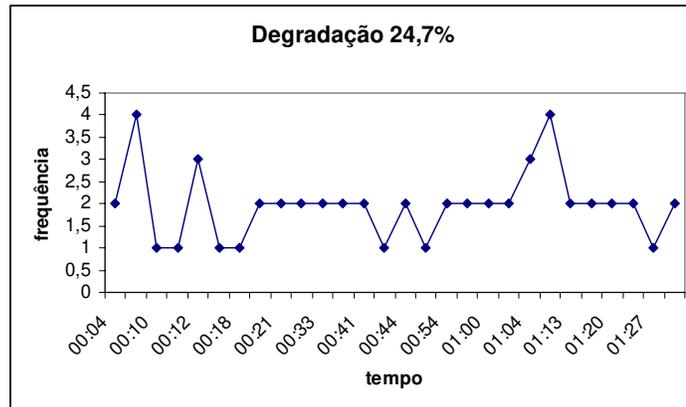
- **Nível de qualidade de 12%**

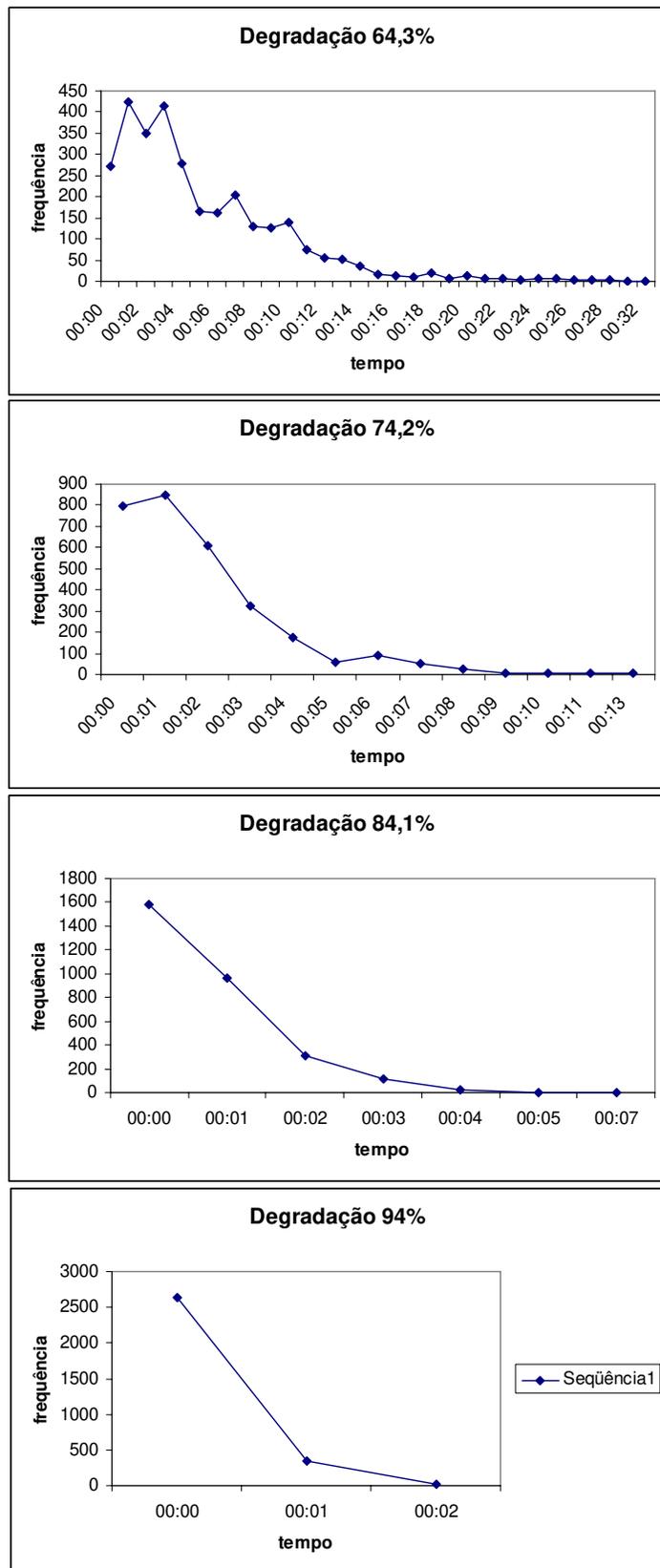




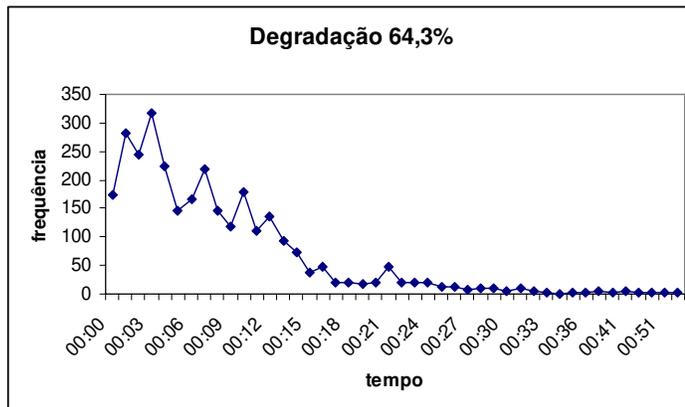
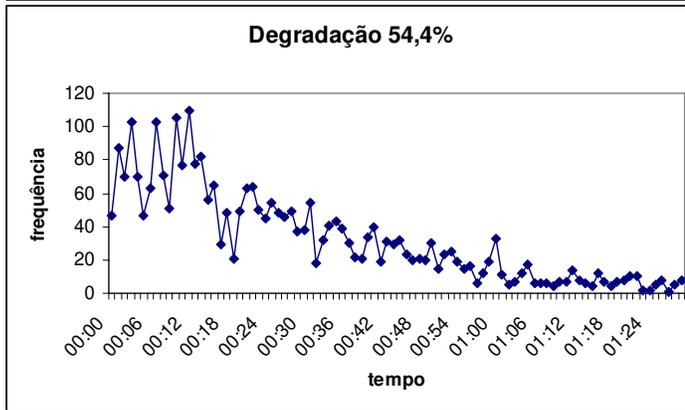
