

**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE TECNOLOGIA**  
**DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**APLICAÇÃO DE REDES NEURAS À MIGRAÇÃO DE  
SERVIÇOS DE TELECOMUNICAÇÕES DE REDES  
LEGADAS PARA REDES IP**

**ANGELITA FLEIG KAPP**

**ORIENTADOR: JOÃO MELLO DA SILVA**

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA**

**PUBLICAÇÃO: PPGENE\_DM-044/07**

**BRASÍLIA/DF: DEZEMBRO - 2007**

UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

APLICAÇÃO DE REDES NEURAS À MIGRAÇÃO DE SERVIÇOS  
DE TELECOMUNICAÇÕES DE REDES LEGADAS PARA REDES IP

ANGELITA FLEIG KAPP

DISSERTAÇÃO DE MESTRADO PROFISSIONALIZANTE SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:



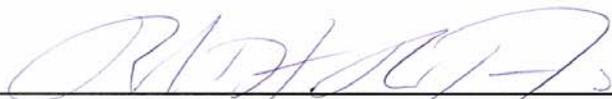
---

JOÃO MELLO DA SILVA, Dr., ENM/UNB  
(ORIENTADOR)



---

LUIZ FERNANDO RAMOS MOLINARO, Dr., ENE/UNB  
(EXAMINADOR INTERNO)



---

RUBEM DUTRA RIBEIRO FAGUNDES, Dr., PUC/RS  
(EXAMINADOR EXTERNO)

BRASÍLIA, 14 DE DEZEMBRO DE 2007.

## **FICHA CATALOGRÁFICA**

KAPP, ANGELITA FLEIG.

APLICAÇÃO DE REDES NEURAS À MIGRAÇÃO DE SERVIÇOS DE TELECOMUNICAÇÕES DE REDES LEGADAS PARA REDES IP. [Distrito Federal] 2007 xix, 136p., 210 x 297 mm (ENE/FT/UnB, Mestre, Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Elétrica.

1. Redes de Próxima Geração

2. Redes Neurais

I. ENE/FT/UnB

II. Título (série)

## **REFERÊNCIA BIBLIOGRÁFICA**

KAPP, ANGELITA FLEIG (2007). Aplicação de Redes Neurais à Migração de Serviços de Telecomunicações de Redes Legadas para Redes IP. Dissertação de Mestrado em Engenharia Elétrica, Publicação PPGENE.DM – 044/07, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 2007, 136p.

## ***CESSÃO DE DIREITOS***

AUTORA: Angelita Fleig Kapp

TÍTULO: Aplicação de Redes Neurais à Migração de Serviços de Telecomunicações de Redes Legadas para Redes IP.

GRAU: Mestre ANO: 2007

É concedida à Universidade de Brasília, permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito deste.

---

Angelita Fleig Kapp

CCSW2, Lote 1, Quadra 1, apartamento 203.

70680-250, Sudoeste, Brasília, DF.

## **DEDICATÓRIA**

Dedico este trabalho ao melhor Engenheiro que conheci na minha vida, por coincidência meu marido, Paulo Eduardo Kapp.

## **AGRADECIMENTOS.**

Agradeço aos meus orientadores, meus colegas de profissão e de empresa, João Mello e Marcelo Blanes, pela paciência e empurrões nas horas certas. À minha filha Ane, por entender minha ausência nas suas brincadeiras, meu marido Kapp pela co-orientação informal e indispensável.

Principalmente, a Deus, por estar comigo em mais esta etapa.

## **RESUMO**

### **APLICAÇÃO DE REDES NEURAS À MIGRAÇÃO DE SERVIÇOS DE TELECOMUNICAÇÕES DE REDES LEGADAS PARA REDES IP.**

**Autora: Angelita Fleig Kapp**

**Orientador: João Mello da Silva**

**Programa de Pós-Graduação em Engenharia Elétrica.**

**Brasília, dezembro de 2007.**

O planejamento e o dimensionamento de recursos, tanto humanos quanto materiais, devem estar suportados por dados que facilitem a sua execução e garantam uma assertividade mínima suficiente em seus resultados. Este trabalho trata deste e outros assuntos relacionados à predição do período – com a melhor assertividade possível – em que se deve migrar o transporte dos atuais serviços de conexão de redes, hoje suportados por redes totalmente Determinísticas, para redes Estatísticas.

Esta estimativa de período será feita a partir de uma rede Neural de aprendizado Hebbiano, baseada na Transformada de *Karhunen-Loève* onde dados de evolução de flexibilidade, custo, confiabilidade e eficiência são comparados para redes *IP (Internet Protocol)* e redes não *IP* e o cruzamento destas duas curvas estima uma data onde os citados custo, confiabilidade e eficiência serão melhores para as redes baseadas em *IP*.

## **ABSTRACT**

### **MIGRATION OF TELECOMMUNICATIONS SERVICES FROM LEGACY NETWORK TO IP NETWORKS USING A NEURAL NETWORK APPLICATION**

**Author: Angelita Fleig Kapp**

**Supervisor: João Mello da Silva**

**Programa de Pós-Graduação em Engenharia Elétrica.**

**Brasília, dezembro de 2007.**

The human and material resources planning must be supported by data that makes this task easier and they must be as correct as possible. This work, based in this directive, tries to predict a period that the service transportation will migrate fro TDM (Time Division Multiplex) to IP (Internet Protocol) based Networks.

This period estimate will use a Hebbian Learning Neural Network, based on Karhunen-Loève Transform where the flexibility evolution data, cost, trustworthiness and efficiency are compared for IP Networks and Non-IP Networks. The crossing of these two curves esteems a cited date where flexibility, cost, trustworthiness and efficiency will be better for IP based networks

## SUMÁRIO

DEDICATÓRIA.....	II
AGRADECIMENTOS.....	II
RESUMO .....	III
ABSTRACT .....	IV
SUMÁRIO.....	V
LISTA DE FIGURAS .....	IX
LISTA DE TABELAS .....	XII
LISTA DE TABELAS .....	XII
LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES .....	XIII
1 INTRODUÇÃO.....	1
1.1 MOTIVAÇÃO.....	3
1.2 ESTRUTURA DA DISSERTAÇÃO .....	4
2 CONCEITOS EMBASADORES SOBRE REDES .....	5
2.1 OS MODELOS DE REFERÊNCIA.....	5
2.1.1 O Modelo de Referência OSI .....	6
2.1.2 O Modelo de Referência TCP/IP ou IETF .....	7
2.2 REDES DE TELECOMUNICAÇÕES .....	10
2.2.1 Rede telefônica pública comutada - RTPC. ....	11
2.2.2 Rede de comunicações móveis ou telefonia celular. ....	11
2.2.3 Rede de dados – Rede Estatística. ....	11

2.2.3.1	Redes X.25 e <i>Frame Relay</i> .....	13
2.2.3.2	Frame Relay.....	13
2.2.3.3	ATM ( <i>Asynchronous Transfer Mode</i> ):.....	17
2.2.3.4	Rede Ethernet .....	23
2.2.3.5	Rede <i>IP</i> .....	26
2.2.3.6	Rede MPLS ( <i>Multiprotocol Label Switching</i> ).....	34
2.2.3.7	Formato do Datagrama .....	38
2.2.4	Rede Inteligente.....	39
2.2.5	Rede Canal Comum para Sistema de Sinalização Número 7.....	40
2.2.6	Rede Digital de Serviços Integrados (RDSI).....	40
2.2.7	RDSI-FL ou B-ISDN.....	41
2.2.8	Rede Determinística ou Rede Dedicada de Circuitos de Dados.....	41
2.2.9	Rede de Gerência ou TMN ( <i>Telecommunications Management Network</i> ).....	41
2.2.10	Rede de Acesso.....	42
2.2.11	Rede de Transporte.....	43
2.3	REDES NEURAIIS .....	43
2.3.1	Aprendizado nas Redes Neurais.....	49
2.3.1.1	Aprendizado Supervisionado.....	50
2.3.1.2	Aprendizado Reforçado.....	51
2.3.1.3	Aprendizado Não-Supervisionado.....	51
2.3.1.4	Aprendizado Competitivo .....	52

2.3.1.5	Aprendizado Hebbiano .....	52
2.3.2	Redes Neurais Artificiais para Decomposição de um Espaço Vetorial em Sub-Espaços .....	53
2.3.3	Mapas Auto-Organizados de Kohonem - SOM .....	56
3	ANÁLISE CONTEXTUAL DOS COEFICIENTES UTILIZADOS NO MODELO DE PREDIÇÃO .....	61
3.1	CRITÉRIOS DE COMPARAÇÃO .....	61
3.1.1	Qualidade .....	64
3.1.1.1	<i>Performance</i> ou Desempenho .....	65
3.1.1.2	Confiabilidade .....	79
3.1.2	Flexibilidade .....	80
3.1.2.1	Flexibilidade de Novos Produtos/serviços .....	81
3.1.2.2	Flexibilidade de <i>Mix</i> de Produto .....	82
3.1.2.3	Flexibilidade de volume .....	82
3.1.2.4	Flexibilidade de entrega .....	82
3.1.3	Desempenho de Entrega .....	83
3.1.4	Custo .....	84
3.1.4.1	Mapeamento dos Equipamentos de Dados .....	86
3.1.4.2	Mapeamento dos Equipamentos de Transmissão .....	86
3.1.4.3	Mapeamento da Última Milha – Par Metálico ou Fibra Óptica .....	86
3.2	JUSTIFICATIVA DO USO DA REDE NEURAL .....	87
4	EXPERIMENTO .....	91

4.1	OBTENÇÃO DOS DADOS.....	91
4.2	TRANSFORMAÇÃO DE DOMÍNIOS.....	97
4.3	LINEARIZAÇÃO E PREDIÇÃO.....	99
5	CONCLUSÃO.....	110
5.1	ANÁLISE DOS RESULTADOS E CONCLUSÃO.....	110
5.2	SUGESTÃO DE TRABALHOS FUTUROS.....	118
5.2.1	Melhoria da Predição.....	118
5.2.2	Subprodutos da base de dados trabalhada.....	118
5.2.3	Aplicações na Empresa.....	120
	REFERÊNCIAS BIBLIOGRÁFICAS.....	121
	APÊNDICES.....	124
	A ALGORITMO HEBBIANO GENERALIZADO PARA EXTRAÇÃO DOS COMPONENTES PRINCIPAIS – GHAPCA.....	125

## LISTA DE FIGURAS

Figura 2.1. O modelo de referência <i>OSI</i> [TANEMB2003] .....	6
Figura 2.2 – Modelo de Referência TCP/IP [TANEMB2003].....	8
Figura 2.3 – Topologia típica <i>Frame Relay</i> [BRITO2006].....	14
Figura 2.4 – Quadro do <i>Frame Relay</i> [TANEMB2003] .....	15
Figura 2.5 – Cabeçalho <i>Frame Relay</i> [BRITO 2006].....	16
Figura 2.6 – Cabeçalho da Célula ATM [ALBERTI1998] .....	20
Figura 2.7 – Formato quadro <i>Ethernet</i> segundo IEEE 802.3 [TANEMB2003].....	24
Figura 2.8 – Organização hierárquica de rede [TANEMB2003] .....	27
Figura 2.9 – Cabeçalho IPv4 (IP) [TANEMB2003] .....	28
Figura 2.10 – Cabeçalho do UDP [TANEMB2003] .....	30
Figura 2.11 – Cabeçalho do RTP [TANEMB2003].....	31
Figura 2.12 – Cabeçalho TCP [TANEMB2003].....	33
Figura 2.13 – Principais elementos MPLS [BRITO2006] .....	36
Figura 2.14 – Estrutura <i>Shim Header</i> [BRITO2006] .....	39
Figura 2.15. Ambiente Usuário-Operadora [KAPP2003] .....	42
Figura 2.16 – Representação do sistema Nervoso em Diagrama de Blocos [CAST2001] .	44
Figura 2.17 – Estrutura básica do neurônio [LABINFO2007].....	44
Figura 2.18 – Neurônio Artificial [BARR2002] .....	45
Figura 2.19 – Função Threshold ou Limiar [CAST2001].....	46
Figura 2.20 - Função Piecewise-linear ou Linear por Partes [CAST2001].....	47

Figura 2.21 – Função Sigmoidal [CAST2001] .....	47
Figura 2.22 Representação de um neurônio Artificial por Grafos [CAST2001].....	48
Figura 2.23 – Superfície de Erro apresentando um ponto mínimo local e um ponto mínimo global. [CAST2001] .....	50
Figura 2.24 Modelo de RNA 64 nós de entrada e K neurônios de saída [CAST2001].....	56
Figura 2.25 – Treliza de Neurônios [CAST2001] .....	57
Figura 2.26 – Vizinhança do Neurônio Vencedor [BENITE2003].....	59
Figura 3.1 – Fluxograma de Entrega de um Serviço de Comunicação de Dados [AUTORA] .....	64
Figura 3.2 – Planejamento e o controle da qualidade [SLACK1996].....	65
Figura 3.3 - Sobreposição genérica dos serviços [AUTORA] .....	69
Figura 3.4 – Amostra do arquivo de Bilhetes de Defeitos .....	80
Figura 3.5- Fatores que influem no desempenho da função de operação [SLACK1996]...	81
Figura 3.6 – Amostra do arquivo de tramitação das Ordens de Serviço .....	84
Figura 3.7 - Treliza de Neurônios e uma conexão sináptica reforçada entre dois neurônios [CAST2001] .....	89
Figura 4.1 – Histograma da quantidade de tarefas de configuração e provisionamento ...	93
Figura 4.2 – Diagrama do vetor de entrada do modelo de predição [AUTORA] .....	96
Figura 4.3 – Arquivo hpar.txt [CAST2001] .....	98
Figura 4.4 – Arquivo de saída do programa GHAPCA.....	99
Figura 4.5 – Tempo de configuração normalizado não IP dopado e não dopado .....	100
Figura 4.6 – Tempo de configuração normalizado IP dopado e não dopado .....	101
Figura 4.7 – Comparativo entre tempo de configuração dopado IP e não IP.....	102

Figura 4.8 – Tempo de configuração IP e não IP .....	104
Figura 4.9 – Custo de implantação e manutenção de serviços não IP dopado e não dopado .....	104
Figura 4.10 – Custo de implantação e manutenção de serviços IP dopado e não dopado	105
Figura 4.11 – Custo de implantação e manutenção de serviços IP e não IP dopados .....	105
Figura 4.12 – Custo de implantação e manutenção de serviços IP e não IP .....	106
Figura 4.13 – Quantidade de reclamações de serviços não IP dopado e não dopado .....	107
Figura 4.14 – Quantidade de reclamações de serviços IP dopado e não dopado .....	107
Figura 4.15 – Quantidade de reclamações de serviços não IP e IP dopado .....	108
Figura 4.16 – Quantidade de reclamações de serviços não IP e IP não dopado.....	108
Figura 5.1 - Comparativo entre tempo de configuração dopado IP e não IP .....	113
Figura 5.2 – Comparativo entre custo dopado IP e não IP .....	113
Figura 5.3 – Comparativo entre quantidade de reclamações dopadas IP e não IP .....	114
Figura 5.4 – Comparação de quantidade de reclamações IP e não IP não “dopada” .....	115
Figura 5.5 – Comparativo custo médio diário não IP e IP não “dopados” .....	116
Figura 5.6 – Comparativo dos tempos de configuração IP normalizados entre Filiais.....	119
Figura A.1 – RNA progressiva com uma única camada de neurônios lineares, a ser treinada pelo Aprendizado Hebbiano Generalizado.....	134

## **LISTA DE TABELAS**

Tabela 2.1. Classes de Serviços (I.362) [BRITO2006] .....	22
Tabela 3.1 – Codec vs Largura de Banda [CISCO2004a].....	67
Tabela 3.2 - Resumo Cálculo da Eficiência de Transmissão [AUTORA] .....	79
Tabela 5.1 - Comparação dos tempos de predição a partir de todos os coeficientes [AUTORA].....	115

## LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES

<i>3G</i>	<i>Third Generation Móbile System</i>
<i>AAL</i>	<i>ATM Adaptation Layer</i>
<i>AAL-PDU</i>	<i>ATM Adaptation Layer – Protocol Data Unit</i>
<i>AC/RO</i>	<i>Acre/Rondônia</i>
<i>ACK</i>	<i>Acknowledgement number</i>
<i>AS</i>	<i>Autonomous Systems</i>
<i>ATM</i>	<i>Asynchronous Transfer Mode</i>
<i>ATM-SDU</i>	<i>ATM – Service Data Unit</i>
<i>B</i>	<i>Bottom of stack flag</i>
<i>Bc</i>	<i>Burst committed</i>
<i>BD</i>	<i>Bilhete de Defeito</i>
<i>Be</i>	<i>Burst excided</i>
<i>BECN</i>	<i>Backwards Explicit Congestion Notification</i>
<i>BGP</i>	<i>Border Gateway Protocol</i>
<i>B-ISDN</i>	<i>Broadband – Integrated Services Digital Network</i>
<i>BNC</i>	<i>British Naval Connector</i>
<i>C/R</i>	<i>Command/Response</i>
<i>CAPEX</i>	<i>Capital expenditures</i>
<i>GHAPCA</i>	<i>Algoritmo Hebbiano Generalizado para Extração dos Componentes Principais</i>

<i>CIR</i>	<i>Committed information rate</i>
<i>CLP</i>	<i>Cell Lost Priority</i>
<i>CODEC</i>	Codificador Decodificador
<i>CPE</i>	<i>Customer Premises Equipment</i>
<i>CPU</i>	<i>Central Processor Unit</i>
<i>CRC</i>	<i>Cyclic redundancy check</i>
<i>CSMA-CD</i>	<i>Carrier Sense Multiple Access with Collision Detection</i>
<i>DE</i>	<i>Discart Eligibility</i>
<i>DF</i>	Distrito Federal
<i>DF</i>	<i>Don't Fragment</i>
<i>DIFFSERV</i>	<i>Differentiated Service</i>
<i>DLCI</i>	<i>Data Link Connection Identifier</i>
<i>DNS</i>	<i>Domain Name Server</i>
<i>DSE</i>	Decomposição em sub-espacos
<i>DWDM</i>	<i>Dense wavelength division multiplexing</i>
<i>EA</i>	<i>Address Field Extension</i>
<i>EXP</i>	<i>Experimental bit</i>
<i>FCS</i>	<i>Frame Check Sequence</i>
<i>FEC</i>	<i>Forward Equivalence Class</i>
<i>FECN</i>	<i>Forward Explicit Congestion Notification</i>
<i>FiFo</i>	<i>First in First out</i>

<i>FIN</i>	<i>Final Flag</i>
<i>FR</i>	<i>Frame Relay</i>
<i>FRAD</i>	<i>Frame Relay Access Device</i>
<i>FTP</i>	<i>File Transfer Protocol</i>
<i>GFC</i>	<i>Generic Flow Control</i>
<i>GO/TO</i>	<i>Goiás/Tocantins</i>
<i>HEC</i>	<i>Header Error Control</i>
<i>HTTP</i>	<i>HyperText Transfer Protocol</i>
<i>IEEE</i>	<i>Institute of Electrical and Electronic Engineers</i>
<i>IETF</i>	<i>Internet Engineering Task Force</i>
<i>IHL</i>	<i>Internet Header Length</i>
<i>IP</i>	<i>Internet Protocol</i>
<i>IP/MPLS</i>	<i>Internet Protocol/Multi Protocol Label Switch</i>
<i>IPv4</i>	<i>Internet Protocol version 4</i>
<i>IPv6</i>	<i>Internet Protocol version 6</i>
<i>ISDN</i>	<i>Integrated Services Digital Network</i>
<i>ISO</i>	<i>International Standardization Organization</i>
<i>ITU</i>	<i>Internacional Telecommunications Unit</i>
<i>ITU-T</i>	<i>The Telecommunication Standardization Sector of the ITU</i>
<i>KLT</i>	<i>Transformada de Karhunen-Loève</i>
<i>LSP</i>	<i>Label Switch Path</i>

<i>LSR</i>	<i>Label Switch Router</i>
<i>MF</i>	<i>More Fragment</i>
<i>MPLS</i>	<i>Multi Protocol Label Switching</i>
<i>MS</i>	Mato Grosso do Sul
<i>MSE</i>	<i>Mean Square Error</i>
<i>MT</i>	Mato Grosso
<i>MTU</i>	<i>Maximum Transfer Unit</i>
<i>NGN</i>	<i>Next Generation Network</i>
<i>NNI</i>	<i>Network Network Interface</i>
<i>OMS</i>	<i>Operational Management Services</i>
<i>OPEX</i>	<i>Operating expenditures</i>
<i>OS</i>	Ordem de Serviço
<i>OSI</i>	<i>Open System Interconnection</i>
<i>OSPF</i>	<i>Open Shortest Path First</i>
<i>PCA</i>	<i>Principal Components Analysis</i>
<i>PDH</i>	<i>Plesiochronous Digital Hierarchy</i>
<i>PR</i>	Paraná
<i>PSA</i>	<i>Principal Components Analysis</i>
<i>PSH</i>	<i>Push Flag</i>
<i>PSTN</i>	<i>Public Switched Telephone Network</i>
<i>PT</i>	<i>Payload Type</i>

<i>PVC</i>	<i>Permanent Virtual Circuits</i>
<i>QoS</i>	<i>Quality of Service</i>
RDSI	Rede Digital Serviços Integrados
RDSI-FL	Rede Digital Serviços Integrados – Faixa Larga
<i>RFC</i>	<i>Request for Comments</i>
RNA	<i>Redes Neurais Artificiais</i>
RS	Rio Grande do Sul
<i>RST</i>	<i>Reset Flag</i>
<i>RSVP</i>	<i>Resource Reservation Protocol</i>
<i>RTCP</i>	<i>Real Time Control Protocol</i>
<i>RTP</i>	<i>Real Time Transport Protocol</i>
<i>RTPC</i>	<i>Real Time Control Protocol</i>
<i>SAP</i>	<i>Service Access Point</i>
SC	Santa Catarina
<i>SDH</i>	<i>Synchronous Digital Hierarchy</i>
SID	Sistema Indicadores de Dados
<i>SMTP</i>	<i>Simple Mail Transfer Protocol</i>
<i>SOM</i>	<i>Self Organizing Map</i>
SP/RJ	São Paulo/Rio de Janeiro
SS7	<i>Signaling System #7</i>
<i>STEM</i>	<i>Strategic Telecoms Evaluation Model</i>

<i>SYN</i>	<i>Synchronization Flag</i>
<i>T1/E1</i>	É um padrão de linha telefônica digital americano/europeu
<i>TCP</i>	<i>Transmission Control Protocol</i>
<i>TCP/IP</i>	<i>Transmission Control Protocol/Internet Protocol</i>
<i>TDM</i>	<i>Time Division Multiplexing</i>
<i>TELNET</i>	<i>Telecommunications Network</i>
<i>TMN</i>	<i>Telecommunications Management Network</i>
<i>TOS</i>	<i>Type of Service</i>
<i>TTL</i>	<i>Time to live</i>
<i>UDP</i>	<i>User Datagram Protocol</i>
<i>UNI</i>	<i>User Network Interface</i>
<i>URG</i>	<i>Urgent pointer</i>
<i>VC</i>	<i>Virtual Circuit</i>
<i>VCC</i>	<i>Virtual Channel Connections</i>
<i>VCI</i>	<i>Virtual Connection Identifier</i>
<i>VOIP</i>	<i>Voice over IP</i>
<i>VPI</i>	<i>Virtual Path Identifier</i>
<i>VPLS</i>	<i>Virtual Private LAN Service</i>
<i>VPN</i>	<i>Virtual Private Network</i>
<i>WAN</i>	<i>Wide Área Network</i>
<i>Wi-Fi</i>	<i>Wireless Fidelity</i>

*WiMAX*

*Worldwide Interoperability for Microwave Access*

# 1 INTRODUÇÃO

O desenvolvimento de uma plataforma capaz de suportar uma grande gama de serviços de telecomunicações sempre foi o desejo de qualquer operadora. Uma rede que fosse capaz de transportar voz, vídeo e dados com a máxima eficiência de transmissão, flexibilidade e rapidez no provisionamento, confiável e de custo baixo foram e ainda são as principais premissas durante a escolha de redes a serem implantadas.

Já em meados da década de 1980 o *ITU (International Telecommunications Union)* criava as recomendações das RDSI (Redes Digitais de Serviços Integrados) e suas respectivas redes de transporte *B-ISDN (Broadband – Integrated Services Digital Network)*. Uma rede de transporte, especificamente, foi e ainda é bastante utilizada por ser a rede que mais se aproximou daquelas premissas originais de flexibilidade e capacidade de transportar qualquer serviço: as redes conhecidas como *ATM (Asynchronous Transfer Mode)*. O *ATM* se mostrou um padrão seguro de transporte e mais próximo da realidade das Operadoras de telefonia convencional, isto sem a necessidade de grandes alterações no modo de prestação e na tarifação destes serviços.

O *ATM*, apesar de o próprio nome trazer a característica da assincronalidade, se ajustava plenamente às premissas de uma rede totalmente convergente, pois possui camadas de adaptação ao transporte praticamente para qualquer serviço. Voz, dados e vídeo poderiam, finalmente, serem comutados e transportados sobre uma rede única, com gerenciamento e tarifação únicos, e os usuários teriam endereçamento próprio, exatamente como nas redes de comutação de voz.

Em 1988, o *ITU* criava a recomendação [*ITU-T.122*] – *Framework for Providing Additional Packet Mode Bearer Services*, esta homologada pela *ANSI* em 1993. Esta nova rede ficou conhecida no mercado como *Frame Relay* sendo, também, mais uma rede orientada à conexão, sem controle de erros e nenhum controle de fluxo [*TANENB2003*], mas com características complementares à antiga rede *X.25/75*. Ela fora desenvolvida para operar em faixas de banda entre 64Kbits/s e 2048Kbits/s, exatamente a lacuna entre o [*ITU-T.25*] e a rede *ATM*. Desta forma, as redes WAN pareciam estar completas para a camada dois do modelo *OSI (Open System Interconnection)*.

Aparentemente, por volta de 1990, segundo [TANENB2003], muitos especialistas acreditavam que o modelo *OSI* seria adotado maciçamente e muito de sua padronização foi criada antes de uma pesquisa mais elaborada. Paralelamente a isto, as universidades já pesquisavam e utilizavam amplamente o modelo de referência do *IETF* (*Internet Engineers Task Force*), mais conhecido como modelo *TCP/IP*, concorrente do *OSI*. Assim, ainda segundo [TANENB2003], a estrutura *OSI*, apesar de trazer a tona os conceitos de protocolos e interfaces de forma sem igual, não foi adotado pelos fabricantes por ter sido padronizado em um momento ruim, por adotar uma tecnologia ruim, por ser considerado um modelo mais político do que técnico e, finalmente, pela má fama de suas primeiras e lentíssimas implementações.

A convergência das redes prevista nos anos 80 para ser sobre o *ATM*, ocorreu no final dos anos 90, mas somente de forma parcial, uma vez que o *ATM* foi utilizado apenas para aplicações de transporte deixando a comutação de circuitos, prevista para substituir a *PSTN* (*Public Switch Telephone Network*), de fora. No Brasil, a utilização ainda ficou restrita às operadoras como *Core* das redes *Frame Relay* e transporte das redes *IP*, uma vez que poucos usuários tiveram condições de suportar os custos de acesso e interfaces.

O aumento da utilização de *VPNs* (*virtual private network*) no início dos anos 2000 como um recurso na interconexão de redes locais, a princípio e ainda sobre *Frame Relay*, *ATM* ou até *TDM* trouxe a tona uma possibilidade bastante promissora.

Como a conexão era sobre a camada 3, independente da camada 2 utilizada, por que não utilizar, ainda que basicamente, os princípios da convergência de redes sobre a camada 2 (*ATM*), mas agora contando com a flexibilidade da camada 3?

O desenvolvimento das *VPLSs* (*Virtual Private LANs Services*) sobre *MPLS* (*Multi-protocol Label Switching*) vem de encontro a esta tão desejada convergência.

Por outro lado, outras perguntas devem ser respondidas: quando os serviços de conexão entre redes locais de usuários, hoje utilizando redes *ATM* ou *FR*, devem migrar para a rede *MPLS*? Esta migração deve ocorrer mesmo que os usuários ainda desejem os antigos serviços *ATM/FR*? Esta rede deve permanecer restrita ao *Core*? Ou deve ser utilizada nas camadas de distribuição ou de acesso? Qual a capilaridade necessária? Quando transferir a voz da *PSTN* para esta nova rede *MPLS*?

Este trabalho, fazendo uso de uma Rede Neural de aprendizado Hebbiano e a luz dos conceitos de eficiência de transmissão, velocidade de provisionamento e custos de implantação dos serviços, procura responder a, pelo menos, uma das perguntas acima, mesmo que de forma indireta: quando se deve iniciar a migrar de forma expedita o transporte e interconexão de serviços para as redes *MPLS*.

## 1.1 MOTIVAÇÃO

A Operação de Comunicação de Dados em uma Operadora de Telecomunicações no Brasil é, usualmente, uma tarefa difícil. A pressão para a redução de custos, proteção de investimentos e, acima de tudo, o retorno do capital investido, levam o paradigma “*obter o máximo com o mínimo*” ao seu extremo. Assim, diante da pressão em se obter resultados imediatos, surge a carência de soluções criativas e inovadoras frente a desafios cada vez maiores a intervalos cada vez menores. Tal situação leva a se considerar que é tão importante a capacidade de remodelação desta operação quanto a qualidade do serviço prestado.

Para [SLACK1999] existem três papéis particularmente importantes para a função da área de Operações na empresa:

- Como apoio para a estratégia empresarial;
- Como implementadora da estratégia empresarial;
- Como impulsionadora da estratégia empresarial.

Ou seja, Operação promove sucesso ao longo do tempo, pois seu papel é o de “*fazer as coisas melhor*” bem como “*entregar serviços melhores*”.

Este trabalho remete ao meio acadêmico alguns dos problemas mais comuns de uma Operadora de Telecomunicações, ou seja, quando “abandonar” tecnologias, teoricamente obsoletas, mas que suportam serviços demandados pelo mercado e migrar de forma transparente ao usuário para novas tecnologias que podem suportar os mesmos serviços a um custo operacional menor.

Com este objetivo o trabalho se propõe a estimar uma data para migração dos serviços que, suportados por redes não IP, passariam a ser transportados por uma rede IP. Utilizou-se

para o experimento dados reais de velocidade de provisionamento, custo por serviço, cálculo da eficiência de transmissão de redes que suportam os serviços e como ferramenta computacional para relacionar os dados e colocá-los na base do tempo, foi utilizado um modelo de redes neurais auto-organizados.

Analisando a saída do sistema neural pretendeu-se estabelecer esta data estimada levando em conta as diferenças geográficas e de procedimento de operação entre as Filiais, a eficiência de transmissão da rede de transporte para cada serviço considerado, a flexibilidade na instalação deste mesmo serviço bem como sua confiabilidade.

## **1.2 ESTRUTURA DA DISSERTAÇÃO**

Este trabalho está dividido em cinco capítulos, incluindo a Introdução. No capítulo 2, onde foram abordados os conceitos embasadores para o desenvolvimento dos capítulos seguintes, procurou-se detalhar todos os tipos de Redes de Telecomunicações a luz das recomendações M3200 e *ITU-T*, além de um breve descritivo das Redes Neurais, principalmente a técnica de decomposição de um espaço vetorial em subespaços, visto que foi através de uma ferramenta que faz uso destes conceitos que se suportou o experimento e se comprovou a pesquisa.

O capítulo 3 apresenta a análise contextual, onde se descreve a situação e o ambiente do experimento e os coeficientes utilizados no cálculo da data estimada da convergência dos serviços. Neste capítulo se fez um paralelo entre os critérios ditos competitivos, segundo as prioridades que as empresas elegem para competir no mercado, como: qualidade, flexibilidade, confiabilidade e custo, com os mesmos critérios aplicados para as Redes de Telecomunicações.

O capítulo 4 explica o experimento realizado, bem como a origem dos dados utilizados e o capítulo 5 trata das conclusões do trabalho, bem como as sugestões de trabalhos futuros.

## **2 CONCEITOS EMBASADORES SOBRE REDES**

Este capítulo revisa a bibliografia necessária aos capítulos seguintes. Não se busca, nesta revisão, aprofundar o assunto mais do que o necessário ao entendimento do experimento e obtenção de dados para este.

Inicialmente, se descreve as redes de telecomunicações à luz de dois modelos: *Telecommunication Standardization Sector of International Telecommunication Union (ITU-T)* e o *Internet Engineering Task Force (IETF)*.

Cada rede será descrita com ênfase nas suas características funcionais e de “*interworking*” com as demais redes a esta associadas. Assim, quando for necessário calcular os dados de entrada para a rede Neural Hebbiana, tais como eficiência de transmissão e custo operacional, estes conceitos serão úteis.

Ainda cabe uma revisão dos conceitos de RNA (Rede Neural Artificial), a qual foi a técnica computacional utilizada para correlacionar os dados da pesquisa e a partir da qual foram tiradas as conclusões do trabalho.

Sendo assim, inicia-se esta revisão pelos Modelos de Referência.

### **2.1 OS MODELOS DE REFERÊNCIA**

Segundo [TANEMB2003] a estratificação em camadas dos Modelos de Interconexão tem por objetivo reduzir a complexidade de projeto de transferência de dados fazendo com que cada camada seja responsável por oferecer determinados serviços às camadas superiores.

Dentro desta filosofia, foram definidas duas importantes arquiteturas de rede, o modelo de referência *OSI (Open Systems Interconnection)* e o modelo de referência *TCP/IP (Transport Control Protocol/Internet Protocol)*.

Embora os protocolos associados ao modelo *OSI* raramente serem usados, o modelo em si é de fato bastante válido e as características descritas em cada camada ainda são muito relevantes. O modelo *TCP/IP* tem características opostas: o modelo não é muito utilizado, mas os protocolos têm uso generalizado.

### 2.1.1 O Modelo de Referência OSI

O modelo *OSI* é mostrado na Figura 2.1. Esse modelo se baseia em uma proposta desenvolvida pela *ISO (International Standards Organization)* como um primeiro passo em direção à padronização internacional dos protocolos empregados nas diversas camadas. O modelo é chamado *ISO OSI*, pois ele trata da interconexão de sistemas abertos à comunicação com outros sistemas.

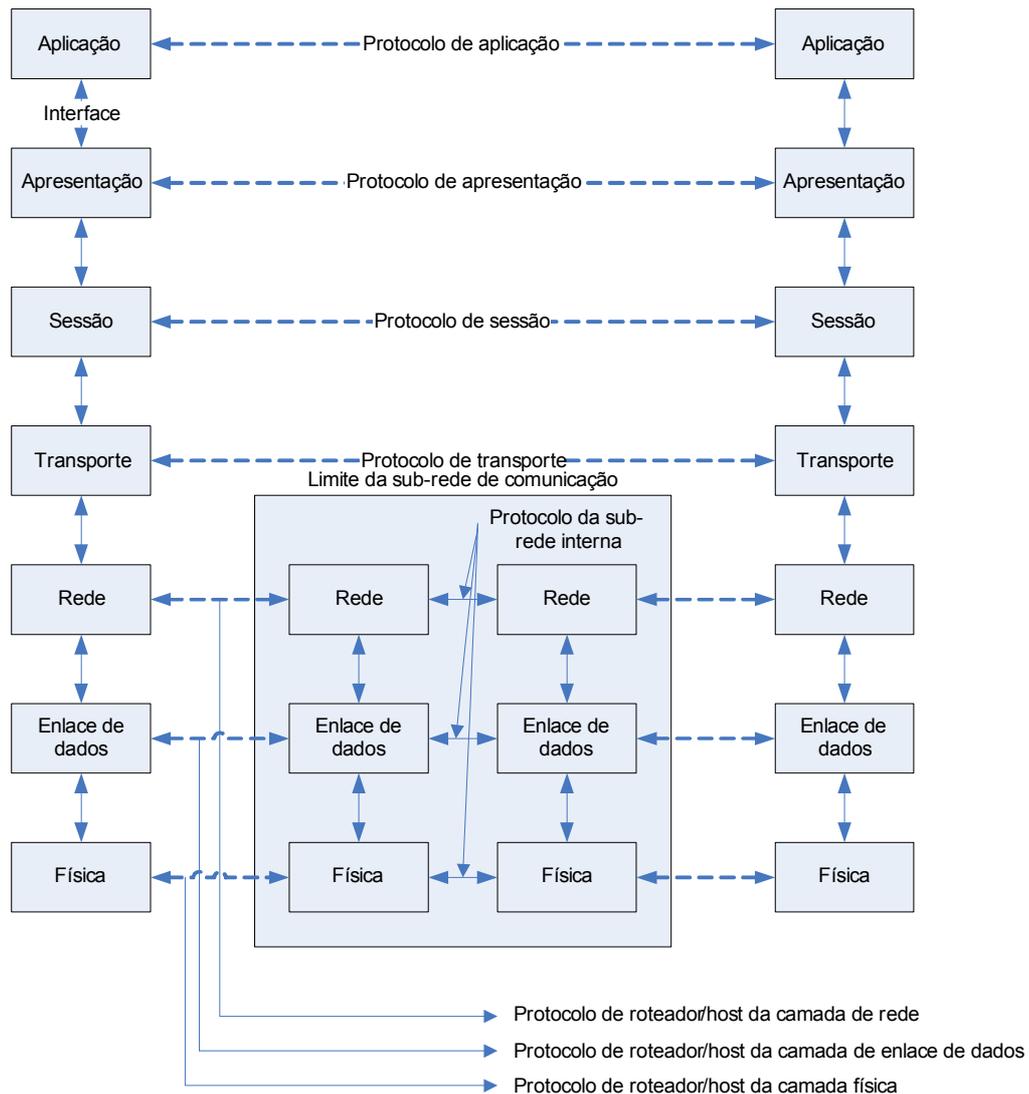


Figura 2.1. O modelo de referência *OSI* [TANEMB2003]

Este modelo tem sete camadas, a saber:

Camada Física – esta camada trata da transmissão de *bits* por um canal de comunicação. A preocupação é com a questão elétrica, mecânica e de sinalização, e com o meio físico de transmissão que se situa abaixo da camada física.

Camada de Enlace – é responsável por dividir os dados de entrada em quadros de dados e transmiti-los sequencialmente. Se for um serviço confiável, o receptor confirmará o recebimento através de um quadro de confirmação. Também é responsável por mecanismos de controle de fluxo, de modo que o transmissor seja informado do tamanho do *buffer* do receptor. Em alguns casos controla o acesso ao meio ou canal compartilhado, de modo a evitar o maior número possível de colisões.

Camada de Rede – é responsável pelo roteamento dos pacotes da rede além do controle do congestionamento. Adicionalmente, cuida da qualidade do serviço (retardo, tempo em trânsito, instabilidade, etc.) e dos ajustes para interoperabilidade com outras redes.

Camada de Transporte – é a verdadeira camada fim a fim. A diferença é que nas camadas 1-3 os protocolos são trocados entre cada máquina e seu vizinho imediato, e não entre as máquinas de origem e destino, as quais podem estar separadas por muitos roteadores. Esta camada também determina que tipo de serviço será fornecido à camada de sessão e, em última análise, aos usuários da rede. A função básica é aceitar os dados da camada de cima, dividi-los em unidades menores e repassá-los para a camada de rede, assegurando que todos os fragmentos chegarão corretos no destino.

Camada de Sessão – permite que usuários de diferentes máquinas estabeleçam sessões entre eles. Oferece diversos serviços como: controle de diálogo (controla quem deve transmitir), gerenciamento de *token* (impede que duas partes gerem operações críticas ao mesmo tempo) e sincronização (permite continuar a transmissão de onde parou em se tratando de falha).

Camada de Apresentação – Está relacionada a sintaxe e a semântica das informações transmitidas.

Camada de Aplicação – contém uma série de protocolos úteis para o usuário, por exemplo *HTTP (Hyper Text Transfer Protocol)*, que constitui a base para a *World Wide Web*. Entre outros, para transferência de arquivos, correio eletrônico e transmissão de notícias pela rede.

### **2.1.2 O Modelo de Referência TCP/IP ou IETF**

[TANEMB2003] afirma que não é errado dizer que a origem da criação desse modelo de referência se deu através da ARPANET. Esta era uma rede de pesquisa patrocinada pelo

Departamento de Defesa dos Estados Unidos e pouco a pouco centenas de universidades e repartições públicas foram a ela conectadas usando linhas telefônicas dedicadas. Quando surgiram as redes via rádio e satélite, os usuários passaram a ter problemas com os protocolos existentes, forçando a criação de uma nova arquitetura de referência.

Diante das premissas do Departamento de Defesa dos Estados Unidos, foi necessário criar uma arquitetura flexível, capaz de se adaptar a aplicações com requisitos divergentes como, por exemplo, a transferência de arquivos e a transmissão de voz em tempo real.

O modelo *TCP/IP* é mostrado na Figura 2.2.

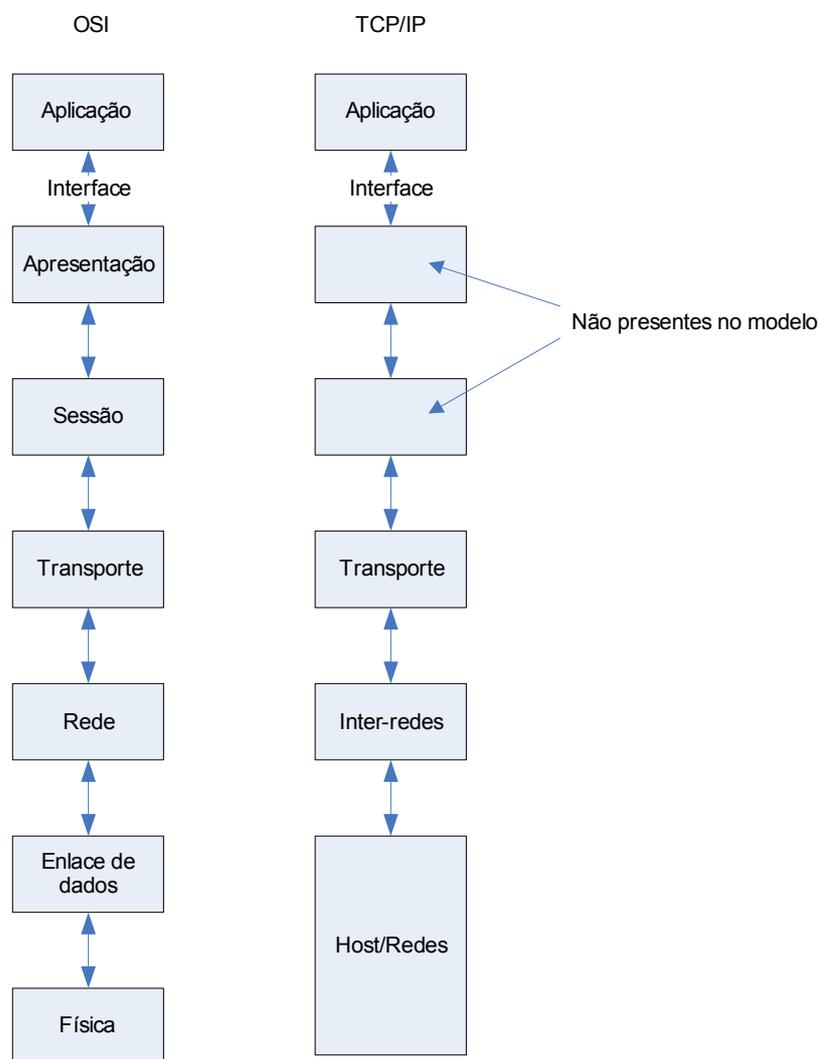


Figura 2.2 – Modelo de Referência TCP/IP [TANEMB2003]

Este modelo tem quatro camadas, a saber:

Camada Inter-Redes – esta camada define um formato de pacote oficial e um protocolo chamado *IP (Internet Protocol)*. A tarefa desta camada é agregar pacotes *IP* onde eles são necessários. Nesta camada é essencial o roteamento de pacotes e o controle de congestionamento.

Camada de Transporte – esta camada tem a função de manter a conversação entre *host* de origem e destino. Dois protocolos foram definidos para isso: o *TCP (Transmission Control Protocol)*, que é orientado à conexão, confiável, que permite a entrega sem erros, fragmentando as mensagens antes de entregá-las à camada de inter-redes. Além disso, faz controle de fluxo, impedindo que um transmissor rápido sobrecarregue um receptor lento. O segundo protocolo é o *UDP (User Datagram Protocol)*, que é não orientado à conexão, destinado a aplicações que não exigem controle de fluxo, nem seqüência de mensagens, portanto, não é confiável. É usado em aplicações onde a entrega imediata é mais importante que a entrega precisa.

Camada de Aplicação – esta camada contém todos os protocolos de nível mais alto. Dentre eles estão o protocolo de terminal virtual (*TELNET- Telecommunication Network*), o protocolo de transferência de arquivos (*FTP - File Transfer Protocol*) e o protocolo de correio eletrônico (*SMTP - Simple Mail Transfer Protocol*), além do *DNS (Domain Name System)*, que mapeia os nomes dos *hosts* para seus respectivos endereços de rede e o *http - HyperText Transfer Protocol*, usado para buscar páginas na *World Wide Web*, entre outros.

Camada *Host/Rede* – o modelo não especifica muito bem o que está abaixo da camada inter-redes, exceto o fato de que o *host* tem de se conectar a rede utilizando algum protocolo para que seja possível enviar pacotes *IP*. Esse protocolo geralmente não é definido na literatura.

