

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE TECNOLOGIA
PROGRAMA DE ENGENHARIA DE PRODUÇÃO

**ANÁLISE DOS INDICADORES DE QUALIDADE VERSUS TAXA DE
ABANDONO UTILIZANDO MÉTODO DE REGRESSÃO MULTIPLA
PARA SERVIÇO DE BANDA LARGA**

por

ANDRÉ PEDRO FERNANDES NETO

ENGENHEIRO ELETRICISTA, UFRN, 1993.

TESE SUBMETIDA AO PROGRAMA DE ENGENHARIA DE PRODUÇÃO DA
UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE

MESTRE EM CIÊNCIAS EM ENGENHARIA DE PRODUÇÃO

NOVEMBRO, 2007

© 2007 ANDRÉ PEDRO FERNANDES NETO

TODOS DIREITOS RESERVADOS.

O autor aqui designado concede ao Programa de Engenharia de Produção da Universidade Federal do Rio Grande do Norte permissão para reproduzir, distribuir, comunicar ao público, em papel ou meio eletrônico, essa obra, no todo ou em parte, nos termos da Lei.

Assinatura do Autor: _____

APROVADO POR:

Prof. Dr. José Alfredo Ferreira Costa, D.Sc. – Orientador, Presidente

Prof. Dr. Manoel Veras de Sousa Neto - Co-orientador.

Prof. Paulo César Formiga Ramos, D.Sc - Membro Examinador.

Prof. Marcelo Sampaio de Alencar, D.Sc - Examinador Externo.

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Divisão de Serviços Técnicos

Catálogo da Publicação na Fonte. UFRN / Biblioteca Central Zila Mamede

Fernandes Neto, André Pedro.

Análise dos indicadores de qualidade versus taxa de abandono utilizando método de regressão múltipla para serviço de banda larga / André Pedro Fernandes Neto. – Natal, RN, 2007.

208 f.

Orientador: José Alfredo Ferreira Costa.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte. Centro de Tecnologia. Programa de Pós-Graduação em Engenharia de Produção.

1. Medição de desempenho – Dissertação. 2. Qualidade de serviço – Dissertação. 3. Bases de dados – Conhecimento – Dissertação. 4. Taxa de abandono (Churn) – Dissertação. 5. Análise de regressão multivariada – Dissertação. I. Costa, José Alfredo Ferreira. II. Universidade Federal do Rio Grande do Norte. III. Título.

RN/UF/BCZM

CDU 658.53(043.3)

CURRICULUM VITAE RESUMIDO



EXPERIÊNCIA PROFISSIONAL:

- Gestão na Operação, Manutenção Corretiva e Preventiva de Grandes Clientes;
- Elaboração do Orçamento de Despesas Operacionais;
- Sistema de Gestão TELEMAR-SIG.
- Gestão dos Indicadores Operacionais de Comunicação de Dados;
- Engenharia de Planejamento de Comunicação de Dados;
- Sistema de Coleta e Visualização Gráfica de Dados (Centro de Lançamento Alcântara);
- Implantação e Manutenção de TV a Cabo.

EXPERIÊNCIA ACADÊMICA:

- Professor Assistente IV no Curso de Ciência da Computação UERN;
- Professor na cadeira de Circuitos Eletrônicos;
- Professor na cadeira de Dispositivos Semicondutores;
- Professor na cadeira de Transmissão de Dados.
- Trabalho Publicado no XIV SIMPEP (Nov./2007) Simpósio de Engenharia de Produção com o título “Sistema de Análise de Desempenho de Vendas em Telecomunicações Utilizando Técnicas de Mineração de Dados”.

FORMAÇÃO ACADÊMICA:

- Especialização – 2004. Fundação Getúlio Vargas, MBA em Gestão Empresarial;
- Especialização – 1996. Universidade Federal do Maranhão (UFMA), Departamento de Engenharia Elétrica, Curso: Engenharia Elétrica (Telecomunicações);
- Graduação – 1992. Universidade Federal do Rio Grande do Norte (UFRN), Departamento de Engenharia Elétrica, Curso: Engenharia Elétrica (Telecomunicações).

CURSOS RELEVANTES:

Cursos de aprimoramento profissional com ênfase em: Sistema Gráfico do Centro de Lançamento de Alcântara, Administração TCP-IP, Comunicação via Satélite e Aplicações, HDSL, Instalação/Administração de Site Internet Baseado em *Windows* NT, PCM, Análise de Investimento, Enlaces Ópticos, Business Object, PERT-PCM Controle e Manutenção de Estoque, Operação e Monitoração de Fibra Óptica, Wireless, Rede Óptica Primária, Transmissão Digital em Rede Metálica, PDCA Plano de Ação – Fundação Getulio Vargas, Sistema TELEMAR de Gestão – SIG, CEQUAL – Certificado de Qualidade Técnicos de Telecomunicações, CEQUAL – Certificado de Qualidade Técnicos de Comunicação de Dados.

DEDICATÓRIA

Aos meus pais, **Emanuel e Tânia** pelo seu carinho, a minha esposa, **Airleth** pelo seu amor incondicional. As minhas filhas **Ana Letícia e Ana Júlia** por serem a alegria da minha vida, ao meu irmão **Jean** e a minha irmã **Fernanda** pela amizade. As minhas cunhadas e cunhados pelos seus frutos, os meus sobrinhos, **Cinara, Felipe, Maria Luiza, Matheus, Felipe e Artur**, pois são fundamentais para a minha felicidade de viver, obrigado. Ao meu sogro **Airton** e sogra **Eliete**, por terem me emprestado uma flor, que torna a minha vida mais bonita, perfumada e melhor. E por fim ao meu avô **Antonio**, que apesar da saudade, a certeza de está sempre presente na minha vida.

AGRADECIMENTOS

Aos meus orientadores, Prof. Dr. José Alfredo Ferreira Costa e Prof. Dr. Manoel Veras de Sousa Neto, pelos seus valiosos ensinamentos e o essencial apoio que tornou esse trabalho em uma realidade.

Aos professores Dr. Paulo Formiga Ramos e Dr. Marcelo Sampaio Alencar pela sua presteza em concederem seu tempo e conhecimento, por terem aceitado participar da banca examinadora.

A todos os professores, alunos e funcionários do programa de Pós-graduação em Engenharia de Produção da UFRN, como parte fundamental e apoio necessário para realização dessa pesquisa. Agradecimento especial ao Departamento de Engenharia da Elétrica, seus professores e alunos e funcionários, pela maravilhosa formação e estrutura que me forneceram ao longo desses anos de convívio.

Aos meus colegas de turma, cujo trabalho de equipe contribuiu, sem dúvida, para êxito desse trabalho. E a todos aqueles que, de maneira direta ou indireta, contribuíram para a realização desse trabalho.

Resumo da Tese apresentada à UFRN/PEP como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências em Engenharia de Produção.

ANÁLISE DOS INDICADORES DE QUALIDADE VERSUS TAXA DE ABANDONO UTILIZANDO MÉTODO DE REGRESSÃO MULTIPLA PARA SERVIÇO DE BANDA LARGA

ANDRÉ PEDRO FERNANDES NETO

Novembro/2007

Orientador: José Alfredo Ferreira Costa

Curso: Mestrado em Ciências em Engenharia de Produção

A área de telecomunicações é uma das mais estratégicas e dinâmicas do mundo atual. Esse fato se deve a inúmeras inovações tecnológicas que afetaram a forma como as informações trafegam. O conhecimento deixou de ser percebido como um acúmulo linear, lógico e cronológico de informações e passou a ser visto como uma construção em rede, consequentemente a massificação da Internet banda larga em alta velocidade teve grande influência sobre esse fenômeno.

Essa dissertação aborda um estudo sobre medição de desempenho e serviços de manutenção em telecomunicações, com o uso de ferramentas de descoberta de conhecimento em base de dados (KDD). Objetiva-se transformar informações, armazenadas nas bases de dados de uma grande empresa de telecomunicações do país, em conhecimento útil. A metodologia de pesquisa utilizada focou no uso de análise de regressão múltipla como ferramenta para estimar a taxa de abandono de clientes em serviços de Internet de banda larga, como variável dependente, e indicadores de qualidade de serviço como variáveis independentes.

Modelos foram testados e analisados estatisticamente. O trabalho apresenta resultados que permitem analisar e identificar quais os indicadores de qualidade que exercem maior influência na taxa de abandono dos clientes. São propostas sugestões que possam ser aplicadas para melhoria de qualidade do serviço percebido e consequentemente diminuições das perdas com a taxa de abandono.

Palavras-chave: Medição de Desempenho; Qualidade de Serviço; Descoberta de Conhecimento em Bases de Dados; Taxa de Abandono (Churn). Análise de Regressão Multivariada.

Abstract of Master Thesis presented to UFRN/PEP as fulfillment of requirements to the degree of Master of Science in Production Engineering

November / 2007

Thesis Supervisor: José Alfredo Ferreira Costa
Program: Master of Science in Production Engineering

Telecommunication is one of the most dynamic and strategic areas in the world. Many technological innovations has modified the way information is exchanged. Information and knowledge are now shared in networks. Broadband Internet is the new way of sharing contents and information.

This dissertation deals with performance indicators related to maintenance services of telecommunications networks and uses models of multivariate regression to estimate churn, which is the loss of customers to other companies. In a competitive environment, telecommunications companies have devised strategies to minimize the loss of customers. Loosing customers presents a higher cost than obtaining new ones.

Corporations have plenty of data stored in a diversity of databases. Usually the data are not explored properly. This work uses the Knowledge Discovery in Databases (KDD) to establish rules and new models to explain how churn, as a dependent variable, are related to a diversity of service indicators, such as time to deploy the service (in hours), time to repair (in hours), and so on. Extraction of meaningful knowledge is, in many cases, a challenge.

Models were tested and statistically analyzed. The work also shows results that allows the analysis and identification of which quality services indicators influence the churn. Actions are also proposed to solve, at least in part, this problem.

Keywords: Performance measurement; Quality of Services; Knowledge Discovery in Databases; Churn; Models of Multivariate Regression; Decision Support Systems.

SUMÁRIO

LISTA DE TABELAS	xii
LISTA DE FIGURAS E GRAFICOS	xiv
LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS.	xvi
RESUMO	vii
ABSTRACT	viii
Capítulo I – INTRODUÇÃO	1
1.1 Problema	3
1.2 Objetivo	4
1.3 Justificativa	4
1.4 Estrutura geral do trabalho	5
Capítulo II - REVISÃO DE LITERATURA	
2.1 MEDIÇÃO E DESEMPENHO	6
2.1.1 Características e Elementos de Serviços	7
2.1.2 Qualidade de Serviço	9
2.1.2.1 A Qualidade Percebida em Serviço	10
2.1.2.2 <i>Marketing</i> de Serviços	15
2.1.2.3 Valor	17
2.1.2.4 Serviços Internos	25
2.1.3 Indicadores de Qualidade e Medidas de Qualidade	28
2.1.3.1 Histórico dos Indicadores de Qualidade	29
2.1.3.2 Medidas de Qualidade	30
2.1.3.3 Finalidade de Medir	32
2.1.4 Indicadores de Gestão	37
2.2 SERVIÇOS DE MANUTENÇÃO EM TELECOMUNICAÇÕES	40
2.2.1 O Setor de Telecomunicações no Brasil	41
2.2.2 As Empresas Operadoras de Telecomunicações	42
2.2.3 O Serviço de Telecomunicações de Internet Banda Larga	44
2.2.4 A Engenharia de Manutenção	46
2.2.5 Modelo de Gestão para as Operações de Manutenção	60
2.3 DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS	80
2.3.1 Dados	82
2.3.1.1 <i>Data Warehouse</i>	82

2.3.1.2	Análise Múltipla de Dados	83
2.3.2	Caracterização do Processo de Descoberta de Conhecimento em Base de Dados	85
2.3.2.1	Pré-processamento	86
2.3.2.2	Mineração de Dados	89
2.3.2.3	Pós-processamento	105
2.3.2.4	Controle	107
2.3.3	Metodologias de Descoberta de Conhecimento em Base de Dados	107
2.3.3.1	Tipos de Metodologias	108
2.3.3.1	Avaliação das Metodologias	118
Capítulo III – METODOLOGIA UTILIZADA		118
3.1	Introdução	118
3.2	Objetivo (Estágio 1)	122
3.3	Planejamento da Pesquisa (Estágio 2)	138
3.4	Suposições em Análise de Regressão Linear Múltipla (Estágio 3)	140
3.5	Estimação do Modelo de Regressão e Avaliação do Ajuste Geral do Modelo (Estágio 4)	149
3.6	Interpretação das Variáveis Estatísticas (Estágio 5)	154
3.7	Validação dos Resultados (Estágio 6)	156
Capítulo IV – ANÁLISE DOS RESULTADOS		159
4.1	Introdução	159
4.2	Estimação do Modelo de Regressão, Avaliação do Ajuste Geral do Modelo e da Significância do Modelo.	160
4.2.1	Método Combinatório	160
4.2.2	Método Sequencial <i>STEPWISE</i>	165
4.2.3	Identificação de Observações Influentes	170
4.3	Interpretação das Variáveis Estatísticas e Validação dos Resultados	171
4.4	Conclusão	173
Capítulo V – CONCLUSÕES E RECOMENDAÇÕES		
5.1	Conclusões	174
5.2	Recomendações para Estudo Futuro	177

Referências Bibliográficas	178
Anexo I	187
Anexo II	189
Anexo III	190
Anexo IV	194
Anexo V	206

LISTA DE TABELAS

Tabela 2.1	Percentual de empregos em serviços em alguns países industrializados
Tabela 2.2	Indicadores de negócio
Tabela 2.3	Indicadores de Qualidade e Desempenho
Tabela 2.4	Estatísticas gerais do Brasil como população, PIB e pesquisas do IBGE (PNAS, PAS). <i>Ranking</i> internacional e estatístico de serviços de telecomunicações no Brasil.
Tabela 2.5	Quantitativo de telefones fixos nas maiores economias do mundo.
Tabela 2.6	Evolução trimestral das tecnologias de acesso à banda larga.
Tabela 2.7	Evolução trimestral em milhares de terminais nas operadoras ADSL
Tabela 2.8	Tabela de Priorização dos Fatores de Avaliação do Serviço de Manutenção, Segundo a Técnica de Mudge.
Tabela 3.1	Indicadores de Qualidade.
Tabela 3.2	Indicadores de Negócio.
Tabela 3.3	Tabela de priorização dos fatores de avaliação do serviço de manutenção segundo a técnica de Mudge.
Tabela 3.4	Fatores de avaliação do serviço de banda larga X escala de graduação da importância relativa para o cliente.
Tabela 3.5	“Atributos de Valor” versus processos críticos nas operações de manutenção.
Tabela 3.6	Causas da <i>taxa de abandono</i> em uma empresa.
Tabela 3.7	Descrições Estatísticas.
Tabela 3.8	Descrições Estatísticas Normalizada.
Tabela 4.1	Sumário do modelo Combinatório.
Tabela 4.2	Coefficiente do modelo Combinatório.
Tabela 4.3	<i>ANOVA</i> do modelo Combinatório.
Tabela 4.4	Sumário do modelo <i>STEPWISE</i> .
Tabela 4.5	Coefficientes do modelo <i>Stepwise</i> .
Tabela 4.6	<i>ANOVA</i> do modelo <i>Stepwise</i> .
Tabela 4.7	Correlações.

Tabela 4.8	Diagnóstico <i>Case wise</i> .
Tabela AII.1	Níveis de significância.
Tabela AIII.1	Estimativa de perda.
Tabela AIII.2	Valores médios gastos com abandono.
Tabela AIII.3	Estimativa de despesas com abandono.
Tabela AIII.4	Parâmetros de qualidade.

LISTA DE FIGURAS E GRÁFICOS

- Figura 2.1 Vantagens e desvantagens de se deslocar à linha de visibilidade.
- Figura 2.2 Qualidade percebida versus qualidade esperada.
- Figura 2.3 Modelo de percepção de satisfação ao cliente.
- Figura 2.4 *Marketing* Integrado em Serviços.
- Figura 2.5 Valor como vantagem competitiva na relação cliente x fornecedor.
- Figura 2.6 Fatores de Percepção do Valor.
- Figura 2.7 O Processo de Criação do Valor e da Satisfação.
- Figura 2.8 A Cadeia de Valor.
- Figura 2.9 Preconizações para Utilização Critérios MBNQA da "Orientação/Satisfação do Cliente" em Serviços Internos.
- Figura 2.10 Distribuição das áreas de telefonia fixa no Brasil.
- Figura 2.11 Interação entre as fases.
- Figura 2.12 Evolução da postura da manutenção.
- Figura 2.13 Tipos de manutenção.
- Figura 2.14 Custo de manutenção em relação ao tempo.
- Figura 2.15 Dados do Sistema Genérico de Operações e Manutenção.
- Figura 2.16 Organização das Operações de Manutenção por Níveis de Atividades, Perímetros e Funções.
- Figura 2.17 Fatores da Percepção do Valor e Indicadores Associados, para as Operações de Manutenção.
- Figura 2.18 Fatores de Avaliação do Serviço de Manutenção X Escala de Graduação da Importância Relativa para o Cliente.
- Figura 2.19 Análise de Conflito dos "atributos de valor" das Operações de Manutenção Ab Fb Cs Ct Cp Cf Va.
- Figura 2.20 "Atributos de valor" X Processos Críticos nas Operações de Manutenção
Figura.

- Figura 2.25 Metodologia *SEMMA*.
- Figura 2.26 Metodologia ADRIAANS e ZANTINGE.
- Figura 2.27 Metodologia KLEMENTINEN.
- Figura 2.28 Metodologia FELDENS.
- Figura 3.1 Estágios no diagrama de decisão de regressão múltipla.
- Figura 3.2 Caracterização dos serviços e entendimento do valor oferecido ao cliente dentro do quadro geral de competitividade em serviço.
- Gráfico 3.3 Análise dos gráficos de regressão parcial (TMR, TRP, QRPR, TMI, IGT, RAV).
- Gráfico 3.4 Análise dos gráficos de regressão parcial (RRT, RPT, RRDC, RPAV, RPDC, IPGC).
- Gráfico 3.5 Resíduos (estudentizados) versus os valores dependentes.
- Gráfico 3.6 Resíduos em relação a qualquer variável sequencial possível.
- Gráfico 3.7 Normal dos resíduos.
- Gráfico 3.8 Histograma da taxa de abandono.
- Gráfico 4.1 Identificação de pontos atípicos.
- Figura AV.1 Importação de Dados.
- Figura AV.2 Regressão linear múltipla por meio do *SPSS*.
- Figura AV.3 Representação da seleção das variáveis.
- Figura AV.4 Seleção dos dados estatísticos.
- Figura AV.5 Seleção dos gráficos estatísticos.
- Figura AV.6 Escolha da probabilidade “F”.

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.

ABNT	Associação Brasileira de Normas Técnicas.
ADSL	<i>Asymmetric Digital Subscriber Line.</i>
ANATEL	Agencia Nacional de Telecomunicações.
CBM	Condition Based Maintenance.
CTB	Companhia Telefônica Brasileira.
DDD	Discagem Direta a Distância.
FAA	<i>Federal de Aviação Americana.</i>
FMEA	<i>Failure Mode and Effect Analysis.</i>
IBGE	Instituto Brasileiro de Geografia e Estatística.
IGT	Instalação em Garantia Total.
IPGC	Instalação no Prazo Grandes Clientes.
KDD	<i>Knowledge Discovery in Database.</i>
NBR	Normas Brasil.
MTBF	<i>Medium Time Between Failure.</i>
MTTR	<i>Medium Time to Repair.</i>
PAS	Pesquisa Anual de Serviços.
PNAS	Política Nacional de Assistência Social.
PDCA	<i>Plan Do Check Action.</i>
PGMQ	Plano Geral de Metas de Qualidade.
PIB	Produto Interno Bruto.
QPC	Qualidade Percebida Pelo Cliente.
Qos	Qualidade de Serviço.
QRPR	Reparos Preventivos.
RCM	Manutenção Centrada em Confiabilidade.
RDSI	Rede Digital de Serviço Integrada.
RPAV	Reparo no Prazo Alto Valor.
RPDC	Reparo no Prazo Demais Clientes.
RPT	Reparo no Prazo Total.
RAAV	Reparo Reincidente Alto Valor.
RRDC	Reparo Reincidente Demais Clientes.

RRT	Reparo Reincidente Total.
SPSS	<i>Statistical Package for the Social Sciences.</i>
STFC	Sistema de Telefonia Fixa Comutada.
SQE	Soma dos Quadrados Explicada.
SQR	Soma dos Quadrados dos Resíduos.
SQT	Soma dos Quadrados Total.
TBM	<i>Time Based Maintenance.</i>
TPM	<i>Total Productive Maintenance.</i>
TI	Tecnologia da Informação.
TMR	Tempo Médio de Reparo.
TRP	Taxa de Reparo.
TMI	Tempo Médio de Instalação.
VIF	<i>Variance Inflation Factor.</i>
VPC	Valor Percebido pelo Cliente.

Capítulo 1

Introdução

A informação desconhece as fronteiras geográficas, propagando-se com facilidade e rapidez pelo mundo, o setor da economia ligado ao segmento de telecomunicações transformou-se nos últimos anos em um dos mais promissores, estratégicos e dinâmicos, propondo constantemente desafios á indústria, as empresas e aos profissionais que consolidam as tecnologias e serviços, de uma forma eficiente, segura e economicamente viável. De acordo com SIQUEIRA, (2004) esse novo mundo é em grande medida, moldado pela convergência digital, resultante da fusão das tecnologias da informação, das telecomunicações e de multimídia, ou seja, para traduzir essa era em uma única palavra seria infocomunicação. A Infocomunicação encontra-se no cerne desse novo paradigma que tornou o suporte sem o qual a globalização, tal como se apresenta, não seria possível.

O mundo globalizado vem exigindo, a cada ano, um crescente aumento de taxa de transmissão nas redes de computadores, a banda larga viabiliza aumento de velocidade nessa estrada de conhecimento, informação, e com isso proporciona uma maior possibilidade de coleta e armazenamentos múltiplos de dados. Por sua vez os pesquisadores vêm tentando desenvolver uma forma mais inteligente para analisar, interpretar e relacionar bases de maneira útil em benefício das organizações. Surgiram assim as análises múltiplas, que são técnicas de análise de dados, com as quais se procura uma forma mais adequada para atacar um problema envolvendo uma única variável dependente com duas ou mais variáveis independentes e contínuas. A descoberta de conhecimento em base de dados por meio de análise múltiplas, é um processo em várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grande conjunto de dados [FAYYAD et al., (1996)].

Já o aperfeiçoamento constante, em mercados competitivos é condição para o crescimento e desenvolvimento das empresas. O aumento da concorrência e dos custos associados à atração de novos clientes impõe o desafio de reter clientes cada vez mais exigentes, fazendo com que as empresas alterem seu comportamento tático e estratégico. Como resposta a esse cenário, uma das práticas de gestão que começou a ser empregada foi o gerenciamento da taxa de abandono. Em função disso, muitos setores, em diversos segmentos, estão realizando investimentos para reformular seus sistemas de medição de desempenho no sentido de acompanhar melhor os movimentos dos clientes e adaptar as novas estratégias aos seus processos. Uma forma de garantir a eficiência na gestão é obter uma base de medidas de indicadores de qualidade de serviço, que sejam operacionais e significativamente úteis aos responsáveis pelas decisões. Esses indicadores precisam refletir a forma como o cliente enxerga o serviço prestado.

O relacionamento entre indicadores de desempenho e qualidade da gestão de serviços é relevante para o ambiente empresarial atual. As prestadoras de serviços estão incluídas nesse contexto, já que a formulação das estratégias, o planejamento e aplicação das ações, programas e processos de melhoria e mudanças nas operações iniciam-se na elaboração de métricas que possam monitorar seus estágios, passados e atuais, permitindo previsões para o futuro (ARNETT, 2000). Diante desse cenário, escolheu-se analisar os indicadores de desempenho operacional do segmento de manutenção, melhoria e de instalação. Essa análise foi feita verificando-se o comportamento dos indicadores de qualidade, peça fundamental no planejamento de mercado e da estratégia de qualquer empresa (NEELY, 1997).

Nesta pesquisa procurou se relacionar a taxa de abandono e a correta gestão de qualidade dos serviços, pois se por um lado devem ser garantidos níveis de satisfação que tornem o cliente fiel, por outro se gasta tempo e dinheiro na procura de novos clientes. Esse direcionamento foi feito por meio da análise de uma empresa de telecomunicações operadora de serviço em banda larga, tendo o objetivo de auxiliar na gestão das áreas operacionais das prestadoras de serviços, bem como no eficiente tratamento e aproveitamento dos dados armazenados.

1.1 Definição do Problema

Foi nesse contexto globalizado em que a telemática exige cada vez mais elevadas taxas de transmissão que se direcionou essa pesquisa para empresas operadoras de banda larga, tendo o objetivo de auxiliar na gestão das áreas operacionais das prestadoras de serviços, com a apresentação de uma análise de desempenho da qualidade de serviços versus a taxa de abandono. Buscar-se a explicação de um problema de gerenciamento da relação com o cliente, com a criação, estudo e otimização do processo de mineração de dados.

A relação entre a taxa de abandono de cliente e a correta gestão de qualidade dos serviços, caracteriza-se como um problema, pois se por um lado se tenta garantir níveis de satisfação que tornem o cliente fiel, por outro se gasta tempo e dinheiro na procura de novos clientes. Pretende-se avaliar a relação entre taxa de abandono e qualidade dos serviços, em busca de uma explicação satisfatória para o problema.

O processo de descoberta e conhecimento de base de dados (KDD) inicia-se com o estudo do problema, passando pelo tratamento dos dados e culminando na identificação e explicação do modelo mais adequado para a tarefa de combate e diminuição da taxa de abandono.

1.2 Objetivo

Essa pesquisa tem como objetivo de analisar, descrever e explicar em parte o comportamento do cliente no setor de telecomunicações, usuário de Internet de banda larga, por meio da análise de uma equação em que a taxa de abandono da planta pode ser relacionada com alguns indicadores de qualidade dos serviços. Tenta-se entender como o cliente percebe a qualidade dos serviços prestados e como isto pode afetar diretamente a taxa de abandono da planta. Os resultados foram obtidos por meio de análise dos indicadores, que foram obtidos a partir de ligações telefônicas para um *call center* efetuadas pelos clientes. O estudo detém-se no período referentes ao ano de 2006.

O objetivo principal pode ser desdobrado nos seguintes objetivos específicos:

- Descrever e explicar uma forma de relação entre os serviços de banda larga disponibilizados pelas empresas de telecomunicações e a taxa de abandono da planta;

- Analisar um conjunto de variáveis que possuam a maior relevância para o problema em estudo.

1.3 Justificativa

Muitas empresas em diversos segmentos estão realizando altos investimentos para reformular seus sistemas de medição de desempenho no sentido de acompanhar melhor os movimentos dos clientes e adaptar as novas estratégias aos seus processos, buscando continuamente uma melhor relação custo/benefício. O segmento de telecomunicações não é diferente sendo um dos mais proeminentes da economia, respondendo atualmente por 7% do PIB nacional. Sua intensiva busca tecnológica leva-o a assumir um perfil essencialmente globalizado, sendo que a Internet de banda larga está ampliando a sua quantidade de clientes, possibilitando com isso a implantação de novos serviços e tecnologias. Diante desse cenário, escolhe-se analisar os indicadores de qualidade de serviço prestado do segmento. A análise é feita verificando o comportamento dos indicadores de qualidade, peça fundamental no planejamento de mercado e da estratégia de qualquer empresa.

Conforme citado por MARTINS & NETO, (1998) a evolução da gestão da qualidade atrelada à evolução do conceito da qualidade tem tornado essa filosofia de gestão um importante diferencial para organizações que atuam num contexto cada vez mais competitivo.

1.4 Estrutura geral do trabalho

Essa dissertação é composta por cinco capítulos, contando com esse:

O Capítulo 2 é composto pela revisão bibliográfica, tendo sido dividido em três assuntos. Inicialmente, trata-se da medição de desempenho, e se procura caracterizar a Qualidade de Serviço e os seus Indicadores de Qualidade, por meio da contextualização do processo de qualidade e sua importância na economia. A seguir trata-se das questões da gestão de qualidade nos serviços de manutenção, com ênfase no setor de telecomunicações em que a pesquisa essa inserida, caracterizando com isto a necessidade de um modelo de gestão para as operações de manutenção em telecomunicações. Por fim apresentam-se os conceitos da descoberta de conhecimento em bases de dados (KDD), discutindo-se pontos como metodologias de descoberta de conhecimento, tipos e avaliação de metodologias bem como as ferramentas de *data mining* utilizadas.

O Capítulo 3 apresenta a metodologia utilizada na pesquisa. Para isto procura-se detalhar o problema, o planejamento e as suposições da pesquisa abrangendo suas técnicas e métodos utilizados. Por fim, de forma conceitual, procurou-se estimar o modelo de regressão bem como interpretar as variáveis estatísticas e a validação dos resultados.

O Capítulo 4 está dividido na estimação do modelo de regressão e na avaliação do ajuste geral do modelo. Para isto selecionou-se um método para especificar um modelo. A seguir avaliou-se a significância desse modelo bem como a identificação das observações influentes para avaliação do ajuste geral e interpretação das variáveis estatísticas na validação do modelo.

O Capítulo 5 apresenta as conclusões ao trabalho desenvolvido, identificando as contribuições para área e recomendações para trabalhos futuros.

Capítulo 2

Revisão da Literatura

2.1 Medição e Desempenho

Introdução

FNQ (Fundação Nacional de Qualidade) conceitua que a qualidade é expressa por critério e fundamentos que compõem os atributos de um modelo de gestão, porém, a qualidade baseada em serviços remete à teoria de EDWARDS, (1968) na qual qualidade consiste na capacidade de satisfazer os desejos dos clientes. O relacionamento entre indicadores de desempenho e qualidade da gestão de serviços é relevante para o ambiente empresarial atual, e as empresas prestadoras de serviços estão incluídas nesse contexto, já que a formulação das estratégias do planejamento e aplicação das ações, programas e processos de melhoria e mudanças nas operações inicia-se, na elaboração de métricas que possam monitorar seus estágios, passados e atuais, permitindo previsões para o futuro.

A importância das atividades de serviços na economia ocidental apresenta uma ciclicidade histórica, elas existem, enquanto categoria de atividade econômica, desde a Grécia Clássica, a importância dada pela sociedade à educação dos jovens é conhecida, porém essa atividade acabava tendo um papel econômico secundário, numa sociedade escravocrata e agrícola.

Na baixa idade média, os serviços de transporte de especiarias e tecidos pela rota da seda, atravessava a Europa e o Oriente até a China, fizeram à riqueza de cidades-estado, como Veneza, sendo a principal atividade econômica de países como Portugal e Holanda com suas companhias de navegação.

A partir do século XVIII, com a primeira revolução industrial, o setor de serviços perdeu sua relativa importância, para somente ser retomada em meados do século XX de maneira crescente e consistente até os tempos atuais.

Em todos os países desenvolvidos, o setor de serviços ocupa posição de destaque na economia [CORRÊA, (2002)]. O principal indicador utilizado para evidenciar esse fato tem sido a participação do setor de serviços na ocupação da mão-de-obra e na geração de riquezas, refletida pelo Produto Interno Bruto.

Na Tabela 2.1 pode-se notar o aumento no percentual de empregos gerados na área de serviço, bem como o elevado grau de importância que esse mercado ocupa na economia de países desenvolvidos.

Tabela 2.1 Percentual de empregos em serviços em alguns países industrializados.

Países	1980	1987	1993	1999	2005
EUA	67,1	71,0	74,3	80,4	82,5
Canadá	67,2	70,8	74,8	73,9	73,9
Japão	54,5	58,1	59,9	72,4	73,0
França	56,9	63,6	66,4	70,8	71,0
Israel	63,3	66,0	68,0	70,7	70,7
Brasil	45,6	49,3	53,2	60,0	68,0
Itália	48,7	57,7	60,2	61,1	62,0
China	13,1	17,8	21,2	26,4	30,4

Fonte: UNITED NATIONS, (2005).

O Brasil estatisticamente acompanha o ritmo apresentado no mundo industrializado. Em termos de empregos o setor de serviços vem aumentando a sua participação a cada década, sendo responsável em 2005 por 68% da ocupação da população economicamente ativa do País. Esse número é muito próximo ao valor encontrado nos EUA e superiores a países como Itália e Canadá, sendo, porém bem superiores aos valores encontrados na China.

2.1.1 Características e Elementos de Serviços

Atribui-se a Taylor a sistematização do estudo e análise do trabalho o que se tornaria mais tarde a área de gestão de operações. Taylor estabeleceu princípios que passaram a ser conhecidos como os princípios da administração científica, sistematizando técnicas e

princípios que, em seu conjunto contribuíram para o aumento substancial dos níveis de eficiência da indústria no século XX.

Nas empresas modernas apenas os princípios de Taylor associados aos mandamentos de intercambialidade de peças introduzidos por Ford, com o seu lendário Modelo T não são capazes de fazer uma organização se sobressair em um ambiente competitivo. Fala-se agora em oferecer pacotes de valor aos clientes. Esses pacotes são comumente formados por parcelas de serviço e produtos. O gestor de operações tem que gerenciar tanto a parte de serviço como gerenciar a parte do produto físico, pois o cliente enxerga a empresa fornecedora como uma geradora de soluções para as suas necessidades.

Tradicionalmente, os autores colocam três principais diferenças entre produto e serviço:

- Serviços seriam simultâneos;
- Serviços seriam intangíveis;
- Serviços precisariam de presença do cliente para ser produzidos.

A simultaneidade caracteriza-se pelo fato da inseparabilidade do serviço com a produção, ou seja, os serviços são produzidos e consumidos simultaneamente, sendo difícil de gerenciar o controle de qualidade e praticar um *marketing* no sentido tradicional. Dessa forma não há uma qualidade pré-produzida para ser controlada com antecedência, isto é, antes do serviço ser consumido.

GRONROOS, (1993) e CORRÊA, (2002) concordam que existem duas situações distintas em relação à visibilidade do cliente em função do serviço consumido. Uma das situações é que o cliente está presente e consome o serviço (ex: cabeleireiro). A outra situação ocorre quando da entrega de bens, apenas parte do processo de produção do serviço é vivenciado e, portanto consumido simultaneamente pelo cliente. Entretanto é importante compreender que a parte de produção visível é a que vai influenciar a perspectiva do cliente.

Outro aspecto importante é a intangibilidade dos serviços. CORRÊA, (2002) salienta que uma implicação da intangibilidade maior ou menor dos serviços está na maior ou menor facilidade com que se avalia a qualidade do pacote de valor a ser recebido. Essa característica faz com que o serviço seja de difícil avaliação em sua qualidade pelo gestor da operação.

Quando se necessita da presença do cliente para realização da produção, faz com que o cliente deixe de ser um simples receptor, para torná-lo um recurso da produção. Por causa desse e de outros fatores, aponta [GRONROOS, (1993)] como sendo características para não se ter serviços em estoque, da mesma forma que se mantêm os bens.

2.1.2 Qualidade de Serviço

Quando se afirma que em serviço o cliente tem algum nível de contato com algum aspecto do processo produtivo, obrigatoriamente o mesmo não tem contato com todo o processo produtivo. O grau de contato que o cliente tem com o processo depende muito da sua interação e extensão com o processo da empresa, variando também em relação ao tempo de contato e que tipo de contato se estabelecerá.

As atividades que pertencem ao processo de prestação de serviço que ocorrem com a participação ativa do cliente podem ser definidas como atividades de *front Office* ou atividades de alto contato [CORRÊA, (2002)].

Nem sempre as pessoas ou os recursos podem ser divididos como de *front Office* ou de retaguarda, porém têm-se vantagens em usar essa classificação, pois se identifica as variáveis que influenciam na gestão dos serviços. A seguir algumas dessas variáveis:

- Estocabilidade;
- Intensidade e extensão da interação no contato com o cliente;
- Objetivação possível na avaliação do desempenho.

Segundo TEBOUL, (1995) um serviço não consumido é um serviço perdido. Salienta que as atividades de *front Office* tendem a ter um baixo grau de estocabilidade, maior grau de intensidade com o cliente bem como o serviço tende a ser um alto grau de subjetividade, sendo essas variáveis que se encontram as maiores complexidades, tradicionalmente associadas aos serviços.

O gestor do processo é responsável por outro aspecto importante para avaliação da qualidade de *front Office*, ele determina a linha de visibilidade, ou seja, ele define a linha que separa o que o cliente tem contato do que o cliente não tem contato. Existem serviços que estão deslocando a sua linha de visibilidade para a esquerda, para reduzir as atividades de

linha de frente e com isto trabalha com níveis menores de variabilidade e incertezas, ter mais controle sobre a operação e com isto atingir níveis mais altos de eficiências e redução dos custos. Por outro lado, há operações que resolvem aumentar a área de contato com os clientes deslocando sua linha de visibilidade para a direita aumentando os benefícios e custos decorrentes [CORRÊA, (2002)].

A figura 2.1 ilustra as possíveis vantagens e desvantagens de se deslocar à linha de visibilidade para aumentar ou diminuir a área de contato com o cliente.

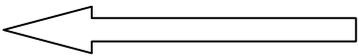
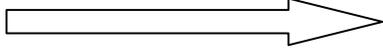
Atividades de Frente Office	Atividade de Retaguarda 	
Vantagens Maior Eficiência operacional Libera tempo do Front Office para relacionamento Maior consistência de qualidade	Desvantagens Menor relacionamento com o cliente Aumenta distancia da gestão para o cliente Imagem do serviço impessoal	
Atividades de Front Office 		Atividades de Retaguarda
Vantagens Maior interação e relacionamento com o cliente Imagem do serviço pessoal e customizado	Desvantagens Menor eficiência operacional Menor consistência na qualidade Aumenta dependência da força de trabalho	

Figura 2.1 Vantagens e desvantagens de se deslocar à linha de visibilidade.

Fonte: Adaptado de CORRÊA, (2002).

2.1.2.1 A Qualidade Percebida em Serviços

A qualidade é, para as operações, condição para a produção de bens e serviços competitivos. Em operações de fabricação de bens, os produtos podem ser verificados, inspecionados e modificados ao longo do processo. Esses procedimentos não são visualizados pelos clientes, mas permitem atingir a satisfação dos mesmos, com a oferta de produtos sem nenhum defeito e que atenda às suas necessidades.

O produto de uma operação de serviço não pode ser apresentado, possuído ou comprado da mesma forma que um bem, fruto das suas especificidades. Para que as operações de serviços produzam atividades de qualidade é necessário primeiramente entender como o cliente de serviços percebe a qualidade em função das suas características específicas. As especificidades de serviços, particularmente a ausência de critérios tangíveis e a dificuldade de entender os critérios do cliente, tornam a avaliação da qualidade de serviços mais delicada que a de bens manufaturados. Basicamente a qualidade em serviços é difícil de ser assegurada, pois:

- Há simultaneidade entre a produção e consumo;
- É dependente de relações Inter-Pessoais;
- A sucessão de encontros de serviços entre o cliente e os membros da empresa, encarregados do contato (linha-de-frente), que constituem os momentos da verdade [NORMANN, (1993)], onde se formam as percepções do serviço;
- A descentralização dos agentes prestadores de serviço, tornando difícil o controle sobre as atitudes e comportamentos Inter-Pessoais.

Como os serviços possuem como característica principal a intangibilidade, função de sua subjetividade principalmente em relação à qualidade intrínseca do serviço, normalmente define-se a qualidade como percebida pelo cliente (QPC). A qualidade percebida é o resultado da comparação entre a qualidade experimentada e a esperada pelo cliente. Dessa forma, boa qualidade ou qualidade satisfatória percebida é atingida quando a qualidade experimentada atende às expectativas do cliente.

O processo de percepção da qualidade é complexo, pois além de ser uma parte bem subjetiva do cliente leva em conta a qualidade esperada do serviço.

A qualidade experimentada em serviços possui duas dimensões básicas: a qualidade técnica ou "o que" o cliente recebe como solução técnica, e a qualidade funcional ou "como" o cliente recebe o serviço e como ele vivencia o processo de produção e consumo simultâneo [GRONROOS, (1993)]. A imagem da organização, nesse caso, funciona como um filtro para a percepção da qualidade pelo cliente, amenizando o impacto de fatores que possam influenciá-la negativamente. Por sua vez, a qualidade esperada é influenciada por alguns fatores de interação com o mercado:

- Imagem: apresenta a função básica na formação da qualidade esperada, criando expectativas a partir de experiências anteriores;
- Comunicação boca a boca: influencia na expectativa e na criação da mensagem sobre a organização, sua credibilidade e confiabilidade;
- Comunicação com o mercado: a qualidade esperada é influenciada por meio de campanhas publicitárias, treinamento do cliente, comunicações, etc.;
- Necessidades dos consumidores: que é um fator inerente à busca da satisfação pelo cliente;
- Experiências anteriores: as quais induzem diretamente o cliente na formação da sua expectativa em relação à qualidade.

Algumas dessas variáveis podem ser diretamente controladas pela empresa sendo exemplos desses controles a comunicação com o mercado: publicações, mala direta, relações públicas e campanhas de vendas, já os fatores de imagem e comunicação boca a boca são apenas indiretamente controlados. O impacto externo desses fatores também pode ocorrer, mas eles são basicamente funções do desempenho prévio da empresa, Finalmente, as necessidades dos clientes têm impacto nas suas expectativas.

Na figura 2.2 ilustra-se a forma de relacionamento entre a qualidade esperada e qualidade percebida e como os fatores podem se relacionar.

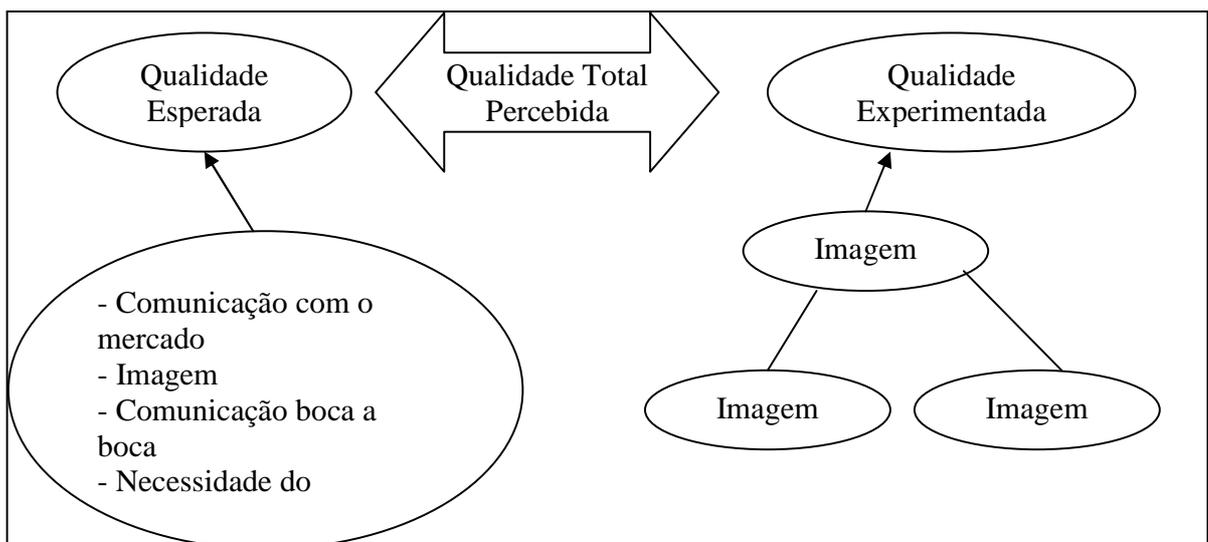


Figura 2.2 Qualidade percebida versus qualidade esperada.

Fonte: Adaptado de GRONROOS, (1993).

O nível de qualidade total percebida não é determinado pelo nível das dimensões da qualidade técnica e funcional apenas, mas sim pela diferença entre a qualidade esperada e a qualidade experimentada. Conseqüentemente um programa de qualidade para ser percebido pelo cliente tem que envolver todos os atores responsáveis como produção, *marketing* e pela comunicação com o mercado.

Segundo BERRY, (1983) existem dez fatores de qualidades percebidos pelos clientes, são eles:

1 - Confiabilidade envolve coerência no desempenho e constância: a empresa desempenha o serviço de forma correta na primeira vez; precisão no faturamento; manutenção correta dos serviços e execução dos serviços no tempo correto.

2 – Rapidez de resposta diz respeito à taxa de transmissão empregada para fornecer o serviço: serviço prestado em tempo adequado e prestação imediata dos serviços

3 – Competência consiste em ter as habilidades e o conhecimento necessário: habilidade em interagir com o cliente e conhecimento e habilidade do pessoal de retaguarda.

4 - Acesso envolve facilidade de abordagem e de contato: facilidade de acesso ao serviço, o tempo de espera é conveniente e local conveniente de execução.

5 – Cortesia envolve boas maneiras, respeito, consideração e contato amigável: consideração pela propriedade do consumidor e aparência pessoal.

6 – Comunicação significa manter os clientes informados, bem como ouvir os clientes: explicar o serviço.

7 – Credibilidade envolve confiança, crédito, honestidade e comprometimento com os interesses do cliente: reputação da empresa e envolvimento nas interações.

8 - Segurança é estar livre de perigos, riscos ou dúvidas: Segurança física e financeira e confiabilidade.

9 – Compreender e conhecer o cliente esforço empregado para compreender as necessidades do cliente: conhecer as exigências específicas do cliente e fornecer atenção individualizada.

10 – Tangíveis incluem evidências físicas do serviço: instalações físicas e ferramentas e serviços utilizados.

A partir da análise dos dez fatores BERRY, (1983) desenvolveu o modelo de análise do gap da qualidade, que se destinam a serem usados para análise das fontes dos problemas da qualidade no auxílio aos gerentes a tomadas de decisão.

Gap 1 Percepção gerencial: esse gap pode ser causado pela imprecisão das informações geradas para os gerentes ou por um erro na interpretação gerencial

Gap 2 Especificação da qualidade: esse gap significa que as especificações da qualidade por serviço não são coerentes com as percepções da gerencia das expectativas da qualidade. Sendo esse gap ligado a problemas com o planejamento.

Gap 3 Entrega do serviço: mostra-se que as especificações da qualidade não são atendidas pelo desempenho do processo da produção e entrega dos serviços.

Gap 4 Comunicação com o mercado: significa que promessas feitas não são coerentes com os serviços prestados.

Gap 5 Qualidade percebida do serviço: significa que o serviço percebido não é coerente com o serviço esperado.

TEBOUL, (1995) o processo de satisfação não pode ser entendido por Gap e sim conforme a satisfação do cliente de acordo com a figura 2.3. Nesse modelo associam-se as três tarefas acima da linha as interpretações físicas e técnicas, ou seja, estão ligadas aos padrões, procedimentos e regras do serviço, enquanto as duas atitudes abaixo da linha estão ligadas às interações com as pessoas sendo associadas a comportamento e padronização.

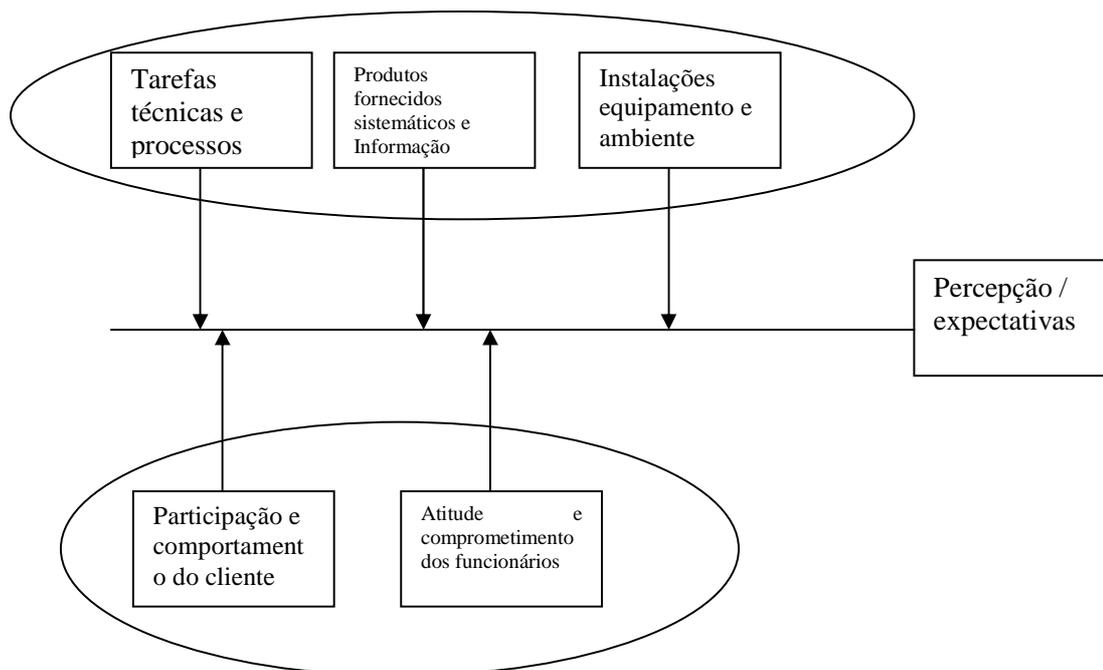


Figura 2.3 Modelo de percepção de satisfação ao cliente.

Fonte: Adaptado de TEBOUL, (1995).

2.1.2.2 Marketing de Serviços

Qualquer uma das características de serviços, em particular a simultaneidade entre o consumo e a fabricação e a impossibilidade de estocagem, identificam a mais importante especificidade de serviços: o fato de que o resultado de um serviço é o próprio processo. Um elemento central do *marketing* de serviços reside no fato que o consumo é antes um consumo de processo que um consumo de resultado. Em serviços, como os processos de prestação e o de consumo são simultâneos e em momentos de interação com os clientes, a abordagem tradicional de *marketing* "mix" não se aplica. Mesmo a tentativa de inclusão de mais um "P", para "pessoas", por JUDD (1987), na reconceituação da função do *marketing* de serviços, ou mesmo para seis "Ps", por TÉBOUL, (1995) onde foram incluídos "participantes" (clientes/empregados) e o "processo de entrega" (momentos de interação), não evitaram que o conceito de *marketing* "mix" caísse em desuso, principalmente para serviços. No triângulo do *marketing* de serviços, os elementos característicos são muito diferentes em relação ao *marketing* de produto, de acordo com a figura 2.4.

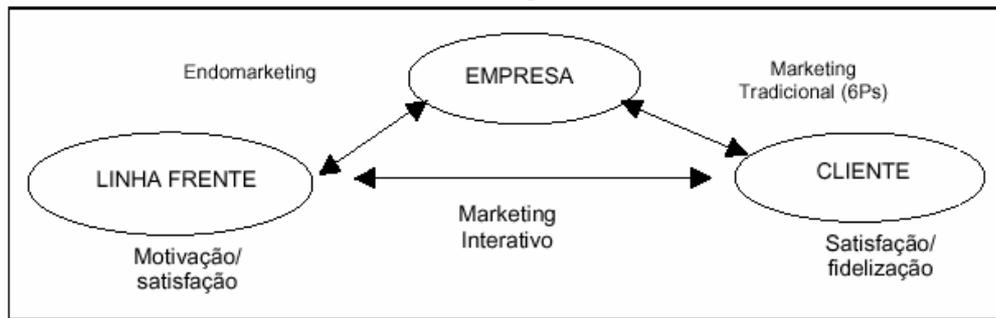


Figura 2.4 *Marketing Integrado em Serviços*.