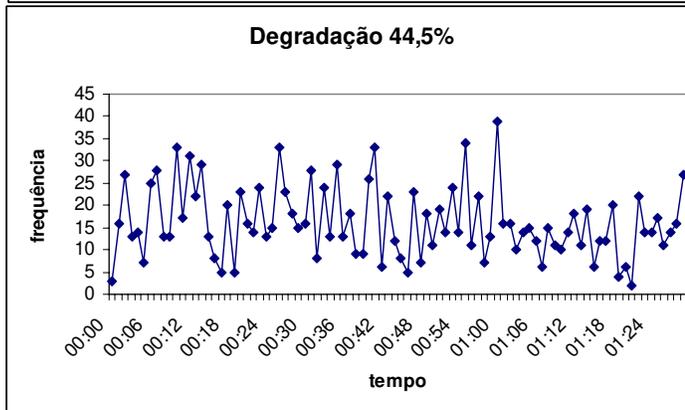
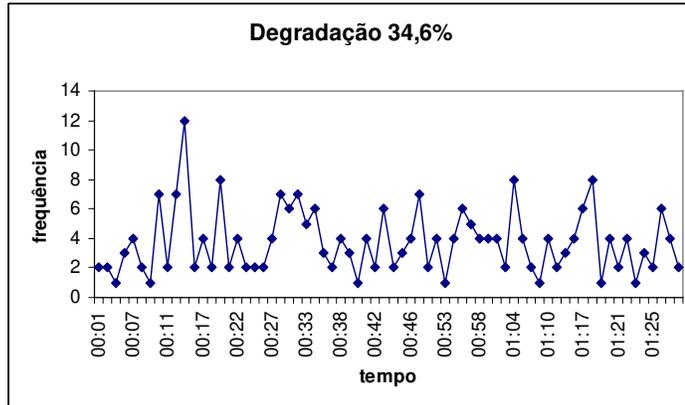


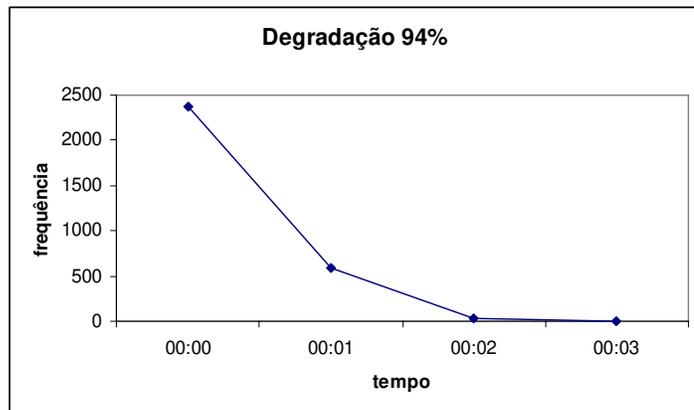
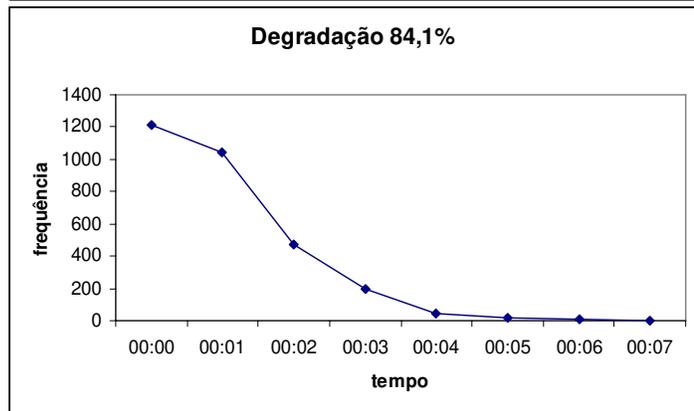
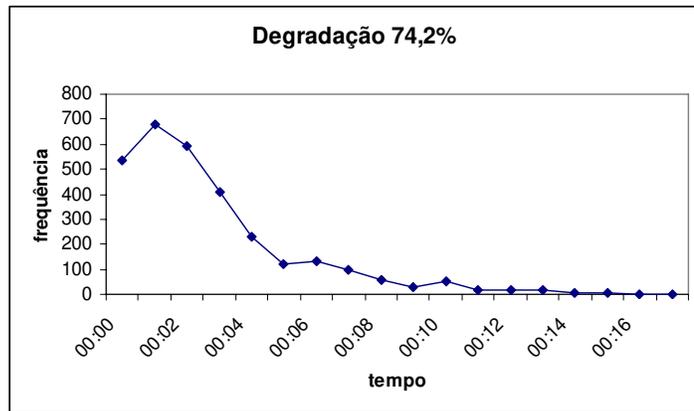
- **Nível de qualidade de 17%**