Os modelos *OSI* e *TCP/IP* têm muito em comum. Os dois se baseiam no conceito de camadas independentes. No entanto, também possuem muitas diferenças, a saber:

O modelo *OSI* tem três conceitos fundamentais: serviços, interfaces e protocolos.

A definição do serviço informa o que a camada faz, e não a forma como as entidades acima dela a acessam ou como a camada funciona. A interface de uma camada informa como os processos acima dela podem acessá-la. Ou seja, quais são os parâmetros e os resultados a serem esperados. Não revelando o funcionamento interno da camada.

Os protocolos, por sua vez, são de responsabilidade da camada que os utiliza.

Originalmente, o modelo *TCP/IP* não distinguia com clareza a diferença entre serviço, interface e protocolo. Por esta razão os protocolos no modelo *OSI* são mais bem encapsulados que os do modelo *TCP/IP* e podem ser substituídos com relativa facilidade, conforme as mudanças de tecnologia.

O modelo *OSI* foi concebido antes de os protocolos correspondentes terem sido criados. Isso significa que o modelo não foi desenvolvido com base em um determinado conjunto de protocolos, o que o deixou mais flexível e genérico.

Com modelo *TCP/IP*, ocorreu exatamente contrário: como os protocolos foram definidos primeiro, o modelo foi criado como uma descrição desses protocolos.

Outra diferença está na comunicação. No modelo *OSI*, a camada de rede pode ser sem conexão ou orientada à conexão; no entanto, a camada de transporte só aceita comunicação orientada à conexão, o que de fato é mais importante já que a camada de rede é visível para o usuário. O modelo *TCP/IP* só tem um modo de comunicação na camada de rede (sem conexão), mas aceita ambos na camada de transporte, deixando a escolha por conta do usuário.

## **2.2 REDES DE TELECOMUNICAÇÕES**

As empresas de telecomunicações no mundo e particularmente no Brasil, até bem pouco tempo atrás, priorizavam critérios técnicos para seu planejamento de redes e serviços. Os usuários de pequeno, médio e grande porte normalmente adaptavam seus sistemas internos às capacidades e serviços disponíveis prestados pelas operadoras.

Hoje, ocorre uma inversão nesta linha de planejamento e as empresas dividiram suas redes de acordo com os serviços fornecidos. Assim, o que passou a determinar o planejamento das redes foram as necessidades e a demanda por serviços do mercado. Em outras palavras, as operadoras tiveram de se adequar, tanto em custos quanto em tipos de serviços visando adaptar-se às necessidades dos usuários.

Assim, se definirão, agora, os principais tipos de redes empregadas, de um modo geral, nas operadoras de telecomunicações. O objetivo é situar o leitor sobre a localização e a necessidade específica de cada rede, onde ela se localiza nas camadas *OSI* ou *Internet*

vistas acima. Uma abordagem estratificada das redes facilitará o entendimento que algumas redes são “clientes” de outras. Será introduzido o conceito de rede fornecedora de serviço.

A seguir, descrevem-se as principais redes fornecedoras de serviços ou que a elas suportam, nas operadoras de telecomunicações.

### **2.2.1 Rede telefônica pública comutada - RTPC.**

Como o nome já diz, esta rede é a prestadora do serviço original das operadoras telefônicas, é formada pelas centrais de comutação, sejam elas simples centrais locais ou trânsitos nacionais e internacionais de grande capacidade. É uma rede totalmente hierarquizada, ou seja, uma central encaminha as chamadas conforme rotas de encaminhamento pré-definidas e numeração discada pelo usuário. Atualmente, além dos serviços de voz, com o crescimento nos anos 90 dos serviços de dados comutados (*Dial up*), hoje estabilizado, estas redes estão fornecendo outros serviços que não os originalmente especificados. Pontos de consulta e venda por cartão de crédito (que utilizam o protocolo assíncrono X.28 do *ITU*, um subprotocolo do X.25) é um exemplo de como antigas redes de pacotes ainda utilizam a rede comutada como acesso a serviços.

### **2.2.2 Rede de comunicações móveis ou telefonia celular.**

A Rede de Comunicações Móveis ou Telefonia Celular é uma rede bastante similar, no que se refere a voz, à rede telefônica pública comutada – RTPC, pois é formada por centrais de comutação bastante semelhantes às utilizadas por aquela rede. A diferença é o conjunto de sistemas para tratamento da mobilidade e a inexistência de uma rede física de acesso, uma vez que esta é formada por Estações Rádio Base. Esta rede está fora do escopo deste trabalho.

### **2.2.3 Rede de dados – Rede Estatística.**

A rede de pacotes pode ser considerada uma rede fornecedora de serviços e é formada por um conjunto de elementos comutadores (*switches*) e elementos roteadores (*routers*) necessários a prover o serviço de dados estatísticos de camada dois e três (modelos *OSI* e *TCP/IP*). Necessitam de outras redes para prover seus serviços, sendo estes elementos interconectados entre si via Rede de Transporte e com os usuários via Rede de Acesso.

Esta rede recebe ainda outras denominações como rede de pacotes, rede estatística ou os nomes específicos dos protocolos utilizados: Rede X.25, Rede *Frame Relay*, Rede *IP*. As recomendações do ITU-T regulamentam algumas delas: X.25, X.32, X.75, X.28, X.29, *ISDN* (X.31) e *Frame Relay* enquanto as *RFCs* do *IETF* padronizam o *IP*.

Neste ponto é importante distinguir dois tipos de redes Estatísticas quanto ao tráfego dos dados ou pacotes, as **redes Orientadas a conexão** e as **redes Não orientadas a conexão** [TANEMB2003].

O conceito de redes não orientadas a conexão, surgiu por volta dos anos 50, com o surgimento da ARPANET, como já citado no item 2.1.2, uma rede de dados financiada pelo Departamento de Defesa Americano às Universidades daquele país. Nesta época, no auge da Guerra Fria, o governo norte-americano almejava por uma rede de controle e comando capaz de sobreviver a uma guerra nuclear. Isto é, uma rede que continuasse a funcionar mesmo depois de vários ataques diretos que destruíssem numerosos roteadores e linhas de transmissão. Ou seja, uma rede altamente tolerante a falhas, mas não necessariamente comercial. Esta abordagem resultou em um projeto sem conexões, no qual cada pacote é roteado de modo independente. Em consequência disso, se alguns roteadores ficarem inativos não haverá nenhum dano desde que o sistema possa se reconfigurar dinamicamente, para que os pacotes subseqüentes possam encontrar alguma rota até o destino, mesmo diferente daquela que foi usada por pacotes anteriores.

Enquanto isto, no sistema de telefonia, um visitante deve discar o número do telefone chamado e esperar por uma conexão antes de falar ou enviar os dados. Essa configuração de conexão estabelece um “caminho” pelo sistema que é mantido até a chamada ser encerrada. Todos os pacotes seguem a mesma rota. A vantagem deste sistema sobre o outro é, basicamente: Qualidade de Serviço e Faturamento.

Ou seja, configurando um recurso com antecedência, a sub-rede pode reservar espaço em *buffer* e capacidade de *CPU* do roteador. Deste modo, uma vez estabelecida, a conexão receberá um bom serviço. Em uma rede sem conexão, se um número excessivo de pacotes chegarem ao mesmo tempo no mesmo roteador, este será sufocado e pode perder pacotes. Eventualmente, o transmissor percebendo isto, poderá retransmitir os pacotes, mas a qualidade do serviço será instável e inadequada para algumas aplicações como áudio e vídeo que requerem tempo real.

A segunda razão para as empresas de telefonia apreciarem o serviço orientado a conexões é o fato de cobrarem pelo tempo de conexão. Se uma conexão for estabelecida antes dos dados serem enviados, terá início o tempo de cobrança. Se não há conexão, não é possível nenhuma cobrança.

As redes de transporte *IP* também vêm buscando fornecer melhor qualidade de serviço, principalmente para os serviços de áudio e vídeo. Este assunto será mais bem descrito no item 2.2.3.5.

#### 2.2.3.1 Redes X.25 e *Frame Relay*:

[TANEMB2003] afirma que a rede X.25 foi a primeira rede pública de dados. Para ser usada, primeiro um computador estabelecia uma conexão com o computador remoto, isto é, fazia uma chamada telefônica. Esta conexão recebia um número de conexão que seria usado em pacotes de transferência de dados. Os pacotes de dados eram muito simples, consistindo em um cabeçalho de 3 *bytes* e até 128 *bytes* de dados. O cabeçalho tinha um número de conexão de 12 *bits*, um número de seqüência de pacote, um número de confirmação e alguns bits variados.

Na década de 80, as redes X.25 foram substituídas em grande parte por redes *Frame Relay*. A essência do *Frame Relay* é o fato de ser uma rede orientada a conexões sem controle de erros e nenhum controle de fluxo. Os pacotes são entregues em ordem, quando entregues. Sua aplicação mais importante é a interconexão de *LAN's* (*Local Area Network*) instaladas em vários escritórios de uma empresa. O *Frame Relay* ainda é extensivamente usado nas Operadoras e será melhor detalhado.

#### 2.2.3.2 *Frame Relay*

Segundo [BRITO2006] o objetivo da rede *Frame Relay* era o de prover o usuário com uma rede privativa virtual (*VPN*) capaz de suportar aplicações que requeriam altas taxas de transmissão (T1/E1). No protocolo *Frame Relay* uma série de funções, que existiam nos protocolos anteriores, são minimizados ou eliminados.

A topologia típica de uma rede *Frame Relay* possui o dispositivo do usuário conectado através de um roteador ou algum outro dispositivo de acesso, que pode ser chamado de *FRAD* (*Frame Relay Access Device*), que implementa o protocolo *UNI* (*User Network*

*Interface*). Esta, apesar de não ser uma abordagem obrigatória, provê operação transparente para o usuário.

Na Figura 2.3, está ilustrada a topologia típica de uma rede *Frame Relay*.

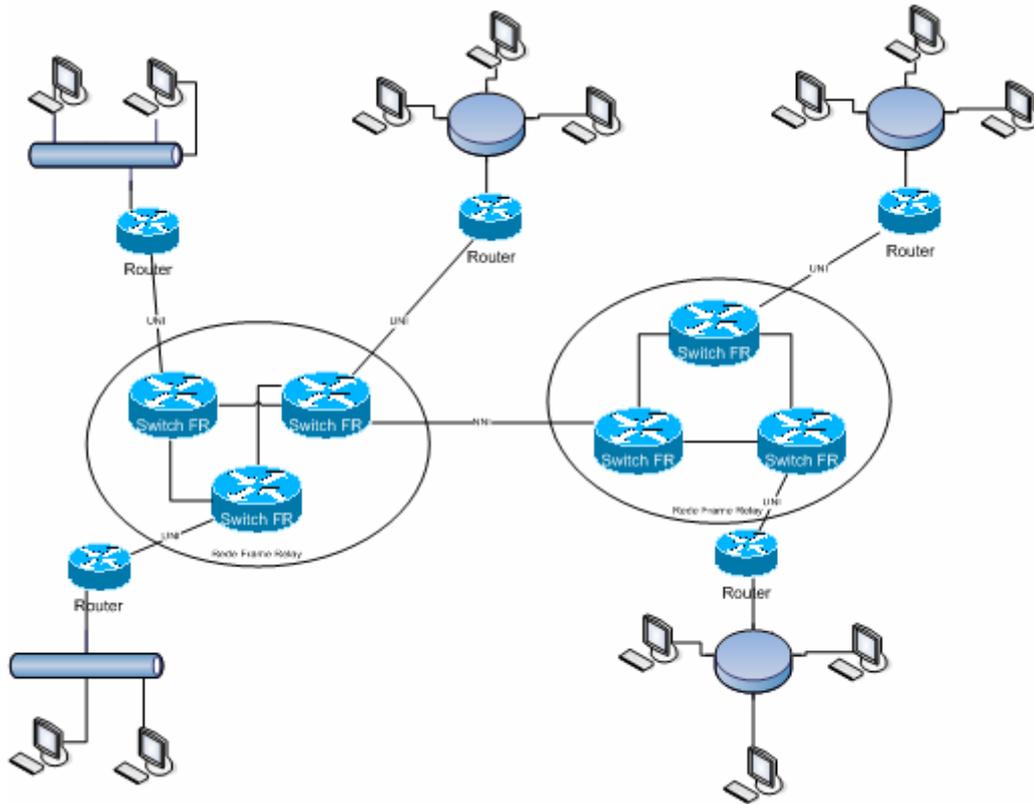


Figura 2.3 – Topologia típica *Frame Relay* [BRITO2006]

A rede *Frame Relay* é concebida para eliminar ou combinar muitas operações residentes nas camadas 2 e 3 do modelo *OSI*. Esta abordagem resulta em aumento de vazão e diminuição de atraso (se comparado com o X.25).

#### 2.2.3.2.1 *Quadro do Frame Relay*

Como pode ser visto na Figura 2.4 as funções de endereçamento e controle estão em um mesmo campo.

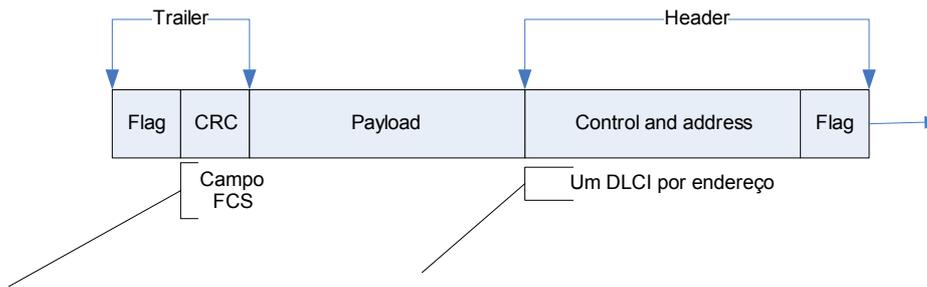


Figura 2.4 – Quadro do *Frame Relay* [TANEMB2003]

Os campos *flag* do início e final têm a função de delimitação e alinhamento do quadro. O campo de controle contém o *DLCI* (*Data Link Connection Identifier*) do quadro. Este identificador é responsável pela multiplexagem e demultiplexagem de circuitos virtuais, uma vez que os campos de *payload* podem conter tráfego de múltiplos usuários, logo, cada campo de *payload* é identificado com um único *DLCI*.

O sistema deve verificar se o quadro possui um número inteiro de octetos. O campo *FCS* (*Frame Check Sequence*) dá ao sistema a capacidade de detectar erros de transmissão, erros de formatação e outros erros operacionais.

O *Frame Relay* usa um identificador de conexão de enlace de dados que permite identificar o usuário de destino. O *DLCI* é um número de 10 *bits* semelhante ao número de circuito virtual de um protocolo de camada de rede. Os 10 *bits* podem ser expandidos com a opção de endereçamento estendido.

Os *DLCIs* são pré-mapeados para o nó de destino. O roteador encaminha o tráfego com base no valor do *DLCI*, que serve como parâmetro de entrada para consulta a sua tabela de roteamento.

A operação dentro da rede pode ser através do *DLCI* ou de outra forma qualquer (por exemplo, operação não-orientada à conexão).

A função de conversão de endereços é fundamental para a operação da rede, embora não definida pelo *Frame Relay*.

Uma técnica de controle de erro - *CRC* (*Cicle Redundance Check*) - é utilizada para detectar erros nos quadros transmitidos.

A verificação pode ser feita nó-a-nó ou apenas no último nó da rede, a opção é do implementador. Um quadro errado é descartado pela rede, mas o transmissor não é notificado do erro.

#### 2.2.3.2.2 Cabeçalho Frame Relay

O cabeçalho do quadro *Frame Relay* possui os seguintes campos, conforme Figura 2.5 abaixo:

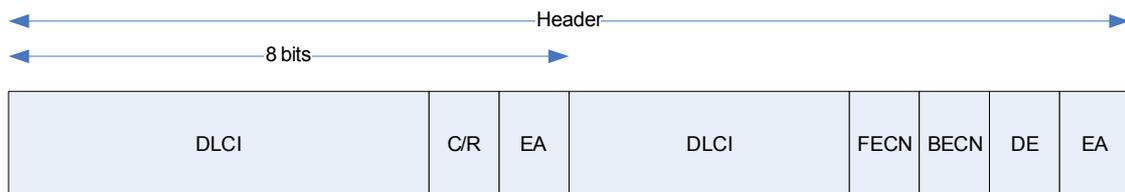


Figura 2.5 – Cabeçalho Frame Relay [BRITO 2006]

*DLCI* – identificador de conexão de enlace de dados.

*C/R (Command/Response)* – *bit* de comando e resposta. Não é usado pela rede Frame Relay;

*EA (Address Field Extension Bits)* – *bit* de extensão de endereço. Utilizado para estender o cabeçalho para 3 ou 4 octetos;

*FECN (Forward Explicit Congestion Notification)* – *bit* de notificação (para frente) explícita de congestionamento; usada para notificar um usuário a *upstream* de problemas de congestionamento na rede;

*BECN (Backwards Explicit Congestion Notification)* – *bit* de notificação (para trás) explícita de congestionamento; usada para notificar um usuário a *downstream* de problemas de congestionamento na rede;

A implementação da funcionalidade oferecida por estes *bits* não é obrigatória. Alguns fabricantes adotam a filosofia de confiar na camada de transporte para controlar o fluxo dos dispositivos dos usuários, não implementando o controle de fluxo pelos *bits FECN* e *BECN*.

*DE (Discard Eligibility)* – bit de elegibilidade de descarte; utilizado para marcar quadros para possíveis descartes.

#### 2.2.3.2.3 Funcionamento do Serviço

O usuário estima a taxa média de tráfego gerado em condições normais, que é denominada de *CIR (Committed Information Rate)*. Este parâmetro é incorporado ao contrato de serviço entre o usuário e a rede.

A rede monitora o tráfego do usuário. Se o tráfego ultrapassa o valor do *CIR*, os *bits DE* mudam o status para um, tornando os quadros elegíveis para descarte em caso de congestionamento. Portanto, a rede transporta o tráfego que ultrapassa *CIR* desde que não haja congestionamento.

Ainda existem dois outros parâmetros na rede que o implementador pode fazer uso: *Bc (Comitted Burst Size)* e o *Be (Comitted Burst Excess)*.

*Bc* é a máxima quantidade de dados (em *bits*) que o usuário pode entregar à rede em um dado intervalo de tempo (*T*). Logo:

$$CIR = \frac{Bc}{T} \quad \text{Equação 2.1}$$

O valor de *Bc* é definido durante o estabelecimento da chamada ou pré-definido no caso de *PVC (Permanent Virtual Circuits)*.

*Be* é a máxima quantidade de dados (em *bits*) que o usuário pode enviar, excedendo *Bc*, durante o intervalo de tempo *T*. Ou seja, o valor de *Be* também identifica o número máximo de *bits* que a rede tentará entregar, excedendo *Bc*, durante o intervalo *T* e este valor também é negociado no momento do estabelecimento da conexão. No entanto, a probabilidade de *Be* é menor do que a de *Bc*.

#### 2.2.3.3 ATM (*Asynchronous Transfer Mode*):

##### 2.2.3.3.1 Características Gerais

Segundo [ALBERTI1998], *ATM* é uma tecnologia de transmissão, multiplexação e chaveamento usado para transportar pequenos pacotes de tamanho fixo, chamados de

células, sobre uma rede de alta velocidade. Uma célula tem 53 *bytes*, sendo 5 *bytes* de cabeçalho e 48 *bytes* para o campo de informações. Toda a informação (voz, vídeo, dados, etc) é transportada pela rede através de células *ATM*.

O cabeçalho da célula *ATM* garante o processamento rápido dentro da rede que pode ser feito a taxas muito altas (155,52 Mbps até Gbps), porque possui funcionalidades limitadas. Sua principal função é de identificar uma conexão virtual (lógica), selecionados em uma fase de estabelecimento de conexão e que garantirão o encaminhamento das células. Além desta função, foi inserido um campo de proteção contra erros no cabeçalho (*HEC - Header Error Control*) - da célula *ATM*. Este campo evita o encaminhamento errado das células.

Outra característica importante da tecnologia é o campo de informações da célula, padronizado pelo *ITU-T* em 48 *bytes*, a partir de um compromisso firmado entre vários grupos de interesse e levando-se em conta uma série de fatores conflitantes, dos quais se destacam: Atrasos da rede; Eficiência de transmissão e Complexidade de implementação.

O fluxo de informações é estabelecido através de percursos pré-definidos, chamados canais virtuais (*VCs - Virtual Channels*). As células de um canal virtual seguem o mesmo percurso através da rede e são entregues ao destino na mesma ordem que foram inseridas na rede. Este fluxo é fragmentado em células no ponto de ingresso na rede e recuperado no ponto de egresso da rede. Esta adaptação possibilita o tratamento diferenciado para cada tipo de serviço.

Não é feita correção de erro no nível de enlace, uma vez que se considera que os enlaces utilizados nas redes *ATM* apresentam alta qualidade e, portanto, baixa taxa de erro de *bits*. Ainda, não é feito controle de fluxo no nível de enlace, em vez disso as redes *ATM* utilizam um controle de taxa de entrada que limita o tráfego entregue à rede.

*ATM* provê um serviço orientado à conexão, o que significa que antes do envio de dados é estabelecida uma conexão lógica/virtual entre os *hosts ATM*, que tem por objetivo reservar os recursos necessários para transmissão. Se os recursos forem insuficientes a conexão será recusada. Ao final da transmissão a conexão é desfeita e os recursos desalocados. Mesmo o serviço não orientado à conexão, quando prestado sobre a rede *ATM*, o fluxo de dados será transmitido sobre um ou mais caminhos pré-estabelecidos.

Em caso de congestionamento, a única ação da rede ATM é o descarte de células até que o problema seja neutralizado. As células menos prioritárias poderão ser marcadas para descarte primeiro. O nó final da rede não será notificado das células perdidas. Portanto, cabem as funções de adaptação detectar e recuperar as informações perdidas devido a congestionamentos.

O ATM usa conexões virtuais (*VCC - Virtual Channel Connections*) como mecanismo de transporte de dados da fonte para o destino. Esta conexão não poderá ser compartilhada, mas, pode haver várias conexões virtuais para um mesmo enlace físico.

#### 2.2.3.3.2 *Formato da célula ATM*

As células ATM possuem 48 bytes de informação e 5 bytes de cabeçalho, totalizando 53 bytes de tamanho total. A ilustração do cabeçalho da célula ATM pode ser visto em Figura 2.6.

O ITU-T definiu dois formatos diferentes de célula de acordo com o cabeçalho. Uma para o cabeçalho da interface usuário-rede (*UNI - User Network Interface*) e outra para o cabeçalho da interface rede-rede (*NNI - Network Network Interface*).

A diferença entre o formato da célula para as duas interfaces está no campo, de 4 bits usado para o controle genérico de fluxo (*GFC - Generic Flow Control*) no cabeçalho da célula UNI. No cabeçalho da célula NNI este campo é alocado para identificador de caminho virtual (*VPI - Virtual Path Identifier*). Por definição o controle de fluxo é feito apenas na interface UNI e o aumento do campo VPI permite que redes com grande quantidade de comutadores ATM interconectados possam suportar mais caminhos virtuais.

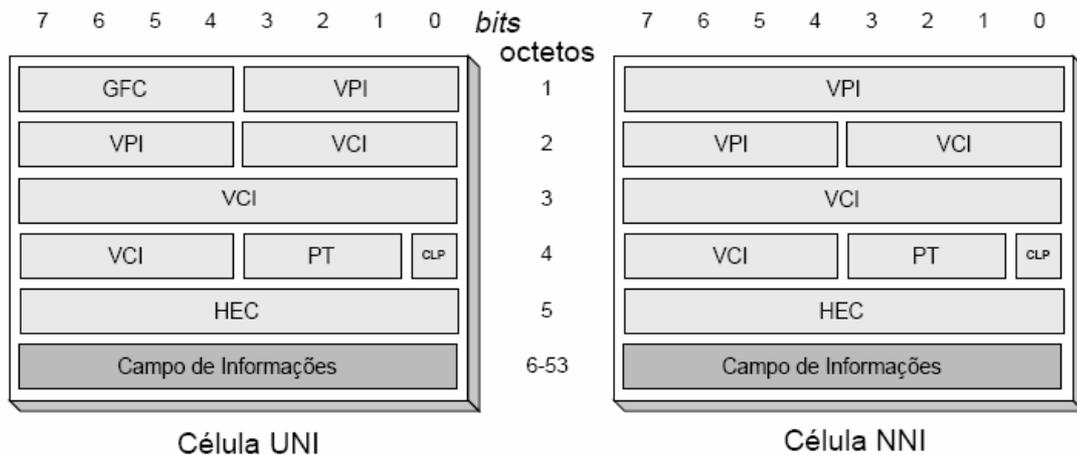


Figura 2.6 – Cabeçalho da Célula ATM [ALBERTI1998]

Segue a definição dos campos da célula ATM:

Controle Genérico de Fluxo (*GFC – Generic Flow Control*): foi concebido como mecanismo para priorização de células.

Identificador de Caminho Virtual (*VPI – Virtual Path Identifier*): possui 8 *bits* na interface *UNI* e 12 *bits* na *NNI*.

Identificador de Conexão Virtual (*VCI – Virtual Channel Identifier*): possui 16 *bits* para ambas interfaces.

Tipo de Carga (*PT – Payload Type*): usado para identificar o conteúdo da célula. É um campo de 3 *bits*.

Prioridade de Perda de Célula (*CLP – Cell Loss Priority*): sinaliza a célula que poderá ser descartada, em caso de congestionamento, por estar em desacordo com o contrato de tráfego pré-estabelecido.

Controle de Erro do Cabeçalho (*HEC – Header Error Correction*): usado para detectar e corrigir erros no cabeçalho da célula ATM.

### 2.2.3.3.3 Arquitetura das Redes ATM

Conforme [ALBERTI1998], a arquitetura das redes *ATM* é baseada em um modelo tridimensional composto por três planos e três camadas: plano de usuário (*User Plane*),

plano de controle (*Control Plane*) e plano de gerenciamento (*Management Plane*); camada física (*Physical Layer*), camada *ATM* (*ATM Layer*) e camada de adaptação (*ALL – ATM Adaptation Layer*).

O Plano de usuário é responsável por prover transferência, controle de fluxo e recuperação de informações dos usuários. Então, contém uma camada física, uma camada *ATM* e várias *AALs* suportando diferentes serviços.

O Plano de controle é responsável pelo controle e sinalização no estabelecimento, gerenciamento e finalização das conexões virtuais comutadas. Não sendo necessário nas conexões permanentes. Compartilha com o plano de usuário as camadas físicas e *ATM* e possui uma *AAL* específica de sinalização.

O Plano de gerenciamento, como o próprio nome diz, é responsável pelo gerenciamento do plano de usuário e do plano de controle. As funções são: gerenciamento de planos e gerenciamento de camadas, esta última sendo responsável pela execução de serviços de operação, administração e manutenção.

#### 2.2.3.3.4 *Camada Física*

É a camada responsável pela transmissão das células entre dois equipamentos *ATM*, através de um meio físico. O *ATM* pode usar qualquer meio físico capaz de carregar suas células.

#### 2.2.3.3.5 *Camada ATM*

É a camada acima da camada física responsável por um grande número de funções que manipulam o cabeçalho da célula, a exceção do campo *HEC* que é controlado pela camada física.

Basicamente, a camada *ATM* insere os campos de cabeçalho, incluindo os identificadores virtuais *VPI* e *VCI*, em cada *AAL-PDU* (*ATM Adaptation Layer – Protocol Data Unit*) proveniente da *AAL*. Na estação de egresso, o cabeçalho das células é removido e as *ATM-SDUs* (*ATM – Service Data Unit*) são passadas para *AAL*.

Esta camada também é responsável pela multiplexação/demultiplexação de células. Ou seja, as células provenientes de vários caminhos virtuais e conexões virtuais são

combinadas aleatoriamente na camada física para serem transmitidas. No destino, estas células são recombinaadas, com base nos conteúdos dos campos *VPI* e *VCI*, em caminhos e conexões virtuais específicas.

Em comutadores ou *cross-connects* a camada *ATM* faz a translação do *VPI* e *VCI*.

Outra função importante desta camada é a discriminação de células. Por exemplo, o campo *PT* habilita a camada a discriminar informações de células. Esta característica da camada *ATM* habilita outra função, que é de gerenciamento de tráfego, que tem por objetivo suportar *QoS* durante a conexão e proteger o usuário de possíveis congestionamentos de rede.

#### 2.2.3.3.6 Camada de Adaptação

A camada de adaptação *AAL*, executa funções solicitadas pelos planos de usuário, controle e gerenciamento e funciona como uma camada de ligação entre os serviços oferecidos pela camada *ATM* e os serviços solicitados pelas camadas superiores da rede.

A Recomendação [ITU.I.361] do *ITU-T* classifica os serviços e os protocolos designados para cada classe de serviço a serem atendidos pela *AAL*. Estas definições são mostradas na Tabela 2.1.

Tabela 2.1. Classes de Serviços (I.362) [BRITO2006]

	Classe A	Classe B	Classe C	Classe D
Sincronização entre extremos	Sim		Não	
Taxa	Constante	Variável		
Modo de conexão	Orientado à conexão			Sem conexão
Exemplos	Voz	Vídeo codificado	X.25 e FR	Datagrama IP
AAL	1	2	3/4, 5	3/4, 5

As *AALs* possuem algumas funções comuns a todos os tipos de protocolos, a saber:

As *AALs* estão localizadas no *CPE* do cliente *ATM*;

As *AALs* dependem dos aplicativos das camadas superiores;

As informações das camadas superiores passam para a *AAL* através de um ponto de acesso conhecido como *SAP – Service Access Point*, no formato de *frames*, que podem ter até 64Kbytes;

A *AAL5* é a mais importante das camadas de adaptação, pois verifica e corrige erros no nível de *frame* o que a torna mais “barata” de ser implementada;

O campo de informações disponível na *AAL5* tem 48 *bytes*, que apresenta um ganho significativo em termos de eficiência de transmissão, uma vez que a célula *ATM* tem 53 *bytes* e ainda possui mecanismos de proteção contra erros.

#### 2.2.3.4 Rede Ethernet

Segundo [TANENB2003], tanto a Internet quanto o *ATM* foram criadas para redes geograficamente distribuídas. No entanto, muitas empresas, universidades e outras organizações têm grande número de computadores que devem estar conectados. Essa necessidade deu origem à rede local.

O *IEEE (Institute of Electrical and Electronic Engineers)* padronizou várias redes locais e metropolitanas com o nome *IEEE 802*. A rede local mais popular é conhecida como *Ethernet*, conhecido padrão 802.3.

A tecnologia *Ethernet* introduziu um aperfeiçoamento importante em relação aos padrões vigentes: antes de transmitir, um computador inspeciona o meio de transmissão para ver se alguém mais já está transmitindo. Nesse caso, a interface *ethernet* não transmite seus dados até que a transmissão em andamento seja concluída. Isto evita interferências e aumenta a eficiência de transmissão. O sistema funciona da seguinte forma: cada máquina se mantém na escuta durante sua própria transmissão, se detectar interferência, gera um alerta a todos os transmissores bloqueando-os. Em seguida, aguarda um tempo aleatório antes de tentar novamente. Caso ocorra uma segunda colisão, o tempo aleatório será duplicado e assim por diante, até separar as transmissões concorrentes e dar a uma delas a chance de iniciar sua transmissão.

##### 2.2.3.4.1 Protocolo da subcamada MAC Ethernet

A estrutura do quadro *Ethernet* é mostrada na Figura 2.7.

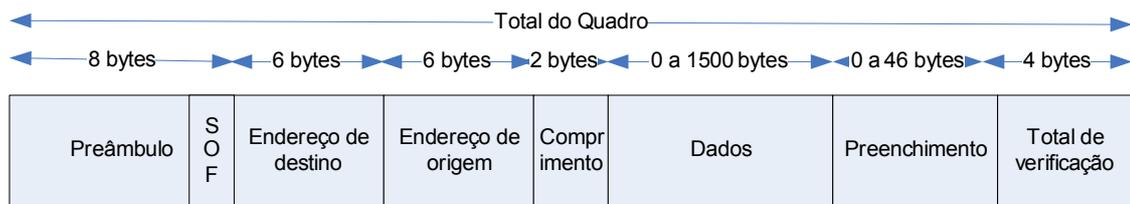


Figura 2.7 – Formato quadro *Ethernet* segundo IEEE 802.3 [TANEMB2003]

Cada quadro contém um **Preâmbulo** de *8bytes*, cada um contendo o padrão de *bits* 10101010. A codificação Manchester desse padrão produz uma onda quadrada, a fim de sincronizar o *clock* do transmissor com o *clock* do receptor. O quadro contém dois **endereços**, um para o destino e um para origem. O padrão prevê endereços de 2 e de 6 *bytes*. O bit de alta ordem do endereço de destino é 0 (zero) para endereços comuns e 1 para endereços de grupos. O endereço de grupo permite que várias estações escutem um único endereço. A transmissão para um grupo de estações é conhecida como multidifusão ou *multicast*. Um endereço contendo somente *bits* 1 é reservado para difusão ou *broadcast*.

Um quadro quando enviado para um endereço contendo apenas *bits* 1's é aceito por todas as estações da rede.

Em seguida vem o campo **Comprimento**, que indica ao receptor o tamanho do quadro que está sendo transmitido. Depois vêm os **Dados**, com até 1.500 *bytes*. Existe também um tamanho mínimo do quadro de Dados, para evitar que, ao detectar uma colisão, o transceptor trunque o quadro, o que significará *bits* perdidos e fragmentos de quadros. Logo, quadros válidos devem ter, no mínimo, 64 *bytes*. Se a parte de dados for menor que 46 *bytes*, o campo **Preenchimento** será usado para preencher o quadro até o tamanho mínimo.

Outra razão para manutenção de um quadro mínimo é impedir que uma estação conclua a transmissão de um quadro curto antes de o primeiro *bit* ter atingido a outra extremidade do cabo, onde ele poderá colidir com outro quadro.

O último campo *Ethernet* é o **Total de Verificação**. Este campo utiliza o algoritmo *CRC* (*Cyclic Redundancy Check*) que é um código de 32 *bits* que realiza a detecção de erro antecipada, não fazendo a correção de erros.

#### 2.2.3.4.2 *Fast Ethernet*

A idéia básica por trás do *Fast Ethernet* era simples: manter os antigos formatos de quadros, interfaces e regras de procedimentos e apenas reduzir o tempo de *bit* de 100ns para 10ns [TANEMB2003]. Como o projeto do *Fast Ethernet*, ou 802.3u como chamou o *IEEE*, se baseou inteiramente no cabo 10Base-T, este sistema só utiliza *hubs* e *switches*, ou seja, cabos multiponto com conectores de pressão ou cabos *BNC* não podiam mais ser usados.

#### 2.2.3.4.3 *Gigabit Ethernet*

Em 1998 o *IEEE* padronizou o *Gigabit Ethernet*, como o padrão 802.3z. O objetivo do comitê era, essencialmente, tornar a *Ethernet* 10 vezes mais rápida, mantendo a compatibilidade retroativa com os demais padrões [TANEMB2003].

Logo, ficou definido que *Gigabit Ethernet* pode operar baseado em *switch* ou *hub*, todas as configurações da *Gigabit Ethernet* são ponto-a-ponto e não mais multiponto como o padrão original 10Mbps. Permite dois modos de operação: *full-duplex* e *half-duplex*. O meio físico é baseado em fibra ou par metálico (categoria 5).

Ainda, o *Gigabit Ethernet* utiliza o mesmo formato de quadro e mesmo comprimento de quadro da *Ethernet*, e no modo *half-duplex* também o mesmo protocolo de acesso (*CSMA-CD - Carrier Sense Multiple Access With Collision Detection*).

Permite a integração imediata com a rede existente, facilitando processo de migração. Inicialmente utilizado no *backbone* da rede, mantendo-se o padrão *Ethernet* (ou *Fast Ethernet*) no *desktop*.

#### 2.2.3.4.4 *Ethernet na Metro*

A tecnologia *Ethernet* tem sido amplamente difundida e aceita no mundo *Enterprise*, onde milhares de portas já foram implementadas. A simplicidade desta tecnologia permite um aumento escalar de largura de banda enquanto é mantido o custo efetivo da interface [BRITO2006].

Este custo, performance medida e as facilidades da *Ethernet* são motivadores para as Operadoras de rede fazerem uso desta tecnologia. No novo modelo o usuário recebe uma interface *Ethernet* no lugar de uma interface *TDM (Time Division Multiplex)* com as seguintes vantagens:

**Escalabilidade de Largura de Banda:** o baixo custo das interfaces *ethernet* no *CPE* (*Customer Premises Equipment*) e no equipamento de acesso do lado da operadora favorece a instalação de interface *ethernet* de alta velocidade. As interfaces *TDMs* oferecidas na maioria dos *CPEs* variam de acordo com a velocidade o que dificulta no provisionamento de novas velocidades, já que isto implica na substituição do equipamento do cliente.

**Granularidade de Largura de Banda:** a interface *ethernet* pode ser provisionada para entregar a largura de banda que escala a máxima velocidade da interface. É importante notar que a granularidade de largura de banda não é uma função específica da *ethernet*, mas é implementada em qualquer interface de pacotes.

**Aprovisionamento mais rápido:** o desenvolvimento do serviço *ethernet* implica em um diferente modelo operacional, no qual um “pacote” é entregue, no lugar do “circuito”. O modelo de provisionamento de pacotes pode ser feito sem mudanças nos equipamentos de rede e suas interfaces. O provisionamento de pacotes é uma função simples de mudar os parâmetros de software que regulam a vazão dos pacotes e podem aumentar ou diminuir a largura de banda, estabelecendo uma conexão em minutos e construindo novos serviços.

#### 2.2.3.5 Rede IP

A Internet pode ser vista como uma coleção de sub-redes independentes, chamadas de Sistemas Autônomos (*AS – Autonomous Systems*), interconectados através de redes regionais, *backbones* nacionais, etc. [TANENB2003].

Um exemplo de sistemas autônomos pode ser uma rede local de uma corporação, rede corporativa (*WAN – Wide Area Network*) interconectadas, ainda um conjunto de redes locais (sub-redes) interconectadas por roteadores e conectadas ao *backbone* através de um único *Exterior Gateway*.

Um esquema dessa organização semi-hierárquica é mostrado na Figura 2.8.

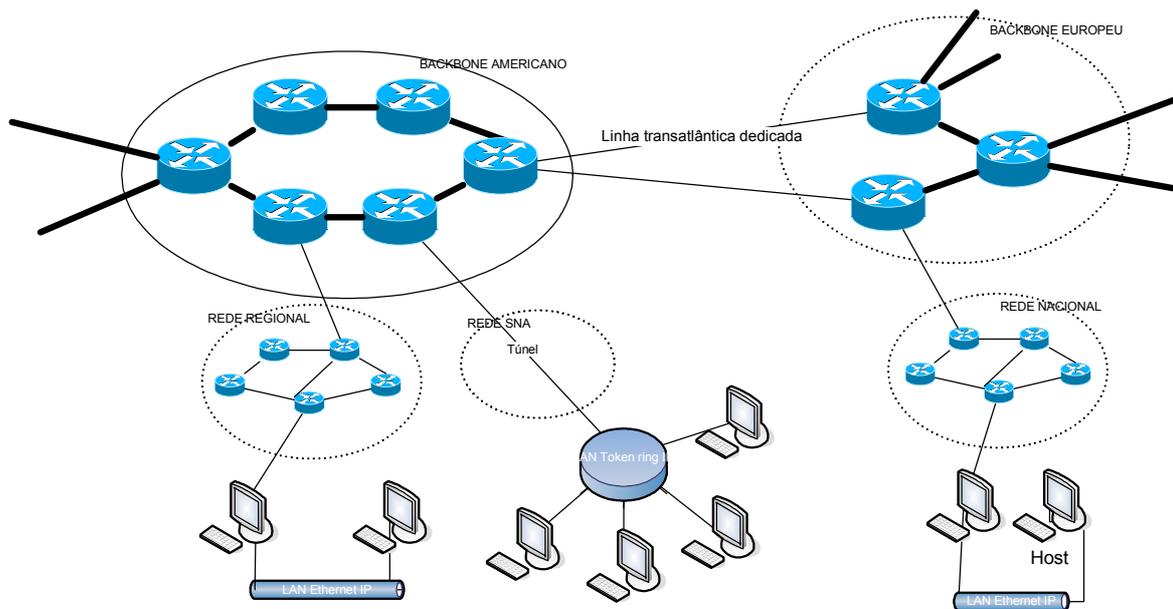


Figura 2.8 – Organização hierárquica de rede [TANEMB2003]

Importante entender que um *Interior Gateway* é utilizado para interconectar redes de um mesmo sistema autônomo, enquanto um *Exterior Gateway* é utilizado para interconectar os sistemas autônomos aos *backbones*.

O elemento que mantém a Internet unida é o protocolo da camada de rede, o *IP (Internet Protocol)*. Ao contrário da maioria dos protocolos da camada de rede mais antigos, o *IP* foi projetado desde o início tendo como objetivo a interligação de redes. Portanto, o objetivo do *IP* é fornecer a melhor forma possível, sem garantias, de transportar datagramas da origem para o destino, independentemente de essas máquinas estarem na mesma rede ou de haver outras redes entre elas.

Na Internet a camada de transporte recebe os fluxos de dados e os divide em datagramas. Na teoria, cada datagrama pode ter até 64 *Kbytes*, no entanto, geralmente eles têm no máximo 1500 *bytes*, cabendo em um único quadro *Ethernet*.

Cada datagrama pode ser fragmentando durante o transporte em unidades menores. Quando todos os fragmentos chegarem ao destino, são remontados pela camada de rede no datagrama original. Em seguida, esse datagrama é entregue à camada de transporte, que o insere no fluxo de entrada do processo de recepção.

#### 2.2.3.5.1 Protocolo IP

Conforme [TANEMB2003], o cabeçalho do datagrama *IP* segue o seguinte formato, apresentado na Figura 2.9. O campo *Version*, de 4 *bits*, controla a versão do protocolo à qual o datagrama pertence. Atualmente, a maioria dos equipamentos trabalha na versão IPv4 (*Internet Protocol version 4*) com uma provável transição prevista para versão IPv6 (*Internet Protocol version 6*).

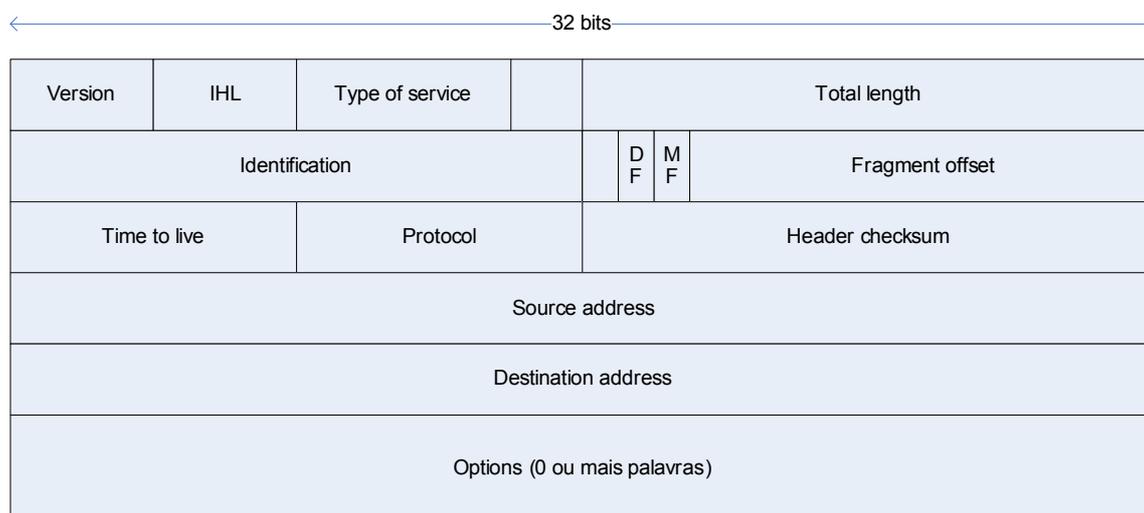


Figura 2.9 – Cabeçalho IPv4 (IP) [TANEMB2003]

Como o tamanho do cabeçalho não é constante, existe um campo, *IHL* (*Internet Header Length*), que informa o tamanho do cabeçalho, em palavras de 32 *bits*. O valor mínimo é 5 e o valor máximo é 15 *bits*, que limita o cabeçalho a 60 *bytes*, e o campo *Options* a 40 *bytes*.

O campo *Type of service*, campo de 8 *bits* destinado a distinguir entre diferentes classes de serviço. São possíveis várias combinações de confiabilidade e velocidade. Por exemplo, para o serviço de voz a velocidade supera a segurança ou qualidade da entrega, por outro lado, na transferência de dados, uma transmissão sem erros é mais importante do que uma transmissão rápida.

O campo *Total Length*, de 16 *bits* que contém o comprimento total do datagrama, isto é: cabeçalho mais dados. O tamanho máximo é de 65.535 *bytes*. O campo *Identification*, igualmente um campo de 16 *bits* permite que o *host* destino determine qual datagrama pertence um fragmento recém-chegado. Todos os fragmentos de um datagrama contêm o mesmo valor de *Identification*.

Em seguida está o campo *DF (Don't Fragment)*, campo de 1 *bit* que indica que o datagrama não deve ser fragmentado. E o campo *MF (More Fragments)*, que serve para indicar a ordem dos fragmentos de um mesmo datagrama. Apenas o último fragmento não tem esse conjunto de *bits*.

O campo de 13 *bits* de *Fragment offset*, informa a que ponto do datagrama atual o fragmento pertence. Todos os fragmentos de um datagrama, com exceção do último, devem ser múltiplos de 8 *bytes*. A unidade elementar de fragmento.