Fonte: Adaptado de TEBOUL, (1995).

Define-se como uma das características básicas dos serviços, a interação do cliente com o fornecedor de serviços. Logo, uma das primeiras diferenciações em relação à produção de bens é possuir pessoal capacitado e habilitado a interagir com os clientes na resolução de seus problemas e/ou atendimento de suas necessidades.

Nessas situações de interação, o prestador de serviços tem a oportunidade de demonstrar a qualidade de seus serviços, bem como avaliar o comportamento e necessidades dos consumidores, atividades, que conforme GRÖNROOS, (1993) denominadas de *marketing interativo*.

Em uma organização de fabricação de bens, essas ações são normalmente desempenhadas por um departamento específico de *marketing*, sendo que o resto da organização não tem responsabilidade sobre os clientes e o *marketing*.

Diferentemente da manufatura, em organizações de serviços, todos os empregados são pessoas de *marketing* e devem ser treinadas e terem o apoio necessário, para agirem como consultores que estejam preparados para desempenhar suas tarefas e tomar decisões quando o cliente deles necessitar e da forma que os clientes desejarem. As decisões do dia-a-dia devem ser descentralizadas ao nível operacional, ou "*Front - Office*" e o resto da organização, bem como a gerência e pessoal de apoio, ou "*Back - Office*", deve criar condições necessárias para a produção e entrega de serviços em interações com os clientes.

A prestação de serviços é uma atividade exercida em interações entre *Front - Office* o cliente, ou momentos da verdade, nos quais existe um relacionamento e que por natureza é uma função de *marketing*. O que acontece em serviços, é que essas interações são realizadas com os clientes por profissionais, que não são normalmente especialistas de *marketing*. O

empregado de *Front - Office* fornece, controla e eventualmente comercializa a prestação, realizando suas tarefas de forma que o cliente deseje voltar, estabelecendo um relacionamento sólido com o cliente. GUMMESON, (1991) denominou esses empregados de "marqueteiros em tempo integral", identificando que como serviços é um processo de relacionamento, todo funcionário de *Front - Office*, em potencial, realiza o *marketing* diretamente e de acordo com as suas interações com o cliente.

Entretanto, essa função deve ser complementada com a outra função tradicional do *marketing mix*, a qual deve ser coordenada por um setor especializado de *marketing* e que executa tarefas inerentes ao processo como publicidade, promoções e esforços de vendas. Segundo GUMMESON, (1991) os departamentos de *marketing* e vendas – os "marqueteiros a tempo integral" não podem gerir que uma parte limitada do *marketing*, na medida em que suas equipes não podem estar em contato permanente com os consumidores.

Essas duas funções são integradas, em uma terceira função de *marketing*, que se pode definir de *endomarketing* [GRÖNROOS, (1990)], como o gerenciamento do *marketing* interno dos serviços, ou seja, a gestão da motivação, comportamento e das atitudes dos funcionários nos relacionamentos com os clientes. O *endomarketing* identifica a noção de que os empregados constituem o primeiro mercado para as organizações e se os bens ou serviços não conseguirem atingir o mercado interno, o *marketing* pouco fará na satisfação do cliente externo. O objetivo do *endomarketing* nesse sentido é assegurar que os empregados se motivem para uma orientação ao cliente e um desempenho consciente de serviços e, portanto desempenhem com êxito suas atividades de "marqueteiros a tempo parcial" em suas tarefas de *marketing* interativo [BERRY e GRÖNROOS, (1983)].

2.1.2.3 Valor

Satisfação é o nível de sentimento de uma pessoa, resultante da comparação do desempenho de um serviço ou produto em relação as suas expectativas. Um cliente pode em função das suas expectativas e o desempenho percebido experimentar três níveis diferentes de satisfação [KOTLER, (1996)]. Essa avaliação pode ir desde ficar longe das expectativas até exceder suas expectativas passando por atender as expectativas.

Um cliente em que as suas expectativas são atendidas é um cliente meramente satisfeito não voltará necessariamente e não é um agente de *marketing*, porém um cliente em que a empresa tenha excedido as suas expectativas esse sim é um cliente muito satisfeito, esse cliente se tornará um cliente fiel e na maioria das vezes se torna um entusiasmado “vendedor” [CORRÊA, (2002)]. A alta satisfação ou encanto gera uma afinidade emocional com a marca, e isto gera uma grande lealdade dos consumidores [KOTLER, (1996)].

Na visão de satisfazer o cliente às empresas estão pretendendo desenvolver confiança e lealdade mais forte junto a seus consumidores finais, assim exercem uma maior retenção. As empresas atuais estão dedicando mais atenção a suas taxas de abandono, pois essa movimentação envolve custos.

A satisfação do cliente é o critério mais importante para a lealdade, no sentido da recompra de um bem e/ou serviço [HESKETT et al., (1994)], e mais do que isto, é a chave para a garantia para um relacionamento em longo prazo. O modelo tradicional de qualidade prioriza a qualidade percebida pelo cliente para a sua satisfação [PARASURAMAN et al., (1988)].

O fato é que, conforme pesquisas recentes há um criticismo crescente em relação ao modelo da qualidade para a garantia da satisfação do consumidor de serviços [ANDERSON et al., (1994); IACOBUCCI et al., (1994); LILJANDER e STRANDVIK, (1995)]. Nesse modelo, os custos envolvidos na relação cliente/fornecedor, apesar de serem descritos como critérios de avaliação da qualidade não têm a mesma proporção, nem talvez a mesma influência sobre a satisfação dos consumidores, que as demais dimensões (segurança, credibilidade, confiança) de percepção da qualidade, exaustivamente exploradas por [PARASURAMAN et al., (1988)].

Complementando então a qualidade percebida, tem-se o conceito de "valor". Valor é basicamente definido como a relação entre os benefícios e os sacrifícios de um processo de compra de um bem ou serviço. De acordo com esse conceito, sacrifícios ou custos são critérios igualmente importantes na determinação da satisfação do cliente. Os consumidores, na maioria das situações de compra, usam referências de preço e valor na avaliação de uma oferta. Dessa forma, o valor não mede somente a satisfação do cliente em relação aos aspectos da qualidade de um produto ou serviço, mas mede também a satisfação relacionada aos sacrifícios do cliente, na sua aquisição e uso.

Pelo exposto, ao invés de focar-se a garantia da satisfação do cliente pela gestão da qualidade percebida, a idéia é gerir as operações de serviço, como forma de agregar valor, com as quais os benefícios e sacrifícios envolvidos na relação. Isto é gerir o valor percebido pelo cliente.

A) Valor Percebido pelo Cliente (VPC)

Valor é um termo com várias definições e tem sido aplicado de várias formas nas literaturas relacionadas ao gerenciamento, estratégia, finanças, sistemas de informações e *marketing* [WIKSTROM e NORMANN, (1994)]. De acordo com HUBER, HERMANN e MORGAN (2001), "o conceito de valor é multifacetado e possui inúmeras interpretações". Nesse trabalho utiliza-se a definição de [RAVALD e GRÖNROOS, (1996)], a qual o valor é algo percebido pelo cliente, em um processo de aquisição de um bem ou serviço e é a relação imediata entre os benefícios e sacrifícios. Essa definição dá origem à equação geral do valor:

$$\text{VPC} = \text{Benefícios percebidos} / \text{Sacrifícios percebidos}$$

Já ZEITHAML, (1988) define valor percebido pelo cliente como "sendo a taxa de utilidade de um produto e/ou serviço baseado na percepção do que é recebido e o que é dado na relação". De acordo com RAVALD e GRÖNROOS, (1996), se define os benefícios percebidos na compra de um bem ou serviço como sendo uma combinação de atributos físicos, atributos do serviço e suporte técnico disponível ao uso do produto ou serviço. Já os sacrifícios percebidos podem ser definidos como todos os custos que o cliente dispensa com o fornecedor no momento da aquisição: preço de compra, transporte, instalação, reparos e manutenção, riscos de falha ou baixa performance.

Por outro lado, existem muitos incidentes que podem elevar os esforços do cliente para a longevidade da relação, envolvendo custos adicionais, inesperados e desnecessários, e que podem ser chamados de custos indiretos e psicológicos [GRÖNROOS, (1993)]. Custos indiretos são relacionados ao prazo de entrega, a emissão errada de uma fatura, etc. Por sua vez, custos psicológicos estão relacionados aos esforços cognitivos do cliente quanto à incerteza se o fornecedor vai cumprir as promessas preestabelecidas.

O valor percebido pelo cliente deve adquirir um conceito mais profundo, o qual não considera somente o conceito imediato da relação cliente/ fornecedor, mas sim as expectativas do cliente e a responsabilidade da empresa em tratar essas expectativas em um horizonte mais amplo. Isto significa que os fatores ou determinantes (benefícios/sacrifícios), que maximizam o valor percebido pelo cliente, quando se considera a relação cliente/fornecedor como um todo, devem ser investigados. As organizações necessitam entender quais fatores criam valor para os clientes na ordem de construir vantagem competitiva [LICHTENTHAL, et al., (1998)].

Pode-se sintetizar o exposto, conforme esquema da figura 2.5 adaptada de GRÖNROOS, (1993) que se forma a lealdade a longo prazo, na relação cliente/fornecedor, a partir da agregação de valor pela gestão dos benefícios e sacrifícios, desenvolvendo-se uma relação mutuamente vantajosa. Então, é reconhecido que para obter sucesso nas organizações, é preciso satisfazer os clientes, gerando lealdade e a retenção dos mesmos, pela gestão do valor.

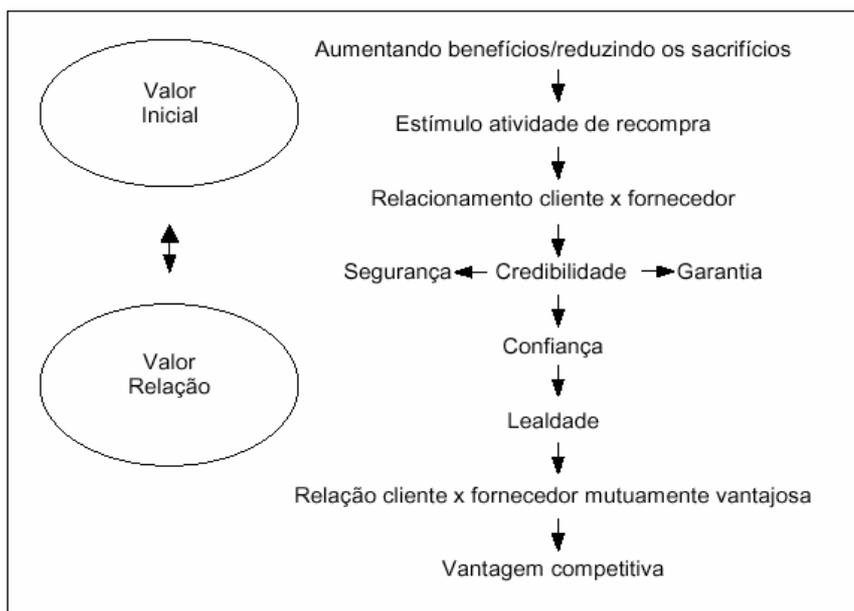


Figura 2.5 Valor como vantagem competitiva na relação cliente x fornecedor.
 Fonte: adaptada de RAVALD E GRONROOS, (1996).

Entretanto, poucas empresas têm o conhecimento e a capacidade de gerir o valor para os seus clientes, como forma de construir alguma vantagem competitiva.

Nesse sentido, o primeiro passo, é tentar responder às seguintes questões básicas, conforme [ANDERSON et al., (1994)]:

- Quais são os fatores influenciadores para a percepção de valor, quando um consumidor escolhe um entre vários fornecedores?
- Como os consumidores comparam esses fatores de percepção do valor entre os diversos fornecedores?
- Qual o grau de importância de cada um desses fatores relativamente ao valor percebido pelo cliente?

B) Fatores Influenciadores do Valor Percebido pelo Cliente

Os fatores formadores da percepção do valor, os quais por definição são os benefícios e os sacrifícios da aquisição de um bem ou serviços, tendem a ser de características altamente pessoais e variarem largamente de um consumidor para outro, especialmente nas operações de serviços [ZEITAHML, (1988)].

De acordo com WALTERS, (1999) esses fatores ou "atributos de valor" são os atributos de um produto ou serviço, os quais constituem a razão primária de seleção, e que agregam valor ao estilo de vida do consumidor. Os "atributos de valor", de acordo com LAPIERRE, (2000) podem ser identificados em categorias genéricas, porém, características específicas dentro dessa classificação geral devem também ser investigadas. Nessa pesquisa, serão adotados genericamente 13 fatores formadores da percepção do valor, os quais foram selecionados e analisados empiricamente por LAPIERRE, (2000) entre 2400 executivos na área de serviços e produtos industriais de tecnologia da informação, no Canadá:

Relacionado ao Produto:

- **Soluções alternativas:** diz respeito à gama de alternativas, à capacidade de atender às expectativas e de resolução de problemas;
- **Qualidade:** a qualidade intrínseca, ou seja, a durabilidade, a confiabilidade e a performance do produto ao longo do tempo;
- **Customização:** relacionado ao uso e o consumidor ao qual o produto foi concebido. Nesse sentido, é a habilidade de expandir o uso e a clientela de um determinado produto.

Relacionado ao Serviço:

- Receptividade: significa prover respostas e soluções rápidas para os problemas;
- Flexibilidade: a flexibilidade relacionada às solicitações dos clientes. Habilidade para se ajustar as necessidades mutantes dos consumidores;
- Confiabilidade: significa a habilidade para cumprir as promessas, e manter boas relações;
- Competência técnica: ter o conhecimento em sua atividade de serviço. Está relacionado com o modo de solucionar problemas, ou seja, a habilidade em prover soluções;
- Imagem: reputação, credibilidade;
- Confiança: sentimento de parceria entre cliente e fornecedor, ou seja, transparência do fornecedor quanto à capacidade de cumprir as promessas;
- Solidariedade: significa a habilidade para assumir e partilhar os problemas e as suas soluções.
- Preço: quantidade monetária a ser dispensada pelo produto e/ou serviço ofertado. Está relacionado ao uso e vantagens do produto ou serviço;
- Tempo/esforço/energia: sacrifício imposto na manutenção da relação;
- Conflito: esforços para resolver situações controversas da relação.

Segundo o estudo de RAVALD e GRÖNROOS, (1996) que considera o valor como uma medida expressa não somente do produto e/ou serviço em questão, mas também do relacionamento cliente/fornecedor estabelecido. A figura 6 sintetiza os treze "atributos de valor" para a percepção do valor.

ESCOPO DOMÍNIO	PRODUTO	SERVIÇO	RELACIONAMENTO
BENEFÍCIOS	Soluções alternativas Qualidade intrínseca Customização	Receptividade Flexibilidade Confiabilidade Competência técnica	Imagem Confiança Solidariedade
SACRIFÍCIOS	Preço	Preço	Tempo/esforço/energia Conflito

Figura 2.6 Fatores de Percepção do Valor.

Fonte: Adaptado de LAPIERRE, (2000).

C) Marketing do Valor

Até recentemente, sob o âmbito de uma economia de subsistência, não havia o menor esforço no sentido da satisfação dos clientes, ou seja, o objetivo era somente a produção em larga escala para alcançar o maior número de pessoas, sem a preocupação efetiva com as suas necessidades.

Ao contrário, em uma economia competitiva, com a rapidez e amplitude da informação e crescimento das operações de serviços, as pessoas podem escolher entre uma vasta gama de bens e serviços. As empresas que querem ganhar, ou mesmo como condição de continuar competindo, têm de adotar uma nova filosofia de ação orientada para o cliente. A diretriz é então construir uma relação cliente/fornecedor duradoura e fiel, e não somente a elaboração de um produto dentro de suas especificações técnicas e/ou o simples fornecimento de um serviço, com o intuito da satisfação imediata. Assim, o conceito de "valor" assume o aspecto mais importante na satisfação dos clientes e serve como subsídio básico para uma relação duradoura com o mercado. KOTLER, (1996) reforça essa afirmação como: "Nossa premissa é que o mercado irá comprar da empresa da qual há a percepção do maior valor criado". O cliente, no limite de seus esforços, de suas informações e mobilidade tende a maximizar a percepção desse valor. No momento que o produto e/ou serviço adquirido atende efetivamente o valor esperado, nasce à satisfação. Na figura 2.7, adaptada de BEST, (1997) mostra-se uma identificação com o processo de criação da satisfação.

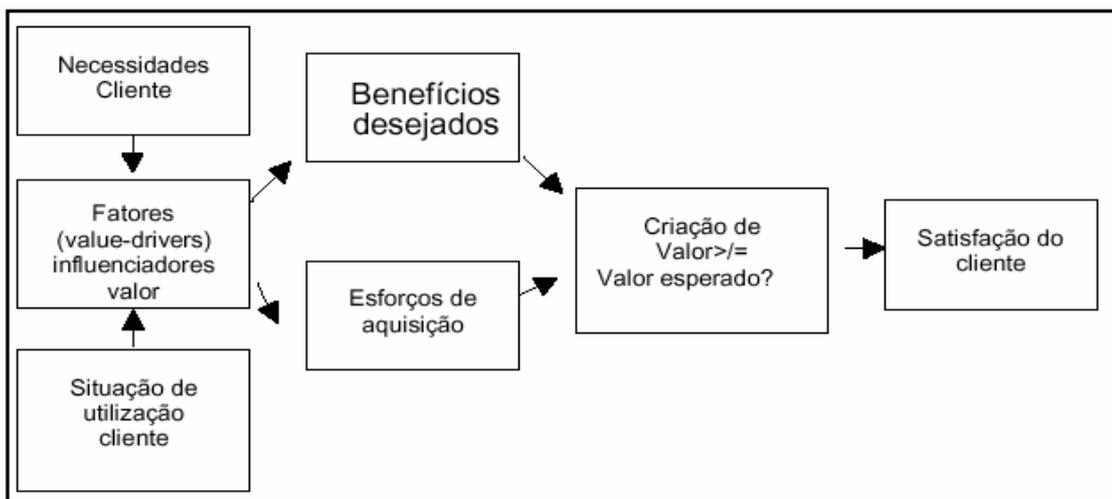


Figura 2.7 O Processo de Criação do Valor e da Satisfação.

Fonte: Adaptado de BEST, (1999).

A partir do exposto, tem-se essa nova filosofia de *marketing*, que é buscar a satisfação em uma relação duradoura com o mercado, pela maximização do valor entregue ao cliente, o "*marketing* do valor".

D) Cadeia de Valor

PORTER, (1985) introduziu o conceito de cadeia de valor para facilitar a identificação dos modos de criação de valor. Toda organização constitui-se de um conjunto de funções para criar, desenvolver ou fabricar e comercializar produtos e/ou serviços. A cadeia de valor define essas funções em nove pólos, os quais servem ao mesmo tempo, de centros de custo e fontes de valor. Essas nove funções estão divididas em atividades essenciais e de apoio.

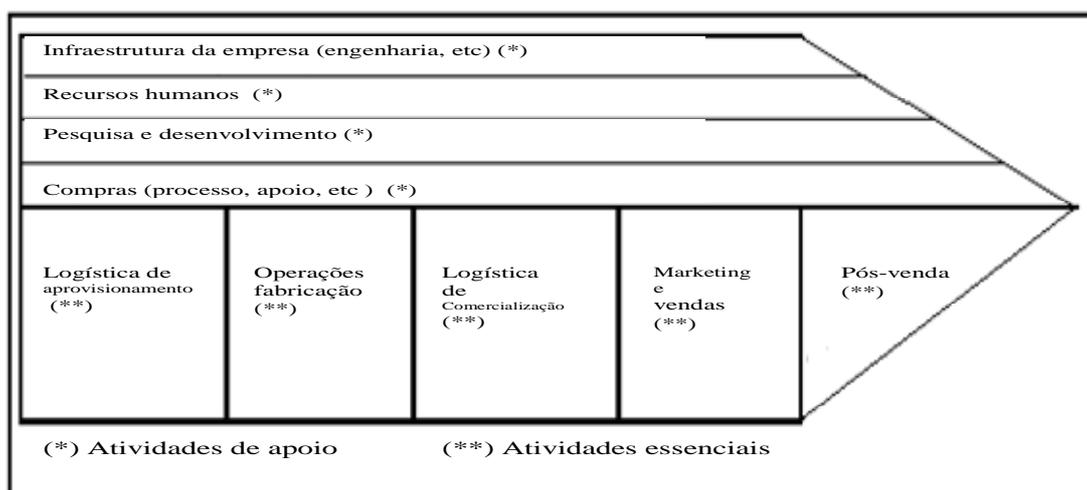


Figura 2.8 A Cadeia de Valor.

Fonte: Adaptado de PORTER, (1985).

As atividades ou funções essenciais ou básicas são articuladas pela seqüência: provisionamento – operações de manufatura ou serviços – comercialização. As atividades ou funções de apoio, por sua vez, dizem respeito àquelas necessárias à sustentação das funções essenciais. Assim, engenharia, manutenção, compras, pesquisa e desenvolvimento são atividades sem as quais é impossível a execução das funções essenciais. O processo de criação de valor se dá pela análise e gestão dos custos e desempenho em todas as suas funções

componentes, a qual é feita comparativamente aos concorrentes potenciais. Porém, o sucesso definitivo desse sistema só se dará se houver uma concatenação e coordenação dos objetivos e missões das diversas funções da cadeia no objetivo maior da satisfação do cliente externo. WALTERS e LANCASTER, (1999) reiteram: "o valor entregue compreende todas as atividades envolvidas nos atributos de fornecimento do produto e/ou serviço, os quais são necessários para criar satisfação e manter uma relação de longo-prazo com os clientes e assim construir vantagem competitiva". WALTERS e LANCASTER, (1999) definem *marketing* em torno do conceito da cadeia de valor.

Marketing é o processo de definir, desenvolver e criar valor. Com esse conceito, a função *marketing* deve não somente ser realizada por um departamento especializado, com as atribuições tradicionais de *marketing*, o "*mix-marketing*", para se tornar uma função a ser exercida por todos os departamentos. Como WALTERS e LANCASTER, (1999) sugerem: "a proposição de valor deve ser o mais importante princípio organizacional", ou seja, que a criação de valor, seja a missão principal da organização e todos os departamentos e funções da organização estejam engajados nessa proposição de valor.

2.1.2.4 Serviços Internos

Pela definição da cadeia de valor PORTER, (1985) a criação de valor e conseqüente a satisfação do mercado, se dará por meio da coordenação das funções internas da organização. Ou seja, o gerenciamento da organização com foco na satisfação do cliente externo, é primeiramente implementado internamente. Isto quer dizer que, para atingir-se a satisfação do cliente externo, é preciso que cada unidade gerencial da organização satisfaça seu cliente direto.

As funções essenciais e de apoio da cadeia de valor podem ser divididas em unidades organizacionais e seus empregados como fornecedores de serviços, os usuários desses serviços como clientes internos, e as atividades definidas como operações de serviços internos. Conforme WALTERS, (1999): "serviços internos podem ser definidos como serviços prestados pelas distintas unidades organizacionais ou por pessoas trabalhando nesses departamentos para outras unidades ou empregados na organização".

Justamente pelo fato de cada unidade intra-organizacional ou fornecedor interno é um micro processo independente, em relação às próprias metas, prioridades e procedimentos, o grande desafio é minimizar nas relações internas essa dificuldade de integração, mútuo contato e troca de informações.

Estabelecendo-se uma cadeia de processos internos como sistema de gestão intra-organizacional, não há a garantia que a prática do gerenciamento orientado para o cliente ocorra sem problemas. STAUSS, (1995) identifica pela correlação com os critérios identificados no MBNQA - *Malcolm Bridge National Quality Awards* – o prêmio nacional americano para a qualidade, as preconizações para a gestão intra-organizacional orientada para o cliente e sua satisfação, conforme figura 2.9:

- Evidência do cliente: o fornecedor deve conhecer qual é o seu cliente;
- Evidência do fornecimento: que os clientes percebam os fornecedores e as suas ofertas;
- Evidência do processo de aquisição do serviço/segmentação: medições das necessidades e expectativas evidenciam o segmento de cliente (s);
- Características especificáveis do serviço interno, assim como as obrigações do fornecedor para com o cliente (s);
- Disposição dos clientes para exporem a sua insatisfação.

CRITÉRIOS "ORIENTAÇÃO CLIENTE/SATISFAÇÃO" MBNQA	PRECONIZAÇÕES UTILIZAÇÃO PARA SERVIÇOS INTERNOS
7.0 Foco no cliente e satisfação 7.1 Expectativas do cliente 7.2 Gestão da relação cliente-fornecedor 7.3 Obrigações para com os clientes 7.4 Determinação da satisfação do cliente 7.5 Resultados da satisfação do cliente 7.6 Comparação cliente perante satisfação	Evidência do cliente Evidência do fornecimento Evidência do processo de aquisição do serviço/segmentação Características especificáveis Disposição dos clientes para exporem a sua insatisfação

Figura 2.9 Preconizações para Utilização Critérios MBNQA da "Orientação/Satisfação do Cliente" em Serviços Internos.

Fonte: Adaptado de STAUSS, (1995).

A - Classificação dos Serviços Internos

A classificação de serviços externos tem sido utilizada amplamente para a descrição e estudo desses tipos de operações LOVELOCK, (1983) enquanto uma classificação específica para serviços internos tem sido um assunto de pesquisa restrita.

Um dos primeiros pesquisadores a trabalhar com esse tema foi Sayles que em 1964, realizou uma extensiva pesquisa baseado em observações, no qual identificou dois tipos fundamentais de relações intra-organizacionais: "processo", "serviços".

Relações de processo envolvem atividades as quais devem ser realizadas por diferentes pessoas em uma seqüência relativamente fixa, de forma que um começa somente depois que o anterior tenha sido completado. Já as relações de serviço são caracterizadas por múltiplas ligações ou "obrigações" para com outras unidades na organização, na qual não é necessariamente o próximo estágio.

Mais recentemente, DAVIS, (1993) combinou os tipos de relações internas definidas por Sayles e os padrões administrativos intra-organizacionais existentes e classificou os serviços internos em: "de processo", "de apoio ou suporte" e "de avaliação ou auditoria".

B - Serviços Internos de Processo

Nos serviços de processo, normalmente há um grande número de operações de serviços rotineiras desenvolvidas segundo procedimentos preestabelecidos, isto é, padronizadas e mais facilmente especificáveis quantitativa e qualitativamente.

Como há transparência quanto às obrigações e à performance para com os clientes internos na cadeia de processos, normalmente não há barreiras sociais nem psicológicas para a exposição da insatisfação.

Em síntese, serviços de processo, os quais podem ser exemplificados como as diversas operações padronizadas de manufatura e de logística de uma atividade industrial, são características identificadas por meio de uma gestão orientada para o cliente com:

- Um alto grau de evidência do cliente e uma clara situação definida de fornecimento e pouca ou quase nenhuma segmentação;
- Uma alta evidência da oferta;

- Características altamente especificáveis, em função da repetibilidade;
- Voluntariedade do consumidor interno para expor a insatisfação.

C - Serviços Internos de Auditoria

Serviços de auditoria envolvem os serviços relacionados às responsabilidades de auditoria e avaliação interdepartamental da organização [DAVIS, (1993)]. Podem também ser classificados segundo as atividades de apoio da cadeia de valor, porém é difícil manter-se uma relação cooperativa, já que os clientes não podem esperar benefícios diretos desse tipo de relação, a menos que o serviço seja solicitado pelo cliente. Podem ser classificadas dentro desse grupo, as atividades de controle da qualidade e controle de gestão (controladoria contábil/fiscal) da organização.

D - Serviços Internos de Apoio ou Suporte

Contrariamente aos serviços de processo, as operações internas de apoio ou suporte não possuem as mesmas boas condições para uma gestão focada no cliente. Especificamente esses tipos de serviços menos padronizados e menos repetitivos são caracterizados por:

- Baixa evidência da existência de clientes: muitos fornecedores de serviços de suporte não estão envolvidos com os mesmos clientes todo o tempo, não criando uma relação cliente/fornecedor estável;
- Existência de diferentes segmentos de consumidores: os fornecedores de suporte atendem a diversos segmentos de clientes, os quais variam fortemente em expectativas e percepções;
- Baixa evidência do fornecimento: freqüentemente essa evidência é limitada porque os clientes internos não reconhecem que para um determinado problema, um serviço específico pode ser a solução;
- Dificuldade na especificação das características por não serem processos repetitivos e padronizados;

- Apatia para expressar a satisfação ou insatisfação: o cliente pode se sentir coagido ao expor sua insatisfação, por não ser um serviço repetitivo, em função das conseqüências que isto possa trazer para o ambiente interno à organização.

Assim, conforme evidenciado, os serviços internos de suporte, ou de operações menos padronizáveis que as de processo, têm uma dificuldade maior quanto à aplicação da gestão orientada para o cliente. De acordo com a cadeia de valor de PORTER, (1985) nesse tipo de serviços estão relacionadas às atividades de apoio às atividades essenciais de uma organização, tais como as operações de manutenção, recursos humanos, compras, engenharia, etc.

2.1.3 Indicadores de Qualidade e Medidas de Qualidade

Destaca-se a importância das medidas de qualidade, tendo em vista o papel decisivo por elas assumido em face dos processos de globalização, da abertura dos mercados e da conseqüente competição entre organizações [CORRÊA, (2002)].

Os indicadores de qualidade e as medidas que fazem parte do universo da gestão da qualidade possibilitam o incessante aprimoramento das prestadoras de serviço, que a todo instante são impelidas alterar suas sistemáticas e procedimentos na tentativa de atender a satisfação desejada do cliente [FITZSIMMONS, (2004)].

Atualmente, a gestão de qualidade abrange uma visão macro das prestadoras de serviço, influenciando com isto o modo de pensar e de agir. Qualidade não significa apenas o controle do serviço prestado, a qualidade intrínseca associada a serviços, essa ligada ao uso de ferramentas e métodos de gestão. Num sentido mais amplo, o conceito de qualidade passou a significar modelo de gerenciamento que busca a eficiência para a total satisfação das necessidades dos clientes.

2.1.3.1 Histórico dos Indicadores de Qualidade

O tema gestão de qualidade é dinâmico, sendo sua evolução fruto de interação dos diversos fatores que compõem a estrutura organizacional e sua administração. Há 300 anos no início da primeira revolução industrial observou-se a primazia do setor têxtil, pioneiro na

mecanização. Logo se estabeleceu um novo paradigma de qualidade, baseado no modo de produção mecânico. O conhecimento começou a ser introduzido na vida industrial e novas técnicas de trabalho eram exigidas para fazer frente ao notável crescimento observado nas empresas e a crescente demanda dos mercados em expansão.

Foi no início do século XX, com a segunda revolução industrial já consolidada, que se firmou o campo do conhecimento da administração e foram delineados os primeiros modelos de gestão racionalmente construídos, com o objetivo de dotar as organizações de maior eficiência produtiva. Foi nessa época que se observou a primeira onda do Controle Estatístico de Processo CEP, iniciando o período do controle de qualidade. Na *Bell System*, com profissionais vindos da *Bell Telephone*, o uso de alguns métodos estatísticos para controlar a qualidade na produção de itens de telefonia. O mercado de telecomunicações nessa época tinha apenas a visão da indústria sem a visão moderna de uma empresa de serviço.

Shewhart criou e desenvolveu o ciclo PDCA (*Plan, Do, Check, Act*), difundido por Deming anos depois. Embora o principal interesse de Shewhart fosse à metodologia estatística ele estava consciente a respeito dos princípios de gerenciamento da qualidade.

A segunda Guerra acelerou o passo da tecnologia da qualidade e nesse ambiente. Os conceitos básicos de controle da qualidade se expandiram rapidamente. Sendo nessa época o desenvolvimento de técnicas de análise de falhas para resolver problemas.

Em 1950, Deming um estatístico que trabalhou na *Bell System*, foi ao Japão, onde proferiu palestras para líderes industriais.

Já Juran visitou o Japão em 1954, introduzindo uma nova era no controle da qualidade, liderando a passagem de uma fase, na qual as atividades relativas à qualidade baseavam-se nos aspectos tecnológicos das fábricas, para uma nova, na qual a preocupação com a qualidade passou a ser global e holística, abarcando todos os aspectos do gerenciamento da organização.

A partir do início da década de 80, começou a ficar evidente que as crescentes exigências do mercado, nos aspectos mercadológicos e de qualidade aliada a uma maior preocupação ecológica gerando um novo conceito da qualidade, orientada também a qualidade de vida.

Hoje a qualidade essa relacionada às necessidades e aos anseios dos clientes. Seja qual for o porte da empresa, observam-se programa de qualidade e de melhoria de processos em todos os setores da organização.

2.1.3.2 Medidas de Qualidade

Os avanços nas últimas duas décadas transformaram os conceitos tradicionais sobre prestação de serviço em conceitos obsoletos. As empresas não conseguem mais vantagens competitivas sustentáveis apenas com a rápida locação de ativos físicos e novas tecnologias. É necessário entender e gerenciar eficientemente todos os recursos [KAPLAN & NORTON, (1996)].

Para garantir a eficiência na gestão necessita-se obter uma base de medidas de Qualidade de Serviço (Qos). Essas medidas devem ser usadas para avaliar todos os tipos de serviços. Nesse contexto está incluso os serviços de telecomunicações nas suas diversas áreas (telefonia fixa, dados e telefonia móvel). Procura-se encontrar indicadores que sejam cada vez mais operacionais e significativamente úteis aos responsáveis pelas decisões. Eles precisam refletir a forma como o cliente enxerga o serviço prestado pela empresa. Os conceitos a seguir mostram duas percepções de medidas. A primeira avaliação evidencia medidas quantitativas dos serviços de manutenção que afeta a percepção do usuário. A segunda avaliação por meio da taxa de abandono do cliente na planta, que refleti diretamente a insatisfação dos usuários e a forma como suas expectativas não foram atendidas com um serviço em particular.

Nos últimos anos, diversos trabalhos têm considerado a importância da mensuração da qualidade por clientes externos às organizações. No entanto, qualquer organização industrial, não prescinde da avaliação de desempenho dos serviços internos como um meio para a melhoria contínua dos seus processos. Por isso, algumas pesquisas têm sugerido a utilização de instrumentos geralmente utilizados para mensuração da qualidade por clientes externos na avaliação dos serviços [SALOMI et al., (2005)].

A avaliação da qualidade dos serviços é um fator abstrato e complexo, em função das suas características de intangibilidade, heterogeneidade e da inseparabilidade da produção e consumo [MARGEM, (2005)]. Ao contrário da manufatura, onde a qualidade dos produtos pode ser avaliada de forma objetiva. A qualidade dos serviços é avaliada por meio da comparação entre as expectativas prévias do cliente e sua real percepção, o que se denomina qualidade percebida dos serviços.

À medida que a competição entre as organizações aumenta, o cliente torna-se cada vez mais exigente e crítico em relação aos serviços prestados [FREITAS, (2005)]. Ou seja, os padrões de qualidade de atendimento estabelecidos pelo mercado estão cada vez mais severos. Devido às características intrínsecas aos serviços, um modelo de gerenciamento visando a excelência em serviços pode ser composto por três elementos fortemente relacionados, no qual o cliente é considerado como peça fundamental para o sucesso empresarial:

- **Foco do serviço:** consiste nos procedimentos e ações que devem ser planejados e criados para que retratem os reais desejos e necessidades do cliente;
- **Funcionários:** são todas as pessoas pertencentes à organização que, uma vez conhecido o foco do serviço, devem estar preparados para realizá-los;
- **Suporte técnico:** nesse aspecto estão englobados sistemas, técnicas e procedimentos que auxiliam a prestação dos serviços.

Saber fazer e adaptar *benchmarking* no processo da organização pode nos permitir vislumbrar oportunidades e também ameaças competitivas, constituindo um atalho seguro para a excelência, com a utilização de todo um trabalho intelectual acumulado por outras organizações evitando os erros e armadilhas do caminho [MORAIS e GODOY, (2005)]. Reduzindo-se o número de falhas potencialmente prejudiciais ao bom andamento dos serviços, mediante uma atuação pro ativa junto as suas principais causas, estar-se-á diminuindo o número de reprogramações e, por conseguinte, reduzindo os custos. Salienta-se, que em alguns setores, ainda prevalece à indiferença do administrador em relação ao cidadão (usuário de serviço) e que a única forma de se tentar resgatar o papel do administrador em seu exercício do poder voltado para a cidadania, é a qualidade nos serviços.

Após essas definições de medidas e com a finalidade de sempre melhorar o serviço prestado ao cliente, as empresas passaram a questionar:

- O que poderá ser feito para melhorar a satisfação do usuário?
- Os usuários podem encontrar alguma mudança na qualidade intrínseca procurando ver o valor do serviço prestado?

De uma forma geral, os indicadores têm a função de dar subsídios aos especialistas para responder e realizar ação em prol de resolver essas questões.

2.1.3.3 A Finalidade de Medir

O que é sugerido, é que os indicadores de qualidade são medidas cujo objetivo final é produzir respostas específicas às perguntas específicas. Esse ponto de vista é predicado na afirmação de HARDY, (2001) no qual a única boa razão para medir qualquer coisa é reduzir a incerteza com respeito a algum curso de ação que deve ser decidido. Entretanto, tudo que é realmente possível conseguir com indicadores de qualidade é uma avaliação do momento, ou seja, pode ser interpretado o comportamento do cliente naquele instante. A análise desse momento pode ser útil apenas quando o resultado ajudar a alguém a se sentir mais confortável sobre alguma decisão a respeito de o que, como e quando fazer determinadas ações, tal como a quantidade de preventivas que devem ser executadas por mês, ou quantos técnicos serão necessários para diminuir o nosso tempo de reparo, como corrigir um problema reconhecido, como reconhecer que um problema emergiu.

Nesse contexto MANTOVANI, (2001) menciona os serviços fornecidos por empresas de telecomunicações que atendem exclusivamente a outras empresas, aos consumidores finais e ao auto-serviço. Por exemplo, uma empresa de telecomunicações com diversas filiais possui um sistema de informações interno suprido pelo seu próprio serviço. Os serviços de uma prestadora de telecomunicações atendem outras empresas, tal qual, uma rede de dados de um banco de âmbito nacional. E, para o consumidor final, tem-se o exemplo de um usuário de telefonia realizando uma ligação telefônica a outro usuário.

Os avanços nas telecomunicações estão ocorrendo paralelamente ao advento dos computadores pessoais, a introdução e a disseminação da tecnologia de transmissões via satélite foi o primeiro passo no encurtamento das distâncias [GOUVEIA e JÚNIOR, (2004)]. O termo “em tempo real” define exatamente a taxa de transmissão com que as informações circulam pelo globo. Também a utilização de fibras óticas revolucionou os sistemas de comunicação. Esses cabos possuem uma capacidade muito maior de transmissão de dados, proporcionando, além do avanço nas telefônicas fixa e móvel, as conexões de Internet de alta taxa de transmissão, a chamada “banda larga”. Esse potencial, quando amplamente disseminado num futuro próximo, vai permitir a criação de novos negócios, explorando a alta performance do acesso dos usuários à rede.

Para NEELY et al., (1997) medição de desempenho pode ser compreendida como a técnica usada para quantificar a eficiência e a eficácia das atividades de negócio. A eficiência vai tratar da relação entre utilização econômica dos recursos, levando em consideração um determinado nível de satisfação. Por sua vez, a eficácia avalia o resultado de um processo onde as expectativas dos diversos clientes são ou não atendidas.

A medição de desempenho é apenas um dos aspectos que suportam a melhoria contínua sendo destacada em função de seu forte poder retórico como parte da linguagem de melhoria contínua para induzir os membros da organização à ação [ATTADIA e MARTINS, (2003)]. No entanto afirmam que a medição de desempenho não apóia completamente o processo de melhoria contínua. Tal falta de sincronia entre as duas práticas está relacionada principalmente pelo pouco conhecimento das pessoas sobre as habilidades e competências da melhoria contínua no uso da medição de desempenho.

Já SÓRIA, (2005) percebe que a melhoria contínua envolve vários componentes estratégicos da organização e para que haja uma melhoria efetiva é necessário que a organização esteja bem estruturada em relação ao desdobramento das suas estratégias e ao sistema de medição de desempenho, ou seja, seus indicadores de qualidade.

A importância da Medição de Desempenho nas organizações é incontestável atualmente [BUOSI, et al., (2003)]. Diversas metodologias, novas propostas e publicações têm surgido. Porém observa-se que as falhas no desenvolvimento de sistemas de medição de desempenho ainda são altas. Diversos aspectos relacionados a esse processo necessitam ser mais explorados, tanto teoricamente como na prática.

Nesse contexto FIGUEIREDO, et al. (2005) define nove atributos desejáveis e sessenta e três variáveis observáveis para auxiliar a auto-avaliação de Sistema de Medição de Desempenho Organizacional. Os atributos Desejáveis foram: Aprendizado Organizacional, Análise Crítica, Balanceamento, Clareza, Dinamismo, Integração, Alinhamento, Relacionamento Causal e Participação. O Atributo desejável Dinamismo foi decomposto em três sub-atributos: Agilidade, Flexibilidade e Monitoramento.

Um dos sistemas de medição de desempenho que teve mais repercussão e sucesso em suas implantação foi o “*Balanced Scorecard*” (BSC) proposto por KAPLAN & NORTON, (1996) devido a dois motivos: o uso de indicadores não-financeiros complementando os financeiros e pelo grande vínculo do sistema à estratégia empresarial. O BSC demonstra-se

muito adequado como um sistema de medição de desempenho para apoiar um processo de melhoria contínua, pois fornece as informações necessárias e mantém o alinhamento aos objetivos estratégicos da organização.

A medição de desempenho é um conceito multidisciplinar, trabalhado por diversas áreas do conhecimento tais como: Engenharia, Administração, Economia, entre outras. Sua operacionalização ocorre por meio de indicadores de desempenho, os quais buscam quantificar o desempenho do objeto de estudo [KIYAN, (2001)].

Os indicadores de desempenho interferem no comportamento dentro da organização. Ou seja, pessoas modificam seu comportamento na tentativa de assegurar um resultado positivo da medição, mesmo que isso resulte em inadequados cursos de ação. Assim, os indicadores de desempenho devem ser escolhidos de forma que induzam há comportamentos planejados [GEROLAMO *et al.*, (2002)]. Esse é um ponto fundamental dos indicadores de desempenho, porque se eles forem escolhidos de forma errônea podem provocar reações inapropriadas dentro da organização. Por outro lado, se bem definidos, estarão impulsionando a organização na direção de seus objetivos estratégicos e na antecipação de possíveis problemas.

Um Sistema de Indicadores de Desempenho não controla a performance da organização como um todo, mas, quando usado corretamente, é uma importante ferramenta do Gerenciamento de Processos na busca pelo aperfeiçoamento [GEROLAMO *et al.*, (2002)]. O controle é a dimensão central do gerenciamento, mas possui uma grande gama de estratégias a serem escolhidas para uso. O crítico é saber escolher qual estratégia em particular se encaixa numa determinada situação e num estilo gerencial, tendo em mente que quanto melhor a estratégia combinar com o estilo gerencial melhor para a organização. A busca pelo aprimoramento de processos com o uso dos indicadores de desempenho irá, portanto, não só atender a uma determinada necessidade premente da organização, definida por uma estratégia de ação, como também em paralelo, aprimorar o clima organizacional, sendo esse, portanto, outro importante “papel” dos Indicadores de Desempenho.

É necessário contar com um Sistema de Gestão muito bem estruturado e ancorado firmemente em mecanismos de medição e avaliação capazes de fornecer, a qualquer momento, o correto diagnóstico, com indicadores que expressem a realidade de maneira didática e possibilitem a escolha dos caminhos estratégicos a percorrer para passar da situação atual

(“Onde estamos”) para a posição desejada (“Onde queremos chegar”). Essa análise crítica, vital para qualquer atividade empresarial, ganha maior importância no ambiente da Manutenção, onde as dificuldades são ainda mais presentes e contundentes [CARVALHO *et al.*, (2004)]. Ao se posicionar como principal avalista da competitividade, produtividade, lucratividade e sustentabilidade empresariais, a Manutenção chama para si a responsabilidade de se (auto) avaliar constantemente, para saber se seu desempenho continua a contribuir para o alcance daqueles objetivos, necessitando, dessa forma, de instrumentos que aporem índices e resultados fidedignos, direcionados para a melhoria contínua e a excelência.

A importância de se fazer medidas, consiste em aferir as possíveis análises da qualidade de serviços de telecomunicações [HARDY, (2001)]. Como descrito abaixo, há pelo menos cinco classes distintas de responsáveis pelas decisões para ações de eficiência, depende da confiabilidade das informações pode ser verificada uma provável percepção do usuário Qos, e a avaliação das medidas necessária para fazer os resultados dos indicadores, úteis aos especialistas da empresa em tomadas de decisões.

- Usuários de serviço. A classe mais óbvia compreende os usuários reais do serviço, que está testando continuamente sua qualidade.
- Representantes do usuário. As empresas cliente de serviços de telecomunicações designam pessoas responsáveis para selecionar serviços e características de telecomunicações, escolhendo entre empresas fornecedoras de serviços de banda larga, negociar preços e qualidade dos serviços, esses técnicos em informática, gerentes de Tecnologia da Informação, estão sempre em contato com os *call center*, sendo os mesmos responsáveis pela interface da empresa.
- Vendas, fornecedor de serviço e pessoal de *marketing*. No outro lado, um dos consumidores principais de análises de Qualidade de serviço é o pessoal de vendas e *marketing*, que não é necessariamente responsáveis pelas decisões, mas deve responder aos interesses da qualidade do serviço prestado levantando informação junto aos usuários.
- Prestadores de serviços de manutenção. Os prestadores de serviços podem geralmente mostrar as expectativas do cliente em relação aos indicadores de Qualidade de serviço, são as prestadoras de serviço as responsáveis pelo

pessoal da manutenção, que desempenham a operação no cotidiano dos sistemas, ou seja, a manutenção e operação dos serviços de banda larga que é entregue aos clientes, essas operadoras asseguram que a qualidade de serviço é mantida em níveis aceitáveis.

- Arquitetos e coordenadores do sistema. Último em nossa lista de consumidores possíveis de Qos. Essas pessoas são as que devem tomar as decisões a respeito da tecnologia a ser empregada em executar vários serviços de telecomunicações, e recursos de como devem ser configuradas e entregues os serviços de banda larga.

A qualidade é resultante da integração das diversas visões de qualidade, sejam elas da produção, do usuário, transcendentais, de produto ou de valor [BACARIM *et al.*, (2005)]. E essa integração deve ocorrer pelo comprometimento das diversas gerências e áreas de uma empresa em atingir esses níveis de excelência em qualidade. Somente com essa integração de conceitos, as empresas de telecomunicações conseguirão realmente prover à sociedade um produto ou serviço visto por todos os clientes e segmentos como algo que realmente agrega qualidade diferenciada.

Seguindo a vertente de BACARIM, (2005) e mesclando com as cinco classes distintas de HARDY, (2001) entende-se o mercado de telecomunicações de Internet banda larga como sendo um mercado segmentado, e dessa segmentação pode-se destacar clientes alto valor para atender a demanda de representante do serviço e os demais clientes para atender aos usuários do serviço. Para mapear as necessidades de um sistema de gestão bem estruturado e ancorado em mecanismos de medição e avaliação capazes de fornecer, um correto diagnóstico, coleta-se indicadores de serviço de instalação e reparo diariamente, esses indicadores são utilizados para sedimentar as análises feitas com relação à qualidade prestada pelo serviço de Internet banda larga.

2.1.4 Indicadores de Gestão

A abordagem aqui utilizada para os indicadores de gestão sugere que os mesmos sejam definidos, como parte integrante do modelo de gestão do negócio, podendo assim estruturar-se

como uma relação entre duas variáveis, na forma de numerador e denominador, em que seus atributos e valores são factíveis de medição.

Segundo TAKESHY et al., (2003) conceitualmente pode-se dizer que um modelo de gestão depende das medições, informações e análises. As medições precisam ser decorrentes da estratégia da organização, abrangendo os principais processos, bem como seus resultados. As informações necessárias para avaliação e melhoria do desempenho incluem, entre outras, as relacionadas com o cliente, o desempenho de produtos, as operações, o mercado, as comparações com as concorrentes ou referências de excelência, os fornecedores, os colaboradores e os aspectos de custo e financeiro.

Os indicadores de gestão da qualidade são utilizados com propósito de selecionar os objetivos e métricas relacionadas à medição da eficácia de uma organização, sendo importantes para que as decisões sejam tomadas de uma forma científica, eles geralmente representam uma relação matemática na qual mensura atributos de um processo ou dos resultados empresariais, com o objetivo de comparar os valores advindo de eventos reais com as metas preestabelecidas. Características descritivas dos indicadores de gestão: É uma relação matemática que resulta em medida quantitativa, Identifica-se um estado do processo ou resultado desse e Associa-se a metas numéricas pré-estabelecidas.

O estado-da-arte da gestão pode ser entendido como um modelo em que um de seus elementos estruturais é a chamada inferência científica. Esse elemento responde pela maneira como as decisões acontecem nas organizações, as quais são baseadas em fatos, dados e informações quantitativas [TAKESHY et al., (2003)].

Partindo dessas premissas também é importante levar-se em conta que os indicadores podem ser divididos em três níveis de abrangência: Indicadores de Negócio, Indicadores de Desempenho Global e Indicadores de Qualidade.

Os indicadores de negócio destinam-se a avaliar a organização como uma entidade fornecedora de produtos e prestadora de serviços a seus clientes atuais e potenciais, por meio de mensuração dos parâmetros estratégicos, principalmente em seu processo de interação como o meio ambiente.

Na Tabela 2.2 os indicadores de negócio de uma prestadora de serviço podem ser identificados como a forma que a empresa mede seus resultados e processos: serviços de compra, contábeis, financeiros, crescimento e taxa de abandono.

Tabela 2.2 Indicadores de negócio.

	Indicadores	Formula dos Indicadores
1	Ganho	$\frac{(\text{Quant de Instalações} - \text{Quant de Retiradas}) \times 100}{\text{Base em Serviço do Mês Anterior}}$
2	Índice do Resultado Financeiro	$\frac{\text{Receita}}{\text{Custo Despesas}}$
3	Índice de Produtividade Empregado	$\frac{\text{Somatório Valores Faturado}}{\text{Quantidade de empregados}}$
4	Índice de Satisfação dos Empregados	Pesquisa de Clima Organizacional
5	Taxa de Abandono	$\frac{\text{Quant de Retiradas} \times 100}{\text{Base em Serviço do Mês Anterior}}$
6	Retiradas	Quantidade total de retiradas efetuadas no mês em questão.

Fonte: Adaptado pelo autor a partir de TACHIZAWA, (2003).

Já os indicadores de desempenho global visam avaliar o desempenho de toda a organização, embora também possam ser utilizados para a avaliação de clientes institucionais. Sendo bastante utilizados pelo corpo gerencial da empresa.

Os indicadores de qualidade e de desempenho que por sua vez, destinam-se a avaliar a qualidade relativa a cada processo ou tarefa a ele associado. São indicadores que normalmente, medem qualidade na entrega de serviços, bem como na manutenção dos serviços, portanto, voltados e relacionados a clientes.