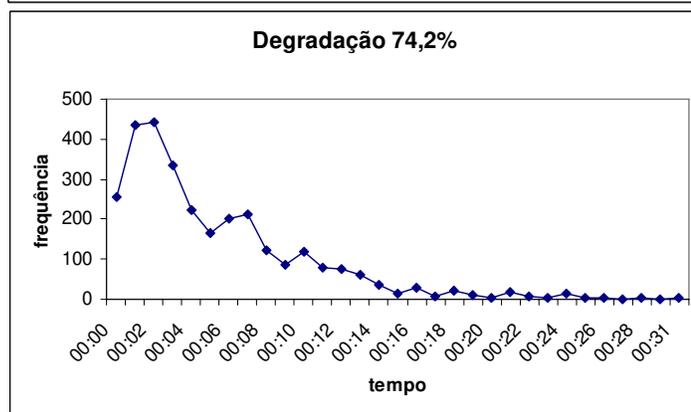
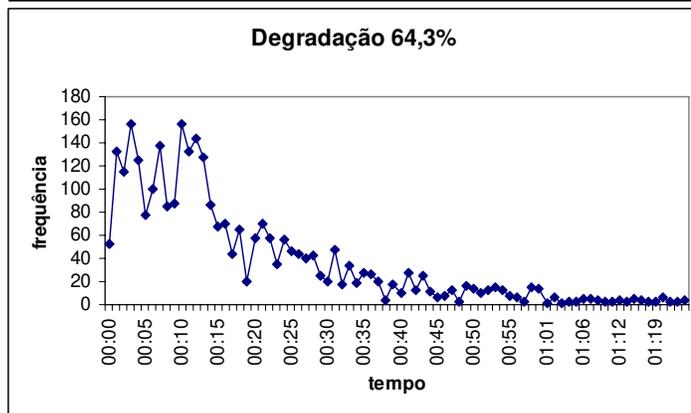
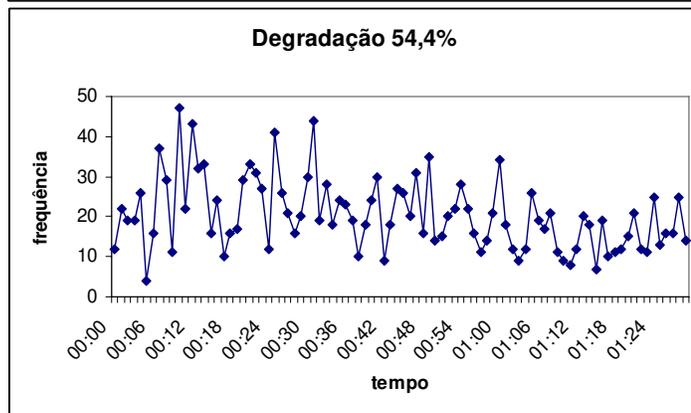
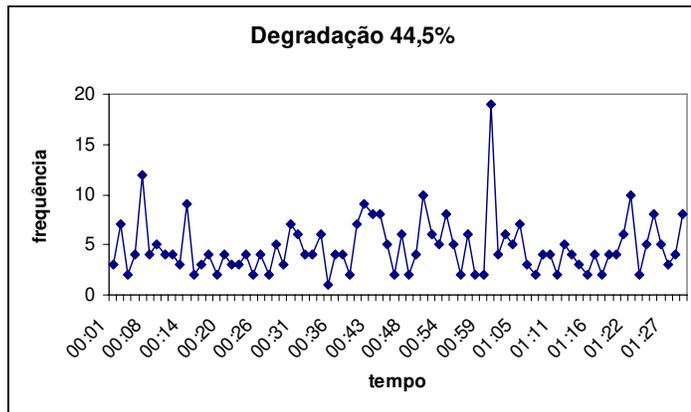