O campo *Time to live* é um contador de 8 *bits* usado para limitar a vida útil dos pacotes, permitindo o máximo de 255 segundos. Esse contador deve ser incrementado a cada *hop* e supõe-se que ele seja decrementado diversas vezes quando estiver enfileirado durante um longo tempo em um roteador.

O campo *Protocol*, igualmente de 8 *bits*, informa a que processo de transporte o datagrama deve ser entregue. O *TCP (Transmission Control Protocol)* é uma opção, mas também há o *UDP (User Datagram Protocol)*, entre outros.

O campo *Header checksum* confere apenas o cabeçalho. Este campo serve como verificador de erros gerados por palavras de memória incorretas em um roteador, possui 16 *bits*.

Os campos *Source address* e *Destination address*, cada um com 32 *bits* indicam o número da rede e o número do *host*.

O campo *Options* foi projetado para permitir que versões posteriores do protocolo incluam informações inexistentes no projeto original.

#### 2.2.3.5.2 *Protocolos de Transporte*

A camada de Transporte é o ponto central de toda a hierarquia de protocolos. O objetivo desta camada é a qualidade do serviço que será entregue à camada de aplicação, vindo da camada de rede [TANEMB2003].

A internet possui dois protocolos principais na camada de Transporte, um protocolo sem conexão, que é o *UDP* e outro orientado à conexão, o *TCP*.

##### 2.2.3.5.2.1 *UDP (User Datagram Protocol)*,

É descrito na *RFC 768*. É um meio para as aplicações enviarem datagramas *IP* encapsulados sem que seja necessário estabelecer uma conexão. Isto é, fornece uma interface para o protocolo *IP* com o recurso adicional de demultiplexação de vários processos que utilizam as portas.

O *UDP* transmite segmentos que consistem em um cabeçalho de 8 *bytes*, seguido pela carga útil.

A Figura 2.10 mostra o Cabeçalho do protocolo *UDP*.



Figura 2.10 – Cabeçalho do UDP [TANEMB2003]

A principal vantagem do *UDP* em relação ao *IP* simples são as facilidades das portas de fonte e destino. Uma vez que, estas portas servem para identificar os pontos extremos nas máquinas de origem e destino. Sem estes campos a camada de transporte não saberia o que fazer com o pacote.

O campo *UDP length* inclui o cabeçalho de 8 *bytes* e os dados. O campo *UDP checksum* é opcional, sendo 0 quando não calculado.

Importante mencionar que o *UDP* não implementa algumas funcionalidades como: controle de fluxo e controle de erros ou retransmissão após a recepção de um segmento incorreto.

#### 2.2.3.5.2.2 RTP – (Real-time Transport Protocol)

É descrito na [RFC1889]. É um protocolo de transporte inserido no plano do usuário (aplicação) isto é, executado sobre *UDP*.

A função básica do *RTP* é multiplexar diversos fluxos de dados de tempo real sobre um único fluxo de pacotes *UDP*. Como o *RTP* utiliza apenas o *UDP* normal, seus pacotes não são tratados de maneira especial pelos roteadores, a menos que alguns recursos de qualidade de serviço normais do *IP* estejam ativos.

O mecanismo usado pelo *RTP* é o de numerar seus pacotes com um número maior que seu predecessor. Desta forma, se o destino detecta a falta de algum pacote, pode estimar, por interpolação, o valor faltante e completá-lo, já que retransmissões não são úteis neste caso porque o pacote retransmitido provavelmente chegaria tarde demais ao destino. Como consequência, o *RTP* não tem nenhum controle de fluxo, de erros, confirmação e retransmissão.

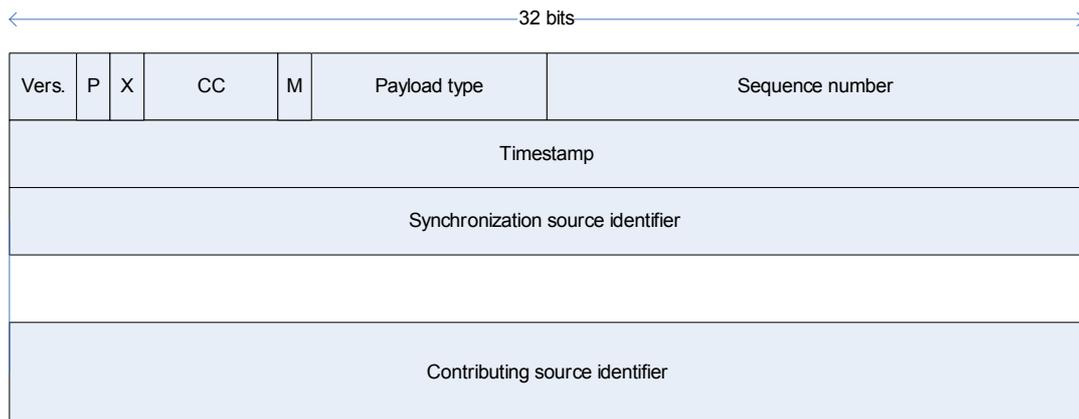


Figura 2.11 – Cabeçalho do RTP [TANEMB2003]

A Figura 2.11 mostra o Cabeçalho do protocolo *RTP*.

Consiste em três palavras de 32 *bits* e algumas extensões.

A primeira palavra contém o campo *Version*, campo de 2 *bits* que indica a versão do protocolo. Um *bit P* indicando que o pacote foi completado até chegar a um múltiplo de 4 *bytes*. Um *bit X* indicando que um cabeçalho de extensão está presente.

O campo de 4 *bits CC* que informa quantas origens estão presentes, de 0 a 15. Um *bit M*, que é um marcador da aplicação. Pode ser usado para indicar o começo de um quadro de vídeo, por exemplo.

O campo *payload type*, de 7 *bits*, informa que algoritmo de codificação está sendo usado. Importante comentar que o *RTP* possui este campo para permitir especificar um perfil de codificação usado pela aplicação para cada amostra de carga útil. Para interoperação, o *RTP* define vários perfis e, para cada perfil, são permitidos vários formatos de codificação.

O campo *sequence number* de 16 *bits* é apenas um contador incrementado em cada pacote *RTP*. Serve para detectar pacotes perdidos.

Todos os demais campos são de 32 *bits*. O *timestamp* é gerado pelo fluxo de origem para apontar quando a primeira amostra do pacote foi realizada. Este valor pode ajudar a reduzir a latência no receptor, já que a latência é medida pela diferença dos tempos das amostras no receptor. Outra vantagem do *timestamp* é a sincronização dos fluxos. Ou seja, dois fluxos de fontes diferentes, mas que estejam marcados como o mesmo *timestamp*, poderiam ser reproduzidos de modo sincronizado.

O *synchronization source identifier* informa a que fluxo o pacote pertence. Este é o método usado para multiplexar e demultiplexar vários fluxos de dados em um único fluxo de pacotes *UDP*.

Os campos *contributing source identifiers* são opcionais. Caso estejam presentes serão usados como misturadores, por exemplo: *mixer* de áudio no estúdio.

Não serão considerados, para fins do trabalho, os protocolos de controle. Desta forma, não será feito nenhum comentário sobre o algoritmo *RTCP* (Real Time Transport Control Protocol).

#### 2.2.3.5.2.3 TCP – (*Transmission Control Protocol*)

Segundo [TANEMB2003], pode também ser chamado de protocolo de transporte da Internet. Está definido na RFC 1122 e suas extensões na RFC 1323. Foi projetado para oferecer um fluxo de dados fim a fim confiável quando servir de transporte inter-redes não confiáveis.

Uma inter-rede é uma rede que possui várias partes interconectadas de maneira aleatória. Estas partes podem ter topologias, larguras de banda, latência, tamanhos de pacotes e outros parâmetros completamente diferentes. Logo, o *TCP* foi projetado para se adaptar dinamicamente a estas diferenças e às falhas que podem ocorrer durante o transporte.

Uma característica geral do *TCP* é que cada *byte* em uma conexão *TCP* tem seu próprio número de seqüência de 32 *bits*. As entidades transmissoras e receptoras trocam dados na forma de segmentos. Um segmento *TCP* consiste em um cabeçalho fixo de 20 *bytes*, seguido por zero ou mais *bytes* de dados. O software *TCP* decide o tamanho de cada segmento, restrito por dois fatores: cada segmento, incluindo o cabeçalho, deve caber na carga útil do *IP*, que é de 65.515 *bytes*; cada rede tem uma unidade máxima de

transferência, ou *MTU (Maximum Transfer Unit)* e, portanto, cada segmento deve caber na *MTU*. Na prática, o *MTU* é da ordem de 1.500 *bytes*.

A Figura 2.12 mostra o cabeçalho do segmento *TCP*:

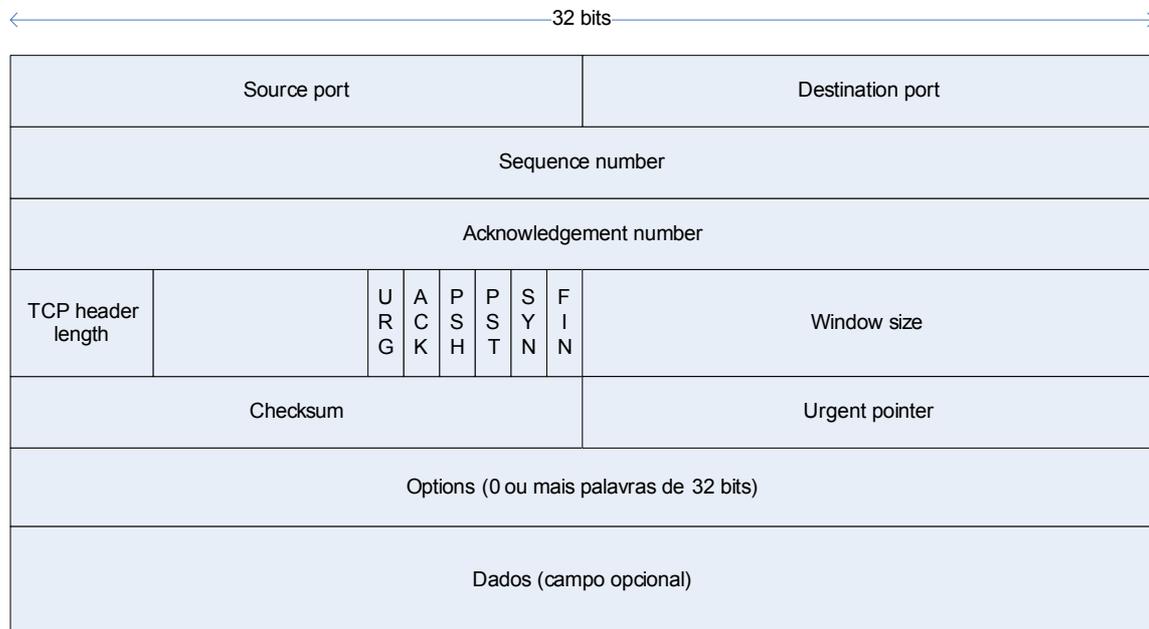


Figura 2.12 – Cabeçalho TCP [TANEMB2003]

Os campos *source port* e *destination port* identificam os extremos da conexão (origem e destino). Uma porta e o endereço *IP* de seu *host* formam um único ponto terminal de 48 *bits*. Os números dos pontos terminais identificam a conexão.

Os campos *sequence number* e *acknowledgement number*, têm 32 *bits*, pois cada *byte* de dados é numerado em uma conexão *TCP*. O segundo especifica o próximo *byte* esperado.

O campo *TCP header length* informa quantas palavras de 32 *bits* existem no cabeçalho *TCP*.

O campo *checksum* é um campo de 6 *bits* não utilizado.

Em seguida existem seis *flags* de 1 *bit* cada um: *URG (Urgent pointer)*, este recurso substitui a mensagem de interrupção, permitindo ao transmissor enviar uma mensagem ao receptor sem envolver o serviço *TCP* no motivo da interrupção; *ACK (Acknowledgement number)* é o *bit* de confirmação; *PSH (Push)* informa ao receptor que ele deve entregar o fluxo de dados a aplicação e não armazená-lo na espera de completar o *buffer*; *RST (Reset)*

usado para reinicializar uma conexão; *SYN* (*Synchronization*) usado para estabelecer uma conexão, significa em outras palavras, *connection request* e *connection accepted*; *FIN* (*Final*) é utilizado para encerrar uma conexão.

O campo *window size* indica quantos *bytes* podem ser enviados a partir do *byte* confirmado. Importante comentar que o controle de fluxo do *TCP* é feito por meio do mecanismo de Janela Deslizante, que não será objeto de estudo nesta pesquisa.

O campo *checksum* serve para aumentar a confiabilidade. Confere o total de verificação do cabeçalho e dos dados.

O campo *Options* é um recurso extra. O uso de segmentos grandes é mais eficiente do que a utilização de segmentos pequenos, pois o cabeçalho de 20 *bytes* pode ser diluído em um maior volume de dados; porém, é possível que *hosts* pequenos não sejam capazes de administrar segmentos muito grandes. Durante a configuração da conexão, cada lado pode anunciar sua capacidade máxima, se não usar esta opção, o valor padrão é de 536 *bytes* para carga útil. Ou seja, todos os *hosts* da Internet devem ser capazes de aceitar segmentos *TCP* de 556 *bytes* (536 + 20).

#### 2.2.3.6 Rede MPLS (*Multiprotocol Label Switching*)

Segundo [BRITO2006], aumentar a velocidade e eficiência de redes *IP* é essencial para desenvolvimento de novas aplicações. Neste contexto, o propósito do *MPLS* é melhorar o desempenho do roteamento de datagramas e evitar congestionamentos na rede.

Inicialmente, as tecnologias de comutação utilizavam hardware *ATM* para comutação de *IP* por rótulos e eram soluções aplicadas nos protocolos de enlace de redes. Essas tecnologias, na metade dos anos noventa, sofreram mudanças em sua utilização, continuaram com hardware *ATM*, por sua velocidade e incorporaram sinalização *IP* pensando em roteamentos integrados com endereços *IP* [BRITO2006].

*Multiprotocol Label Switching* (*MPLS*) é um padrão outrora emergente do *IETF* e hoje bastante disseminado, o qual permite ampliar o desempenho de tecnologias de redes já existentes. O *MPLS* evolui combinando tecnologias de comutação da camada 2 (camada de enlace) com as tecnologias de roteamento da camada 3 (camada de rede).

O que significa dizer também que o *MPLS* separa os planos de roteamento e encaminhamento. No plano de roteamento os pacotes são enviados pela rede com base no endereço *IP* de destino, utilizando-se do roteamento *IP* normal, enquanto no plano de encaminhamento os pacotes são enviados com base em um *label* acrescentado ao pacote. [BRITO2006]

No modelo *MPLS*, quando um datagrama chega ao elemento de entrada da rede, denominado *LSR* (*Label Switch Router*) ingresso, este busca classificá-lo em uma *FEC* (*Forward Equivalence Class*) específica, para a qual já existia um *LSP* (*Label Switch Protocol*) definido. Seu *LSP* tem sentido de fluxo unidirecional e define uma seqüência ordenada de *LSRs* onde o pacote irá passar, sendo que o último é chamado de *LSR* egresso. [BRITO2006]

A seguir será detalhada a arquitetura e nomenclatura da rede, bem como o funcionamento do padrão *MPLS*.

#### 2.2.3.6.1 *LSR – Label Switching Routers, LSP – Label Switch Path*

Roteadores de comutação por rótulos (*LSR*) são equipamentos que realizam a comutação no protocolo *MPLS*. Um *LSR* é um dispositivo que aumenta a velocidade de encaminhamento no núcleo de uma rede *MPLS*. Quando um *LSR* localiza-se na periferia da rede *MPLS* denomina-se *LSR* de borda (*Edge*), enquanto que os situados no núcleo da rede denominam-se *LSR* de núcleo (*Core*). Ao conjunto de *LSRs* denomina-se nuvem *MPLS*.

Os *LSRs* participam no estabelecimento do *LSP* (*label switch path*) ou caminhos comutados por rótulos. Os *LSPs* são determinados por ação do protocolo do plano de controle ou por ação de gerência de rede. As determinações das rotas para um *LSP* podem ser definidas com auxílio de protocolos de roteamento convencionais. Os *LSPs* são unidirecionais, isto é, suportam encaminhamentos de datagramas em um único sentido de modo que um *LSP* pode definir uma seqüência ordenada de *LSPs*. Ao primeiro *LSR* chamamos de *LSR* ingresso (*Ingress*) e o último de *LSR* egresso (*Egress*).

No estabelecimento de uma sessão para distribuição de rótulos necessitamos sempre de *LSRs* pares. *LSR* par é o conjunto de *LSR* adjacentes, onde é estabelecida uma sessão para fins de distribuição de rótulos. Dependendo do sentido do fluxo de datagramas no *LSP*, o

*LSR* a jusante, seleciona o rótulo que o *LSR* a montante, deve utilizar no encaminhamento dos datagramas.

Sendo que, *LSR* jusante é o primeiro equipamento que faz parte do *LSR* par, responsável por selecionar o rótulo que será utilizado. O *LSR* montante é o segundo equipamento que faz parte do *LSR* par e que utilizará os rótulos que foram selecionados [BRITO2006].

Na Figura 2.13 é ilustrada a topologia da Nuvem *MPLS*.

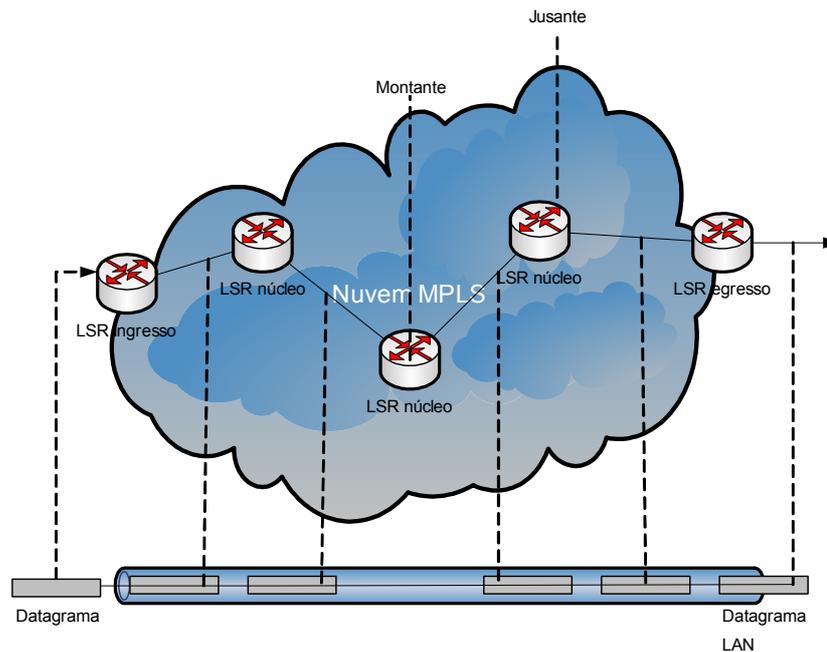


Figura 2.13 – Principais elementos *MPLS* [BRITO2006]

#### 2.2.3.6.2 *FEC* – *Forward Equivalence Class*

Uma *FEC* representa um conjunto de condições, pré-determinadas, para classificar se um dado datagrama pertence ou não a *FEC* estabelecida, ou seja, se compartilham das mesmas exigências para seu transporte [BRITO2006].

Em todo o pacote há um grupo específico e a este é dado o mesmo tratamento da origem ao destino, ao contrário da remessa convencional do *IP*, em *MPLS*, as atribuições de um determinado pacote a uma *FEC* são baseadas em exigências de serviços ou simplesmente no prefixo do endereço.

Cada *LSR* constrói uma tabela para especificar como um pacote deve ser enviado. Existem dois tipos de elementos *FEC*:

Prefixo de rede: no caso *IPv4*, possui comprimento arbitrário de 0 a 32 *bits*;

Endereço de nó: é um endereço *IP* de classe A, B ou C.

Um elemento *FEC* pode contemplar ainda informações adicionais sobre:

A origem do datagrama (endereço *IP* de origem);

A carga do datagrama (protocolo e *ports* de transporte);

Parâmetros de qualidade de serviços tais como:

Precedência do quadro, conforme definido pelo padrão *IEEE 802.1d*; classe de serviço, conforme definida pela arquitetura de serviços diferenciados (*DiffServ*);

Precedência do datagrama, conforme definida pelo campo *TOS (Type of Service)* do protocolo *IP*.

#### 2.2.3.6.3 Funcionamento dos Rótulos no MPLS

Ainda segundo [BRITO2006], um rótulo, em sua forma mais simples, é atribuído na camada de enlace e identifica o trajeto que um pacote deve seguir. O roteador de recepção examina o pacote e verifica o índice do rótulo para determinar o próximo *hop*.

Uma vez o pacote rotulado, o encaminhamento deste para o destinatário está baseado na troca do rótulo, isto é, cada valor do rótulo é de significado apenas local pertencendo somente aos “saltos” (*hops*) entre *LSRs*. Quando um pacote for classificado com uma *FEC* nova ou existente, um rótulo será distribuído ao pacote. Os valores dos rótulos são responsáveis pelo envio dos pacotes e são baseados na camada de dados.

O rótulo é encaixado entre a camada de dados e a camada de rede. Atribuição do rótulo permite uma melhor qualidade de serviço de remessa nos seguintes casos:

- Distribuição *unicast*;
- Engenharia de Tráfego;
- Multicast;
- Rede Privativa Virtual (*VPN*).

Há diversos métodos usados na criação dos rótulos:

Método baseado na topologia – usa protocolo de distribuição normal como o *OSPF (Open System Protocol Forward)* e o *BGP (Border Gateway Protocol)*;

Método baseado na requisição – usa processo baseado nas requisições de fluxo de controle de dados como o *RSVP (Resource ReSerVation Protocol)*.

Método baseado no tráfego – usa a recepção do pacote para produzir a atribuição e distribuição da etiqueta.

Uma coleção de dispositivos *MPLS* forma um domínio *MPLS*. Dentro deste domínio, o pacote de dados recebe um rótulo que define o trajeto, bem como uma *FEC* que define um grupo específico ao qual o pacote pertence. O componente responsável por determinar a distribuição de trajetos antes da transmissão é o *LSP*. Quanto à forma de distribuir o *LSP* pode fazer:

Distribuição *hop-by-hop*: Cada *LSP* seleciona independentemente do próximo *hop* aqueles pertencentes a uma *FEC* determinada. Esta metodologia é similar às usadas em redes *IP* como *BGP* e *OSPF*;

Distribuição Explícita: O *LSR* de ingresso determina uma lista de nós em que o pacote irá passar. Seu trajeto determinado pode assegurar qualidade de serviço no tráfego de dados, possibilitando otimização do fluxo da rede. O estabelecimento do *LSP* para uma *FEC* é unidirecional.

Ainda, pode haver fluxo de pacotes de diferentes interfaces rotulados e comutados juntos, usando um rótulo comum. Isto é conhecido como agregação dos fluxos. [BRITO2006]

#### 2.2.3.7 Formato do Datagrama

Para tecnologias de enlace que não empregam identificadores de conexão, o *MPLS* define uma rotulação denominada encapsulamento genérico. Esta estrutura chamada *shim header* armazena o rótulo ao datagrama de forma posicionada entre o cabeçalho de enlace e o datagrama *IP*.

Esta estrutura é composta por um rótulo de 20 *bits*, um campo experimental (*EXP - Experimental*) de 3 *bits*, um campo *TTL (time to live)* de 8 *bits* e um campo *B (Bottom of*

*stack flag*) de 1 bit. Este último tem por objetivo indicar se o rótulo corresponde ou não ao último de uma pilha de rótulos, o que permite o encapsulamento de múltiplos rótulos. O formato do datagrama é mostrado na Figura 2.14:

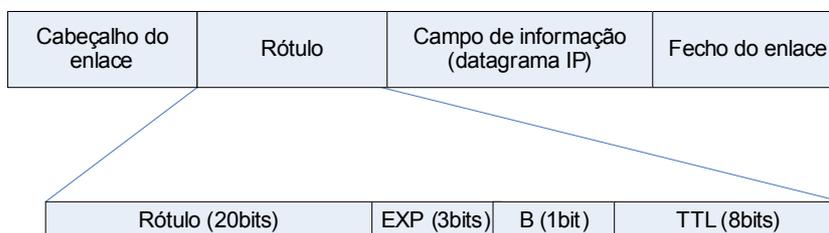


Figura 2.14 – Estrutura *Shim Header* [BRITO2006]

A arquitetura *MPLS* possibilita organização de vários rótulos utilizando-se o método *FIFO* (*First in First out*). O ponto fundamental de envolvimento de pilhas de rótulos é o encaminhamento realizado nos *LSRs*, onde é sempre baseado no rótulo encontrado no topo da pilha. Esta possibilidade permite definir uma hierarquia de rótulos na pilha e implica em importantes aspectos de roteamento e tunelamento de informações.

#### 2.2.4 Rede Inteligente

A Rede Inteligente é a rede que suporta os serviços diferenciados e adicionais sobre a RTPC (Rede de Telefonia Pública Comutada). As suas principais características são: centralização lógica dos serviços; uso intensivo de técnicas de processamento de informação; eficiente uso dos recursos da rede; modularização e reutilização de funções da rede para determinados casos; integração no aprovisionamento de serviços no sentido de reutilização das funções da rede; capacidade de delegar ao usuário do serviço a autonomia de controlar por si só os serviços contratados; e uma padronização intuitiva para gerência de serviços [KAPP2003 apud ITUM3200]. São exemplos destes serviços: “Número Único Nacional”, “0800”, “0300”, entre outros.

Esta rede não ficou restrita à rede de comutação telefônica, mas inclui outros serviços interligados. Por exemplo: se um usuário estiver conectado à Internet, por via discada, utilizando o padrão ITU V.92, a rede inteligente é capaz de enviar uma mensagem ao usuário, a qual aparecerá na tela do computador uma nova chamada entrante, avisando o número que está chamando e perguntando se o usuário deseja atendê-la.

### **2.2.5 Rede Canal Comum para Sistema de Sinalização Número 7**

O sistema de Sinalização Canal Comum número 7 (*SS7 - Signaling System #7*) é um padrão global para telecomunicações definido pelo *ITU-T*, o qual determina os procedimentos e protocolos utilizados por uma rede pública de comutação para a troca de informações sobre uma rede digital de sinalização, com o objetivo de, dinamicamente, estabelecer, gerenciar, tarifar uma conexão na rede física ou celular [SCHÄFF1999].

Sinalização por Canal Comum, também chamada sinalização Fora de Banda, é aquela onde se dispõe de um canal específico e exclusivo para a troca de informações de sinalização entre dois centros de comutação. Este canal de sinalização é completamente independente dos canais utilizados para o estabelecimento de uma conexão de voz ou de dados.

Através deste canal de sinalização trafegam todos os sinais necessários para o estabelecimento de uma conexão na rede, suas informações de tarifação, controle e desconexão. Também trafegam todos os demais sinais utilizados para uma gama de outros serviços que podem ser propiciados, por exemplo, através de consultas a bancos de dados específicos [KAPP2003].

### **2.2.6 Rede Digital de Serviços Integrados (RDSI)**

A RDSI – ou *ISDN (Integrated Services Digital Network)* – é uma rede que surgiu em meados da década de 1980, como uma rede de faixa estreita integrando serviços de dados e voz por meio da digitalização de dois canais de voz sobre os acessos metálicos. Previa a interconexão destes usuários separando os serviços de voz, os quais iriam para as centrais de comutação de voz, e os serviços de dados iriam para uma rede de pacotes. Este sistema funcionou bem nos Estados Unidos em quatro operadoras locais, na Sprint e na AT&T. Na Europa, apenas a Alemanha, Áustria, Dinamarca e França operaram desta forma. Entretanto, uma adaptação realizada na Suécia se espalhou para o resto do mundo, adaptando as centrais de comutação existentes de modo a trocarem sinalização RDSI. Assim, foi possível disponibilizar esta forma de RDSI sem uma rede de pacotes ou uma rede *ATM*. A RDSI funciona em todas as operadoras do Brasil desta forma diferenciada [KAPP2003].

### **2.2.7 RDSI-FL ou B-ISDN**

A RDSI Faixa Larga ou Rede Banda Larga é uma rede desenvolvida, segundo [STALLINGS1997] como uma rede integrada de comutação, ou seja, trabalhando na camada dois, a fim de transportar diferentes tipos de serviços de telecomunicações e baseada nas tecnologias *SDH* (*Synchronous Digital Hierarchy*) e *ATM*. Comercialmente estão implantadas na Austrália, Estados Unidos, França e Singapura. A integração dos serviços seria possível através das Camadas de Adaptação *ATM – AALs* – onde cada serviço seria adaptado ao transporte via células comutáveis na camada dois e de tamanho fixo (53 octetos). Como utiliza agregados de tamanho equivalente ao *SDH*, pode ser facilmente transportada pela Rede de Transporte.

### **2.2.8 Rede Determinística ou Rede Dedicada de Circuitos de Dados**

Esta rede é formada pelo conjunto de equipamentos que fornecem linhas dedicadas como serviço ou subserviço, assim dizendo, como acesso a outras redes fornecedoras de serviços. Pode transportar voz, dados, imagem e vídeo. É totalmente transparente a qualquer protocolo [ITUM3200].

O nome Rede Determinística deve-se ao fato do usuário contratar um serviço de banda fixa e de camada um (modelo *OSI* e *IETF*). Para determinados casos, ela faz parte da rede de acesso e, assim como a Rede Estatística e de Pacotes, tem a rede de acesso como tributários e a rede de transporte como agregados.

Os principais equipamentos são os multiplexadores determinísticos e os protocolos *OSI* são o X.50 e X.51, com padrões de interfaces V.24, V.35 e V.36. Existem padrões proprietários, o que foge ao escopo deste trabalho.

### **2.2.9 Rede de Gerência ou TMN (Telecommunications Management Network).**

A Rede de Gerência é a entidade que propicia o gerenciamento das demais redes de telecomunicações, pode ou não suportar totalmente o Gerenciamento de Telecomunicações. O termo *TMN* está mais associado ao Modelo *OSI* do *ITU*, mas nada impede que seja utilizado para qualquer rede que suporte serviços de telecomunicações, mesmo fazendo uso do modelo *Internet* [ITUM3200].

### 2.2.10 Rede de Acesso

Segundo [KAPP2003], uma rede de acesso pode ser definida como sendo o conjunto de equipamentos e meios destinados a interconectar um sistema autônomo a outro, e que deles não faça parte. Assim dizendo, duas centrais telefônicas, quando não pertencentes a uma mesma operadora, são ligadas via uma rede de acesso. Da mesma forma, dois roteadores, dois *FRADs* (*Frame Relay Access Device*), dois *modems*, quando não pertencentes a uma mesma empresa, são ligados via uma rede de acesso. A Figura 2.15 mostra dois sistemas autônomos típicos de um ambiente usuário-operadora, onde este cliente deseja interligar um computador pessoal de uma filial (Unidade Remota) à sua matriz (Unidade Principal).

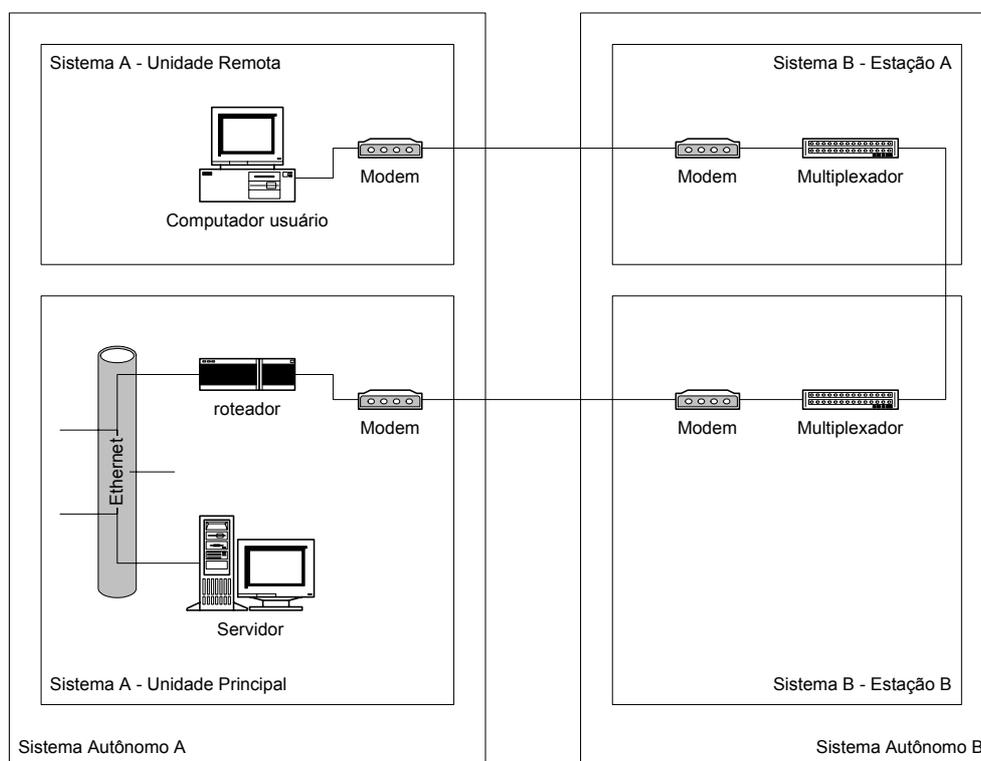


Figura 2.15. Ambiente Usuário-Operadora [KAPP2003]

Não é difícil identificar a rede de acesso neste exemplo como sendo o conjunto formado pelos *modems* e pelos pares de cobre que os interligam. O meio de transmissão poderia ser além do cobre, fibra óptica ou o ar, desde que fosse utilizado um modem óptico ou um rádio, mas o princípio continuaria sendo o mesmo.

Uma outra maneira de se identificar a rede de acesso seria identificá-la na camada de serviço. Um usuário que deseja se conectar a um serviço de Camada 2 do tipo *Frame Relay* ou *ATM*, teria sua rede de acesso identificada como o conjunto de meios de transmissão e

equipamentos entre seu *FRAD* (*Frame Relay Access Device*) e o *Switch Frame Relay* ou *ATM* da operadora fornecedora do serviço, não interessando a qual centro de fios este está ligado. Assim, da rede de acesso fariam parte inclusive estações do sistema autônomo da operadora que estivessem entre o serviço fornecido e o usuário.

Outra maneira de se identificar a Rede de Acesso é observar se ela é totalmente transparente ao serviço por ela transportado (ou acessado).

Nas operadoras, esta definição tem sofrido alterações baseadas no uso, e há quem chame a rede de acesso como sendo a rede formada por equipamentos exclusivamente da chamada última milha. De todas as redes componentes de um sistema de telecomunicações, a Rede de Acesso é a que possui maior diversidade de equipamentos e de fabricantes, tanto pela sua capilaridade como pelas constantes evoluções, tecnológicas, de demanda, e, principalmente, no tipo de serviços oferecidos (o que influi nas demais), sendo assim, considerada a mais heterogênea das redes.

### **2.2.11 Rede de Transporte**

Segundo [ITUM3200], a Rede de Transporte é formada pelo conjunto de equipamentos de transmissão destinados a interconectar as demais redes, tanto entre si, quanto em locais diferentes, com as outras redes. É a principal rede de qualquer operadora, pois dela todas as demais dependem. Falhas na rede de transporte provocam interrupções nas várias redes fornecedoras de serviços relacionados à função de transporte. Talvez, por este motivo, foram as primeiras redes a dispor de qualidade de serviço, Áreas Gerenciadas de Telecomunicações e sistemas redundantes tanto como rede quanto elemento de rede.

Os principais equipamentos que dela fazem parte são os equipamentos *SDH* (*Synchronous Digital Hierarchy*), *PDH* (*Plesiochronous Digital Hierarchy*), *DWDM* (*Dense wavelength division multiplexing*) e, para o caso de transporte de dados, algumas malhas *ATM*, emulando circuitos para a Rede Determinística ou Estatística. São, usualmente, de grande largura de banda.

## **2.3 REDES NEURAIAS**

Uma vez descritos os conceitos embaixadores sobre as redes relacionadas aos serviços de telecomunicações usualmente prestados pelas Operadoras, cabe, agora, um embasamento

teórico sobre a ferramenta utilizada no experimento descrito no capítulo 4: as Redes Neurais.

Redes Neurais são técnicas computacionais baseadas na forma altamente complexa, não-linear e paralela de operação do cérebro humano.

O cérebro é o centro do sistema nervoso humano, pois sempre está recebendo, nos receptores, os impulsos elétricos e convertendo-os em respostas, as quais podem ser consideradas como as saídas deste sistema.

A Figura 2.16 ilustra este funcionamento em forma de diagrama de blocos.

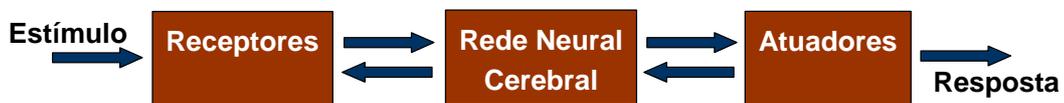


Figura 2.16 – Representação do sistema Nervoso em Diagrama de Blocos [CAST2001]

Conforme mostrado na figura Figura 2.17, os componentes estruturais do sistema nervoso, os neurônios, são formados pelos Dendritos, pelo corpo central e pelos Axônios. Os Dendritos são terminais de entrada e os Axônios são longos terminais de saídas codificadas como uma série de breves pulsos de tensão ou potenciais de ativação. Tais potenciais de ativação podem impor excitação ou inibição e são passados entre as células nervosas (Neurônios) através das conexões sinápticas, as quais conectam os Dendritos com sua membrana pós-sináptica e aos axônios com sua membrana pré-sinápticas. [BARR2002]



Figura 2.17 – Estrutura básica do neurônio [LABINFO2007]

De forma comparativa, no diagrama de blocos da Figura 2.18, é mostrado o modelo de um neurônio artificial, onde se tem:

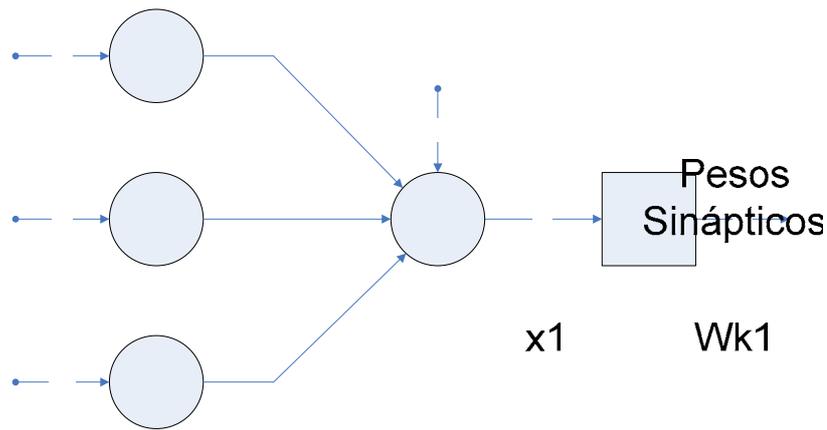


Figura 2.18 – Neurônio Artificial [Ribeiro 2002]

Sinais de entrada

$x_j$ , representando os sinais de entrada;

$W_{kj}$ , representando os pesos sinápticos do neurônio K;  $x_2$   $W_{k2}$

$v_k$ , representando a saída do combinador linear;

$b_k$ , representando o bias;

$\varphi(\cdot)$ , representando a função de ativação;

$y_k$ , representando o sinal de saída.

$x_m$   $W_{km}$

Um sinal de entrada  $x_j$  conecta-se a um neurônio  $k$  através de uma sinapse  $j$ . O combinador linear soma todos os sinais de entrada, multiplicados pelos respectivos pesos sinápticos  $W_{kj}$ , com a polarização externa ou bias, denotada por  $b_k$ , podendo ser positiva ou negativa. E a função de ativação é introduzida para limitar a amplitude do sinal de saída  $y_k$ .

De acordo com Castro apud [HAYKIN1996], um neurônio  $k$  pode ser descrito pelas equações:

$$v_k = \sum_{j=1}^m W_{kj} \cdot x_j \quad \text{Equação 2.2}$$

$$y_k = \varphi(v_k + b_k) \quad \text{Equação 2.3}$$

A função de ativação, que define o potencial da saída do neurônio, como foi descrito acima, pode ser de três diferentes tipos:

Função *Threshold* (Função Limiar), que permite a saída assumir apenas os valores 0 e 1 como descrito por:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad \text{Equação 2.4}$$

E a saída do neurônio para esta função de ativação é descrita por:

$$y_k = \begin{cases} 1, & \text{se } v_k \geq 0 \\ 0, & \text{se } v_k < 0 \end{cases} \quad \text{Equação 2.5}$$

Segue na Figura 2.19, um exemplo de uma função Limiar.

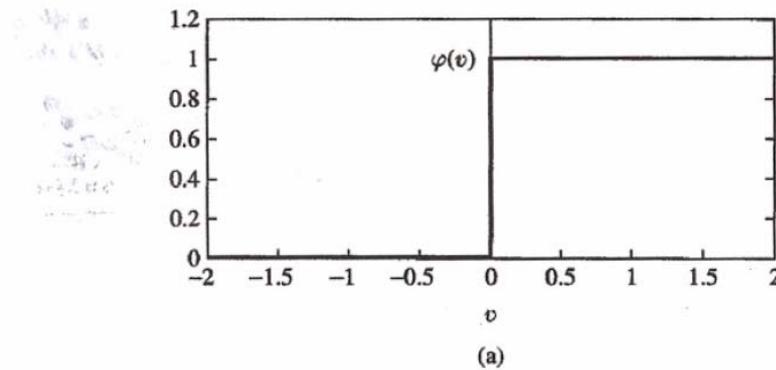


Figura 2.19 – Função Threshold ou Limiar [CAST2001]

Função *Piecewise-linear* (Linear por Partes), que pode ser vista como uma amplificação não linear, demonstrada por:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq \frac{1}{2} \\ \frac{1}{2}, & \text{se } -\frac{1}{2} < v < \frac{1}{2} \\ 0, & \text{se } v \leq -\frac{1}{2} \end{cases} \quad \text{Equação 2.6}$$

Segue na Figura 2.20, um exemplo da função Linear por Partes.

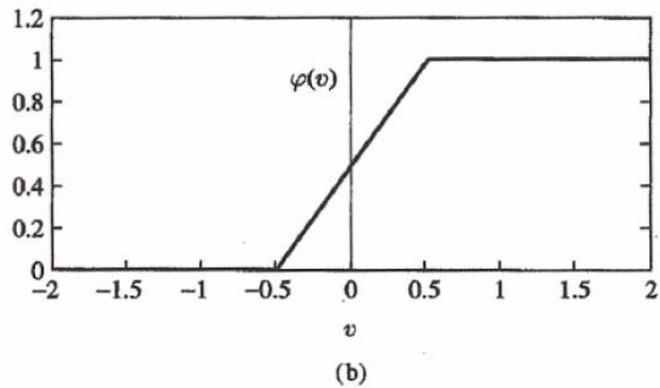


Figura 2.20 - Função Piecewise-linear ou Linear por Partes [CAST2001]

Função *Sigmoide*, que pode assumir diferentes declividades de acordo com a variação do parâmetro  $a$ .

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad \text{Equação 2.7}$$

Esta é a função de ativação mais usada das três e pode assumir valores negativos de acordo com a definição abaixo:

$$\varphi(v) = \begin{cases} 1 & \text{se } v > 0 \\ 0 & \text{se } v = 0 \\ -1 & \text{se } v < 0 \end{cases} \quad \text{Equação 2.8}$$

Na Figura 2.21, está representada a função Sigmoide.

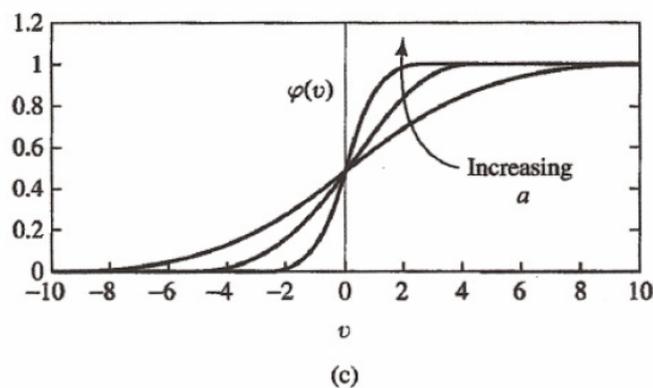


Figura 2.21 – Função Sigmoide [CAST2001]

Conforme [CAST2001], outra forma mais simples de representar o neurônio artificial é através de ramos orientados e interconectados aos nós, que são os grafos. Nos grafos, a direção que flui o sinal no ramo é representada por setas, o nó representa a soma algébrica de todos os sinais dos ramos terminados nele e todo sinal gerado no nó é transmitido para todos os ramos originados nele.

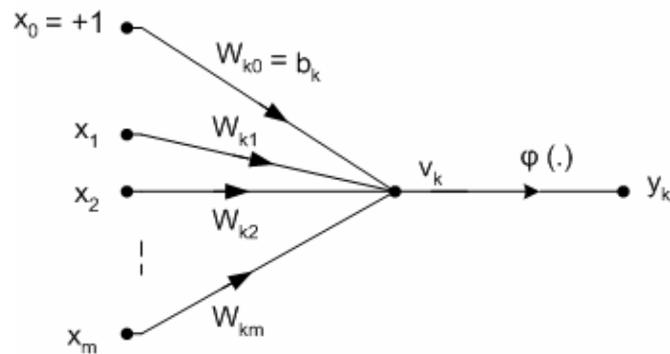


Figura 2.22 Representação de um neurônio Artificial por Grafos [CAST2001]

Pode-se denominar as Redes Neurais Artificiais (RNAs) como sendo um agrupamento organizado de sinapses que servem para desempenhar uma operação funcional específica, da mesma forma que um micro circuito neural.