A Tabela 2.3 ilustra os indicadores de qualidade com as suas respectivas formulas mostrando como se calcular cada indicador.

Tabela 2.3 Indicadores de Qualidade e Desempenho.

	Indicadores	Formula dos Indicadores
1	Garantia de Instalação	$\frac{\text{Reparos Aberto nos últimos 30 dias} \times 100}{\text{Total de Instalações Realizadas nos Últimos 30 dias}}$
2	Instalação no Prazo	$\frac{\text{Instalações Executada no prazo} \times 100}{\text{Total de Instalação no Período}}$
3	Taxa de Preventivas	$\frac{\sum \text{de todos os reparos preventivos realizadas}}{\text{Total de Reparos no Mês}}$
4	Reparo no Prazo	$\frac{\text{Reparos Executados no Prazo} \times 100}{\text{Total de Reparos no Mês}}$
5	Reparos Reincidentes	$\frac{\text{Reparos Reincidentes} \times 100}{\text{Total de Reparos no Mês}}$
6	Taxa de Reparo	$\frac{\text{Quantidade Reparos em 30 dias} \times 100}{\text{Base em Serviço Mês}}$
7	Tempo Médio de Instalação	$\frac{\sum \text{de todos os tempos de Instalação no mês}}{\text{Base de Terminais instalados no mês}}$
8	Tempo Médio de Reparo	$\frac{\sum \text{de todos os tempos de Reparo no mês}}{\text{Base de Terminais Reclamados no mês}}$

Fonte: Adaptado pelo autor a partir de TACHIZAWA, (2003).

Os indicadores de qualidade comumente refletem a relação de produtos serviços/insumos, ou seja, buscam medir a eficiência de um processo ou operação em relação à utilização de um recurso ou insumo específico, devendo existir à medida que estes são necessário ao controle de qualidade e do desempenho no âmbito do processo/tarefa.

Os indicadores de qualidade são indicadores que buscam relacionar a percepção do cliente, em relação ao produto ou serviço recebido, ao grau de expectativa do mesmo em relação a esse produto ou serviço [TAKESHY et al., (2003)]. Ou seja, o indicador mede o grau de satisfação do cliente em relação a determinado produto adquirido ou serviço executado.

Qualidade é o grau de excelência a um preço aceitável e o controle da variabilidade a um custo aceitável [GAVIN, (1998)].

Os indicadores de desempenho de um processo são os índices numéricos estabelecidos sobre as principais causas, que afetam determinados indicadores de qualidade.

2.2 Serviços de Manutenção em Telecomunicações

Parece óbvio que a tecnologia essa mudando de forma acelerada o mundo, o século XX foi marcado por profundas transformações na base material da sociedade. A globalização aliada a rápidas e profundas transformações nos sistemas produtivos tem levado às indústrias a preocupação, cada vez maior, com o aumento da produção, redução de lucros e ganhos de qualidade e produtividade. Nessa busca acirrada, a manutenção, identificada como um meio eficaz de assegurar sua competitividade e sobrevivência vem recebendo ao longo das cinco últimas décadas, maior investimento em tecnologia, treinamento do pessoal técnico e desenvolvimento de novas técnicas de gerenciamento.

2.2.1 O Setor de Telecomunicações no Brasil

Os princípios sobre os quais se baseia o telefone foram descobertos pelo escocês Alexandre Gram Bell, em 1875 [ALENCAR, (1998)]. Em pouco mais de um século, a telefonia evoluiu do primitivo equipamento com que Bell fez suas experiências básicas, até os atuais aparelhos com comunicação por meio de voz sobre IP. Porém a comunicação entre os homens vem de bem antes de Gram Bell conforme anexo IV.

No Brasil a telefonia foi trazida pelo Dom Pedro II, mecenas incentivador das artes e das ciências, em 15 de novembro de 1879 surgia no Rio de Janeiro o primeiro telefone, construído para D. Pedro II nas oficinas da *Western and Brazilian Telegraph Company*, sendo instalado no Palácio de São Cristóvão, na Quinta da Boa Vista, hoje, Museu Nacional.

Em 1883, o Rio de Janeiro já possuía cinco estações de 1000 assinantes cada uma e, ao terminar o ano, estava pronta a primeira linha interurbana ligando o Rio de Janeiro a Petrópolis. Em 13 de outubro de 1888, foi criada a empresa *Telephone Company of Brazil*, com capital de US\$ 300 mil, integralizado por 3 mil ações de US\$ 100. Porém apenas em 1910 foi inaugurado o primeiro cabo submarino para ligações nacionais entre Rio de Janeiro e Niterói. Nessa época, 75% dos telefones instalados no país pertenciam ao Rio de Janeiro e São Paulo *Telephone Company* e, o restante, distribuía-se por outras 50 empresas menores. Em 11

de janeiro de 1923 a empresa Rio de Janeiro e São Paulo *Telephone Company* passou a denominar-se *Brazilian Telephone Company*, facultada a tradução do nome para o português em 28 de novembro de 1923, a *Brazilian Telephone Company* passou a denominar-se Companhia Telefônica Brasileira - CTB.

Foi instalada, em São Paulo, a primeira central automática do País, que dispensava o auxílio da telefonista. Em 28 de janeiro de 1932 foram inaugurados os circuitos rádio telefônicos Rio de Janeiro para Buenos Aires, Rio de Janeiro para Nova York e Rio de Janeiro para Madri. Em 27 de julho de 1939 chegou-se no País a marca de 200.000 telefones automáticos instalados pela Companhia Telefônica Brasileira - CTB. Em 28 de novembro de 1956 foi nacionalizada a CTB fixando sua sede no Rio de Janeiro, com serviços extensivos a São Paulo. Introduzido o sistema de micro-ondas e de Discagem Direta a Distância - DDD.

Em 27 de agosto 1962 foi criado o Código de Telecomunicações do Brasil através da Lei 4.117. A Empresa Brasileira de Telecomunicações - Embratel foi criada com a tarefa de construir o Sistema Nacional de Telecomunicação, e explorar as telecomunicações de âmbito internacional e interestadual, empregando recursos provenientes do Fundo Nacional de Telecomunicações, constituído de uma taxa de 30% sobre as tarifas telefônicas. Em 28 de fevereiro de 1967 foi criado o Ministério das Comunicações, tendo como patrono o Marechal Cândido Mariano da Silva Rondon. A necessidade de se estabelecer padrões de qualidade para as linhas telefônicas no País levou o ministério das comunicações, no final da década de 60, a formar um grupo constituído por engenheiros da Embratel e CTB e com coordenação da secretaria geral do ministério. A dificuldade ou, ate mesmo, a impossibilidade do estabelecimento de comunicações telefônicas com a necessária inteligibilidade, em ligações locais, interurbanas ou internacionais não só pelos baixos níveis dos sinais recebidos, como também por fatores adversos que degradam as ligações, contribuindo para a diminuição da qualidade da transmissão, dentre os quais: ruído, eco, instabilidade e distorções dos sinais, isto exigia a criação de normas e processos de qualidade para regulamentar o setor.

Com a criação da Telebrás, o sistema de discagem direta à distância (DDD) começou a ser introduzido no País. Aumentando com isto as exigências dos clientes em relação à qualidade das ligações. Esse sistema foi o grande responsável pelo desenvolvimento das telecomunicações até a segunda metade da década de 90.

A partir da Lei 9472 de julho de 1997, dispõe sobre a organização dos serviços de telecomunicações, a criação e funcionamento de um órgão regulador e outros aspectos institucionais. A lei geral das telecomunicações foi aprovada em votação na Câmara dos Deputados em 18/06/1997 e no Senado Federal em 10/07/97. No dia 16 de julho de 1997 a lei foi sancionada pelo Presidente da República Fernando Henrique Cardoso.

Em 1998, o ministério das comunicações decidiu dividir a Telebrás em doze companhias: três empresas concessionárias regionais da telefonia fixa, uma empresa operadora de longa distância e oito empresas concessionárias de telefonia móvel banda A. A agência nacional de telecomunicações (ANATEL) foi criada como uma autarquia especial, sendo administrativamente independente, financeiramente autônoma, não se subordinando hierarquicamente a nenhum órgão de governo. Sua missão é promover o desenvolvimento das telecomunicações no País de modo a dotá-lo de uma moderna e eficiente infra-estrutura de telecomunicações, capaz de oferecer a sociedade serviços adequados, diversificados e a preços competitivos como os valores mundiais. Nesse contexto é importante ter uma visibilidade do crescimento econômico brasileiro nos últimos quatro anos conforme Tabela 2.4.

Tabelas 2.4 Estatísticas gerais do Brasil como população, PIB e pesquisas do IBGE (PNAS, PAS) *Ranking* internacionais e estatísticas de serviços de telecomunicações no Brasil.

Brasil	2003	2004	2005	2006
População IBGE	178.985.306	181.586.030	184.007.699	186.770.562
Área total	8.514.215,3 km ²			
PIB.	1.699.948	1.941.498	2.147.944	2.322.818
PIB (Crescimento anual).	1,1%	5,7%	2,9%	3,7%
PIB per capita	R\$ 9.500	R\$10.692	R\$ 11.660	R\$ 12.440
PIB per capita (Crescimento anual real)	-0,3%	4,2%	1,5%	2,3%
PIB (US\$ Bilhões)	507	604	796	1101
Telefones Fixos em Operação (Milhões)	39,2	39,6	39,6	38,5
Celulares (Milhões)	46,4	65,6	86,2	99,9
TV por assinatura (Milhões).	3,5	3,8	4,2	4,6
Usuários de Internet Residenciais (Milhões) .	20,5	17,9	18,9	22,1
Conexões de Banda Larga (Milhões)	1,2	2,3	3,8	5,6

Fonte: Site IBGE (Agosto - 2007).

Após essa breve descrição das mudanças ocorridas no Brasil e no mundo nas últimas décadas, é importante mostrar o mercado brasileiro das operadoras de telecomunicações.

2.2.2 As Empresas Operadoras de Telecomunicações

Telecomunicações é a área do conhecimento que trata da transmissão de voz, dados e imagem por meio de uma rede. Ela surgiu da necessidade das pessoas que estão distantes se comunicarem. Os sistemas telefônicos se difundiram pelo mundo atingindo em 2001 mais de 1 bilhão de linhas e índices de penetração apresentados na Tabela 2.5.

Tabela 2.5 Quantitativo de telefones fixos nas maiores economias do mundo.

	Telefones	Fixo	Densidade
	(Milhões)		(100 hab) Fixo
1	China	350	26,6
2	US	178	59,7
3	Rússia	40	27,9
4	Japão	59	45,9
5	Alemanha	55	66,6
6	Brasil	40	21,6
7	Índia	49	4,4
8	Itália	25	43,1
9	Reino Unido	34	56,5
10	França	36	59,0

Fonte: Revista TELECOM, (Março - 2007).

Com o aparecimento dos sistemas de comunicação móvel com a telefonia celular e a banda larga o termo telefonia fixa passou a ser utilizado para caracterizar os sistemas telefônicos tradicionais que não apresentam mobilidade para os terminais. A operadora que presta o serviço local é aquela que possui a central local e a rede de acesso à qual o terminal

do assinante está conectado. É considerado serviço local aquele destinado à comunicação entre dois terminais fixos em uma área geográfica contínua de prestação de serviços, definida pela Agência, segundo critérios técnicos e econômicos, como uma área local.

As empresas privatizadas do sistema Telebrás receberam as concessões para a prestação de STFC. Essas concessionárias têm que cumprir metas, mais abrangentes de atendimento tendo em vista as metas de universalização de serviço. Isto decorre do fato de serem operadoras estabelecidas em posição de quase monopólio em suas áreas de atendimento, principalmente quanto ao serviço local.

A privatização das operadoras de telefonia fixa no Brasil causou um aumento na oferta de telefones, atingindo em 2002 mais de 49 milhões de acessos instalados. Entretanto, o número de acessos em serviço era menor que 40 milhões em dezembro de 2002, demonstrando que a oferta dos acessos precisa ser melhor, adequada ao que algumas camadas da população podem pagar.

Na figura 2.10 se identifica uma divisão das três regiões da telefonia fixa, em que se tem na região I TELEMAR, na região II Br TELECOM e na região III a TELEFONICA.

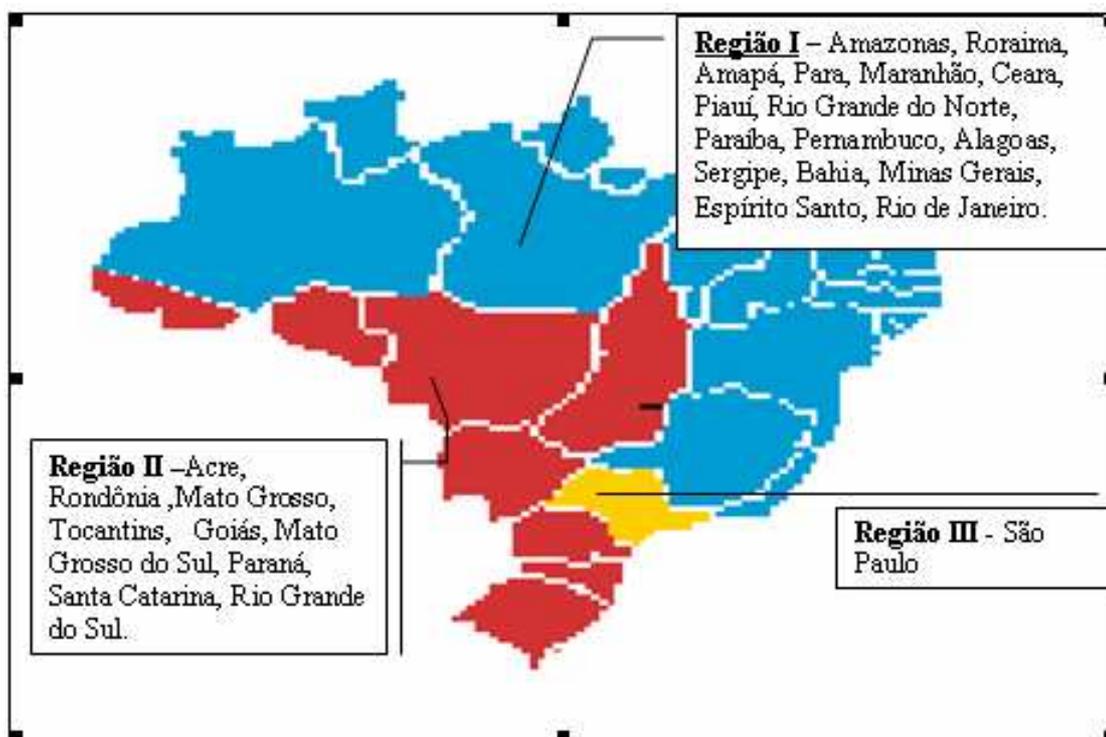


Figura 2.10 Distribuição das áreas de telefonia fixa no Brasil

Fonte: Site ANATEL, (Agosto - 2007).

2.2.3 O Serviço de Telecomunicações de Internet Banda Larga

Um usuário pode acessar a Internet por ligações telefônicas via *modem*, rede digital de serviço integrada (RDSI), circuitos dedicados ou via ADSL. Essa última possibilidade é a que proporciona a melhor relação custo benefício, utilizando os pares metálicos de cobre da rede telefônica existente. As linhas ADSL foram desenvolvidas nos anos 80, para suportar vídeos conferências usando linhas telefônicas. Tendo em vista o grande crescimento dos sistemas ADSL, a expectativa de sucesso dessa tecnologia foi extrapolada. Devido à explosão do uso da Internet, os sistemas ADSL foram introduzidos com sucesso entre a casa do usuário e o ponto de presença da Internet (POP), o que garantiu uma nova alternativa, que foi a conexão de múltiplos usuários em um único acesso a Internet.

A tecnologia ADSL surgiu fora das companhias de telefonia e foi projetada para concorrer com o sistema de TV a cabo, o qual fornece o serviço por meio de seu sistema de cabo coaxial e fibra ótica, sendo que também esta explorando o fornecimento de acesso à Internet. Essa tecnologia foi idealizada para trabalhar simultaneamente com o sistema de telefonia, o mesmo foi projetado de tal forma que a taxa de transmissão no sentido rede para usuário “*downlink*” é maior que o no sentido usuário para a rede “*uplink*”.

No contexto mundial onde o Brasil está inserido a tecnologia de banda larga surgiu como solução para a taxa de transmissão de acesso a preços acessíveis para o cliente da telefonia fixa. Esse serviço vem sendo vendido como um valor agregado das linhas telefônicas fixas, revertendo com isso um pouco da tendência mundial de diminuição da base de clientes na telefonia fixa. Por meio deste novo horizonte as empresas já visualizam um futuro mais promissor devido ao crescente mercado de banda larga. Esse crescimento é cada vez maior, pois a necessidade de aumento da taxa de transmissão no acesso às informações faz com que a banda larga deixe de ser um produto acessível a poucos para se tornar uma necessidade.

A Tabela 2.6 ilustra como essa tendência de crescimento vem refletindo nas quantidades totais de conexões banda larga no Brasil no quarto trimestres de 2006 e no primeiro trimestre de 2007.

Tabela 2.6 Evolução trimestral das tecnologias de acesso à banda larga.

Milhares	1T06	2T06	3T06	4T06	1T07
ADSL	3.432	3.685	3.997	4.341	4.573
TV Assinatura	789	914	1.057	1.200	1.347
Outros(Rádio)	80	92	105	115	120
Total	4.301	4.691	5.159	5.656	6.040

Fonte: Revista TELECOM (Março - 2007).

A Tabela 2.7 ilustra o comportamento trimestral das no segmento banda larga nas Principais Operadoras ADSL.

Tabela 2.7 Evolução trimestral em milhares de terminais nas operadoras ADSL.

Milhares	1T06	2T06	3T06	4T06	1T07
Telemar	896	970	1.046	1.128	1.182
Brt	1.084	1.155	1.252	1.318	1.384
Telefônica	1.300	1.378	1.480	1.607	1.689
GVT	79	95	117	137	162
CTBC	73	87	102	131	N.D.
Sercomtel				20	ND
Total ADSL	3.432	3.685	3.997	4.321	4.573

Fonte: Revista TELECOM, (Março - 2007).

2.2.4 A Engenharia de Manutenção

Em qualquer setor no qual a competição e a busca por cliente é questão de sobrevivência, a manutenção é considerada estratégica, pois sua ausência resulta em perdas e redução do lucro, afetando clientes, funcionários, investidores e a sociedade.

Assim, para que não ocorram prejuízos para os clientes de uma organização, inclusive para os próprios acionistas, em se tratando de manutenção, deve-se investir tempo e recursos para assegurar e gerenciar seus serviços, a fim de obter-se melhor eficiência em seus processos. Por esse motivo, dentro de uma nova visão estratégica, de influência na

competitividade da empresa FACINA, (1999) afirma: "... a manutenção agora faz parte do negócio. E não só faz parte, como também influencia em seus resultados".

As atividades da manutenção dos serviços não se limitam apenas a mantê-los. Mas, dada a sua importância no setor de serviços, suas atividades e responsabilidades devem se revestir de uma maior abrangência. Nesse sentido, XENOS, (1998) subdivide as atividades de manutenção em dois tipos: as atividades de manutenção propriamente ditas e as atividades de melhoria. As atividades de manutenção, também chamadas de execução da manutenção, têm, num sentido mais restrito, a responsabilidade de "manter suas condições originais de operação e seu desempenho por intermédio do restabelecimento de eventuais deteriorações dessas condições". Essas atividades devem ser executadas pelo pessoal da manutenção e operação, com o auxílio de normas e padrões sistematizados.

As atividades de melhoria dos serviços, num sentido mais abrangente, conforme observa XENOS, (1998) "visam melhorar suas condições originais de operação, desempenho e confiabilidade intrínseca, por meio da incorporação de modificações ou alterações no seu projeto ou configuração original". Essas atividades são chamadas de Engenharia da Manutenção.

A – Definição de Manutenção

Segundo MOLBRAY, (1997) a palavra manter, contida na definição de vários dicionários, sugere que manutenção significa preservar alguma coisa. Associação Brasileira de Normas Técnicas (ABNT), na norma NBR-5462 de 1994, define manutenção como "a combinação de todas as ações técnicas e administrativas, incluindo as de supervisão, destinadas a manter ou recolocar um item em um estado no qual possa desempenhar uma função requerida".

O termo manutenção é mais amplamente entendido, quando se questiona a sua missão. MOLBRAY, (1997) levanta as seguintes e importantes premissas a serem consideradas, para uma abrangente formulação da declaração da missão da manutenção:

- As organizações possuem ativos físicos, logo, devem ser preservados;
- Os ativos devem continuar a fazer tudo que os seus usuários querem que façam;

- Os clientes do serviço de manutenção são os proprietários dos ativos, os usuários dos ativos (usualmente operadores) e a sociedade como um todo;
- A tecnologia de manutenção deve ser voltada para encontrar e aplicar modos apropriados de gerenciar a falha dos ativos e evitar suas conseqüências;
- Os custos de propriedade dos ativos devem ser minimizados ao longo de suas vidas úteis, e não apenas visando o final do período contábil seguinte;
- A manutenção depende das pessoas, não somente dos que a executam, mas também dos operadores, projetistas e vendedores.

A partir dessas premissas MOLBRAY, (1997) formula uma declaração para a missão da Manutenção. Preservar as funções dos ativos físicos, ao longo das suas vidas tecnicamente úteis:

- Para a satisfação dos seus proprietários, dos seus usuários e da sociedade como um todo;
- Selecionando e aplicando as técnicas de melhor custo-benefício, para gerenciar as falhas e suas conseqüências;
- Com o ajuda ativa de todo o pessoal envolvido.

Para PINTO e XAVIER, (1998), “manutenção existe para que não haja manutenção” e a missão da manutenção está atualmente focada, na função que o ativo deve desempenhar e não no ativo em si. Isto provoca uma mudança radical no entendimento da missão da manutenção, e, por conseqüência, nas ações do pessoal da manutenção.

A NBR 5462 de 1994 define confiabilidade como a probabilidade de um item desempenhar sua função requerida, por um intervalo de tempo estabelecido, sob condições definidas de uso.

O desempenho da manutenção é fortemente dependente da interação entre as fases do processo de implantação de uma prestadora de serviço, desaguando nos seus dois principais indicadores; a disponibilidade e a confiabilidade. Essas fases compreendem o projeto, a fabricação, a montagem e a manutenção [PINTO E XAVIER, (1998)], conforme mostra a Figura 2.11.



Figura 2.11 Interação entre as fases.

Fonte: Adaptada PINTO E XAVIER, (1998).

É importante ressaltar que, segundo MONCHY, (1989) “por maior que seja a importância do serviço manutenção, ele não é um fim, mas um meio para ajudar a produção”. Nisto CASTELLA, (2001) concorda em que a responsabilidade pelo sucesso do desempenho do ativo ou da instalação, deve-se também a uma interação entre as áreas da empresa, durante as fases do ciclo de vida do serviço prestado, e não somente a manutenção.

Por fim, como conceito adotado para essa pesquisa, a Engenharia de Manutenção, preocupa-se com a produção do serviço, atuando na capacitação do pessoal e no desenvolvimento de técnicas para atuação nos serviços e instalações e manutenção em geral. Sendo assim identifica-se por meio de planejamento e gerenciamento de sistemas a manutenção desses serviços e instalações, de modo a operar dentro dos requisitos de confiabilidade e disponibilidade, associadas ao menor custo, preservando a vida e a segurança das instalações. Busca ainda, o aumento contínuo da produtividade e da qualidade dos serviços.

B – Evolução da Manutenção

Como resultado da evolução, no último século, da área tecnológica e das técnicas de gerenciamento, a indústria passou por drásticas transformações. A manutenção, seguindo essa mudança, também evoluiu principalmente nas últimas cinco décadas. Essa evolução, de acordo com MOLBRAY, (1997) deu-se em três gerações.

A primeira geração cobre o período da Segunda Guerra Mundial. Esse período caracteriza-se pela existência de poucas máquinas nas indústrias, geralmente simples, super dimensionadas e de fácil manutenção, não exigindo, portanto, uma manutenção sistematizada. Por estar á exigência de produtividade ainda no início, os tempos de reparo não tinham relevância e a manutenção restringia-se a limpeza, assistência e lubrificação.

A *segunda geração* recebe todo o impacto da Segunda Guerra Mundial, trazendo a sofisticação das máquinas e o aumento do volume de produção, devido à demanda por todo o tipo de produto. Em face dessa necessidade de produção, as falhas começaram a ser motivo de preocupação, bem como o tempo de parada para reparo das máquinas. Como conseqüência da sofisticação das máquinas, aliada a falta de mão de obra especializada, os custos de manutenção começaram a ser representativos na composição do preço final dos produtos.

Nesse período aparece o interesse pelo aumento da vida útil das máquinas. Surge então, a necessidade do planejamento da manutenção, como também as revisões periódicas, dando origem ao conceito de manutenção preventiva. Essa geração estende-se até fins da década de 60.

Finalmente, a *terceira geração* surge, a partir da década de 70, com o crescimento da automação e da mecanização. Aumenta-se a preocupação com a paralisação da produção e conseqüentemente com a qualidade dos produtos. Questões novas são levantadas no seio da sociedade, com implicações diretas tais como: segurança, preservação do meio ambiente e satisfação do cliente.

Todas essas questões levaram a procura pela eliminação da falha dos serviços ou instalações. É nesse período que os conceitos de manutenção preditiva, disponibilidade e confiabilidade aparecem. Também, nesse período, em face da necessidade de alta capacitação do homem de manutenção, o trabalhador, o trabalho em equipe e as técnicas de gerenciamento são valorizados.

Essas mudanças ocorridas ao longo das três gerações provocaram uma evolução na visão e na postura do homem de manutenção, em relação a sua missão e a missão da manutenção, que passou da consideração da manutenção como um mal necessário, para a visão da manutenção como função estratégica.

PINTO e XAVIER, (1998) concordam sobre a evolução da postura do homem de manutenção, e a Figura 2.12 mostra a mudança do seu comportamento ao longo do tempo.

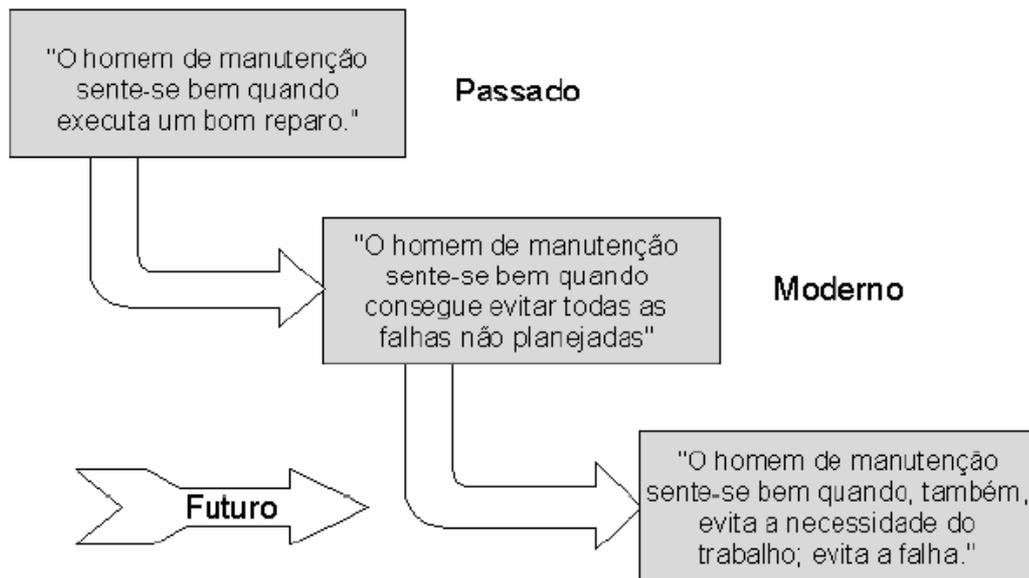


Figura 2.12 Evolução da postura da manutenção.

Fonte: Adaptada de PINTO e XAVIER, (1998).

A Engenharia de Manutenção surge juntamente com a evolução da manutenção, da chamada primeira geração. Essa nova disciplina surgiu, em decorrência da simples observação, por parte dos gerentes, entre o tempo gasto para diagnosticar as falhas e o reparo das mesmas, gerando a necessidade da criação de equipes especializadas, para planejar e controlar a manutenção [TAVARES, (1999)].

PINTO e XAVIER, (1998) observam que a prática da Engenharia de Manutenção significa "uma mudança cultural, perseguir *benchmarks*, aplicar técnicas modernas". Segundo TAVARES, (1999) com a difusão dos computadores, a partir de 1966, e dos microcomputadores, a partir de 1980, possibilitou-se o desenvolvimento de programas de gerenciamento e de acompanhamento e controle da manutenção, bem como de critérios de previsão das falhas.

C - Tipos de Manutenção

Muitas são as variações da nomenclatura e das definições para classificar os tipos de manutenção encontrados na literatura, chamados ainda de métodos ou formas de manutenção.

A classificação adotada nesse trabalho abrange as mais utilizadas na literatura corrente, contribuindo para o entendimento dessa matéria.

Existem duas maneiras principais de se classificar e entender os tipos ou formas usuais de manutenção. A primeira delas é classificar a manutenção de acordo com o tipo de intervenção que se faz no serviço ou instalação.

Assim, é identificada a manutenção corretiva, a preventiva baseada no tempo, a preventiva baseada em condição ou preditiva e a manutenção de melhoria, com as suas definições como se segue:

- Manutenção corretiva - intervenção decorrente de uma falha, ou mau funcionamento. Um serviço exige manutenção corretiva quando é necessário intervir no mesmo porque interrompeu ou degradou sua função;
- Manutenção preventiva baseada em tempo - intervenção feita a intervalos regulares de tempo corrido (por exemplo, semanas) ou de funcionamento (por exemplo, horas trabalhadas). Em inglês conhecida como *TBM (time based maintenance)*;
- Manutenção preventiva baseada em condição ou preditiva – intervenção feita de acordo com o acompanhamento de determinados parâmetros do serviço (por exemplo, medição de desgaste ou elevado grau de degradação no serviço). Em inglês conhecida como *CBM (condition based maintenance)*;
- Manutenção de melhoria - intervenção feita para alterar as condições de um serviço com o objetivo de aumentar o seu rendimento, a qualidade dos produtos processados ou melhorar algum parâmetro operacional.

Os dados estatísticos, disponibilizados pela empresa, caracterizam que os recursos de manutenção de telecomunicações são distribuídos da seguinte forma:

- 78,5 % em manutenção corretiva;
- 8,0 % em manutenção preditiva;
- 13,5 % em execução de projetos e melhorias.

Pode-se perceber que, em média, ainda 3/4 dos recursos da manutenção são aplicados em corretivas, o que significa quebras e falhas de serviços e paradas não programadas.

A segunda maneira de se entender e classificar os tipos de manutenção, adotado por MONCHY, (1989) é conceituá-los por meio da condição da ocorrência ou não da falha no serviço ou instalação, conhecida também como formas de manutenção. Em manutenção falha

e defeito, têm significados distintos. Para BRANCO, (1996) define falha como "o término da capacidade de um serviço desempenhar a função requerida", enquanto que "um defeito não torna o serviço indisponível".

Assim, em caso de ocorrência de falha do serviço, a sua condição é alterada por meio da aplicação da manutenção corretiva. Caso não haja detecção de falha no serviço ou haja detecção de defeito, a condição é administrada pela aplicação da manutenção preventiva.

Nesse sentido NUNES, (2001) define a manutenção corretiva como "todo trabalho de manutenção, realizado após a falha, visando restabelecê-lo à sua função requerida, eliminando o estado de falha", e manutenção preventiva em trabalhos realizados no serviço, "quando se apresentam situações de defeito, ou não se caracterizou um estado de falha". Normalmente programada e aplicada para redução da probabilidade de ocorrência da falha. A Figura 2.13 elucida a classificação acima exposta. Cabe destacar que, de acordo com a descrição acima e a Figura 2.13, a literatura corrente converge para os seguintes tipos de manutenção: manutenção corretiva e manutenção preventiva. Os demais tipos constituem-se variações ou evoluções dessas duas.

A Figura 2.13 apresenta dois tipos de intervenção para a manutenção corretiva. A manutenção paliativa, como intervenções executadas provisoriamente, a fim de tirar o serviço do estado de pane e colocá-lo em funcionamento, para, de maneira programada, executar o reparo definitivo. A manutenção curativa, como intervenções executadas em caráter definitivo, para restabelecimento da função requerida do serviço. Da mesma forma, a Figura 2.13 apresenta três variações para a aplicação da manutenção preventiva, em função do conhecimento do comportamento do serviço, ou da evolução do desgaste ou deterioração do serviço, chamada de lei de degradação do serviço.

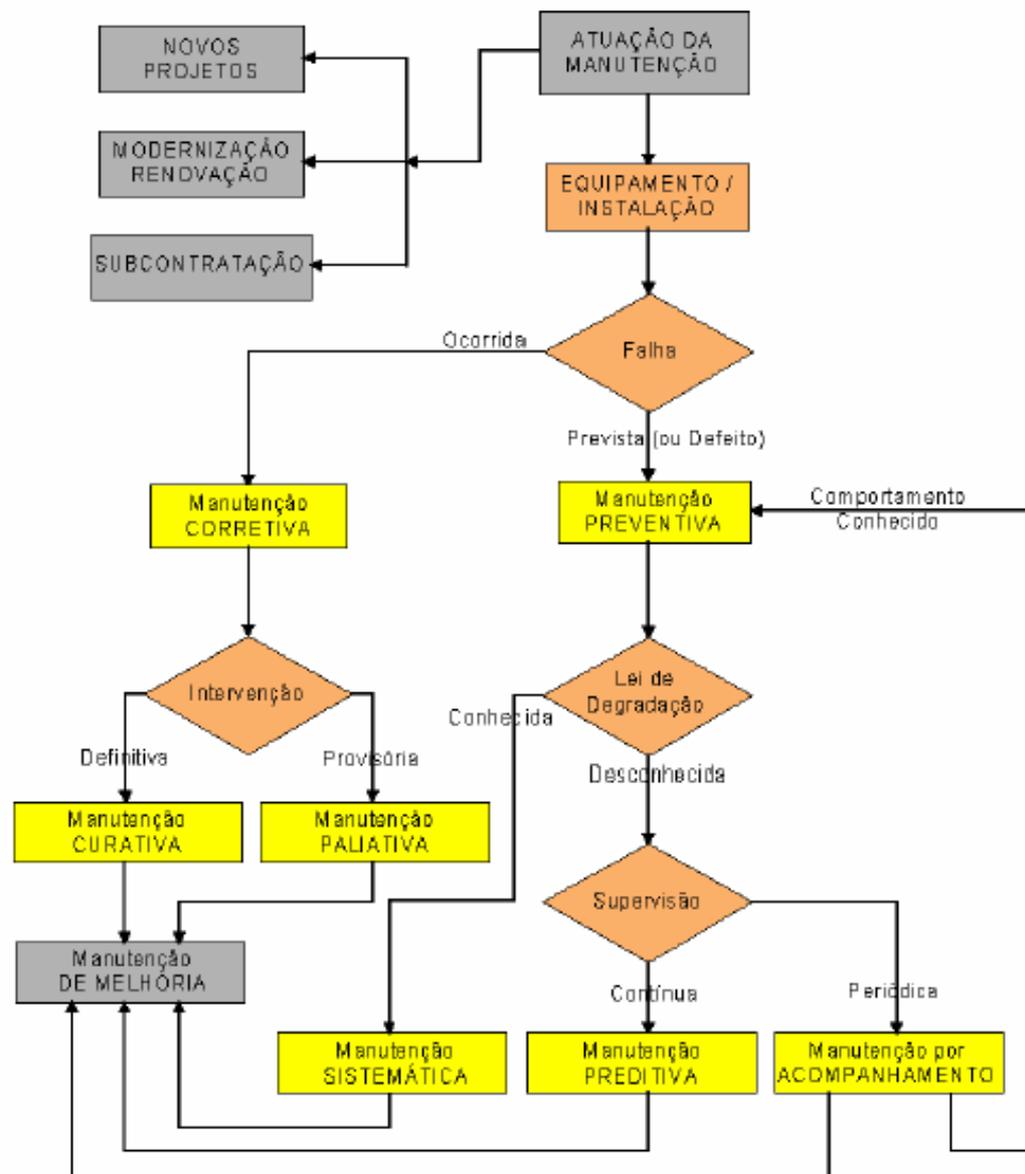


Figura 2.13 Tipos de manutenção.

Fonte: Adaptada de MONCHY, (1989).

A primeira, em caso de conhecimento da lei de degradação, é conhecida como manutenção preventiva sistemática, aplicada, de maneira periódica, em serviços cuja falha causará grandes prejuízos.

Em caso de desconhecimento da lei de degradação do serviço, a manutenção preventiva evolui para dois tipos. O segundo tipo é manutenção preventiva por

acompanhamento ou manutenção de rondas, aplicada na forma de intervenções e/ou inspeções leves e de curta frequência.

O terceiro tipo de manutenção preventiva é manutenção preventiva condicionada, mais conhecida como manutenção preditiva ou técnicas preditivas de manutenção. O termo condicionado é aplicado por estar à execução da manutenção do serviço subordinada a um evento, decorrente de informações monitoradas de maneira constante ou periódica, geralmente pelas inspeções realizadas por meio dos sentidos humanos, por instrumentos manuais ou automatizados, sem, contudo, necessitar retirá-lo de operação. É a forma mais moderna de manutenção, por atuar no serviço no momento adequado e anterior à falha do serviço, reduzindo riscos e minimizando custos.

Uma outra maneira de especificar os tipos de manutenção poderia ser reativa e pró-ativa, conforma abaixo:

1. Atividades de manutenção reativa ou corretiva: a manutenção corretiva pode ser descrita como uma técnica de "combate ao fogo", ou seja, somente há a intervenção no momento em que o serviço falha, o objetivo do serviço de manutenção é minimizá-las, por meio de um trabalho forte de análise de falhas e confiabilidade;
2. Atividades de manutenção pró-ativa: a manutenção pró-ativa é uma estratégia para manutenção onde as falhas são minimizadas por meio de atividades de monitoramento da deterioração dos serviços e intervenções para restaurá-los às condições próprias de funcionamento.

Manutenção preventiva: inicialmente, esse tipo de operação de manutenção baseava-se na estimativa da probabilidade que o serviço irá falhar para um determinado intervalo de tempo. Porém, em função do aumento dos custos de manutenção, essas intervenções passaram a ser realizadas baseadas na condição dos serviços, por meio de inspeções regulares.

Manutenção condicionada ou preditiva: especificamente, a manutenção preditiva é uma resposta a uma condição física momentânea como temperatura, vibração, ruído, etc. Segundo EADE, (1997) quando um desses indicadores atinge um determinado nível, uma intervenção é realizada para restaurar o serviço às condições próprias.

A escolha do tipo de manutenção a ser aplicado no serviço ou instalação, ou até mesmo o seu emprego combinado, deve passar, necessariamente, por uma análise de custo-benefício [NUNES, (2001)]. Sobre esse assunto TAVARES, (1999) apresenta a evolução do custo das manutenções corretivas e preventivas ao longo do tempo, de acordo com a Figura 2.14.

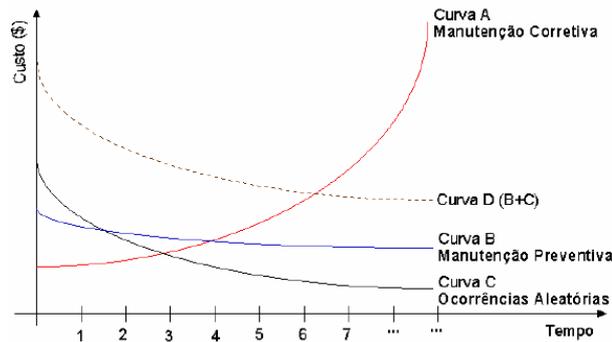


Figura 2.14 Custo de manutenção em relação ao tempo.

Fonte: Adaptada de TAVARES, (1999).

Definidos e descritos os conceitos de manutenção possíveis (corretiva, preventiva e preditiva), a pergunta comum é como fazer a escolha entre esses? Qual é a melhor opção? Nenhum conceito de manutenção é intrinsecamente "melhor" que qualquer outro. Cada um possui o seu lugar e o grande desafio consiste em encontrar qual é esse lugar. Isto tem levado os sistemas de produção e prestação de serviços, a buscarem novas estratégias possíveis para melhorar o atendimento das necessidades de seus clientes.

Cabe ainda ressaltar que, devido à preocupação constante, por parte das empresas, com a qualidade dos serviços prestados e a confiabilidade das instalações, foram desenvolvidas mais recentemente, na área de manutenção, algumas metodologias e ferramentas para auxiliar na sua gestão. As mais difundidas são a Manutenção Produtiva Total e a Manutenção Centrada em Confiabilidade.

A Manutenção Produtiva Total (em inglês, *Total Productive Maintenance - TPM*) é uma técnica desenvolvida no Japão na década de 70, para melhoria da qualidade dos produtos e serviços. Autores como MIRSHAWAKA e OLMEDO (1994), PINTO e XAVIER (1998) e TAVARES (1999), a definem como uma filosofia de manutenção que envolve o conjunto de todos os empregados da organização, pela reformulação e melhoria da sua estrutura, por meio

de uma maior qualificação das pessoas e melhoria dos serviços, buscando a maximização da eficiência dos serviços ou instalações.

O TPM, ou "*Total Productive Maintenance*" é uma filosofia de gestão da manutenção desenvolvida inicialmente em plantas industriais japonesas como suporte à implementação do método de fabricação "*Just-in-Time*", frente à adoção de avançadas tecnologias de fabricação, principalmente na indústria automobilística, e como apoio aos esforços para a melhoria da qualidade do produto [NAKAJIMA, (1989)].

As atividades de TPM são focalizadas na eliminação das "seis maiores perdas", ou as maiores causas de perda de rendimento operacional da manufatura, que normalmente são as falhas de serviços ou confiabilidade da linha, tempo de partida e ajuste de linha, paradas menores e desnecessárias, perda de tempo de ciclo, defeitos no produto e redução da capacidade produtiva. O TPM tem sido descrito como uma abordagem de parceria para a função manutenção, relacionada à cooperação do serviço de manutenção, a fabricação e os demais serviços de apoio, como engenharia de processo, no intuito de aperfeiçoar o desempenho de fabricação [MAGGARD e RHYNE, (1992)].

Com o TPM, os operadores de fabricação tornam-se envolvidos na manutenção dos serviços, por meio de treinamento intensivo, passam a executar as tarefas básicas de monitoramento, limpeza e pequenos ajustes, e que são definidas como operações de manutenção autônoma.

A Manutenção Centrada em Confiabilidade - RCM (em inglês, *Reliability Centered Maintenance*), surgiu na década de 60, na indústria aeronáutica americana, como uma ferramenta de manutenção, para eliminação dos modos de falha, ou redução dos riscos a elas associados, no auxílio das empresas aéreas em seus programas de manutenção.

NUNES, (2001) observa que a RCM "visa racionalizar e sistematizar a determinação das tarefas adequadas a serem adotadas no plano de manutenção, bem como garantir a confiabilidade e a segurança operacional dos serviços e instalações ao menor custo".

A RCM, ou "*Reliability Centred Maintenance*" é uma estratégia de manutenção com origem em uma frustração da autoridade federal de aviação americana (FAA) e das companhias aéreas comerciais na impossibilidade de se controlar a taxa de falha em classes de serviços não confiáveis, por quaisquer mudanças possíveis, tanto no conteúdo, como na frequência das revisões programadas. Como resultado, implantou-se uma equipe, denominada

de força-tarefa, com representantes desses dois segmentos, para investigar a capacidade técnica da manutenção preventiva. A definição de RCM é de "um processo usado para determinar as exigências de manutenção de um bem físico no seu contexto operacional atual" [NOWLAN e HEAP, (1978) e MOLBRAY, (1991)]. É uma metodologia estruturada, a qual se empenha em permitir que um serviço preencha as funções e os padrões de desempenho pretendidos, no seu contexto operacional atual.

Desse modo, o RCM parte do serviço e o que ele deve realizar, mostrando às equipes de manutenção e fabricação quais as rotinas de manutenção são necessárias para o alcance dos objetivos. O processo RCM consiste em responder sete questões principais sobre os serviços e seus subsistemas: Quais são as suas funções? Quais as formas (funções) de falhas? Os que as fazem falhar? O que importa se elas falham? Pode alguma coisa ser feita para predizer ou prevenir a falha? E o que se deve fazer se não se pode predizer ou prevenir a falha?

Para a identificação dos modos e efeito das falhas em potenciais dos serviços, sistemas ou processos, a RCM utiliza a ferramenta designada de Análise do Efeito e Modo de Falhas - FMEA (em inglês, *Failure Mode and Effect Analysis*). O FMEA é uma ferramenta definida como sendo um processo sistemático, que permite identificar potenciais falhas de um sistema, projeto e/ou processo, com o objetivo de eliminar ou minimizar o risco associado, antes que aquelas aconteçam.

Infelizmente, parece existir uma pressão a qual está forçando as organizações produtivas e de serviços a escolherem uma estratégia em detrimento a outra. Na verdade, as estratégias se complementam. De acordo com GERAGHETY, (2001) a função manutenção tem dois objetivos básicos: determinar as exigências de manutenção dos bens físicos dentro de seu contexto operacional atual, e assegurar que essas exigências sejam as mais fáceis de atender, com o menor custo e efetivas possíveis. O RCM está direcionado para ao alcance do primeiro objetivo e o TPM está focado no segundo objetivo.

D - Atribuições da Manutenção

A Operação e Manutenção têm, dentro de uma organização, a função de propiciar o desenvolvimento técnico de toda a área da manutenção industrial, por meio do emprego de

ferramentas que proporcionem a melhoria do desempenho, tanto do pessoal da manutenção, como dos serviços e instalações.

Para a atuação da Operação e Manutenção, além da necessidade dos conhecimentos técnicos específicos de engenharia de cada planta a ser mantida, são necessários, também, os conhecimentos de administração, economia e estatística. Portanto, essa área tem caráter multidisciplinar e com um forte componente de natureza criativa. Conseqüentemente, as suas atribuições são abrangentes, e a sua descrição, além de ser extensa, varia de acordo com a filosofia e os objetivos de cada empresa.

De forma resumida e de acordo com a literatura pesquisada, as principais atribuições da Operação e Manutenção:

a) Na esfera gerencial:

- Definir política de manutenção, filosofia de gerenciamento, planejamento estratégico, orçamento e investimentos;
- Coordenar e implantar sistemas de gerenciamento de manutenção.

b) Na área de controle:

- Padronizar os procedimentos de manutenção;
- Levantar e disponibilizar indicadores;
- Aplicar as técnicas especializadas de manutenção para monitoramento;
- Desenvolver fornecedores de materiais, serviços e serviços utilizados na manutenção;
- Dimensionar o material sobressalente;
- Analisar e dar parecer sobre o desempenho dos serviços e instalações, bem como as causas e conseqüências das falhas.

c) Junto à execução:

- Estudar e propor soluções para problemas críticos e crônicos de serviços;
- Formar e atualizar a equipe técnica da manutenção;
- Apoiar tecnicamente a execução da manutenção, tornando-a mais simples, proporcionando melhorias;
- Identificar as oportunidades de melhoria no processo e práticas de gestão.

2.2.5 Modelo de Gestão para as Operações de Manutenção

A perspectiva da manutenção como um provedor de serviços internos, no qual a gestão do valor entregue ao cliente deve ser otimizado, torna-se um fator primordial para a eficiência e produtividade, na busca da satisfação dos mesmos.

A) Representação do Serviço de Manutenção

A representação do serviço de manutenção em telecomunicações pode ser construída a partir da definição, referente ao sistema genérico de operações. "Conhecendo-se" o serviço de manutenção, é possível definir as variáveis de influência da proposição de valor, que são os "atributos de valor", os indicadores associados e os processos críticos dessas operações. Portanto a descrição do serviço de manutenção é uma condição à operacionalização do valor.

São considerados como dados de entrada do sistema de representação de operações e manutenção, conforme figura 2.15, as definições relativas às entradas, saídas, processos, e interferências do meio-ambiente ao qual esse sistema de operações está inserido.

1- A missão manutenção:

A missão do serviço de manutenção deve ser realizada a partir da definição da função manutenção em um sistema de operações. Nesse trabalho a definição de GERAERDS, (1999): "são todas as atividades relacionadas em manter, um serviço ou instalação, ou restaurá-lo a um estado físico considerado necessário, para garantir a sua função de produção".

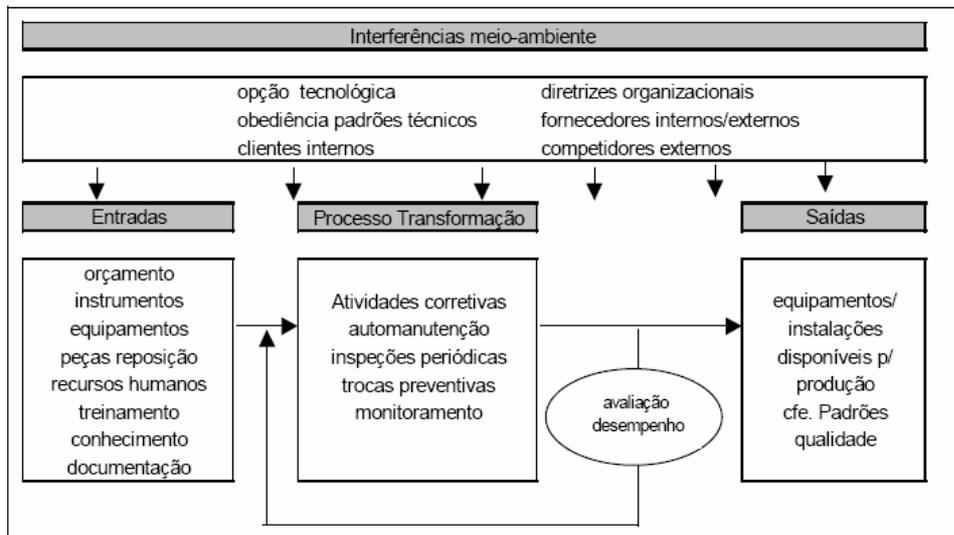


Figura 2.15 Dados do Sistema Genérico de Operações e Manutenção

Fonte: Adaptada de MEREDITH, (1992).

A definição da missão da manutenção deve partir do pressuposto que um serviço ou instalação desgasta-se com o uso, porém a manutenção minimiza a probabilidade de falha, quando o mesmo é solicitado a produzir, estabelecendo e executando ações pró-ativas. Ao mesmo tempo, a definição da missão da manutenção deve estar em sintonia com o plano estratégico estabelecido para a organização, para a correta orientação, no sentido dos esforços e recursos utilizados, da função manutenção, em relação às diretrizes organizacionais.

2- O desdobramento do objetivo e diretriz:

O estabelecimento da diretriz e objetivo é básico para qualquer atividade ou negócio, uma vez que define aonde e como a organização quer chegar. Os objetivos são desenvolvidos a partir dos anseios e expectativas quanto a uma situação futura que satisfaça esses anseios. Segundo REBOUÇAS, (1998), existe dois elementos importantes no estabelecimento de quaisquer objetivos:

- "O elemento psicológico", que envolve os valores, atitudes, motivações e desejos;
- Os "recursos" humanos e materiais envolvidos na obtenção desses objetivos.