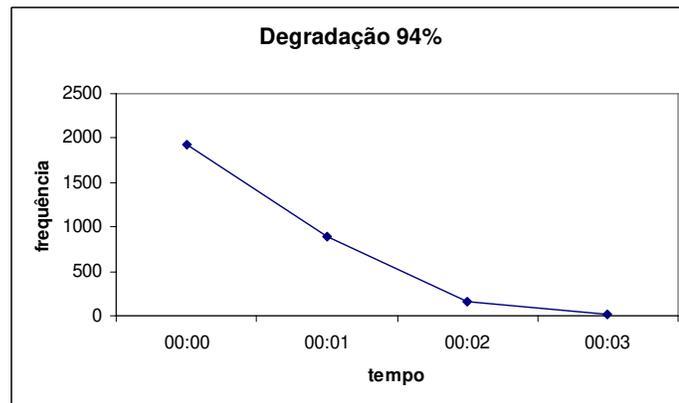
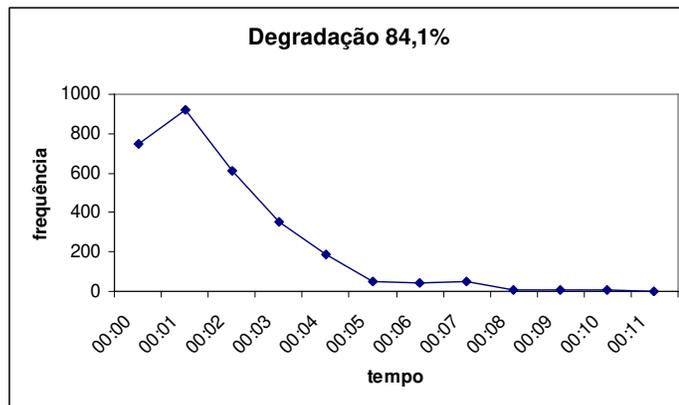
- **Nível de qualidade de 22%**





- **Nível de qualidade de 27%**





Apêndice C. Sinalização entre Centrais

A seguir são apresentados dois tipos de sinalização utilizados em telefonia, são eles: Sinalização MFC e Sinalização por Canal Comum nº 7. A compreensão das sinalizações conduz ao entendimento do que é possível acontecer com uma chamada telefônica.

C.1 - Sinalização MFC

Na Telefonia, há troca de informações (por meio de sinais) entre o assinante e a central e também entre centrais. Essa sinalização [98][99] é dividida em três grupos:

- Sinalização *acústica*;
- Sinalização de linha;
- Sinalização de registro.

C.1.1 - Sinalização Acústica

A finalidade da sinalização acústica é a de informar aos usuários do sistema telefônico as condições de estabelecimento de chamadas, ou seja, fornece informações sobre os estados da conexão. Compreende os seguintes tipos:

- Tom de discar;
- Corrente de toque;
- Tom de controle de chamada;
- Tom de ocupado;
- Tom de nível vago.

C.1.2 - Sinalização de Linha

Esse tipo de sinalização é responsável pelo estabelecimento da troca de informações entre as centrais, relacionados com o estágio da chamada e supervisão dos juntores. A sinalização de linha atua durante toda a chamada e os sinais podem ser para frente e para trás, compreendendo os seguintes sinais:

- Ocupação  (para frente);
- Atendimento  (para trás);

Sinalização entre Centrais

- Desligar para frente —————> (para frente);
- Desligar para trás —————> (para trás);
- Confirmação de desconexão —————> (para trás);
- Desconexão forçada —————> (para trás);
- Bloqueio —————> (para trás);
- Tarifação —————> (para trás);
- Re-chamada —————> (para frente).

C.1.3 - Sinalização de Registro

É a sinalização entre os registradores das centrais para a troca de informações a respeito da origem e do destino necessários para o estabelecimento da chamada. Os sinais de registro estão divididos em sinais para frente, grupos I e II, e sinais para trás, grupos A e B. A sinalização entre registradores pode ser classificada nos seguintes tipos:

- Decádica: Sinalização empregada nos sistemas tipo passo a passo;
- Impulsos Inversos: Sinalização empregada nos sistemas Rotary e AGF;
- Multifrequencial Compelida (MFC): Sinalização que emprega combinação de sinais de frequências. No Brasil, devido à existência de telefonia terrestre, bem como, por satélite, são adotados os seguintes tipos: Terrestre, R2(variantes 5B e 5C) e Satélite (5S).

No caso de sinalização terrestre é utilizada a Multifrequencial Compelida, ou seja, é necessário sinais para trás, como resposta aos sinais para frente, sem os quais a chamada será interrompida (até o limite de tempo permitido pelos órgãos de temporização dos equipamentos de comutação). Os sinais para frente são formados por combinações de duas frequências dentre seis, compreendidas entre 1380 a 1980 Hz (frequências altas) e os sinais para trás formados analogamente, com as frequências compreendidas entre 540 a 1140 Hz (frequências baixas).

A sinalização 5S não é compelida devido aos tempos de propagação, sendo por isso, conhecida como 5S-MF.