Ou seja, segundo [CAST2001], uma Rede Neural Artificial (RNA) é uma estrutura computacional projetada para mimetizar a maneira pela qual o cérebro desempenha uma particular tarefa de seu interesse.

A capacidade de adaptação ao ambiente, através da criação de novas conexões sinápticas ou da modificação das sinapses entre neurônios existentes, permite que o cérebro construa suas próprias regras pelas experimentações contínuas. Este processo de aprendizagem também é possível nas Redes Neurais Artificiais por causa da mudança adaptativa dos parâmetros livres da rede feita pelo Algoritmo de Aprendizagem.

Outras características importantes e influenciadoras na escolha das RNAs para este trabalho, são:

- A possibilidade de ser uma rede linear ou não-linear, dependendo do comportamento dos sinais de entrada;

- Não necessita de muito conhecimento estatístico sobre o ambiente no qual a rede está inserida;
- Flexibilidade em ambientes extremamente dinâmicos;
- Tolerância à falhas de alguns neurônios.

Em projetos de redes neurais artificiais são identificadas três diferentes classificações de arquiteturas [SANT2001]:

Rede *Single-Layer Feedforward*, onde a camada de entrada é conectada a camada de saída, formada pelas unidades processadoras, os neurônios. Estas são redes progressivas de uma única camada.

Redes *Multilayer Feedforward*, as quais são redes progressivas compostas por uma ou mais unidades escondidas ou neurônios escondidos, classifica um espaço dimensional maior.

Redes Recorrentes, as quais não são progressivas por possuírem pelo menos um *loop* de realimentação, tendo uma grande melhoria na aprendizagem e no desempenho da rede.

### **2.3.1 Aprendizado nas Redes Neurais**

Segundo [CAST2001], a melhora no desempenho de uma RNA ocorre ao longo do aprendizado, de acordo com critérios específicos para atingir tal propósito. Ou seja, Aprendizagem é um processo pelo qual os parâmetros livres de uma RNA são adaptados através de um processo de estimulação do ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela forma através da qual é efetuada a mudança dos parâmetros.

A maneira de aprender das RNAs depende das regras de treinamento, que determinam como os pesos das conexões são ajustados. Sendo assim as RNAs são capazes de formular regras a partir de dados reais, diferindo da computação programada visto que nestas é necessário um conjunto de regras pré-definidas e algoritmos.

São destacados os seguintes tipos de aprendizado:

### 2.3.1.1 Aprendizado Supervisionado

No aprendizado supervisionado tem-se a intervenção de um tutor, que detém o conhecimento do ambiente e é dispensado quando este conhecimento é transferido para a RNA. Para tanto a rede deve possuir entradas associadas a saídas desejadas, a fim de que sejam conferidas com as saídas obtidas e, se não forem coincidentes o algoritmo de aprendizagem ajusta os pesos, armazenando o conhecimento adquirido.

O Aprendizado por Correção de Erro supervisionado pelo tutor pode ser representado como um ponto sobre a superfície de erro, e este deve se mover em direção ao ponto mínimo global para que o desempenho da rede seja melhorado. Assim, após medir a distância entre a resposta atual e a desejada, denotada por  $d_k(n)$ , ajustam-se os pesos das conexões, através da Regra Delta, de modo a atingir o ponto mínimo global, diminuindo a distância entre as respostas medidas. [BARR2002]

Na Figura 2.23, pode-se entender melhor o funcionamento deste aprendizado.

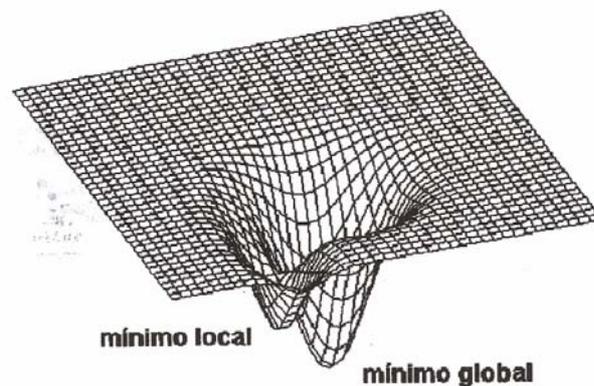


Figura 2.23 – Superfície de Erro apresentando um ponto mínimo local e um ponto mínimo global.

[CAST2001]

O sinal de erro  $e_k(n)$  é representado como:

$$e_k(n) = d_k(n) - y_k(n) \quad \text{Equação 2.9}$$

Se o erro não for satisfatoriamente pequeno, aplica-se o ajuste  $\Delta w_{kj}(n)$  ao peso  $w_{kj}$ , definido por:

$$\Delta w_{kj}(n) = \eta e_k(n) \cdot x_j(n) \quad \text{Equação 2.10}$$

Onde  $\eta$  determina positivamente a razão de aprendizado, devendo ser cuidadosamente escolhido. Conseqüentemente o valor atualizado do peso sináptico será:

$$\Delta w_{kj}(n+1) = \Delta w_{kj}(n) + \Delta w_{kj}(n-1) \quad \text{Equação 2.11}$$

### 2.3.1.2 Aprendizado Reforçado

O Aprendizado Reforçado ou “*Reinforcement learning*” consiste no aprendizado através do método da tentativa e erro, de modo a otimizar um índice de performance chamado sinal de reforço, sem a utilização de um tutor, mas sim de um juiz.

Este sistema tem profunda influência biológica, pois os sinais de reforço primários satisfatórios, recebido pelo ambiente, reforçam as conexões que os produziram, e aqueles insatisfatórios modificam o valor das correspondentes conexões. [CAST2001]

### 2.3.1.3 Aprendizado Não-Supervisionado

No aprendizado não-supervisionado ou auto-supervisionado ou ainda auto-organizado não há saídas desejadas e por isso não precisa de um tutor externo ou crítico. No processo de treinamento a rede utiliza apenas os valores de entrada, que são classificados de acordo com o seu algoritmo de aprendizado.

No aprendizado baseado em Memória, os algoritmos de aprendizado respondem através da análise da vizinhança local do vetor a ser classificado pela primeira vez, definindo qual é a vizinhança local e a qual regra de aprendizagem deve ser submetida.

A regra do vizinho mais próximo e sua variante a regra dos k-vizinhos mais próximos definem a vizinhança local do ponto de teste. Na regra do vizinho mais próximo a vizinhança é determinada através da mínima distância Euclidiana entre os vetores  $x_i$  e  $x_{teste}$ . Em termos matemáticos, tem-se:

$$\min dd(x_i, x_{teste}) = d(x_N, x_{teste}) \quad \text{Equação 2.12}$$

Na regra dos k-vizinhos mais próximos define-se o inteiro k. Este valor é o número dos vizinhos avaliados, dentre estes atribui-se o valor de maior ocorrência ao vetor de teste, reconhecido por ser um ponto com valor diferente dos demais [CAST2001].

#### 2.3.1.4 Aprendizado Competitivo

Neste aprendizado, ocorre uma competição baseada na força de cada neurônio do conjunto de neurônios aleatórios e iguais, exceto por alguns pesos sinápticos, do conjunto padrão de entrada.

O neurônio vencedor chamado de neurônio *winner-takes-all*, é o que demonstrou ter maior potencial de ativação para o padrão de entrada específico. Então o sinal de saída deste vencedor recebe o valor 1 e os demais, que perderam a competição, recebem o valor 0 [CAST2001].

#### 2.3.1.5 Aprendizado Hebbiano

Segundo [CAST2001] a idéia central da teoria de Hebb está na afirmação, datada de 1949, e feita no livro “*The Organization of Behavior*”, escrito por Hebb:

“Se um neurônio B é repetidamente estimulado por outro neurônio A, ao mesmo tempo em que ele está ativo, ele ficará mais sensível aos estímulos de A, e a conexão sináptica de A para B será mais eficiente. Deste modo, A achará mais fácil estimular B para produzir uma saída”.

[CAST2001] interpreta o postulado de Hebb da seguinte maneira: se dois neurônios em cada um dos lados de uma sinapse (conexão) são ativados simultaneamente, então a força daquela sinapse é seletivamente aumentada.

Para formular a idéia de que se duas unidades são ativadas simultaneamente, suas interconexões tendem a se fortalecer também, considerando-se o ajuste do peso sináptico  $\Delta W_{kj}$ , no intervalo de tempo  $n$ , com os sinais pré-sinápticos e pós-sinápticos denotados respectivamente por  $x_j$  e  $y_j$ . Com  $\eta$  sendo a função dos sinais pré e pós, o ajuste aplicado é expresso na forma geral:

$$\Delta W_{kj}(n) = \eta(x_k(n), y_k(n)) \quad \text{Equação 2.13}$$

Substituindo os sinais pré e pós sinápticos pelos seus respectivos valores médios sobre um determinado intervalo de tempo, tem-se a hipótese da Covariância descrita por:

$$\Delta W_{kj}(n) = \eta \cdot \lambda(x_j - \bar{x})(y_j - \bar{y}) \quad \text{Equação 2.14}$$

Uma vez descrito matematicamente o aprendizado Hebbiano cabe uma correlação entre este e o tema proposto deste trabalho.

A escolha de um particular algoritmo de aprendizado é motivada pela tarefa que se deseja que a RNA desempenhe. Uma das tarefas de aprendizagem é a Aproximação de Funções. Aproximação de Funções é um processo através do qual a RNA apresenta uma solução para o mapeamento entrada/saída de interesse, cuja função que o descreve tem sua expressão analítica desconhecida. A rede projetada deve aproximar a função desconhecida, de tal forma que o mapeamento entrada/saída efetivamente realizado pela rede aproxime a função para todas as entradas.

Para construir o conhecimento sobre o mapeamento entrada/saída, a rede é provida de um conjunto de exemplos representativos deste mapeamento. Desde que o tamanho do conjunto seja suficientemente grande e a rede disponha de um número adequado de parâmetros livres, o erro de aproximação obtido será relativamente pequeno para tarefa proposta.

Considera-se o aprendizado Hebbiano o mais adequado para aplicar no contexto deste trabalho, uma vez que ele retrata a influência dos eventos entre si, por exemplo: o tempo de provisionamento de um serviço em uma filial pode influenciar o tempo de provisionamento em outra filial e quanto mais esta influência ocorrer mais fácil ela poderá induzir nos tempos de provisionamento entre as duas filiais.

No próximo item será explicado o modelo para decomposição de um espaço vetorial em subespaços. Onde o aprendizado Hebbiano será aplicado.

### **2.3.2 Redes Neurais Artificiais para Decomposição de um Espaço Vetorial em Sub-Espaços**

Segundo [CAST2001], a análise dos componentes principais (*Principal Components Analysis – PCA*) ou decomposição em subespaços (DSE) é aplicada com o objetivo de reduzir a dimensionalidade dos dados nas RNAs.

No processo de Análise dos Componentes Principais, utiliza-se o Algoritmo Hebbiano Generalizado baseado no aprendizado auto-organizado que será detalhado neste tópico.

No aprendizado não supervisionado ou auto-organizado, a rede cria padrões de atividade de acordo com os diferentes sinais de entrada e modifica os pesos sinápticos da rede em resposta aos padrões de atividade criados. Dessa forma a rede atinge um estado de ordem global a partir de interações locais, pois a mudança do peso sináptico de um neurônio é limitada à sua vizinhança.

Como já dito no item 2.3.1.5, se dois neurônios de lados opostos da uma sinapse são ativados simultaneamente, a força da sinapse é aumentada. Assim os aumentos de sinapses ocorrentes nas RNAs devem ser compensados com decréscimos em outras sinapses.

A redução dimensional dos sinais de entrada é feita pela Transformação *Karhunen-Loève* (*KLT*) que resulta em uma matriz de  $M$  auto-vetores associados aos autovalores, representando as  $M$  direções principais ao longo das quais a variância é máxima. Lembrando que a variância de uma variável aleatória é uma medida de sua dispersão estatística, indicando quão longe em geral os seus valores se encontram do valor esperado (média).

Para chegar à matriz de  $M$  auto-vetores, a *KLT* projeta o conjunto de vetores de entrada sobre uma base ortonormal de auto-vetores, orientada pela direção de maior variância, descartando os subespaços de menor variância. Assim a variância dos subespaços é a máxima variância da projeção realizada em todas as possíveis direções.

As variações no auto-vetor  $\underline{e}$ , que indicam a direção de projeção, podem influenciar no resultado de máximo  $f(\underline{e})$ . Por isso duas condições devem ser satisfeitas: as variações em  $\underline{e}$  devem ser ortogonais durante o processo de busca da direção de máximo  $f(\underline{e})$  e a norma do vetor influenciado  $\underline{e} + \delta\underline{e}$  deve permanecer unitária.

A título de interpretação da *KLT*, caso fosse de interesse obter a *KLT* através de um processo manual e experimental, tomar-se-ia um vetor arbitrário de módulo unitário  $\underline{e} \in \mathbb{R}^M$  com origem em  $0 \in \mathbb{R}^M$ . Este vetor definiria uma direção arbitrária na qual o conjunto  $U$  de vetores de dados,  $u_i \in \mathbb{R}^M$ , seria projetado.

Todo o conjunto  $U$  de vetores de dados seria projetado na direção dada por  $\underline{e}$  e seria medida a variância  $\lambda$  da projeção. Após serem testadas todas as direções possíveis no espaço  $\mathbb{R}^M$ , deverá ser selecionada aquela que apresentar maior variância. Esta será  $\lambda_0$ .

O vetor  $\underline{e}$ , que define tal direção, é igual ao auto-vetor  $\underline{e}_0$  associado ao maior autovalor, obtidos pela *KLT*. Com o valor do maior autovalor dado por  $\lambda_0$ .

O processo é repetido novamente para a obtenção do segundo maior autovalor, que será chamado  $\lambda_1$ . Com restrição de que a busca da direção de maior variância em  $R^M$  sempre seja feita em direções ortogonais à do auto-vetor  $\underline{e}_0$  associado ao maior autovalor  $\lambda_0$  recém determinado e assim sucessivamente para os demais autovalores que serão determinados.

Ou seja, a busca da direção de maior variância em  $R^M$  para obtenção do  $n$ -ésimo autovalor  $\lambda_n$  é feita com a restrição de que as direções testadas sejam ortogonais às direções dos  $n-1$  auto-vetores associados aos maiores autovalores  $\lambda_{n-1}$  previamente encontrados.

Assim, o processo é recursivo até que os  $M$  auto-vetores e autovalores sejam determinados.

Na *KLT* realiza-se uma compressão com perdas e uma redução dimensional através da supressão dos subespaços de menor energia.

O Aprendizado Hebbiano, já mencionado anteriormente, é um aprendizado auto-organizado, aonde o conhecimento vai sendo adquirido através de observações repetidas de parâmetros estatísticos e os pesos vão sendo aumentados de acordo com o definido. Mas o constante aumento do peso sináptico pode levar a rede à saturação. Assim, [OJA1992] propôs adicionar um decremento proporcional  $y_0^2(n)$  ao peso sináptico como forma de restringir o crescimento do peso.

Em uma RNA, formada com todos os processos descritos, são armazenados os seus pesos sinápticos, os auto-vetores e autovalores da matriz de covariância do conjunto de dados de entrada preservados. E o número da camada de saída é definido pelo número de componentes que se deseja extrair do conjunto de dados representativos.

A Figura 2.24 representa os pesos sinápticos de uma RNA.

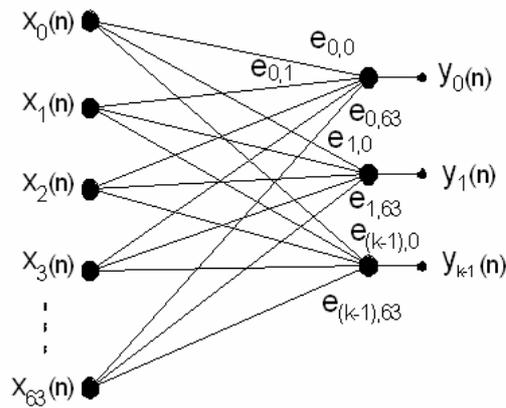


Figura 2.24 Modelo de RNA 64 nós de entrada e K neurônios de saída [CAST2001]

Outro problema encontrado resultante da variação dos pesos sinápticos é a possível estabilização dos pesos em valores correspondentes aos auto-vetores e auto-valores associados, gerando uma convergência do algoritmo. Para resolver este problema foi desenvolvido o Critério para Atualização da Razão de Aprendizado baseado no aumento da razão de aprendizado, quando ocorre a convergência da primeira componente do vetor de teste.

O princípio da deflação diz que quando um neurônio converge, todos os subsequentes tendem a convergir também. Assim para evitar processos redundantes de aumento de aprendizado, é proposto o GHAPCA (Algoritmo Hebbiano Generalizado para Extração dos Componentes Principais de um Conjunto de Dados no Domínio Real) um algoritmo de janelamento de convergência. O detalhamento do algoritmo encontra-se no Apêndice deste trabalho. Este janelamento consiste no deslocamento da janela de convergência para frente depois da convergência do primeiro auto-vetor [DECASTRO1996].

### 2.3.3 Mapas Auto-Organizados de Kohonem - SOM

As RNAs *SOM* (*Self-Organizing Map*) são algoritmos de aprendizado não-supervisionado que foram desenvolvidos para construir uma matriz de neurônios ou mapa computacional com o intuito de facilitar a visualização e identificação das similaridades em grandes bancos de dados. Estas, através da simulação do funcionamento do cérebro na armazenagem de conhecimento e no desempenho de tarefas, resolvem problemas não-lineares de alta dimensionalidade.

Por este motivo diversos campos da ciência adotaram os *SOMs* como uma ferramenta analítica padrão. Dentre estes se encontram, entre outras, a estatística, o processamento de sinais, a teoria de controle, a análise financeira, a física experimental, a química e a medicina. Um exemplo de utilização em Telecomunicações seria a atribuição de importância, ou não, a parâmetros de alarmes de equipamentos durante a correlação entre estes alarmes.

O objetivo maior do *SOM* é a ordenação e o arranjo entre os neurônios envolvendo os processos de competição, cooperação e adaptação sináptica. Um determinado padrão de entrada, quando apresentado à rede, ativa regiões da camada de neurônios e reforça as conexões das camadas de entrada com as regiões ativadas. Se um determinado neurônio é ativado, provavelmente outros em sua vizinhança serão ativados e dessa forma gera-se um arranjo baseado em localização.

Assim, o modelo apresentado envolve a adaptação das ponderações das conexões em reflexo ao aprendizado obtido, tendo todos os vetores de entrada ligados a um arranjo de neurônios normalmente unidimensional ou bidimensional, como o mostrado na Figura 2.25.

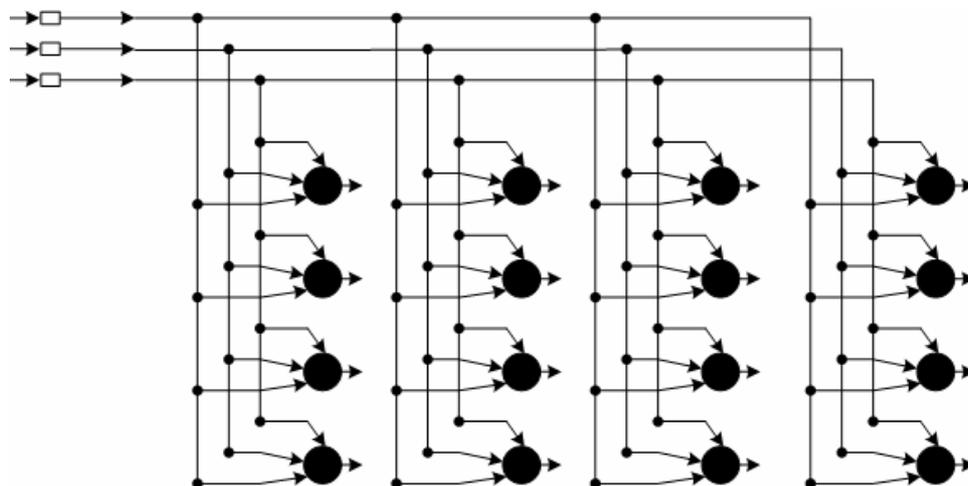


Figura 2.25 – Treliça de Neurônios [CAST2001]

Cada neurônio da treliça é completamente conectado a todos os nós fonte na camada de entrada. A rede da Figura 2.25 representa uma estrutura progressiva, com uma única

camada computacional, consistindo de neurônios arranjados em linhas e colunas [CAST2001].

Cada padrão de entrada apresentado à rede consiste de uma região localizada de atividade. Todos os neurônios na rede devem, portanto, ser expostos a um número suficiente de diferentes realizações dos padrões de entrada, para garantir que o processo de auto-organização ocorra de forma apropriada.

Depois de comparados os produtos internos dos dados de entrada com os respectivos pesos sinápticos dos neurônios da rede, é eleito um neurônio vencedor, que é o maior produto interno, e a localização da vizinhança que será excitada. Dessa forma as redes *SOM* são baseadas no aprendizado competitivo.

O neurônio vencedor localiza-se no centro da vizinhança provendo, desta forma, as bases para a cooperação entre tais neurônios vizinhos, conforme Figura 2.26 que representa uma estrutura de vizinhança retangular.

De acordo com [CAST2001] e [DECASTRO1996], a melhor forma de representar matematicamente a vizinhança topológica é pela função Gaussiana variante no tempo:

$$h_{L1\sigma}(n) = \exp(-d_{L1}^2 / 2\sigma^2(n)) \quad \text{Equação 2.15}$$

Onde o parâmetro  $\sigma$  representa a largura efetiva da vizinhança e mede o grau com que os neurônios excitados e vizinhos do neurônio vencedor participam no processo de aprendizagem (variância). E  $d_{j,i}$  é a distância lateral, que decai com a vizinhança topológica ao redor do neurônio vencedor, pois a região de vizinhança deve diminuir gradativamente para que a resposta da rede se torne mais localizada.

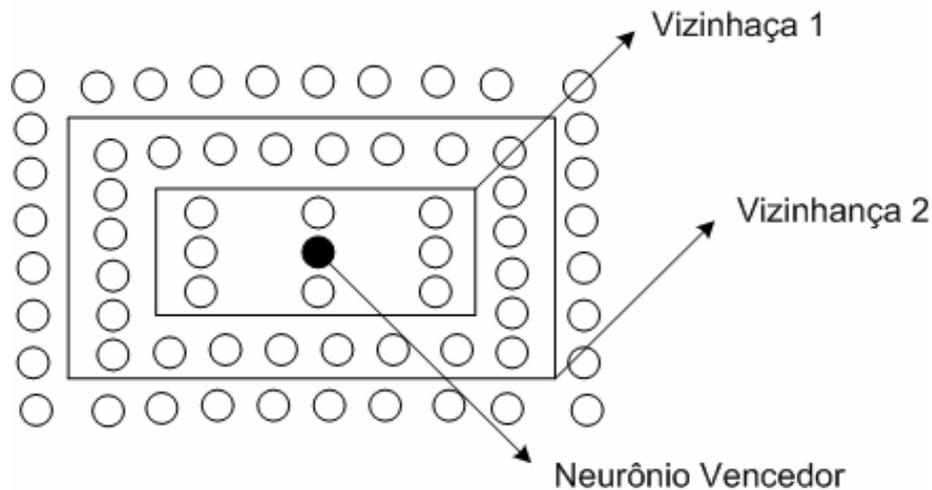


Figura 2.26 – Vizinhança do Neurônio Vencedor [BENITE2003]

O Processo Adaptativo utilizado nas RNAs *SOM* é uma adaptação do postulado de Hebb – item 2.3.1.5. Isto é, permite aos neurônios excitados aumentar seus valores individuais da função discriminante em relação ao padrão de entrada, através de ajustes adequados aplicados a seus pesos sinápticos. Os ajustes feitos são tais que a resposta do neurônio vencedor à subsequente aplicação de um padrão similar de entrada é realçada.

Inclui-se o termo  $g(y_i)\underline{w}_j$ , onde  $\underline{w}_j$  é o vetor de pesos sinápticos de neurônio  $j$  e  $g(y_i)$  é uma função escalar positiva da resposta  $y_j$ , então a mudança no vetor de pesos sináptico do neurônio  $j$  na treliça pode ser expresso como:

$$\Delta \underline{w}_j = \eta y_j \underline{x} - g(y_i) \underline{w}_j \quad \text{Equação 2.16}$$

Onde,  $\eta$  é o padrão razão de aprendizado e  $g(y_i)$  é:

$$g(y_i) = \eta y_j \quad \text{Equação 2.17}$$

$$y_j = h_{j,i(x)} \quad \text{Equação 2.18}$$

O parâmetro de aprendizado deve decair no decorrer do tempo, mas não deve assumir o valor zero para não apresentar um defeito topológico. O ideal é iniciar com um valor próximo a 0,1 e permanecer acima de 0,01.

Resumindo, para a aplicação do algoritmo em estudo, deve-se ter: uma inicialização, onde são escolhidos valores aleatórios; uma definição de amostragem, associando a amostra do

espaço de dados de entrada e associando-a a probabilidade de ocorrência; e a finalização com a definição de vizinhança junto com seu aprendizado.

Depois de repetidas passagens de dados de teste, os vetores de pesos sinápticos tendem a seguir a distribuição dos vetores de entrada, devido à atualização da vizinhança. O algoritmo, portanto, conduz a uma ordenação topológica do mapa de características no espaço de entrada, no sentido de que os neurônios que são adjacentes na treliça tenderão a ter vetores de pesos sinápticos similares [CAST2001].

### **3 ANÁLISE CONTEXTUAL DOS COEFICIENTES UTILIZADOS NO MODELO DE PREDIÇÃO**

O capítulo anterior estabeleceu os conceitos de Redes de Telecomunicações, com ênfase nas redes comparadas por este trabalho e Redes Neurais, também com mais importância ao aprendizado dito Hebbiano.

Dentro do objetivo de se montar curvas comparativas entre os serviços transportados por Redes *IP* e não *IP*, ainda falta o estabelecimento de critérios de comparação entre estes dois conjuntos de redes e, uma vez estabelecidos estes critérios cabe um explicativo do Porquê de se usar uma Rede Neural para a esta comparação.

Uma vez cumpridas estas etapas, o experimento pode ser, enfim, descrito no capítulo 4.

Assim, este capítulo descreve os critérios para popular a base de dados do modelo de comparação.

Para simplificar a comparação entre as redes, uma vez que o que se busca é descobrir o tempo estimado onde os serviços passarão a serem prestados por uma única rede, são definidos os dois conjuntos que serão comparados a partir destes critérios, as Redes *IP* e as Redes não *IP*.

Considerou-se como parte do conjunto de redes não *IP* todos os serviços relacionados às redes de camadas 1 e 2, tais como: serviços sobre *ATM*, *Frame Relay* e redes Determinísticas. A rede *IP*, por sua vez, agrupou todos os serviços de camada 3 por ela transportados.

#### **3.1 CRITÉRIOS DE COMPARAÇÃO**

Considerando que, segundo [PAIVA2004], critérios competitivos são definidos como um conjunto consistente de prioridades que uma empresa elege para competir no mercado. Ou seja, em função das competências internas da empresa, do tipo de mercado onde ela deseja atuar, do grau de concorrência desse mercado e do tipo de produto ou serviço que produz, ela terá que priorizar alguns critérios aproveitando seus recursos, competências e oportunidades de mercado com o intuito de se tornar competitiva.

Desta forma, [PAIVA2004] elege cinco critérios ditos competitivos na área de operações que se relacionam à estratégia de negócios. O primeiro deles considera que a empresa pode competir baseada em **custo**, ou seja, a empresa pode produzir com margens de lucros maiores ou produzir grandes quantidades com margens menores.

No entanto, a empresa pode optar por competir com base na **qualidade** dos produtos, ou seja, de forma que seus produtos tenham desempenho superior ao de seus concorrentes. É importante distinguir o que é qualidade real e qualidade percebida. A segunda mais relacionada a *marketing* e publicidade.

A terceira possibilidade é o **desempenho na entrega**, isto é, a relação entre fornecedor e cliente. A empresa deve ser capaz de mobilizar recursos para garantir a entrega no prazo prometido ou ter prazo de entrega mais curto que dos competidores.

Um outro critério é a **flexibilidade**, tanto para *mix* de produtos como para volume produzido. Uma empresa que compete com base em flexibilidade deve ser capaz de absorver rapidamente mudanças em lotes de produção não-padronizados, bem como a mudanças no tipo de produto a produzir.

O quinto e último critério refere-se à **inovação**. Este critério historicamente está relacionado a habilidade que a empresa possui em lançar novos produtos no mercado em curto espaço de tempo.

Em outras palavras as empresas devem eleger critérios que as qualifiquem para estar no mercado, ou seja, seus produtos devem satisfazer um padrão mínimo. Enquanto que, para vencerem a concorrência é preciso “fazer melhor” para aumentar a competitividade e ganhar mercado.

O conceito de “fazer melhor” segundo [RIPOLL1993] significa:

- Fazer certo: vantagem de **qualidade**.
- Fazer rápido: vantagem de **velocidade**.
- Fazer pontualmente: vantagem de **confiabilidade**.
- Mudar o que está sendo feito: vantagem de **flexibilidade**.

- Fazer barato: vantagem de **custo**.

Utilizando-se do conceito de “fazer melhor” para competir, estabelecido por [RIPOLL1993], foi feito um paralelo para o contexto de Redes de Comunicação de Dados, pelo uso comum, um subconjunto das Redes de Telecomunicações, composto por todas as redes que não foram concebidas para transportar unicamente voz..

Neste novo contexto, entendem-se como critérios competitivos o conjunto de prioridades eleitas por um tipo de Rede quando comparada a outra. Ou seja, em função das capacidades da rede ou eficiência, do tipo de serviço que irá trafegar por ela, do grau de concorrência entre as tecnologias, as redes terão competências que se destacarão, gerando melhor aproveitamento do recurso disponível, tornando-as mais ou menos competitivas.

Prioritariamente, tanto para comparação entre empresas como para comparação entre redes, no lugar de tentar atender todas as dimensões simultaneamente, deve-se enfatizar a combinação de alguns critérios. Isto porque deve ser observada a questão dos *trade-offs* entre os critérios, ou seja, as situações onde a melhoria de um critério poderá implicar em impacto negativo no outro.

Assim, seguirá uma abordagem dos critérios competitivos contextualizados para o presente trabalho.

No entanto, antes de partir para abordagem de tais critérios é importante descrever o processo, adotado pelas Operadoras em geral, de entrega de um serviço de Comunicação de Dados qualquer, desde o pedido do cliente até a entrega propriamente dita.

No fluxograma da Figura 3.1 observa-se quatro etapas bem definidas do processo: primeiramente o cliente solicita um serviço à área Comercial, que internaliza o pedido verificando viabilidade e retornando com uma proposta comercial. Depois de concluída a negociação comercial é gerada a Ordem de Serviço. A Ordem de Serviço gerada dispara a instalação do serviço solicitado pelo cliente. Neste momento a célula de provisionamento irá executar várias tarefas relevantes para configuração do serviço na rede, como a montagem da topologia, alocação de facilidades de rede, configuração da rede de comunicação de dados, etc. Após esta etapa são geradas as atividades (bilhetes de atividades) para os técnicos de campo concluírem a instalação. Esta última etapa é padronizada e os *SLA's* (*Service Level Agreement*) assegurados por contrato.

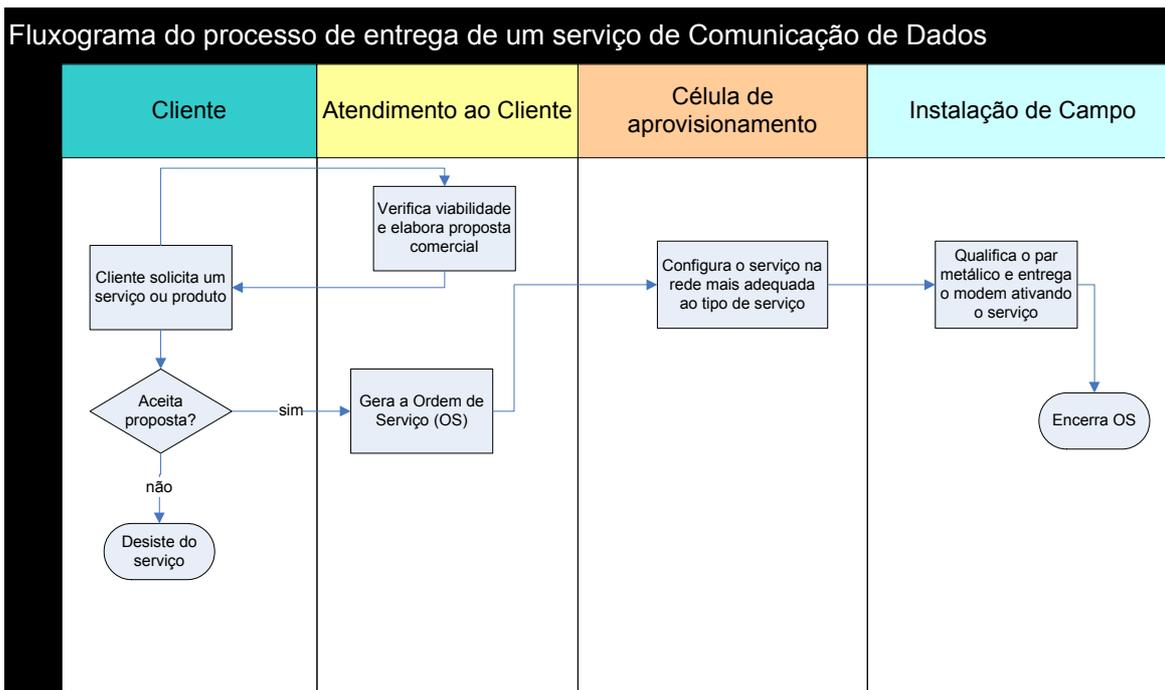


Figura 3.1 – Fluxograma de Entrega de um Serviço de Comunicação de Dados [AUTORA]

Conforme dito acima, segue o detalhamento dos critérios contextualizados para as Redes de Telecomunicações.

### 3.1.1 Qualidade

Existe crescente consciência de que serviços de alta qualidade podem fornecer vantagem competitiva a uma empresa.

Conforme esquematizado na Figura 3.2, de um lado estão os produtos e serviços com os quais a empresa trabalha e que para serem entregues consomem determinado recurso de produção. Enquanto do outro lado está o consumidor, quem demanda os produtos. Para [SLACK1996], investir em planejamento e controle de qualidade pode levar a redução de custos. Principalmente aqueles ligados ao retrabalho, refugo, devoluções e/ou *churn* de clientes. E ainda, e mais importante boa qualidade gera consumidores satisfeitos.

Do ponto de vista de Redes, investir em planejamento de qualidade também pode trazer benefícios. Redes mais bem projetadas geram menos retrabalho, consumidores mais satisfeitos com o serviço e redução de *churn* de clientes.

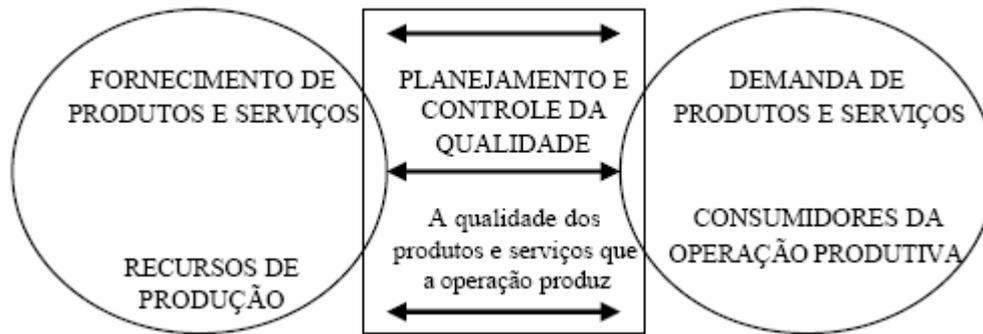


Figura 3.2 – Planejamento e o controle da qualidade [SLACK1996]

Na verdade, é difícil associar “qualidade” a um conceito único e fechado. O conceito que melhor parece adequar-se ao contexto deste trabalho, seguindo a própria tendência de diversos autores, é o conceito proposto por [GIANESI1996]: qualidade pode ser definida como o grau em que as expectativas dos clientes são atendidas ou excedidas por sua percepção do serviço prestado.

No entanto, segundo Garvin apud [PAIVA,1994], existem pelo menos oito dimensões da qualidade, que são: desempenho ou *performance*, *features* ou características secundárias, confiabilidade, conformidade, durabilidade, serviços agregados, estética, qualidade percebida. Sendo possível identificar uma ou mais dimensões como oportunidade de mercado para empresa competir.

Quando se comparam Redes de Comunicação de Dados, há, pelo menos, dois critérios que se destacam sob o aspecto de competitividade: a *Performance* ou Desempenho e a Confiabilidade.

#### 3.1.1.1 *Performance* ou Desempenho

[PAIVA2004] refere-se às características operacionais primárias, envolvendo atributos mensuráveis de um produto ou serviço, e em sua grande maioria estão relacionadas às funções básicas para as quais o produto foi originalmente projetado.

Neste quesito, os serviços são construídos segundo uma topologia de rede, para a qual se calcula a eficiência de transmissão. Este valor serve de comparativo de desempenho entre redes. Este critério será mais bem detalhado no item 3.1.1.1.2 abaixo.

### 3.1.1.1.1 Modelo de Cálculo Comparativo para Eficiência de Enlace (Transmissão) entre Produtos

Por definição, eficiência é a virtude de produzir um efeito desejado ou um bom resultado, isto é, um sistema é mais eficiente quando altera o mínimo possível na saída as características da informação de entrada.

Extrapolando esta definição para um sistema de transmissão, não seria errado dizer que um sistema é mais eficiente quando agrega menos banda à realmente necessária para transmissão da informação. Isto é, eficiência de transmissão é a relação entre o tamanho de um quadro ou pacote transmitido em uma rede, ou em várias redes, desde a origem até o destino, com o tamanho do *payload* deste quadro ou pacote. Para redes *IP*, por exemplo, quando um conjunto de dados é empacotado para ser transmitido, é adicionado a este um *overhead*, que consiste nos endereçamentos de origem e destino, entre outros *bits* de controle do protocolo, isto é, são os dados efetivos do *payload* que representam os dados de interesse do usuário ao final da transmissão em relação ao total transmitido.

Para este trabalho, onde se deseja comparar os serviços *IP* e os serviços não *IP*, considera-se um sistema mais eficiente àquele que introduzir menos mecanismos de controle e, por conseguinte, menos banda à informação transmitida.

Igualmente, se considera esta eficiência como sendo a razão entre a banda necessária para transmitir informação na camada de Aplicação do modelo *OSI* e a efetivamente utilizada na camada Física.

Segundo [CISCO2004a], pode-se calcular a banda necessária para transmissão de voz tanto sobre um canal *TDM* quanto sobre um canal *VoIP* (*Voice over IP*) definindo-se o Codec utilizado, o tamanho da amostra e a quantidade de cabeçalho (*overhead*) em todas as camadas. Também se considera a possibilidade de compressão de cabeçalho, aspectos de segurança e tunelamento.

A diferença básica entre o *TDM* e o *VoIP* é que o *TDM* está diretamente sobre a camada 2, enquanto o *VoIP* necessita de outras camadas para seu funcionamento.

Ainda segundo [CISCO2004a], o total de banda necessária para transmissão de voz sobre um canal é mostrado na equação Equação 3.1:

$$Larg. banda = \left[ \frac{Cab.camada2 + Cab.camada3e4 + Tam.amostra}{Tam.amostra} \right] \times Codec \quad \text{Equação 3.1}$$

Sendo:

**Larg. Banda** é a banda total utilizada na camada 1;

**Cab. Camada 2** é o tamanho em bytes do maior cabeçalho utilizado na camada 2. Para o caso de células *ATM*, deve-se considerar o somatório de todos os cabeçalhos para o transporte do pacote de camada 3;

**Cab. Camada 3 e 4** é o tamanho em bytes da soma dos cabeçalhos utilizados nas camadas 3 e 4;

**Tam. Amostra** é o *payload* de informação. Por exemplo: para o Codec G.711 [ITU.711] onde a largura de banda necessária é de 64Kbps, considerando um tempo de espera de 20ms o tamanho da amostra é de 160 bytes.

**Codec**, por definição, converte sinais analógicos em sinais digitais para transmissão de dados na rede. Para equação em questão, trata-se da banda após a conversão do sinal digital.

Segue na Tabela 3.1, todos os tipos de Codec e as respectivas larguras de banda requeridas para cada um:

Tabela 3.1 – Codec vs Largura de Banda [CISCO2004a]

Codec	Largura de Banda (Kbps)
G.711	64
G.726r32	32
G.726r24	24
G.726r16	16
G.728	16
G.729	8
G.723r63	6,3
G.723r53	5,3

Assim, pode-se considerar a eficiência de transmissão ou de enlace para transmitir informação de voz como sendo a razão entre a largura de banda e a velocidade do Codec [CISCO2004a].

### 3.1.1.1.2 Cálculo da Eficiência de Transmissão

Neste item pretende-se calcular a eficiência de transmissão para cada um dos principais serviços de comunicação de voz.

Segundo [CISCO2004a], quando se deseja comparar a eficiência de transmissão entre redes diferentes só faz sentido fazê-lo quando comparadas dentro de um mesmo serviço. Como este trabalho busca comparar redes *IPs* e redes não *IPs*, o serviço utilizado para esta comparação foi o serviço de transporte de voz, pelos seguintes motivos: é o único serviço transportado por todas as redes com os mesmos requisitos de qualidade, é um serviço que demanda amostras pequenas, logo, exige grande quantidade de *overhead* sendo, talvez, um dos piores eficiências de transmissão.

Para tanto, levantou-se a topologia dos serviços usualmente oferecidos pelas Operadoras de Telecomunicações e calculou-se a eficiência de cada trecho conforme o caminho percorrido pelo circuito de acordo com a Figura 3.3.

Esta figura é um diagrama das camadas envolvidas durante o transporte da amostra de voz. A amostra de voz é digitalizada, através dos Codecs G.711 ou G.729 sendo, na prática, os Codecs mais utilizados pelas Operadoras que prestam este tipo de serviço. De forma geral as Operadoras utilizam G.711 porque devem permitir a passagem da sinalização de FAX (*Facsimile*) e URA (Unidade de Resposta Audível) de modo transparente. Já o G.729 mesmo não permitindo as sinalizações acima, é de fácil implementação e baixo consumo de banda, sendo preferido pelas Operadoras.

Após esta etapa a amostra é transportada ou por rede determinística ou por rede estatística.

A rede determinística considera os dados por ela transmitidos de forma transparente, adicionando à transmissão dos mesmos apenas mecanismos de sinalização e sincronismo.

A rede estatística, por sua vez, dependendo de cada camada ou tecnologia, adicionará mais ou menos *overhead* à informação transmitida.

O diagrama da Figura 3.3 considera todas as possibilidades de transporte da amostra de voz por todas as redes, os quais se denominaram caminhos.

Foi excluído do cálculo o elemento *SDH*, uma vez que se entende que todos os serviços o utilizam como meio de transporte e, portanto, o valor deste trecho seria sempre o mesmo.

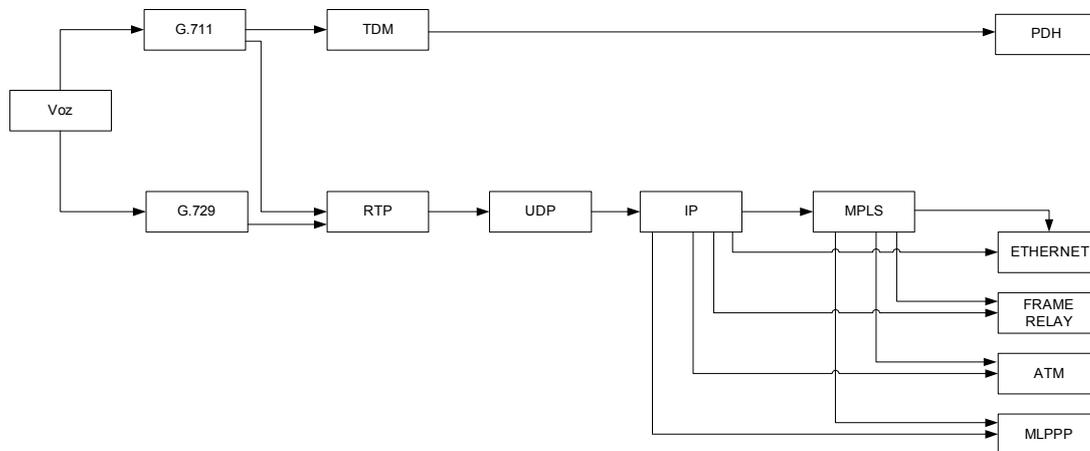


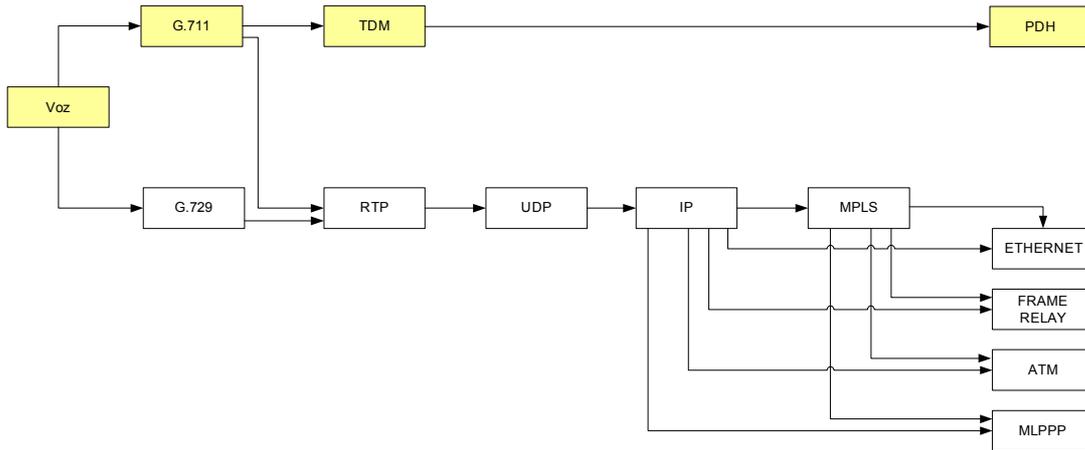
Figura 3.3 - Sobreposição genérica dos serviços [AUTORA]

Como exemplo, um cliente que solicita um serviço de transporte de voz utilizando como acesso o *Frame Relay* precisa adicionar ao *payload* ou informação que se deseja transmitir os seguintes *overheads*: Codec (G.711 ou G.729) + *RTP* (*Real Time Transport Protocol*) + *UDP* (*User Datagram Protocol*) + *IP* (*Internet Protocol*) + *Frame Relay*.