Na prática, a diretriz e objetivo posicionam a organização frente à concorrência em uma situação futura. Esse posicionamento deve ser repassado para os demais níveis da

organização, e, portanto cada departamento deve ter o estabelecimento claro e concreto dos objetivos de desempenho a serem alcançados.

Os objetivos de desempenho são ainda fortemente influenciados por fatores inerentes ao ambiente. No caso específico das operações de manutenção em telecomunicações, pode-se dizer que essas operações sofrem influências no ambiente interno, por meio da opção tecnológica das instalações e serviços, dos fornecedores e clientes internos. Também há influências do ambiente externo, por meio de normas técnicas e regulamentações, de fornecedores externos e de uma possível concorrência de prestadores desse mesmo tipo de serviços. Tais referências são:

A - Opção tecnológica de produção: as organizações de produção têm optado por novas tecnologias no sentido de melhorar o desempenho, as quais são normalmente mais complexas em serem mantidas, e ao mesmo tempo, as falhas na prestação dos serviços tendem a ter um custo maior, bem como provocar perdas maiores.

B - Fornecedores e clientes: com o avanço da implantação tecnológica nas plantas produtivas, os serviços internos de apoio de recursos humanos, de compras e o de engenharia de processos parecem ter um papel ainda mais importante, o mesmo acontecendo para os fornecedores externos, por exemplo, de insumos para os processos de manutenção.

C - Competidores externos: o serviço de manutenção pode ser fornecido por uma empresa externa à organização, sendo uma ameaça real. Face a essa força competitiva, se defini estratégias básicas de neutralização, definidas de uma forma geral para a gestão de serviços por GIANESI e CORRÊA, (1994) e que podem ser utilizadas especificamente para as operações de manutenção:

- Diferenciação de serviços: ter a melhor competência da gestão do serviço.
- Custos de troca: tornar o serviço de manutenção o mais especializado possível, de forma a aumentar os custos de troca.
- Tecnologia de informação e base de dados: possuir um sistema de informações, que seja fiel ao nível do desempenho das operações.
- Relação entre capacidade disponível e variação da demanda: possuir recursos humanos e ferramentas flexíveis à variação da demanda.

D - *Normas e regulamentações*: os objetivos de desempenho do serviço de manutenção são também influenciados por normas técnicas e regulamentações internas, quanto definidas

pelo ambiente a que está sujeita a organização, como por exemplo, à legislação governamental.

3 - A proposição de valor:

O valor no serviço de manutenção é proposto a partir dos objetivos desdobrados da organização, ou seja, a equação do valor (equação inicial) deve contemplar os "atributos de valor" que influenciam o alcance desses objetivos. Assim, a proposição de valor é baseada em:

- Objetivos desdobrados para o serviço de manutenção, a partir do planejamento estratégico da organização.
- São relacionados a esses objetivos, os benefícios e sacrifícios que influenciam o seu alcance, ou seja, os "atributos de valor" genéricos da composição do valor em serviços.

4- Conceito de manutenção adotado:

Consiste em descrever as práticas usuais ou o conceito de manutenção relacionado à habilidade de manter um serviço ou instalação em condições de operação. O conceito de manutenção a se adotar para um sistema de produção e prestação de serviços, depende, entre outros fatores, da disponibilidade requerida dos serviços, da sua complexidade tecnológica, e dos custos envolvidos em cada prática de manutenção, que podem ser reativas e pró-ativas, conforme detalhamento mostrado no item 2.2.4-C.

5- A organização das operações:

A organização do serviço de manutenção de um sistema produtivo e de serviços pode ser realizada de acordo com uma divisão em níveis de atividades de manutenção e conforme o conceito de manutenção adotado.

Normalmente, as atividades de manutenção corretiva, preventiva e preditiva interagem no TPM, e as operações do serviço de manutenção podem ser divididas em cinco níveis.

Isto é realizado de acordo com as responsabilidades sobre as tarefas, seja do cliente, o qual não somente participa do serviço, mas tem responsabilidade sobre uma parte dessas

operações, dos mantenedores da operação e manutenção, definindo-se funções e um perímetro claro de cada um na organização das atividades, como pode ser visualizado no figura 2.16.

ATIVIDADES	NÍVEL	RESPONSABILIDADE	CLASSIFICAÇÃO OPERAÇÃO
Condução equipamentos Limpeza Inspeções (utilizando os sentidos)	N1	Operadores de fabricação	Operações de manutenção de "linha-de-frente" ou "front-office"
Manutenção autônoma Ajustes Pequenos reparos/intervenções	N2		
Intervenções emergenciais e programadas Manutenção corretiva Manutenção preventiva Manutenção preditiva/monitoramento	N3	Operadores de manutenção	
Planejamento/control operacional Análise da confiabilidade Plano manutenção preventiva Listas peças de reposição	N4	Engenharia de manutenção	Operações de manutenção de "retaguarda" ou "back-office"
Especialidade/apoio às operações: Especialidade técnica Padronização peças de reposição Definição política de manutenção	N5		

Figura 2.16 Organização das Operações de Manutenção por Níveis de Atividades, Perímetros e Funções.

Fonte: Adaptada de MEREDITH, (1992).

6- A execução das atividades de manutenção:

A execução das atividades de manutenção, para garantir as condições de operação dos serviços e instalações, depende dos recursos humanos e materiais envolvidos:

Recursos humanos: os quais devem ser motivados e qualificados por meio de treinamento intensivo, são:

- Operadores de serviço: os quais executam os níveis 1 e 2 de manutenção;
- Mantenedores: os quais exercem as funções relacionadas ao nível 2 (corretiva e preventiva);
- Engenharia de manutenção: recursos responsáveis pela execução dos níveis 3 e 4. É o apoio às operações, pelo planejamento e controle, e especialidade técnica;
- Sub-Contratados: eventualmente podem ser contratadas empresas externas que dispõem de pessoal técnico e ferramental especializado.

Recursos materiais: têm por finalidade subsidiar, por meio da utilização racional, as atividades de manutenção:

- Documentação Técnica: é o conjunto das especificações e procedimentos técnicos normalmente fornecidos na fase de recepção dos novos serviços e instalações.
- Peças de reposição/consumíveis: são as peças e componentes constituintes do maquinário e instalações de fabricação, as quais sofrem desgaste e que são repostas corretiva ou preventivamente.
- Ferramentas/instrumentos: são essenciais para o diagnóstico, reparação, previsibilidade e planejamento quanto às atividades de manutenção.

B) Valor do Serviço de Manutenção

As operações de Manutenção têm um papel fundamental como parcela de contribuição do valor total produzido, e a performance dessas atividades e conseqüentemente dos serviços de produção são elementos vitais nesse processo. Isto se evidencia na participação dos custos de Manutenção e de perdas de clientes, decorrentes da performance das operações de Manutenção. Atualmente é admitido por vários autores que a Manutenção é o maior contribuinte para a performance e lucratividade dos sistemas de serviços [MAGGARD e RHYNE, (1992); PEHANICH, (1995); COETZEE, (1998)].

As operações de Manutenção, originalmente pouco exploradas quanto ao potencial de ganho de performance, passaram de meras atividades reativas, para um contexto de desenvolvimento gerencial e ganhos crescentes de produtividade.

As razões principais, que levaram as organizações de serviço a gerirem efetivamente as operações de Manutenção, são:

- Com o fim do monopólio, uma competição crescente exigiu um controle irrestrito de custo, com a crescente participação da Manutenção nos custos operacionais;
- Maiores taxas de instalações exigiram o aumento da confiabilidade e disponibilidade dos serviços;
- A perda de qualidade ligada ao mau funcionamento levando as maiores taxas de abandono;

- Redefinição das atividades de Manutenção em Telecomunicações fomentada por um maior gerenciamento das redes.

O fato é que o serviço de manutenção nas prestadoras de serviços não pode mais ser visto como "mal necessário", existindo somente como um sorvedouro de recursos e que não agrega valor aos processos. Justamente o contrário, as operações de manutenção têm uma correlação muito forte com a produtividade desse setor, interferindo não somente sobre a disponibilidade dos serviços, mas também sobre a qualidade, a segurança, enfim sobre as condições que levam aos ganhos de desempenho.

Pelas diversas razões já enumeradas, a gestão do serviço de manutenção deve ser realizada, primeiramente considerando os aspectos das especificidades das operações de serviços e da perspectiva de contribuição do valor desse serviço, no intuito de se alcançar vantagem competitiva para a organização.

C) A Operacionalização do Valor

A operacionalização do valor, como componente complementar do modelo de gestão do valor do serviço de manutenção, consiste na identificação dos critérios de percepção do valor pelo cliente, o tratamento desses critérios com as quais sua priorização, análise de possíveis conflitos e correlação com os processos de manutenção, e também a avaliação de desempenho baseado nos critérios priorizados. O conjunto das ações anteriores permite o estabelecimento do plano de ações prioritário para a operacionalização do valor.

1- Identificar os "atributos de valor" e indicadores associados:

A percepção do valor é influenciada por critérios, os quais podem ser conceituados como os atributos de um produto ou serviço, que constituem a razão primária de seleção, e que agregam valor ao consumidor. No caso de um serviço, e especificamente no serviço de manutenção, esses fatores ou "atributos de valor" têm uma natureza extremamente subjetiva, mas que podem ser quantificáveis a partir dos indicadores de desempenho das suas operações. Nessa proposição, para a identificação dos "atributos de valor" específicos do serviço de manutenção foram utilizados como referência, os trabalhos de identificação desses critérios,

inicialmente chamados de critérios gerais da percepção da Qualidade em serviços – "SERVQUAL", por PARASURAMAN et al., (1988) e mais especificamente realizados por LAPIERRE, (2000). Basicamente a identificação desses critérios, a ser desenvolvida pelo gestor do serviço, desenvolve-se por meio de:

- Entrevistas, que são conduzidas com "pessoas-chave" relacionadas ao serviço, sejam clientes ou fornecedores;
- Primeiramente essas entrevistas devem abordar o conceito de valor percebido, a fim de se relacionarem as diferenças de percepção entre clientes e fornecedores;
- Devem também identificar junto à população entrevistada as formas de agregação de valor, ou seja, quais são os "*atributos de valor*" do serviço;
- Em uma segunda etapa, podem ser acrescentadas nas entrevistas, questões específicas sobre cada "*atributos de valor*" identificado, ou seja, qual é a percepção de cada pessoa sobre os critérios.

Como os trabalhos de PARASURAMAN et al., (1988) e LAPIERRE, (2000) possuíram como população pesquisada, os serviços de manutenção e assistência técnica, então a seguir, baseado nessas pesquisas, apresenta-se os fatores influenciadores específicos na percepção do valor do consumidor das operações de manutenção.

A - Taxa de transmissão de Atendimento (VA): A taxa de transmissão de atendimento refere-se à prontidão, ou seja, o tempo dispensado para a execução das atividades de manutenção, às quais têm maior evidência para o cliente, àquelas relacionadas a intervenções corretivas e que normalmente geram perda de produção. Também dizem respeito ao tempo previsto e efetivamente gasto nas operações planejadas, ou seja, as intervenções corretivas programadas e os serviços de melhoria. Os indicadores normalmente utilizados para esse critério de avaliação são:

- *MTTR* ou "*Medium time to repair*": é o tempo médio para reparo corretivo, ou seja, o tempo médio para restabelecer as condições de funcionamento de uma instalação ou serviço.
- Taxa de intervenções programadas: é a relação entre o prazo planejado e o executado das atividades programadas de natureza corretiva e melhorias.

B – Confiabilidade (CF): Define-se a confiabilidade de um serviço, como uma das dimensões da qualidade ou valor percebido, de acordo com PARASURAMAN et al., (1988) como "a habilidade de prestar o serviço de forma confiável, precisa e consistente".

Na prática, para as operações de manutenção, a confiabilidade desse serviço se traduz na disponibilidade e no desempenho das instalações e serviços. Ou seja, quanto mais disponível está um serviço ou sistema para a produção, mais se percebe ser confiável o serviço de manutenção, considerando-se evidentemente as condições de funcionamento ou de concepção (projeto e instalação).

Já a disponibilidade de um item (componente, subsistema ou sistema) é expressa, de acordo com a NBR5462 (1994), como: capacidade de um item estar em condições de executar certa função em um dado instante ou durante um intervalo de tempo determinado, levando-se em conta os aspectos de sua confiabilidade, e suporte de manutenção, supondo que os recursos externos requeridos estejam assegurados. Ou seja, por essa definição evidencia-se que a disponibilidade de um item é uma função da confiabilidade (de um item e não de um serviço) e da manutenibilidade, cujas definições segundo NBR5462 são apresentadas a seguir:

- "Confiabilidade é a capacidade de um item desempenhar uma função requerida sob condições especificadas, durante um intervalo de tempo";
- "Manutenibilidade é a capacidade de um item em ser mantido ou recolocado em condições de executar suas funções requeridas, sob condições de uso especificadas, quando a manutenção é executada sob condições determinadas e mediante os procedimentos e meios prescritos".

Assim, a disponibilidade é uma função resultante das condições de projeto e instalação e de manutenção de um item. Quando se fez a correlação entre a confiabilidade do serviço de manutenção e a disponibilidade de um item, considera-se um grau aceitável para a confiabilidade desse item na fase de recepção do projeto. Dessa forma, os "atributos de valor" confiabilidade do serviço de manutenção podem ser expressos:

- Número de falhas: é o número médio da ocorrência de falhas em um serviço ou instalação, para um período ou quantidade produzida.
- *MTBF* ou "*Medium time between failure*": é o tempo médio entre a ocorrência de falhas de um serviço ou instalação.

- Disponibilidade: percentualmente, exprime o quanto uma instalação ou serviço está disponível para a produção, dentro dos limites de segurança e qualidade de fabricação. É definida por:

$$Disponibilidade (\%) = \frac{MTBF}{(MTBF + MTTR)} \cdot \quad (2.1)$$

- Rendimento operacional: é a relação entre o volume efetivamente produzido e a capacidade de produção de um serviço ou instalação. As perdas de rendimento podem ter várias justificativas como perdas por não-qualidade, logística, e as perdas por não confiabilidade das instalações exprimem a parcela de não-confiabilidade do serviço de manutenção.

C – Competência (CP): A competência para as operações de manutenção está relacionada ao grau de conhecimento técnico e experiência para o planejamento, organização e execução das atividades de manutenção:

- Taxa de competência: a avaliação é definida por meio da média dos percentuais estabelecidos entre a competência necessária e a adquirida.
- Plano de treinamento: estabelecido pelas necessidades expressas na tabela de competência. É medido por meio da taxa de cumprimento do mesmo, em relação aos prazos.
- Turn-Over: corresponde ao percentual de recursos humanos substituídos na equipe para um determinado período e que afeta diretamente a competência do serviço de manutenção.

D - Custos (CT): Os custos envolvidos nas operações de manutenção, como em qualquer relação cliente-fornecedor têm dois componentes, um monetário e outro não monetário. Os custos monetários de manutenção podem ser indicados a seguir:

- Custos totais: devem ser contabilizados os gastos totais, de materiais, sejam peças de reposição, ferramentas, consumíveis, instrumentos, bem como os gastos com recursos humanos diretos e indiretos.
- Custos por unidade instalada: utilizado para expressar a contribuição da manutenção nos custos de instalação, ou seja, os gastos de manutenção por unidade instalada.

- Perda por descontinuidade de serviço: é o custo relativo à perda de serviço gerada pela ocorrência de falhas. Pode ser expresso monetariamente ou por unidade de reparo, o que acaba por traduzir o esforço psicológico, enfim o "stress" da relação manutenção-serviço produzido.

E - Consistência (CS): A consistência para as operações de serviço reflete a previsibilidade do resultado esperado, ou mesmo do seu processo de criação. Para o serviço de manutenção a consistência é percebida pela ausência de falhas, indicando que as intervenções preventivas e programadas foram realizadas e são consistentes. No caso de ocorrência de falhas, a consistência é percebida pela sua não-reincidência, ou seja, indicando a eficácia das ações tomadas.

- Taxa de manutenção preventiva: corresponde á relação entre o planejado e o executado, para as intervenções preventivas.
- Quantidade de falhas reincidentes: é o somatório, para um período, das panes ocorridas nos mesmos serviços e pelas mesmas causas.

F - Flexibilidade (FB): A flexibilidade é percebida como a capacidade de adaptação às necessidades mutantes dos clientes. A adaptação do serviço a essa variação deve ser a busca da manutenção da sinergia dos prestadores desse serviço no propósito das metas de fabricação, independentemente de uma possível variação no suprimento de recursos materiais e humanos. Pode ser medido por:

- Taxa de polivalência: ou seja, é o grau de homogeneidade em relação à competência da equipe de manutenção.

G - Acessibilidade (AB): A acessibilidade ao serviço de manutenção tem implicação sobre o tempo de resposta principalmente para as intervenções corretivas, e, portanto está estreitamente relacionada à taxa de transmissão de atendimento.

- Tempo de reatividade: é o tempo dispensado entre a localização e chamado do serviço de manutenção e a sua chegada ao local da intervenção.

De acordo com a definição da equação do valor a qual é expressa pela relação direta entre os benefícios e sacrifícios percebidos, é estabelecida a equação inicial do valor das operações de manutenção, a partir dos "atributos de valor" identificados no item anterior:

$$\text{Valor percebido} = \frac{f(VA, CF, CP, CS, FB, AB)}{f(CT)}. \quad (2.2)$$

A equação inicial aqui definida mostra de uma forma genérica, os fatores que influenciam a percepção do valor para o cliente das operações de manutenção, sem priorizá-los, ou seja, nesse momento inicial todos esses fatores têm igual importância. Observa-se que "custos - CT" são aqui definidos como um critério de percepção do valor, refletindo os sacrifícios da relação fabricação-manutenção de forma qualitativa e não quantitativa, o mesmo acontecendo para os benefícios (VA, CF, CP, FB, CS, AB), apesar de serem apresentados indicadores quantitativos para os mesmos (figura 2.17).

FATORES DE PERCEPÇÃO DO VALOR (VALUE-DRIVERS) PARA OPERAÇÕES DE MANUTENÇÃO	INDICADORES USUAIS ASSOCIADOS
Velocidade de atendimento - VA	MTTR ou tempo médio de reparo Tempo médio de intervenções programadas: corretivas e serviços de melhoria
Confiabilidade - CF	MTTR ou tempo médio de reparo MTBF ou tempo médio entre falhas Número médio de falhas Disponibilidade própria Rendimento operacional
Competência - CP	Taxa de treinamento Plano de treinamento Grau de rotatividade ou "turn-over"
Consistência - CS	Quantidade de falhas repetitivas Taxa realização manutenção preventiva
Custos - CT	Custo total de manutenção Custo de manutenção por unidade produzida Perda de produção relacionada a panes
Flexibilidade - FB	Taxa de polivalência
Acessibilidade - AB	Tempo de reatividade (intervenções corretivas)

Figura 2.17 Fatores da Percepção do Valor e Indicadores Associados, para as Operações de Manutenção.

Fonte: Adaptada de MEREDITH, (1992).

O próximo passo é justamente criar uma forma para priorizar os "atributos de valor", a fim de definir quais critérios os clientes desse serviço são mais influentes para a percepção do valor.

2 - Priorizar os "Atributos de valor" junto aos Clientes:

Nessa etapa é apresentada uma forma de avaliar e priorizar grandezas que possuem características eminentemente subjetivas, e que possam ser medidas diretamente a partir das percepções dos clientes. Dentre as diversas técnicas de avaliação de grandezas com natureza intangível, utiliza-se o método de Mudge, que é descrito a seguir.

A - O método de Mudge: Mudge foi um dos expoentes nas técnicas de análise do valor, que com a avaliação das relações funcionais, elaborou uma técnica que permite comparar cada função já definida com as outras, a fim de determinar uma importância relativa entre elas. Segundo CSILLAG, (1995) de acordo com a análise do valor, função pode ser definida como uma característica de um bem ou serviço, que atinge as necessidades do comprador e/ou usuário.

O método de Mudge é uma técnica de comparação entre as funções ou critérios de avaliação de um produto ou serviço, de forma a valorizá-las e priorizá-las. Pode ser dividido em: a definição de graduação das funções e a Tabela de comparação ponderada.

Definidas as funções ou critérios de avaliação de um produto ou serviço, precisa-se estabelecer uma escala apropriada de medição. Nesse trabalho adota-se a escala de graduação de importância ponderada, ou seja, de comparação mútua do grau de importância para o cliente: pouca importância, média e muita importância relativa.

A seguir temos a figura 2.18 que representa os fatores de avaliação do serviço de manutenção e a escala de graduação comparativa.

FATORES DE PERCEPÇÃO DO VALOR POR OPERAÇÕES DE MANUTENÇÃO	GRAU DE IMPORTÂNCIA RELATIVA
Velocidade de Atendimento - VA	
Confiabilidade - CF	n = 1 para "pouca"
Competência - CP	n = 2 para "média"
Consistência - CS	n = 3 para "muita"
Custos- CT	
Flexibilidade - FB	importância relativa
Acessibilidade - AB	

Figura 2.18 Fatores de Avaliação do Serviço de Manutenção X Escala de Graduação da Importância Relativa para o Cliente.

Fonte: Adaptada de MEREDITH, (1992).

A "Tabela de comparação ponderada" definida por Mudge consiste em uma ferramenta que permite a comparação mútua entre as funções ou nesse caso, os critérios de avaliação de um produto ou serviço, de forma a se determinar uma importância relativa entre elas.

A comparação deve ser executada cruzando-se, conforme a graduação predeterminada no item anterior, os critérios de avaliação entre si, conforme Tabela 2.8.

Agora, a partir dos critérios da percepção do valor, os "*atributos de valor*", devidamente priorizados a partir da perspectiva dos clientes, se possível reescrever a equação do valor para as operações de manutenção:

$$Valor\ percebido = \frac{f(\%VA, \%CF, \%CP, \%CS, \%FB, \%AB)}{f(\%CT)} \quad (2.3)$$

Ou seja, o valor efetivamente percebido pelo cliente é uma função percentual dos critérios que influência de avaliação nessa percepção.

Tabela 2.8 Tabela de Priorização dos Fatores de Avaliação do Serviço de Manutenção, Segundo a Técnica de Mudge.

VA	CF	CP	CT	CS	FB	AB	TOTAL	%
VA	VA1, 2, 3 ou CF1, 2, 3	VA1, 2, 3 ou CP1, 2, 3	VA1, 2, 3 ou CP1, 2, 3	VA1, 2, 3 ou CS1, 2, 3	VA1, 2, 3 ou CP1, 2, 3	VA1, 2, 3 ou CP1, 2, 3	Soma Van	%VA
	CF	CF1, 2, 3 ou CP1, 2, 3	CF1, 2, 3 ou CT1, 2, 3	CF1, 2, 3 ou CS1, 2, 3	CF1, 2, 3 ou FB1, 2, 3	CF1, 2, 3 ou AB1, 2, 3	Soma CFn	%CF
		CP	CP1, 2, 3 ou CT1, 2, 3	CP1, 2, 3 ou CS1, 2, 3	CP1, 2, 3 ou FB1, 2, 3	CP1, 2, 3 ou AB1, 2, 3	Soma CPn	%CP
			CT	CT1, 2, 3 ou CS1, 2, 3	CT1, 2, 3 ou FB1, 2, 3	CT1, 2, 3 ou AB1, 2, 3	Soma CTn	%CT
				CS	CS1, 2, 3 ou FB1, 2, 3	CS1, 2, 3 ou AB1, 2, 3	Soma CSn	%CS
					FB	FB1, 2, 3 ou AB1, 2, 3	Soma FBn	%FB
						AB	Soma ABn	%AB

Fonte: Adaptada de MEREDITH, (1992).

3- Analisar o Conflito entre os "Atributos de valor":

É preciso também se identificar os possíveis conflitos entre os critérios de percepção do valor priorizados pelo cliente. De acordo com a equação final do valor estabelecida no item anterior, os "atributos de valor" devem também ser correlacionados mutuamente, segundo o método proposto por GIANESI e CORRÊA, (1994) para evitar que uma ação sobre um benefício percebido aumente também o sacrifício percebido, e o resultado final do valor seja o mesmo, ou até inferior.

Portanto, a correlação dos "atributos de valor" deve ser sempre realizada com base na equação do valor percebido, e na figura 2.19, tem-se o resultado da análise do conflito dos critérios do valor percebido para as operações de manutenção.

	AB	FB	CS	CT	CP	CF	VA
VA				●			
CF				●			
CP		●		●			
CT	●	●					
CS							
FB							
AB							

onde: ● Conflito baseado na eq. do valor

Figura 2.19 Análise de Conflito dos "atributos de valor" das Operações de Manutenção *Ab Fb Cs Ct Cp Cf Va*.

Fonte: Adaptada de MEREDITH, (1992).

3 - Relacionar "Atributos de valor" com os Processos de Manutenção:

Definem-se como áreas de decisão de sistema de operações, como os itens de decisões gerenciais relacionados aos recursos humanos, materiais e de sistema que influenciam o desempenho de um sistema de operações. Ou, podem também ser caracterizadas como os processos críticos, que afetam diretamente a competência do sistema de operações e, por conseguinte, sua vantagem competitiva [GIANESI e CORRÊA, (1994)].

Os processos críticos que afetam o sistema de operações de manutenção, baseado na sua representação sistemática apresentada na figura 2220 desse trabalho, são:

- Gestão de materiais: relacionado à gestão das peças de reposição, consumíveis de manutenção, instrumentos específicos e ferramentas.
- Gestão de recursos humanos: referente à política de contratação, treinamento e qualificação, bem como de motivação dos recursos humanos envolvidos.
- Conceito manutenção adotado: filosofia de manutenção, se reativa, baseada em intervenções corretivas ou pró-ativa, relacionada à eliminação, antecipação e prevenção das falhas (TPM).
- Organização: definição de funções, perímetros e organograma relacionado aos cinco níveis de manutenção decorrentes do conceito de manutenção adotado.

- Sistemas de informação: relacionado aos sistemas de coleta de dados referentes a panes, registro de intervenções para formação do histórico de manutenção, apontamento de gastos com mão-de-obra e materiais.
- Gestão do cliente: interface com o cliente, na gestão da comunicação e das expectativas, bem como da co-responsabilidade sobre os resultados e do treinamento do cliente na execução de uma parte das operações (TPM).
- Planejamento/control: atividades de retaguarda relacionadas ao planejamento e controle das operações.

A correlação entre os fatores de percepção do valor priorizados e os processos críticos de um sistema de operações, define os processos prioritários, de modo a se concentrar os esforços nos aspectos de desempenho priorizados pelo cliente. O método utilizado para relacionar os critérios que influenciam a percepção do valor e os processos críticos das operações de manutenção é a matriz desenvolvida por GIANESI e CORRÊA, (1994) conforme figura 2.20.

Value Drivers / Áreas de decisão	VA	CP	CF	CT	CS	FB	AB
Gestão de materiais	■	□	■	■	■	■	■
Gestão de RH	■	■	■	■	■	■	■
Conceito manutenção	□	■	■	■	■	■	■
Organização operações	■	■	■	■	■	■	■
Sistemas de informação	■	□	■	■	□	□	■
Gestão do cliente	■	■	■	■	■	□	■
Planejamento/control	■	□	■	■	■	■	□
Onde:	□	relação fraca	■	relação média	■	relação forte	

Figura 2.20 "Atributos de valor" X Processos Críticos nas Operações de Manutenção.

Fonte: Adaptada de MEREDITH, (1992).

5- Avaliar, Junto aos Clientes, o Desempenho das Operações de Manutenção.

Define-se desempenho como o nível o qual uma meta é atingida [DWIGHT, (1995)]. O problema gerado nessa definição parece óbvio, desde que a "meta" deva ser definida e normalmente é algo de natureza subjetiva. Assim, o grande desafio de uma avaliação do

desempenho é minimizar a subjetividade de seus resultados, possibilitando ao gestor que o utiliza, ter uma noção clara das ações a serem desenvolvidas.

ROSE, (1995) declarou que a avaliação da performance é a linguagem do progresso para a organização. Isto indica onde a organização está e para onde ela deve ir, e funciona como um guia para o alcance das metas organizacionais [MEEKINGS, (1995)].

O desempenho do serviço de manutenção é avaliado pela medição não somente dos indicadores quantitativos, mas também dos fatores subjetivos que influenciam a percepção do valor para o consumidor desse serviço. Para isto, é apresentado o método de avaliação da figura 2.21, cuja sistemática é:

- Inicialmente (1ª coluna), estão indicados todos os "atributos de valor" correspondentes ao serviço de manutenção;
- Na segunda coluna deve classificar os "atributos de valor" de acordo com o percentual de criticidade K (%) encontrada na priorização por meio do método de Mudge;
- Na 3ª coluna encontram-se os indicadores associados aos "atributos de valor";
- Na 4ª coluna se devem avaliar junto ao cliente os indicadores de processo, especificando uma cotação (Cot) para cada indicador, conforme segue: Cot = 1 para avaliação "péssima", Cot = 2 para avaliação "insuficiente", Cot = 3 para avaliação "regular", Cot = 4 para avaliação "boa" e Cot = 5 para avaliação "ótima";
- Na 5ª coluna faz-se uma avaliação considerando a cotação de cada indicador e o grau de criticidade dos "atributos de valor" correspondente, multiplicando-se esses dois elementos (K e Cot);
- Por fim, a avaliação de desempenho instantânea do serviço de manutenção é realizada por meio da equação:

$$Avaliação\ de\ desempenho(\%) = \frac{\sum(K \times Cot)}{\sum(K \times 5)}. \quad (2.4)$$

"VALUE DRIVER"	CRITICIDADE (K)	INDICADOR ASSOCIADO	COTAÇÃO (Cot)	AVALIAÇÃO (K X Cot)
Velocidade Atendimento VA	%VA	Tempo de resposta intervenções corretivas	1 a 5	
		Tempo de resposta intervenções programadas	1 a 5	
		Tempo de resposta serviços melhoria	1 a 5	
Confiabilidade - CF	%CF	MTTR ou tempo médio de reparo	1 a 5	
		MTBF ou tempo médio entre falhas	1 a 5	
		Número falhas por mil unidades produzidas	1 a 5	
		Disponibilidade própria	1 a 5	
Competência - CP	%CP	Taxa de competência	1 a 5	
		Plano de treinamento	1 a 5	
		Índice de rotatividade ou "turn-over"	1 a 5	
Consistência - CS	%CS	Quantidade de panes repetitivas	1 a 5	
		Taxa manutenção corretiva x preventiva	1 a 5	
Custos - CT	%CT	Custo total de manutenção	1 a 5	
		Custo de manutenção por unidade produzida	1 a 5	
		Perda de produção relacionada a falhas	1 a 5	
Flexibilidade - FB	%FB	Taxa de polivalência	1 a 5	
Acessibilidade - AB	%AB	Tempo de resposta intervenções corretivas	1 a 5	

Figura 2.21 Avaliação de Desempenho nas Operações de Manutenção.

Fonte: Adaptada de MEREDITH, (1992).

6- Estabelecer o Plano de Ações Prioritárias:

O plano de ações de gestão do valor nas operações de manutenção decorre naturalmente do resultado das etapas anteriores do presente modelo. Inicialmente são dispostos todos os "atributos de valor" com os graus de prioridade estabelecidos pelos clientes. Em seguida são relacionados os objetivos desdobrados para o serviço de manutenção, de acordo com o "atributo de valor" correspondente. Por sua vez, é indicada a correlação entre o "atributo de valor" e os processos críticos do serviço de manutenção e em seguida, são apresentados os indicadores associados aos "atributos de valor" e a cotação.

A partir da correlação dos "atributos de valor", dos objetivos desdobrados relacionados, dos processos críticos e dos indicadores de desempenho com as respectivas cotações, no sentido vertical são estabelecidas as ações prioritárias, com os respectivos prazos e responsáveis.

Finalmente, para verificação da eficiência do plano estabelecido, faz-se uma avaliação periódica de desempenho no sentido de se medir numericamente a evolução do valor percebido pelo cliente das operações de manutenção de telecomunicações.

7- Verificar os Resultados:

Estabelecido e executado o plano de ações prioritárias de operacionalização do valor, uma nova avaliação de desempenho é necessária, no intuito da verificação da eficácia das ações e do valor resultante alcançado.

O modelo de gestão do valor para as operações de manutenção em telecomunicações, traz uma forma de identificar o que realmente o cliente valoriza nessas operações. A identificação dos fatores que influenciam a percepção do valor é o ponto de partida para a aplicação do modelo de gestão, permitindo o estabelecimento da equação do valor dessas operações. Entretanto, a definição de tais fatores somente pode ser realizada havendo um profundo conhecimento das operações em questão, fato que gerou a necessidade da representação sistemática desse serviço.

Porém, somente a identificação dos "atributos de valor" não é suficiente para a correta gestão das expectativas do cliente. A partir disto, deve ser estabelecido um método de priorização de tais fatores, que definirá quais critérios são prioritários na formação da percepção do valor, estabelecendo-se assim a equação final priorizada do valor percebido do serviço de manutenção.

Essa equação é também utilizada para a análise dos possíveis conflitos entre os "atributos de valor" formadores dessa equação, a fim de se evitar uma proposição de ações ineficazes quanto à evolução do valor percebido.

Para a elaboração do plano de ações prioritárias de gestão do valor, identificaram-se os processos críticos do serviço de manutenção, que em seguida foram correlacionados com os fatores influenciadores da percepção do valor. Essa tomada de decisões considera também uma avaliação de desempenho instantânea das operações, baseada nos "atributos de valor" priorizados, bem como em uma cotação para os indicadores associados a esses fatores.

2.3. Descoberta de Conhecimento em Base de Dados

Introdução

Os constantes avanços na área de tecnologia da informação têm viabilizado uma elevada quantidade de armazenamento de dados nunca antes imaginado. Em termos históricos divide a descoberta de conhecimento em base de dados em quatro gerações [GOLDSCHMIDT et al., (2005)].

A primeira geração apareceu em meados dos anos 80. Consistia em ferramentas de análise voltadas a uma única tarefa, sem portar as demais etapas do processo [PIATETSKY-SHAPIRO, (2000)]. Essas tarefas incluíam em geral, a construção de classificadores usando ferramentas de construção de regras (por exemplo, C4.5) ou de redes neurais (*backpropagation*), a descoberta de grupos (*clusters*) nos dados (*k-means*) ou ainda a visualização dos dados.

O desenvolvimento de ferramentas chamadas de “suítes” foi o propulsor da segunda geração, essas ferramentas eram dirigidas ao fato que o processo de descoberta do conhecimento requer múltiplos tipos de análise de dados. São exemplos de “suítes” *softwares* como o *SPSS*, *Clementine*, *Intelligent Miner*, e *SAS Enterprise Miner*. Essas ferramentas permitem ao usuário realizar diversas tarefas suportando transformações de dados.

A terceira geração surgiu no final dos anos 90 essa geração foi orientada para a resolução de um problema específico diretamente por um usuário, não necessitando mais do auxílio de especialistas em análise de dados. Nessa geração as interfaces são orientadas para os usuários e procura-se esconder toda complexidade da mineração de dados. O *software* HNC Falcon especializado em fraudes no cartão de crédito é um exemplo dessa geração.

A quarta geração compreende o desenvolvimento e a aplicação de técnicas e ferramentas que auxiliam o homem na própria condução do complexo processo de descoberta de conhecimento em base de dados [GOLDSCHMIDT et al., (2005)].

O conceito de descoberta de conhecimento em bases de dados pode ser resumido como o processo não-trivial de identificar padrões novos, válidos, potencialmente úteis e, principalmente, compreensíveis em meio às observações presentes em uma base de dados. Contudo o objetivo último da descoberta do conhecimento em bases de dados não é o de

simplesmente encontrar padrões e relações em meio à imensa quantidade de informação disponível em bases de dados, e sim a extração de conhecimento inteligível e imediatamente utilizável para o apoio às decisões [FAYYAD, (1996)].

As bases de dados podem ser tratadas seguindo um processo de descobertas do conhecimento em etapas conforme ilustração de [GOLDSCHMIDT et al., (2005)]. A Figura 2.22 ilustra o ciclo de descoberta do conhecimento em bases de dados e suas etapas.

Nesse caso a origem diversa dos dados que serão utilizados, coletados em diferentes instantes de tempo em um mesmo lugar e sistema, cria um esforço inicial de consolidação e agrupamento de toda a informação que irá servir de base para o processo. Sendo um ponto crítico a compreensão do negócio e do ambiente em que os dados estão inseridos para o entendimento dos mesmos. Dada essa diversidade e heterogeneidade dos dados, esforços de pré-processamento e limpeza dos mesmos são cruciais na geração de dados que possam vir a ser trabalhados em busca de conhecimento útil.

É essencial que seja realizada a investigação de inconsistências e problemas devido a diferenças de escalas, assim como o tratamento de valores fora da normalidade (*outliers*) e observações errôneas [GOLDSCHMIDT et al., (2005)]. Realizadas essas tarefas iniciais, que tornam os dados tratáveis e homogêneos, mineração dos dados pode ser iniciada, na busca por padrões e relações que com sentido e sejam úteis para o problema a ser resolvido ou objetivo a ser alcançado.

Por último, a interpretação, compreensão e aplicação dos resultados encontrados é o passo que torna o conhecimento adquirido por meio de bases de dados um real insumo para o apoio às decisões.

Na Figura 2.22 ilustra-se como o conhecimento inicialmente essa na forma de dados sendo passada para uma segunda fase denominada de informações e finalmente gera-se dos dados o conhecimento. Todo esse processo de transformação do conhecimento é útil para a pesquisa se for utilizado pelas cinco etapas que se encontram no ápice da figura.

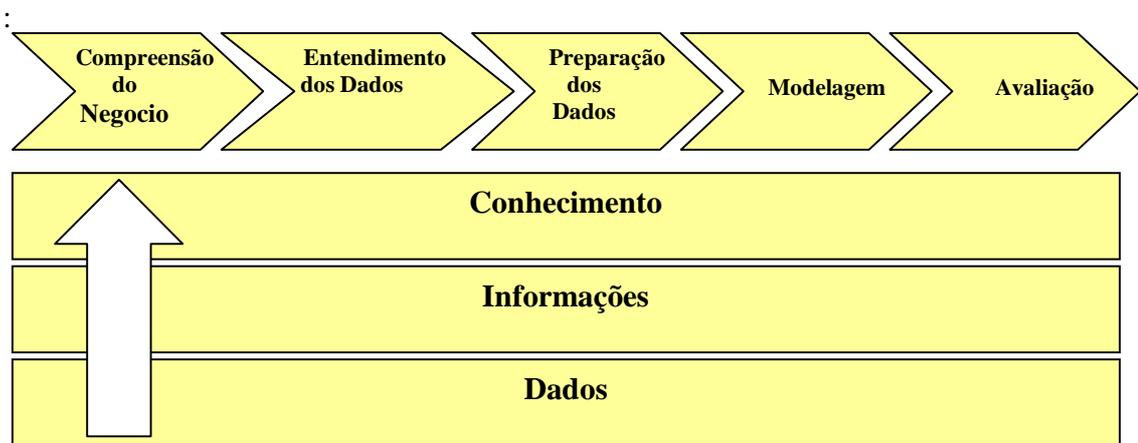


Figura 2.22 Hierarquia da descoberta e conhecimento de dados.

Fonte: Adaptado pelo autor de GOLDSCHMIDT, (2005).

2.3.1 Dados

Para que qualquer conhecimento seja gerado a partir de dados, o primeiro passo é que tais dados existam e estejam disponíveis em algum lugar [BOZDOGAN, (2003)]. Empresas e outras organizações fazem uso de uma grande infra-estrutura de tecnologia da informação para garantir a disponibilidade de dados, mas nem sempre a obtenção coleta e armazenamento desses dados são de forma simples, para se usar adequadamente as informações desses dados para apoio à decisão como foi dito anteriormente, as etapas da figura 2.22 devem ser cumpridas.

2.3.1.1 Data Warehouse

Uma das formas de organização, consolidação e disponibilidade de dados são os chamados “armazéns de dados” (*data warehouse*) [BALLOU, (1999)]. Uma *data warehouse* é um grande repositório de bases de dados alimentado por muitos sistemas operacionais. *Data warehouse* nascem da necessidade de empresas e organizações possuírem uma visão de dados e operações centralizadas em um único ponto, ao invés de ter seus dados espalhados por diversos locais ou departamentos sem muita coesão e por vezes incomunicáveis entre si. Um dos objetivos primordiais de *data warehouse* é garantir a integridade e consistência de todos os dados coletados dos sistemas operacionais de uma empresa, além do acesso seguro e

consistente de todas as partes da empresa a esses dados. Em geral, ao serem encaminhados para um *data warehouse*, dados operacionais são limpos e transformados em uma primeira instância, principalmente para garantir que dados de diferentes fontes e formatos passem então a possuir as mesmas definições e obedeçam às mesmas regras. Por vezes algumas das transformações padrões realizadas nos dados para se encaixarem no *data warehouse* podem danificar ou até mesmo destruir informação que poderia vir a ser valiosa no processo de descoberta de conhecimento. Normalização, agregação e sumarização dos dados são algumas dessas transformações que podem vir a atrapalhar a análise e mineração dos dados.

2.3.1.2 Análise Múltipla de Dados

Segundo HAIR, (2005) não é fácil definir análise multivariada de dados. De um modo geral, ela se refere aos métodos estatísticos que simultaneamente analisam múltiplas medidas sobre cada medida ou objeto sob investigação. Qualquer análise simultânea de mais de duas variáveis de certo modo pode ser considerada análise multivariada. Os conceitos básicos de Análise Multivariada são originários dos estudos de estatística univariada e bivariada, a extensão para o domínio multivariado introduz conceitos adicionais e questões que tem particular relevância, os quais se citam abaixo:

1 - Variáveis estatísticas – É uma combinação linear de variáveis com os pesos empiricamente determinados. As variáveis estatísticas são especificadas, sendo os pesos determinados pela técnica multivariada para atingir um objetivo específico. Uma variável estatística de n variáveis ponderadas (X_1 até X_n) pode ser enumerada matematicamente como:

$$\text{Valor da Variável Estatística} = W_1X_1 + W_2X_2 + \dots + W_nX_n . \quad (2.5)$$

Em que: X_n é a variável observada e W_n é o peso determinado pela técnica multivariada utilizada.

O resultado como foi visto anteriormente é um único valor que representa uma combinação do conjunto inteiro de variáveis que melhor atinge o objetivo da análise multivariada específica.

2 - Escalas de Medidas – A análise de dados envolve a participação, a identificação e a medida de variação em um conjunto de variáveis, seja entre elas ou entre uma variável dependente e uma ou mais variáveis independentes. A mensuração é importante para representar com precisão o conceito de interesse e é fundamental na seleção do método de análise apropriado.

3 - Tipos básicos de dados - Não-métricos (qualitativos) e métricos (quantitativos). Dados não-métricos são atributos, características ou propriedades categóricas que identificam ou descrevem um objeto. Descrevem diferenças em tipos ou espécies, indicando a presença ou ausência de uma determinada característica. Já as medidas de dados métricos são feitas de modo que os indivíduos podem ser identificados como diferenciados em quantia ou grau. Variáveis metricamente medidas refletem quantidade relativa ou grau. Medidas métricas são apropriadas para os exemplos que envolvem quantia ou magnitude, como nível de satisfação ou indicadores de qualidade.

4 - Erro de medida - É o erro em que os valores observados não são representativos dos valores verdadeiros. O impacto do erro de medida é o acréscimo de ruído as variáveis observadas ou medidas. O objetivo de reduzir o erro de medida pode seguir diversos caminhos. Ao avaliar o grau de erro presente em qualidade medida, deve-se levar em conta validade e a confiabilidade da medida e a garantia da validade da medida começa com uma compreensão direta do que deve ser medido e então realizar a medida tão correta e precisa quanto possível.

5 - Significâncias estatísticas versus poder estatístico – A maioria das técnicas multivariadas, exceto análise de agrupamento e escalonamento multidimensional, são baseados na inferência estatística dos valores. Interpretar inferências estatísticas requer que sejam especificados os níveis de erro estatísticos aceitáveis. A abordagem mais comum é especificar o nível do erro tipo I, conhecido como alfa (α). O erro alfa é a probabilidade de rejeitar a hipótese nula quando a mesma é verdadeira, ou seja, a chance de o teste exibir significância estatística quando na verdade não existe tal significância. Ao especificar um nível alfa, se estabelecido os limites para o erro, especificando a probabilidade de se concluir que a significância exista quando na verdade essa não ocorre. Quando se especifica o nível de erro tipo I, também se determina um erro associado, chamado de erro tipo II ou beta (β). O erro beta é a probabilidade de não rejeitar a hipótese nula quando na realidade essa é falsa.

6 - Poder – É a probabilidade significância estatística ser indicada se estiver presente, ou seja, o poder é igual a $(1-\beta)$. Apesar de a especificação de alfa estabelecer o nível de significância estatística aceitável, é o nível de poder que determina a probabilidade de sucesso em encontrar as diferenças se elas realmente existem. Portanto reduzir o erro alfa é reduzir o poder do teste estatístico. Assim sendo, se deve trabalhar com o equilíbrio entre o nível alfa e o poder resultante.

2.3.2 Caracterização do Processo de Descoberta de Conhecimento em Base de Dados

O processo de descoberta de conhecimento em base de dados pode ser dividido em três etapas: Pré-processamento, Mineração de Dados e Pós-processamento [GOLDSCHMIDT et al., (2005)].

A etapa de pré-processamento compreende as funções relacionadas à captação, à organização e ao tratamento de dados. São na etapa de pré-processamento que se preparam os dados para a etapa de mineração onde é realizada a busca efetiva por conhecimentos úteis no contexto da pesquisa. A etapa de pós-processamento abrange o tratamento do conhecimento obtido na mineração.

O tratamento dispensado na fase de pós-processamento nem sempre é necessário, porém tem como objetivo a avaliação da utilidade do conhecimento [FAYYAD, (1996)]. Todo conjunto de dados pode ser observado sob os aspectos intencionais e extensionais, sendo o aspecto intencional referente à estrutura do conjunto de dados, encontra-se nesse aspecto as características ou atributos dos dados, já os registros representam o aspecto extensional [DATE, (1991)].

Pode-se estender o conceito das três etapas para um conceito onde se coloca uma nova etapa conhecida como controle, nessa etapa o especialista tem a finalidade de corrigir distorções em quaisquer das etapas anteriores e deve ser realizado por um especialista com elevado conhecimento do problema [MALHOTRA, (2004)]. Deve-se até mudar o rumo da pesquisa caso sinta necessidade da sua interferência para que a pesquisa atinja a utilidade do conhecimento desejado, conforme figura 2.23.

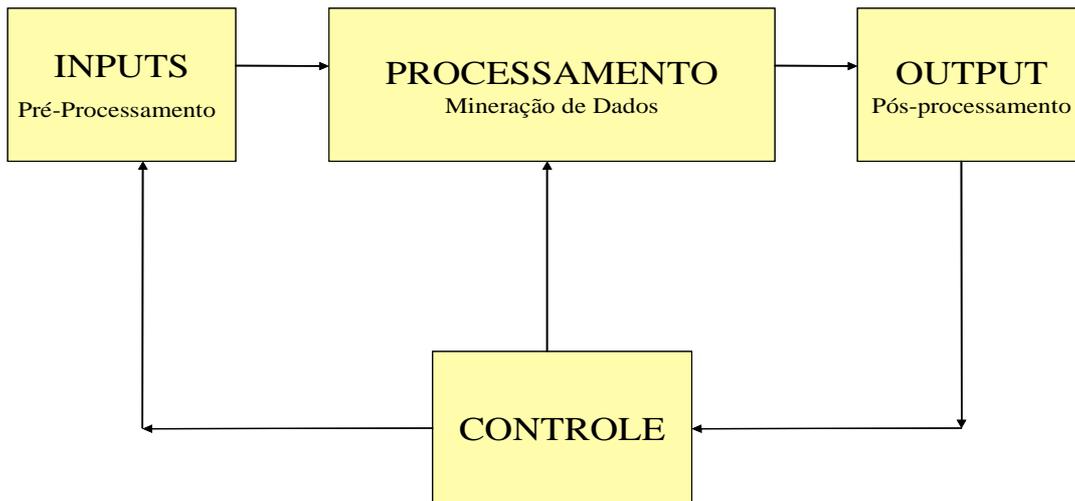


Figura 2.23 Etapas operacionais no processo de descoberta do conhecimento.

Fonte: Adaptado pelo autor de GOLDSCHMIDT, (2005).

2.3.2.1 Pré-processamento

A etapa de pré-processamento compreende, conforme foi apresentado anteriormente, às funções relacionadas à captação, organização, tratamento e a preparação dos dados para etapa seguinte de mineração, possuindo relevância no processo de descoberta, pois compreende desde a correção de dados errados até o ajuste da formatação a serem utilizadas. A seguir se descreve as etapas do pré-processamento:

- Seleção dos Dados
- Tratamento e Limpeza dos Dados
- Normalização dos Dados
- Enriquecimento dos Dados

A - Seleção dos Dados

Essa etapa compreende na identificação e no planejamento de todas as atividades para se chegar ao ponto final de carga dos dados no ambiente de mineração de dados. A seleção deve levar em consideração quais informações dentre os dados existentes são realmente relevantes para a pesquisa.. Dependendo dessa escolha, os dados serão formatados de

maneiras diferentes. O primeiro requisito para que a seleção seja bem sucedida é possuir dados de qualidade.

O processo de seleção dos dados é de grande relevância na busca do conhecimento em bases de dados. Mesmo já existindo *data warehouse* com os dados disponíveis para análise e já pré-processados, é essencial criar-se uma representação dos dados que satisfaça os objetivos da análise de dados a ser realizada e que se encaixe de forma ótima na resolução do problema enfrentado [FERREIRA, (2005)].

O conceito de seleção dos dados engloba: consultas (*queries*) iniciais a *data warehouse* ou outros repositórios de dados em busca dos dados procurados; consolidação de toda a informação de interesse em um local ou base única [CONTRERAS (2002)]. O processo de seleção dos dados tem uma razão muito maior para existir do que simplesmente fornecer uma base para os modelos a serem utilizados: ao selecionar os dados, o responsável pela análise também é “preparado” pelos dados [PYLE, (1999)].

Nesse processo o maior beneficiado é o pesquisador, pois o mesmo ao despender esforço para obter a melhor representação possível para os dados, convive com as outras formas de visualizar e detalhes. Compreendo com isso melhor o problema que está sendo estudado, o que contribui em etapas futuras para um melhor desempenho, tanto do mesmo quanto dos dados.

B - Tratamento e Limpeza

Limpeza de dados visa detectar e remover anomalias presentes nos dados com o objetivo de aumentar e melhorar a sua qualidade. Tipicamente o processo de limpeza de dados não pode ser executado sem o envolvimento de um perito no negócio ao qual correspondem os dados, uma vez que a detecção e correção de anomalias requerem conhecimento especializado.