Veja, na Tabela 8.6 e Tabela 8.7, todos os sinais utilizados:

Sinalização entre Centrais

	GRUPO I	GRUPO II
1	Algarismo 1	Assinante Comum
2	Algarismo 2	Assinante com tarifação imediata
3	Algarismo 3	Equipamento de manutenção
4	Algarismo 4	Telefone público
5	Algarismo 5	Operadora
6	Algarismo 6	Equipamento de transmissão de dados
7	Algarismo 7	TPIU
8	Algarismo 8	Serviço internacional
9	Algarismo 9	Serviço internacional
10	Algarismo 0	Serviço internacional
11	Acesso a posição de operadora. Inserção de semi-supressor de eco na origem.	Reserva
12	Pedido recusado indicação de trânsito	Reserva
13	Acesso a equipamento de manutenção.	Reserva
14	Inserção de supressor de eco no destino.	Reserva
15	Fim de número	Reserva

Tabela 8.6 – Sinais para frente

	GRUPO A	GRUPO B
1	Enviar próximo algarismo	Assinante livre com tarifação
2	Enviar primeiro algarismo	Assinante ocupado
3	Preparar p/recepção do grupo B	Assinante com número mudado
4	Congestionamento	Congestionamento
5	Enviar categoria e identidade do assinante chamador.	Assinante livre sem tarifação
6	Reserva	Assinante livre com tarifação, colocar retenção sob assinante chamado.
7	Enviar o algarismo n-2	Número ou nível vago
8	Enviar o algarismo n-3	Assinante com defeito
9	Enviar o algarismo n-1	Reserva
10	Reserva	Reserva
11	Enviar indicação de trânsito internacional.	Serviço internacional
12	Serviço internacional	Serviço internacional
13	Serviço internacional	Serviço internacional
14	Serviço internacional	Serviço internacional
15	Serviço Internacional	Serviço internacional

Tabela 8.7 – Sinais para trás

No grupo B, sinais para trás, existe uma pequena descrição do que pode acontecer com uma chamada. A sinalização MFC possui pouca flexibilidade, dessa maneira, existem poucos sinais possíveis, porém como veremos, na sinalização por canal comum o leque de sinais possíveis é muito

Sinalização entre Centrais

maior. Apesar de estar apresentando a sinalização MFC, ela está em desuso, sendo substituída pelo Canal Comum.

Veja, na Figura 8.3, um exemplo de chamada entre dois assinantes, A (originador) e B(destino).

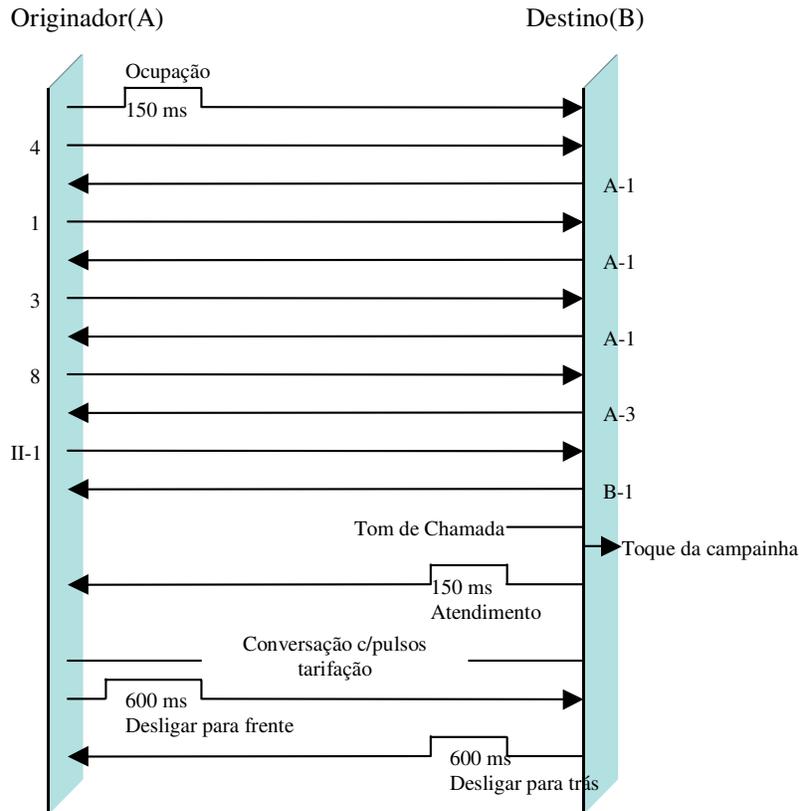


Figura 8.3 – Exemplo de troca de sinalização

Toda a sinalização, linha e registro, utilizam os mesmos canais que posteriormente irão transmitir a voz, o que permite dizer que existe uma associação entre a sinalização de linha, sinalização de registro e voz. Esse tipo de sinalização é definida como sinalização por canal associado. Pode-se ter a sinalização por canal associado em circuitos analógicos ou em circuitos digitais. Embora nos circuitos digitais a sinalização de linha seja transmitida no *time slot* 16 do *PCM-30*, considera-se esse caso como sinalização por canal associado, pois dentro do *time slot* 16 cada bit diz respeito a um canal. Os *time slots* 1 a 15 e 17 a 31, transmitem a sinalização de registro e voz.