Para os cálculos consideraram-se duas amostras por pacote, ou seja, 8.000 *bytes* por segundo para o G.711 e 1000 *bytes* para o G.729. Em 20 ms uma amostra (*payload*) terá 160 *bytes* para G.711 e 40 *bytes* para G.729 [CISCO2004a].

Caminho 1:

Este é o caminho mais simples que a amostra de voz pode percorrer. A amostra é codificada (Codec) e transportada em *time slots* de 64Kbps (*TDM – Time Division Multiplexing*) de um feixe E1 (2048Kbps) do *PDH* (*Plesiochronous Digital Hierarchy*).

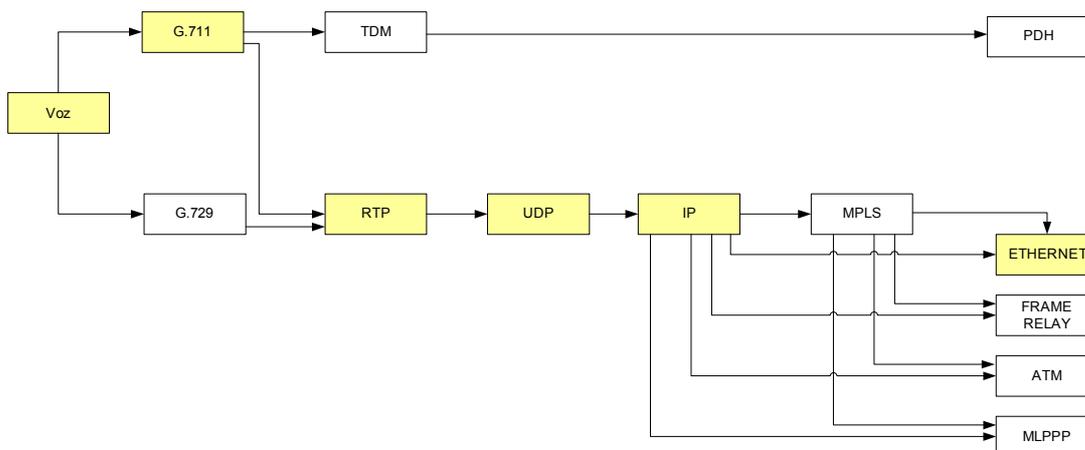


Sabe-se que o *PDH* utiliza dois *time slots* para controle e sinalização dos 32 *time slots* disponíveis no E1 (G.703), assim, tem-se:

$$Eficiência = \frac{2.048 - 128}{2.048} = 93,75\%$$

Caminho 2:

A amostra é codificada (Codec) e transportada sobre uma rede estatística *IP*. Para tal à amostra serão adicionados os cabeçalhos do *RTP* (*Real-time Transport Protocol*), do *UDP* (*User Datagram Protocol*) e do *IP* (*Internet Protocol*). Este modelo de transporte de voz sobre rede estatística, até o *IP*, é o adotado pela maioria das Operadoras no Brasil. A partir deste ponto a pilha de protocolos utilizados dependerá do tipo de acesso contratado.



Como visto na Equação 3.1:

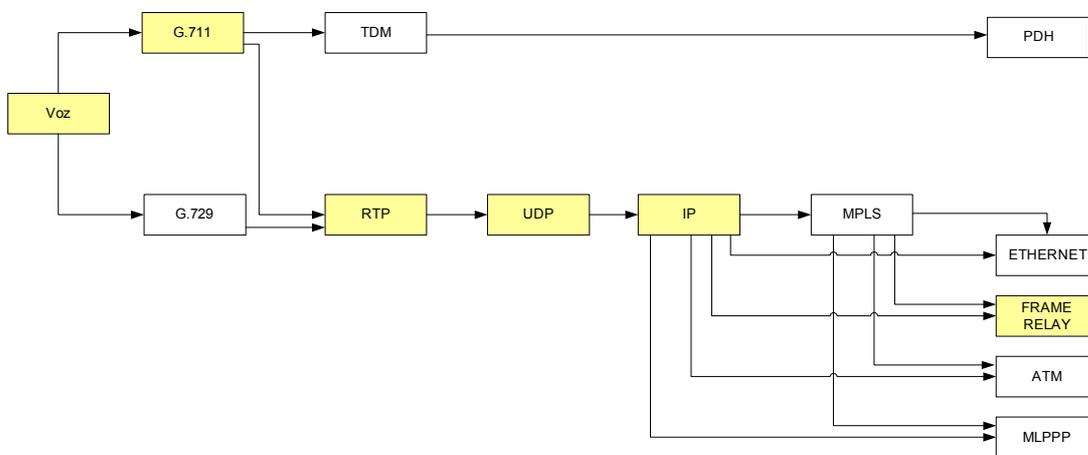
$$Eficiência = \frac{Codec}{L arg.banda} = \left[ \frac{Tam.amostra}{Cab.camada2 + Cab.camada3e4 + Tam.amostra} \right] \quad \text{Equação 3.2}$$

Então,

eficiência para o Caminho 2 vale:

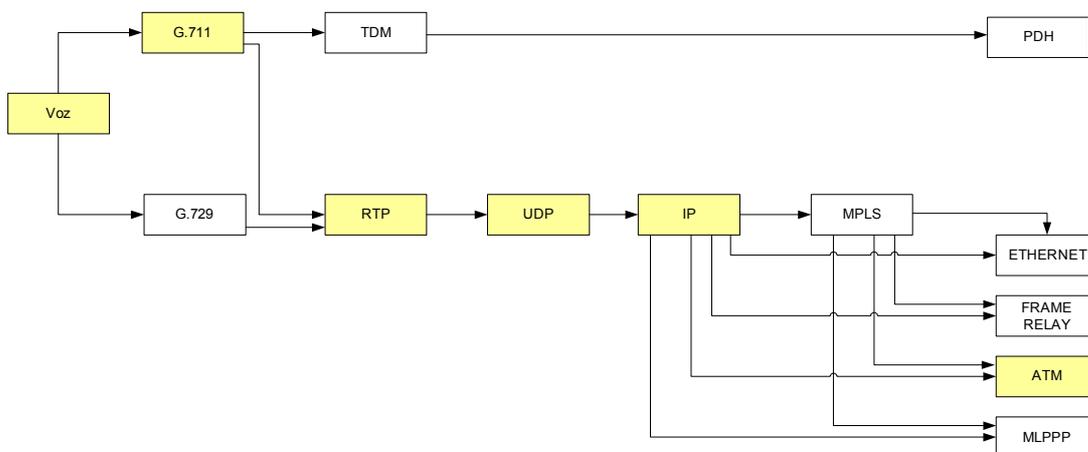
$$Eficiência = \left[ \frac{160bytes}{18bytes + 40bytes + 160bytes} \right] = 73,39\%$$

Caminho 3:



$$Eficiência = \left[ \frac{160bytes}{6bytes + 40bytes + 160bytes} \right] = 77,66\%$$

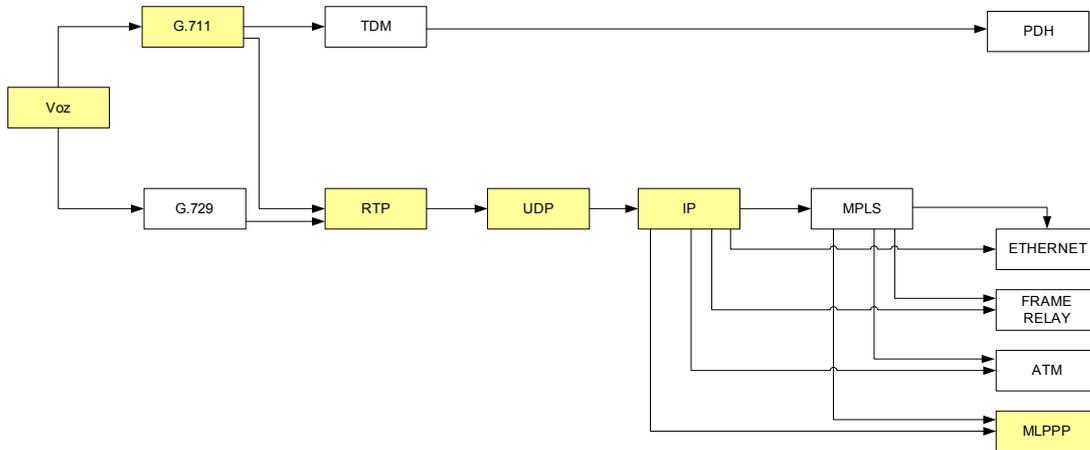
Caminho 4:



$$Eficiência = \left[ \frac{160bytes}{25bytes + 40bytes + 160bytes} \right] = 71,11\%$$

Vale lembrar que foram utilizadas 5 células de 48 *bytes*, cada célula com 5 *bytes* de cabeçalho, para o transporte do pacote *IP* (160 *bytes* + 40 *bytes*).

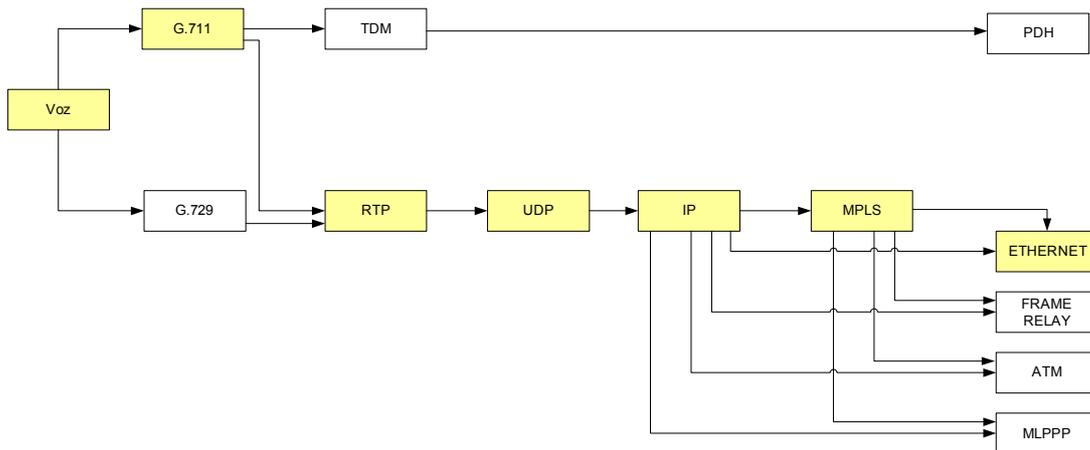
Caminho 5:



$$Eficiência = \left[ \frac{160bytes}{6bytes + 40bytes + 160bytes} \right] = 77,66\%$$

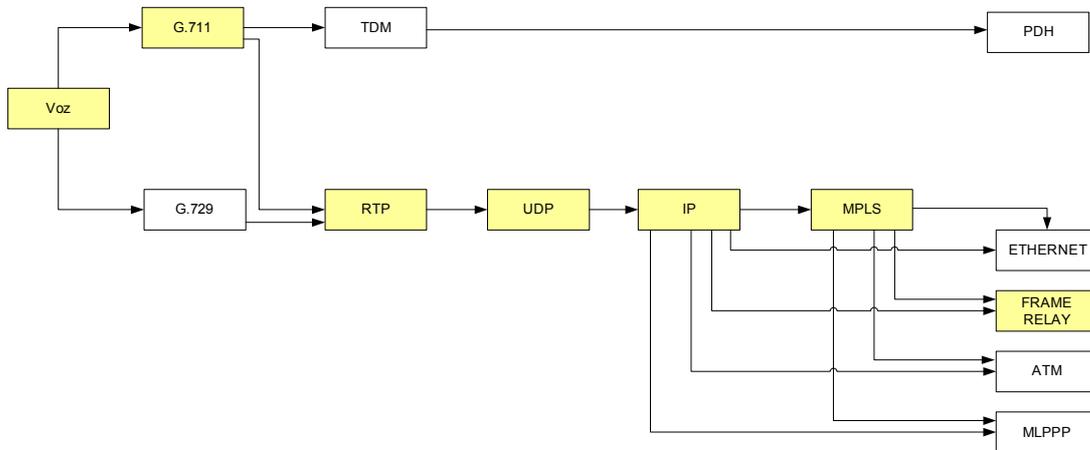
Caminho 6:

Este caminho utiliza o protocolo *MPLS* (*Multiprotocol Label Switching*) o qual agrega características de camadas 2 e 3.



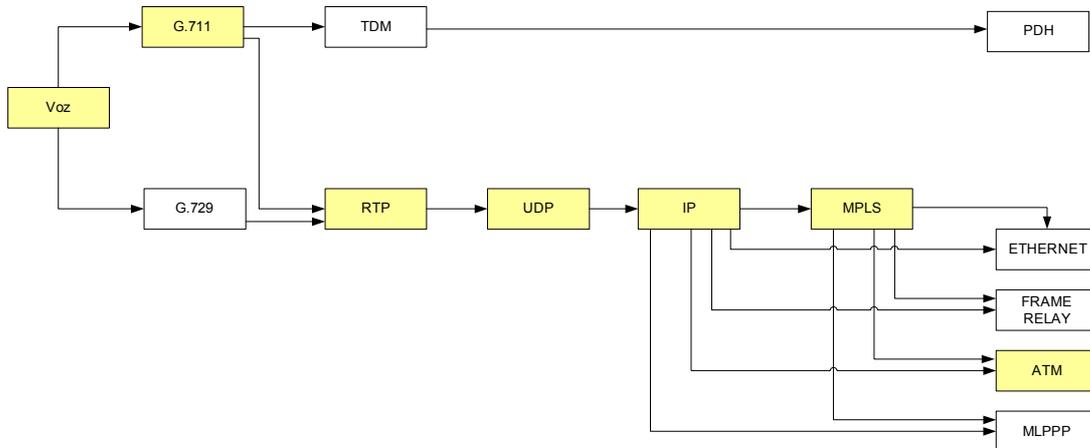
$$Eficiência = \left[ \frac{160bytes}{18bytes + 44bytes + 160bytes} \right] = 72,07\%$$

Caminho 7:



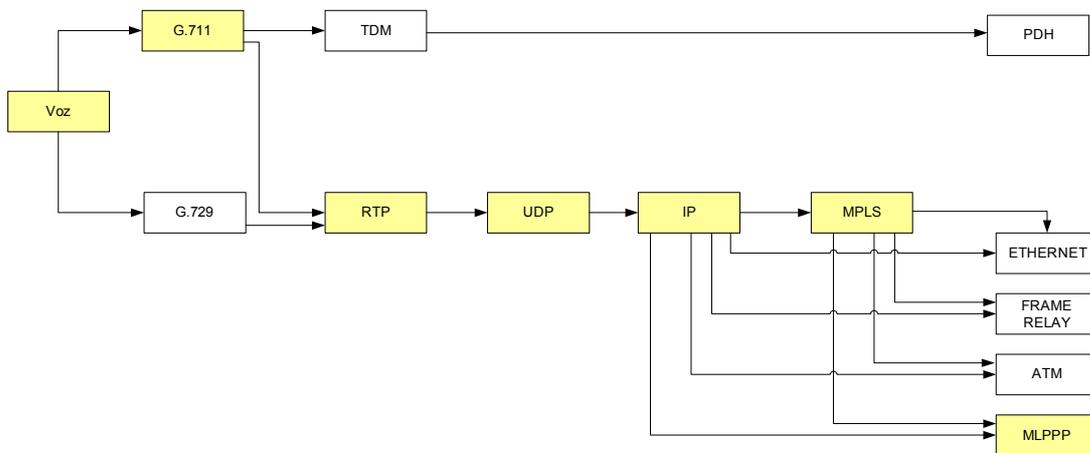
$$Eficiência = \left[ \frac{160bytes}{6bytes + 44bytes + 160bytes} \right] = 76,19\%$$

Caminho 8:



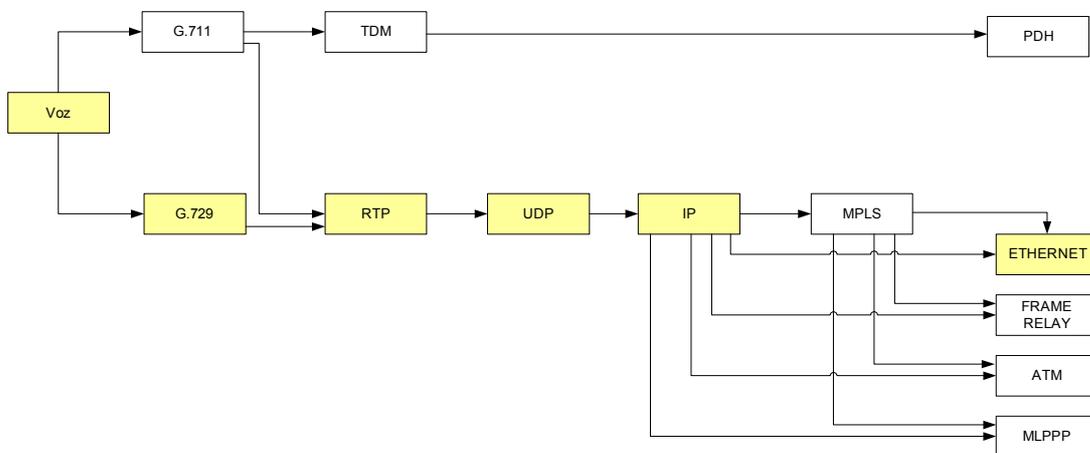
$$Eficiência = \left[ \frac{160bytes}{25bytes + 44bytes + 160bytes} \right] = 69,86\%$$

Caminho 9:



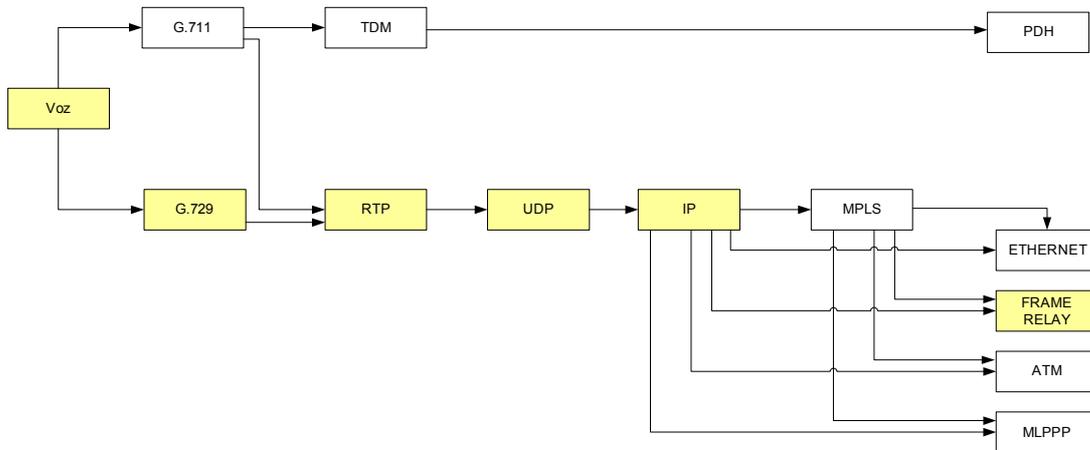
$$Eficiência = \left[ \frac{160bytes}{6bytes + 44bytes + 160bytes} \right] = 76,19\%$$

Caminho 10:



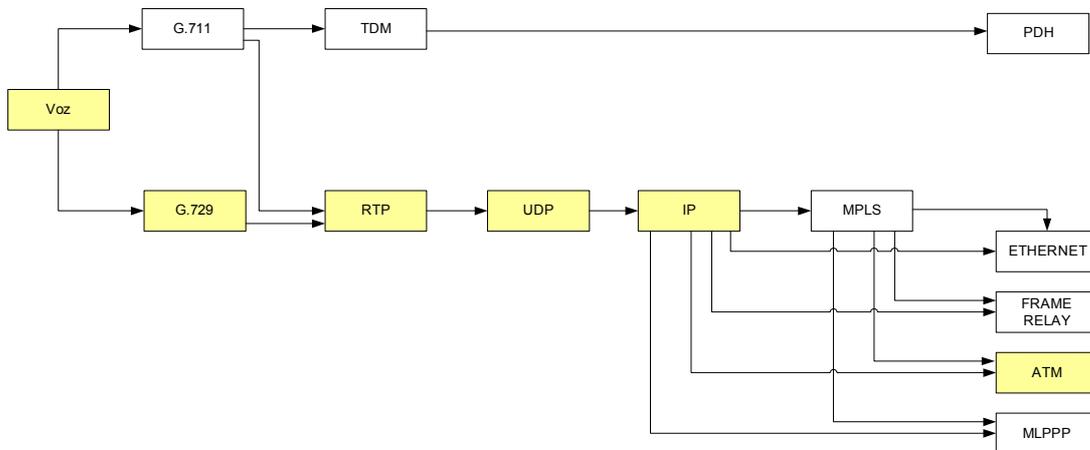
$$Eficiência = \left[ \frac{40bytes}{18bytes + 40bytes + 40bytes} \right] = 40,81\%$$

Caminho 11:



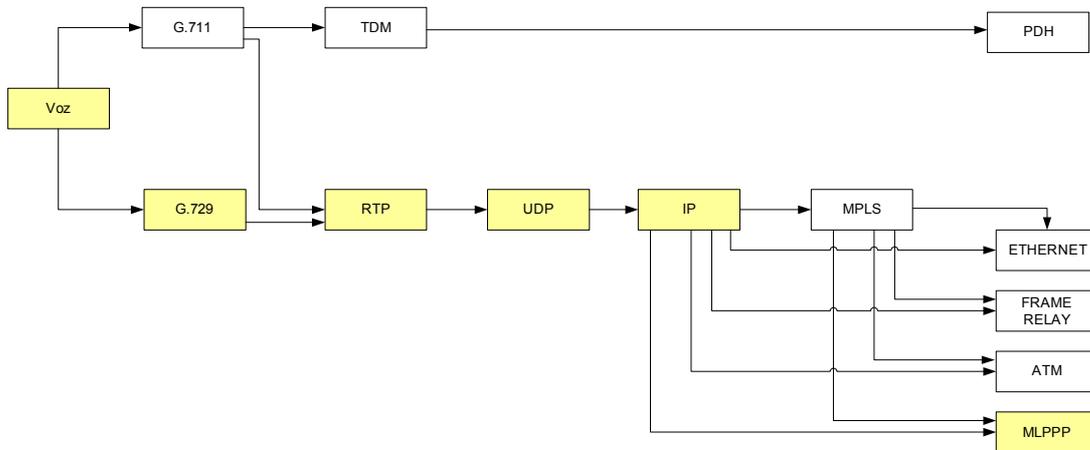
$$Eficiência = \left[ \frac{40bytes}{6bytes + 40bytes + 40bytes} \right] = 46,51\%$$

Caminho 12:



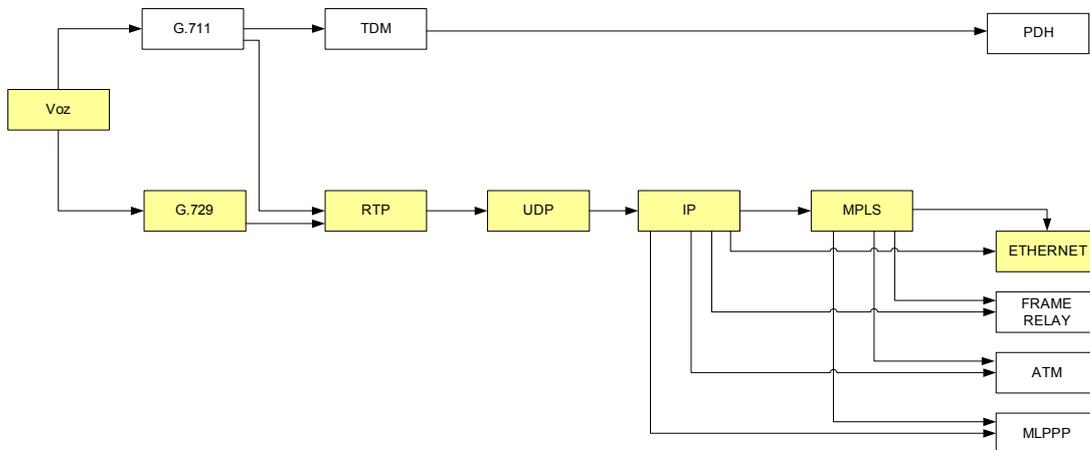
$$Eficiência = \left[ \frac{40bytes}{10bytes + 40bytes + 40bytes} \right] = 44,44\%$$

Caminho 13:



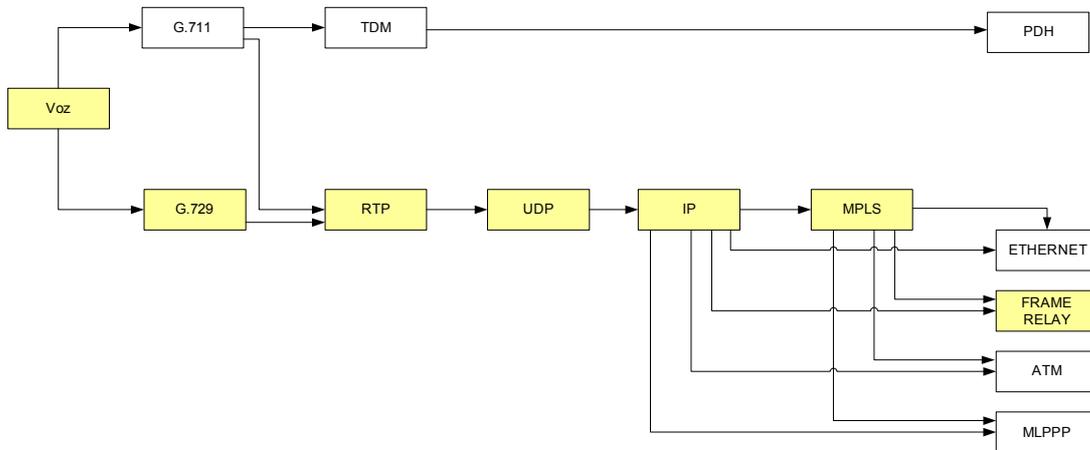
$$Eficiência = \left[ \frac{40bytes}{6bytes + 40bytes + 40bytes} \right] = 46,51\%$$

Caminho 14:



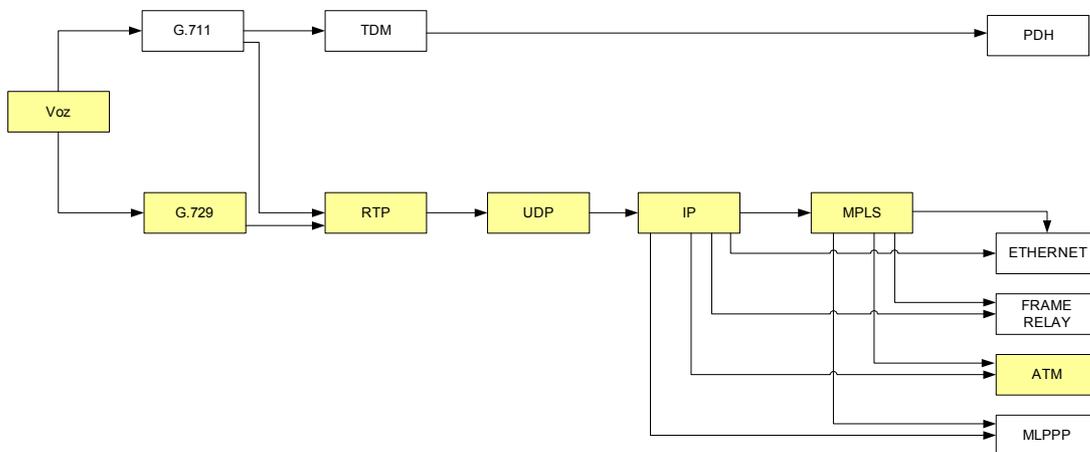
$$Eficiência = \left[ \frac{40bytes}{18bytes + 44bytes + 40bytes} \right] = 39,21\%$$

Caminho 15:



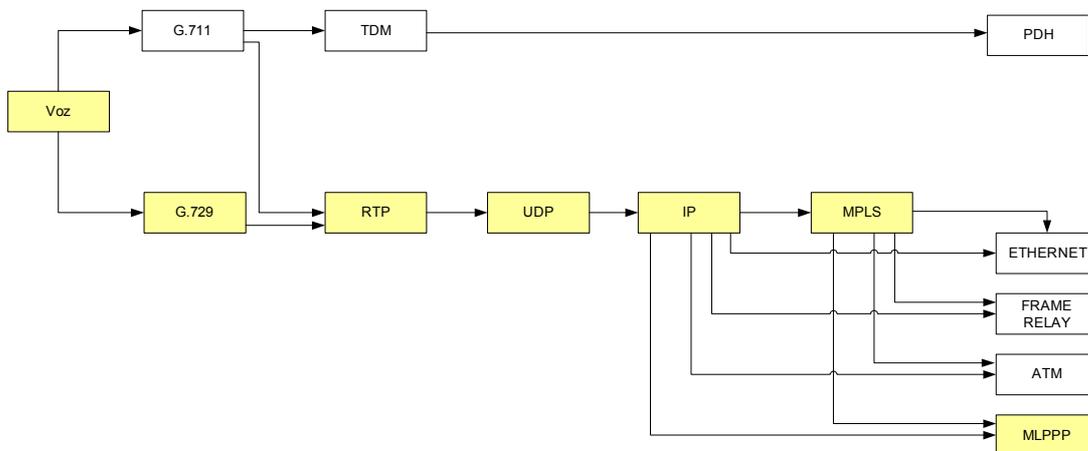
$$Eficiência = \left[ \frac{40bytes}{6bytes + 44bytes + 40bytes} \right] = 44,44\%$$

Caminho 16:



$$Eficiência = \left[ \frac{40bytes}{10bytes + 44bytes + 40bytes} \right] = 42,55\%$$

Caminho 17:



$$Eficiência = \left[ \frac{40bytes}{6bytes + 44bytes + 40bytes} \right] = 44,44\%$$

No entanto, observa-se que esta é a eficiência para uma amostra de voz quando transmitida sozinha no canal. Para efeitos de determinar o caminho ou topologia de rede mais eficiente é preciso calcular quantas conversações simultâneas pode-se ter em um mesmo canal.

Para tanto, se:

$$L \text{ arg. de banda} = \left[ \frac{Cab.camada2 + Cab.camada3e4 + Tam.amostra}{Tam.amostra} \right] \times Codec \quad \text{Equação 3.3}$$

e

$$Eficiência = \frac{Tam.amostra}{Cab.camada2 + Cab.camada3e4 + Tam.amostra} \quad \text{Equação 3.4}$$

Então:

$$L \text{ arg. de banda} = \frac{1}{Eficiência} \times Codec \quad \text{Equação 3.5}$$

Pode-se dizer que:

$$Quant.conversaçõessimultâneas = \frac{Tamanhocanal}{L \text{ arg. de banda}} \quad \text{Equação 3.6}$$

Logo,

$$EficiênciarelativaaoPCM = \frac{Quant.conversaçõessimultâneas}{30timeslots} \quad \text{Equação 3.7}$$

O resultado deste cálculo foi resumido na Tabela 3.2 a seguir:

Tabela 3.2 - Resumo Cálculo da Eficiência de Transmissão [AUTORA]

Caminhos	Amostra	G.711	G.729	Eficiência	1/Eficiência	Largura de Banda	Quantidade de canais de voz suportados por 2Mbps	Eficiência temporal relativa ao TDM
1	160	64		93,75%			30	1
2	160	64		73,39%	1,362583458	87,20534133	23,4848	0,78282667
3	160	64		77,66%	1,287664177	82,41050734	24,8512	0,82837333
4	160	64		71,11%	1,406271973	90,00140627	22,7552	0,75850667
5	160	64		77,66%	1,287664177	82,41050734	24,8512	0,82837333
6	160	64		72,07%	1,387539892	88,80255307	23,0624	0,76874667
7	160	64		76,19%	1,312508203	84,000525	24,3808	0,81269333
8	160	64		69,86%	1,431434297	91,61179502	22,3552	0,74517333
9	160	64		76,19%	1,312508203	84,000525	24,3808	0,81269333
10	40		8	40,81%	2,450379809	19,60303847	104,4736	3,48245333
11	40		8	46,51%	2,150075253	17,20060202	119,0656	3,96885333
12	40		8	44,44%	2,250225023	18,00180018	113,7664	3,79221333
13	40		8	46,51%	2,150075253	17,20060202	119,0656	3,96885333
14	40		8	39,21%	2,550369804	20,40295843	100,3776	3,34592
15	40		8	44,44%	2,250225023	18,00180018	113,7664	3,79221333
16	40		8	42,55%	2,350176263	18,80141011	108,928	3,63093333
17	40		8	44,44%	2,250225023	18,00180018	113,7664	3,79221333

### 3.1.1.2 Confiabilidade

Reflete a probabilidade de mau funcionamento ou falha no produto ou serviço, em um determinado período de tempo.

No contexto das Redes de Comunicação de Dados, esta dimensão é importante para quem está contratando um serviço ou produto, pois reflete o quantitativo de falhas percebidas pelo cliente, ou seja, a indisponibilidade do serviço.

Assim, para esta dimensão da qualidade, foram analisados os bilhetes de defeitos dos serviços, onde se coletou a quantidade de defeitos reclamados naqueles circuitos que popularam a base de dados do experimento no período de janeiro/2005 a julho/2007.

A fim de distribuir os bilhetes de defeito no tempo foi aplicada a média de defeitos diários por circuitos no período de janeiro/2005 a julho/2007 porque, naturalmente, um serviço ativado há mais tempo pode ter um histórico de defeitos maior que aquele ativado há menos tempo. Por exemplo: um circuito que teve sua configuração encerrada em 14/12/2005 foram consideradas todas as reclamações entre o período de 14/12/2005 à 28/07/2007. Esta quantidade foi dividida pelo número de dias deste período de modo a se ter uma quantidade média diária de reclamações

A título de ilustração, segue na Figura 3.4 uma parte do arquivo de BD's (Bilhetes de Defeito), que foram levantados para o período.

FILIAL	DATA ENCERRAMENTO O	LOCAL CIRCUIT	NUMERO CIRCUITO	DATA ABERTURA	CODIGO ENCIERRAM ENTO	TIPO VELOCIDADE	DEGRAU TARIFARIO	CLIENTE PONTA A	LOCALIDADE PONTA A	LOCALIDADE PONTA B	FAMILIA PRODUTO
PR	15/12/2006 13:15	CTA	485274	7/6/2006 10:40	3636	128		3 EMPRESA BRASILEIRA DE TELECOMUNICACOES S/A EMBRATE	LDA	APQ	EILD
PR	1/12/2006 11:37	CTA	713741	4/10/2006 21:35	504	64		5 TELEFONICA EMPRESAS S/A	CTA	SLOU	EILD
DF	15/12/2006 13:16	BSA	463478	11/11/2006 16:27	206	64		99 EMPRESA BRASILEIRA DE TELECOMUNICACOES S/A EMBRATE	BSA	BSA	EILD
PR	8/12/2006 12:17	CTA	472059	17/11/2006 10:21	1206	64		99 HSBC BANK BRASIL S/A BANCO MULTIPLO	MGA		FR
PR	8/12/2006 10:37	CTA	471960	24/11/2006 14:57	87953	64		1 PRIMESYS SOLUCOES EMPRESARIAIS S/A	MGA	MGC	EILD
DF	22/12/2006 16:57	BSA	487841	25/11/2006 11:12	504	256		99 POLICIA MILITAR DO DISTRITO FEDERAL	SOBD		VEFOR ADSL
BRTCM	30/11/2006 22:36	SPO	417918	27/11/2006 16:19	52112	2048		1 PRIMESYS SOLUCOES EMPRESARIAIS S/A	BRE	SPO	
CRT	1/12/2006 12:13	PAE	498562	28/11/2006 20:21	82142	256		99 COMPANHIA ZAFFARI COMERCIO E INDUSTRIA	PAE		VEFOR FR
BRTCM	1/12/2006 13:01	SPO	415623	29/11/2006 00:07	50313	1024		99 INTELIG TELECOMUNICACOES LTDA.	SPO		BRTLINK
CRT	1/12/2006 11:36	PAE	452619	29/11/2006 15:08	104	2048		99 TELET S/A	EIO		
CRT	7/12/2006 11:20	PAE	444061	29/11/2006 15:45	82892	64		99 SECRETARIA FAZENDA RIO GRANDE SUL	MEL		FR
GO	30/11/2006 21:05	GNA	444276	29/11/2006 16:08	82608	64		99 DEPARTAMENTO ESTADUAL DE TRANSITO	MRR		FR
CRT	30/11/2006 19:44	PAE	475986	29/11/2006 16:34	81192	64		99 TRIBUNAL SUPERIOR ELEITORAL	SVC		FR
GO	30/11/2006 19:30	GNA	441949	29/11/2006 18:43	82972	64		99 TRIBUNAL DE JUSTICA DO ESTADO TOCANTINS	ICJ		FR
GO	30/11/2006 19:12	GNA	452939	30/11/2006 06:19	81121	64		3 TELEFONICA EMPRESAS S/A	ANU	GNA	EILD
MT	30/11/2006 19:11	CBA	419544	30/11/2006 06:52	82672	128		99 GAZIN IND E COM DE MOVEIS E ELETRODOMESTICOS LTDA	GDN		INTERLAN FR
DF	1/12/2006 10:14	BSA	463276	30/11/2006 09:28	82622	64		99 BRB BANCO DE BRASILIA S/A	SOBD		FR
GO	30/11/2006 21:11	GNA	420573	30/11/2006 10:10	82192	64		3 SECRETARIA DA FAZENDA	PMP		FR
CRT	30/11/2006 19:18	PAE	460001	30/11/2006 10:54	82508	2048		1 TIM CELULAR S/A	SDE	BSR	EILD
GO	1/12/2006 11:16	GNA	433514	30/11/2006 10:59	84523	64		99 CIA ENERGETICA DE GOIAS - CELG	SFZ		FR
DF	1/12/2006 10:15	BSA	465691	30/11/2006 11:00	3686	NAO AFE		99 BRASIL TELECOM S/A	BSA		PLUS
MT	8/12/2006 20:44	TCS1	508267	30/11/2006 11:13	51101	512		8 AT&AMP	NORMAL	BENTO B DE ANDMT	
CRT	4/12/2006 12:21	PAE	600095	30/11/2006 11:28	504	512		99 MAGAZINE LUIZA S/A	BGV		FR
CRT	1/12/2006 14:51	PAE	482123	30/11/2006 11:32	524	64		99 PRIMESYS SOLUCOES EMPRESARIAIS S/A	PAE	PAE	EILD
MT	30/11/2006 16:33	CBA	415564	30/11/2006 09:40	81153	128		99 HSBC BANK BRASIL S/A BANCO MULTIPLO	CBA		FR
PR	1/12/2006 20:24	CTA	463992	30/11/2006 11:56	811G2	256		99 HSBC BANK BRASIL S/A BANCO MULTIPLO	CTA		FR
GO	1/12/2006 16:45	GNA	433530	30/11/2006 12:07	82142	256		99 TRIBUNAL DE JUSTICA DO ESTADO TOCANTINS	CGO		FR
CRT	1/12/2006 06:51	PAE	491554	30/11/2006 14:01	205	256		99 MELBAR PRODUTOS DE LIGNINA LTDA	OSKR		IP
SC	30/11/2006 19:16	FNS	475722	30/11/2006 14:02	82532	512		99 FUNDO PARA MELHORIA DA SEGURANCA PUBLICA FMSP	NLA		VEFOR ADSL
PR	30/11/2006 22:48	CTA	414603	30/11/2006 14:24	82152	256		99 CONFEDERACAO INTERESTADUAL DAS COOPERATIVAS LIGADA	PAC		VEFOR FR
PR	1/12/2006 11:27	CTA	725645	30/11/2006 14:58	85523	128		99 CONFEDERACAO INTERESTADUAL DAS COOPERATIVAS LIGADA	FMS		VEFOR FR
SC	30/11/2006 21:16	FNS	467837	30/11/2006 15:00	1306	2048		2 TIM SUL S/A	JVE	BNU	EILD
SC	30/11/2006 21:18	FNS	467838	30/11/2006 15:04	1306	2048		2 TIM SUL S/A	JVE	BNU	EILD
BRTCM	30/11/2006 19:38	SPO	419718	30/11/2006 15:07	2803	NAO AFE		99 AGENCIA NACIONAL DE TRANSPORTES TERRESTRES ANTT	BHE		PLUS
SC	30/11/2006 21:18	FNS	467839	30/11/2006 15:08	1306	2048		2 TIM SUL S/A	JVE	BNU	EILD
BRTCM	1/12/2006 00:33	SPO	421849	30/11/2006 15:10	85194	256		99 INDEPENDENCIA ALIMENTOS LTDA	JUA		VEFOR FR OFF
BRTCM	1/12/2006 01:55	SPO	417491	30/11/2006 15:12	84195	8000		99 SO DO BRASIL TELECOM LTDA	RSD		IP
BRTCM	1/12/2006 02:00	SPO	417624	30/11/2006 15:13	84195	128		99 CIA INDUSTRIAL H. CARLOS SCHNEIDER	NLA		INTERLAN FR ON
BRTCM	22/12/2006 11:32	SPO	420873	30/11/2006 15:15	3686	2048		99 UNIAO BRASILENSE DE EDUCACAO E CULTURA	CRF		
BRTCM	1/12/2006 01:50	SPO	418962	30/11/2006 15:15	84195	144 MBP!		99 WAY TV BELO HORIZONTE S/A	BHE		IP
BRTCM	30/11/2006 19:03	SPO	422303	30/11/2006 15:19	2803	144 MBP!		99 WAY TV BELO HORIZONTE S/A	BHE		
BRTCM	1/12/2006 02:33	SPO	418488	30/11/2006 15:19	1206	256		99 WANT COMMUNICATIONS LTDA	CEM		IP
BRTCM	22/12/2006 16:04	SPO	418751	30/11/2006 15:19	105	2048		99 INTERBET LTDA	SLA		

Figura 3.4 – Amostra do arquivo de Bilhetes de Defeitos

### 3.1.2 Flexibilidade

Segundo [PAIVA, 2004], dentro de uma abordagem operacional, flexibilidade é a capacidade de um sistema de adotar uma gama de estados diferentes.

Em uma abordagem mais estratégica, flexibilidade é a capacidade que um determinado sistema produtivo tem de responder a variáveis externas e internas. Como variáveis externas citam-se: mercados turbulentos, novas necessidades dos consumidores, avanços tecnológicos, necessidades de entregas cada vez mais rápidas. Como variáveis internas citam-se: falhas em sistemas ou equipamentos, quebra no suprimento e falhas de fornecedores.

A capacidade de um sistema se adaptar as variáveis internas ou externas não define o grau de flexibilidade deste. Porém, certamente, um sistema que se move rapidamente e com baixo custo de um estado para o outro é mais flexível.

A necessidade da função de operações de uma empresa ser flexível justifica-se por:

1. Variedade de atividades: a operação deve lidar no seu cotidiano com a variedade de produtos ou serviços, variedade de canais de entrega ou variedade de prazos de entrega.
2. Incertezas de curto prazo: manter o desempenho, principalmente a confiabilidade, quando as coisas não ocorrem conforme o planejado.
3. Incertezas de longo prazo: a operação deve ser flexível o suficiente para operar em novas condições, tais como: novos produtos, novos mercados ou novos concorrentes.
4. “Ignorância”: a operação deve ser flexível o suficiente para agir, quando na empresa não existe um direcionamento estratégico coerente.