A limpeza dos dados envolve uma verificação da consistência das informações, a correção de possíveis erros e o preenchimento ou a eliminação de valores nulos e redundantes. Nessa fase são identificados e removidos os dados duplicados e corrompidos. A execução dessa fase corrige a base de dados eliminando consultas desnecessárias que seriam executadas pelos modelos e que afetariam o seu desempenho. A procura de valores absurdos que não deveriam existir na base simplesmente por serem impossíveis é um das atividades

desempenhas na limpeza de dados, na prática. No esforço para limpeza e consistência dos dados, os campos com valores absurdos, mesmo sendo raros, devem ser preenchidos com valores possíveis, utilizando-se, por exemplo, médias ou medianas da variável. Outra opção seria a eliminação do registro que contém tal valor. A filosofia por trás dessas ações é evitar que tal valor atrapalhe a compreensão dos dados pelos modelos, levando-os a tomar conclusões errôneas. Outro caso interessante de limpeza de dados é o tratamento de valores ausentes (*missing*). Se o número de observações ausentes for significativo, o desempenho de grande parte dos modelos de análise de dados pode ser seriamente comprometido.

Para lidar com valores ausentes, em geral utiliza-se uma das seguintes abordagens: Ignorar a descrição do indivíduo ou mesmo eliminar o descritor e preencher os valores ausentes manualmente.

C - Normalização dos Dados

Essa etapa consiste em ajustar a escala dos valores de cada atributo de forma que os valores fiquem em pequenos intervalos. Tal ajuste faz-se necessário para evitar que alguns atributos, por apresentarem uma escala de valores maior que os outros, influenciem de forma tendenciosa em determinados métodos de mineração de dados [GOLDSCHMIDT et al., (2005)]. Existem muitas maneiras de normalização de dados, Linear, Por Desvio Padrão, Pela Soma dos Elementos, Pelo Valor Máximo dos Elementos e Normalização por Escala Decimal. A seguir algumas dessas técnicas são conceituadas:

- Normalização Linear – Essa técnica também é conhecida como interpolação linear, consiste em considerar os valores mínimos e máximos de cada atributo no ajuste da escala. Mapeia os valores de um atributo no intervalo fechado de 0 a 1. Mantém distâncias entre os dados normalizados que sejam proporcionais as distâncias entre os dados originais. Essa técnica é recomendada apenas quando se tem certeza que o domínio dos atributos está entre os valores mínimos e máximos considerados.
- Normalização por Desvio Padrão – Normalmente conhecida como *Z-Score* ou *Z-Mean*, a padronização por desvio padrão considera a posição média dos valores de um atributo, assim como os graus de dispersão desses valores em relação à posição média. Essa técnica de normalização é útil quando os valores mínimo e máximo são desconhecidos.

- Normalização Pela Soma dos Elementos – Consiste em dividir cada valor do atributo que esteja sendo normalizado pelo somatório de todos os valores de tal atributo. Uma desvantagem é que certos valores podem ser muito pequenos comparado com os demais.
- Normalização Pelo Máximo dos Elementos. – Essa técnica consiste em dividir cada valor do atributo que esteja sendo normalizado pelo maior valor dentre todos os valores de tal atributo.
- Normalização por Escala Decimal - Essa técnica consiste em realizar o processo de normalização por meio do deslocamento do ponto decimal dos valores dos atributos a ser normalizados. O número de casas decimais depende do maior valor absoluto do atributo em questão.

D - Enriquecimento dos Dados

Essa etapa consiste em conseguir agregar mais informações aos registros existentes para que esses forneçam mais elementos para a pesquisa. Estão incluídas nessa operação todas as iniciativas que envolvam captação e transformações junto à fonte de dados original. Difere-se da etapa de limpeza, pois não se restringe a preencher informações ausentes. Busca-se, agregar novas informações, muitas vezes essa etapa esbarra no alto custo de implementação.

Uma das formas de enriquecimento de dados pode ser executada mediante a incorporação de informações fornecidas por outros sistemas ou fontes, sendo muito comum à importação de informações advindas de outras bases em adição aos dados das bases já existentes. Tais informações podem ser úteis em um contexto de análise fornecendo um indicador para uma eventual tomada de decisão.

2.3.2.2 Mineração de Dados

O termo mineração de dados (*data mining*) define a exploração e análise de grandes quantidades de dados com o objetivo de encontrar padrões, regras e relações interessantes e significativas para algum fim [BERRY, (2000)]. Sendo essa a principal etapa do processo, nela ocorre à busca efetiva por conhecimento por esse motivo entende-se que a etapa de

mineração engloba as mais variadas áreas do conhecimento sendo válida, tanto na compreensão de clientes de uma empresa quanto no desenvolvimento de uma nova vacina para alguma doença.

Existem dois tipos de mineração de dados: a direta e a indireta. A mineração de dados direta tenta explicar ou categorizar a variável alvo, definida como receita proveniente de algum esforço de vendas ou a resposta aos indicadores de qualidade. Em geral toma a forma de modelagem preditiva, onde se sabe o que se quer prever. Mineração de dados indireta, por sua vez, procura encontrar padrões ou similaridades entre grupos de registros de uma base de dados sem o uso de um alvo particular ou de alguma coleção de classes pré-definida. Ambas as abordagens não são mutuamente exclusivas e, na verdade, freqüentemente as tarefas de mineração de dados envolvem as duas. Esse trabalho é focado na Mineração de Dados direta, ou seja, existe um alvo e ele é delineado uma modelagem desse indicador.

Existem muitas técnicas de modelagem em mineração de dados. A escolha depende, muitas vezes, do tipo da tarefa de mineração a ser realizada [GOLDSCHMIDT et al., (2005)].

2.3.2.2.1 Tarefas de Mineração de Dados

Uma tarefa equivale a uma operação de descoberta de conhecimento estando ligada a etapa de mineração de dados. Conforme visto anteriormente essas tarefas podem ser divididas em diretas e indiretas.

Os métodos de mineração de dados podem ser classificados pela função que executam ou de acordo com a classe de aplicação em que podem ser usados [DILLY, (1995)]. Cada classe de aplicação tem como base um conjunto de algoritmos a serem utilizados na extração de relações relevantes de uma base de dados, diferindo uma das outras quanto aos tipos de problemas que o algoritmo é capaz de resolver.

Nessa sessão é apresentada uma breve introdução as principais tarefas de mineração de dados: Associação, Classificação, *Clustering*, Regressão e Sumarização. Assim sendo descreve-se abaixo as tarefas mais comuns na realização do processo de descoberta de conhecimento.

- Classificação
- Agrupamento por afinidade ou regras de associação
- Regressão
- *Clustering*
- Sumarização

Duas dessas tarefas (classificação, regressão) consistem em exemplos de mineração de dados direta. As outras três são exemplos de atividades da mineração de dados indireta.

A - Classificação

Consiste em descobrir a função que mapeie um conjunto de registros em um conjunto de rótulos categóricos predefinidos denominados classes que, uma vez descoberta, tal função pode ser aplicada a novos registros de forma a prever a classe em que tais registros se enquadram.

Essa modalidade é também conhecida como regras de classificação, decisão supervisionada, aprendizado supervisionado ou processo direto [GROTH, (1997)]. Para classificar é necessário selecionar um atributo alvo, chamado variável dependente ou classe, cujo valor é usado para elaborar regras de classificação e as variáveis independentes ou atributos preditores.

A classificação utiliza dados sobre o passado para encontrar padrões significantes de forma a induzir regras sobre o futuro, isto é, regras que predizem o valor do atributo alvo, pela combinação dos valores dos atributos preditores.

O processo inicia-se com um conjunto de treinamento e com os registros pré-classificados espera-se associar cada inclusão a um código de classe, fundamentado nos valores dos atributos preditores. O sistema deve inferir regras para classificar e encontrar a descrição da classe. Ao final do processo, tem-se um modelo da base de dados capaz de classificar um número maior de registros. A precisão do resultado da classificação é medida pela taxa de erro que é o percentual de registros classificados incorretamente [BERRY et al., (2000)].

B - Agrupamento Por Afinidade ou Regras de Associação

Abrange a busca por itens que frequentemente ocorram de forma simultânea em transações do banco de dados. Um exemplo clássico e didático da aplicação dessa tarefa é na área de *marketing* durante um processo de descoberta de associação em sua vasta base de dados. A associação ou afinidade de grupos visa combinar itens importantes, tal que, a presença de um item em uma determinada transação pressupõe a de outro na mesma transação. Isto foi inicialmente proposto por AGRAWAL em 1993.

A tecnologia possibilitou às organizações coletar e armazenar grandes quantidades de dados, como é o caso da tecnologia de código de barras sobre os dados de venda [AGRAWAL et al., (1992)]. Segundo a revista *Business Miner* (1997) as grandes redes varejistas estudam as compras dos clientes para descobrir quais as vendas são normalmente realizadas ao mesmo tempo, chamando isso de *market basket analysis*. Essa análise pode determinar, por exemplo, os produtos que devem estar expostos juntos, objetivando incrementar as vendas.

A regra de associação é uma expressão representada na forma $X \Rightarrow Y$ (X implica em Y), em que X e Y são conjuntos de itens da base de dados; X é o antecedente da regra (lado esquerdo) e Y é o conseqüente da regra (lado direito) e pode envolver qualquer número de itens em cada lado da regra [DILLY, (1995)]. O significado dessa regra é que as transações da base que contêm X tendem a conter Y . Um exemplo prático é afirmar que "30% dos registros que contêm X também contêm Y ; 2% dos registros contêm ambos" [AGRAWAL et al., (1992)].

A regra de associação possui dois parâmetros básicos: o suporte e a confiança. Esses parâmetros limitam a quantidade de regras que serão extraídas e descrevem a qualidade delas [Para GOLDSCHMIDT, (2005)]. Considerando que os conjuntos de itens X e Y estão sendo analisados, o **suporte** é definido como a fração de registros que satisfaz a união dos itens no conseqüente (Y) e no antecedente (X), correspondendo à significância estatística da regra [AGRAWAL et al., (1992)]. A **confiança** é expressa pelo percentual de registros que satisfaz o antecedente (X) e o conseqüente (Y), medindo a força da regra ou sua precisão [AGRAWAL et al., (1992)]. No exemplo anteriormente citado, 30% é o fator de confiança e 2% é o suporte da regra. BERRY & LINOFF, (2000) definem a confiança como a freqüência com que o

relacionamento mantém-se verdadeiro na amostra de treinamento e o suporte como a frequência com que a combinação acontece. Assim, uma associação pode se manter 100% do tempo e ter a mais alta confiança, porém pode ser de pouca utilidade se a combinação ocorrer raramente.

O problema das regras de associação é encontrar todas as que possuem o suporte e a confiança acima de um determinado valor mínimo, pois, na prática os usuários normalmente estão interessados somente num subconjunto de associações [AGRAWAL et al., (1992)]. Um dos algoritmos mais referenciados para esse método é o *Apriori*, nas diversas variações, tais como, o *Apriori*Tid, DHP e *Partition*.

C - Regressão

Essa tarefa compreende, fundamentalmente, a busca por funções lineares ou não, que mapeiam os registros de um banco de dados em valores reais. Essa tarefa é similar à tarefa de classificação, sendo restrita apenas a atributos numéricos [GOLDSCHMIDT, (2005)]. A regressão linear é a forma mais simples de regressão, em que a função a ser abstraída a partir dos dados é uma função linear. O número de variáveis, ou atributos, envolvidos no problema varia de um caso para outro. O caso mais simples é conhecido como regressão linear bivariada, na qual uma variável aleatória Y , denominada de variável dependente é função de uma outra variável denominada independente.

$$Y = \alpha + \beta X . \quad (2.6)$$

Nesse caso a variância da variável Y é assumida como constante, e α e β são os coeficientes de regressão linear. Esses coeficientes podem ser obtidos a partir dos dados analisados pelo método dos mínimos quadráticos, que busca minimizar o erro entre os dados reais e os dados estimados pela função.

A regressão linear múltipla é uma extensão da regressão linear bivariada envolvendo mais de uma variável independente. Nesse tipo de regressão, a variável dependente Y deve ser modelada como função linear de um vetor de características multidimensional, conforme equação a seguir.

$$Y = \alpha + \beta X + \beta X + K \beta X. \quad (2.7)$$

O método dos mínimos quadráticos pode ser estendido para obter os coeficientes para a regressão múltipla [HAN&KEMBER, (2001)].

Existem muitos problemas em que os dados não apresentam dependência linear entre si. Nesses casos, podem ser aplicadas técnicas de regressão não linear. Pode-se citar como exemplo a regressão polinomial na qual se adiciona termos polinomiais à expressão do modelo linear. Assim, aplicando transformações às variáveis, um modelo não linear pode ser convertido em um modelo linear, que pode então, ser resolvido pelo método dos mínimos quadráticos.

D – Clustering

Esse método é utilizado para separar os registros de uma base de dados em subconjuntos ou clusters, de tal forma que os elementos de um cluster compartilhem de propriedades comuns que os distingam de elementos em outros clusters [GOLDSCHMIDT, (2005)]. O objetivo nessa tarefa é maximizar similaridades *intra-cluster* e minimizar similaridades *inter-cluster*.

Diferente da tarefa de classificação, que tem rótulos pré-definidos, a *clustering* precisa automaticamente identificar os grupos de dados aos que o usuário deverá atribuir rótulos [FAYYAD et al., (1996)].

É um exemplo de aprendizado não supervisionado ou indireto, cujo objetivo é agrupar tipos similares de dados ou identificar exceções [GROTH, (1997)]. O sistema tem que descobrir suas próprias classes, isto é, agrupar os dados e descobrir subconjuntos de objetos relacionados ao conjunto de treinamento, encontrando descrições de cada um desses subconjuntos [DILLY, (1995)]. Um cluster pode ser definido como um conjunto de objetos agrupados pela similaridade ou proximidade e a *clustering* como “a tarefa de segmentar uma população heterogênea em um número de subgrupos (ou clusters) mais homogêneos possíveis, de acordo com alguma medida” [BERRY&LINOFF, (2000)].

Quando o processo é bem sucedido, os objetos do cluster têm alta homogeneidade interna e alta heterogeneidade externa. Um exemplo disso é a geração de clusters de sintomas

de pacientes, que podem indicar diferentes doenças baseadas nas suas características. Na *clustering*, diferentemente da classificação, não há classes pré-definidas. Na classificação, a população é subdividida e associa cada registro a uma classe pré-definida, com base no modelo desenvolvido por meio de treinamento e exemplos pré-classificados. A *clustering* é mais geral e frequentemente realizada como primeira etapa de outros métodos de Mineração de Dados ou de modelagem. Assim, aplica-se o modo direto para reconhecer relações nos dados e o indireto para explicar essas relações [BERRY&LINOFF, (2000)].

É aplicada em atividades de *marketing* para identificar os segmentos de mercado, para encontrar estruturas significantes nos dados e na descoberta de fraudes ou dados incorretos [GROTH, (1997)].

E - Sumarização

A tarefa de sumarização, também é conhecida como descrição de conceitos, consiste em identificar e apresentar, de forma concisa e compreensível, as principais características dos dados contidos em um conjunto analisado. São exemplos de aplicações envolvendo a tarefa de sumarização:

- Identificar as características dos assinantes de uma revista que reside em uma região do país, identificando a faixa salarial, o nível de escolaridades dos clientes e se possui ou não residência própria;
- Descrever o perfil dos meninos de rua, a faixa etária, se eles usam ou não drogas, e se possuem ou não irmãos.

Um conceito normalmente se refere a uma coleção de dados com pelo menos uma característica em comum. Por exemplo, clientes inadimplentes, pacientes com cardiopatias, alunos de graduação dentre muitos. Sendo assim não se pode definir sumarização como uma simples enumeração dos dados. Busca-se gerar descrições para características resumidas dos dados e possível comparação entre eles. Tais descrições são denominadas descrições de classe, quando o conceito a ser descrito se refere a uma classe de objetos.

A descrição de conceitos pode ser interpretada como uma generalização dos dados a partir das características mais relevantes dentre os registros analisados. Um método muito utilizado na descrição de conceitos denomina-se decisão orientada a atributos. Esses método

consiste da análise de medidas da teoria da informação, faz parte do algoritmo tradicional C4.5.

2.3.2.2 Ferramentas de *Data Mining*.

Atualmente existe uma grande variedade de produtos comerciais para mineração de dados. Tem-se nesse item o objetivo de apresentar algumas das principais ferramentas utilizadas para auxiliar na execução do processo de Mineração. Entre algumas das principais dificuldades para criação de uma ferramenta de mineração pode-se citar a necessidade de manipulação de grandes e heterogêneos volumes de dados, o tratamento de resultados representados em diferentes formatos e a dificuldade de interação de diversos algoritmos específicos.

A – *Weka*

O *Weka* é uma ferramenta de código aberto, aplicável em mineração de dados, flexível, desenvolvida na linguagem Java pelo curso de ciência da computação da Universidade de *Waikato* na Nova Zelândia. A utilização da ferramenta pode ser realizada de diversas maneiras. Possui quatro diferentes implementações de interface, que permitem que todos os seus algoritmos sejam chamados diretamente via código Java. As interfaces são:

- *Simple Client* – Nessa interface, a interação do usuário com o *Weka* ocorre por meio de linhas de comando. Requer um profundo conhecimento do programa, porém é extremamente flexível e ágil para usuários avançados;
- *Explorer* – Trata-se da interface de utilização mais comum, e enquadra separadamente as etapas de pré-processamento, mineração de dados e pós-processamento;
- *Experimenter* – Constitui um ambiente de experimentação, em que testes estatísticos podem ser conduzidos a fim de avaliar o desempenho de diferentes algoritmos de aprendizado;
- *Knowledgeflow* – É uma ferramenta gráfica, ainda em desenvolvimento, que permite o planejamento de ações, construção do fluxo de processos.

O *Weka* possui implementados diversos métodos de associação, classificação e *clustering*. A inclusão ou remoção de novos métodos pode ser realizada de forma simples e rápida, o que torna a ferramenta versátil e expansível. O *Weka* suporta a abertura direta de arquivos ARFF, CSV, C45. Porém apenas consegue manipular os ARFF. Esse é um arquivo ASCII usado para definir atributos e seus valores. O *Weka* também permite a visualização gráfica dos dados em forma de histograma, e a apresentação de resultados em árvores de decisão, diagrama de dispersão, além de prover modelos gráficos para montagem de redes neurais.

B – Darwin

Desenvolvido pela *Thinking Machines Corp*, o Darwin é uma ferramenta disponível para operação nas plataformas *Windows*, *Sun Solares* e *HP-UX*. Fornece interfaces simples, baseada no modelo *Windows*, e oferece também uma programação de vários algoritmos de mineração, que podem ser executados paralelamente. Pode ser utilizado em conjunto com diversos bancos de dados, sendo capaz de implementar o modelo de mineração por completo.

C - Intelligent Miner

Ferramenta fabricada pela IBM que possui versões para operação nas plataformas *Windows*, *Solaris*, *AIX*, *OS/390* e *OS/400*. O *Inteligente Miner* não é dependente do sistema IBM, podendo também ser utilizado junto a outros bancos de dados de outros fabricantes. A IBM também oferece o *Intelligent Miner* para texto, que realiza atividades de mineração em dados de texto, incluindo a filtragem de *e-mail* e páginas *Web*. Escalável e com suporte para várias plataformas, o pacote *Intelligent Miner for Data* oferece um conjunto de ferramentas apto a fornecer uma estrutura que suporta o processo iterativo de mineração.

O *Intelligent Miner* permite a utilização de algoritmo de mineração, de forma individual ou combinada, para solucionar problemas de mineração também possui uma interface de programação de aplicativos que permite o desenvolvimento de aplicação personalizada.

D - *Bramining*

É uma ferramenta nacional de mineração produzida ao longo de três dissertações de mestrado realizadas pela PUC e no IME. Essa ferramenta disponibiliza um ambiente para a realização do processo e conceitualmente trabalha em três níveis denominados níveis funcionais: Nível dos Métodos, Nível das Operações e Nível das Etapas.

O nível dos métodos é o nível funcional mais baixo e contém os métodos que se encontram disponíveis no *Bramining* para utilização durante a realização de aplicações de mineração. Os métodos são classificados em operações. Uma operação de mineração é uma especificação lógica de um grupo de métodos que tem a mesma finalidade. As operações disponíveis no *Bramining* compõem o nível funcional intermediária da ferramenta, denominado nível de operações. O nível das etapas é o nível funcional mais elevado. Nele as operações de descoberta e conhecimento de dados são agrupadas nas etapas do processo de descoberta e conhecimento: pré-processamento, mineração de dados, pós-processamento.

O processamento iterativo e interativo de uma aplicação de descoberta e conhecimento pelo *Bramining* requer, a cada ação, que seja definido o método que deverá implementar tal ação. A hierarquia mencionada anteriormente é utilizada para auxiliar usuários inexperientes a identificar, por meio de filtros, os métodos disponíveis que viabilizam a operação desejada. Uma vez selecionada o método desejado, a sua execução demanda a especificação de estrutura de entrada e dos parâmetros específicos do referido método. Cada método, após ser executado, apresenta uma interface específica com os resultados produzidos. Esses resultados podem ser novos conjuntos de dados ou mesmo modelos de conhecimento descobertos.

E - *Oracle Data Mining (ODATA MINING)*

É um *software* de mineração em que todas as atividades de descobertas de conhecimento ocorrem no mesmo ambiente do gerenciador de banco de dados *Oracle*, provendo uma plataforma integrada simples, segura e escalável. Tal integração representa um diferencial importante em favor de sua utilização, pois ao contrário das demais ferramentas, não requerendo a extração prévia dos dados para que esses sejam processados pelos métodos de mineração.

O *ODATA MINING* permite a realização de tarefas de classificação, regressão, associação, *clustering*, e mineração de texto. A infra-estrutura de análise de dados e desenvolvimento de aplicações integrada a mineração de dados é suportada por meio do *software Java API, graphical user interface*.

F – SPSS

O *software* estatístico *SPSS* para *Windows* é um poderoso sistema de análise estatística e manuseamento de dados, em um ambiente gráfico, em que a utilização mais freqüente, para a maioria das análises a efetuar, se resume à seleção das respectivas opções em menus e caixas de diálogos [FERREIRA, (2005)]. Além disso, o sistema dispõe de um editor de comandos, a que o utilizador mais avançado poderá recorrer a fim de realizar determinados tipos de análises mais complexas e elaboradas.

O *SPSS* implementa a técnica de classificação baseada em associações, cujo objetivo é gerar subconjuntos de regras de associações, em que fica restrito ao lado direito das regras, o atributo alvo da classificação. Além de produzir regras de classificação, também pode ser aplicado para extrair regras normais de associação e categorização de textos, executa árvore de decisão para analisar e gerar modelos de predição, com diagramas de árvore de fácil entendimento, também tem produtos para rede neurais que provê modelagem e predição, séries temporais e *clustering* [GROTH, (1997)].

É um pacote dos programas para manipular, analisar, e apresentar dados; o pacote é usado extensamente nas ciências sociais e estatísticas. Há diversos formulários de *SPSS*. O programa do núcleo é chamado base de *SPSS* e há a número dos módulos *add-on* que estendem a escala da introdução de dados, estatística, ou relatando potencialidades. Em nossa experiência, o mais importantes desses para a análise estatística são os modelos avançados *SPSS* e da regressão de *SPSS* os módulos do *add-on* dos modelos.

2.3.2.2.3 Métodos de Mineração

Cada método de mineração requer necessidades de pré-processamento diferentes. Tais necessidades variam em função do aspecto extensional da base de dados em que o método é

utilizado [MORIK, (2000)]. Em decorrência da grande diversidade de métodos de pré-processamento de dados, são muitas as alternativas possíveis de combinações entre os métodos. Sendo assim essa escolha pode influenciar na qualidade da pesquisa. A seguir se definem alguns dos métodos utilizados na mineração dos dados:

- Redes Neurais
- Algoritmos Genéticos
- Instâncias
- Estatísticos
- Específicos
- Arvore de Decisão
- Lógica Nebulosa

A) Redes Neurais

Diversos modelos de redes neurais podem ser utilizados na implementação de métodos de descoberta e conhecimento de dados. Classificação, regressão, previsão em series temporais e *clustering* são exemplos de tarefas de mineração que podem ser implementadas por redes neurais. A topologia da rede neural varia em função do problema e da representação adotadas para os dados. Em geral, em aplicações de mineração, a camada de entrada do modelo neural recebe os dados pré-processados de cada registro de um banco de dados. A rede processa esses dados produzindo uma saída cuja natureza também varia em função da aplicação.

Em redes neurais com aprendizado supervisionado, a entrada corresponde aos atributos preditivos enquanto a saída do modelo corresponde ao atributo objetivo do problema. Assim sendo, o algoritmo de aprendizado pode estimar o erro, ou distância, entre a saída produzida e a saída desejada. Em função do erro calculado, o algoritmo ajusta os pesos das conexões da rede a fim de tornar a saída rela tão próxima quanto seja possível da saída desejada. Esses modelos são muito úteis para reconhecimento de padrões, e em particular, para tarefas de mineração que envolve predição. Por outro lado, redes neurais com aprendizado não supervisionado são adequadas para tarefas que envolvam descrição dos dados, ou seja, tarefas de *clustering*.

A seguir se descreve os algoritmos de aprendizado *back-propagation*: algoritmo de retro propagação do erro e de aprendizado supervisionado, cuja aplicação é adequada à tarefa de mineração tais como classificação, regressão ou previsão de series temporais.

Esse algoritmo tem como objetivo minimizar a função de erro entre a saída gerada pela rede neural e a saída desejada, utilizando o método do gradiente descendente. A topologia de uma rede neural não linear cujo comportamento seja codificado pelo algoritmo *back-propagation* é, em geral, composto de uma camada de entrada, uma camada de saída e um número arbitrário de camadas intermediária. Cada neurônio de uma camada, com exceção da camada de entrada, encontra-se conectado a todos os neurônios presentes na camada intermediariamente anterior a sua.

B) Algoritmos Genéticos

São modelos computacionais de busca e otimização de soluções em problemas complexos, inspirados em princípios da teoria da evolução natural de Charles Darwin e da reprodução genética. Segundo o princípio básico da evolução natural de Darwin, indivíduos mais aptos possuem maiores chances de sobrevivência e, conseqüentemente, mais oportunidades de gerarem descendentes e perpetuarem seus códigos genéticos pelas gerações seguintes. A identificação é expressa individualmente pelo seu código genético que fica representado nos cromossomos desse indivíduo.

Genericamente os algoritmos genéticos são extremamente úteis em problemas complexos que envolvam otimização. Em particular, podem ser aplicado a diversas tarefas de mineração de dados. Essa técnica procura obter soluções por meio da evolução de populações, por meio de soluções codificadas em cromossomos artificiais.

Algoritmos genéticos empregam um processo adaptativo e paralelo de busca de soluções em problemas complexos. O processo é adaptativo, pois as soluções existentes a cada instante influenciam a busca por futuras soluções. O paralelismo do processo é decorrência natural do fato de que diversas soluções são consideradas a cada momento pelos algoritmos genéticos.

Rule-Evolver: Baseado na teoria de algoritmos genéticos [MICHALEWICZ, (1996)] procura regras SE-ENTÃO para melhor descrever as classes de um problema. A parte SE de

cada regra dita qual intervalo de cada atributo da base de dados melhor caracteriza a classe descrita após o ENTÃO. Regras compostas por diferentes atributos podem ser agregadas com a utilização do conectivo E.

C) Instâncias

Indica que o método, ao processar um novo registro, leva em consideração as instâncias ou os registros existentes na base de dados. Um dos principais métodos baseados em instância é denominado de *K-NN*, ou seja, *K* vizinho mais próximo.

K-NN – o método é muito utilizado em aplicações envolvendo a tarefa de classificação. Trata-se de um método de fácil entendimento e que não requer treinamento e implementação que não necessita de treinamento prévio para ser aplicado. O funcionamento do *K-NN* consiste na tarefa de classificação por atributos de uma base de dados conhecida também como base de referência de um problema, e cada novo registro a ser classificado segue os passos descritos a seguir:

- Determinar o cálculo da distância do novo registro, a cada um dos registros existentes na base de referência;
- Identificar os *k* registros da base de referência que apresentaram menor distância em relação ao novo registro;
- Apuração da classe mais frequente entre os *k* registros identificados no passo anterior;
- Comparação da classe apurada com a classe real, computando erro o acerto do algoritmo.

D – Estatísticos

Existem diversos métodos estatísticos para a classificação de padrões, alguns clássicos e outros mais recentes. Todos assumem a existência de uma variável (atributo) resposta, *y*, e uma coleção de variáveis de entrada, $x = (x_1, x_2, \dots, x_j)$, além da disponibilidade de dados para treinamento. A meta é encontrar um modelo para prever $(x \hat{=} y)$ = que funcione bem quando aplicado a um novo dado. Um resumo dos principais modelos aplicados a todas as técnicas de decisão pode ser encontrado em FRIEDMAN, (1995).

Entre as Metodologias mais usadas estão os classificadores *Bayesianos*. O princípio básico de classificadores *Bayesianos* está fundamentado na teoria da probabilidade *Bayesianos* [DUDA et al., (2000)].

K-means – Método mais popular de *Clustering*, nesse método toma-se, randomicamente, k pontos de dados como sendo os centróides dos clusters. Em seguida, cada ponto é atribuído ao cluster cuja distância desse ponto em relação ao centróide de cada cluster é a menor dentre todas as distâncias calculadas. Um novo centróide para cada cluster é computado pela média dos pontos do cluster, caracterizando a configuração dos clusters para a iteração seguinte. O processo termina quando os centróides dos clusters param de se modificar, ou após um número limitado de iterações que tenha sido especificada pelo usuário.

E) Específicos

São métodos desenvolvidos especificamente para implementar alguma tarefa de descoberta e conhecimento de dados, no caso do *Apriori* é um método desenvolvido para descoberta de associações.

Apriori – É um algoritmo clássico de mineração que inspirou a criação de diversos outros algoritmos como exemplo DHP, GSP entre outros. O *Apriori* baseia-se no princípio da antimonotonicidade do suporte segundo o qual um *k-itemset* somente pode ser freqüente se todos os seus $(k-1)$ -*itemset* forem freqüentes. Assim sendo, a combinação de *itemsets* para gerar um novo *itemset* somente ocorre quando esses são freqüentes.

F) Árvore de Decisão

A árvore de decisão é uma ferramenta completa e bastante conhecida para classificar os dados e apresentar os resultados sob a forma de regras [BERRY & LINOFF, (2000)]. A maioria das árvores de decisão executa a classificação em duas fases: construção da árvore e *prunning* [AGRAWAL et al., (1992)]:

1. Construção da árvore: a árvore vai se ramificando por meio de sucessivas divisões dos dados com base nos valores dos atributos. O processo é repetido recursivamente até que todos os registros pertençam a uma classe:

2. *Prunning* (poda): remove as ramificações que não tem valor significativo para criar o modelo de classificação, selecionando a sub-árvore que contém a menor taxa de erro estimada. Os nós são rotulados pelos nomes dos atributos; os galhos são os valores possíveis de cada atributo e as folhas são os valores das classes. Os registros são classificados seguindo um caminho para baixo na árvore, sendo desenhada com a raiz no topo e as folhas embaixo. Um registro entra na árvore pelo nó raiz. Na raiz, é aplicado um teste para determinar o próximo nó aonde o registro irá se posicionar. Há diferentes algoritmos para escolher o teste inicial, mas o objetivo é sempre o mesmo: escolher aquele que melhor descreve a classe alvo. O processo é repetido até que o registro chegue a uma folha. Assim, todos os registros que terminam na mesma folha são classificados da mesma forma. Há somente um caminho da raiz até cada folha e esse caminho é a expressão da regra usada para classificar os registros [BERRY & LINOFF, (2000)].

Há vários algoritmos para construir árvores de decisão. Os mais conhecidos são CART, CHAID (*Chi-Squared Automatic Interaction Detection*), ID3 (*Iterative Dichotomiser 3*) e C4.5.

G) Lógica Nebulosa

Diversos métodos de mineração foram adaptados de forma a incorporar a flexibilidade proporcionada pela lógica nebulosa. A lógica nebulosa (*Fuzzy Logic*) é uma teoria matemática que tem como principais objetivos permitir a modelagem do modo aproximado de raciocínio, imitando a habilidade humana de tomar decisões em ambientes de incerteza e imprecisão. Com o conceito de lógica nebulosa constroem-se sistemas inteligentes de controle e suporte a decisão que lidem com informações imprecisas e subjetivas.

H) Sistemas *Neuro-Fuzzy*

Sistemas *neuro-fuzzy* (SNF) são sistemas híbridos que combinam as vantagens das redes neurais, no que se refere ao aprendizado, com o poder de interpretação lingüístico dos sistemas baseados em lógica nebulosa.

Os sistemas *neuro-fuzzy* realizam, internamente, um mapeamento entre regiões do espaço de entrada em regiões *fuzzy* do espaço de saída, por meio de regras *fuzzy* do sistema. As regiões *fuzzy* do espaço de E/S são determinadas no processo de identificação da estrutura. Nesse processo, os espaços de entrada e/ou saída são divididos segundo um determinado método de partição. As variáveis de entrada e saída dos sistemas *neuro-fuzzy* são divididas em vários termos lingüísticos (por exemplo: baixo, alto) que são utilizados pelas regras de lógica nebulosa.

2.3.2.3 Pós-Processamento

Após serem óbitos os modelos de conhecimentos úteis a partir das bases de dados pesquisadas, as atenções se voltam para aplicabilidade dos resultados no contexto da pesquisa. Essa fase compreende a análise final onde todos os dados criados são direcionados para as tomadas futuras decisões. Exemplos comuns da aplicação dos resultados são as alterações em estratégias de negócios que tenham como objetivo procurar tomar proveito dos conhecimentos obtidos. Tais alterações podem variar desde tomadas de decisões sobre em que posição às gôndolas de supermercados, até políticas estratégicas de mercado.

O desenvolvimento de sistemas que utilizam conhecimentos extraídos de base de dados tem propiciado uma valiosa ferramenta na tomada de decisões empresariais [AGRAWAL et al., (1992)].

A - Simplificação do Modelo de Conhecimento

Consiste em remover detalhes desse modelo de conhecimento de forma a torná-lo menos complexo, sem perda de informação relevante. A representação de conhecimento por meio de regras usado em descoberta e conhecimento de dados. Para a compreensão humana,

um conjunto com grande quantidade de regras é de difícil aceitação. E nesse contexto que métodos voltados ao corte de regras.

Esses métodos se baseiam em medidas de qualidade das regras tais como precisão e abrangência. Para ilustrar considere um modelo de conhecimento composto por regra da forma: $X \rightarrow Y$, onde X e Y são predicados, ou seja, condições, verdadeiro ou falso em função de cada registro da base [Segundo HAN e KEMBER, (2001)].

Precisão ou acurácia da regra: é o percentual de registros na base de dados que ao satisfazerem aos antecedentes da regra, satisfazem também o conseqüente.

$$A_{cc} = \frac{[X \wedge Y]}{[X]}. \quad (2.8)$$

Abrangência da regra: é o percentual de registros da base de dados que ao satisfazerem ao conseqüente da regra, satisfazem também ao antecedente.

$$A_{br} = \frac{[X \wedge Y]}{[Y]}. \quad (2.9)$$

Sendo comum em mineração que o usuário estabeleça limites de acurácia e abrangência para as regras, de tal forma a excluir do modelo gerado todas as regras que não satisfaçam a tais limites.

B - Transformações de Modelo de Conhecimento

Com a finalidade de facilitar a análise muitas vezes se utiliza métodos de transformação sobre os modelos. Esses métodos consistem basicamente na conversão da forma de representação do conhecimento de um modelo para outra forma de representação do mesmo modelo, tornando-o mais entendível.

Um exemplo comum é a conversão de árvores de decisão em regras ou vice-versa. Conforme visto a fase de pós-processamento envolve a visualização, a análise e a interpretação do modelo de conhecimento gerado pela etapa de mineração [GOLDSCHMIDT et al (2005)]. Em geral, é nessa etapa que o especialista em conhecimento e o especialista no

domínio da aplicação avaliam os resultados obtidos e definem novas alternativas de investigação dos dados.

2.3.2.4 Controle

Nessa etapa o papel do pesquisador é de fundamental importância na condução do processo, pois diversos fatores podem influenciar nas etapas de desenvolvimento e conhecimento de dados. O controle é uma etapa que depende exclusivamente do conhecimento do especialista, nessa árdua etapa o mesmo influencia na tarefa de orientar a execução de todo o processo de mineração. Para tanto exige conhecimento profundo do cenário, nesse o pesquisador utilizará experiências anteriores, seu conhecimento, sua intuição e o poder de combinar subjetivamente os fatos de forma a decidir qual a melhor estratégia a ser adotada.

Descreve que para as intervenções nas etapas de pré-processamento, mineração de dados e pós-processamento deve se envolver geralmente em inúmeros aspectos tais como: fatos observados cuja origem e os níveis de detalhamentos são diversos e difusos, o conhecimento e análise dos dados intermediários obtidos, opiniões de outros especialistas no contexto da aplicação, e o máximo de conhecimento prévio do problema a ser controlado [FAYYAD et al., (1996)]. Fica claro o grau de dificuldade na formação de um pesquisador em mineração, sendo essa tarefa árdua, longa e exaustiva, pois requer não somente o conhecimento de uma ferramenta teórica sobre a área, mas também a participação desses em inúmeras experiências práticas.

2.3.3 Metodologias de Descoberta de Conhecimento em Base de Dados

A escolha e aplicação de técnicas de descoberta de conhecimento em base de dados (KDD), consistem em encontrar uma melhor aplicação para o seu problema, entre diferentes técnicas a serem aplicadas. A construção da Metodologia varia de técnica para técnica. Para a classificação, o conjunto de treinamento é usado para gerar uma explicação da variável alvo em relação às variáveis independentes. Essa explicação pode ser na forma de uma árvore de decisão, de rede neural ou de outra modalidade de relação entre a variável que se deseja

classificar e as demais variáveis da base de dados. A classificação utiliza as ocorrências passadas para construir um modelo futuro. Para isto, é necessário dados pré-classificados, oriundos de dados históricos ou de um outro processo de descoberta de conhecimento.

2.3.3.1 Tipos de Metodologias

À medida que as técnicas de análise de dados se multiplicam, fica aparente que a conclusão bem sucedida de uma pesquisa envolve mais do que a seleção do método correto. Segundo HAIR, (2005) questões que variam desde a definição do problema até o diagnóstico crítico dos resultados que devem ser abordados. A seguir se discute algumas das metodologias de pesquisas encontradas atualmente.

A - Metodologia *CRISP-DATA MINING*

A Metodologia *CRISP-DATA MINING* foi concebida em finais de 1996 e o seu desenvolvimento foi motivado pelo interesse crescente e generalizado, por um lado pelo mercado de *DATA MINING*, e por outro, pelo consenso de que a indústria necessitava de um processo padronizado [WIRTH, (2000)]. Os fundamentos dessa Metodologia vão além dos princípios acadêmicos e teóricos baseiam-se na prática, na experiência daqueles que desenvolvem de fato, projetos de *DATA MINING*. O conhecimento prático foi assim incorporado de forma a dar resposta aos requisitos dos utilizadores, não se centrando unicamente na tecnologia, mas antes na resolução de problemas do negócio HAN e KEMBER, (2001). A Metodologia *CRISP-DATA MINING* é descrita em termos de um processo hierárquico, com um ciclo de vida que se desenvolve em seis fases:

- Estudo de Negócios
- Estudo dos Dados
- Preparação dos Dados
- Modelação
- Implementação
- Avaliação

As fases não têm uma seqüência fixa, dependendo do resultado e do desempenho das outras fases ou das tarefas particulares de determinada fase [CHANPMAN et al., (2000)].

A Figura 2.24 apresenta o ciclo de vida da Metodologia *CRISP-DATA MINING*, em que as setas indicam a seqüência, ligações e interligação entre as fases.

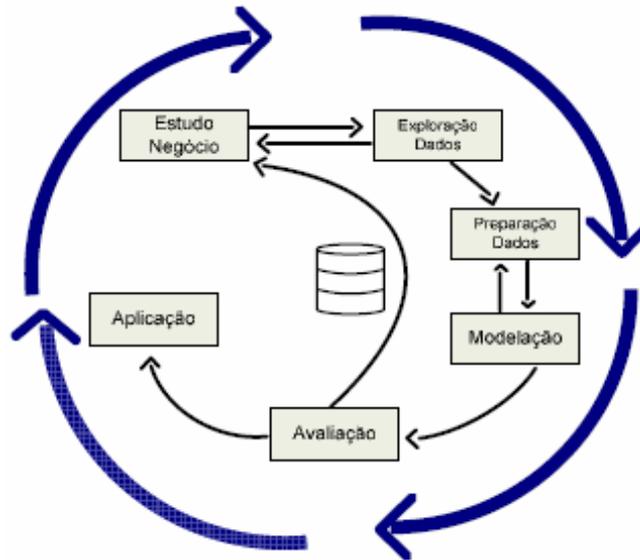


Figura 2.24 Metodologia *CRISP-DATA MINING*.

Fonte: Adaptado de CHANPMAN et al. (2000).

B - Metodologia *SEMMA*

A Metodologia *SEMMA* foi proposta pelo Instituto SAS (SAS *Institute Inc.* <http://www.sas.com>), que se dedica ao desenvolvimento de soluções para estatística, análise de dados, *business intelligence*, *Data Mining* SAS, 2005. Essa Metodologia surge como resposta à necessidade de definição, padronização e de integração dos processos de *DATA MINING* nos ciclos de produção, para que a solução seja aceita mais facilmente no ambiente do negócio [GROTH, (2000)]. Mais do que uma Metodologia de *DATA MINING*, é considerada como um auxiliar para conduzir um projeto em todas as suas etapas, desde a especificação do problema do negócio até à sua implementação.

DATA MINING é definida pelo Instituto SAS como o “processo de extrair informação valiosa e relações complexas de um grande volume de dados” e foi nesse sentido, que

dividiram o processo de *DATA MINING* em cinco etapas da Figura 2.25 – dando origem ao acrônimo *SEMMA* [SAS, (2005)][GROTH, (2000)]:

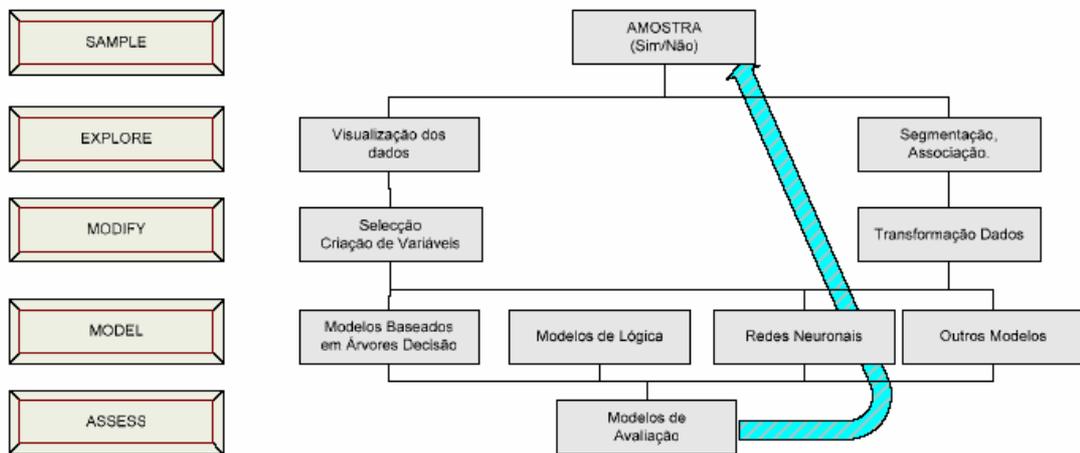


Figura 2.25 Metodologia *SEMMA*.

Fonte: Adaptado de SAS, (2005).

Numa forma resumida é possível apresentar essa Metodologia como um processo com cinco fases, que se inicia com uma amostra (*Sample*) representativa dos dados à qual se aplicam técnicas estatísticas de exploração e de visualização dos dados (*Explore*).

Posteriormente são selecionadas e transformadas as variáveis (*Modify*) consideradas mais significativas (as variáveis que sobressaíram na fase anterior), as que são mais relevantes em termos de projeto, e sobre as quais se constroem os modelos (*Model*) (aplicam-se algoritmos no sentido de alcançar os objetivos) e por fim se avalia o modelo (*Assess*). Cada uma das etapas é distinta e corresponde a um ciclo, e as suas tarefas internas podem ser executadas repetidamente sempre que necessário, pode-se atualizar e ajustar quando surgir nova informação. A caracterização mais completa das etapas da Metodologia *SEMMA*.

C - Metodologia **ADRIAANS** e **ZANTINGE**

Essa abordagem do processo de mineração baseia-se na necessidade das organizações em obterem continuamente novas informações sobre seus dados, por isso não deve ser executado apenas uma vez, mas repetido sempre que novas necessidades de informações

aparecerem. Portanto, nessa abordagem não existe uma etapa específica para entendimento dos dados. É pressuposto que já exista um conhecimento prévio do domínio da base de dados e, conseqüentemente, do objetivo do processo. O processo é composto por seis etapas conforme a Figura 2.26 [ADRIAANS, (1997)].

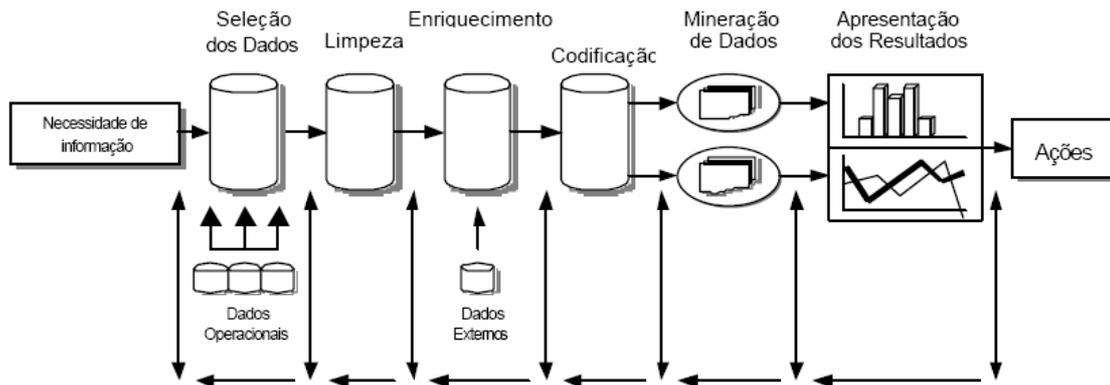


Figura 2.26 Metodologia ADRIAANS e ZANTINGE.

Fonte: Adaptado de Adriaans et al., (1997).

Seleção dos dados: Nessa etapa ocorre uma análise de todos os dados operacionais da base de dados e são selecionados apenas aqueles que são necessários para alcançar os objetivos do processo. Podem ser feitas novas seleções quando houver outra iteração, ou seja, podendo incluir dados anteriormente descartados, pois o processo é iterativo, permitindo a retomada de qualquer etapa, independente da etapa em que se encontra.

Limpeza: Nessa etapa são utilizadas diversas operações de limpeza nos dados, como por exemplo: a remoção de dados duplicados e a correção dos dados. A etapa de limpeza pode ser executada inúmeras vezes, já que é impossível prever com antecedência todos os problemas de qualidade existente na base.

Enriquecimento: Algumas informações podem ser incluídas na base de dados para que seja possível atingir os objetivos do processo. Esses dados podem estar disponíveis em outros locais, ou até mesmo podem ser gerados a partir de dados existentes na base de dados e transformados para se obter a informação.

Codificação: A forma que os dados estão armazenados nas bases de dados pode não ser a representação mais apropriada para a utilização no processo de mineração. Geralmente, os dados têm sua representação apropriada ao contexto da aplicação. Por exemplo, um atributo

com valores literais pode não ser adequado a determinados algoritmos mineradores utilizados na etapa de mineração de dados. Para adequá-lo pode ser necessário normalizar esses valores dentro de um determinado intervalo. A codificação é um procedimento criativo, existem diversas maneiras de codificação, assim é difícil descrevê-las, pois cada caso deve ser analisado individualmente e sua codificação pode variar de acordo com a escolha do algoritmo minerador da próxima etapa.

Mineração de dados: Essa é a etapa onde os dados são manipulados para que seja extraído o conhecimento. É a etapa que mais exige dos recursos computacionais. O autor afirma que, utilizando inicialmente uma ferramenta de consulta SQL, pode ser possível ter uma visão geral dos dados para então partir para uma análise menos trivial. Nessa primeira tarefa, 80% do conhecimento são extraídos e já podem revelar alguma informação interessante. Entretanto, as informações extraídas por essas consultas podem não ser suficientes, surgindo à necessidade de se utilizar técnicas avançadas.

Apresentação dos resultados: Finalizada a etapa de mineração de dados, resultam informações num formato específico de acordo com a técnica utilizada. Deve-se levar em conta que os dados podem estar codificados ou mesmo que o método utilizado na etapa de mineração que gere, como saída, informações em algum formalismo ou representação muito específicas.

Esses resultados devem ser exibidos de forma clara para que sejam de fácil entendimento para quem irá utilizá-los, geralmente pessoas que necessariamente não interpretarão os resultados tão facilmente quanto aquela que conduziu o processo de mineração.

D - Metodologia KLEMENTINEN

KLEMENTINEN et al., (1997) apresentam uma Metodologia que pode ser usada para automatizar aquisição de conhecimento. As fases dessa Metodologia são aquelas já definidas pelo autor FAYYAD, (1996): pré-processamento, transformação, descoberta, apresentação e utilização na Figura 2.27.

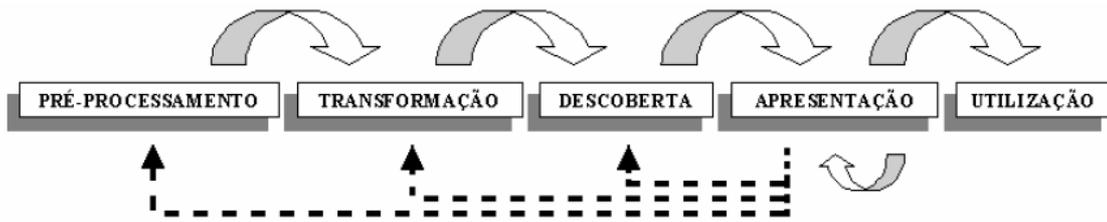


Figura 2.27 Metodologia KLEMENTINEN.

Fonte: Adaptado de KLEMENTINEN et al., (1997).

No entanto, a maior ênfase é dada nas duas fases centrais dessa Metodologia:

- Fase de descoberta de padrões: onde são encontrados todos os padrões potencialmente relevantes para algum critério bastante livre;
- Fase de apresentação: onde são fornecidos métodos flexíveis para iterativa e interativamente criar diferentes visões para os padrões descobertos.

Nas duas primeiras fases do processo, os dados são coletados e preparados de forma adequada para descoberta de padrões. Uma visão geral sobre os dados pode ser produzida nessa fase. Os atributos identificados como irrelevantes são removidos e novos atributos podem ser derivados. Na fase de descoberta de padrões, todos os padrões potencialmente interessantes são gerados do conjunto de *data set*. A apresentação do conhecimento descoberto é uma parte principal dessa Metodologia. Nessa fase, os padrões relevantes podem ser localizados de grandes coleções de padrões potencialmente relevantes.

E - Metodologia FELDENS

FELDENS et al. (1998) propõem uma Metodologia integrada, na qual as tecnologias de mineração de dados e *data warehouse*, bem como questões de visualização têm papéis muito importantes no processo. Também supõe uma forte interação entre mineradores de dados e pessoas da organização para questões de modelagem e preparação de dados. As fases definidas para essa Metodologia são: pré-processamento, mineração de dados e pós-processamento, conforme Figura 2.28.