C.2 - Sinalização por Canal Comum

Na sinalização por canal [100]-[102] comum podemos dizer que, numa central telefônica, a voz e parte da sinalização acústica são transmitidas por um canal e as informações correspondentes as sinalizações de linha e de registro passam por um canal reservado exclusivamente para esse fim, chamado de canal de sinalização. A grande diferença para o canal associado é que na sinalização por canal comum, o canal de sinalização transmite as informações dos sinais de linha e de registro de todos os juntores que interligam as duas centrais, enquanto no canal associado apenas de um juntor específico, por isso mesmo, é chamado de canal comum, é um canal comum a todos os circuitos de voz (juntors) entre duas centrais. Tudo isso se tornou possível graças à introdução de um novo tipo de central chamada CPA (Central por Programa Armazenado). Veja abaixo algumas vantagens da sinalização por canal comum:

- A simplificação nos juntores de entrada e saída que não são mais responsáveis pela sinalização de linha;
- Eliminação de emissores e receptores para sinalização MFC;
- Possibilidade de utilização de novos sinais, permitindo a introdução de novas funções como a manutenção à distância, a gerência de rede, etc;
- Possibilidade de utilização de canais bidirecionais, com aumento na eficiência, especialmente em rotas com pequenas quantidades de troncos.

Para dar suporte ao canal comum, devemos ter um ou mais enlaces de transmissão de dados entre as centrais, Figura 8.4. Esses enlaces podem ser juntores analógicos ligados por modem a uma velocidade de 4800 bps (bits por segundo). No entanto o mais comum é usar o *time slot* número 16, ficando os *time slots* restantes para voz (isso possibilita uma transmissão a 64 Kbps). Uma ligação por canal comum utilizando um canal de 64 Kbps é suficiente para a ligação de aproximadamente 1300 circuitos (juntors) entre duas centrais. Portanto, podemos ter 42 enlaces *PCM* (31 circuitos digitais em cada um, num total de 1302 circuitos) utilizando apenas um circuito para a transmissão de dados e ficando os outros 1301 para a transmissão de voz.

Em relação ao tempo gasto, a sinalização 5C gasta alguns segundos para enviar os sinais de linha e de registro para o estabelecimento da chamada, por outro lado a sinalização por canal comum gasta apenas frações de segundo, pois todos os dados de uma chamada vão de uma vez só a 64 Kbps. Apesar de serem apenas alguns segundos de diferença, isso em telefonia é significativo, trazendo um melhor aproveitamento da rede telefônica.

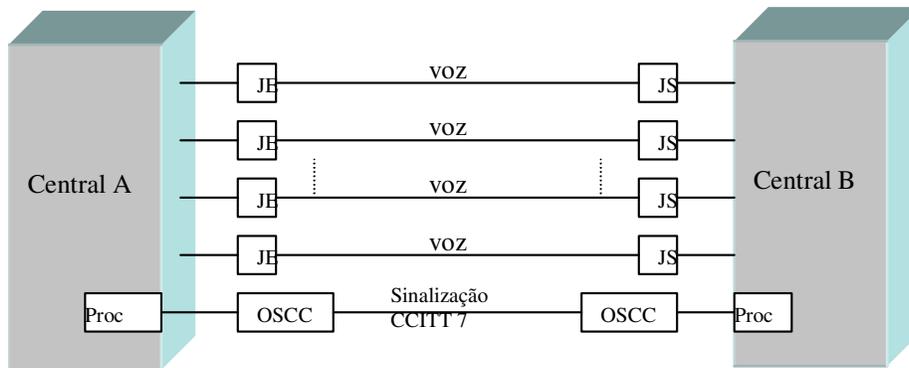


Figura 8.4 – Sinalização por Canal Comum

O CCITT (agora *ITU-T*) elaborou uma série de recomendações sobre um sistema de sinalização denominado “Sinalização por Canal Comum nº 6”. Esse sistema, no entanto, mostrou-se insuficientemente flexível nos anos que se seguiram, em face da difusão de uma tecnologia que favorecia cada vez mais a digitalização e integração das redes de telecomunicações. Desse modo foi desenvolvido um sistema de sinalização *ITU-T* nº 7 aplicável à interligação entre centrais telefônicas temporais, fazendo uso de um canal *PCM* com velocidade de 64 Kbps.

C.2.1 - “Causes” Válidos

Segue abaixo uma descrição dos “*Causes*” válidos pela norma. Não será feita nenhuma tradução, haja visto que o sentido sempre fica um pouco prejudicado. Esses dados foram retirados da norma *ITU-T Q767* [102][103]. Cada um destes “*Causes*” válidos é uma diferente maneira de uma chamada não ser completada. Observe que são muito mais respostas possíveis que na sinalização *MFC*.

a) *Normal class*

- Cause 1* – Unallocated (unassigned) number;
- Cause 2* – This cause is not applicable to the international interface;
- Cause 3* – No route to destination;
- Cause 4* – Send special information tone;
- Cause 5* – This cause is not applicable to the international interface;
- Cause 16* – Normal call clearing;
- Cause 17* – User busy;
- Cause 18* – No user responding;
- Cause 19* – No answer from user, user alerted;
- Cause 21* – Call rejected;
- Cause 22* – Number changed;
- Cause 27* – Destination out of order;

Cause 28 – Address incomplete;

Cause 29 – Facility rejected;

Cause 31 – Normal, unspecified.

b) Resource unavailable class

Cause 34 – No circuit available;

Cause 38 – Network out of order;

Cause 41 – Temporary failure;

Cause 42 – Switching equipment congestion;

Cause 44 – Requested circuit / channel not available;