As razões acima podem ser mais bem compreendidas por meio da Figura 3.5 seguinte:

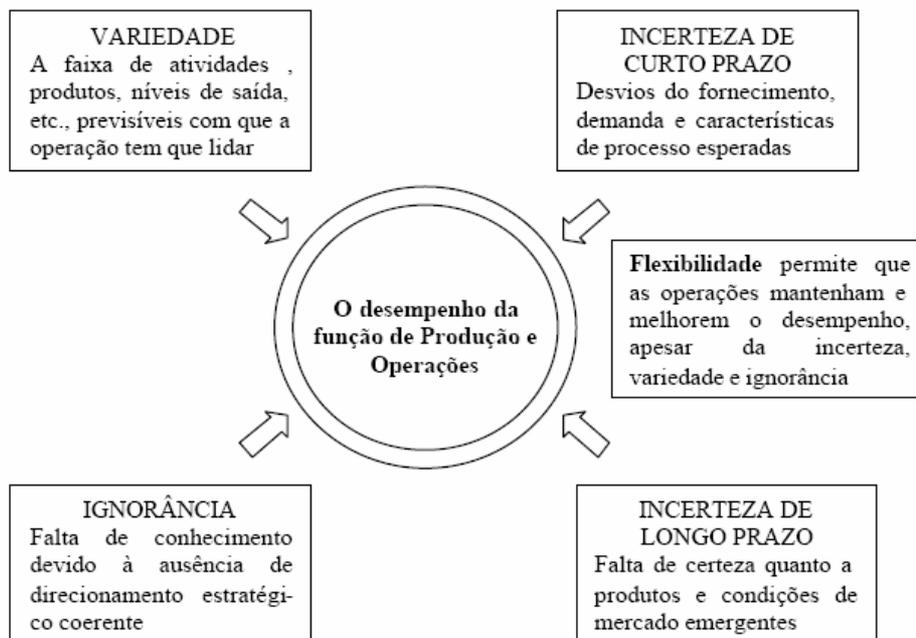


Figura 3.5- Fatores que influem no desempenho da função de operação [SLACK1996]

A seguir, descreve-se cada tipo de flexibilidade apresentando suas principais características.

### 3.1.2.1 Flexibilidade de Novos Produtos/serviços

A flexibilidade de novos produtos está intimamente relacionada à habilidade que a empresa possui de lançar no mercado e produzir novos produtos e serviços, bem como modificar aqueles já existentes.

### 3.1.2.2 Flexibilidade de *Mix* de Produto

É definida como a quantidade de produtos diferentes que um determinado sistema produtivo pode produzir sem a necessidade de alterar suas características fundamentais. [PAIVA, 2004]

### 3.1.2.3 Flexibilidade de volume

É a capacidade de alterar o volume total de produção, seja este de produtos ou serviços.

Para se alcançar a flexibilidade de volume é importante que a empresa seja ágil na capacidade de reprogramar fornecedores e administrar os níveis de estoque.

### 3.1.2.4 Flexibilidade de entrega

A flexibilidade de entrega é a habilidade que a empresa tem em alterar a data de entrega assumida antecipadamente com o cliente. Na grande maioria das vezes, flexibilizar a entrega significa antecipar a entrega do produto ou serviço, o que requer maior empenho das operações da empresa, para o cumprimento do novo prazo. Postergar o prazo de entrega de um pedido, ou seja, flexibilizar a entrega para além do prazo inicial, normalmente é uma tarefa bem mais fácil para a empresa.

[PAIVA2004] diz que a correta mensuração da flexibilidade é importante à medida que atende aos interesses de pesquisadores e administradores, porém, uma das dimensões da flexibilidade, o tempo, ainda não foi corretamente estudado. É possível identificar algumas dificuldades para operacionalizar as medidas de flexibilidade:

- Falta de um critério rigoroso para classificar os tipos de flexibilidade;
- Multidimensionalidade da flexibilidade gerando dificuldades para estabelecer uma escala adequada para teste e coleta de dados;
- Dificuldade na generalização dos resultados;

Em função das dificuldades relatadas acima e da orientação de alguns autores de que pode ser perigoso tentar atender todas as dimensões simultaneamente foi descartada deste estudo a comparação entre redes segundo o critério da flexibilidade.

### **3.1.3 Desempenho de Entrega**

Conforme [PAIVA2004], este critério possui duas dimensões, cada uma podendo ser considerada como um critério competitivo em si.

A primeira relaciona-se a capacidade da empresa em lançar produtos ou serviços que funcionem conforme especificado, entregar dentro dos prazos estipulados e corrigir qualquer defeito imediatamente após ocorrer. Também pode ser entendido como a confiabilidade da entrega, uma vez que está relacionado com a capacidade de entregar serviços no prazo prometido.

A segunda é a velocidade de entrega. A velocidade está relacionada a capacidade de fazer as operações rapidamente e entregar com prazos menores que os concorrentes.

Neste sentido, o desempenho de entrega está condicionado a quão rápido ou quanto é possível apressar a data de entrega de um serviço. Em redes este critério está relacionado ao tempo gasto no provisionamento, isto é, o tempo de configurar o serviço contratado na rede que suporta aquele determinado produto. Para utilização deste critério foi feito o levantamento da base de dados de tramitação das ordens de serviço, no período de janeiro/2005 a julho/2007, onde foram coletados os tempos das tarefas de configuração dos produtos, para os quais são conhecidas as topologias.

Ilustrativamente, segue na Figura 3.6 resumo do arquivo de tramitação das Ordens de Serviço, com as tarefas executadas durante a fase de provisionamento do serviço.

Filial	Data Início	Data Fim	Origem	Tarefas
PR	28/3/2005 14:32	28/3/2005 15:07	OMS	S49 T0 Gerenciar Pendências
PR	28/3/2005 15:07	5-APR-2005 16:23:00	OMS	S69 T0 Liberar Facilidades de Par Metálico
PR	5-APR-2005 16:23:00	5-APR-2005 16:27:00	OMS	S49 T0 Gerenciar Pendências
PR	5-APR-2005 16:27:00	6-SEP-2005 9:34:00	OMS	S49 T01 Gerenciar Pendências Cliente
PR	6-SEP-2005 9:34:00	6-SEP-2005 9:51:00	OMS	S49 T0 Gerenciar Pendências
PR	6-SEP-2005 9:51:00	13-SEP-2005 11:30:00	OMS	S69 T0 Liberar Facilidades de Par Metálico
PR	13-SEP-2005 11:30:00	13-SEP-2005 11:45:00	OMS	S49 T0 Gerenciar Pendências
PR	13-SEP-2005 11:45:00	13-SEP-2005 16:15:00	OMS	S69 T0 Liberar Facilidades de Par Metálico
PR	13-SEP-2005 16:15:00	14-SEP-2005 10:41:00	OMS	P99 T01 Atualizar Status da OS para o CRM
DF	14/3/2005 17:25	15/3/2005 08:56	OMS	S188 T0 Montar Topologia
DF	15/3/2005 08:56	15/3/2005 09:02	OMS	S188 T01 Alocar Facilidades Auto
DF	15/3/2005 09:02	15/3/2005 09:23	OMS	S188 T03 Validar/Montar Topologia
DF	15/3/2005 09:23	15/3/2005 09:24	OMS	S112 T0 Alocar Facilidades de Redes de Dados
DF	15/3/2005 09:24	15/3/2005 09:28	OMS	S453 T0 Configurar Rede de Dados/Validar Ativações Automáticas
DF	15/3/2005 09:24	15/3/2005 09:24	OMS	S452 T0 Ativar Facilidades
DF	15/3/2005 09:28	16/3/2005 17:45	OMS	S68 T0 Designar e qualificar par metálico
DF	16/3/2005 17:45	16/3/2005 18:39	OMS	S63 T0 Instalar, Configurar e Testar Modem Digital
DF	16/3/2005 18:39	17/3/2005 08:42	OMS	S85 T0 Gerenciar Pendências
DF	17/3/2005 08:42	15/7/2005 16:13	OMS	S85 T02 Gerenciar Pendências Comercial
DF	15/7/2005 16:13	15/7/2005 17:22	OMS	S85 T0 Gerenciar Pendências
DF	15/7/2005 17:22	18/7/2005 18:09	OMS	S63 T0 Instalar, Configurar e Testar Modem Digital
DF	18/7/2005 18:09	18/7/2005 18:17	OMS	S85 T0 Gerenciar Pendências
DF	18/7/2005 18:17	19/7/2005 18:46	OMS	S63 T0 Instalar, Configurar e Testar Modem Digital
DF	19/7/2005 18:46	19/7/2005 18:47	OMS	S85 T0 Gerenciar Pendências
DF	19/7/2005 18:47	13-SEP-2005 19:11:00	OMS	S85 T01 Gerenciar Pendências Cliente
DF	13-SEP-2005 19:11:00	13-SEP-2005 19:41:00	OMS	S85 T0 Gerenciar Pendências
DF	13-SEP-2005 19:41:00	14-SEP-2005 9:58:00	OMS	S63 T0 Instalar, Configurar e Testar Modem Digital
DF	14-SEP-2005 9:58:00	14-SEP-2005 10:02:00	OMS	S751 T0 Confirmar ativação (OBJ)
DF	14-SEP-2005 10:02:00	14-SEP-2005 10:03:00	OMS	P98 T01 Atualizar Status da OS para o CRM
CRT	21/3/2005 16:17	21/3/2005 16:19	OMS	S185 T0 Consultar Topologia Atual
CRT	21/3/2005 16:19	22/3/2005 09:00	OMS	S185 T05 Inserir Dados
CRT	22/3/2005 09:00	22/3/2005 16:38	OMS	S185 T02 Identificar Topologia
CRT	22/3/2005 16:38	24/3/2005 17:22	OMS	S70 T0 Retirar Modem Digital e Jumpers
CRT	22/3/2005 16:38	22/3/2005 16:38	OMS	S185 T03 Liberar Facilidades
CRT	24/3/2005 17:22	28/3/2005 14:10	OMS	S49 T0 Gerenciar Pendências
CRT	28/3/2005 14:10	8-SEP-2005 17:39:00	OMS	S49 T01 Gerenciar Pendências Cliente
CRT	8-SEP-2005 17:39:00	9-SEP-2005 8:57:00	OMS	S49 T0 Gerenciar Pendências
CRT	9-SEP-2005 8:58:00	12-SEP-2005 17:17:00	OMS	S70 T0 Retirar Modem Digital e Jumpers
CRT	12-SEP-2005 17:17:00	13-SEP-2005 11:04:00	OMS	S69 T0 Liberar Facilidades de Par Metálico
CRT	13-SEP-2005 11:04:00	13-SEP-2005 13:52:00	OMS	S106 T0 Desalocar redes de Dados
CRT	13-SEP-2005 11:04:00	13-SEP-2005 13:55:00	OMS	S107 T0 Desconfigurar Rede de Dados/Validar desativação automática
CRT	13-SEP-2005 13:55:00	13-SEP-2005 13:55:00	OMS	P99 T0 Confirmar Desativação
CRT	13-SEP-2005 13:55:00	13-SEP-2005 14:01:00	OMS	P99 T03 Monitorar Erro
CRT	13-SEP-2005 14:01:00	13-SEP-2005 14:02:00	OMS	P99 T0 Confirmar Desativação
CRT	13-SEP-2005 14:02:00	13-SEP-2005 14:02:00	OMS	P99 T01 Atualizar Status da OS para o CRM
SC	24/3/2005 09:44	24/3/2005 14:16	OMS	S187 T0 Montar Topologia

Figura 3.6 – Amostra do arquivo de tramitação das Ordens de Serviço

### 3.1.4 Custo

Para Slack apud [VIEIRA,2001], o custo está no centro dos objetivos da manufatura como atributo que causa impacto direto no resultado financeiro. Melhorar a qualidade dos produtos ou serviços, o tempo de entrega, a pontualidade da entrega e flexibilidade operacional, terá por certo impacto nos resultados da empresa em termos de receita, porém a influência na redução dos custos é imediata e direta.

É usual que as empresas de Telecomunicações tenham áreas responsáveis por desenvolver metodologias de Mapeamento e Levantamento periódico dos custos operacionais de cada

serviço ou produto e este estudo passa a ser a base para a elaboração da precificação destes produtos e serviços.

Como critério competitivo entre Redes o cálculo do custo se torna essencial. Uma vez que os serviços são disponibilizados sob topologias, tecnologias e capilaridade de redes diferentes.

Este trabalho utilizou-se de metodologia usualmente desenvolvida pelas Operadoras Brasileiras para definição do custo dos produtos ou serviços prestados. A metodologia detalhada neste trabalho está baseada no mapeamento da topologia média para cada serviço com as seguintes variantes:

- Por filial;
- Por velocidade;
- Por distância ou degrau tarifário.

A topologia média de um determinado produto é definida com base em uma amostra de circuitos, selecionada através de critérios estatísticos que garantam a representatividade, como:

- Intervalo de confiança: define a probabilidade de uma amostra “x”, retirada aleatoriamente de uma população “y” representar a população;
- Precisão: erro máximo;
- Tamanho da população;
- Tamanho da amostra: é considerado intervalo de confiança de 90% e de precisão de 25%.

Cada circuito da amostra é avaliado individualmente e são mapeados todos os equipamentos de rede que compõem o circuito.

#### 3.1.4.1 Mapeamento dos Equipamentos de Dados

A partir da amostra dos circuitos e da topologia média dos produtos, são levantados todos os equipamentos que compõem as Redes de Comunicação de Dados e seus custos de reposição levando-se em conta os seguintes critérios:

- Capacidade do equipamento;
- Quantidade de placas e portas;
- Composição média;
- Taxa de ocupação;
- Valores contratuais para aquisição e manutenção atualizados. Estão incluídos os custos com a manutenção da última milha, do suporte técnico e da reposição de material sobressalente.

#### 3.1.4.2 Mapeamento dos Equipamentos de Transmissão

Da mesma forma que é feito para os equipamentos de dados, são mapeados os equipamentos de transmissão (Rede de Transporte conforme item 2.2.3.5.2) usados como transporte para os serviços de comunicação de dados, bem como seu custo de reposição.

Devido à diversidade de equipamentos e a complexidade da rede, foram consideradas somente parte das Redes de Transporte, a saber:

- Equipamentos das localidades da amostra;
- Degrau médio entre *SDH* e *PDH*;
- Uso de rádio e fibra óptica;
- Uso de *PDH* e *SDH*.

#### 3.1.4.3 Mapeamento da Última Milha – Par Metálico ou Fibra Óptica

Na composição da solução padrão são consideradas duas situações para a última milha:

Velocidade até 2Mbps: utiliza-se par metálico na última milha;

Velocidades entre 34Mbps e 155Mbps: utiliza-se de fibra óptica na última milha, considerando um comprimento médio de três quilômetros de cabo com doze fibras;

Com os dados das amostras e os dados coletados junto aos sistemas de inventário, é possível montar a topologia de cada produto, por faixa de velocidade, degrau (distância) médio entre equipamentos de transmissão (*PDH* e *SDH*), fator de utilização e capacidade dos equipamentos.

Com estes dados é possível calcular, com grau de confiabilidade bastante próximo do real, o custo médio por produto ofertado, que pode variar conforme a velocidade, filial e distância.

O desdobramento destes custos não será explicitado no trabalho, porém, foram usados na para compor a base de dados do experimento. Estes dados são considerados confidenciais e por isso não serão publicados.

### **3.2 JUSTIFICATIVA DO USO DA REDE NEURAL**

Conforme visto nos parágrafos anteriores, todos os critérios apresentados até aqui dizem respeito às estratégias de operação de empresas de Manufatura. O que estas empresas normalmente fazem para determinar sua posição frente às concorrentes, são pesquisas de mercado.

Traçando um paralelo entre as empresas de manufatura semelhante e as Redes de Telecomunicações, observa-se que uma comparação pode ser feita sobre o produto ou serviço entregue: duas empresas concorrentes podem comercializar o mesmo produto, assim como duas redes podem prover o mesmo serviço.

No item 3.1, foram descritos os critérios competitivos usados comumente para comparar uma empresa com outra ou um processo operacional com outro, ao mesmo tempo foram definidos os critérios contextualizados para comparação de dois conjuntos de redes: *IP* e não *IP*.

Entretanto, surge um novo problema: de encontrar uma ferramenta capaz de comparar tais critérios para os dois conjuntos de redes. Ferramenta esta que não dependa de métodos subjetivos, como “pesquisas de mercado”, e não represente uma comparação estática, mas sim uma evolução comparativa, pois o objetivo final deste trabalho não é provar que uma

rede é melhor que a outra e sim, a partir de alguns critérios, por quanto tempo uma rede continuará sendo melhor que a outra.

Como comparar redes cujos critérios competitivos: nem sempre são lineares e dependem fortemente do comportamento dos atributos geradores? Não possuem regra de formação estatística sobre o ambiente no qual a rede está inserida? Estão contextualizados em ambientes extremamente dinâmicos?

Segundo o item 2.3, as Redes Neurais são ferramentas computacionais capazes de simular um processo de aprendizagem extremamente adaptativo aos parâmetros livres a ela impostos como entrada. Mais especificamente, o Aprendizado Hebbiano, pode ser facilmente adaptado quando se deseja comparar dois conjuntos de entradas sem qualquer relação aparente entre elas.

No contexto de redes visto no item 3.1, pode-se notar que o tempo de configuração de um serviço em uma filial aparentemente não deveria influenciar em nada o tempo de configuração de outro serviço em outra filial. Ou ainda, o custo de um serviço em uma filial não teria correlação aparente com a qualidade (confiabilidade) em outra filial.

Por outro lado, há situações em que esta aparente desconexão entre eventos nem sempre ocorre na prática. Por exemplo: o tempo de configuração de um serviço em duas filiais deferentes pode ser influenciado uma vez que uma mesma pessoa pode receber as duas tarefas para executar as configurações ao mesmo tempo. A quantidade de reclamações também pode influenciar no custo, uma vez que são necessárias mais pessoas para resolver os problemas reclamados no tempo de recuperação do serviço contratado.

Quando todos estes parâmetros, aparentemente sem relação entre si, forem usados como entrada de um algoritmo de aprendizado não-supervisionado ou auto-organizado, as RNAs conhecidas como Mapas Auto-Organizados de *Kohonen* (*Self-Organizing Map – SOM*), como visto no item 2.3.1.3, dão início ao princípio do postulado de Hebb.

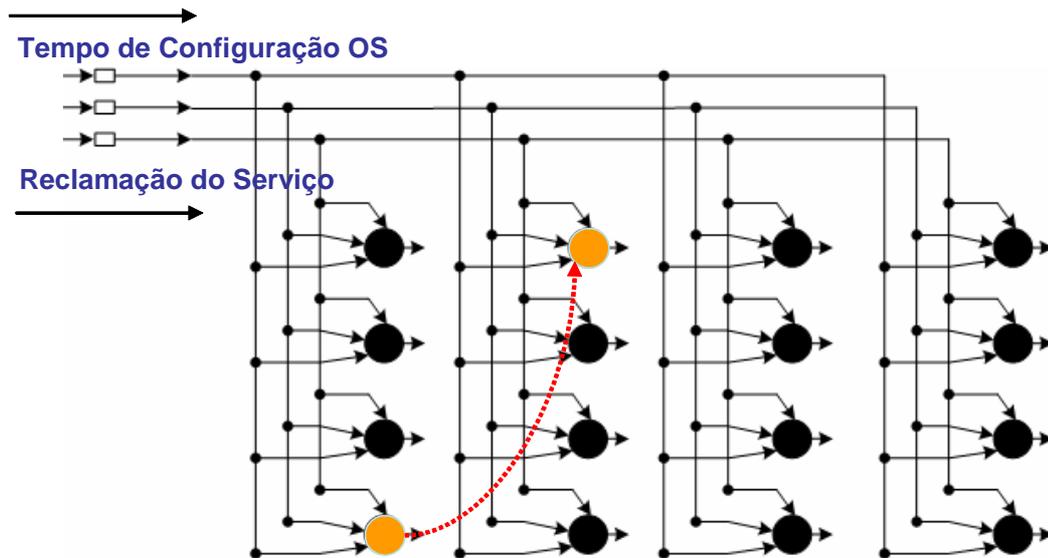


Figura 3.7 - Treliça de Neurônios e uma conexão sináptica reforçada entre dois neurônios [CAST2001]

Exemplificando, na Figura 3.7 imagina-se que a RNA recebeu ao mesmo tempo dois componentes quaisquer que serão comparados, para fins deste exemplo o tempo de configuração e a quantidade de reclamação. Estas duas entradas vão para vários neurônios, entretanto, dois deles estavam ativos no momento deste estímulo seja por ter recebido outras entradas ou qualquer outro motivo. Logo, a rede vai se comportar como se já existisse uma conexão sináptica entre eles e esta será reforçada a cada vez que a situação dos estímulos se repetirem.

Na prática, tempos de configuração muito altos podem ser um indício de que houve problema na configuração. Conseqüentemente pode gerar maior quantidade de reclamações, as quais levam ao aumento do custo de manutenção do serviço.

No próximo capítulo dois conjuntos de critérios competitivos serão apresentados sob forma de conjuntos de vetores. Estes vetores serão utilizados como entrada em uma ferramenta computacional – GHAPCA (Algoritmo Hebbiano Generalizado para Extração dos Componentes Principais) – a qual aplicará a *KLT*.

Para os critérios influenciados, depois de submetidos a RNA, atribui-se o termo “dopado”. Este conceito será aplicado apenas a este trabalho e significará o critério, que poderá ser: tempo de provisionamento, custo do serviço, quantidade de reclamação ou eficiência de transmissão, influenciados uns pelos outros no domínio da *KLT*. Assim, entende-se como

“tempo dopado” o resultado do tempo influenciado pelos outros critérios competitivos quando aplicada a *KLT*, uma vez que ele foi “dopado” pelos outros critérios. A nomenclatura se estende para “custo dopado” e “quantidade de reclamações dopada”.

Este procedimento será detalhado no capítulo 4.

## 4 EXPERIMENTO

De posse de todas as ferramentas, de todo o embasamento teórico e de toda a situação contextualizada, finalmente, pode-se agora iniciar o experimento de predição.

Este experimento, para melhor estruturação e entendimento, foi dividido em três fases:

- Obtenção de dados;
- Transformação de domínios;
- Linearização e predição.

A obtenção de dados está embasada como descrito no item 3.1, a partir do levantamento das dimensões competitivas no contexto de Telecomunicações.

A transformação de domínios nada mais é do que a normalização e aplicação dos dados decorrentes das ditas dimensões competitivas como entrada no Algoritmo GHAPCA e a obtenção destes transformados para o domínio de *Karhunen-Loève*.

A linearização e predição são as últimas atividades e nestas são obtidos os períodos estimados da migração dos serviços.

A seguir, estas fases serão detalhadas.

### 4.1 OBTENÇÃO DOS DADOS

Com relação às dimensões competitivas abordadas até aqui – qualidade (confiabilidade e desempenho), velocidade e custo, traduzidos para o contexto de Telecomunicações, pode-se esperar que todas contribuam para a variabilidade dos custos operacionais de uma empresa. Entende-se como cenário ideal o desenvolvimento de uma Rede única capaz de suportar toda gama de serviços da empresa, com procedimentos operacionais padronizados e aprovisionamento o mais automatizado possível.

Na área de Operações das Empresas de Telecomunicações, observa-se que a revisão dos fluxos operacionais e melhoria contínua dos processos visam principalmente à redução de custos e o aprimoramento da qualidade do serviço prestado. Entende-se como imperativa a

manutenção das redes legadas em função de serem estas redes que suportam os serviços ditos “ganhadores de pedidos”. Ou seja, apesar dos avanços tecnológicos, o mercado, principalmente o mercado brasileiro, é bastante conservador na escolha das soluções para transferência de informações entre suas redes privadas.

Assim, a redução no custo operacional com manutenção de contratos de suporte, sobressalentes e mão-de-obra especializada de várias redes diferentes que prestam serviços semelhantes tornou-se um grande desafio ao longo do tempo.

Com este objetivo, procuraram-se reunir todas estas informações de custo de implantação e manutenção dos serviços de comunicação de dados, da eficiência de transmissão das topologias de entrega dos serviços, dos tempos de provisionamento dos serviços e da quantidade de reclamação dos clientes que usam os serviços. Informações que, aparentemente, não têm relação entre si e compará-las no tempo, utilizando como ferramenta as RNAs de Aprendizado Hebbiano, como já detalhado nos capítulos anteriores.

Espera-se com isso predizer, de forma estimada, a data em que os serviços transportados por Redes não *IP* passem a ser transportados unicamente por Redes *IP*. Esta estimativa será construída a partir da sobreposição de gráficos no domínio do tempo.

Com base nas informações coletadas obteve-se um histograma da quantidade de tarefas de configuração e provisionamento executadas no período de 28/12/2004 a 28/07/2007, conforme mostrado na Figura 4.1, onde o eixo vertical representa a quantidade de tarefas executadas e o eixo horizontal representa a cronologia destas execuções.

Conforme a legenda, a curva superior mostra a quantidade de tarefas executadas referentes a serviços não *IP* e a curva inferior, por sua vez, mostra a quantidade de tarefas executadas referentes a serviços *IP*.

Durante o período observado, percebe-se uma tendência de diminuição da quantidade de tarefas de configuração para serviços *IP* e um pequeno aumento para os serviços não *IP*. Isto denota, claramente, que ainda existe forte demanda por serviços não *IP* o que pesa seriamente na decisão de diminuir investimentos de expansão e manutenção das redes que suportam exclusivamente serviços não *IP*.

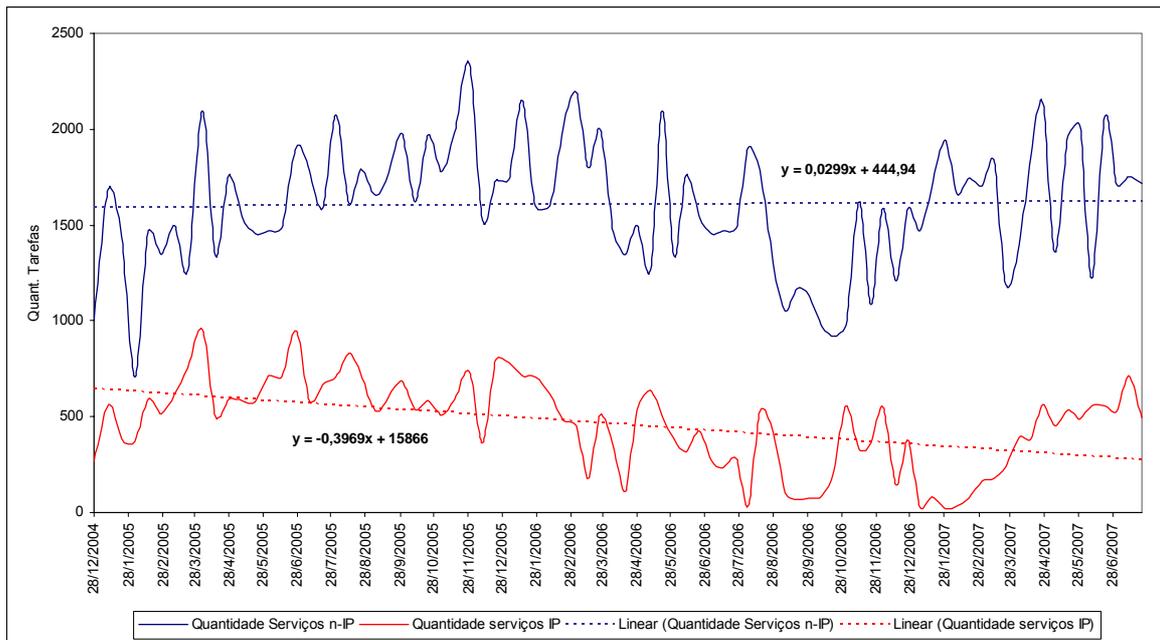


Figura 4.1 – Histograma da quantidade de tarefas de configuração e provisionamento

Dentro do objetivo de se comparar os critérios competitivos dos dois conjuntos de redes, serão comparados os tempos de configuração, a quantidade de reclamações e os custos dos serviços não *IP* e as mesmas dimensões “dopadas” entre si. Lembrando que o termo “dopado” é atribuído ao resultado de um determinado critério influenciado pelos outros critérios competitivos no domínio da *KLT*. Da mesma forma será feita a comparação para os serviços *IP* e o ponto onde estes dois gráficos “dopados”, para cada dimensão, se cruzarem no tempo será a data estimada de migração.

Adicionalmente, vê-se que os fatores que estarão sendo comparados têm uma influência diferente no tempo para cada filial, de modo que a migração deve ser progressiva e em tempos diferentes respeitando a particularidade de cada uma.

Observa-se que a execução de tarefas de provisionamento, quando aplicado no domínio do tempo, não espelha o seu grau de influência na entrega do serviço a não ser que esteja acompanhada de outros fatores. Por exemplo: a filial, a hora do dia, o dia da semana ou o dia do mês em que a tarefa está sendo executada, ou até mesmo a banda contratada para o produto.

Por outro lado, se, de alguma forma, fosse possível atribuir um valor adicional ao tempo de execução da tarefa, de modo que um único valor demonstrasse a importância da duração deste evento e este valor levasse em consideração tantos fatores quantos fossem

necessários, seria mais fácil determinar o período no mês onde as atividades têm maior rendimento ou a hora do dia em que o rendimento é pior. Ou ainda, qual filial estaria apta para iniciar a convergir todos os serviços para uma única rede.

Da mesma forma, esta ferramenta, quando aplicada aos custos e quantidade de reclamações, poder-se-ia descobrir a que quantidade de reclamações estão associados os maiores custos.

De acordo com os itens 2.2, 3.1.1.1.2, 3.1.1.2, 3.1.3 e 3.1.4 foram obtidos os dados da seguinte forma:

Somatório dos tempos de configuração de serviço: para obtenção do tempo médio para a tarefa de configuração do serviço, parâmetro relevante para medir os tempos de provisionamento de serviços *IP* e não *IP* foi necessário pesquisar na base de dados do sistema *OMS (Operational Management Services)*, o qual é o sistema de gerenciamento das Ordens de Serviço (OSs) de uma Operadora Brasileira, os arquivos de tramitação das OSs no período de janeiro/2005 a julho/2007 e posteriormente filtrar as tarefas relevantes para esta medição.

Foram obtidos dois arquivos: total de Ordens de Serviço no período e a tramitação de cada Ordem de Serviço. Isto por que, no banco de dados de tramitação não constava a família do produto, dado relevante para determinar a topologia de rede para aquele serviço. O objetivo deste levantamento foi obter a evolução dos tempos de configuração para todos os serviços.

O Banco de Dados consolidado até aqui possuía o conjunto de coeficientes considerados para a predição da data de migração, a saber:

- Filial;
- Número do Circuito;
- Início da tarefa (hora, dia, mês e ano);
- Final da tarefa (hora, dia, mês e ano);
- Tempo total na tarefa;

- Família do Produto.

Conforme a metodologia aplicada, o conjunto de coeficientes foi considerado um vetor e, portanto, teve de ser normalizado para que todos ficassem em uma mesma faixa de magnitudes, de modo que um componente do vetor não tenha magnitude maior e assim uma influência maior. Desta forma foi feita a normalização dos dados.

De posse dos dados acima foi feito o primeiro teste no Algoritmo GHAPCA (ver Apêndice). Esta primeira tentativa de predição não teve sucesso pela falta de um critério único de normalização.

Posteriormente, foram inseridos os demais coeficientes contextualizados no capítulo 3 (custo, eficiência de transmissão e quantidade de reclamações).

Quantitativo de reclamações: para obtenção do quantitativo de reclamações ou bilhetes de defeito (BDs) pesquisou-se na base de dados do sistema SID (Sistema de Indicadores de Dados) da mesma Operadora na qual se aplicou o experimento. Foram levantados todos os bilhetes de defeito registrados no período janeiro/2005 a julho/2007, para cada serviço configurado. Desta forma, obteve-se a média de reclamações para este serviço configurado entre o encerramento de sua configuração e julho de 2007.

Igualmente ao que foi feito para tramitação das OSs, foi selecionado o conjunto de coeficientes mais relevantes:

- Filial;
- Circuito;
- Início da falha (hora, dia, mês e ano);
- Encerramento da falha (hora, dia, mês e ano);
- Tempo de reparo;
- Produto;
- Quantitativo de falhas.

Custo de operação e uso de rede: foram levantados os dados de evolução do custo por produto. Estes dados foram obtidos conforme metodologia apresentada no capítulo 3 e não serão explicitados aqui por requisitos de confidencialidade.

Eficiência de transmissão: foram calculados os dados de eficiência de transmissão por serviço conforme descrito e apresentado no capítulo 3.

Um breve resumo dos dados apresentados até aqui pode ser visto na Figura 4.2:

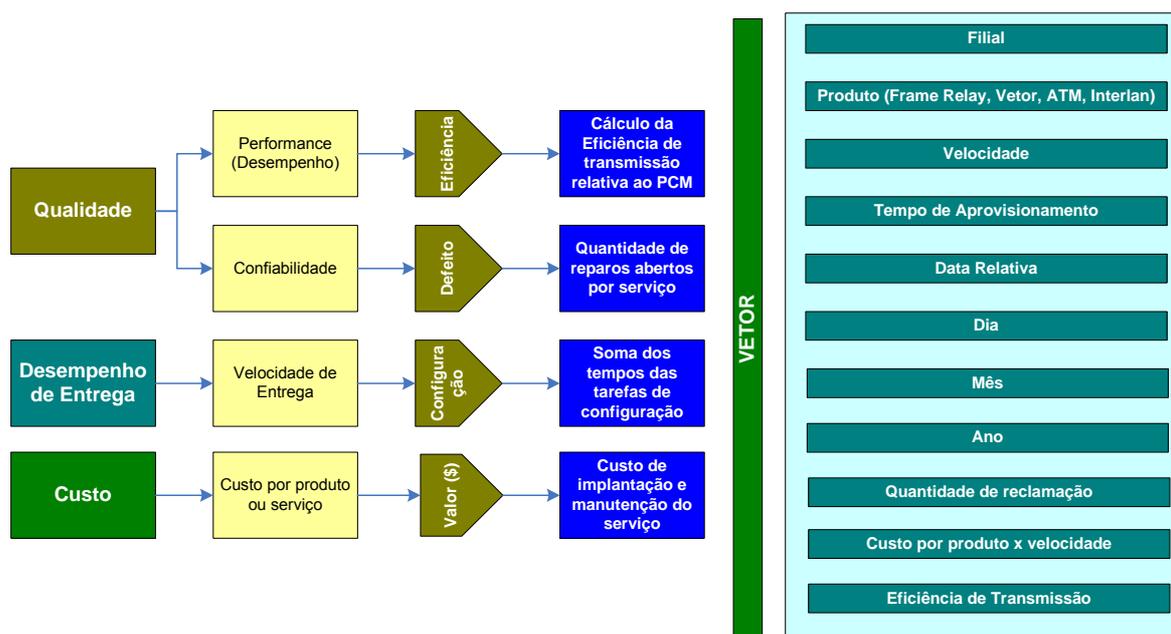


Figura 4.2 – Diagrama do vetor de entrada do modelo de predição [AUTORA]

Na citada Figura 4.2, observa-se que os três fatores de entrada são os critérios competitivos de operações usados em qualquer empresa como tradução da estratégia competitiva da empresa. Estes critérios foram desdobrados nas suas dimensões e quando contextualizados para Telecomunicações (Eficiência de Transmissão, Quantidade de defeitos, Tempo de configuração e Custo de operação e uso de rede) foram selecionadas aquelas dimensões que mais se adequavam ao novo cenário. As dimensões contextualizadas geraram um vetor.

Este vetor será o elo entre esta e a próxima fase: Transformação de domínios.

## 4.2 TRANSFORMAÇÃO DE DOMÍNIOS

Uma vez construído o vetor observou-se que as magnitudes de cada componente: produto, filial, velocidade, tempo de provisionamento, etc. não podiam ser comparadas uma vez que possuíam referenciais de magnitude diferentes. A solução encontrada para este problema foi a normalização baseada na divisão de segmentos paralelos, onde a magnitude máxima de cada componente não ultrapassa o valor unitário e é feita de acordo com a Equação 4.1.

$$ValorNormalizado = \frac{Valor - Valormenor}{Valormaior - Valormenor} \quad \text{Equação 4.1}$$

Onde:

Valor Normalizado é o valor entre zero e um aplicado à Transformada; valor maior é o valor de maior magnitude de um critério competitivo, da mesma forma que o valor menor é o de menor magnitude; e o valor é a magnitude do critério competitivo para o vetor de entrada em questão.

Cada linha do arquivo de entrada foi considerada um vetor de doze dimensões. Esta normalização nada mais é do que deixar os dados em uma magnitude semelhante. Conforme já mencionado, a necessidade desta atividade leva em conta que os valores, quando apresentados como entrada no treino de uma rede auto-organizada, tendem a influenciar uns aos outros de modo a construir um eixo onde os autovalores tendem igualmente a serem máximos. Assim, a normalização faz com que, desta forma, as influências sejam proporcionais.

Com a devida permissão do autor: Fernando de Castro [CAST2001] foi obtido o algoritmo de treino para a transformada de *Karhunen-Loève* em forma de código fonte. Igualmente, o autor forneceu o compilador mais apropriado para a construção.

O processo de instalação do DJGPP se iniciou com o *download* do arquivo DJGPP.zip obtido através da página <http://www.ee.pucrs.br/~decastro/download.html>. O código foi adaptado para este trabalho, uma vez que o original foi desenvolvido para 65.536 linhas.

Os resultados foram plotados em gráficos para melhor visualização.

Foram criados dois arquivos de vetores de entrada: arquivo *IP* e arquivo não *IP*. Cada linha (vetor) foi composta por doze números com magnitudes entre zero e um. Vale ressaltar que após a normalização da eficiência de transmissão outro ajuste se fez necessário. Em todas as magnitudes quanto maior era esta, pior era a condição de rede, o que não é verdade para a eficiência de transmissão. Assim, os componentes de eficiência foram multiplicados por menos um (-1), de modo que, quanto menor a eficiência maior a magnitude do componente vetorial, ficando todos os componentes do vetor com regra de formação semelhante.

O Algoritmo GHAPCA é acionado de forma bastante simples. A chamada de um arquivo de MS-DOS seguido de um arquivo com dados de configuração abre o programa e este transforma a matriz com os componentes vetoriais em uma nova matriz do mesmo tamanho da matriz de entrada.

A matriz vetorial resultante recebe o nome, neste trabalho, de matriz “dopada”. Esta dopagem é a influência dos componentes vetoriais entre si.

O arquivo de configuração (hpar.txt) tem o formato mostrado na Figura 4.3:

```
saida02.txt % Training Data Vector Set (Input File).
saida02r.txt % Rebuilt Data Vector Set from the S sub-spaces (Output File).
cnnsf.txt % Converged Neural Net specification (Output File).
7 % The desired number S of sub-spaces used to decompose the input Data Vector
Set (= number of neurons in the Neural Net).
1E-3 % Initial Learning Rate.
3 % Training window size (= number of neurons simultaneously trained).
1000 % Ratio Alpha = [(Eigenvalue)^-1]/LearningRate {Note1: Alpha=0 keeps constant
the neurons learning rate (they are not adaptively updated according to the instant
eigenvalue), however they can be manually
changed along training by invoking the change menu. Note2: A lower Alpha value yields a
higher convergence speed - usually 500 < Alpha < 2000 }.
0.5 % Convergence criterion C ( C is the percent synapses vector squared Euclidean
norm change value, below which the neuron synapses vector is regarded as converged to an
eigenvector).
100000 % Max number of training epochs.
1.0 % The value R which defines the synapses random init interval [-R,R] (Note: R
is further averaged by the number of synapses per neuron).
```

Figura 4.3 – Arquivo hpar.txt [CAST2001]

Os dois primeiros parâmetros da configuração são os arquivos de vetores de entrada e saída.

O arquivo “cnnsf.txt” carrega os dados de resultados da convergência da Rede Neural.

O parâmetro “7” foi o número de subespaços usados na decomposição dos dados dos vetores de entrada, em outras palavras, número de neurônios de entrada. Foram feitas

várias tentativas mudando este parâmetro, no entanto, com este número de subespaço os resultados foram alcançados com a menor quantidade de épocas de treino.

O parâmetro “1E-3” é a taxa inicial de aprendizado. É o quanto se atribui de variação de aprendizado para o primeiro vetor de subespaço.

O parâmetro “3” é o número de neurônios treinados ao mesmo tempo. Do mesmo modo que o parâmetro “7” este foi o que melhor apresentou resultados em termos de quantidade de épocas de treino.

O próximo parâmetro alterado foi o máximo número de épocas, “10.000”. O valor utilizado foi o de “1.000”, mas no pior caso o máximo número de treinos dos neurônios (épocas) nunca passou de 271. Este parâmetro garante que mesmo que a transformada não venha a convergir, o algoritmo não fique calculando os vetores de saída indefinidamente.

O resultado desta fase são dois conjuntos de vetores listados em dois arquivos: vetores **IP dopados** e vetores **não IP dopados**, a serem utilizados na próxima fase.

### 4.3 LINEARIZAÇÃO E PREDIÇÃO

Os arquivos de saída do Algoritmo GHAPCA têm o formato mostrado na Figura 4.4:

Filial	Produto	Tipo de Prod.	Efic. TX	Velocidade	Custo	Ano	Mes	Dia	Temp. Conf.	Data	Quant. Rec.
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,300226	0,200327	0,0403823	-0,159538	0,000340853	0,00130705	-0,00026362	0,80974	0,899775	0,0337375	0,151505	0,0292065
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349
0,896964	0,200355	0,0427741	-0,158314	0,0123932	0,00903263	-0,00985752	0,811767	0,899672	0,0275538	0,142879	0,00957349

Figura 4.4 – Arquivo de saída do programa GHAPCA

Especificamente o décimo componente de cada linha é o tempo de configuração “dopado”. O décimo primeiro componente é a data na qual foi realizada a tarefa de configuração. A partir daí se adotou o seguinte procedimento:

Do arquivo de entrada podem ser retiradas diversas informações para compará-las no domínio da *KLT* (informação dopada). Um dos parâmetros que mais interessam para o

presente trabalho é o tempo de configuração normalizado e do arquivo de saída o tempo de configuração “dopado”. Tanto para produtos *IP* quanto não *IP*.

Como várias atividades foram executadas, em alguns casos, no mesmo dia, uma plotagem direta destes valores geraria um gráfico com várias barras verticais onde em um mesmo dia seria apresentado o valor máximo e mínimo dos tempos de configuração. Com o intuito de evitar este problema, optou-se por calcular a média diária dos tempos de configuração não “dopados” e “dopados”, tanto para produtos *IP* quanto para os não *IP*.

A partir desta decisão plotou-se os gráficos comparativos dos tempos de configuração *IP* e não *IP* para os tempos de configuração “dopados” e não “dopados” conforme a Figura 4.5 e a Figura 4.6:

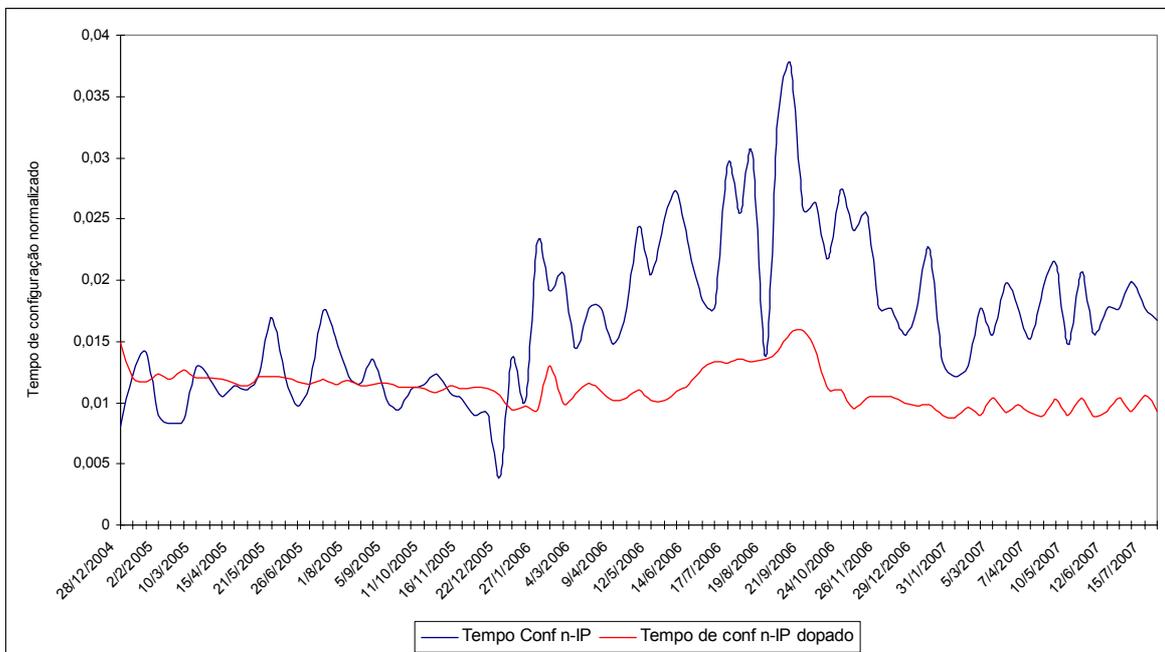


Figura 4.5 – Tempo de configuração normalizado não IP dopado e não dopado

Onde, o eixo horizontal representa a evolução das datas de encerramento das OS (Ordens de Serviço) e o eixo vertical representa os tempos de configuração normalizados.

A decisão de usar a data de encerramento das OSs resulta do fato de que apenas a partir desta data são contabilizados os Bilhetes de Defeito (BDs) ou reclamações que, por ventura, o cliente venha a fazer.

As linhas em vermelho são as curvas dos tempos de configuração normalizados e “dopados”, enquanto as linhas azuis representam os tempos de configuração normalizados não “dopados”. Optou-se por manter os valores normalizados a fim de facilitar a comparação.