A fase de pré-processamento inclui tudo o que é feito antes da descoberta e conhecimento de dados. Nesta fase é feita uma análise na organização, para focar no projeto

de descoberta e conhecimento de dados, a análise dos dados existentes, integração de fontes de dados, transformações de dados, etc.

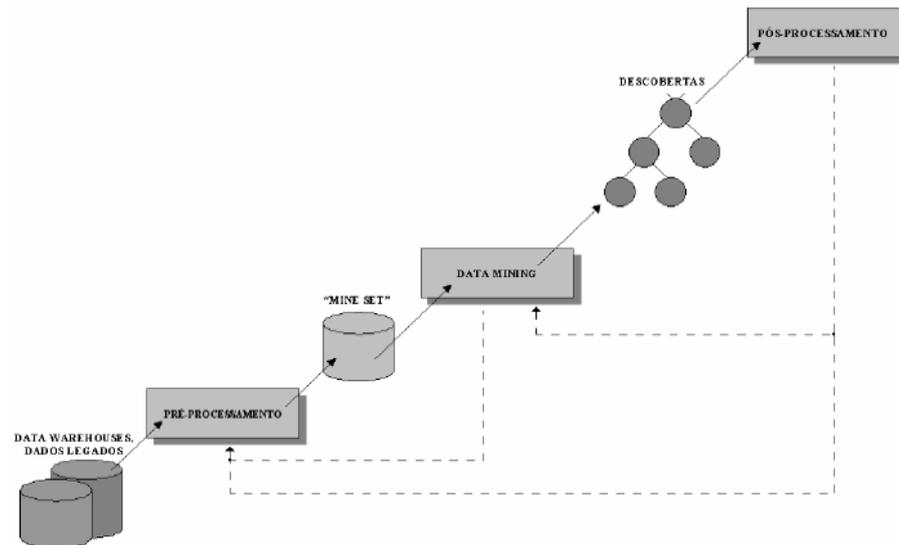


Figura 2.28 Metodologia FELDENS.

Fonte: Adaptado de FELDENS et al., (1998).

A fase de mineração de dados inclui a aplicação de algoritmos, possivelmente a aplicação repetida. A escolha dos algoritmos pode ser realizada baseando-se na análise que é feita na fase de pré-processamento. A fase de pós-processamento pode ser definida por operações de filtragem, estruturação e classificação. Somente após essa fase, o conhecimento descoberto é apresentado ao usuário. O conhecimento descoberto pode ser filtrado por alguma medida estatística, por exemplo, suporte, confiança ou outro critério definido pelo usuário. Estruturação significa que o conhecimento pode ser organizado de forma hierárquica.

F - Metodologia HAIR

Uma grande contribuição foi dada por HAIR et al., (2005) que apresentou uma metodologia baseada em seis estágios. A meta do autor não foi de fornecer um conjunto rígido de procedimentos, mas sim orientações que enfatizem maneira de se construir um modelo.

Um processo com seis estágios para construir modelos fornece uma estrutura para desenvolver, interpretar e validar qualquer análise multivariada. O processo discutido por HAIR consiste nos estágios a seguir:

Estágio 1: Definição do problema da pesquisa, dos objetivos e da técnica multivariada a ser usada.

O ponto de partida para qualquer análise múltipla é definir o problema da pesquisa e os objetivos de análise em termos conceituais, antes de especificar qualquer variável ou medida. Um modelo conceitual não precisa ser complexo e detalhado; pode ser uma simples representação das relações a serem estudadas. Se uma relação de dependência é proposta como objetivo da pesquisa, devem ser especificados os conceitos de dependentes e independentes. Já para uma aplicação de uma técnica de independência, as dimensões de estrutura ou similaridade devem ser especificadas. Devem ser observados que um conceito, diferentemente de uma variável sempre ser definidas, independente da relação.

Inicialmente devem ser identificadas idéias ou os tópicos de interesse, em vez de se concentrar nas medidas específicas a serem utilizados. Isso minimiza as chances de conceitos relevantes serem omitidos no esforço de desenvolver medidas e de definir as especificidades do plano de pesquisa.

Com o objetivo e o modelo conceitual especificados, deve-se escolher a técnica multivariada a ser utilizada. Após escolher entre um método de dependência ou independência, a última decisão é selecionar a técnica em particular com base nas características de medidas das variáveis dependentes e independentes. As variáveis podem ser especificadas antes do estudo em seu planejamento ou depois que os dados foram coletados, quando são definidas análises específicas.

Estágio 2: Planejamento da pesquisa.

Com o modelo conceitual estabelecido e a técnica multivariada selecionada, a atenção se volta para a implementação. Para cada técnica deve ser desenvolvido um plano de análise que aborde as questões particulares a seu propósito e projeto. As questões incluem considerações gerais, como tamanho mínimo da amostra, tipos permitidos ou exigidos de variáveis e métodos de estimação, além de aspectos específicos, como o tipo de medidas de associação de resultados agregados ou desagregados em análise conjunta ou uso de formulações especiais de variáveis para representar efeitos não-lineares ou interativos em

regressão. Em cada caso, essas questões resolvem detalhes específicos e exigências para a coleta dos dados.

Estágio 3: Suposições em análise.

Nesse estágio devem-se fazer várias suposições sobre as relações entre as variáveis dependentes e independentes que afetam o procedimento estatístico (mínimos quadráticos). A seguir discutem-se testes para as suposições e possíveis ações para corretivas.

Atender as suposições da análise é essencial para garantir que os resultados obtidos sejam realmente representativos nas observações, obtendo com isto os melhores resultados possíveis. As questões básicas a serem respondidas nesse estágio estão ligadas às suposições abaixo:

- Linearidade do fenômeno medido;
- Variância constante dos termos de erro;
- Independência dos termos de erro;
- Normalidade da distribuição dos termos de erro.

Todas as técnicas multivariadas têm suposições inerentes, estatísticas e conceituais, que influenciam muito suas habilidades para representar relações multivariadas. Cada técnica tem também uma série de suposições conceituais que lidam com questões como a formulação de modelo e os tipos de relações representadas. Antes de qualquer estimação de modelo, deve ser garantido que as suposições estatísticas e conceituais estejam satisfeitas.

Estágio 4: Estimação do modelo e avaliação do ajuste geral do modelo.

Com a análise da pesquisa especificada em termos de variáveis dependentes e independentes, a amostra considerada adequada para os objetivos do estudo e as suposições avaliadas para as variáveis individuais, o processo de construção do modelo agora segue para a estimação do modelo a ser pesquisado e a avaliação do ajuste geral do modelo. Nesse estágio devem ser cumpridas três tarefas básicas:

- Selecionar um método para especificar o modelo a ser estimado;
- Avaliar a significância estatística do modelo geral na previsão da variável estatística;
- Determinar se algumas das observações exercem uma influencia indevida nos resultados.

No processo de estimação, se dispõe de opções para atender características específicas dos dados ou maximizar o ajuste dos dados. Depois que o modelo é estimado, o seu ajuste geral é avaliado para estabelecer se atinge níveis aceitáveis sobre os critérios estatísticos, se identifica às relações propostas e se tem significância prática. Muitas vezes o modelo é reespecificado, em uma tentativa de atingir melhores níveis de ajustes e explicação geral. Em todos os casos, um modelo aceitável deve ser obtido antes de se prosseguir.

Estágio 5: Interpretação das variáveis estatística pesquisadas.

Nesse estágio deve-se examinar a equação preditiva, e com isto avaliar a importância relativa que as variáveis individuais na previsão geral do produto. O pesquisador nesse estágio tem como função interpretar a variável estatística de regressão pela avaliação dos coeficientes de regressão estimados em termos de sua explicação da variável dependente, não se deve avaliar tão somente o modelo de regressão estimado, mas também as variáveis independentes omitidas, se uma busca seqüencial ou abordagem combinatória foi empregada. Além disso, algumas técnicas também estimulam múltiplas variáveis estatísticas que representam dimensões latentes de comparações ou associações. A interpretação também pode levar a reespecificações adicionais as variáveis e/ou formulação do modelo, onde o modelo é reestimado e então novamente interpretado. O objetivo é identificar evidências empíricas de relações multivariadas nos dados da amostra que possam ser generalizados para a população total.

Estágio 6: Validação dos Resultados.

Essa fase consiste em generalizar a aplicação do modelo, demonstrando que ele não é específico as observações utilizadas na estimação. Nesse estágio final se deve garantir que ele representa a população geral e que sejam apropriadas as situações nas quais é usada. As

tentativas de validar o modelo são direcionadas no sentido de demonstrar a generalidade dos resultados para a população total, sendo que essas análises de diagnósticos acrescentam pouca interpretação dos resultados, mas podem ser vistas como uma garantia de que os resultados são os melhores descritivos dos dados e generalizáveis a população.

2.3.3.2 Avaliação da Metodologia

Nesse passo, a Metodologia construída deverá ser criteriosamente avaliada visando a sua aplicação no problema sugerido. Objetiva determinar se algum conhecimento adicional foi descoberto ou se as hipóteses existentes foram confirmadas. Devem ser definidas as regras selecionadas no estudo que agregam valores úteis à predição. A medida dos resultados se refere especificamente ao valor para o negócio e se esse resultado pode ser usado no futuro.

Identificar as informações úteis, sua incorporação aos processos de negócio e mais importante, quem usará essas informações é o grande segredo da mineração de dados [BERRY (2000)].

Para conferir a performance da Metodologia, aplica-se uma estimativa à coleção final de registros pré-classificados. A taxa de erro do conjunto de treinamento é um bom preditor da taxa de erro dos demais dados. Após esse passo, fecha-se o ciclo de mineração de dados. Novas hipóteses podem ser formuladas, reiniciando o processo.

3.1 Introdução

Para efeito dessa pesquisa utilizou-se da metodologia descrita por HAIR et al (2005) com as técnicas de regressão múltipla. A análise de regressão múltipla é uma técnica estatística, que pode ser usada para se analisar a relação entre uma única variável dependente (critério) com várias variáveis independentes (preditoras).

Regressão múltipla é um processo poderoso e flexível para a análise de relações associadas entre uma variável métrica dependente e uma ou mais variáveis independentes, podendo ser utilizada conforme descrito a seguir [MALHOTRA, (2004)]:

- Determinar se as variáveis independentes explicam uma variação significativa na variável dependente: se existe uma relação;
- Determinar quanto da variação na variável dependente pode ser explicado pelas variáveis independentes: intensidade da relação;
- Determinar a estrutura ou a forma da relação: a equação matemática que relaciona as variáveis independentes e dependentes;
- Predizer os valores da variável dependente;
- Controlar outras variáveis independentes quando da avaliação das contribuições de uma variável ou conjunto de variáveis específicas.

Um dos objetivos da análise de regressão múltipla é usar as variáveis independentes cujos valores são conhecidos para prever ou explicar os valores da variável dependente. Cada variável independente selecionada pelo especialista é ponderada pelo procedimento da análise de regressão para garantir a máxima relação de previsão a partir do conjunto de variáveis independentes. Os pesos denotam a contribuição relativa das variáveis independentes para a previsão geral e facilitam na interpretação sobre a influência de cada variável em fazer a

previsão, apesar da correlação entre as variáveis independentes complicarem o processo interpretativo. A variável estatística de regressão é formada pela ponderação das variáveis independentes.

O processo de construção do modelo em seis estágios, como foi visto no Capítulo 2, é utilizado como referência para discutir os fatores que impactam na criação, estimação, interpretação e validação da análise de regressão a ser analisada em seguida.

A Figura 3.1 mostra os estágios que a pesquisa seguiu.

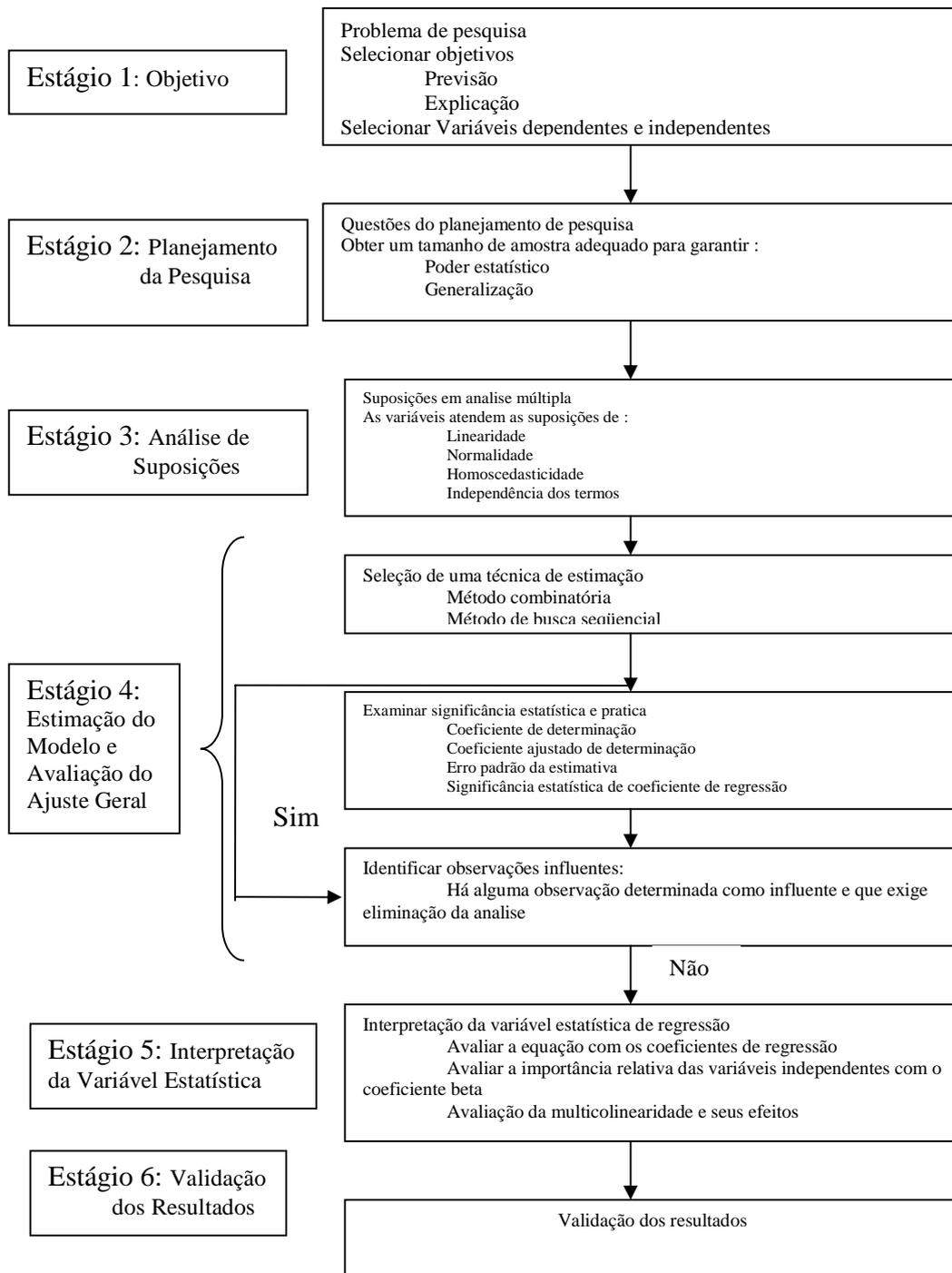


Figura 3.1 Estágios no diagrama de decisão de regressão múltipla.
Fonte: Adaptado pelo autor do HAIR, (2005).

3.2 Objetivo (Estágio 1)

Nesse estágio trabalha-se com adequação do problema da pesquisa, especificações de uma relação estatística e a seleção das variáveis dependentes e independentes. Todo o planejamento inicial da pesquisa deve ser fundamentado na seleção das variáveis a serem utilizadas, sendo necessário nessa técnica que o especialista identifique qual variável é dependente e quais as independentes. Para isto torna-se importante o conhecimento das variáveis, pois a seleção deve ser baseada em questões conceituais ou teóricas.

A regressão múltipla fornece um meio de avaliar objetivamente o grau e caráter da relação entre as variáveis dependentes e independentes, pela formação da variável estatística a partir da variável independente. Na seleção da variável dependente se deve estar conhecedor de todas as variáveis, para isto descreve-se a seguir sucintamente todas as variáveis coletadas.

A) Objetivo da Pesquisa

O objetivo principal é entender o comportamento dos clientes de telecomunicações que são usuários de Internet de banda larga, por meio da análise da taxa de abandono associado com as medidas de qualidade de serviços prestados. Tenta-se associar como a qualidade de serviço prestado ao cliente pode ser percebida por ele e como isto pode afetar diretamente a taxa de abandono. Para isto utilizou-se da análise múltipla de dados com a técnica de regressão linear com a utilização da ferramenta estatística *SPSS*, visando com isto encontrar uma equação linear que correlacione os indicadores de qualidade com a rotatividade da planta, a qual é conhecida mundialmente como Taxa de Abandono. Esse objetivo principal é desdobrado nos seguintes objetivos específicos:

- 1) Descrever e explicar uma forma de relação entre os serviços de banda larga disponibilizados pelas empresas de telecomunicações e a taxa de abandono da planta;
- 2) Analisar um conjunto de variáveis que possua a maior relevância para o problema em estudo.

As duas hipóteses propostas para essa pesquisa, considerando-se que a taxa de abandono é a variável dependente, e os indicadores de qualidade de serviço são as variáveis independentes estão descritas abaixo:

H0: Não existe relação entre a taxa de abandono e os indicadores de qualidade de serviço;

H1: Existe uma relação entre a taxa de abandono e os indicadores de qualidade de serviço;

B) Problema da Pesquisa

A regressão linear múltipla tem suas aplicações em duas grandes classes de problema de pesquisa: previsão e explicação, as quais não são mutuamente excludentes, e uma aplicação da análise de regressão linear múltipla pode abordar qualquer um desses problemas de pesquisa ou ambos.

Nessa pesquisa tem-se como aplicação apenas uma destas classes, explicação, que é feita por meio de uma relação entre o problema, existência de abandono no serviço de Internet de banda larga, medida pela sua taxa de abandono com os indicadores de qualidade de serviço de banda larga.

C) Seleção das variáveis dependentes e independentes

Como o objetivo da pesquisa proposta é uma relação de dependência, há necessidade de se especificar e selecionar as variáveis dependente e independente. Para tanto, é descrito os indicadores de Internet banda larga utilizados nas empresas de telecomunicações, em seguida é feita análise dos atributos de valor utilizando-se a técnica de Mudge (método de priorização das variáveis) para, respaldado em informações substanciadas, se caracterizar os parâmetros no sentido de que fazer uma seleção adequada das referidas variáveis.

C.1 Descrição dos Indicadores de Internet Banda Larga

Com o aparecimento dos serviços Internet de banda larga e a obrigatoriedade por parte da ANATEL em controlar os indicadores de qualidade de desempenho para os Serviços

Telefônicos Fixos Comutados (STFC), a área responsável pela qualidade nas empresas de telecomunicações no segmento banda larga foram motivadas a criar indicadores de qualidade similares aos indicadores da telefonia fixa para esse novo serviço. Esses indicadores são coletados após os clientes realizarem ligações telefônicas efetuadas a um *call center* por meio de um número 0800.

A definição do que medir deve estar associado aos fatores críticos que influenciam o comportamento e a sobrevivência de uma organização. Tais fatores críticos estão ligados diretamente ao setor ao qual pertence, influenciando estrategicamente em seu resultado na sobrevivência e, portanto no posicionamento competitivo da organização.

Para definir o que medir as empresas de telecomunicações criaram a área de gestão e controle de qualidade. Essa área tinha a função de analisar as exigências do órgão regulador, ANATEL - Agencia Nacional de Telecomunicações, sobre os indicadores de qualidade que são obrigatórios, bem como procurar posicionar adequadamente de forma estratégica os mesmos em sintonia com as exigências, tendências e necessidades do mercado. Sendo assim dividiu-se os indicadores em indicadores de negócio usados para medir a satisfação do cliente, e em indicadores de qualidade usados para medir o desempenho das operações de *Front - Office*.

A figura 3.2 ilustra onde os indicadores que estão inseridos no contexto da empresa, os indicadores de taxa de abandono, ganho e retiradas são utilizados para medir a satisfação do cliente, enquanto os outros indicadores estão ligados ao desempenho das operações de *Front - Office* da empresa associados à qualidade do serviço.



Figura 3.2 Caracterização dos serviços e entendimento do valor oferecido ao cliente dentro do quadro geral de competitividade em serviço.

Fonte: Adaptado de CORRÊA, (2002).

O referencial a ser estabelecido, na forma de indicadores de desempenho deve ser, em primeira instância, externa à organização, sendo de fácil compreensão no intuito de dar suporte ao processo decisório.

Na Tabela 3.1 mostra-se os indicadores de qualidade que são coletados, bem como uma breve descrição dos mesmos.

Tabela 3.1 Indicadores de Qualidade.

Indicador	Formula dos Indicadores	Valores	Origem
IGT	$\frac{\text{Reparos Aberto nos últimos 30 dias} \times 100}{\text{Total de Instalação Realizada nos Últimos 30 dias}}$	Percentual	Ligações Telefônicas 0800
TEMPO MEDIO REPARO	$\frac{\sum \text{Tempos de Reparo no mês}}{\text{Base de Terminais Reclamados no mês}}$	Percentual	Ligações Telefônicas 0800
IPGC	$\frac{\text{Instalações Executados em até 3 dias corridas} \times 100}{\text{Total de Instalação no Período}}$	Percentual	Ligações Telefônicas 0800
RPAV	$\frac{\text{Reparos Executados no Prazo (4 h)} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
RPDC	$\frac{\text{Reparos Executados no Prazo (8 h)} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
RPT	$\frac{\text{Reparos Executados no Prazo (4 ou 8 h)} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
RRAV	$\frac{\text{Reparos Reincidentes} < 90 \text{ dias (Alto V.)} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
RRDC	$\frac{\text{Reparos Reincidentes} < 90 \text{ dias (D. C.)} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
RRT	$\frac{\text{Reparos Reincidentes} < 90 \text{ dias} \times 100}{\text{Total de Reparos no Mês}}$	Percentual	Ligações Telefônicas 0800
TAXA DE REPARO	$\frac{\text{Quantidade de Reparos em 30 dias} \times 100}{\text{Base em Serviço Mês}}$	Percentual	Ligações Telefônicas 0800
TEMPO DE INSTALAÇÃO	$\frac{\sum \text{Tempos de Instalação no mês}}{\text{Base de Terminais instalados no mês}}$	Percentual	Ligações Telefônicas 0800

Fonte: Elaborado pelo autor.

Na Tabela 3.2 mostra-se os indicadores de negócio que são coletados bem como uma breve descrição dos mesmos.

Tabela 3.2 Indicadores de Negócio.

Indicador	Formula dos Indicadores	Valores	Origem
GANHO	$\frac{(\sum \text{das Instalações} - \sum \text{das Retiradas})100}{\text{Base em Serviço no Mês Anterior}}$	Percentual	Informações Coletados do Sistema de Controle Empresarial
TAXA PREVENTIVA	$\frac{\sum \text{dos Reparos Preventivos}}{\sum \text{dos Reparos Corretivos do Mês Anterior}}$	Percentual	Informações Coletados do Sistema de Controle Empresarial
TAXA DE ABANDONO	$\frac{\text{Quantidade de Retiradas X 100}}{\text{Base em Serviço do Mês Anterior}}$	Percentual	Informações Coletados do Sistema de Controle Empresarial

Fonte: Elaborado pelo autor.

A seguir, se descreve, de forma sucinta, todas as variáveis citadas nas Tabelas 3.1 e 3.2:

Ano: Refere-se ao ano em que foram coletados os dados para análise, tendo sido trabalhado no processo de coleta no ano de 2006.

Base: A quantidade de clientes banda larga em serviço, ou seja, a quantidade de clientes na planta que estão gerando fatura. Existe uma diferença entre clientes em serviço e clientes instalados, o primeiro são todos os clientes que estão gerando conta o segundo refere-se à base de clientes que se pode instalar. Para efeito desse trabalho, refletir a verdadeira quantidade de clientes que podem gerar uma reclamação, sendo trabalhado com a base de clientes em serviços.

Taxa de Abandono: É a troca constante de clientes, por eles não estarem satisfeitos com o produto/serviço, forçando a empresa a sempre conseguir novos clientes para sobreviver, quando deveria ser para aumentar a clientela, ou seja, é a taxa de cancelamentos em % calculada como o número de cancelamentos do período considerado entre o tempo médio desse período.

Estado: Unidade da federação ao qual o indicador refere-se, no nosso trabalho todos os indicadores tem índices coletados por filial (unidade da federação) de atuação dessa empresa de telecomunicações com atuação em 16 Estados do Brasil (Amazonas, Roraima, Pará, Amapá, Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia, Minas gerais, Espírito Santo e Rio de Janeiro).

Ganho: Esse indicador mede o ganho na planta em relação a sua base em serviço, podendo ser bastante útil na análise do crescimento.

Instalação em Garantia Total (IGT): Esse indicador mede a quantidade das instalações, ou seja, ele indica se as instalações estão sendo feitas abaixo da qualidade desejada, devido à baixa quantidade de reclamações não foi viável segmentar esse indicador, pois era um número muito pequeno de usuários. Nesse indicador quanto menor melhor.

Instalações: Quantidade de circuitos instalados na planta, valor acumulado de todas as instalações do mês, coletado no ultimo dia de cada mês.

Tempo Médio de Reparo (TMR): Tempo médio que se leva para realizar um reparo. Esse tempo tem a finalidade de auditoria na produtividade do técnico. Nesse indicador quanto menor maior a produtividade.

Instalação no Prazo Grandes Clientes (IPGC): Esse indicador mede todas as instalações feitas no prazo de 3 dias referentes ao segmento alto valor.. Nesse indicador quanto maior melhor.

Mês: Mês do ano no qual foram coletados os indicadores. Todos os indicadores têm periodicidade diária menos o taxa de abandono, que tem periodicidade mensal, no caso dos indicadores com periodicidade diária, deve ser escolhido o último dia de cada mês para efetuar a coleta, considerando o indicador o acumulado mensal.

Taxa de Preventivas (QRPR): Quantidade de reparos preventivos realizadas em circuitos de banda larga. Seleciona-se uma quantidade igual ou maior que 8% da quantidade de reparos efetuadas no mês anterior. Quanto maior melhor

Reparos: Quantidade total de clientes que reclamaram algum problema do seu serviço de banda larga. Entra nesse quantitativo todas as reclamações, procedentes ou improcedentes, por se tratar da visão cliente do serviço. Nesse indicador quanto menor melhor

Retiradas: Quantidade total de retiradas efetuadas pela prestadora de serviço no mês em questão. Quanto menor melhor

Reparo no Prazo Alto Valor (RPAV): Esse indicador mede todos os reparos feitos no prazo de 4 horas referentes ao segmento alto valor. Nesse indicador quanto maior e melhor.

Reparo no Prazo Demais Clientes (RPDC): Esse indicador mede todos os reparos feitos no prazo de 4 horas referentes ao segmento demais clientes. Nesse indicador quanto maior e melhor.

Reparo no Prazo Total (RPT): Esse indicador mede todos os reparos feitos no prazo de 8 horas para demais clientes e 4 horas para clientes alto valor. A segmentação desse indicador essa no diferencial de horas entre a marcação alto valor e demais clientes. Nesse indicador quanto menor melhor. Nesse indicador quanto maior melhor.

Reparo Reincidente Alto Valor (RRAV): Esse indicador mede os reparos que se reincidentes em 90 dias para a segmentação alto valor. Tendo a função de verificação da qualidade do reparo executado, mede-se a quantidade de reparos que necessita ir mais de uma vez para resolver qualquer que seja o problema do cliente. Nesse indicador quanto menor melhor. Nesse indicador quanto menor melhor.

Reparo Reincidente Demais Clientes (RRDC): Esse indicador mede os reparos que se reincidentes em 90 dias para a segmentação demais clientes. Tendo a função de verificação da qualidade do reparo executados, mede-se a quantidade de reparos que seja necessário irem mais de uma vez para ser resolvido qualquer que seja o problema do cliente. Nesse indicador quanto menor melhor.

Reparo Reincidente Total (RRT): Esse indicador mede os reparos que se reincidentes em 90 dias para a segmentação nas duas segmentações (demais clientes e alto valor). Tendo a função de verificação da qualidade do reparo executados, mede-se a quantidade de reparos que seja necessário irem mais de uma vez para se resolver qualquer que seja o problema do cliente. Nesse indicador quanto menor melhor.

Taxa de Reparo (TRP): Esse indicador mede a quantidade de reparos executados na visão cliente sobre a base em serviço referente a todos os segmentos. Nesse indicador quanto menor melhor.

Tempo Médio de Instalação (TMI): Tempo médio que se leva para realizar uma instalação. Esse tempo tem a finalidade de auditoria na produtividade do técnico. Nesse indicador quanto menor maior a produtividade.

C.2) Método de priorização das variáveis (Técnica matricial de Mudge)

Procura-se selecionar os indicadores que podem ser relevantes. Nessa fase é feita uma adaptação nos passos propostos por PARASURAMAN et al., (1998) e LAPIERRE (2000) para os indicadores mostrados nas Tabelas 3.2 e 3.3.

- Taxa de transmissão de Atendimento – refere-se ao tempo dispensado para execução do serviço ou reparo:

Tempo Médio de Reparo (TMR)

Tempo Médio de Instalação (TMI)

- Confiabilidade – traduz-se como a disponibilidade e no desempenho dos serviços e reparos, podendo ser medido pelo número médio de reparos, ou pelo atendimento do cliente no prazo no caso de instalação.

Taxa de Reparo (TRP)

Instalação no Prazo Grandes Clientes (IPGC)

- Manutenibilidade - é a capacidade de um item ser mantido ou recolocado em condições de executar suas funções requeridas, sob condições de uso especificadas, quando a manutenção é executada sob condições determinadas e mediante os procedimentos e meios prescritos em um determinado prazo.

Reparo no Prazo Alto Valor (RPAV)

Reparo no Prazo Demais Clientes (RPDC)

Reparo no Prazo Total (RPT)

- Consistência – reflete a previsibilidade do resultado esperado, sendo percebida pela ausência de erro.

Instalação em Garantia Total (IGT)

Reparo Preventivo (QRPR)

Reparo Reincidente Alto Valor (RRAV)

Reparo Reincidente Demais Clientes (RRDC)

Reparo Reincidente Total (RRT).

O principal objetivo dessas definições foi criar condições para a operacionalização do valor por meio do conhecimento dessas operações, resultando na representação sistemática do

serviço de manutenção, obtida pelo desenvolvimento dos conceitos relacionados ao sistema genérico de operações.

Devem ser priorizados os critérios de percepção do valor junto aos clientes, de forma a focar as ações gerenciais nesses critérios. Essa priorização é realizada por meio dos "atributos de valor" e da equação inicial do valor utilizando-se o método das relações funcionais desenvolvido por Mudge que foi explicado no Capítulo 2.

- Critério de Avaliação: nessa fase estabelece-se uma escala apropriada de importância, sendo que nessa pesquisa adotou-se uma escala de graduação ponderada, comparação mútua de grau de importância.

Em função da existência da priorização dos indicadores na empresa em estudo, no ano de 2006 foi realizada por seus especialistas na área de manutenção junto aos clientes uma adequação dos parâmetros da metodologia preconizada por Mudge. Esses resultados encontram-se consolidados na Tabela 3.3 onde foram preenchidas todas as linhas da matriz sempre registrando na célula quem é mais importante e qual o peso da importância, se é muito (peso 3), médio (peso 2) ou pouco (peso 1). Na coluna total se faz a soma de todos os pontos obtidos dos indicadores tanto na horizontal como na vertical.

A partir dessas informações montou-se a equação do valor percebido, em que se caracteriza por uma equação em os termos referentes à confiabilidade participam com os indicadores, reparo no prazo alto valor com grau de importância de 15%, reparo no prazo demais clientes com 14%, e reparo no prazo total 12%. Tendo outro fator de percepção de valor consistência no qual, reparo recorrente alto valor e reparo recidência total com 15% de grau de importância para ambos e reparo recorrente demais clientes com 11%.

Tabela 3.3 Tabela de priorização dos fatores de avaliação do serviço de manutenção segundo a técnica de Mudge.

TMR	TMI	TRP	IPGC	RPAV	RPDC	RPT	IGT	QRPR	RRAV	RRDC	RRT	TOTAL	%
TMR	TMR2	TMR2	IPGC2	RPAV3	RPDC3	RPT3	TMR3	TMR3	RRAV3	RRDC3	RRT3	10	7
	TMI	TMI1	IPGC2	RPAV3	RPDC3	RPT3	TMI2	TMI2	RRAV3	RRDC3	RRT3	5	2
		TRP	IPGC2	RPAV3	RPDC3	RPT3	TREP2	TREP2	RRAV1	RRDC3	RRT3	4	3
			IPGC	RPAV3	RPDC3	RPT3	IPGC2	IPGC2	RRAV3	RRDC3	RRT3	6	4
				RPAV	RPDC3	RPT2	RPAV1	RPAV3	RRAV3	RPAV2	RRT3	23	15
					RPDC	RPDC3	RPDC2	RPDC1	RRAV3	RRDC2	RRT2	21	14
						RPT	RPT2	RPT2	RRAV2	RRDC1	RPT2	18	12
							IGT	IGT2	RRAV1	RRDC1	RRT1	2	1
								QRPR	RRAV1	QRPR1	RRT1	1	1
									RRAV	RRAV2	RRT1	22	15
										RRDC	RRT1	16	11
											RRT	23	15
												151	100

Fonte: Elaborado pelo autor.

$$Valor\ percebido = \left(\begin{array}{l} 7\%TMR, 2\%TMI, 3\%TRP, 4\%IPGC, 15\%RPAV, 14\%RPT, \\ 1\%IGT, 1\%QRPR, 15\%RRAV, 11\%RRDC, 15\%RRT \end{array} \right). \quad (3.1)$$

Na Tabela 3.4 tem-se a representação dos fatores de avaliação do serviço de manutenção.

Tabela 3.4 Fatores de avaliação do serviço de banda larga X escala de graduação da importância relativa para o cliente.

	FATORES DE PERCEPÇÃO DE VALOR	INDICADORES ASSOCIADOS	GRAU DE IMPORTANCIA (%)
1	Taxa de transmissão de Atendimento (VA)	Tempo Médio de Reparo (TMR)	7
2		Tempo Médio de Instalação (TMI)	3
3	Confiabilidade (CF)	Taxa de Reparo (TRP)	3
4		Instalação no Prazo Grandes Clientes (IPGC)	4
5		Reparo no Prazo Alto Valor (RPAV)	15
6		Reparo no Prazo Demais Clientes (RPDC)	14
7		Reparo no Prazo Total (RPT)	12
8	Consistência (CS)	Instalação em Garantia Total (IGT)	1
9		Reparo Preventivo (QRPR)	1
10		Reparo Reincidente Alto Valor (RRAV)	15
11		Reparo Reincidente Demais Clientes (RRDC)	11
12		Reparo Reincidente Total (RRT).	15

Fonte: Elaborado pelo autor.

As ações gerenciais previstas e baseadas nos critérios de percepção do valor junto aos clientes devem ainda ser analisadas do ponto de vista da existência de possíveis conflitos, segundo a equação final do valor com as quais o método da análise de conflitos, sugerido por GIANESI e CORRÊA, (1994). Como a definição e caracterização de uma série histórica no espaço temporal requerido dos custos não são valores fáceis de se identificar, os conflitos entre os indicadores não puderam ser analisados e, conseqüentemente, percebidos.

Para o estabelecimento das ações gerenciais não são suficientes o conhecimento dos "atributos de valor" priorizados e sua análise de conflito. As ações gerenciais normalmente são aplicadas em processos gerenciais, portanto, é preciso correlacionar os "atributos de valor" priorizados com os processos de manutenção.

Essa correlação é realizada nesse passo, por meio do método sugerido por GIANESI e CORRÊA, (1994) descrito no Capítulo 2, que classifica os processos prioritários como áreas de decisão do negócio, ou seja, (Gestão de Material, Gestão de RH, Conceituação de Manutenção, Organização Operacional, Gestão do Cliente, Sistema de Informação e Planejamento e Controle), os quais são descritos especificamente para o serviço de manutenção.

Na Tabela 3.5 mostra-se a correlação entre os fatores de percepção de valor que foram priorizados e os processos críticos de um sistema de operações.

Tabela 3.5 “Atributos de Valor” versus processos críticos nas operações de manutenção.

	TMR	TMI	TREP	IPGC	RPAV	RPDC	RPT	IGT	QRPR	RAV	RRDC	RRT
Gestão de materiais	●	●	●	●	●	●	●	○	○	○	○	○
Gestão de RH	●	●	●	●	●	●	●	●	●	●	●	●
Conceito de Manutenção	○	○	●	●	●	●	●	●	●	●	●	●
Organização Operacional	●	●	●	●	●	●	●	●	●	●	●	●
Sistema de Informação	●	●	●	●	●	●	●	○	○	○	○	○
Gestão do Cliente	○	○	●	●	●	●	●	○	○	○	○	○
Planejamento Controle	○	○	●	●	●	●	●	●	●	●	●	●

Fonte: Elaborado pelo autor.

○ Relação Fraca ○ Relação Média ● Relação Forte

A avaliação de desempenho é também necessária ao estabelecimento das ações gerenciais para operacionalização do valor proposto, na medida em que quantifica e define um valor numérico para o desempenho dessas operações. O objetivo dessa etapa é medir instantaneamente o desempenho do serviço de manutenção, de forma a identificar quais os "atributos de valor" devem ser focados, baseado na comparação com a priorização realizada

anteriormente. Além disto, essa avaliação de desempenho resulta em uma noção quantitativa do valor proposto.

C.3)Identificação das variáveis independentes e da variável dependente

Após análise dos dados conforme os critérios anteriormente descritos, conclui-se que se deve trabalhar, nessa pesquisa, com 12 variáveis independentes, as quais foram consideradas as mais representativas segundo os atributos de valor (definidos anteriormente), quais sejam: Tempo Médio de Reparo (TMR), Tempo Médio de Instalação (TMI), Taxa de Reparo (TRP), Instalação no Prazo Grandes Clientes (IPGC), Reparo no Prazo Alto Valor (RPAV), Reparo no Prazo Demais Clientes (RPDC), Reparo no Prazo Total (RPT), Instalação em Garantia (IGT), Reparo Preventivo (QRPR), Reparo Reincidente Alto Valor (RRAV), Reparo Reincidente Demais Clientes (RRDC), Reparo Reincidente Total (RRT).

Devido à natureza do problema, é selecionada a taxa de abandono como a variável dependente, tendo em vista que o objetivo dessa pesquisa é verificar a correlação das outras variáveis com a taxa de abandono. Para melhor entendimento dessa variável é feito, a seguir, uma breve descrição das suas características, sendo importante salientar que os dados referentes à taxa de abandono da planta são confidenciais.

C.4) Caracterização da Taxa de Abandono

Conhecendo a estrutura do mercado de banda larga no Brasil, é fácil perceber que ele apresentou um crescimento muito rápido e ainda está em fase de crescimento, diferentemente de mercados mais antigos como o de telefonia fixa. A concorrência entre empresas de banda larga, empresas de cabo e empresas que provêm o serviço rádio, fazendo desse mercado competitivo e dinâmico. De qualquer forma, à medida que o mercado se expande e o número de clientes realmente desprovidos da tecnologia banda larga diminui, em algum tempo só restará um único tipo de consumidor para as empresas atraírem: aquele que já utiliza o serviço de outro concorrente. Nesse ponto, surge a preocupante questão da taxa de abandono que já custa milhões de reais por ano às principais empresas do País. Por esse indicador quanto menor seu valor, melhor.

A) O Significado do Taxa de Abandono

Segundo MATT, (2001) a taxa de abandono consiste no ato de um cliente abandonar uma empresa em favor de um concorrente, terminando toda a sua relação com a empresa antiga e iniciando uma nova relação com uma outra. A origem do termo, muito provavelmente, se encontra no sentido do verbo *to churn* na língua inglesa, que significa “mexer, agitar violentamente”. O fenômeno da taxa de abandono nas empresas de banda larga causa exatamente o que o verbo quer dizer: uma grande “agitação” de clientes no mercado, trocando de empresa a todo o momento, o que, por sua vez, leva as empresas a se “mexerem” em busca de novas formas de manter seus clientes no seu negócio, ao mesmo tempo em que buscam seduzir os clientes da concorrência.

B) Possíveis Causas

Existem dois tipos de *taxa de abandono*: o voluntário e o involuntário. O *taxa de abandono* voluntário ocorre quando o cliente por vontade própria termina os serviços com a empresa [FERREIRA, (2005)]. Podendo ainda ser dividido em *taxa de abandono* deliberado, quando o consumidor possui razões relacionadas diretamente ao seu uso da tecnologia para trocá-la por uma concorrente, ou *taxa de abandono* acidental, causado por razões que fogem ao controle do cliente, como mudanças de localidade de moradia, problemas financeiros pessoais e outras grandes mudanças de vida. O *taxa de abandono* involuntário é resultado de uma ação da própria empresa, que, por motivos de fraude, falta de pagamentos e até mesmo utilização do produto indevidamente, pode vir a romper sua relação com o cliente. Na Tabela 3.6 apresenta-se as principais causas para o *taxa de abandono* em uma empresa.

Tabela 3.6 Causas da *taxa de abandono* em uma empresa.

VOLUNTÁRIO	DELIBERADO	TECNOLOGIA	INVOLUNTÁRIO	FRAUDE
		ECONOMIA		
		QUALIDADE		NÃO-PAGAMENTO
		SOCIO PSICOLOGICO		
		CONVENIENCIA		
ACIDENTAL	FINAÇAS	UTILIIZAÇÃO FOR A DO ESCOPO CONTRATUAL		
	LOCALIDADE			
	GRANDES MUDANÇAS			

Fonte: Adaptada de FERREIRA, (2005).

Como se pode perceber, a *taxa de abandono* voluntário deliberado pode variar muito, indo desde a insatisfação com o serviço adquirido, podendo ter suas causas na qualidade de serviço prestado, passando por problemas com o valor pago pelo serviço, e chegando até mesmo a ter causas como a influência de família e amigos insistindo para que o usuário troque de empresa.

C) Relação entre retenção do cliente e lucratividade

Uma correta gestão da qualidade dos serviços pode produzir um pacote de valor que gere níveis de satisfação que garantam a permanência do cliente na empresa criando assim um laço de fidelidade. Um cliente satisfeito é um agente de *marketing* da empresa, fazendo propaganda boca a boca e divulgando sua satisfação em um círculo de influência, auxiliando a ampliar a conquista de novos clientes. Porém um cliente insatisfeito tem um potencial mais devastador ainda, pois segundo HESKETT et al., (1994) um cliente insatisfeito tende a reportar sua insatisfação para outros 11 clientes. É importante salientar que essa pesquisa ocorreu antes da Internet por esse motivo os números hoje devem ser muito mais devastadores.

Um dos modelos mais presentes na literatura é o que relaciona o grau de satisfação do cliente a uma comparação entre a expectativa que ele tem, antes de passar pela experiência do serviço, e a percepção obtida depois que passou pela experiência do serviço [CORRÊA, (2002)].

As expectativas do cliente são formadas com base em alguns fatores como a necessidade e desejos do cliente a sua expectativa à comunicação que foi feita boca a boca e finalmente o preço do serviço [ZEITHAML et al., (1990)].

D) Custo da taxa de abandono de cliente

Para analisar o custo da taxa de abandono deve-se dividir a questão em quatro etapas conforme a análise a seguir [KOTLER, (1996)]:

- Inicialmente deve-se definir e mensurar a taxa de abandono;
- Em seguida, devem ser distinguidas as várias causas que levam um cliente a abandonar a empresa;
- Estima-se o volume de lucro perdido com a taxa de abandono;
- Para por último, calcula-se quanto custaria reduzir a taxa de abandono.

3.3 Planejamento da Pesquisa (ESTÁGIO 2)

Os passos iniciais foram dados na direção de se entender o problema da pesquisa, selecionar os objetivos e explicar a taxa de abandono por meio da regressão múltipla. A definição de variáveis dependentes e independentes, no planejamento de uma pesquisa, faz com que o especialista considere questões como o tamanho da amostra e a natureza das variáveis, que deve ser feito com base no entendimento do negócio, do problema e das informações existentes que possam contribuir para uma melhor caracterização.

A obtenção dos dados é feita por meio da coleta em um sistema de tratamento ao cliente de uma empresa de telecomunicações sempre tendo em vista a confiabilidade dos dados. É importante ressaltar que nessa primeira etapa, entende-se que para um melhor desempenho seria fundamental que as empresa armazenassem os dados em uma *data*

warehouse. Nesse caso não seria necessária a coleta mensal e armazenamento feito em planilhas para só depois se realiza a limpeza e o tratamento.

A determinação do tamanho da amostra é complexa e envolve várias considerações de ordem quantitativas e qualitativas [MALHOTRA, (2004)]. Os fatores qualitativos mais importantes na determinação desse tamanho são: importância da decisão, natureza da pesquisa, o número de variáveis, natureza da análise, os tamanhos amostrais em análises similares, as taxas de incidência, taxas de preenchimento e a restrição de recursos.

De um modo geral, para decisões mais importantes é necessário obter um número maior de informações e essas informações devem ser obtidas de maneira mais precisa, sendo que uma quantidade maior de amostra requer um custo maior da obtenção dessas observações.

A) Tamanho da Amostra

No planejamento de uma pesquisa devem ser consideradas questões como o tamanho da amostra e a natureza das variáveis independentes. O tamanho da amostra tem impacto direto sobre a adequação e o poder estatístico na análise múltipla de variáveis. Amostras pequenas, não são indicadas para análise múltiplas. Do mesmo modo, amostras muito grande, mais de 1000 observações, tornam os testes de significância estatística sensíveis, muitas vezes indicando que qualquer relação seja estatisticamente significativa. Com amostras muito grandes, deve-se garantir que o critério de significância prática seja atendido junto com a significância estatística. O poder é influenciado diretamente e em magnitude pelo tamanho da amostra.

Nessa pesquisa procurou-se medir e verificar o impacto direto e a magnitude considerável sobre o poder, que o tamanho da amostra poderá afetar a pesquisa, ou seja, a probabilidade de se detectar como estatisticamente significativa um nível específico de R^2 . Para isso Foi coletado um total de 168 observações ocorridas em uma grande empresa do setor de telecomunicações, que atua em 14 estados nas regiões nordeste e sudeste do País.

No intuito de responder a questão sobre o tamanho da amostra é utilizada a Tabela estatística do anexo II, na qual o nível de significância $\alpha = 0,05$ e 12 variáveis independentes com 168 observações é possível detectar um o valor de R^2 maior ou igual a igual a 15 % a um poder de 0,80.

Além do poder estatístico, o tamanho da amostra afeta a generalização dos resultados pela proporção entre observações e variáveis independentes. Existe uma regra geral para que a razão observações/variáveis independentes não deve ser menor que 5 para 1, sendo o nível desejado de 15 a 20 observações por variável independente. Quando esse nível é alcançado, os resultados devem ser generalizáveis se a amostra é representativa. Para o caso dessa pesquisa, se tem: Razão (Observações/ Variáveis Independentes) = $168/12 = 14$ observações por variável independente.

B) Generalização no tamanho da amostra

Além do papel na determinação do poder estatístico, o tamanho da amostra afeta a generalização dos resultados pela proporção entre as observações e variáveis independentes.

Segundo HAIR et al., (2005) uma regra geral é que a razão jamais deve ficar abaixo de 5 para 1, o que significa que deve haver cinco observações para cada variável independente na variável estatística. Se as quantidades de observações forem inferiores a cinco existirá o risco de super ajustar a variável estatística com a amostra, perdendo a generalização. Apesar da relação mínima aconselhada de 5 para 1, o nível desejado para a relação é de 15 a 20 observações para variável independente.

3.4 Suposições em Análise de Regressão Linear Múltipla (ESTÁGIO 3)

Todas as técnicas multivariadas têm suposições inerentes, estatísticas e conceituais, que influenciam muito suas habilidades para representar relações multivariadas. Para as técnicas baseadas em inferência estatística, as suposições de normalidades multivariadas, linearidade, independência de termos de erro e igualdade de variâncias em uma relação de dependência devem ser satisfeitas.

As suposições inerentes à análise de regressão múltipla se aplicam as variáveis individuais e a relação como um todo. No capítulo 2 examinou-se os métodos disponíveis para avaliar as suposições para as variáveis individuais. No entanto, na regressão múltipla, uma vez que a variável estatística tenha sido determinada, ela atua coletivamente na previsão da variável dependente. Com isto é necessário avaliar as suposições não apenas para as variáveis

individuais, mas também para a variável estatística em si. Para isto deve-se examinar a variável estatística e a sua relação com a variável dependente para atender as suposições da regressão múltipla. A questão básica é se, no processo de cálculo dos coeficientes de regressão e de explicação da variável dependente, as suposições da análise de regressão serão atendidas.

As suposições a serem atendidas são:

- Linearidade do fenômeno medido;
- Variância constante dos termos de erro;
- Independência dos termos de erro;
- Normalidade da distribuição dos termos de erro.

Linearidade do fenômeno medido – Usada para expressar o conceito de que o modelo possui as propriedades de aditividade e homogeneidade. A linearidade da relação entre variáveis dependentes e independentes representa o grau em que a variação na variável dependente é associada a variável independente. Em termos gerais, os modelos lineares prevêem valores que recaem em uma linha reta que tem uma mudança com unidades constantes (coeficiente angular) da variável dependente em relação a uma mudança com unidade constante da variável independente. A linearidade pode ser examinada por meio de gráficos de resíduos.

O conceito de correlação é baseado em uma relação linear, o que a torna uma questão crucial na análise de regressão. A linearidade pode ser examinada por meio do gráfico de resíduo. Qualquer padrão curvilíneo consistente nos resíduos indicará a existência de uma relação não-linear. Em regressão múltipla com mais de uma variável independente um exame dos resíduos mostra os efeitos combinados de todas as variáveis independentes, mas não se podem ser examinados todas as variáveis independentes separadamente. Para isto se utiliza os gráficos de regressão parcial, os quais mostram a relação de uma única variável independente com a dependente.

Linearidade é o fenômeno que relaciona as variáveis dependentes e independentes representando o grau em que a variável dependente é associada a variável independente [HAIR, (2005)]. Na regressão múltipla uma vez que a variável estatística tenha sido

determinada, ela atua coletivamente na previsão da variável dependente, sendo necessárias avaliações das suposições não apenas para as variáveis individuais, mas também para a variável estatística em si.

Um resíduo é a diferença entre o valor observado de Y e o valor de Y prognosticado pela equação de regressão. A representação gráfica dos resíduos versus as variáveis independentes é um método básico para identificar violações de suposições para a relação geral [MALHOTRA, (2004)].

Em análise múltipla um exame dos resíduos mostra os efeitos combinados de todas as variáveis independentes, mas não se podem ser examinadas quaisquer variável independente separadamente em um gráfico de resíduos. Para isto utiliza-se o gráfico de regressão parcial, os quais mostram a relação de uma única variável independente com a dependente. Eles diferem dos gráficos de resíduos, pois a reta que passa pelo centro dos pontos que era horizontal nos gráficos de resíduos, agora tem uma inclinação para cima ou para baixo, dependendo do coeficiente de regressão para aquela variável independente ser positivo ou negativo.