Cause 47 – Resource unavailable, unspecified .

c) Service or option not available class

Cause 50 – This cause is not applicable to the international interface;

Cause 55 – Incoming calls barred within CUG;

Cause 57 – Bearer capability not authorized;

Cause 58 – Bearer capability not presently available;

Cause 63 – Service or option not available, unspecified.

d) Service or option not implemented class

Cause 65 – Bearer capability not implemented;

Cause 69 – This cause is not applicable to the international interface;

Cause 70 – This cause is not applicable to the international interface;

Cause 79 – Service or option not implemented, unspecified.

e) Invalid message (e.g. parameter out of range) class

Cause 87 – This cause is not applicable to the international interface;

Cause 88 – Incompatible destination;

Cause 91 – This cause is not applicable to the international interface;

Cause 95 – Invalid message, unspecified.

f) Protocol error (e.g. unknown message) class

Cause 97 – This cause is not applicable to the international interface;

Cause 99 – This cause is not applicable to the international interface;

Cause 102 – Recovery on timer expiry;

Cause 103 – This cause is not applicable to the international interface;

Cause 111 – Protocol error, unspecified.

g) Interworking class

Cause 127 – Interworking, unspecified.

Apêndice D. Tempos de Ocupação do Sistema

Todos as temporizações foram fruto da síntese dos dados de no mínimo duas centrais distintas.

D.1 - Tempo de Conversação de uma Chamada OK

O gráfico na Figura 8.5 fornece uma visão do comportamento das chamadas após serem completadas, ou seja, a função densidade de probabilidade para o tempo de conversação.

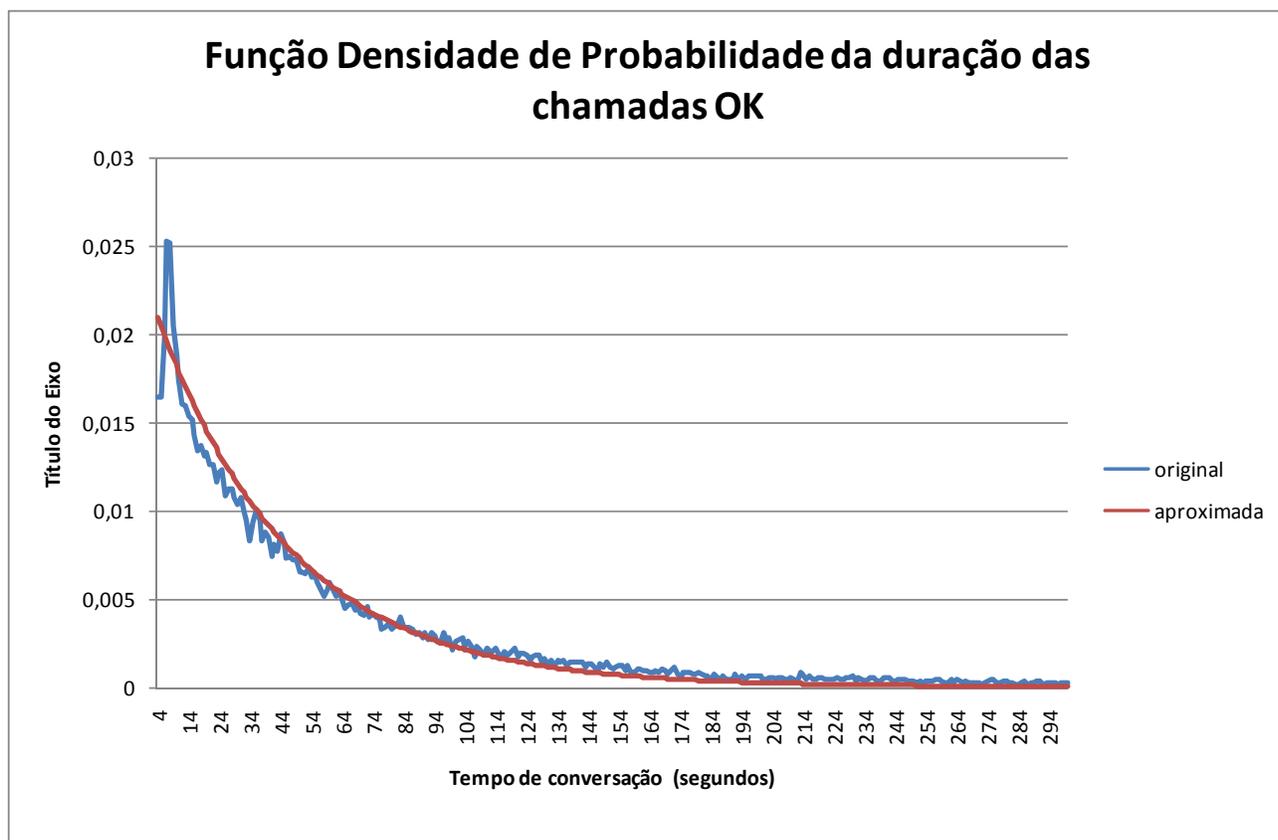


Figura 8.5 – Tempo de conversação para chamada OK

Fomos obrigados a adotar uma função diferente da gaussiana para caracterizar o comportamento dessas chamadas. Optamos então por encontrar uma função polinomial de grau 10 (dez) que melhor representasse esse evento. Devemos tomar cuidado com o fato de que quando integramos essa função num determinado intervalo, ela nos retornará a quantidade de eventos ocorridos nesse intervalo. Para obtermos esse resultado em termos percentuais basta dividirmos esse valor pelo total de ocorrências. O total de ocorrências é obtido quando integramos a função no intervalo de 3 a 400 segundos. Outra solução é normalizar os coeficientes da equação.