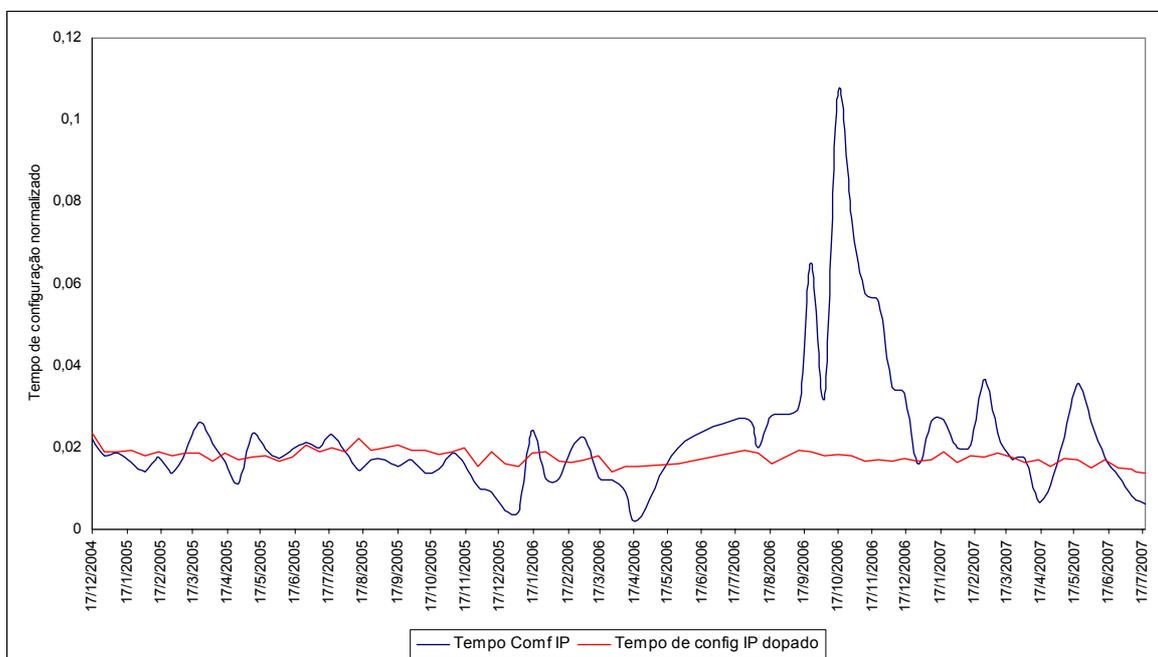


Figura 4.6 – Tempo de configuração normalizado IP dopado e não dopado

Percebe-se que as curvas “dopadas” parecem mais “imunes” a variações diretas nos tempos de configuração, mas demonstram uma tendência mais estável uma vez que levam em conta outros componentes que não exclusivamente o próprio tempo de configuração.

Apenas como exemplo, no período de junho de 2006 a setembro de 2006, houve uma alteração estrutural na empresa onde se fez a medição de modo que as atividades de configuração, antes pulverizadas pelas várias filiais, foram centralizadas em um único Centro de Operação e Manutenção de Rede. Este fato pode ser facilmente percebido nos tempos de configuração não “dopados”, e de certa forma, também nos “dopados” mesmo que de forma mais suave.

O que de importante se comprova, é que a Rede Neural conseguiu influenciar o tempo de configuração com outros componentes que não exclusivamente ele próprio.

No tempo de configuração não “dopado” não se percebe uma tendência de queda, entretanto, impulsionado pela redução de custos em contratos, evolução da eficiência de

transmissão por troca de Codecs, novos usuários adquirindo serviços *IP* em detrimento de serviços não *IP*, quantidade de reclamações, banda transportada, entre outros, acabaram por forçar uma leve tendência de queda no tempo de configuração “dopado”, tanto *IP* quanto não *IP*.

A predição da migração dos serviços não seria possível utilizando este método caso fosse feita a partir dos tempos de configuração não “dopados”. Já a comparação entre as retas de tendência pode ser feita muito mais facilmente sobre os gráficos dos tempos de configuração “dopados” *IP* e não *IP*, como mostrado na Figura 4.7, onde se observa, particularmente, duas retas com suas respectivas equações. Elas representam a tendência linear de cada curva “dopada”.

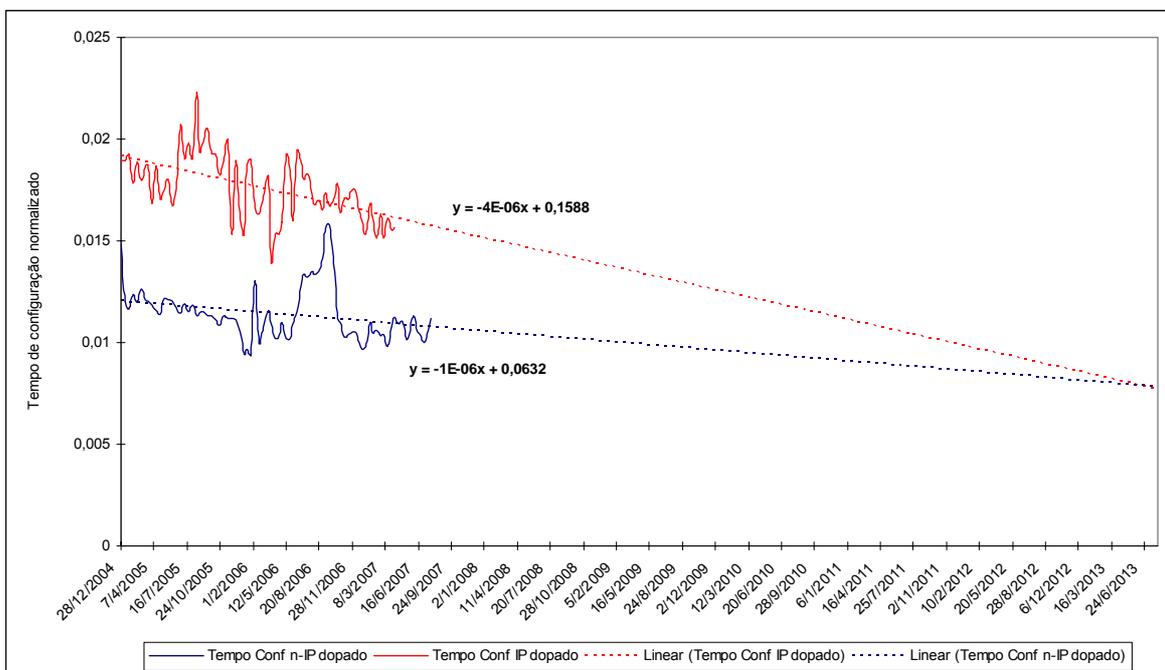


Figura 4.7 – Comparativo entre tempo de configuração dopado IP e não IP

A decisão de utilização da curva de tendência linear está suportada pela importância da predição a longo prazo. Uma curva de tendência polinomial mostraria uma tendência relativa apenas a última parte da curva proporcional ao polinômio empregado, em outras palavras, uma curva de tendência de segundo grau mostraria a tendência baseada apenas a segunda metade da curva, o que não é o objetivo desta predição.

A próxima etapa seria a determinação da data onde as duas retas se cruzam. Isto pode ser obtido a partir das equações das retas de tendência.

Conforme [SPIEGEL1992], a equação simplificada da reta é:

$$y = mx + b \quad \text{Equação 4.2}$$

Sendo “ $m$ ” a inclinação da reta. Considerando dois pontos  $(x_1, y_1)$  e  $(x_2, y_2)$ , tem-se:

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{Equação 4.3}$$

e

$$b = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1} \quad \text{Equação 4.4}$$

Logo, tem-se a equação da **reta IP**  $y = -4 \cdot 10^{-6} x + 0,1588$  e a equação da **reta não IP**  $y = -10^{-6} x + 0,0632$ .

Resolvendo o sistema de equações o valor de  $x$ , ou seja, a data de cruzamento entre as retas, ficou em torno de **junho de 2013**.

Uma tentativa de predição poderia ser feita a partir dos dados de tempo de configuração não dopados, ou seja, não influenciados pelos demais critérios de comparação. Usando da mesma metodologia, uma nova plotagem desta predição pode ser vista na Figura 4.8.

As retas de tendência têm a mesma inclinação, tanto para os tempos de configuração *IP* quanto para o não *IP*, impossibilitando a predição.

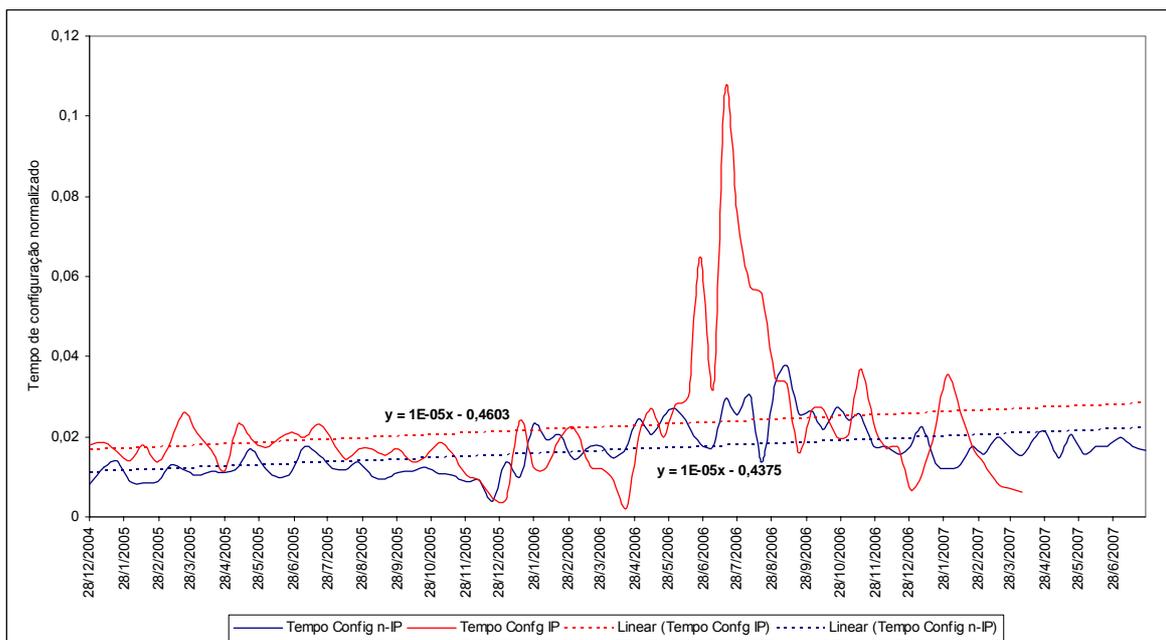


Figura 4.8 – Tempo de configuração IP e não IP

Da mesma forma que foi calculada a data de predição a partir do tempo de configuração “dopado” pode-se obter as predições a partir do custo de implantação e manutenção de um serviço e a quantidade de reclamações deste mesmo serviço após ele ter sido implantado.

Desta forma, o gráfico da Figura 4.9 representa o custo médio diário dos serviços não *IP* e o custo médio diário dos serviços não *IP* “dopados”. Foi calculado o custo médio diário levando-se em conta todos os serviços que tiveram sua configuração encerrada na data conforme o eixo horizontal.

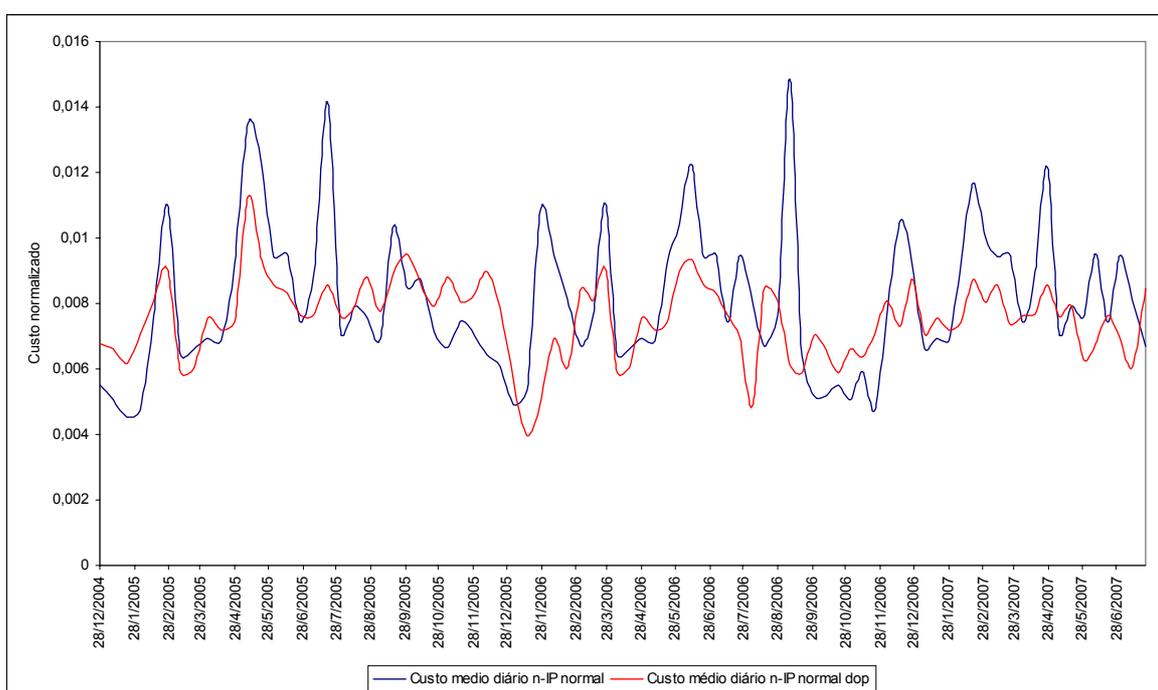


Figura 4.9 – Custo de implantação e manutenção de serviços não IP dopado e não dopado

Na Figura 4.10, se representa o custo médio diário dos serviços *IP* e o custo médio diário dos serviços *IP* “dopados”. O modo de cálculo foi o mesmo descrito para os serviços não *IP*.

Assim como foi calculado para os tempos de configuração e pelo mesmo motivo, os dados foram normalizados. Logo, o eixo vertical representa o custo normalizado para os dois conjuntos de redes.

Dando continuidade à metodologia de predição a Figura 4.11 representa a comparação entre o custo de implantação e manutenção de serviços *IP* e não *IP* dopados e normalizados.

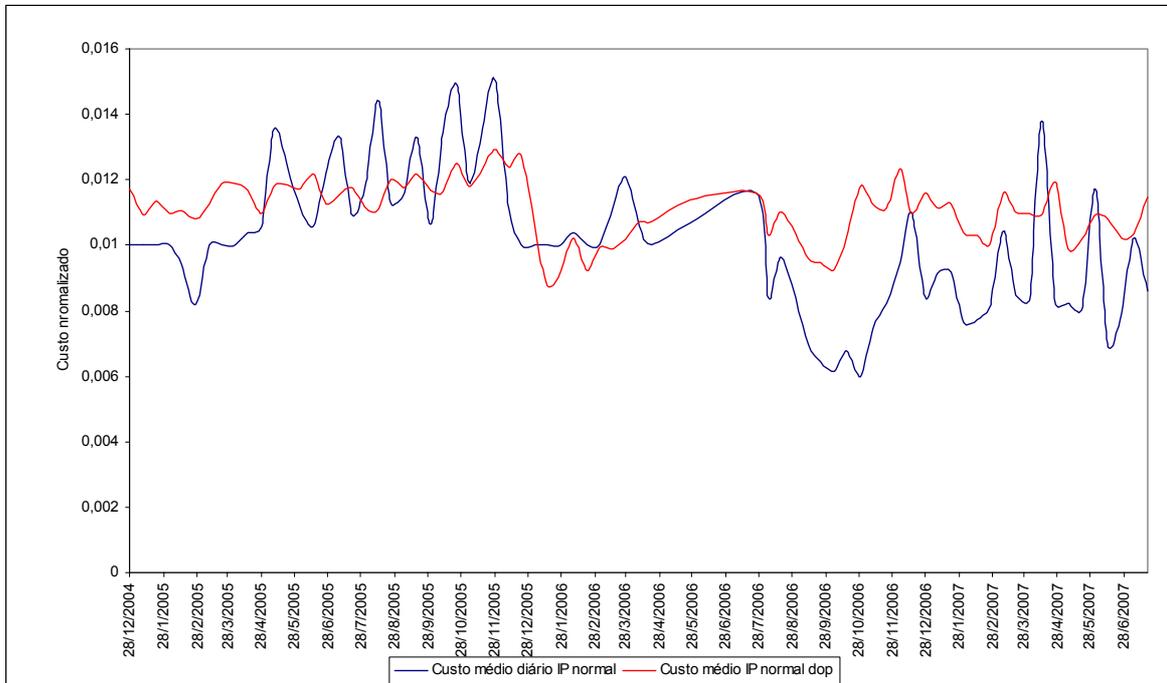


Figura 4.10 – Custo de implantação e manutenção de serviços IP dopado e não dopado

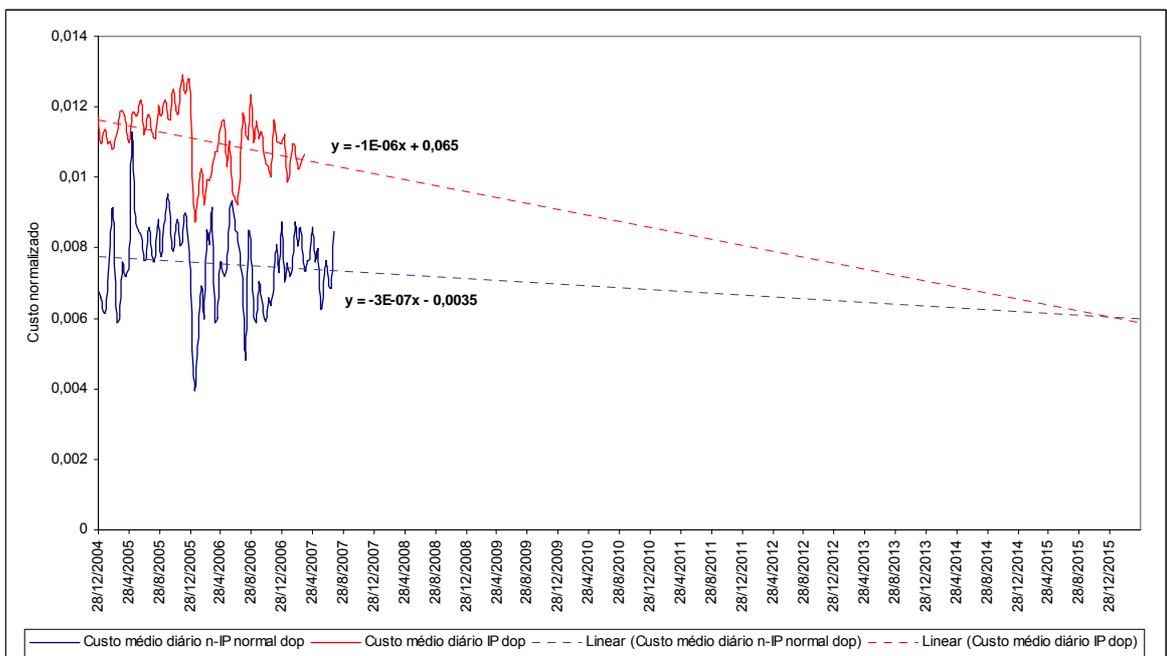


Figura 4.11 – Custo de implantação e manutenção de serviços IP e não IP dopados

Usando [SPIEGEL1992] e com base nas equações das retas  $y = -10^{-6}x + 0,065$  e  $y = -3 \cdot 10^{-7}x - 0,0035$ , a data de cruzamento entre as retas, ficou em torno de **novembro de 2015**. Praticamente dois anos após a predição baseada no tempo de configuração “dopado”.

É válido, ainda, o comparativo entre as curvas de custo de médio diário não *IP* e *IP*, ambos não “dopados”, conforme mostra a Figura 4.12.

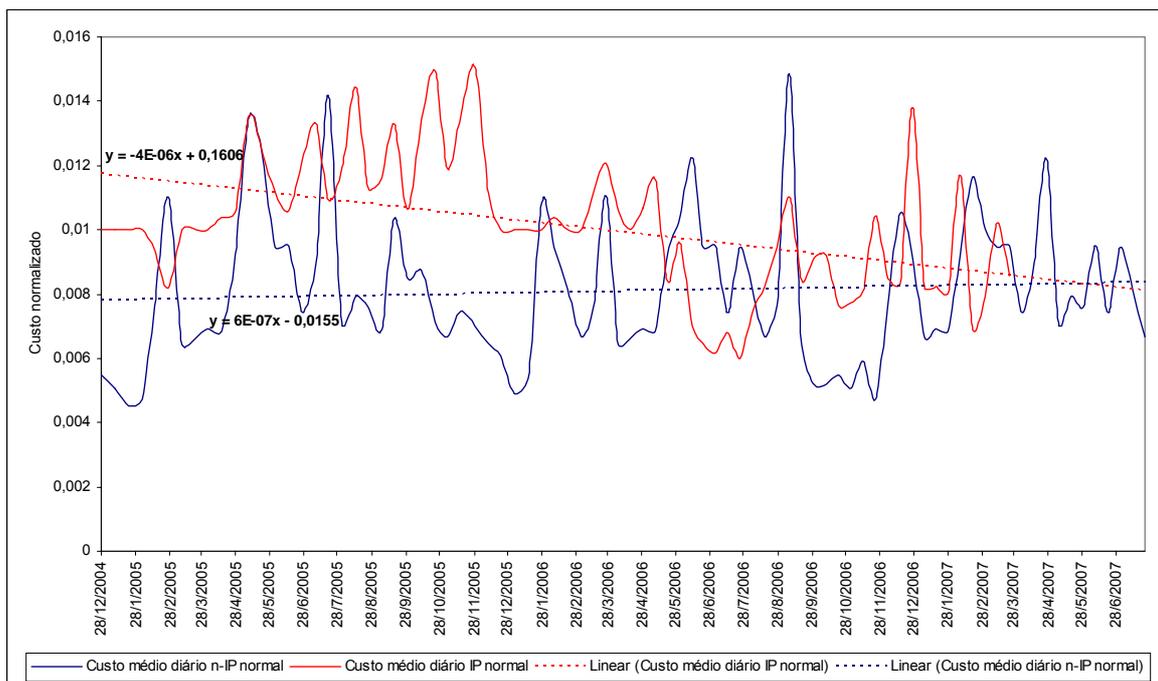


Figura 4.12 – Custo de implantação e manutenção de serviços IP e não IP

Para este, a migração poderia ter iniciado em **maio de 2007**.

Ainda seguindo a metodologia, a Figura 4.13 apresenta a quantidade de reclamações de serviços não *IP* dopado e não dopado. Enquanto a Figura 4.14 representa a quantidade de reclamações de serviços *IP* dopado e não dopado. Lembrando que esta medida expressa a quantidade de reclamações de defeitos no período de janeiro/2005 a julho/2007 e que foi aplicada à média de defeitos diários por circuitos.

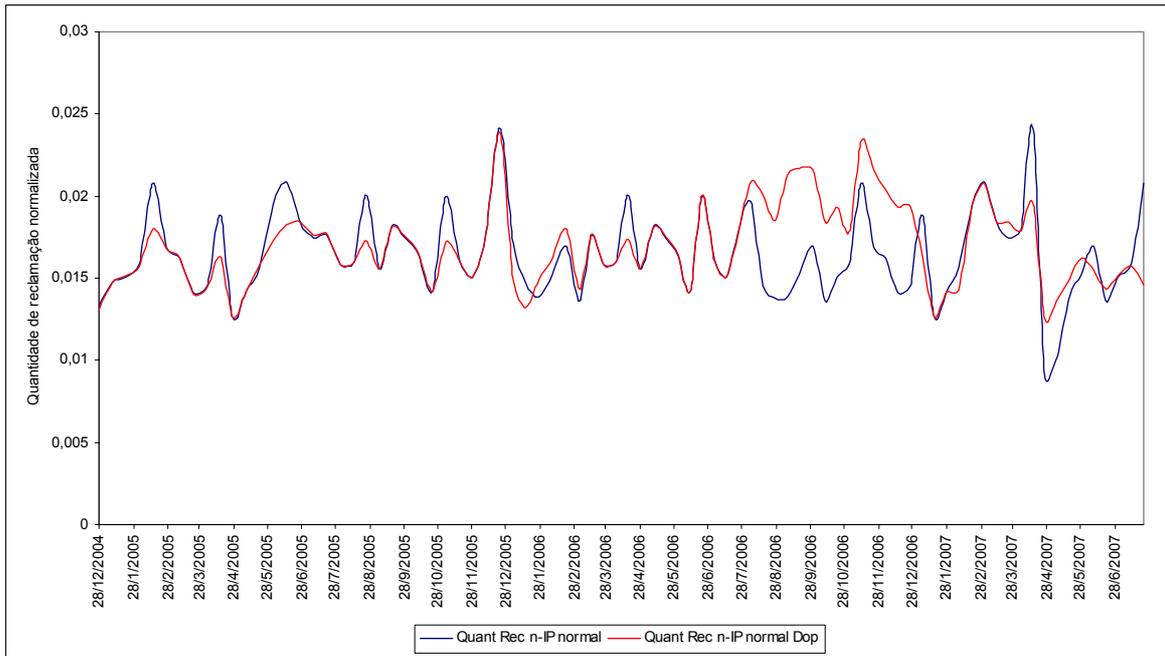


Figura 4.13 – Quantidade de reclamações de serviços não IP dopado e não dopado

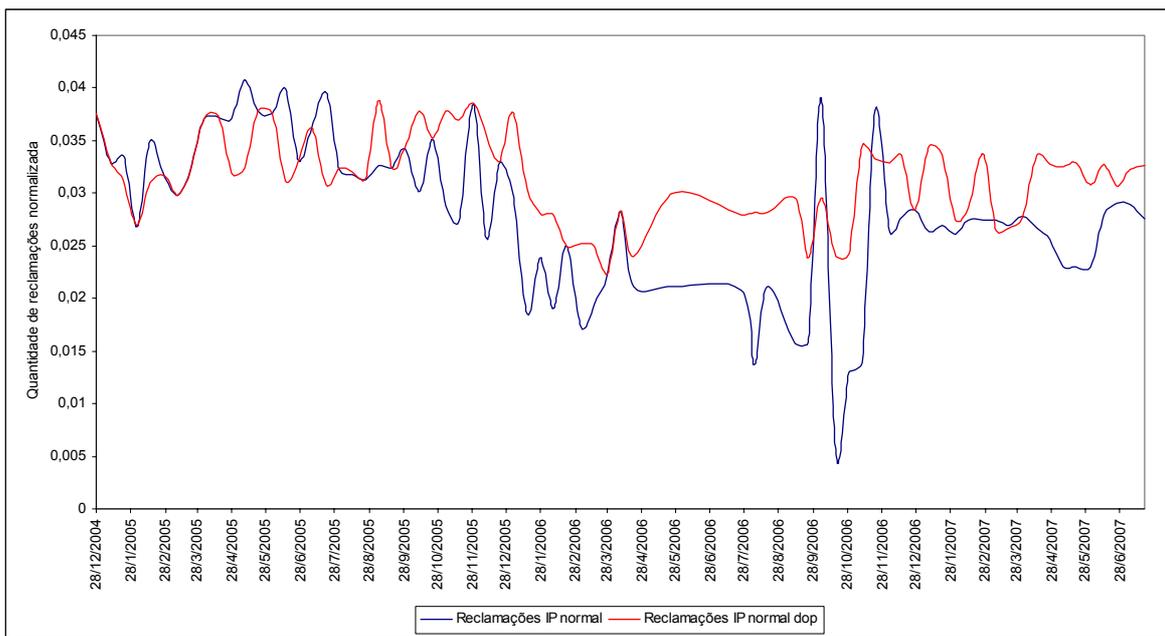


Figura 4.14 – Quantidade de reclamações de serviços IP dopado e não dopado

Repetindo o uso de [SPIEGEL1992] e com base nas equações das retas  $y = -4 \cdot 10^{-6} x + 0,1677$  e  $y = 10^{-6} x - 0,0388$ , da Figura 4.15, a data de cruzamento entre as retas, ficou em torno de **agosto de 2015**. Predição próxima aquela utilizando os tempos de configuração “dopados”.

E, por último, o comparativo entre as curvas de quantidade de reclamações não *IP* e *IP*, ambos não “dopados”, conforme mostra a Figura 4.16.

O cruzamento das curvas ocorre por volta de **março de 2008**.

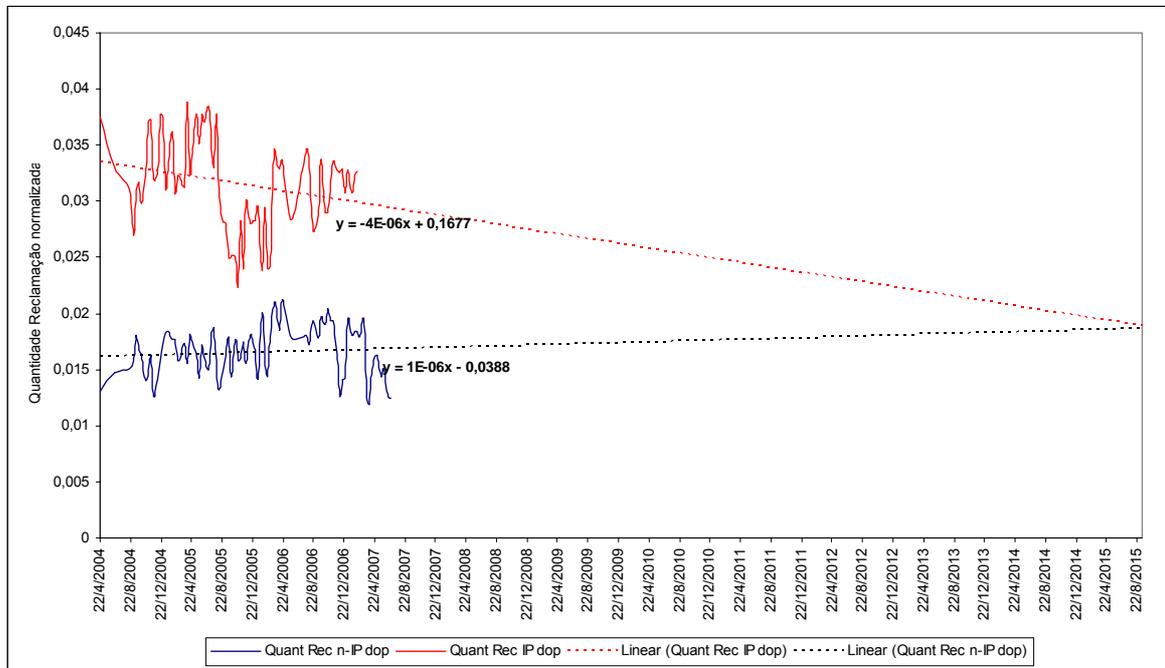


Figura 4.15 – Quantidade de reclamações de serviços não IP e IP dopado

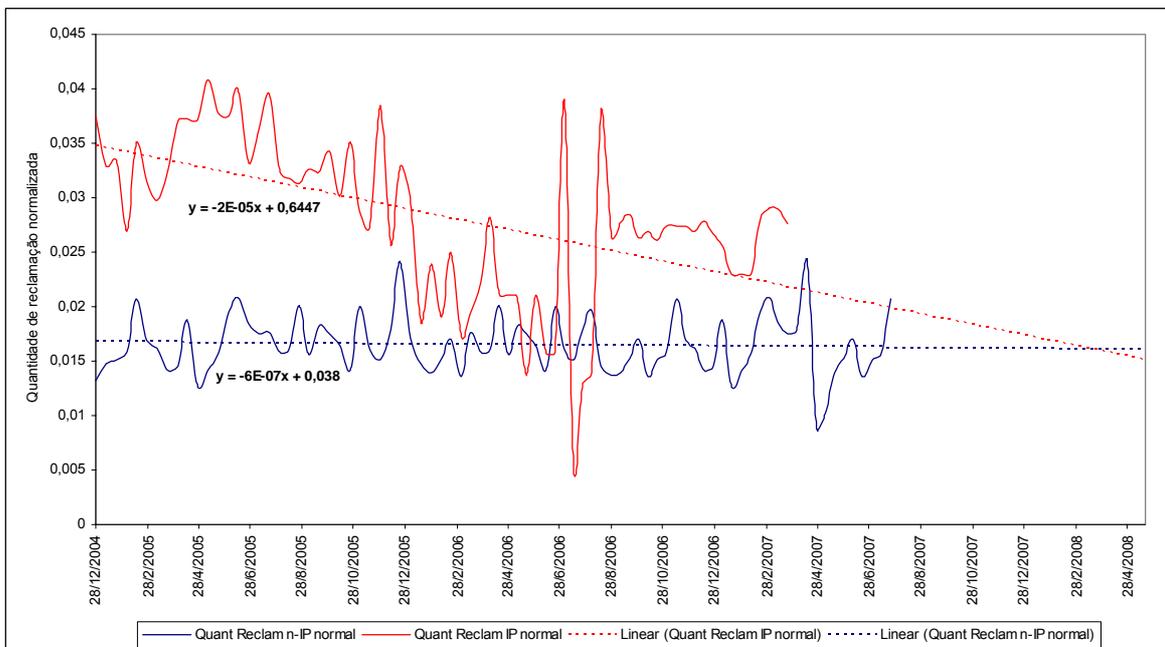


Figura 4.16 – Quantidade de reclamações de serviços não IP e IP não dopado

Além da data de predição é possível obter outros subprodutos das matrizes de vetores “dopadas”. É possível descobrir, por exemplo, em que filial existem mais produtos *IP* sendo configurados que em outras e por onde a migração ocorreria de forma mais viável, pois, em algumas filiais outros componentes vetoriais como: banda transportada maior, menor quantidade de defeito por circuito, menor tempo de configuração, tornariam a migração viável antes desta data.

O próximo capítulo comenta os resultados aqui alcançados e propõem trabalhos futuros mesmo que baseados em partes não executadas neste.

## 5 CONCLUSÃO

### 5.1 ANÁLISE DOS RESULTADOS E CONCLUSÃO

As operadoras de telecomunicações, de uma forma geral, as “*Incumbents*”, ou seja, as antigas operadoras estatais, não só no Brasil, como na Europa, América do Sul e Ásia, mantém certa apreensão na migração do transporte dos seus serviços de conexão para redes unicamente *IP*. Observa-se, igualmente, que a expectativa da migração está quase que totalmente voltada aos serviços de voz, o que era de se esperar, pois a maior parte da receita destas empresas ainda advém destes serviços.

Por outro lado, a migração do transporte de serviços de interconexão de redes de dados para redes unicamente *IP* deve ocorrer de forma diferenciada em relação aos serviços de voz. A chamada *NGN (Next Generation Network)* ainda é tratada de forma análoga as atuais RTPC, ou seja, a convergência existirá muito mais no transporte dos serviços do que na maneira de prestá-los, exatamente como hoje: a rede de voz dispõe de pessoal próprio de operação, com jargões e equipamentos diferenciados. A diferenciação é tanta que, como no caso da *British Telecom*, uma rede de transporte *IP (21th Network)* foi implantada com o único objetivo de transportar voz e serviços de multimídia.

Como falar de convergência a pessoas que operam uma rede de voz (RTPC) com, pelo menos, dez anos de operação a plenos serviços pela frente? Por que mudar se funciona tão bem? Por que mudar se ela atende plenamente aos serviços solicitados? A resposta está muito mais na tendência do mercado e nas exigências de redução de *Opex (Operating expenditures)* do que na necessidade dos clientes. A RTPC literalmente foi abandonada pelos principais fabricantes de equipamentos de comutação e a sua depreciação natural vai tornar sua operação, seja pela carência de sobressalentes e diminuição da disponibilidade de pessoal especializado, cada vez mais cara. Ainda vale dizer que qualquer configuração sistêmica deve ser feita de forma descentralizada e o “consumo de recursos” desta mão-de-obra em redução quantitativa e qualitativa já tem um limite delineado.

Outro fator que levará a migração da RTPC para a *NGN* será a necessidade de oferecer novos serviços como forma de manter ou conquistar clientes visto que estes novos serviços dificilmente seriam possíveis sobre a atual RTPC.

Um artigo publicado no *STEM (Strategic Telecoms Evaluation Model)* Newsletter, em outubro 2002, durante o *ITU Telecom Word Forum* em Genebra por Robin Bailey, afirmou que esta migração ocorreria a partir de 2007 de forma lenta, sendo que a partir de 2013, já não se implantariam centrais que não fossem *NGN* e o transporte de voz se daria totalmente sobre *IP*. Em 2021 já não existiriam centrais que não fossem *NGN*.

E por último, não menos importante, existe a regulamentação da portabilidade numérica, a qual vai obrigar as operadoras a uma interconexão de suas bases numéricas o que é extremamente caro de se implementar sobre os atuais sistemas de comutação de voz.

A migração da voz acontecerá de uma estrutura descentralizada para uma centralizada, baseada na inteligência dos *Softswitches* ao passo que as redes de dados têm sua capilaridade distinta conforme a camada, sendo as Redes Determinísticas as de maior capilaridade, seguidas das redes de camada 2: *Frame Relay*, *ATM*, Banda Larga e, finalmente, a rede *IP/MPLS*. Esta última já é centralizada. Assim, fica a pergunta: como migrar um serviço de camada 2 para ser transportado pela camada 3 se esta sequer existe em algumas localidades mais remotas?

Ainda complementado: o cliente de voz tem um perfil diferente do cliente de interconexão de redes de dados. Este último tem requisitos de *Capex (Capital expenditures)* e *Opex (Operating expenditures)* que por vezes o forçam a manter equipamentos antigos em operação, forçando as operadoras a manterem serviços mesmo com redes bastante obsoletas. Um exemplo disso: a Febrabam mantém sua rede de Caixas eletrônicos 24 horas ainda suportada por uma rede *X.25/X.75*.

Voltando ao item de capilaridade: onde deve primeiro ocorrer a migração dos serviços de interconexão de redes? Tudo leva a crer que esta se dará do centro para as extremidades das redes, pois a rede *IP*, atualmente, só existe no core da rede.

Naturalmente, os roteadores de distribuição terão a sua capilaridade aumentada na medida em que novos serviços suportados por redes *IP* forem sendo solicitados. O problema aparece quando mesmo em locais onde já existe a rede *IP* antigos serviços de camada 2 continuam sendo solicitados, no entanto, a solução aparece da mesma forma uma vez que as recomendações de *interworking* entre as redes de camada 2 e a rede *IP/MPLS* já podem ser implementadas.

Diante do exposto até aqui, resumidamente pode-se concluir o seguinte:

- Que a migração dos serviços de interconexão de redes de dados está diretamente ligado à migração dos serviços de transporte de voz, uma vez que a necessidade de se transportar a voz sobre redes *IP*, naturalmente, aumentaria a capilaridade desta rede;
- Que a necessidade do aumento da capilaridade da rede *IP*, forçaria investimentos massivos, aumento de demanda por fornecimento de equipamentos, aumento de concorrência e queda de preços;
- Que preços menores para contratos de fornecimento diminuiriam os custos dos serviços e conseqüentemente os preços praticados, aumentando a demanda;
- Que a migração, como já dito, se dará do *Core* da rede para as bordas;
- Que o *Core* das redes de camada 2 desaparecerá na medida em que estas serão apenas redes de Acesso à rede *IP/MPLS* e;
- Que alinhando a conclusão com o trabalho desenvolvido nos capítulos anteriores, esta migração se dará tão mais rápido quanto forem: menores os tempos de configuração dos serviços; menores os custos dos contratos de operação e implantação; maiores as bandas dos serviços adquiridos e menores as quantidades de defeitos. Estes parâmetros, ao contrário dos descritos nos parágrafos anteriores, além de serem objeto deste trabalho são bastante mais mensuráveis e justamente por isso, foram utilizados.

Na comparação dos tempos de configuração feita no capítulo 4, o resultado direto apresentado foi o gráfico repetido na Figura 5.1, onde o tempo de configuração médio diário dopado para redes *IP* e não *IP* e sua evolução no período de dezembro/2004 a julho/2007, aparecem como duas curvas e suas reduções lineares.

A primeira conclusão é a de que a reta da rede *IP* tem uma inclinação mais negativa que a da rede não *IP* e estas se cruzam por volta de **junho de 2013**. Percebe-se que este período está bem adequado com as previsões de Robin Bailey para migração dos serviços de voz, conforme *ITU Telecom Word Fórum* e reflete, claramente, a tendência atual de preparação para migração dos serviços dentro da Empresa onde os dados foram coletados.

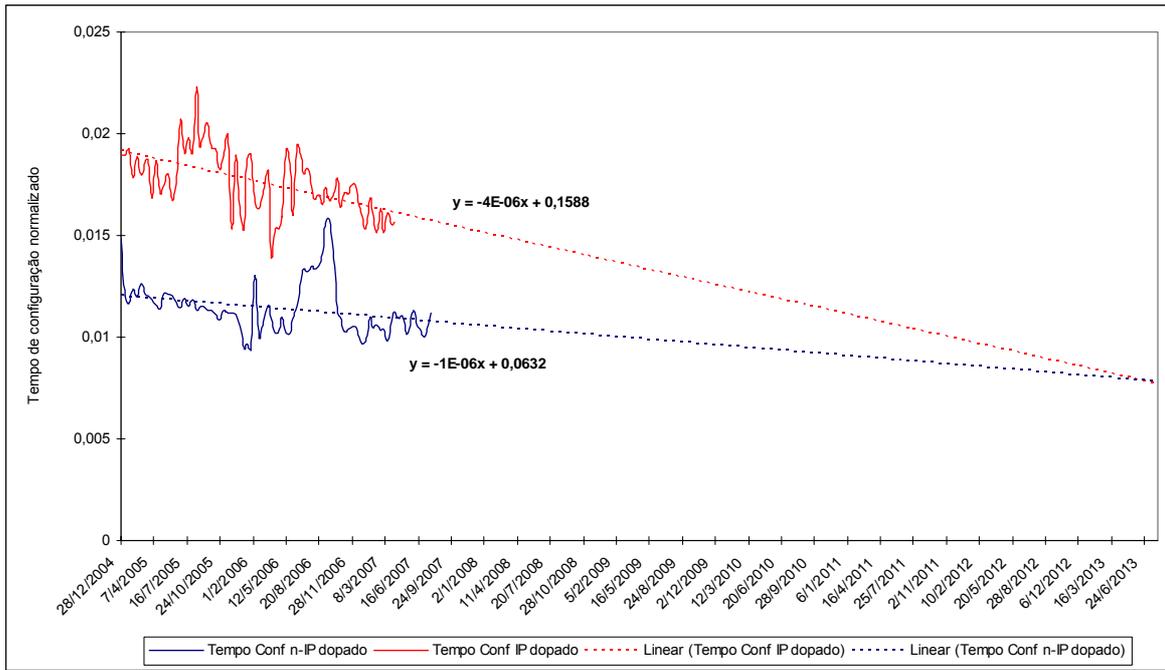


Figura 5.1 - Comparativo entre tempo de configuração dopado IP e não IP

Agora, se a análise for feita baseada no critério de custo, repetido na Figura 5.2, observa-se, igualmente, que a rede *IP* “dopada” tem uma inclinação mais acentuada que a da rede não *IP* “dopada” e o encontro das duas retas se dá por volta de **novembro de 2015**.

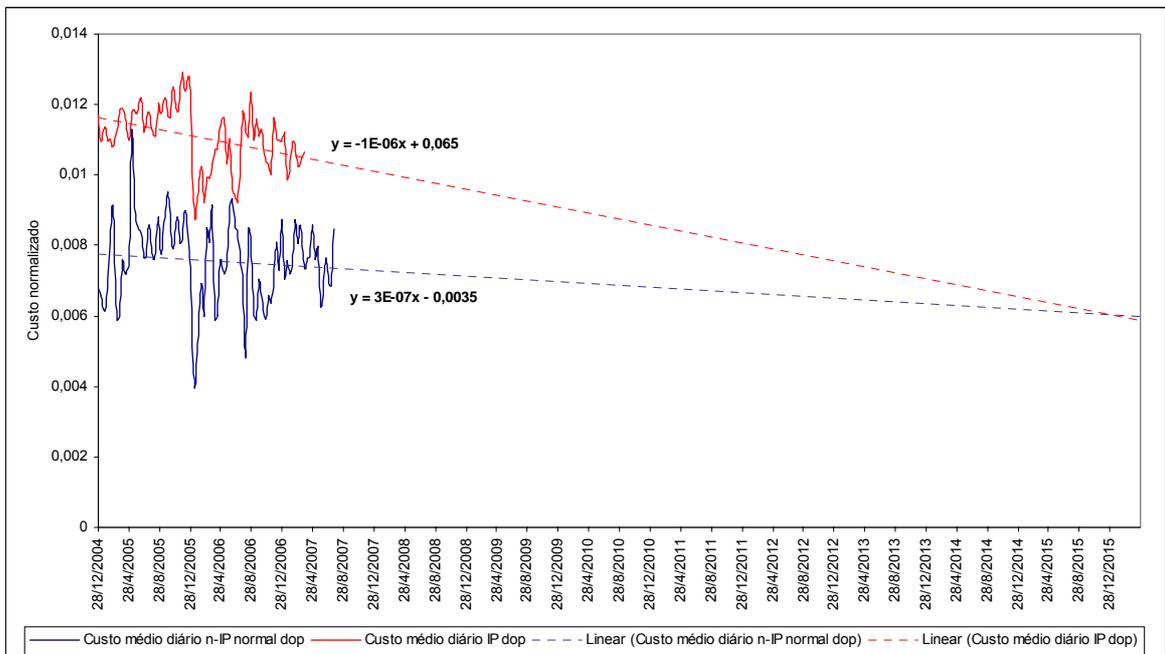


Figura 5.2 – Comparativo entre custo dopado IP e não IP

Concluindo a análise dos critérios “dopados”, ainda resta a comparação entre a quantidade de reclamações “dopada”.