O gráfico de dispersão, mais conhecido no sistema *SPSS* como *scatterplot* é um gráfico de pontos que representa em um plano (x, y) n pares de valores numéricos escalares, sendo possível com isto analisarem a distribuição escalar de duas variáveis. Esse gráfico permite analisar *outliers*, observações fora do padrão, apresentando no eixo vertical a variável dependente e no eixo horizontal as variáveis independentes.

A seguir nos gráficos 3.3 e 3.4 são analisadas as características da regressão parcial para as 12 variáveis independentes escolhidas.

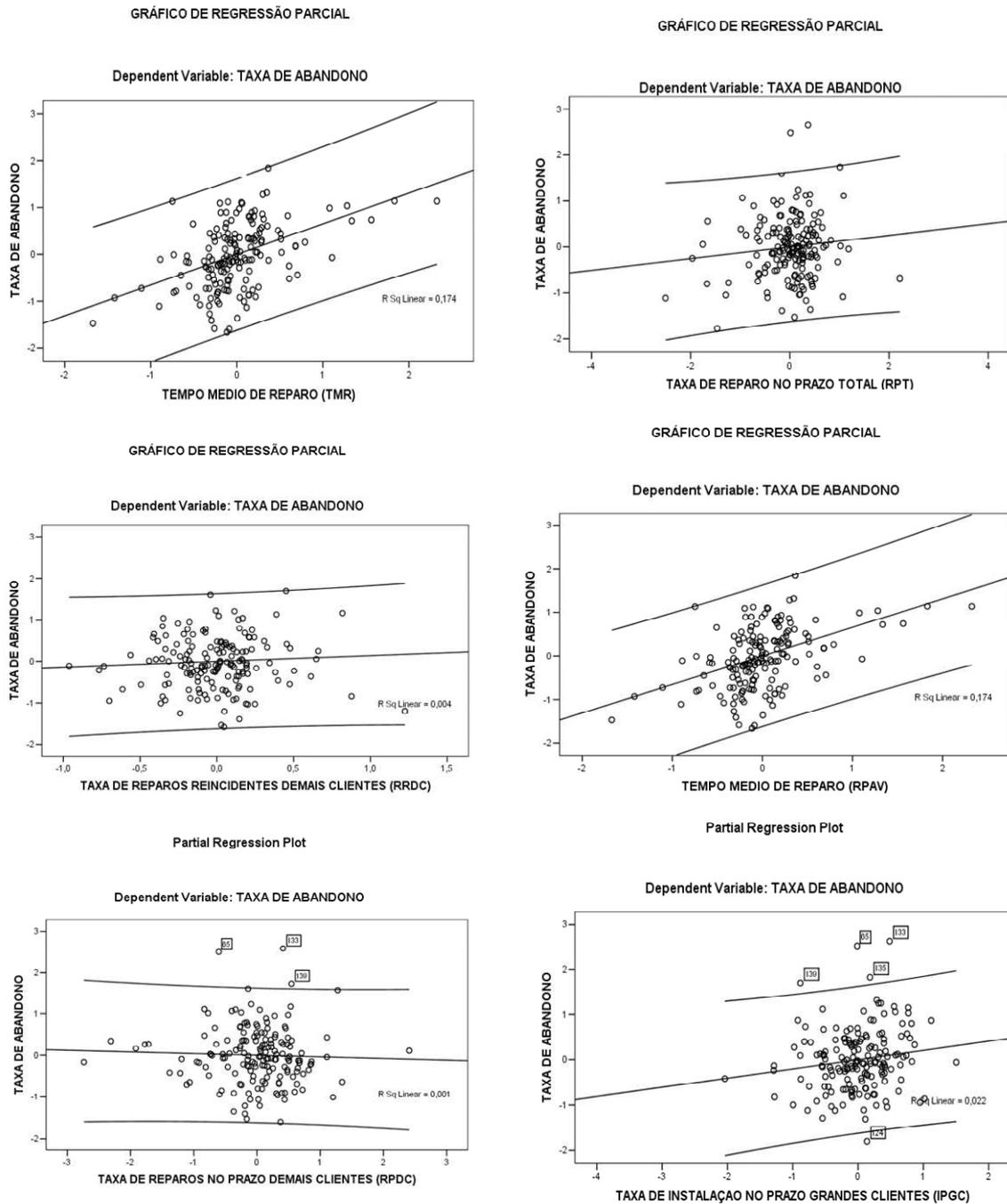


Gráfico 3.3 Análise dos gráficos de regressão parcial (TMR, TRP, QRPR, TMI, IGT, RRAV).

Fonte: Adaptado do SPSS.

GRÁFICO DE REGRESSÃO PARCIAL

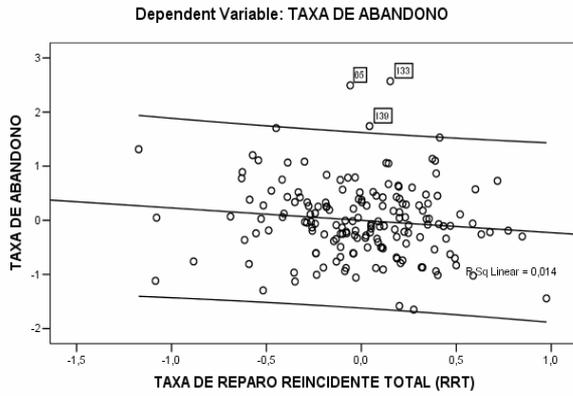


GRÁFICO DE REGRESSÃO PARCIAL

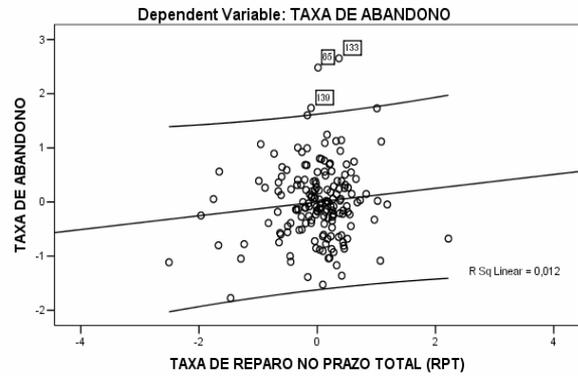


GRÁFICO DE REGRESSÃO PARCIAL

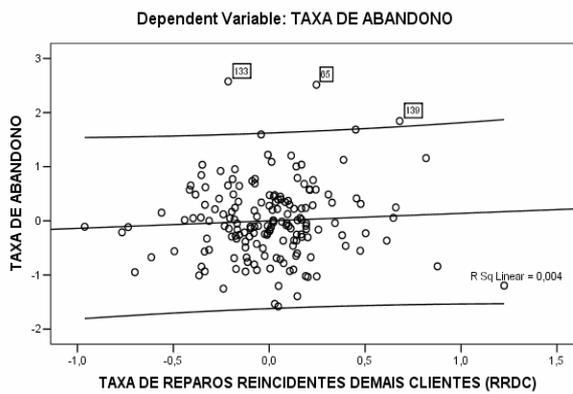
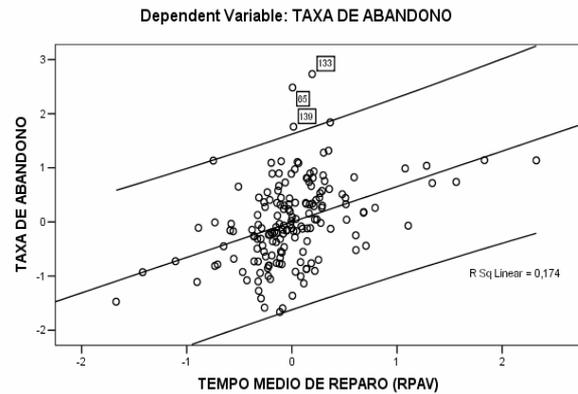
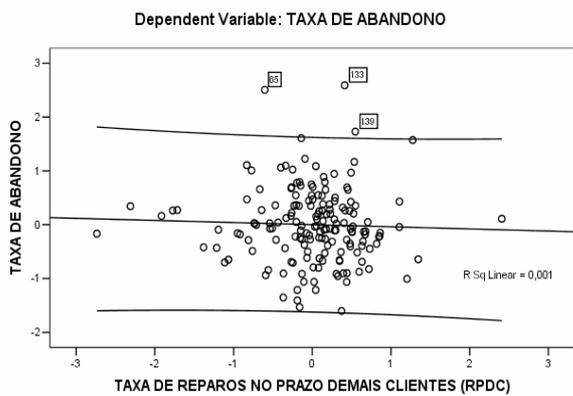


GRÁFICO DE REGRESSÃO PARCIAL



Partial Regression Plot



Partial Regression Plot

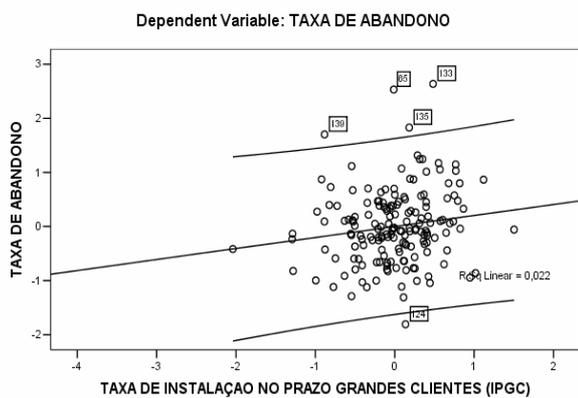


Gráfico 3.4 Análise dos gráficos de regressão parcial (RRT,PT,RRDC,RPAV,RPDC,IPGC).

Fonte: Adaptado do SPSS.

Após análise visual verifica-se que os gráficos dos indicadores possuem retas que passam pelo centro dos pontos, tendo uma inclinação para cima ou para baixo dependendo do coeficiente de regressão para a variável independente.

Variância Constante do Termo de Erro – A presença de variâncias desiguais (heteroscedasticidade) é também uma das violações comuns de suposições, sendo o seu diagnóstico feito com gráficos de resíduos ou teste estatístico simples. A diferença entre a média e um valor observado é o desvio em relação a media. A variância é o desvio quadrático médio em relação a media, esse valor nunca pode ser negativo [MALHOTRA, (2004)]. Quando os dados se apresentam agrupados em torno da media, a variância é pequena. Quando os dados estão dispersos, a variância é grande.

Quando a variância dos erros parece constante ao longo de um domínio de variáveis predictoras, diz-se que os dados são homoscedásticos. A suposição de variância igual do erro da população é crítica para a aplicação correta da regressão linear. Quando os termos de erro têm variância crescente ou flutuante diz-se que os dados são heteroscedásticos. Conforme visto a análise de resíduos é uma boa forma de verificar.

A representação gráfica de resíduos (estudentizados) versus os valores dependentes previstos e a sua comparação com o gráfico nulo não mostra um padrão consistente se a variância for constante. Verifica-se ao analisar o gráfico 3.5 que a representação gráfica dos resíduos (estudentizados) versus os valores dependentes previstos e a sua comparação com um gráfico nulo mostra não um padrão consistente, ou seja, a variância é constante.

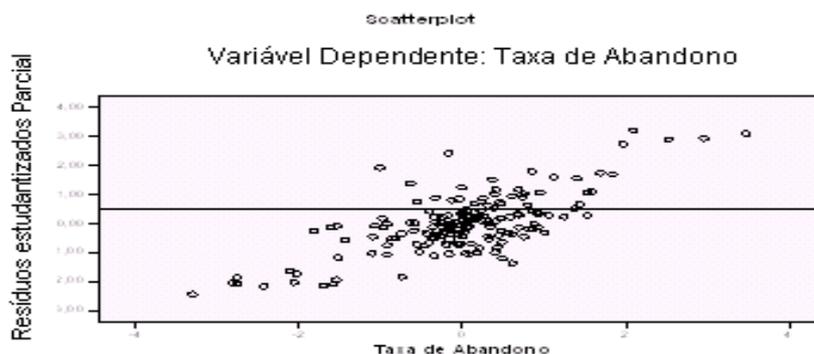


Gráfico 3.5 Resíduos (estudentizados) versus os valores dependentes.
Fonte: SPSS.

Independência dos Termos de Erro – Assume-se que cada valor previsto é independente. Com isto, o valor previsto não é relacionado com qualquer outra previsão, ou seja, eles não são seqüenciados por qualquer variável. Pode-se identificar melhor tal ocorrência fazendo o gráfico de resíduos em relação a qualquer variável seqüencial possível. Se os resíduos forem independentes, o padrão deveria parecer aleatório e semelhante ao gráfico nulo dos resíduos, sendo as violações identificadas por um padrão consistente nos resíduos.

O Gráfico 3.6 mostra-se os resíduos em relação a qualquer variável seqüencial possível, ou seja, em relação valor previsto, nos mostra que os resíduos são independentes, tendo um padrão aleatório semelhante ao gráfico nulo de resíduos, com isto fica caracterizado à independência dos termos de erro.

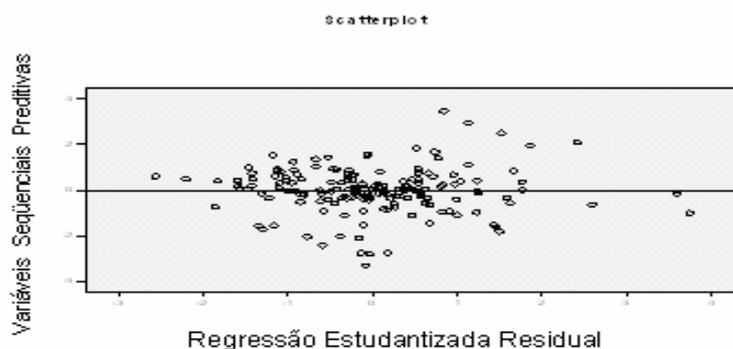


Gráfico 3.6 Resíduos em relação a qualquer variável seqüencial possível.

Fonte: SPSS.

Normalidade da Distribuição dos Termos de Erro – A normalidade da distribuição dos termos de erro, segundo HAIR, (2005) e MALHOTRA, (2004) é a violação de suposição mais freqüentemente encontrada.

O diagnostico mais simples para o conjunto de variáveis independentes na equação é um histograma de resíduos, com uma verificação visual para a distribuição que se aproxima da normal. Uma outra forma é analisar o gráfico de probabilidade da normal onde a distribuição normal forma uma reta diagonal e os resíduos graficamente representados são comparados com a normal.

O gráfico 3.7 mostra a correlação entre a variável estatística e a variável dependente. Verifica-se nesse gráfico sua forte correlação e que a inclinação da regressão mede a direção e

a magnitude da relação. Quando as duas variáveis estão correlacionadas positivamente, a inclinação é positiva conforme é verificado no gráfico. Mostrado no Gráfico 3.7. Normal dos resíduos verifica-se uma tendência de distribuição uniforme, os pontos encontram-se bastantes próximos de uma reta, sem desvios substanciais ou sistêmicos de modo que não há razão para duvidar da normalidade dos erros, sendo assim os resíduos são representativos de uma distribuição normal.

A seguir são mostrados os gráficos de dispersão e o histograma da variável dependente em relação resíduo padrão.

Gráfico de Probabilidade Normal dos Resíduos

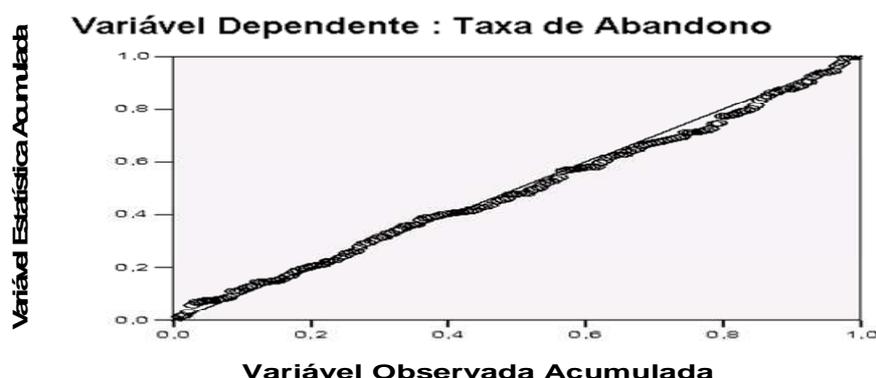


Gráfico 3.7 Normal dos Resíduos.

Fonte: SPSS.

O gráfico 3.8 mostra a distribuição normal dos dados e a sua correlação entre os dados coletados e a Taxa de Abandono, como se verifica visualmente existe uma distribuição que se aproxima da normal.

Histograma de Resíduos da Variável Dependente

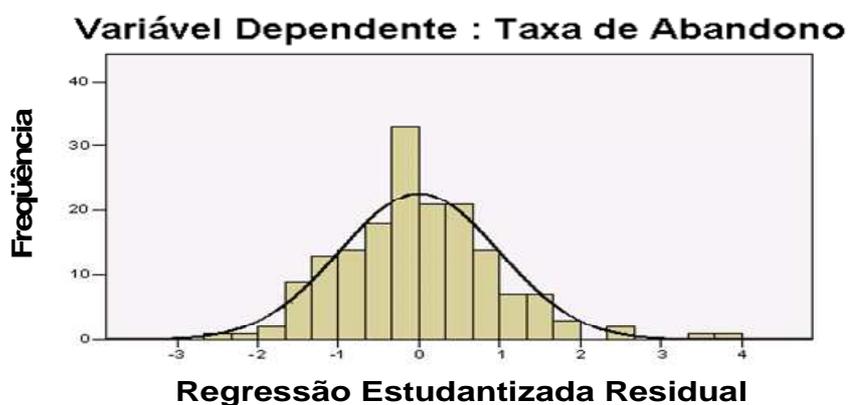


Gráfico 3.8. Histograma da taxa de abandono.

Fonte: SPSS.

Na Tabela 3.7 apresenta-se as descrições estatísticas da variável dependente e das independentes após a normatização, pelo método do desvio padrão.

Tabela 3.7 Descrições Estatísticas.

Quantidade	Indicador	Mínimo	Máximo	Média	Desvio Padrão	Variância
1	RRAV	6,80	99,6	21,3	16,6	277
2	IGT	0,52	97,9	18,0	24,7	612
3	RRT	6,52	34,8	17,9	5,16	26,6
4	RPT	71,8	100	96,6	3,75	14,1
5	IPGC	000	100	88,1	12,7	161,8
6	RRDC	5,74	38,3	18,2	5,25	27,5
7	RPAV	16,9	100	83,5	23,4	549
8	RPDC	57,5	100	97,9	3,65	13,34
9	TMR	000	5,55	4,66	0,70	6,00
10	Taxa de Abandono	000	0,05	0,01	0,01	0000
11	Taxa Reparo (TRP)	0,02	0,24	0,06	0,03	0,001
12	PREVENTIVAS (QRPR)	0,08	0,16	0,12	0,13	0,007
13	Tempo de Instalação (TMI)	000	100	88,8	12,5	156

Fonte: SPSS.

Analisando os valores da curtose para os indicadores na Tabela 3.8, verifica-se que a mesma tem valores positivos, ou seja, os dados têm uma distribuição relativamente elevada comparada com a distribuição normal. Enquanto as taxas de instalação em garantia e do taxa de abandono tem uma distribuição positivamente assimétrica, ou seja, um arco simples acima da diagonal, a taxa de instalação no prazo de grandes contas tem um arco simples abaixo da diagonal tendo uma distribuição assimetricamente negativa, porém nenhum dos indicadores teve valores fora do intervalo entre 1 e -1. Analisando a curtose verifica-se que todos os valores são positivo, o que indica que os indicadores têm uma distribuição relativamente elevada comparada com a normal.

Para padronizarmos os dados utilizou-se a técnica de Normalização por Desvio Padrão normalmente conhecida como *Z-Score* ou *Z-Mean*, considera na posição media dos valores de um atributo, assim como os graus de dispersão desses valores em relação à posição média. Passa-se a analisar os dados gerados por um sistema de análise estatística. Essa normalização foi feita por meio da função a seguir:

$$Z_{if} = \frac{(X_{if} - X_f)}{(S_f)}. \quad (3.2)$$

Em que: X_{if} – Valor a ser normalizado, X_f – Média e S_f – Desvio Padrão.

Tabela 3.8 Descrições Estatística Normalizada.

	Média	Variância	Assimetria		Curtose	
	Estatística	Estatística	Estatística	Erro Padrão	Estatística	Erro Padrão
TAXA DE ABANDONO	0.00	1.00	0.48	0.19	1.46	0.37
RRAV	0.00	1.00	0.85	0.19	0.94	0.37
IGT	0.00	1.00	0.24	0.19	0.79	0.37
RRT	0.00	1.00	0.73	0.19	0.89	0.37
RPT	0.00	1.00	-1.63	0.19	2.58	0.37
RRDC	0.00	1.00	0.89	0.19	1.75	0.37
IPGC	0.00	1.00	-1.00	0.19	1.66	0.37
RPAV	0.00	1.00	-1.73	0.19	3.08	0.37
RPDC	0.00	1.00	-2.74	0.19	8.61	0.37
TMR	0.00	1.00	-0.96	0.19	0.33	0.37
TRP	0.00	1.00	1.64	0.19	4.18	0.37
QRPR	0.00	1.00	2.20	0.19	5.41	0.37
TMI	0.00	1.00	-0.96	0.19	0.33	0.37

Fonte: SPSS.

Após ter sido especificado os objetivos, selecionado as variáveis, avaliado se as variáveis atendem as suposições, deve ser trabalhado o modelo de regressão. Nesse estágio deve-se, selecionar o método para a especificação, avaliar a significância estatística do modelo e determinar se algumas das observações exercem uma influencia indevida nos resultados. Esses tópicos serão detalhados conceitualmente em seguida.

3.5 Estimação do Modelo de Regressão e Avaliação do Ajuste Geral do Modelo (ESTÁGIO 4)

Após terem sido especificados os objetivos da análise de regressão, selecionando as variáveis dependentes e independentes, abordado as questões de planejamento da pesquisa e tendo sido avaliado se as variáveis atendem às suposições da regressão, deve-se ajustar o modelo de regressão e avaliar a precisão explicativa geral das variáveis independentes. Nesse estágio devem ser cumpridas três tarefas básicas: selecionar um método para especificar um

modelo, avaliar a significância estatística do modelo e determinar se alguma das observações exerce uma influência indevida nos resultados.

3.6.1 Seleção de um Método para Especificar um Modelo

Na maioria dos casos, a pesquisa tem várias possíveis variáveis independentes que podem ser escolhidas para inclusão na equação de regressão. Algumas vezes, o conjunto de variáveis independentes pode ser especificado, outras vezes não. Para selecionar um método para especificar o modelo a ser ajustado na pesquisa tem-se que considerar a abordagem a ser escolhida. Algumas vezes, o conjunto de variáveis pode ser especificado e o modelo é essencialmente usado com uma abordagem confirmatória. Em outros casos, podem ser escolhidas algumas variáveis em um conjunto. Existem vários métodos adequados para a tarefa conforme abaixo:

Método Especificações Confirmatórias – É a abordagem mais simples e mais exigente, pois deve ser especificado completamente o conjunto de variáveis independentes a serem utilizadas. Fica na responsabilidade do pesquisador assegurar de que o conjunto de variáveis atinja a previsão máxima. Nessa abordagem se tem total controle sobre a seleção das variáveis, devendo ser assegurado que o conjunto de variáveis atinja a previsão máxima, embora mantendo um modelo parcimonioso.

Método Busca Sequencial – Os métodos de busca sequencial têm em comum a abordagem de estimar a equação com um conjunto de variáveis e então acrescentar seletivamente ou eliminar variáveis até que alguma medida de critério geral seja alcançada. Essa abordagem fornece um método objetivo para selecionar variáveis que maximiza a previsão com o menor número de variáveis empregadas. Há três tipos de abordagem de busca sequencial: *Stepwise e Forward* (adição) e *Backward* (eliminação), as quais serão descritas abaixo:

- **Estimação *Stepwise*** – Talvez seja a abordagem mais comum para a seleção sequencial, pois ela permite examinar a contribuição de cada variável independente para o modelo de regressão sendo que cada variável é considerada para a inclusão antes do desenvolvimento da equação. A variável independente com a maior contribuição é acrescentada em um primeiro

momento. Variáveis independentes são então selecionadas para inclusão, com base em sua contribuição incremental sobre as variáveis na equação. Começa com um modelo de regressão simples no qual a variável mais fortemente correlacionada com a dependente é usada. Examina os coeficientes de correlação parcial para encontrar uma variável adicional que explique a maior parte estatisticamente significativa do erro da primeira equação. Recalcula a equação de regressão usando as duas variáveis independentes sendo então examinando o valor parcial de F. Esse procedimento continua examinando todas as variáveis independentes no modelo para determinar se alguma deveria ser incluída na equação;

- Estimação *Forward e Backward* – Processo de tentativa e erro para encontrar a melhor estimativa de regressão, enquanto o *Forward* vai adicionado as variáveis uma a uma e verificando a equação, o método *Backward* computa a equação com todas as variáveis e depois vai eliminando as que não contribuem significativamente como o processo. A principal diferença da abordagem *Stepwise* em relação aos procedimentos *Forward e Backward* é a sua habilidade em acrescentar ou eliminar variáveis em cada estágio. Uma vez que uma variável é acrescentada ou eliminada nos esquemas de *Forward e Backward* não há como reverter à ação em um estágio posterior.

Método Abordagem Combinatória – É um processo de busca generalizado em todas as possíveis combinações de variáveis independentes. O procedimento mais conhecido é a regressão em todos os possíveis subconjuntos, a qual é exatamente como o nome sugere. Todas as possíveis combinações das variáveis independentes são examinadas e o conjunto de variáveis que melhor se adequar é identificado.

3.6.2 Avaliação da Significância Estatística do Modelo

Para avaliação da significância do modelo estatístico inicia-se testando a hipótese que a quantia de variação explicada pelo modelo de regressão é maior que a variação explicada pela média, ou seja, que R^2 é maior que zero, sendo a estatística F usada para esse fim. Se o

tamanho da amostra for pequeno, o erro amostral poderá ser muito grande para se dizer, com um grau de certeza necessário, que o coeficiente não é igual à zero.

Significância do modelo geral: Para testar a hipótese de que a quantia de variação é maior que a variação explicada pela média (ou seja, que R^2 é maior que zero), usa-se a estatística F, onde:

$$\text{Estatística } F = \frac{\frac{\sum \text{dos quadrados dos erros}}{\text{Graus de liberdade da regressão}}}{\frac{\sum \text{dos quadrados dos erros totais}}{\text{Graus de liberdade dos erros totais}}}$$

Dois aspectos importantes sobre essa razão devem ser observados:

- Cada soma de quadrados dividida por seus graus de liberdades apropriados é simplesmente a variância dos erros de previsão;
- Se a razão da variância explicada pela variância de referência for alta, a variância estatística da regressão deve ser de valor significativo na explicação da variável dependente.

Também se sabe da influência do tamanho da amostra no valor de R^2 . Para resolver isto têm sido proposta várias normas práticas em que deve ocorrer uma variação de 10 a 15 observações por variável independente. Como parte integrante dos softwares de regressão, encontra-se o cálculo do R^2 ajustado, esse coeficiente deve ser analisado igualmente ao coeficiente de determinação, porém o valor de R^2 ajustado, leva em consideração o número específico de variáveis independentes e o tamanho da amostra em que cada modelo é baseado, sendo particularmente útil na comparação entre equações de regressão.

Os testes de significância de coeficiente de regressão fornecem uma avaliação empírica de seu verdadeiro impacto. Apesar desse não ser um teste de validade, ele determina se os impactos representados pelos coeficientes são generalizáveis a outras amostras dessa população.

3.6.3 Identificação de Observações Influentes

Até agora, a pesquisa foi concentrada na identificação de padrões gerais no conjunto inteiro de observações. Agora se desvia a atenção para observações individuais, com o objetivo de encontrar quais observações estão fora dos padrões gerais do conjunto de dados ou que influenciam fortemente nos resultados da regressão. Deve ser lembrado que essas observações não são necessariamente ruins, no sentido que devam ser eliminadas. No entanto deve-se primeiramente identificá-las e avaliar seu impacto antes de seguir adiante. As observações influentes são de três tipos básicos: Atípicas, Pontos de alavancagem e Influentes. As observações atípicas são identificadas apenas em relação a um modelo específico de regressão. As observações atípicas tem sido tradicionalmente a única forma de observação influente considerada em modelos de regressão.

Os pontos de alavancagem são observações diferentes das demais, com base em seus valores para variáveis independentes. Seu impacto é particularmente perceptível nos coeficientes estimados para uma ou mais variáveis independentes. Finalmente, as observações influentes é a categoria mais ampla, incluindo todas as observações que tem um efeito desproporcional sobre os resultados de regressão. As observações influentes incluem tanto as observações atípicas e os pontos de alavancagem.

Os procedimentos para identificar todos os tipos de observações influentes são muito numerosos e ainda não muito bem definidos. Todos os softwares fornecem uma análise de resíduos com os quais, aqueles com grandes valores podem ser facilmente identificados. As observações influentes, as atípicas e os pontos de alavancagem são baseados em quatro condições:

- Um erro em observações ou entrada de dados;
- Uma observação válida, mas excepcional, explicável por uma situação extraordinária;
- Uma observação excepcional sem explicação convincente;
- Uma observação comum em suas características individuais, mas excepcional em sua combinação de característica.

Para análise final, tenta-se identificar observações atípicas que sejam influentes e determinar se elas deveriam ser excluídas ou não. Apesar de procedimentos mais detalhados

estarem disponíveis para identificar observações atípicas, usa-se nessa pesquisa os resíduos para identificar observações atípicas.

3.6 Interpretação das Variáveis Estatísticas (ESTÁGIO 5)

Nesse estágio deve-se interpretar a variável estatística de regressão pela avaliação dos coeficientes de regressão estimados em termos de sua explicação da variável dependente. Avalia-se com isto não apenas o modelo de regressão estimado, mas também as variáveis independentes que foram eventualmente omitidas se uma busca seqüencial ou abordagem combinatória foi empregada. Assim após avaliar os coeficientes estimados, deve-se também avaliar o impacto potencial de variáveis omitidas para garantir que a significância gerencial seja avaliada, juntamente com a significância estatística.

3.6.1 Avaliação da equação com os coeficientes de regressão

Os coeficientes de regressão estimados devem ser usados para calcular os valores previstos para cada observação e para expressar a variação esperada na variável dependente para cada variação unitária nas variáveis independente. A interpretação da variável estatística de regressão acontece com a estimação do modelo concluída, a variável estatística de regressão especificada e os testes diagnósticos que confirmam a adequação dos resultados administrativos, nessa etapa pode-se examinar a equação a qual inclui X_1, X_2, \dots, X_n “variáveis da equação”. A Tabela dos coeficientes fornece a equação de explicação a partir da coluna “coeficientes”, nessa coluna o termo constante e os coeficientes para X_1, X_2, \dots, X_n sendo a equação escrita como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon . \quad (3.3)$$

3.6.2 Avaliação da importância relativa das variáveis independentes com o coeficiente beta

Com a equação definida pode-se verificar o nível de importância relativa das variáveis individuais na explicação geral do problema. Nessa situação todas as variáveis devem ser

expressa em uma escala única, e assim comparações diretas podem ser feitas. Mas na maioria dos casos, os coeficientes β são empregados para que comparações diretas possam ser feitas. Na Tabela de coeficientes na coluna “coeficientes padronizados de regressão”. Assim sendo é possível efetuar comparações diretas entre as variáveis para se certificar de sua importância relativa na variável estatística de regressão.

3.6.3 Avaliação dos efeitos da multicolinearidade

Em todas as análises de variáveis estatísticas de regressão, deve ser observado o impacto da multicolinearidade, ou seja, variáveis altamente colineares podem afetar negativamente o resultado. Para resolver esse problema duas medidas estão disponíveis para testar o impacto da colinearidade: Cálculo dos valores da tolerância e do VIF.

O valor da tolerância é a quantidade da variância da variável explicada pelas outras variáveis, ou seja, uma alta tolerância indica pouca colinearidade. O fator de inflação VIF é o contrario da tolerância, logo são procurados valores VIF pequenos como indicativo de baixa inter-correlação entre as variáveis.

Nessa abordagem a multicolinearidade pode afetar substancialmente as variáveis incluídas por último na equação, assim sendo após avaliar os coeficientes estimados, se deve também avaliar o impacto potencial de variáveis omitidas a fim de garantir a significância gerencial da pesquisa.

Multicolinearidade – é a correlação entre as variáveis independentes. Esse é um problema de dados, e não de especificação do modelo. A situação ideal era ter diversas variáveis independentes altamente correlacionadas com a variável dependente, porém com pouca correlação entre elas próprias. A tarefa fundamental é: avaliar o grau de multicolinearidade e determinar seu impacto sobre os resultados e as ações corretivas se for o caso.

Os efeitos da multicolinearidade podem ser classificados em termos de explicação e estimação. Os efeitos da explicação se referem principalmente a habilidade do procedimento de regressão e do pesquisador em representar e compreender os efeitos de cada variável independente na variável estatística. Quando ocorre multicolinearidade, o processo de separação de efeitos individuais se torna mais difícil. Primeiro, ela limita o tamanho do

coeficiente de determinação e fica cada vez mais difícil acrescentar uma única previsão explanatória a partir de variáveis adicionais. Segundo, ela torna a determinação da contribuição de cada variável independente mais complexa porque os efeitos das variáveis independentes estão confusos, ou seja, a quantia de variância única para as variáveis independentes é reduzida a níveis que torna a estimação de seus efeitos individuais muito problemáticos.

Além dos efeitos na explicação, se podem ter grandes efeitos na estimativa dos coeficientes e em seus testes de significância estatística. Primeiro analisa-se o caso extremo de multicolinearidade em que duas ou mais variáveis são perfeitamente correlacionadas, chamado de singularidade, esse efeito impede a estimação de quaisquer coeficientes. Nesse caso a singularidade deve ser removida antes que a estimação dos coeficientes possa prosseguir.

Identificação da Multicolinearidade – A maneira mais simples e óbvia de se identificar a colinearidade é um exame da matriz de correlação para as variáveis independentes. A presença de altas correlações é a primeira indicação de colinearidade substancial. Duas das maneiras mais comuns para avaliar a colinearidade são: o valor de tolerância e seu inverso o fator de inflação de variância (VIF). Essas medidas dizem-nos o grau em que cada variável independente é explicada pelas demais variáveis independentes. Em termos simples, cada variável independente se torna uma dependente e passa por uma regressão em relação às demais variáveis independentes. Tolerância é a quantia de variabilidade da variável independente selecionada não explicada pelas outras variáveis independentes. Logo valores pequenos de tolerância denotam colinearidade elevada. Segundo HAIR et al., (2005) um valor de referência comum é uma tolerância de 0,10, o que corresponde a um valor de VIF acima de 10.

3.7 Validação dos Resultados (ESTÁGIO 6)

Após identificar o melhor modelo, o passo final é garantir que ele represente a população geral (generalização) sendo com isso apropriado as situações nas quais é usado (transferibilidade). A melhor orientação é a extensão em que o modelo de regressão se ajusta a um modelo teórico existente ou um conjunto de resultados previamente validados sobre o mesmo tópico. Em muitos casos, contudo, não há resultados anteriores ou teorias disponíveis. Assim, também serão feitas abordagens empíricas para a validação de um modelo.

A) Amostras adicionais ou particionadas.

A mais apropriada abordagem empírica de validação é testar o modelo de regressão em uma nova amostra tirada da população geral. Uma nova amostra garantirá representatividade e pode ser usada de diversas maneiras. Primeiro, o modelo original pode prever valores na nova amostra e o ajuste preditivo pode ser calculado. Segundo, um modelo separado pode ser estimado com uma nova amostra e então comparado com a equação original em relação à característica como as variáveis significantes incluídas; sinal, tamanho, e importância relativa de variáveis; e precisão preditiva. Em ambos os casos se determina a validade do modelo original, comparando-o com modelos de regressão estimados com a nova amostra.

Muitas vezes, a habilidade de coletar novos dados é limitada ou impraticável por fatores como custo, pressões de tempo ou disponibilidade de respondentes. Quando esse é o caso, deve ser dividida a amostra em duas partes: uma amostra de estimação para criar o modelo de regressão e a amostra de reserva ou validação, usada para “testar” a equação. Muitos procedimentos, tanto aleatório quanto sistêmicos, estão disponíveis para dividir os dados, de modo que cada um tire duas amostras independentes do conjunto único de dados.

B) Cálculo da estatística PRESS

Uma abordagem alternativa para obter amostras adicionais para fins de validação é em utilizar a amostra original de uma maneira especializada, calculando a estatística PRESS, uma medida semelhante à R^2 , usada para avaliar a precisão preditiva do modelo de regressão estimado. Difere das abordagens anteriores no sentido de que não um, mas $n-1$ modelos de regressão são estimados. O procedimento, semelhante às técnicas *bootsrapping*, omite uma observação na estimação do modelo de regressão e então prevê a observação omitida com o modelo estimado. Logo a observação não pode afetar os coeficientes do modelo usado para calcular seu valor previsto. O procedimento é aplicado novamente, omitindo uma outra observação, estimando um novo modelo e fazendo a previsão. Os resíduos para as observações podem então ser somados para fornecer uma medida geral de ajuste preditivo.

C) Comparação de modelo de regressão

Quando se compara modelos de regressão, o critério mais comum empregado é o ajuste preditivo geral. Foi discutido anteriormente que R^2 nos fornece essa informação, mas apresenta uma desvantagem: à medida que mais variáveis são acrescentadas, R^2 sempre aumenta. Assim, após termos incluído todas as variáveis independentes, não é encontrado um R^2 maior, mas pode-se descobrir que um número menor de variáveis independentes resulta em um valor quase idêntico. Portanto para comparar modelos com diferentes números de variáveis independentes, usa-se o R^2 ajustado, o qual é igualmente útil na comparação de modelos com diferentes conjuntos de dados, uma vez que se faz uma compensação para os diferentes tamanhos de amostras.

Análise dos Resultados

4.1 Introdução

Como já discutido anteriormente, a taxa de abandono é um problema relevante com expressivo valor financeiro que afeta de forma direta as empresas, principalmente as de telecomunicações. Fica evidenciado que a natureza da questão e os impactos nocivos da perda de clientes têm que serem pesquisados, sendo a forma escolhida de se enfrentar a questão o entendimento das causas da taxa de abandono provocadas pela qualidade do serviço. Essa pesquisa teve a finalidade de identificar eventuais correlações dos indicadores de qualidade com a taxa de abandono, tendo sido necessário se compreender os indicadores de qualidades para que, por meio deles, fosse possível identificar as causas para prevenir a taxa de abandono.

Para avaliar o desempenho dos indicadores de qualidade utilizou-se uma base de dados real de uma das quatro maiores empresas de telefonia do Brasil. O nome da empresa não foi considerado relevante e é mantido sob sigilo. A base de dados correspondeu aos dados de 12 meses coletados entre os meses de janeiro a dezembro de 2006.

Para os testes e simulações realizados, foi utilizada a ferramenta *SPSS* sendo a análise de regressão múltipla a técnica escolhida por ser a técnica de dependência mais amplamente usada e versátil, aplicável em cada faceta da tomada de decisões em negócios. Nessa pesquisa descreve-se como a análise de regressão múltipla é usada para resolver problemas de pesquisa importantes, particularmente na área de negócios em telecomunicações. A variável dependente foi à taxa de abandono. As variáveis independentes foram: tempo médio de reparo, tempo médio de instalação, taxa de reparo, instalação no prazo grandes clientes, reparo no prazo alto valor, reparo no prazo demais clientes, reparo no prazo total, instalação em garantia, reparo

preventivo, reparo recorrente alto valor, reparo recorrente demais clientes e o reparo recorrente total.

4.2 Estimação do Modelo de Regressão, Avaliação do Ajuste Geral e da Significância Estatística do Modelo.

Na estimação deve-se dispor de alternativas para atender características específicas dos dados ou maximizar o ajuste dos mesmos. Depois que o modelo é estimado, o seu ajuste geral pode ser avaliado para estabelecer se atinge níveis aceitáveis sobre os critérios estatísticos, se identifica às relações propostas e se tem significância prática.

Na maioria dos casos de regressão múltipla, deve ser feita escolha entre várias possíveis variáveis independentes que podem ser selecionadas para inclusão na equação de regressão. Conforme foi visto no capítulo 3, pode-se escolher entre especificação confirmatória, combinatória e busca seqüencial. Abaixo serão analisadas as variáveis nos dois métodos (combinatória e busca seqüencial), onde é escolhido o que melhor se adapta aos objetivos dessa pesquisa.

4.2.1) Método Combinatório

A abordagem mais simples é a especificação combinatória. É um processo de busca generalizada em todas as possíveis combinações das variáveis independentes. Todas as possíveis combinações das variáveis independentes são examinadas e o conjunto de variáveis que melhor se ajusta é identificado. A primeira simulação é com a inclusão de todas as variáveis disponibilizadas tendo o universo das amostras do ano de 2006.

1) Coeficiente de Determinação

O objetivo é saber se as variáveis explicativas usadas no modelo (a componente determinística) explicam a variabilidade total da variável dependente, ou seja, verificar se é coerente à capacidade explicativa do modelo. Se a capacidade explicativa do modelo é elevada então o resíduo do modelo (efeitos não captados pela componente determinística) é reduzido.

Geralmente, nos modelos de regressão linear a estatística usada para quantificar a capacidade explicativa do modelo é o R^2 , designado por coeficiente de determinação. Essa estatística constitui a proporção da variação explicada face à variação total da variável dependente. Como regra geral o modelo selecionado é o com R^2 mais elevado.

No modelo de regressão linear, a proporção da variação explicada é dada pelo coeficiente de determinação que varia de 0 a 1, $0 \leq R^2 \leq 1$. O seu valor é zero quando SQE (Soma dos Quadrados Explicada) também é zero e $SQR = SQT$ (Soma dos Quadrados Totais). O valor de R^2 é 1 quando $SQE = SQT$ e $SQR = 0$ (Soma dos Quadrados dos Resíduos)

A partir da Tabela 4.1 pode-se observar que $R^2 = 0,54$ significando que 54,0 % da variabilidade total dos dados são explicados pelo modelo de regressão, ou seja, as doze variáveis independentes são responsáveis por explicarem esse percentual da variável dependente (taxa de abandono). Em outras palavras, essas variáveis independentes utilizadas no modelo de regressão linear ajudam a explicar em torno de 54% da variação da taxa de abandono. Entretanto, ficam por explicar 46% das variações da taxa de abandono, que se deve a outros fatores não considerados.

Tabela 4.1 Sumário do modelo Combinatório.

Sumário do Modelo Combinatório

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.725 ^a	.537	.523	.58808

a. Predictors: (Constant), TEMPO MEDIO DE INSTALACAO, TX_INST_GARANTIA, TX_Reparo_DEMAIS_CLIENTES, REPAROS PREVENTIVAS, TAXA DE REPARO, TX_Reparo_ALTO_VALOR, TX_Reparo_Varejo, TX_Reincidente_ALTO_VALOR, TX_INST_Grandes_Clientes, TX_Reincidente_Varejo, TEMPO MEDIO DE REPARO

Fonte: SPSS.

Os valores de R^2 e R^2 ajustado apontam para um bom poder de explicação da equação de regressão, os quais ficaram próximos dos valores obtidos por meio do método *Stepwise*, que é oportunamente elucidado.

2) Equação Estimada

O principal objetivo dos modelos de regressão é a estimação dos parâmetros desconhecidos β . A esse processo é comum chamar-se de “ajuste do modelo aos dados”. A Tabela 4.2 contém as estimativas dos parâmetros e respectivas estimativas do erro padrão, as estimativas dos coeficientes padronizados e o valor da estatística *t-student*.

A partir da Tabela 4.2 pode ser caracterizada a equação de regressão que pode ser vista a seguir:

Tabela 4.2 Coeficiente do modelo Combinatório.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.000	.000										
	RRAV	.144	.388	.144	.886	.714	-.099	.068	.579	-.029	-.010	.453	2.206
	IGT	.171	.079	.171	.159	.032	.015	.327	.882	.171	.061	.129	7.770
	RRT	-.197	.049	-.197	-.534	.127	-.173	.022	-.360	-.122	-.043	.332	3.015
	RPT	.087	.057	.087	-.513	.132	-.201	.027	-.039	-.121	-.043	.243	4.117
	RRDC	.672	.079	.671	.544	.146	.516	.827	.917	.566	.242	.130	7.687
	IPGC	.139	.046	.139	.021	.003	.048	.229	-.293	.236	.086	.381	2.627
	RPAV	-.042	.044	-.042	-.946	.346	-.129	.046	-.220	-.076	-.027	.409	2.447
	RPDC	.390	.041	.390	.043	.965	-.080	.084	.256	.003	.001	.467	2.141
	TMR	.608	.146	.608	3.322	.000	.396	.808	.499	.579	.310	.849	1.160
	TRP	.414	.054	.414	7.466	.000	.304	.505	.575	.507	.438	.977	1.061
	QRPR	-.310	.149	-.310	5.631	.001	-.566	-.128	-.145	-.206	-.179	.266	3.144
	TMI	.525	.065	.525	-3.120	.000	.401	.644	.116	.590	.472	.291	3.723

a. Dependent Variable: TAXA_DE_ABANDONO

Fonte: SPSS.

$$\begin{aligned}
 \text{Taxa de Abandono} = & 0,144(RRAV) + 0,171(IGT) - 0,197(RRT) + \\
 & 0,087(RPT) + 0,672(RRDC) + 0,139(IPGC) - 0,042(RPAV) + 0,39(RPDC) + \\
 & 0,608(TMR) + 0,414(TRP) - 0,310(QRPR) + 0,525(TMI).
 \end{aligned}
 \tag{4.1}$$

A equação estimada pode ser vista como uma tentativa para explicar as variações na variável dependente que resultaram das alterações nas variáveis independentes. Mesmo com bons valores dos coeficientes ajustados de determinação, na estrutura do modelo, permanecem ainda todas as 12 variáveis, o que causa um grande inconveniente gerencial para implementá-lo.

3) Teste de Hipótese para a Significância do Modelo

Esse teste é usado para estabelecer se existe ou não alguma relação entre a variável dependente e o conjunto de variáveis independentes. Consistindo em testar as seguintes hipóteses:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 ;$$

$$H_1 : \beta_k \neq 0 \text{ para algum } k .$$

A estatística do teste tem, sob H_0 (a hipótese H_0 significa que a regressão não é significativa), a distribuição F com p e $(n-p)$ graus de liberdade. A rejeição da hipótese H_0 indica a existência de regressão. As quantidades necessárias para calcular o valor observado dessa estatística estão dispostas na Tabela 4.3, denominada de Tabela de análise de variância - ANOVA.

Como $F_{obs} = 12.572$ é maior que $F_c = F_{(95\% ; 11; 156)} = 1,80$ (valor percentual de uma distribuição F com 11 graus de liberdade no numerador e 156 no denominador) deve ser rejeitada H_0 em favor da hipótese alternativa de que a regressão é estatisticamente significativa. Em outras palavras, pode-se dizer que ao nível de significância de 0,05, rejeita-se a hipótese da regressão não ser significativa. Então há evidência para afirmar que existe um relacionamento linear entre o conjunto de variáveis explicativas e a taxa de abandono.

Tabela 4.3 ANOVA do modelo Combinatório.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47.825	11	4.348	12.572	.000 ^a
	Residual	53.950	156	.346		
	Total	101.776	167			

a. Predictors: (Constant), TEMPO MEDIO DE INSTALACAO, TX_INST_GARANTIA, TX_Reparo_DEMAIS_CLIENTES, REPAROS PREVENTIVAS, TAXA DE REPARO, TX_Reparo_ALTO_VALOR, TX_Reparo_Varejo, TX_Reincidente_ALTO_VALOR, TX_INST_Grandes_Clientes, TX_Reincidente_Varejo, TEMPO MEDIO DE REPARO

b. Dependent Variable: TAXA DE ABANDONO

Fonte: SPSS.

4) Teste de Hipótese para o Parâmetro β_k

Após a verificação de que pelo menos um dos parâmetros β_k é significativo, deve-se testar a significância de cada um deles, isto é, para cada parâmetro β_k ($k=1, \dots, p$), testam-se as hipóteses:

$$H_0 : \beta_k = 0;$$

$$H_1 : \beta_k \neq 0.$$

O teste para cada parâmetro é feito utilizando a estatística *t-student* com $(n-p)$ graus de liberdade, o desvio padrão amostral $S(b_k)$ e o estimador b_k . A estatística do teste é dada por

$$t^* = \frac{b_k}{S(b_k)}. \quad (4.2)$$

Se $|t^*| \leq t(1-\alpha/2; n-p)$, o teste não rejeita H_0 ; caso contrário o teste rejeita H_0 em favor de H_1 . A rejeição de H_0 indica uma contribuição significativa da variável independente X_k no modelo.

A partir dos resultados apresentados na Tabela 4.2 se obtém os dados da estatística t para os 12 parâmetros a serem testados. Para verificar que a variável TMR não é muito útil, isto é, que sua influência sobre a variável dependente não é significativa testa-se a hipótese nula do coeficiente para essa variável cujo valor crítico é $t_c \cong 2,258$ (para os graus de liberdade são 11 e 156).

Como $|t_c| \cong 2,258 < |t| = 3,322$ concluem-se que, ao nível de significância de 0,05, há evidência para dizer que a variável TMR é útil na explicação da taxa de abandono, ou seja, a variável TMR contribui para a explicação da taxa de abandono. Já para a variável RRAV ($|t_c| \cong 2,258 > |t| = 0,886$, sendo feito uma análise similar, rejeita-se a hipótese H_1 em favor da hipótese nula, ou seja, ela não é significativa na explicação da taxa de abandono.

Por meio dos dados observados a partir da referida Tabela e fazendo-se comparação com o valor Tabelado cujo valor crítico de $t_c \cong 2,258$ pode ser concluído que apenas as

estatísticas t correspondentes aos parâmetros β_9 , β_{10} , β_{11} e β_{12} podem ser validadas, ou seja, as variáveis TMR, TRP, QRPR e TMI são estatisticamente significativas.

4.2.2) Método Seqüencial *Stepwise*

Nesse método os regressores incluídos ou excluídos durante a elaboração do modelo podem ser reexaminados. Permite examinar a contribuição de cada variável independente para o modelo de regressão, sendo cada variável considerada para inclusão antes do desenvolvimento da equação. A principal vantagem em relação aos outros métodos seqüenciais é a sua habilidade em acrescentar ou eliminar variáveis em cada estágio.

1) Coeficiente de Determinação

A seguir, são mostrados os dados obtidos por meio do método *Stepwise*. A partir de seus respectivos valores, que se encontram na Tabela 4.4, foi considerado adequado em desconsiderar o modelo de especificação confirmatória e a utilizar, nessa pesquisa, o modelo de busca seqüencial *Stepwise*. Para esse trabalho o método seqüencial mostra-se mais adequado devido ser objetivo para selecionar as variáveis que maximizam o poder de explicação com o menor número de variáveis empregadas.

Na Tabela 4.4 são mostradas os valores obtidos para as observações de 2006 pelo modelo seqüencial *Stepwise*, onde se verifica que se deixa de trabalhar com 12 variáveis independentes para trabalhar apenas com quatro e, apesar dessa redução substancial, os valores de R^2 e R^2 ajustados ficam muito próximos dos valores do modelo anterior (método Combinatório).