D.2 - Perfis de Ocupação do Sistema

Paralelo ao que já foi discutido é importante ter o conhecimento de como é a ocupação das centrais, ou melhor, qual é o perfil de ocupação em função dos horários, pois isso pode acarretar algumas particularidades no sistema. Caso o sistema esteja trabalhando com bilhetes num determinado horário, caracterizado como de alto tráfego e alta ocupação, pode-se esperar que os níveis de congestionamento do sistema sejam elevados, como também que haja uma degradação nos indicadores em função disso. Porém, a maior utilidade desse estudo está na possibilidade, uma vez que é conhecido o perfil de tráfego durante o dia, de prever/estimar o comportamento dos periféricos/elementos em todos os horários.

Foi feito um levantamento dos perfis de tráfego para diferentes centrais. Os resultados desses levantamentos são apresentados a seguir. Na

Figura 8.6 estão ilustrados os gráficos, normalizados entre 0 e 100%, e sobrepostos de maneira a prover uma visão conjunta das diferentes centrais. O eixo “y” representa a ocupação em função do máximo semanal e o eixo “x” o horário durante o dia. A legenda apresenta o nome das centrais ao lado o tamanho em número de habitantes da cidade cuja central se refere.

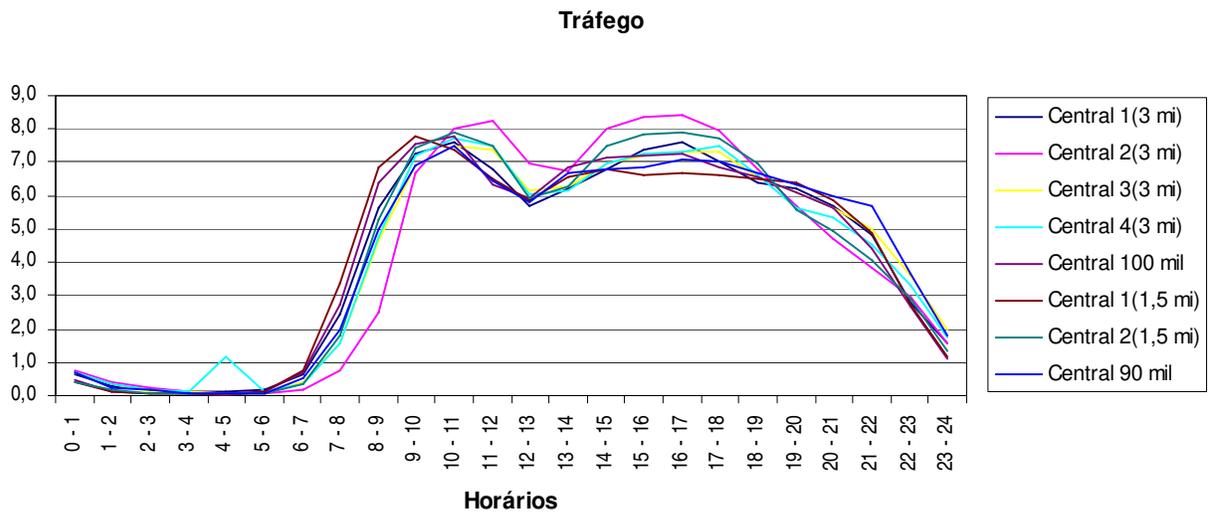


Figura 8.6 – Perfis normalizados

É possível constatar o comportamento das centrais é bastante similar. O horário de maior movimento pela manhã se dá entre o período das 10 – 11 horas, representando 7,67% de todo tráfego do dia, Tabela 8.8. Durante à tarde o período fica entre as 16 –17 horas, representando 7,45% do tráfego do dia.

Horário	%	% Acumulado
0 - 1	0,59	0,59
1 - 2	0,24	0,84
2 - 3	0,13	0,96
3 - 4	0,08	1,05
4 - 5	0,19	1,24
5 - 6	0,09	1,33
6 - 7	0,49	1,82
7 - 8	2,02	3,84
8 - 9	5,13	8,97
9 - 10	7,22	16,19
10 - 11	7,67	23,86
11 - 12	7,10	30,95
12 - 13	6,05	37,00
13 - 14	6,46	43,47
14 - 15	7,13	50,59
15 - 16	7,35	57,94
16 - 17	7,45	65,38
17 - 18	7,24	72,63
18 - 19	6,63	79,26
19 - 20	6,00	85,26
20 - 21	5,47	90,73
21 - 22	4,65	95,38
22 - 23	3,08	98,46
23 - 24	1,54	100,00

Tabela 8.8 – Tráfego por Horário