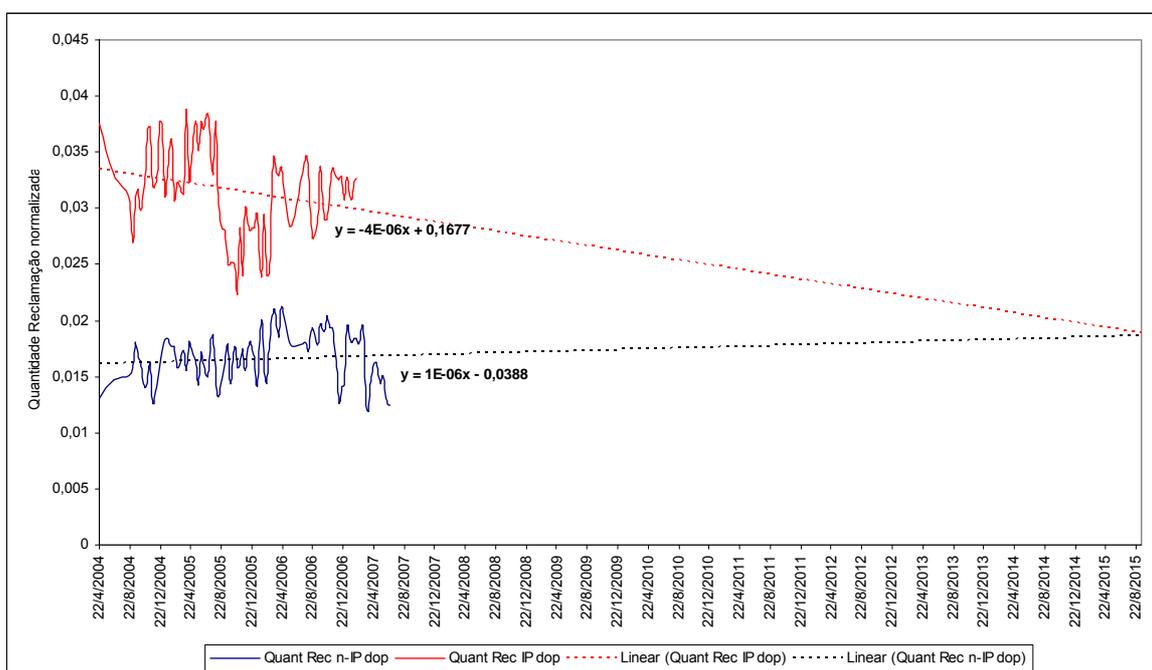


Figura 5.3 – Comparativo entre quantidade de reclamações dopadas IP e não IP

Na Figura 5.3 podem ser vistas as retas de tendência para as quantidades de reclamações *IP* e não *IP* “dopadas”. A quantidade de reclamações média diária por serviços *IP* e não *IP* mostra comportamentos diferentes no que se refere à tendência. As reclamações têm uma tendência de diminuição para as redes *IP* e aumento para redes não *IP*. Reforça-se que esta linha de tendência foi “dopada” pelos outros critérios.

A linha de tendência para a quantidade de reclamações de redes não *IP* não “dopada”, conforme a Figura 5.4, tem uma leve tendência de diminuição, o que não ocorre na reta de tendência “dopada”. Isto só se explica pela influência que um critério causa ao outro no domínio da *KLT*.

Finalmente, observando o gráfico da Figura 5.3, vê-se que o cruzamento das duas retas ocorre em **agosto de 2015**.

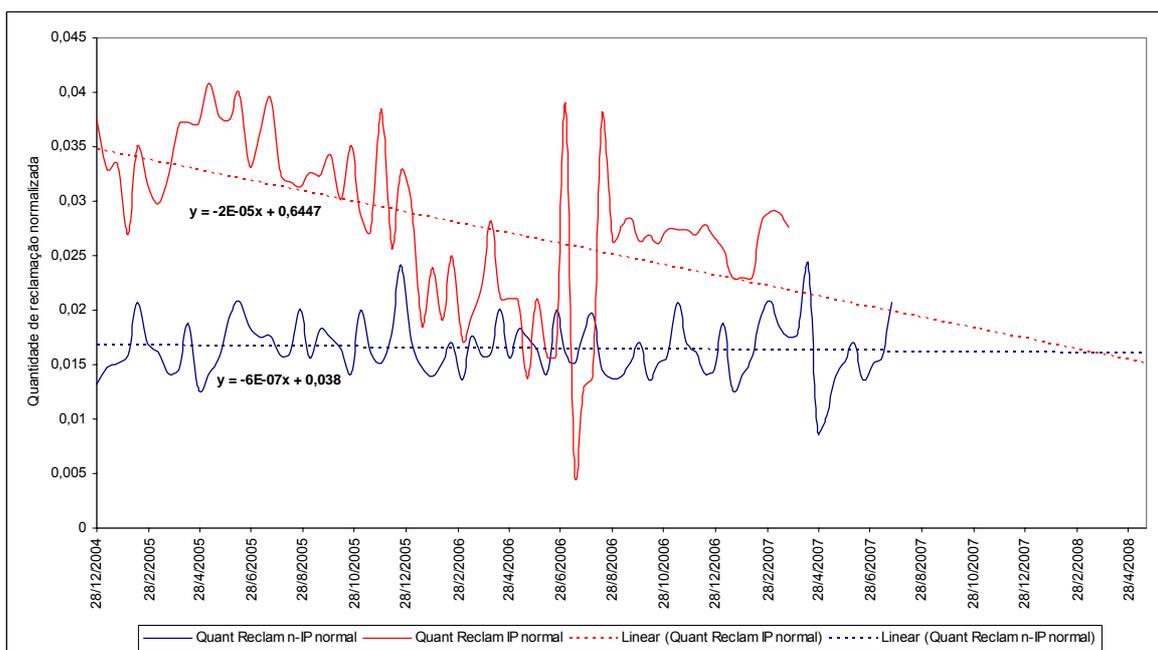


Figura 5.4 – Comparação de quantidade de reclamações IP e não IP não “dopada”

A Tabela 5.1 serve como resumo dos resultados obtidos através da aplicação do Algoritmo Hebbiano Generalizado para Extração dos Componentes Principais – GHAPCA, nos dados de entrada da RNA vistos no capítulo 3.

Tabela 5.1 - Comparação dos tempos de predição a partir de todos os coeficientes [AUTORA]

Conjuntos/Critérios	Tempo de Configuração	Custo	Quantidade de Reclamação
IP dopada	$y = -4.10^{-6}x + 0,1588$	$y = -10^{-6}x + 0,065$	$y = -4.10^{-6}x + 0,1677$
não IP dopada	$y = -10^{-6}x + 0,0632$	$y = -3.10^{-6}x + 0,065$	$y = 1.10^{-6}x - 0,0388$
<b>Predição dopada</b>	<b>jun/2013</b>	<b>nov/2015</b>	<b>ago/2015</b>
IP não dopada	$y = 10^{-5}x - 0,4603$	$y = -4.10^{-6}x + 0,1606$	$y = -2.10^{-5}x - 0,6447$
não IP não dopada	$y = 10^{-5}x - 0,4375$	$y = 6.10^{-7}x - 0,0155$	$y = -6.10^{-5}x - 0,038$
<b>Predição não dopada</b>	<b>retas paralelas</b>	<b>mai/2007</b>	<b>mar/2008</b>

Por inspeção, na Tabela 5.1 observa-se que: quando a predição é feita considerando-se os critérios competitivos “dopados”, isto é, influenciados uns pelos outros, esta fica entre 2013 e 2015, enquanto observando particularmente cada critério comparando os dois conjuntos, rede *IP* e não *IP* não se chega a um período de forma conclusiva.

A validade da utilização da RNA fica clara quando se compara o conjunto das previsões. Uma análise feita separadamente de cada critério poderia levar a uma conclusão errada sobre a melhor data de migração viável.

A *KLT* foi, inicialmente, concebida para se descobrir distúrbios em um conjunto de dados, alterando os referenciais dos componentes vetoriais de modo que a maior parte destes dados tivesse a menor variância possível e, assim, os dados que destoassem do conjunto (distúrbio) seriam mais facilmente encontrados.

Esta transformada, neste trabalho, teve uma utilização em um contexto bastante diferente, onde se buscou uma tendência e não o distúrbio. Aqueles dados (critérios) que ficaram mais distantes da média (maior variância), tiveram menos influência sobre a tendência.

Aliado a este artifício, vale lembrar que não se tem somente a tendência de um único critério, mas de todos os critérios influenciando-se mutuamente.

Ficaria muito difícil se obter uma previsão analisando os critérios isoladamente. Com que grau de confiabilidade se pode dizer que já é mais “barato” provisionar serviços não *IP* sobre redes *IP* conforme a Figura 5.5 está mostrando?

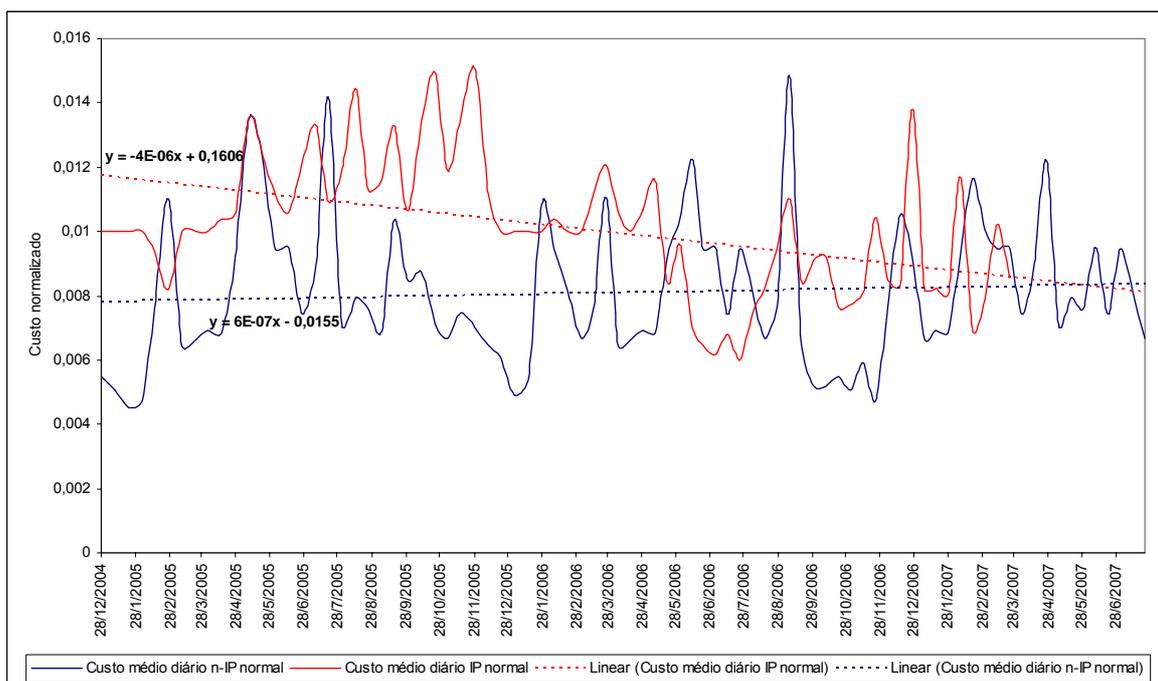


Figura 5.5 – Comparativo custo médio diário não IP e IP não “dopados”

Este gráfico mostra somente que a média de custo dos serviços *IP* e não *IP* configurados em um único dia teve sua relação invertida, ficando a média de custo dos serviços *IP* menor que a média de custo dos serviços não *IP*. Caso a migração fosse feita nesta data (maio de 2007), todos os serviços não *IP* teriam seu tempo de configuração aumentado

para o tempo de configuração *IP* e, conseqüentemente, sua quantidade de reclamações também. Para compensar estes aumentos, mais pessoal especializado seria necessário levando ao aumento do custo.

Existe um ponto ótimo para a migração e este ponto deve levar em conta todos os critérios competitivos. Refazendo a análise com o custo “dopado”, caso a migração fosse feita em novembro de 2015, quando a quantidade de reclamações para os serviços *IP* já estivesse menor que para os serviços não *IP* e o tempo de provisionamento também, a migração se daria muito mais facilmente.

O que pode ser feito para que a migração seja viabilizada antes de 2013? A resposta chega pela alteração na evolução de alguns cenários, os quais influenciam diretamente os critérios competitivos de operações. Alguns deles serão abordados agora:

#### Aumento da capilaridade da Rede *IP*:

Quanto mais capilarizada for a rede *IP*, menos pontos de configuração, portanto menos heterogênea e mais centralizada fica a rede. Isto possibilita a padronização e automação de processos deixando a rede menos sujeita a falhas. Ainda, quanto menos equipamentos de tecnologias diferentes para operar e configurar, menos pessoas especializadas e menores os custos dos serviços prestados, uma vez que são necessários menos contratos de suporte e manutenção, além de redução da necessidade de sobressalentes.

Adicionalmente, uma rede única ou de uma mesma tecnologia proporciona uma topologia mais simples e, portanto, com maior possibilidade de aumento da eficiência na transmissão da banda requisitada.

#### Aprovisionamento automático:

Foi visto no modelo de predição que a velocidade de configuração dos serviços é um dos fatores principais de influência no tempo onde as duas curvas se cruzam e onde o provisionamento sobre uma rede unicamente *IP* passa ser mais vantajosa que provisionar os mesmos serviços sobre redes *IP* e não *IP*.

Entende-se ainda que, quanto mais automático for este processo, menos susceptível a erros e, portanto, menor quantidade de reclamações inerentes às falhas na configuração do serviço. Quantidade de reclamações é outro ponto que interfere na curva de predição.

Assim, concluindo finalmente o trabalho, a previsão inicial para que a migração se torne viável, segundo os critérios competitivos, se dará entre 5 e 7 anos, ocorrendo de forma gradual, primeiro para as plataformas de voz, e, com aumento massivo dos acessos de banda larga e do aumento das bandas contratadas começará pelo *Core*, com o *Core* da rede *IP/MPLS* absorvendo o *Core* das redes de camada 2 – *Frame Relay* e *ATM*.

Isto vale caso permaneça o atual ritmo de evolução dos parâmetros analisados.

## **5.2 SUGESTÃO DE TRABALHOS FUTUROS**

Como sugestão de trabalhos futuros, poder-se-ia dividir em três linhas de desenvolvimento:

- Melhoria da Predição
- Subprodutos da base de dados trabalhada
- Aplicações na Empresa

### **5.2.1 Melhoria da Predição**

Sugere-se em primeiro lugar a implementação de uma Rede Neural preditiva a partir das curvas apresentadas, como forma de apurar as datas de predição e de acompanhar suas evoluções de forma dinâmica.

Segundo, o desenvolvimento de uma ferramenta de cálculo de custo de instalação e manutenção dos serviços de forma dinâmica a partir da montagem da topologia dos serviços no momento de sua configuração. Aproveitando o projeto de OS (Ordem de Serviço) de Viabilidade, poderia ser introduzido o cálculo do custo por serviço para servir de balizador na precificação elaborada pela área Comercial.

Terceiro, somado a esta ferramenta poderia ser desenvolvido módulo para cálculo, através do método apresentado neste trabalho, de localidades (ou grupo de localidades) onde a migração já fosse possível de se realizar.

### **5.2.2 Subprodutos da base de dados trabalhada**

A ferramenta empregada para a predição, apresentou outros resultados aqui não explorados.

A Figura 5.6 ilustra, apenas como exemplo, um dos subprodutos que se pode obter com o modelo empregado. Observa-se que existe uma variação entre os tempos de configuração “dopados” e não “dopados” por filial.

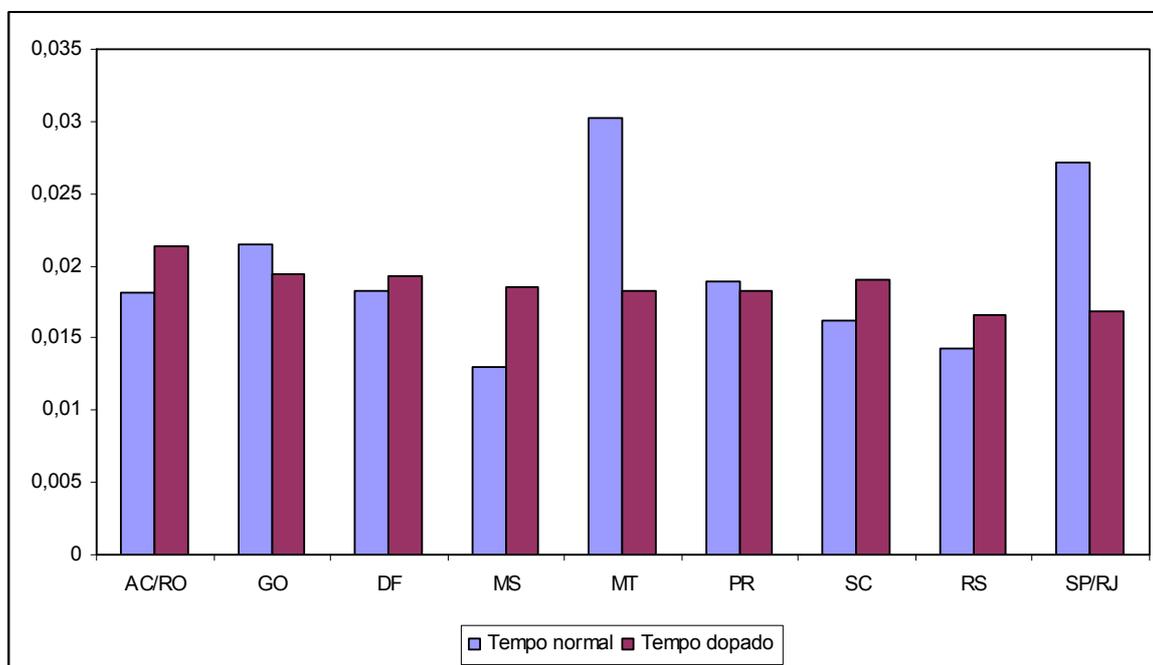


Figura 5.6 – Comparativo dos tempos de configuração IP normalizados entre Filiais

Adotando-se como exemplo a Filial MT, onde o tempo de configuração normalizado e não “dopado”(tempo normal) está bem acima dos tempos de todas as filiais e bem maior que seu tempo de configuração “dopado”. Ele representa, em termos absolutos, qual a média diária do tempo de configuração de um serviço IP para a Filial MT.

Entretanto, a Filial MT não configura a maior quantidade de serviços e os serviços configurados não são os de maior velocidade, ou seja, não são os mais representativos no modelo. Uma vez o tempo “dopado”, ou seja, aplicada a *KLT*, o tempo de configuração diminui drasticamente influenciado pelos outros coeficientes e pelos tempos das demais filiais.

Assim, observando novamente o gráfico as primeiras filiais onde se deveria migrar os serviços para um *Core IP/MPLS* seriam RS e SP/RJ, onde os tempos “dopados” são menores e representativos de todos os demais coeficientes competitivos.

### 5.2.3 Aplicações na Empresa

As áreas de Engenharia e Operação das Operadoras de Telecomunicações em geral, seguem definições estabelecidas pelas áreas de Planejamento Estratégico e Marketing para a aquisição de novas tecnologias ou expansão de tecnologias existente no atendimento de demanda pelos serviços prestados. Estas áreas, porém, baseam-se em estatísticas e tendências mercadológicas para traçar ações de curto, médio e longo prazo. O que nem sempre garante a assertividade destas ações.

Há uma dificuldade muito grande no desenvolvimento de ferramentas ou metodologias que façam de forma sistêmica e evolutiva a comparação entre dados de naturezas diferentes. É intuitivo, mas não é matematicamente comprovada a influência que a quantidade de reclamações de clientes pode exercer no custo do serviço comercializado, por exemplo.

O trabalho propõe uma aplicação direta, como ferramenta de apoio às áreas de Planejamento Estratégico e Planejamento de Rede da Empresa, através da apresentação das curvas de tendência a partir de dados de entrada evolutivos que fazem parte do processo operacional de entrega dos produtos e serviços.

É fato que o mercado caminha para uma rede única, que é a Rede *IP*, porém, ações operacionais de curto e médio prazo podem ser definidas para acelerar o processo de migração, buscando, mais rapidamente, redução dos custos operacionais com as Redes ditas Legadas.

Há necessidade de planejar detalhadamente o lançamento de novos serviços, buscando investir em provisionamento automático, já que, conforme comprovado neste trabalho, o tempo de configuração é o fator que mais contribui para postergar a migração. Além disto, a capilarização da Rede *IP* também é um fator que irá contribuir para a aceleração da migração dos serviços que utilizam rede não *IP* para rede *IP*.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [ALBERTI1998] ALBERTI, A.M.. *Tecnologia ATM*. Monografia. UNICAMP, 1998.
- [BARR2002] BARRETO, C.A.. *Introdução às Redes Neurais Artificiais*. Consultado em <http://www.inf.ufsc.br/~barreto/tutoriais/Survey.pdf>. UFSC. 2002.
- [BENITE2003] BENITE, M.. *Aplicação de Modelos de Redes Neurais na Elaboração e Análise de Cenários Macroeconômicos*. Dissertação de Mestrado. USP, 2003.
- [BRITO2006] BRITO, J.M.C.. *Redes de Comunicação*. Programa de Pós-Graduação Mestrado Profissionalizante – Stricto Sensu Brasil Telecom. Brasília. 2006.
- [CAST2001] CASTRO, F.C.C.. *Redes Neurais Artificiais*. Consultado em [http://www.ee.pucrs.br/~decastro/RNA\\_hp/RNA.html](http://www.ee.pucrs.br/~decastro/RNA_hp/RNA.html). PUC-RS. 2001.
- [CISCO2004a] CISCO SYSTEMS. *Cisco Voice over IP* volume 2, versão 4.2. Cisco Systems Co. 2004.
- [DECAST1996] DE CASTRO, M.C.F.. *Software & C. Source Code*. Consultado em <http://www.ee.pucrs.br/~decastro/download.html>. 1996.
- [GIANESI1996] GIANESI, I.G.N & CORRÊA, H.L.. *Administração Estratégica de Serviços*. Atlas. 1996.
- [HAYKIN1996] HAYKIN, S., *Adaptive Filter Theory*. Prentice Hall. New Jersey. 1996.

- [ITUI.361] ITU I.361.. *B-ISDN ATM Layer Specification*. International Telecommunication Union. Recommendation ITU I.361. 1995.
- [ITUM3200] ITU-T M.3200.. *Principles for a Telecommunication Manager Network*. International Telecommunication Union. Recommendation M.3200. 1996.
- [KAPP2003] KAPP, P. E. H.. *Metodologia de implantação de Rede de Gerência baseada em SNMP para Redes de Acesso*. Dissertação de Mestrado. PUC-RS, 2003.
- [LABINFO2007] LABORATÓRIO DE BIOINFORMÁTICA, Ministério de Ciência e Tecnologia.  
[http://www.labinfo.lncc.br/index.php?option=com\\_content&task=view&id=29&Itemid=114](http://www.labinfo.lncc.br/index.php?option=com_content&task=view&id=29&Itemid=114), acessado em 07.09.2007.
- [OJA1992] OJA, E.. *Principal Components, Minor Components and Linear Neural Networks*. *Neural Networks*. Helsinki. 1992.
- [PAIVA2004] PAIVA, E.L.. *Estratégia de Produção e de Operações*. Bookman. Porto Alegre. 2004.
- [RFC1122] RFC1122.. *Requirements for Internet Hosts*. Internet Engineering Task Force. R. Braden. 1989.
- [RFC1323] RFC1323.. *TCP Extensions for High Performance*. V. Jacobson. Cray Research. 1992.
- [RFC1889] RFC1889.. *A Transport Protocol for Real-Time Applications*. Audio-Video Transport Working Group. Lawrence Berkeley National Laboratory. 1996.
- [RIPOLLI1993] RIPOLLI, G.P.. *Desenvolvimento de um Modelo de Simulação para Dimensionamento de Equipe Polivalente de Manutenção de Produção: uma Abordagem Estratégica*. Dissertação de Mestrado. UFRGS, 1998.

- [SANT2001] DOS SANTOS, C.C.. *Modelagem de Bacias Urbanas com Redes Neurais Artificiais*. Dissertação de Mestrado da Escola Politécnica da Universidade de São Paulo. São Paulo. 2001.
- [SCHÄFF1999] SCHÄFFER, L.A.. *Redes de Alta Velocidade: a Rede de Acesso*. Dissertação de Mestrado. UFRGS, 1999.
- [SLACK1999] SLACK, N.. *Administração da Produção*. Atlas. São Paulo. 1999.
- [SPIEGEL1992] SPIEGEL, M.R... *Manual de Fórmulas, Métodos e Tabelas de Matemática*. Makron Books do Brasil. São Paulo. 1992.
- [STALLINGS1997 ] STALLINGS, W.. *ISDN and Broadband ISDN with Frame Relay and ATM*. Prentice Hall. New York. 1997.
- [TANEMB2003] TANENBAUM, A.S.. *Redes de Computadores*. Campus. 2003.
- [VIEIRA2001] VIEIRA, S.N.. *Estratégia de Operações em Empresas Prestadoras de Serviços de Engenharia Elétrica na Região Metropolitana de Curitiba*. Dissertação de Mestrado. UFRGS, 2001.

## **APÊNDICES**

## **A ALGORITMO HEBBIANO GENERALIZADO PARA EXTRAÇÃO DOS COMPONENTES PRINCIPAIS – GHAPCA**

Este anexo descreve o Algoritmo Hebbiano Generalizado para Extração dos Componentes Principais – GHAPCA, conforme [DECAST1996].

Para o melhor entendimento do GHAPCA é preciso primeiramente entender o que é e como é feita a Análise dos Componentes Principais e o Algoritmo Hebbiano Generalizado.

A análise dos Componentes Principais ou Transformação de *Karhunen-Lòeve* é a análise das principais características de um conjunto de dados. A extração destas características é possibilitada pela determinação das direções no espaço de dados de entrada na quais as amostras exibem a maior variância. Os dados de maior variância obtidos são então projetados em um subespaço de menor dimensão produzindo a redução dimensional desejada.

O truncamento de um vetor de dados causa um erro médio quadrático igual à soma das variâncias dos elementos descartados. Se o vetor de dados for truncado sem critério os elementos eliminados poderão carregar características efetivas dos dados, principalmente se forem elementos de alta variância.

A solução para este problema reside em especificar uma transformação com a característica de que alguns componentes transformados possuam pequena variância para que, em um truncamento criterioso, representem pouca perda de informação.

A Transformação de *Karhunen-Lòeve* é a solução indicada, pois maximiza a razão de decréscimo da variância, concentrando a maior parte da variância dos dados em alguns poucos elementos. Assim, a escolha dos elementos a serem descartados será ótima no sentido do Erro Médio Quadrático (*MSE*).

A seguir será apresentada, matematicamente, a *KLT* para definição dos componentes principais.

Seja  $x$  um vetor aleatório, de dimensão  $p$ , representativo do conjunto de dados de interesse e de média zero. Caso o vetor  $x$  não possua média zero, a média deverá ser subtraída do vetor antes de iniciar a análise.

Seja  $u$  um vetor unitário, igualmente de dimensão  $p$ , sobre o qual será projetado o vetor representativo do conjunto de dados, vetor  $x$ . A projeção é dada pelo produto interno apresentado na Equação A.1.

$$a = x^T u = u^T x \quad \text{Equação A.1}$$

Sendo  $u$  um vetor unitário, a norma de  $u$  é dada pela Equação A.2.

$$\|u\| = (u^T u) = 1 \quad \text{Equação A.2}$$

A projeção  $a$  também é aleatória, com média e variância associadas à estatística do vetor de dados  $x$ .

A média da projeção  $a$  é dada por:

$$E\{a\} = E\{u^T x\} \quad \text{Equação A.3}$$

$$E\{a\} = u^T E\{x\} = 0 \quad \text{Equação A.4}$$

A variância da projeção  $a$  é dada por:

$$\sigma_a^2 = E\{a^2\} \quad \text{Equação A.5}$$

$$\sigma_a^2 = E\{(u^T x)(x^T u)\} \quad \text{Equação A.6}$$

$$\sigma_a^2 = u^T E\{xx^T\}u \quad \text{Equação A.7}$$

Onde o valor esperado do produto externo do vetor  $x$  por ele mesmo ( $E\{xx^T\}$ ) é a matriz covariância  $C_x$ , de dimensões  $p \times p$ , dos vetores de dados  $x$ . Portanto, a Equação A.7 pode ser escrita como:

$$\sigma_a^2 = u^T C_x u \quad \text{Equação A.8}$$

Onde, observa-se, que a variância  $\sigma^2$  da projeção  $a$  é uma função do vetor unitário  $u$ .

$$\sigma_a^2 = f(u) \quad \text{Equação A.9}$$

O intuito agora é determinar os vetores  $u$  ao longo dos quais a variância da projeção  $a$  é máxima, sujeitos à restrição dada pela Equação A.2.

Se  $u$  é um vetor unitário tal que  $f(u)$  tenha um valor máximo, então, para qualquer pequena variação  $\delta u$  do vetor unitário  $u$ ,

$$f(u + \delta u) \approx f(u) \quad \text{Equação A.10}$$

Aplicando-se a definição dada pela Equação A.8 na Equação A.10 e admitindo-se a igualdade devida ao valor de  $\delta u$  ser muito pequeno,

$$f(u + \delta u) = (u + \delta u)^T C_x (u + \delta u) \quad \text{Equação A.11}$$

$$f(u + \delta u) = u^T C_x u + u^T C_x \delta u + \delta u^T C_x u + \delta u^T C_x \delta u \quad \text{Equação A.12}$$

Como a matriz variância do vetor de dados,  $C_x$ , é uma matriz simétrica, tem-se que:

$$u^T C_x \delta u = \delta u^T C_x u \quad \text{Equação A.13}$$

E a Equação A.12 pode ser reescrita como:

$$f(u + \delta u) = u^T C_x u + 2\delta u^T C_x u + \delta u^T C_x \delta u \quad \text{Equação A.14}$$

Como  $\delta u$  é muito pequeno, o termo  $\delta u^T C_x \delta u$  do lado direito da Equação A.14 pode ser desprezado. Além disto, considerando-se as definições das Equação A.8 e da Equação A.9, pode-se escrever que:

$$f(u + \delta u) = f(u) + 2\delta u^T C_x u \quad \text{Equação A.15}$$

Ao substituir a Equação A.10 na Equação A.15, conclui-se que:

$$\delta u^T C_x u = 0 \quad \text{Equação A.16}$$

A outra condição a ser respeitada deriva de que as perturbações  $\delta u$  só são admissíveis para as quais a norma do vetor perturbado  $(u + \delta u)$  permaneça unitária, ou seja:

$$\|u + \delta u\| = 1 \quad \text{Equação A.17}$$

O que equivale a dizer que:

$$(u + \delta u)^T (u + \delta u) = 1 \quad \text{Equação A.18}$$

Expandindo o produto mostrado na Equação A.18, tem-se:

$$(u + \delta u)^T (u + \delta u) = u^T u + u^T \delta u + \delta u^T u + \delta u^T \delta u \quad \text{Equação A.19}$$

Onde:

$$u^T \delta u = \delta u^T u \quad \text{Equação A.20}$$

E onde o termo  $\delta u^T \delta u$  pode ser desconsiderado. Como  $(u^T u + 2\delta u^T u)$  deve ser unitário, conforme Equação A.18, pode-se afirmar que:

$$\delta u^T u = 0 \quad \text{Equação A.21}$$

A análise da Equação A.21 permite concluir que as perturbações  $\delta u$  devem ser ortogonais a  $u$ , sendo esta a única mudança permitida na direção de  $u$ , de maneira que a Equação A.17 seja obedecida.

Na Equação A.16 e na Equação A.21 estão as duas condições a serem respeitadas para a determinação dos vetores  $u$  para os quais  $f(u)$  terá valor extremo. Combinando convenientemente estas equações, pode-se escrever que:

$$C_x (\delta u)^T u - \lambda (\delta u)^T u = 0 \quad \text{Equação A.22}$$

Onde o fator de escala  $\lambda$  foi introduzido para compatibilizar as dimensões dos elementos do vetor unitário  $u$  com a matriz covariância  $C_x$ .

Resolvendo a Equação A.22,

$$(\delta u)^T (C_x u - \lambda u) = 0 \quad \text{Equação A.23}$$

Para que a condição da Equação A.23 seja satisfeita, é necessário e suficiente que:

$$C_x u = \lambda u \quad \text{Equação A.24}$$

Que é a equação que define os vetores  $u$  para os quais  $f(u)$  tem valor máximo.

A Equação A.24 é reconhecida como a equação dos autovalores e tem soluções não-triviais para aqueles valores de  $\lambda$  que são os autovalores da matriz  $C_x$ . Os vetores  $u$  associados a estes  $\lambda$  são auto-vetores de  $C_x$ .

A matriz covariância é uma matriz Hermitiana e, portanto, possui autovalores reais e não-negativos. Assumindo que os autovalores da matriz covariância sejam distintos, os auto-vetores associados a estes autovalores serão ortogonais.

Sejam  $\lambda_0, \lambda_1, \dots, \lambda_{p-1}$  os autovalores da matriz  $C_x$ , de dimensões  $p \times p$ , e  $u_0, u_1, \dots, u_{p-1}$  os auto-vetores associados aos autovalores. Assim, da Equação A.24 pode-se escrever que para  $j = 0, 1, \dots, p-1$

$$C_x u_j = \lambda_j u_j \quad \text{Equação A.25}$$

Seja  $\Lambda$  uma matriz diagonal,  $p \times p$ , composta pelos autovalores da matriz  $C_x$  como segue:

$$\Lambda = \text{diag}[\lambda_0, \lambda_1, \dots, \lambda_j, \dots, \lambda_{p-1}] \quad \text{Equação A.26}$$

Onde os autovalores estão ordenados de forma decrescente, de modo que  $\lambda_0$  seja o maior autovalor.

Seja  $U$  uma matriz  $p \times p$  formada pelos auto-vetores correspondentes como segue:

$$U = [u_0, u_1, \dots, u_j, \dots, u_{p-1}] \quad \text{Equação A.27}$$

De acordo com a Equação A.26 e a Equação A.27, a Equação A.25 pode ser reescrita:

$$C_x U = U \Lambda \quad \text{Equação A.28}$$

Como os auto-vetores de  $C_x$  (vetores coluna de  $U$ ) são ortogonais, a matriz  $U$  é dita ortogonal. Os vetores de  $U$  satisfazem, ainda, a condição de ortonormalidade, pois:

$$u_j^T u_j = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases} \quad \text{Equação A.29}$$

A Equação A.29 também pode ser escrita como:

$$U^T U = I \quad \text{Equação A.30}$$

De onde vem que:

$$U^T = U^{-1} \quad \text{Equação A.31}$$

A Equação A.28 pode ser reescrita, considerando a Equação A.31, sob a forma da transformação de similaridade ortogonal. Pré-multiplicando ambos os lados da Equação A.28 por  $U^T$  fica:

$$U^T C_x U = U^T U \Lambda \quad \text{Equação A.32}$$

Considerando a Equação A.31,

$$U^T C_x U = U^{-1} U \Lambda \quad \text{Equação A.33}$$

De onde resulta a transformação ortogonal de similaridade, mostrada abaixo:

$$U^T C_x U = \Lambda \quad \text{Equação A.34}$$

Que, na forma expandida, pode ser reescrita:

$$u_j^T C_x u_k = \begin{cases} \lambda_j, & k = j \\ 0, & k \neq j \end{cases} \quad \text{Equação A.35}$$

Na Equação A.8 e na Equação A.9 tem-se:

$$\sigma_a^2 = f(u) = u^T C_x u \quad \text{Equação A.36}$$

Comparando a Equação A.35 e a Equação A.36, tem-se que, para  $j = 0, 1, \dots, p-1$ :

$$f(u_j) = \lambda_j \quad \text{Equação A.37}$$

Duas conclusões importantes seguem da interpretação da Equação A.37:

- Que os auto-vetores da matriz covariância  $C_x$  dos vetores de dados de média zero representam as direções principais ao longo das quais a variância é máxima. Os autovalores associados a estes auto-vetores definem os valores máximos de variância. Em outras palavras, os autovalores e auto-vetores representam os modos de energia ou a forma como os dados variam. Cada modo de energia representa uma direção definida pelo respectivo auto-vetor no espaço de dados.

- O conjunto de dados varia de amostra para amostra. Certas variações são semelhantes entre si, de forma que podem ser incluídas em uma mesma classe. As variações pertencentes a mesma classe apresentam energia total associada à classe. Logo, cada classe de variação é um modo de energia. O próprio nível médio, previamente extraído do conjunto de dados, pode ser considerado como um modo de energia cuja forma de variação é constante.

Como há  $p$  possíveis soluções para o vetor unitário  $u$ , também haverá  $p$  projeções  $a$  do vetor representativo do conjunto de dados  $x$  sobre o vetor unitário  $u$ .

Assim como na Equação A.1, para  $j = 0, 1, \dots, p-1$ :

$$a_j = x^T u_j = u_j^T x \quad \text{Equação A.38}$$

Onde os  $a_j$  são chamados componentes principais do vetor de dados  $x$ , que são os modos principais de variação de  $x$ .

Para obter a reconstrução do vetor de dados  $x$  a partir das projeções  $a_j$  parte-se do seguinte:

$$a = U^T x \quad \text{Equação A.39}$$

Onde

$$a = [a_0, a_1, \dots, a_{p-1}]^T \quad \text{Equação A.40}$$

E

$$U^T x = [x^T u_0, x^T u_1, \dots, x^T u_{p-1}]^T \quad \text{Equação A.41}$$

Assim, pré-multiplicando ambos os lados da Equação A.39 por  $U$ , e considerando a Equação A.31, o vetor de dados  $x$  será reescrito por:

$$x = Ua = \sum_{j=0}^{p-1} a_j u_j \quad \text{Equação A.42}$$

Na Equação A.42, os vetores unitários  $u_j$  representam uma base do espaço de dados. Ao considerar os  $p$  termos, a transformação reconstruirá completamente o vetor de dados.

A redução dimensional é obtida através da equação de reconstrução dos dados, descartando as combinações lineares que apresentarem menor variância (energia) e retendo os termos de maior variância.

Sejam  $\lambda_0, \lambda_1, \dots, \lambda_m$  os  $m$  maiores autovalores da matriz covariância do vetor de dados. O vetor de dados poderá, então, ser aproximado para  $m < p$ :

$$\hat{x} = \sum_{j=0}^{m-1} a_j u_j \quad \text{Equação A.43}$$

O que equivale a truncar a reconstrução expressa pela Equação A.42.

O erro resultante da redução das combinações lineares relativas aos autovalores menores é obtido subtraindo o vetor truncado do vetor original.

$$e = x - \hat{x} \quad \text{Equação A.44}$$

Ou, substituindo a Equação A.42 e a Equação A.43 na Equação A.44:

$$e = \sum_{j=m}^{p-1} a_j u_j \quad \text{Equação A.45}$$

A variância total do vetor de dados  $x$ , de acordo com a Equação A.8 e a Equação A.45, pode ser expressa por:

$$\sum_{j=0}^{p-1} \sigma_j^2 = \sum_{j=0}^{p-1} \lambda_j \quad \text{Equação A.46}$$

A variância total do vetor estimado  $\hat{x}$ , referente aos  $m$  autovalores retidos é:

$$\sum_{j=0}^{m-1} \sigma_j^2 = \sum_{j=0}^{m-1} \lambda_j \quad \text{Equação A.47}$$

Quanto mais próximos de zero forem os autovalores descartados, mais eficiente será a redução dimensional, no sentido de preservar a informação contida no conjunto de dados.

Foi visto que para proceder à Análise dos Componentes Principais é necessário computar a matriz covariância do conjunto de dados de entrada para, então, aplicar um procedimento numérico que extraia os autovalores e os correspondentes auto-vetores desta matriz. Os auto-vetores correspondentes aos autovalores mais significativos são usados para extrair os componentes principais dos dados.

No entanto, quando se trata de grandes conjuntos de dados este procedimento onera as dimensões da matriz covariância e neste caso a solução mais eficiente para encontrar os componentes principais é através das Redes Neurais Artificiais (RNAs).

As RNAs têm a capacidade de resolver problemas desta natureza conectando múltiplas unidades de processamento através de um algoritmo de aprendizado, aprendendo padrões diretamente das entradas e generalizando as soluções obtidas.

Esta técnica de usar RNAs para Análise dos Componentes Principais sobre um vetor de dados de interesse é o Algoritmo Hebbiano Generalizado e é um dos mais usados na prática, porque extrai os componentes principais individualmente, em ordem decrescente de variância, além de, após treinada a rede, poder aplicar os resultados em qualquer outro conjunto de dados.

O Algoritmo Hebbiano Generalizado é usado para treinar uma Rede Neural Artificial progressiva, composta de uma camada de neurônios lineares. A RNA treinada por esta regra executará análise dos componentes principais de tamanho arbitrário sobre o vetor de entrada.

A Figura A.1 apresenta o modelo da RNA a ser treinada pelo Algoritmo Hebbiano Generalizado. Esta RNA tem  $p$  nós na camada de entrada e  $m$  neurônios na camada de saída, com  $m < p$ .

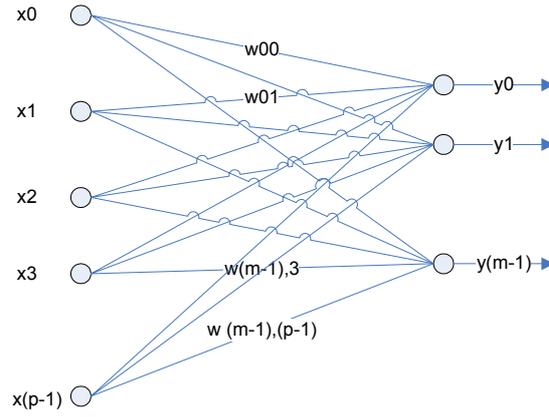


Figura A.1 – RNA progressiva com uma única camada de neurônios lineares, a ser treinada pelo Aprendizado Hebbiano Generalizado

Os pesos sinápticos  $w_{ij}$  conectam os nós de entrada  $i$  aos neurônios da camada de saída  $j$ , com  $i = 0, 1, \dots, p-1$  e  $j = 0, 1, \dots, m-1$  (considerando novamente a notação expandida).

A saída  $y_j(n)$  produzida pelo neurônio  $j$  no tempo  $n$  em resposta ao conjunto de entradas  $x_i(n)$  é dada por:

$$y_j(n) = \sum_{i=0}^{p-1} w_{ji}(n)x_i(n), \quad \text{Equação A.48}$$

Com  $j = 0, 1, \dots, m-1$ .

A adaptação do peso sináptico, devida ao Algoritmo Hebbiano Generalizado é expressa pela Equação A.49.

$$\Delta w_{ji}(n) = \eta \left\{ y_j(n)x_i(n) - y_i(n) \sum_{k=0}^j w_{ki}(n)y_k(n) \right\} \quad \begin{cases} i = 0, 1, \dots, p-1 \\ j = 0, 1, \dots, m-1 \end{cases} \quad \text{Equação A.49}$$

Onde  $\Delta w_{ji}(n)$  é a mudança aplicada ao peso sináptico  $w_{ji}(n)$  no tempo  $n$  e  $\eta$  é o parâmetro razão de aprendizado.

Ao observar a Equação A.49 considerando  $j = 0$  percebe-se que o algoritmo para um simples neurônio é um caso particular do Algoritmo Hebbiano Generalizado ora apresentado.

A equação que expressa a variação no peso sináptico imposta pelo Algoritmo Hebbiano Generalizado pode também ser escrita como:

$$\Delta w_{ji}(n) = \eta y_j(n) x'_i(n) - \eta y_j^2(n) w_{ki}(n) \quad \begin{cases} i = 0, 1, \dots, p-1 \\ j = 0, 1, \dots, m-1 \end{cases} \quad \text{Equação A.50}$$

Onde o vetor  $x'_i(n)$  representa a forma deflacionada do vetor de entrada  $x_i(n)$ , conforme a Equação A.51.

$$x'_i(n) = x_i(n) - \sum_{k=0}^{j-1} w_{ki}(n) y_k(n) \quad \text{Equação A.51}$$

Cabe aqui examinar a operação do algoritmo passo a passo. Para o primeiro neurônio ( $j = 0$ ), a Equação A.50, reduz-se ao caso de um único neurônio, examinado anteriormente, e a rede extrai o primeiro componente principal do vetor de dados de entrada  $x(n)$ .

Para o segundo neurônio ( $j = 1$ ), a Equação A.51 torna-se:

$$x'_i(n) = x_i(n) - w_{0i}(n) y_0(n) \quad \text{Equação A.52}$$

Desde que o primeiro neurônio já tenha convergido para o primeiro componente principal, o segundo neurônio vê um vetor de entrada  $x'_i(n)$  do qual o primeiro auto-vetor da matriz  $C_x$  já foi extraído, conforme pode ser visto na Equação A.49. O segundo neurônio extrairá, portanto, o primeiro componente principal em  $x'_i(n)$ , que é, na verdade, o segundo componente principal do vetor de entrada original,  $x_i(n)$  (segundo autovalor e correspondente auto-vetor da matriz  $C_x$ ).

Repetindo este procedimento para os demais neurônios da rede, cada conjunto de pesos sinápticos convergido representa um auto-vetor da matriz de correlação do vetor de entrada. Os auto-vetores obtidos desta forma encontram-se ordenados em ordem

decrecente do valor dos autovalores associados, sendo o primeiro neurônio a convergir aquele associado ao maior autovalor.