Tabela 4.4 Sumário do modelo *STEPWISE*.

Sumário do Modelo Seqüencial (*Stepwise*)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.522 ^a	.272	.268	.85553	.272	62.179	1	166	.000	1.812
2	.637 ^b	.405	.398	.77577	.133	36.889	1	165	.000	
3	.689 ^c	.475	.466	.73106	.070	21.802	1	164	.000	
4	.718 ^d	.516	.504	.70453	.040	13.581	1	163	.000	

a. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO

b. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPARO

c. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPARO, TEMPO MEDIO DE REPARO

d. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPAROS DA PLANTA, TEMPO MEDIO DE REPARO, REPAROS PREVENTIVOS

e. Dependent Variable: TAXA DE ABANDONO

Fonte: SPSS.

A partir da Tabela 4.4 pode-se observar que $R^2 = 0,52$ significando que 50,4 % da variabilidade total dos dados são explicados pelo modelo de regressão, ou seja, as quatro variáveis independentes são responsáveis por explicarem esse percentual da variável dependente (taxa de abandono). Em outras palavras, essas variáveis independentes utilizadas no modelo de regressão linear ajudam a explicar em torno de 50% da variação da taxa de abandono. Entretanto, ficam por explicar 50% das variações da taxa de abandono, que se deve a outros fatores não considerados. Os valores de R^2 e R^2 ajustado apontam para um bom poder de explicação da equação de regressão.

2) Equação Estimada

O principal objetivo dos modelos de regressão é a estimação dos parâmetros desconhecidos β . A esse processo é comum chamar-se de “ajuste do modelo aos dados”. A Tabela 4.5 contém as estimativas dos parâmetros e respectivas estimativas do erro padrão, as estimativas dos coeficientes padronizados e o valor da estatística *t-student*.

A partir da Tabela 4.5 pode ser caracterizada a equação de regressão que pode ser vista abaixo:

Tabela 4.5 Coeficientes do modelo *Stepwise*.

Model	Coefficients ^a													
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics			
	B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF		
1	(Constant)	.000	.066		-.005	.996	-.131	.130						
	TEMPO MEDIO DE INSTALACAO	.522	.066	.522	7.885	.000	.391	.653	.522	.522	.522	1.000	1.000	
2	(Constant)	.000	.060		-.006	.995	-.119	.118						
	TEMPO MEDIO DE INSTALACAO	.484	.060	.484	8.018	.000	.365	.603	.522	.530	.481	.989	1.011	
	TAXA DE REPARO	.367	.060	.367	6.074	.000	.247	.486	.417	.427	.365	.989	1.011	
3	(Constant)	.000	.056		-.008	.994	-.112	.111						
	TEMPO MEDIO DE INSTALACAO	.568	.060	.568	9.519	.000	.450	.686	.522	.597	.539	.900	1.112	
	TAXA DE REPARO	.394	.057	.394	6.893	.000	.281	.507	.417	.474	.390	.979	1.022	
	TEMPO MEDIO DE REPARO	.279	.060	.279	4.669	.000	.161	.398	.052	.343	.264	.894	1.118	
4	(Constant)	.000	.054		-.007	.994	-.108	.107						
	TEMPO MEDIO DE INSTALACAO	.529	.058	.529	9.051	.000	.414	.644	.522	.578	.493	.870	1.149	
	TAXA DE REPARO	.413	.055	.413	7.486	.000	.304	.523	.417	.505	.407	.970	1.031	
	TEMPO MEDIO DE REPARO	.601	.105	.601	5.745	.000	.395	.808	.052	.410	.313	.271	3.686	
	REPARO PREVENTIVO	-.390	.106	-.390	-3.685	.000	-.599	-.181	-.092	-.277	-.201	.286	3.783	

a. Dependent Variable: TAXA DE ABANDONO

Fonte: SPSS.

$$Taxa\ de\ Abandono = 0,53(TM1) + 0,41(TRP) + 0,60(TMR) - 0,39(QRPR). \quad (4.3)$$

A equação de regressão estimada pode ser vista como uma tentativa para explicar as variações na variável dependente que resultaram das alterações nas variáveis independentes.

3) Teste de Hipótese para a Significância do Modelo

Esse teste é usado para estabelecer se existe ou não alguma relação entre a variável dependente e o conjunto de variáveis independentes. Consistindo em testar as seguintes hipóteses:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0;$$

$$H_1: \beta_k \neq 0 \text{ para algum } k.$$

A estatística do teste tem, sob H_0 (a hipótese H_0 significa que a regressão não é significativa), a distribuição F com p e $(n-p)$ graus de liberdade. A rejeição da hipótese H_0 indica a significância de regressão. As quantidades necessárias para calcular o valor observado dessa estatística estão dispostas na Tabela 4.6, denominada de Tabela de análise de variância - ANOVA.

Como $F_{obs} = 43.367$ é claramente superior a $F_c = F_{(95\% ; 4 ; 163)} = 2,40$ (valor percentual de uma distribuição F com 4 graus de liberdade no numerador e 163 no denominador) deve ser

rejeitada H_0 em favor da hipótese alternativa de que a regressão é estatisticamente significativa. Em outras palavras, pode-se dizer que ao nível de significância de 0,05, rejeita-se a hipótese da regressão não ser significativa. Então há evidência para afirmar que existe um relacionamento linear entre o conjunto de variáveis explicativas e a taxa de abandono.

Tabela 4.6 ANOVA do modelo *Stepwise*

ANOVA ^e						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	45.511	1	45.511	62.179	.000 ^a
	Residual	121.502	166	.732		
	Total	167.013	167			
2	Regression	67.712	2	33.856	56.256	.000 ^b
	Residual	99.301	165	.602		
	Total	167.013	167			
3	Regression	79.364	3	26.455	49.499	.000 ^c
	Residual	87.649	164	.534		
	Total	167.013	167			
4	Regression	86.105	4	21.526	43.367	.000 ^d
	Residual	80.908	163	.496		
	Total	167.013	167			

a. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO

b. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPARO

c. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPARO, TEMPO MEDIO DE REPARO

d. Predictors: (Constant), TEMPO MEDIO DE INSTALAÇÃO, TAXA DE REPAROS, TEMPO MEDIO DE REPARO, REPARO PREVENTIVO

e. Dependent Variable: TAXA DE ABANDONO

Fonte: SPSS.

4) Teste de Hipótese para o Parâmetro β_k

Após a verificação de que pelo menos um dos parâmetros β_k é significativo, deve-se testar a significância de cada um deles, isto é, para cada parâmetro β_k ($k=1, \dots, p$), testam-se as hipóteses:

$$H_0: \beta_k = 0;$$

$$H_1: \beta_k \neq 0.$$

O teste para cada parâmetro é feito utilizando a estatística *t-student* com $(n-p)$ graus de liberdade, o desvio padrão amostral $S(b_k)$ e o estimador b_k . A estatística do teste é dada pela equação (4.2), como foi visto anteriormente.

Se $|t^*| \leq t(1-\alpha/2; n-p)$, o teste não rejeita H_0 ; caso contrário o teste rejeita H_0 em favor de H_1 . A rejeição de H_0 indica uma contribuição significativa da variável independente X_k no modelo.

Por meio dos resultados apresentados na Tabela 4.5 se obtém os dados da estatística t para os 4 parâmetros a serem testados. Para verificar que a variável TMR não é muito útil, isto é, que sua influência sobre a variável dependente não é significativa testa-se a hipótese nula do coeficiente para essa variável cujo valor crítico é $t_c \cong 2,258$ (para os graus de liberdade são 4 e 163).

Como $|t_c| \cong 2,258 < |t| = 9,051$ concluem-se que, ao nível de significância de 0,05, há evidência para dizer que a variável TRM é útil na explicação da taxa de abandono, ou seja, a variável TMR contribui para a explicação da taxa de abandono

Com os dados observados a partir da referida Tabela e fazendo-se comparação, similares ao que foi feito anteriormente, com o valor Tabelado cujo valor crítico de $t_c \cong 2,258$ conclui-se que todas as estatísticas t correspondentes aos parâmetros β_1 , β_2 , β_3 e β_4 podem ser validadas, ou seja, as variáveis TMI, TMP, TMR e QRPR são estatisticamente significativas.

Para a análise a seguir se trabalha com uma variável dependente e quatro variáveis independentes sendo que o método utilizado para cálculo foi o *Stepwise* devido a sua habilidade em acrescentar ou eliminar variáveis em cada estágio, até encontrar a melhor equação, conforme já discutido anteriormente.

Todas as correlações entre as variáveis são mostradas na Tabela 4.7. O exame da matriz indica que o TMI foi a obteve a maior correlação com a variável dependente *taxa de abandono* (0,52), sendo também verificado que a correlação entre o *taxa de abandono* e a variável TRP também foi alta (0,42). Existem outras variáveis correlacionadas, mas que foram eliminadas devido à utilização dessas variáveis que juntas na mesma equação não serem apropriadas. Essas variáveis juntas causam uma alta colinearidade, ficando-se então com as quatro variáveis.

Tabela 4.7 Correlações.

Correlação das Variáveis

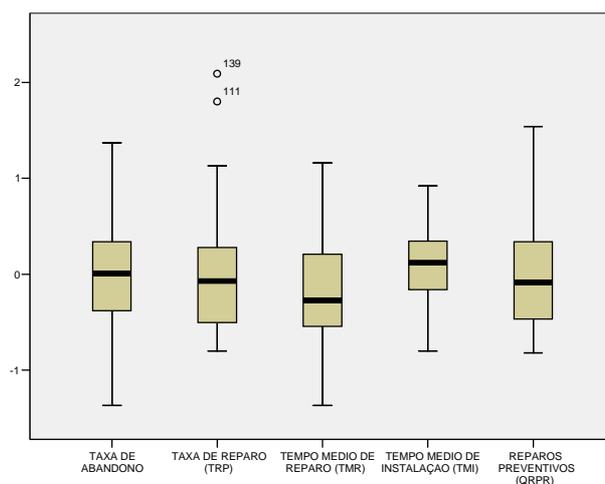
		TAXA DE ABANDONO	TEMPO MEDIO DE REPARO	TAXA DE REPAROS	REPARO PREVENTIVO	TEMPO MEDIO DE INSTALAÇÃO
Correlation	TAXA DE ABANDONO	1.000	.052	.417	-.092	.522
	TEMPO MEDIO DE REPARO	.052	1.000	-.130	.285	-.310
	TAXA DE REPAROS DA PLANTA	.417	-.130	1.000	-.068	.104
	REPARO PREVENTIVO	-.092	.185	-.068	1.000	-.351
	TEMPO MEDIO DE INSTALAÇÃO	.522	-.310	.104	-.351	1.000

Fonte: SPSS.

4.2.3 Identificação de Observações Influentes

No Gráfico 4.1 são identificados dois pontos atípicos sendo eles os pontos 111 e 139, no indicador de Taxa de Reparo. Esses pontos foram identificados na base de dados com sendo da filial Amazonas. Verificou-se por meio da análise realizada que ocorreu uma grande precipitação pluviométrica (chuva). Essas precipitações climáticas são adversas à prática da atividade de telecomunicações, causando paralisações no *backbone* da cidade de Manaus bem como provocando interferência no sistema de transmissão, qual é atendida por meio de satélite. A quantidade média de intervenções fica em torno de 15 acionamentos por mês. Só no mês de março 63 acionamentos e no mês de maio 61.

Gráfico 4.1. Identificação de pontos atípicos.



Fonte: SPSS.

No gráfico 3.2 (capítulo 3) se verificou dois pontos fora do histograma, onde foi identificado, pelo método de diagnóstico *casewise* que foi exposto na Tabela 4.7. Essas observações são as de número 111 e 139, as quais já tinham sido mostradas no gráfico 3.6 anteriormente.

Tabela 4.8 Diagnóstico *Case wise*.

Diagnóstico *Case Wise*

Case Number	Std. Residual	TAXA DE ABANDONO	Predicted Value	Residual
111	3.593	2.41	-.1213	2.53130
139	3.742	1.91	-.7265	2.63651

a. Dependent Variable: TAXA DE ABANDONO

Fonte: SPSS.

4.3 Interpretação das Variáveis Estatísticas e Validação dos Resultados

Com a estimação do modelo concluída, a variável estatística de regressão especificada e os testes diagnósticos confirmando a adequação dos resultados administrados, o próximo passo é examinar a equação preditiva encontrada.

Iniciou-se essa pesquisa com 12 variáveis independentes e uma dependente. Após introduzir os dados na ferramenta *SPSS*, o sistema calculou a equação pelo método *Stepwise*, que foi previamente selecionado, tendo sido verificado, nessa primeira análise valores de $R^2 = 52,0\%$ e R^2 ajustado = $50,0\%$, o que caracteriza valores bastante animadores para a taxa de abandono que foi encontrada na equação (4.3), nesse caso analisa-se individualmente os coeficientes desta equação.

Em que:

TMI – Representa o tempo médio gasto na instalação do circuito, ou seja, é o tempo desde abertura da ordem de serviço até o seu efetivo fechamento.

TRP – Representa a taxa de reparo executado no prazo acordado com o cliente esse prazo varia de 8 horas para clientes Alto Valor e 24 horas para os Demais Clientes geralmente são os clientes varejo.

QRPR – Esse valor representa a quantidade de reparos preventivos efetuados nas empresas.

TMR – Representa o tempo médio gasto no reparo do circuito, ou seja, é o tempo desde abertura do Bilhete de Defeito até o seu efetivo fechamento.

A - Interpretação dos Coeficientes do Modelo *Stepwise*

- O valor do coeficiente para a variável Tempo Médio de Instalação TMI é de 0,53, o que significa que, isolando-se as outras variáveis da equação, para um aumento de 10% normatizado na variável TMI ocorre um aumento normatizado de 5% na taxa de abandono, que em valores não normatizados se refere a um aumento em média de 2500 clientes na taxa de abandono e um aumento em média de 10 horas no tempo de instalação;
- O valor do coeficiente para a variável Tempo Médio de Reparo TMR é de 0,60, o que significa que, isolando-se as outras variáveis da equação, para um aumento de 10% normatizado na variável TMR ocorre um aumento normatizado de 6% na taxa de abandono, que em valores não normatizados se refere a um aumento em média de 2800 clientes na taxa de abandono e um aumento em média de 1 hora no tempo de reparo;
- O valor do coeficiente para a variável TRP é de 0,41, o que significa que, isolando-se as outras variáveis da equação, para um aumento de 10% normatizado na variável TRP ocorre um aumento normatizado de 4,1% na taxa de abandono, que em valores não normatizados se refere a um aumento em média de 1700 clientes na taxa de abandono;
- O valor do coeficiente para a variável QRPR é de 0,39 que significa, isolando-se as outras variáveis da equação, para um aumento de 10% normatizado na variável QRPR ocorre uma diminuição normatizada de 3,9% na taxa de abandono, que em valores não normatizados se refere a uma diminuição em média de 1600 clientes na taxa de abandono.

Nota-se que no modelo escolhido, as variáveis têm pesos bem distribuídos em relação à variável estatística, não tendo nenhuma um peso consideravelmente muito maior do que um outro. Verifica-se também que quanto mais sejam realizados reparos preventivos, mais são reduzidas as saídas de clientes da empresa. Verifica-se, por outro lado, que um tempo médio de instalação elevado causa uma grande insatisfação nos clientes refletindo em sua provável saída. Isto também é percebido com relação à quantidade de reparo o cliente banda larga que não concorda quando fica fora do acesso a Internet e também afeta na sua saída da empresa.

Nesse trabalho os valores de tolerância estão entre 0,57 e 0,97 e os valores do VIF (*Variance Inflation Factor*) entre 1,76 e 1,15 o que caracterizam níveis baixos de colinearidade levando a interpretação dos coeficientes da variável estatística como não sendo afetada desfavoravelmente pela multicolinearidade. A Tabela 4.5 mostrou todos os valores de tolerância e VIF para o modelo selecionado (*Stepwise*).

4.4 Conclusões

Nesse contexto, o modelo estimado chega essencialmente ao resultado em que são observadas duas influências fortes (o tempo médio de reparo e o tempo médio de instalação) e outras duas influências um pouco menores (taxa de reparo na planta e quantidade de reparo preventivo) sendo esse último inversamente proporcional ao valor da Taxa de Abandono. Isso, gerencialmente falando, é bastante favorável, pois significa que quanto mais for realizada a manutenção preventiva, mais se eleva a satisfação do cliente e, conseqüentemente, haverá diminuição da taxa de abandono.

Uma elevação em uma das três variáveis normatizadas (tempo médio de reparo, tempo médio de instalação e taxa de reparo na planta) resulta em um aumento na taxa de abandono (Churn), ou seja, uma maior saída de clientes da empresa, enquanto um aumento na quantidade de reparos preventivos resulta em uma diminuição na taxa de abandono.

5.1 Conclusões

O objetivo desse trabalho foi tentar explicar o comportamento do cliente no setor de telecomunicações, usuário de Internet de banda larga, com o uso de uma equação que representa a taxa de abandono em função dos indicadores de qualidade de serviço. Para isto foi necessário caracterizar a taxa de abandono, bem como suas causas. Foi identificado que há dois tipos de causas associadas à taxa de abandono: causas voluntárias e causas involuntárias. Essa pesquisa procurou se deter apenas nas causas voluntárias, ou desligamento por motivo deliberado ligado à qualidade. Podem-se destacar como principais contribuições desse trabalho:

- 1) Consolidação, em forma de relação, dos indicadores de qualidade de serviços de banda larga, disponibilizados pelas empresas de telecomunicações à taxa de abandono da planta;

Um passo inicial nesse sentido foi realizar um estudo sobre o processo de descoberta de conhecimento em banco de dados para, a partir do estudo, analisar os dados da pesquisa, normaliza-los e selecionar o método para especificação do modelo da equação final. Para o equacionamento houve preocupação com a avaliação da significância estatística e a interpretação das variáveis. O resultado foi uma equação em que foi identificado que alguns indicadores de qualidade devem ter, por partes dos gestores de manutenção das empresas de telecomunicações, uma maior atenção, pois são esses indicadores que efetivamente influenciam a taxa de abandono.

Após o entendimento da aplicabilidade da equação estatística, foi realizada uma análise gerencial dos dados encontrados, em que se verificou a existência de uma correlação entre os indicadores de qualidade de serviço e a taxa de abandono, ou seja, 50,4 % da taxa de abandono no ano de 2006 esta relacionada a causas voluntárias por motivo deliberado ligado à qualidade. Essas causas estão ligadas a quatro indicadores de qualidade: Tempo Médio de Reparo, Tempo Médio de Instalação, Preventivas e Taxa de Reparo, conforme caracterizado no campo da dissertação.

2) Análise do conjunto de variáveis que exercem maior relevância para o problema abordado.

Por meio da análise das 12 variáveis identificou-se que Tempo Médio de Reparo, Tempo Médio de Instalação, Preventivas e Taxa de Reparo possuem uma maior relevância sobre a taxa de abandono. Consequentemente existem indicativos que devem ser focados para diminuir a taxa de abandono:

- Diminuição da Taxa de Reparo, ou seja, o cliente para banda larga é muito exigente com o serviço prestado, não aceitando paralisações. Esse fato pode ser explicado quando se verifica que esse cliente tem como característica ficar conectado à rede o máximo de tempo possível. Para isto tem-se que agir na redundância de sistemas, procurando ao máximo evitar as causas comuns de reparos como quedas de *backbone*, falta de conectividade, etc.;
- Outro fator indicado é que depois de ocorrida uma paralisação, o tempo de restabelecimento do serviço deve ser o menor possível. Esse fator foi medido pelo tempo médio de reparo, e foi demonstrado que quanto maior for o tempo para remover a falha maior é a taxa de abandono. Nesse caso devem ser procuradas ferramentas de recuperação remota de falhas no intuito de diminuir o Tempo de Reparo;
- Em relação aos Reparos Preventivos, a pesquisa sinalizou que é uma boa prática de retenção, pois na antecipação das possíveis falhas conta positivamente. Quanto maior a quantidade de reparo preventivo realizado menor é a taxa de abandono, podendo ser explicado pela satisfação do cliente.

Sendo assim, surge a necessidade de se implementar sistemas de gerência e análise da degradação do sinal, para que, a partir dessas análises, sejam gerados reparos preditivos, com a finalidade de antever as falhas;

- Por fim, o Tempo Médio de Instalação, que afeta a percepção de qualidade do cliente antes mesmo que ele seja cliente. A análise mostra que quanto menor esse tempo, menor serão as taxas de abandono. Nesse caso pode-se deduzir que o cliente que não tem suas expectativas iniciais atendidas estará mais propenso a abandonar a empresa. Tem-se então que manter as expectativas ou superá-las sempre que possível antecipando os prazos de instalação. Caso ocorra algum problema nos processos de instalação ou que as suas expectativas tenham valores diferentes das possibilidades de atendimento das operadoras de serviço, fica como sugestão manter o cliente bem informado e que os atrasos e novos prazos sejam sempre negociados com o ele.

Conforme análise no Anexo III Tabela AIII.3, para uma operadora, que têm uma base de 1.2 milhões de clientes com uma taxa mensal de abandono de 2 %, estima-se uma perda financeira da ordem de 251 milhões de reais. Caso ações derivadas dessa dissertação sejam implementadas poder-se-ia reduzir a taxa de abandono de 2% para 1,59%. Essa redução está referendada nas taxas de melhorias, citadas na Tabela AIII.4, as quais são balizadas em parâmetros de qualidades internacionais, referendados pela empresa em análise. Em termos financeiros essa melhoria pode ser estimada em patamares de uma redução de perda de 45 milhões. Os resultados obtidos demonstram a aplicabilidade da equação, bem como as possibilidades de ganho para as prestadoras de serviço que decidirem utilizar a equação como parte de suas estratégias de manutenção.

Por fim, pode-se entender que a manutenção tem um papel fundamental como parcela de contribuição do valor total produzido, e no desempenho das atividades ligadas ao serviço. Isto se evidencia na participação de perdas de clientes, decorrentes do desempenho em termos da qualidade das operações de manutenção. Por outro lado, pode-se concluir que uma instalação ou reparo feito em menor tempo, uma preventiva bem feita ou a confiabilidade do serviço expressa por menores taxas de reparo, afetam a positivamente a satisfação do cliente fazendo com que ele permaneça na empresa. Para isto os gestores têm que empreender

investimentos em melhoria de qualidade na instalação, reparos e nas manutenções preventivas, que devem sempre ser vistos pela ótica da retenção de cliente e manutenção da base, sendo a taxa de abandono o modo mais prático de se evidenciar a eficiência do todo.

5.2 Recomendações para Estudo Futuro

Propostas para futuros desenvolvimentos sobre o tema:

- Testar a equação envolvendo a base de dados de outras operadoras, visando estabelecer um modelo de avaliação de desempenho da taxa de abandono de forma mais abrangente;
- Inserir na equação um indicador para avaliar os custos da atividade de manutenção, objetivando uma análise do seu desempenho nesse requisito. Pretende com isso, atuar de forma a reduzir os custos da má qualidade no serviço prestado;
- Realizar uma pesquisa de campo para verificar a percepção do cliente sobre a qualidade do serviço de banda larga. O intuito é que com essa percepção se promovam ações de melhoria e conseqüentemente a diminuição nas taxas de abandono;
- Levantar detalhadamente nas bases de dados informações da taxa de abandono, separando os motivos do abandono, para que se possa gerar uma equação com maiores informações entre as entradas e a saída dos clientes.

Referencias Bibliográficas

- ADRIAANS, Pieter; ZANTING, Dolf. “*Data Mining*”. Addison-Wesley, England, 1996.
- ATTADIA, L. C. D. L.; MARTINS, R. A. *Medição de Desempenho Como Base Para Evolução da Melhoria Contínua*. Revista Produção, Vol. 13, 1999.
- ALENCAR, Marcelo Sampaio. *Telefonia Digital*. São Paulo: Érica, 1998.
- ALMEIDA, Adiel Teixeira; WAJSMAN, Dymitr. *Engenharia de Manutenção: Visão Conceitual*. Anais IV Congresso Brasileiro de Manutenção, 1999.
- ANDERSON, E. W.; FORNELL, C.; LEHMANN, D. R. *Customer Satisfaction, Market Share and Profitability: Findings from Sweden*. Journal of Marketing, Vol. 58, July, 1994.
- AGRAWAL, R., S.; GHOSH, T.; IMIELINSKY, B. Yer, A. Swami. *An Interval Classifier for Database Mining Applications*. Proc. 18th Int. Conf. Very Large Databases, pp 560-573, Vancouver, 1992.
- BALLOU, D. G. K. Tayi. *Enhancing Data Quality in Data Warehouse Environments*. *Communications of the ACM*, 1999.
- CALADO, L. R.; BACARIM, D. P. *Qualidade dos Serviços de Telefonia Fixa: A Visão do Usuário Versus Prestadora*. XII SIMPEP, Bauru, SP, Brasil, novembro, 2005.
- BERRY, M.; LINOFF, G. *Mastering Data Mining: The Art and Science of Customer Relationship Management*. John Wiley & Sons, 2000.
- BERRY, L. L. *Emerging Perspectives on Services Marketing*. Chicago, Illinois: American Marketing Association, 1983.
- BEST, R. *Market based management*. Englewood Cliffs, NJ: Prentice-Hall International, 1997.
- BOZDOGAN, H. *Statistical Data Mining & Knowledge Discovery*. CHANPMAN & Hall/CRC, 2003.
- BRANCO, FILHO G. *Dicionário de termos de manutenção, confiabilidade e qualidade*. Rio de Janeiro: ABRAMAN, 1996.
- BUOSI, T.; GANGA G. M. D., et al., *Proposição de um Modelo Conceitual para o Projeto de Sistemas de Medição de Desempenho*. X SIMPEP, 2003.

- FILHO, R. A., Carvalho, N. C. D., *et al.*, (Auto) *Avaliação do Desempenho Empresarial da Manutenção: O Diferencial Competitivo na Busca da Sustentabilidade (Aplicação das Metodologias do PNQ e da ISO 9000 ao Ambiente da Manutenção)*. XI SIMPEP, Bauru, SP, Brasil, novembro, 2004.
- COETZEE, J. L. *Maintenance*. Republic of South Africa: Maintenance Publishers, 1998.
- CORRÊA, H. L. *The Links between Uncertainty, Variability of Outputs and Flexibility in Manufacturing Systems*. University of Warwick, Warwick, 2002.
- CONTRERAS R. J. *Técnicas de Seleção de Características Aplicadas a Modelos Neuro-Fuzzy Hierárquicos*, Dissertação de Mestrado, Departamento de Engenharia Elétrica, Pontifícia, Rio de Janeiro, 2002.
- CHAPMAN, Pete; CLINTON, Julian; KERBER, Randy; KHABAZA, Thomas; REINARTZ, Thomas; SHEARER, Colin; WIRTH, Rüdiger; “CRISP-DM 1.0 – Step-by-Step data mining guide”; CRISP-DM Consortium, 2000.
- CSILLAG, J. M. *Análise do Valor*. 4^o edição, São Paulo: Atlas, 1995.
- DATE, C.J. *Introdução a Sistemas de Banco de Dados*. Rio de Janeiro: Campus, 1991.
- DAVIS, T. R. V. *Internal Service Operations: Strategies for Increasing Their Effectiveness and Controlling Their Cost*. Organizational Dynamics, Vol. 20, Autumn, 1992.
- DAVIS, T. R. V. *Managing Internal Service Delivery in Organizations*. Advances in Services Marketing, Vol. 2, 1993.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. Wiley-Interscience, 2^o Edition, 2000.
- DWIGHT, R. A. *Concepts for Measuring Maintenance Performance*. New Development in Maintenance: An International View. Moret Ernst and Young, 1995.
- EADE, R. *The Importance of Predictive Maintenance*. Iron Age New Steel, 1997.
- EDWARD, Corwin D. *The Meaning of Quality*. Quality Progress, October, 1968.
- FACINA, Taís. *Manutenção Sem Fronteiras*. Revista Manutenção. Julho e Agosto, 1999.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. “*Advances in Knowledge Discovery & Data Mining*”, Cambridge, MA (The AAAI Press/The MIT Press), 1996.

- FERREIRA, J. B. *Mineração de Dados na Retenção de Clientes em Telefonia Celular*. Engenharia Elétrica, Pontifícia, Rio de Janeiro, 2005.
- FERREIRA, J. I. A. X. *Como Medir A Competitividade das Organizações*. Revista Produção on-line, Num. 2, Agosto, de 2006.
- FERREIRA, Aurélio Buarque de Holanda. *Dicionário Aurélio Básico da Língua Portuguesa*. São Paulo: Nova Fronteira, 1995.
- FITZSIMMONS, James A.; MONA, J. *Administração de Serviços – Operações, Estratégias, e Tecnologia da Informação*. 4ª edição, São Paulo: Bookman, 2004.
- FELDENS, M.A. *et al. Towards a Methodology for the Discovery of Useful Knowledge Combining Data Mining, Data Warehousing and Visualization*. CLEI (Conferência Latino americana de Informática), Equador, 1998.
- FREITAS, A. L. P. *A Qualidade em Serviços no Contexto da Competitividade*. Revista Produção On-Line, Vol. 5, 2005.
- FRIEDMAN J. H. *An Overview of Predictive Learning and Function Approximation*. In From Statistics to Neural Networks: Theory and Pattern Recognition Applications. Berlin, Springer, 1995.
- FNQ (Fundação Nacional de Qualidade) site disponível: <http://www.fnq.org.br/site> acesso em: 26/04/2008.
- GALE, B. T. *Gerenciando o Valor do Cliente – Criando Qualidade & Serviços que os Clientes Podem Ver*. São Paulo: Pioneira, 1994.
- GERAERDS, W. M. J. *The Cost of Downtime for Maintenance: Preliminary Considerations*. Maintenance Management International, Vol.5, 1985.
- GERAGHETY, T. *Obtendo a Efetividade do Custo de Manutenção Através da Integração das Técnicas de Monitoramento de Condição, RCM e TPM*. SQL Systems, Brasil, 2001.
- GEROLAMO, M. C.; ESPOSTO, K. F., *et al. Modelo para Identificação de Ações de Melhoria de Desempenho Alinhadas à Estratégia*. XXII Encontro Nacional de Engenharia de Produção Curitiba, PR, outubro, 2002.
- GIANESI, I. G. N.; CORRÊA, H. L. *Administração Estratégica de Serviços – Operações Para a Satisfação do Cliente*. São Paulo: Atlas, 1994.
- GRONROONS, Christian. *Marketing Gerenciamento e Serviço A Competição por Serviço na Hora da Verdade*. 4ª edição, São Paulo: Campus, 1993.

- GAVIN, D. A. *The Process of Organization and Management Sloan Management*. Review: 33, 1998.
- GOLDSCHMIDT, Ronald; PASSOS, Emmanuel. *Data Mining – Um Guia Prático, Conceitos, Técnicas, Ferramentas, Orientações e Aplicações*. São Paulo: Campus, 2005.
- GOUVEIA, J. B.; JÚNIOR J. F. *Um Cenário Organizacional em Constantes Transformações*. (Artigos para publicação). Departamento de Economia, Gestão e Engenharia Industrial, Universidade de Aveiro, Portugal, 2004.
- GROTH, Robert. *Data Mining: Building Competitive Advantage*. Prentice Hall PTR, USA, 2000.
- GROTH, Robert. *Data Mining: a Hands-on Approach for Business Professionals*. New Jersey, EUA: Prentice Hall, 1997.
- GROVER, G. J., S.R. *The Implementation of Business Process Reengineering*. Journal Of Management Information Systems, Vol.12. 1995.
- GUMMESSON, E. *Marketing Revisited: The Crucial Role of the Part-Time Marketers*. European Journal of Marketing, Vol. 25, Num. 2, 1991.
- HAIR, J. J. F. *Análise Multivariada de Dados*. São Paulo: Bookman, Vol.1. 2005.
- HAN, J.; KEMBER, M. *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers, 2001.
- HARDY, W. C. *QoS Measurement and Evaluation of Telecommunications Quality of Service*. New York: John Wiley & Sons, 2001.
- HESKETT, James L. *Service Breakth Roughs: Changing the Rules of the Gam*. The Free Press, 1990.
- HESKETT, J. L.; JONES, T. O.; LOVEMAN, G. W.; SASSER, W.E.; SCHLESINGER, L. A. *Putting the Service-Profit Chain to Work*. Harvard Business Review, March-April, 1994.
- HUBER, F.; HERMANN, A; MORGAN, R. *Gaining Competitive Advantage Through Customer Value Oriented*. Journal of Consumer Marketing, Vol.18, Num.01, 2001.
- IACOBUCCI, D.; GRAYSON, K. A.; OSTROM, A. L. *The Calculus of Service Quality and Customer Satisfaction: Theoretical and Empirical Differentiation and Integration*. Greenwich, CT: JAI Press, Vol. 3, 1994.
- IRELAND, F; DALE, B. G. *A Study of Productive Maintenance Implementation*. Journal of Quality in Maintenance Engineering, Vol.7, Num.3, 2001.
- JUDD, R. C. *Differentiates with the 5th P: People*. Industrial Marketing Management, 1987.

- JURAN, J. M. *A Qualidade Desde o Projeto: Novos Passos Para o Planejamento da Qualidade em Produtos e Serviços*. São Paulo, Vol.2. 1994.
- KAPLAN, R. S. N. *A Estratégia em Ação: Balanced Scorecard*. Tradução de Luiz Euclides Trindade Frazão Filho. Rio de Janeiro, 1996.
- KIYAN, F. M. *Proposta Para Desenvolvimento de Indicadores de Desempenho Como Suporte Estratégico*. Engenharia de Produção, Unidade Escola de Engenharia de São Carlos (EESC), 2001.
- KLEMETTINEN, M.; MANNILA, H.; TOIVONEN, H. *A Data Mining Methodology and Its Application to Semi Automatic Knowledge Acquisition*. DEXA Workshop, 1997.
- KOTLER, Philip. *Administração de Marketing – Análise Planejamento Implantação e Controle*. 3º edição, São Paulo: ATLAS, 1996.
- KOTLER, Philip. *Administração de Marketing – Análise, Planejamento, Implementação e Controle*. 4º edição, São Paulo: ATLAS, 1998.
- LAROSE, D. T. *Discovering Knowledge in Data an Introduction to Data Mining*. New Jersey, 2005.
- LEAL, R. M. D. P. *Atraso e Modernidade no Brasil Globalizado. Uma Análise do Discurso da Mídia na Privatização das Telecomunicações*. Comunicação e Cultura da Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.
- LAPIERRE, J. *Customer - Perceived Value in Industrial Contexts*. Journal of Business & Industrial Marketing, Vol.15, Num. 2/3, 2000.
- LICHTENTHAL, J. D.; WILSON D. T.; LONG, M. M. *Scientific Contribution to the Field from the Journal of Business-To-Business Marketing*. Journal of Business Research, Vol. 38, 1998.
- LILJANDER, V.;STRANDVIK, T. *The Relation between Service Quality, Satisfaction and Intentions*. Managing Service Quality, Vught, 1995.
- LOVELOCK, C. *Classifying Services to Gain Strategic Marketing Insights*. Journal of Marketing, Vol. 47, Num.3, 1983.
- MALHOTRA, N. K. *Pesquisa de Marketing: Uma Orientação Aplicada*. São Paulo, 2004.
- MANTOVANI, C. A. *Uma Sistemática de Gestão da Capacidade Apoiada na Análise e Melhoria dos Processos para Prestadoras de Serviços de Telecomunicações*. Mestrado, Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2001.

- MARCELLI, R. P. *O Papel dos Indicadores de Desempenho na Estratégia das Organizações para o Aprimoramento de Processos: Um Estudo de Caso*. Mestrado, Departamento de Engenharia de Produção e Sistemas, Universidade Federal de Santa Catarina, Florianópolis, 2000.
- MARGEM, F. M. *Uma Análise da Qualidade de Serviço nas Lojas OI Telemar Utilizando Método Servqual*. XII SIMPEP, Bauru, SP, Brasil, Novembro, 2005.
- MARTINS, R. A.; NETO, P. L. D. O. C. *Indicadores de Desempenho Para a Gestão Pela Qualidade Total: Uma Proposta de Sistematização*. Gestão e Produção, Vol. 5, 1998.
- MAGGARD, B. N., RHYNE, D. M. *Productive Maintenance: A Timely Integration of Production and Maintenance*. Production and Inventory Management Journal, 6-11, 1992.
- MCCARTHY, E. J. *Basic Marketing*. Homewood, Ill.: Irvin, 1960.
- MEEKING, A. *Unlocking the Potential of Performance Measurement: a Practical Implementation Guide*. Public Money & Management, 5-12, 1995.
- MEREDITH, J. R. *The Management of Operations: A Conceptual Emphasis*. 4th, ed. Wiley, 1992.
- MUDGE, A. E. *Numerical Evaluation of Functional Relationships*. Texas: SAVE Proc., Vol. 2, 1967.
- MUDGE, A. E. *Value Engineering: A Systematic Approach*. 2º Ed., Pennsylvania, 1981.
- MUDGE, A.E. *Numerical Evaluation of Functional Relationships*. Value Engineering, September, 1968b.
- MUDGE, A.E. *The Preparation of Use of Value Engineering Chart*. Value Engineering, July, 1968a.
- MICHAELIS. *Moderno Dicionário da Língua Portuguesa*. São Paulo: Companhia Melhoramentos, 1998.
- MIRSHAWKA, Victor; OLMEDO, Napoleão L. *TPM À Moda Brasileira*. São Paulo: MAKRON Books, 1994.
- MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd ed. Springer-Verlag, March 1996.
- MOLBRAY, John. *Reliability Centered Maintenance*. Oxford: Butterworth Heinemann, 1997.
- The Responsible Custodianship of Physical Assets. Disponível em: <<http://www.aladon.co.uk/12cust.html>>. Acesso em: 04 set. 2002.

- MONCHY, François. *A Função Manutenção: Formação Para a Gerência da Manutenção Industrial*. São Paulo: Durban, 1989.
- MORAIS, A. P. D.; GODOY, L. P. *Qualidade em Serviços: Uma Abordagem Conceitual*. XII SIMPEP, Bauru, SP, Brasil, Novembro, 2005.
- MORIK, K. *The Representation Race – Preprocessing for Handling Time Phenomena*. Proceeding of the European Conference on Machine Learning, 2000.
- MOSS, Marvin A. *Designing for Minimal Maintenance Expense: A Practical Application of Reliability and Maintainability*. New York, Marcel Dekker, 1985.
- NAKAJIMA, S. *Productive Maintenance Development Program: Implementing Total Productive Maintenance*. Cambridge, Productivity Press, 1989.
- NBR5462. *Confiabilidade e Mantabilidade*. Rio de Janeiro: ABNT, 1994.
- NEELY, A. et. al. *Designing Performance Measures: A Structured Approach*. International Journal of Operations & Production Management, Vol. 17, Num.11, pp. 1131-1152, 1997.
- NETTO, C. A. A. *Proposta de Modelo de Mapeamento e Gestão Por Macroprocessos*. Doutorado, Departamento de Engenharia de Produção, Escola Politécnica da Universidade de São Paulo, São Paulo, 2004.
- NORMANN, R. *Administração de Serviços: Estratégia e Liderança nas Empresas de Serviços*. São Paulo: Atlas, 1993.
- NOWLAN, F. S.; HEAP, H. *Reliability Centered Maintenance*. National Technical Information Service, Springfield, Virginia, US Department of Commerce, 1978.
- NUNES, E. Laércio. *Manutenção Centrada em Confiabilidade (MCC): Análise da Implantação em Uma Sistemática de Manutenção Preventiva Consolidada*. Mestrado, Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis. 2001.
- PARASURAMAN, A., Zeithaml, V.A., Berry, L. L. *A Conceptual Model of Service Quality and Its Implications for Future Research*. *Journal of Marketing, American Marketing Association*, 1985.
- PARASURAMAN, A, ZEITHAML, V.A BERRY, L. L. *SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality*. *Journal of Retailing*, Vol. 64, 1988.
- PIATETSKY-SHAPIRO, “*Measuring Lift Quality in Database Marketing*”. *Journal of SIGKDD Exploration Newsletter*, Vol. 2, New York; 2000.

- PEHANICH, M. *Behind the lines*. Prepared Foods, 164, 12, 87, 1995.
- PORTER, M. E. *Competitive Advantage – Creating and Sustaining Superior Performance*. New York, the Free Press, 1985.
- PINTO, Alan Kardec; XAVIER, Júlio Nassif. *Manutenção: Função Estratégica*. Rio de Janeiro: Qualitymark, 1998.
- PYLE, D. *Data Preparation for Data Mining*. Departamento de Engenharia de Produção, Escola Politécnica da Universidade de São Paulo, São Paulo, 1999.
- RAVALD, A; GRÖNROOS, C. *The Value Concept and Relationship Marketing*. European Journal of Marketing, Vol. 30, Num. 2, 1996.
- REBOUÇAS, D. P. *Planejamento Estratégico – Conceito, Metodologias e Práticas*. São Paulo: ATLAS, 1998.
- ROSE, K. H. *A Performance Measurement Model*. Quality Progress, 1995.
- SALOMI, G. G. E. et al. *Servqual x Servperf: Comparação Entre Instrumentos Para Avaliação da Qualidade de Serviços Internos*. Gestão & Produção, Vol. 12 Num. 2, 2005.
- SAS Enterprise, Site do software de mineração de dados. Disponível: <http://www.sas.com/technologies/analytics/datamining/miner/index.html> [acesso 10/2006].
- SAYLES, L. R. *Managerial Behavior*. New York: Mc Graw-Hill, 1964.
- SCHOELL, W. F.; IVY, J. T. *Marketing: contemporary concepts and practices*. Boston, Mass.: Allyn and Bacon, 1981.
- SLACK, N., CHAMBERS, S., HARLAND, C., HARRISON, A., JOHNSTON, R. *Administração da Produção*. São Paulo: ATLAS, 1997.
- SIQUEIRA, H. Soares Gonçalves. *Formação Interdisciplinar: Exigência Sócio Política Para Um Mundo Em Rede*. 2004, <http://www.angelfire.com/sk/holgonsi/mundorede.html> acesso em julho de 2007.
- SÓRIA, F. G. *Proposta de Alinhamento de Modelos Conceituais Para Melhoria Contínua dos Processos da Manufatura*. XII SIMPEP, Bauru, SP, Brasil, 2005.
- STAUSS, B. *Internal Services: Classification and Quality Management*. International Journal of Service Industry Management, Vol.6, Num. 2, 1995.
- TACHIZAWA, Takeshy et al. *Gestão de Negócios – Visões e Dimensões da Organização*. 2ª edição, São Paulo: ATLAS, 2003.

- TAKASHINA, Newton Tadachi; FLORES, Mario Cesar Xavier. *Indicadores da Qualidade e do Alto Desempenho: Como Estabelecer Metas e Medir Resultados*. Rio de Janeiro: Qualitymark, 1996.
- TAVARES, Lourival. *Administração Moderna da Manutenção*. Rio de Janeiro: Novo Pólo Publicações, 1999.
- TÉBOUL, James. *A Era dos Serviços – Uma Abordagem de Gerenciamento*. Rio de Janeiro: Qualitymark, 1995.
- WALTERS, D. *Marketing and Operations Management: Integrated Approach to New Ways of Delivering Value*. Management Decisions, MCB Press, n.37/3, 1999.
- WALTERS, D; LACASTER, G. *Value and Information – Concepts and Issues for Management*. Management Decisions, MCB Press, n.37/8, 1999.
- WEBSTER, F. E. *Market Driven Management*. New York, Wiley, 1994.
- WIRTH, Ruediger. “*CRISP-DM Position Statement*”. 6th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, USA, 2000.
- WIRTH, R. et al. *Toward Process-Oriented Tool Support for Knowledge Discovery in Database*. 1997.
- WIKSTROM, S.; NORMANN, R. *Knowledge and Value: A New Perspective on Corporate Transformation*. London, Routledge, 1994.
- ZEITHAML, Valarie A.; PARASURAMAN, A.; BERRY, Leonard L. *Delivering Quality Service: Balancing Perception and Expectations*. Free Press, 1990
- ZEITHAML, V. A. *Consumer Perceptions of Price, Quality and Value: A Means-End Model and Synthesis of Evidence*. Journal of Marketing, Vol. 52, July, 1988.
- XENOS, Harilaus Georgius d'Philippos. *Gerenciando a Manutenção Produtiva*. Belo Horizonte: Desenvolvimento Industrial, 1998.

ANEXO I

Indicadores de Gestão no Setor de Telecomunicações no Brasil

O mercado de telecomunicações no Brasil já em 1960 estabeleceu padrões de qualidade para os serviços vinculados às linhas telefônicas. Para regulamentar esses padrões o País criou o Ministério das Comunicações. Nessa época o mercado apresentava um serviço estatal sem concorrência e com pouco foco em serviço de atendimento ao cliente as expectativas dos clientes, chamados de Usuários.

A partir da segunda metade dos anos 90 surgiu à necessidade de investimentos na área de telecomunicações, sendo a privatização do setor uma das premissas do Fundo Monetário Internacional (FMI) que foi imposta ao governo neoliberal do presidente Fernando Henrique Cardoso com intuito de sanear a dívida externa. Para isto foi criada a Lei 9472 de julho de 1997 que dispõe sobre a organização dos serviços de telecomunicações, e cria-se um órgão regulador. Aspectos institucionais e da Lei Geral das Telecomunicações aprovada em votação na Câmara dos Deputados, no Senado Federal e sancionada pelo presidente da República Fernando Henrique Cardoso. Abre-se definitivamente o mercado para a privatização do setor, dando início ao período atual.

Após esses movimentos legais foi aprovado o Plano Geral de Telecomunicações em 1998 que no seu Artigo 1 constava sobre o plano geral de qualidade para o serviço telefônico fixo comutado, esse artigo foi o precursor do regulamento de indicadores de qualidade que tem os seguintes objetivos:

Art. 1º Esse regulamento estabelece as definições, métodos e frequência de coleta de informações, consolidação e envio à Agência Nacional de Telecomunicações - ANATEL, de indicadores de qualidade apresentados no Plano Geral de Metas de Qualidade para o Serviço Telefônico Fixo Comutado – PGMQ - STFC, aprovado pela Resolução nº 341, de 20 de junho de 2003.

Art. 2º O conjunto de informações do presente regulamento tem por objetivo possibilitar o desenvolvimento do processo de gestão, pela ANATEL, da qualidade do Serviço Telefônico Fixo Comutado, destinado ao uso do público em geral - STFC, prestado nos regimes público e privado.

Parágrafo único. Esse regulamento é aplicável a todas as chamadas realizadas no âmbito da rede de suporte do STFC, assim como, a todas as chamadas entre essa rede e redes de suporte a outros serviços de telecomunicações de interesse coletivo a ela interconectada.

Nas disposições gerais em seu artigo cinco estabelece como obrigatória, pelas prestadoras do STFC, a coleta, o cálculo, a consolidação e o envio à ANATEL, das informações relativas a todos os indicadores definidos no presente regulamento, respeitando as particularidades das outorgas de cada prestadora. Sendo o resultado os indicadores consolidados enviados, mensalmente à ANATEL.

O serviço de Internet em banda larga que surgiu no Brasil no ano de 1999, como uma alternativa para agregar valor ao serviço telefônico fixo comutado assumiu como indicadores de qualidade aqueles que se mostraram eficientes no acompanhamento e gestão do serviço telefônico fixo.

ANEXO II

A Tabela AII.1 ilustra o mínimo que pode ser considerado estatisticamente significativo com um poder de 0,80 para diversos números de variáveis independentes e tamanhos de amostras.

Tabela AII. 1: Níveis de significância.

Tamanho da Amostra	Nível de Significância (α)=0,01					Tamanho da Amostra	Nível de Significância (α)=0,05				
	Número de Variáveis Independentes						Número de Variáveis Independentes				
	2	5	10	15	20		2	5	10	15	20
20	45	56	71	NA	NA	20	39	48	64	NA	NA
50	23	29	36	42	49	50	19	23	29	42	42
100	13	16	20	23	26	100	10	12	15	18	21
250	5	7	8	9	11	250	4	5	6	9	8
350	4	5	5	6	7	350	4	5	5	7	9
500	3	3	4	5	6	500	3	4	5	6	7
1000	1	2	2	2	3	1000	1	1	2	2	2

Fonte: Adaptado do HAIR, (2005).

ANEXO III

Estimativa de Custo da Taxa de abandono

Como já foi verificado anteriormente no Capítulo 2, a taxa de abandono gera um custo para a empresa. Aqui se tenta quantificar o lucro trazido por um cliente fiel, com uma média de consumo de R\$ 100,00 por mês com o serviço de banda larga, com uma expectativa de receita de 10 anos a uma taxa de desconto de 13% ao ano. Isso representa teoricamente o fluxo de caixa que esse cliente particular gera.

Inicialmente é estimado o fluxo de receita do cliente, como valor futuro. A Tabela AIII. 1 ilustra quanto se perde em faturamento por ano com a taxa de abandono.

Tabela AIII.1: Estimativa de perda.

Valor Gasto Pelo Cliente Por Ano		Taxa de Juros Ano	VF	Quantidade de Abandono / Ano
Receita	R\$ 1.200,00	13%	R\$ 1.356,00	160.508
Total de Perda de Receita (VF*Quantidade Abandono Ano)		R\$ 217.648.848,00		

Fonte: elaborado pelo autor.

Depois se calcula o valor gasto com uma retirada de cliente da empresa conforme a Tabela AIII. 2. Esta Tabela identifica os valores médios gastos com uma retirada e o total desses em um ano.

Tabela AIII. 2: Valores médios gastos com abandono.

Item	Valor	Quantidade	Total
Retirada	R\$ 59,00	160508	R\$ 9.469.972,00
Instalação	R\$ 111,00	160508	R\$ 17.816.388,00
Material	R\$ 23,00	80254	R\$ 1.845.842,00
Comissão de Vendas	R\$ 25,00	160508	R\$ 4.012.700,00
Postos <i>Call Center</i> HH	R\$ 26,70	26751,3	R\$ 714.260,60
Total	R\$ 33.859.162,60		

Fonte: elaborado pelo autor.

Para finalmente somar os valores da receita estimada perdida com as despesas gastas com as retiradas. Verifica-se, então, na Tabela AIII. 3 os custos com a taxa de abandono.

Tabela AIII. 3: Estimativa de despesas com abandono.

Item	Valor Por Ano
Total de Perda de Receita	R\$ 217.648.848,00
Despesa Total com Retiradas	R\$ 33.859.162,60
Total	R\$ 251.508.010,60

Fonte: elaborado pelo autor.

Verifica-se por meio da tabela AIII, que muito dinheiro é perdido com o abandono dos clientes no serviço de banda larga, a partir das análises efetuadas nesta dissertação, fica clara a importância de atender as necessidades e superar as expectativas dos clientes. Para essa superação identificou-se que uma alta qualidade no serviço influencia positivamente e que se deve medir e acompanhar os indicadores de qualidade relacionando sempre, a melhoria de qualidade do serviço ao aumento de satisfação do cliente. Consequentemente isso causará uma diminuição na taxa de abandono.

Não se tem a pretensão de associar todas as perdas causadas pela taxa de abandono aos problemas com a qualidade no serviço. Nessa pesquisa encontra-se uma explicação para 50,4 % da taxa de abandono realizada no ano de 2006 que foi referente à “falta de qualidade”.

Para sugerir redução nos indicadores de qualidade de serviço procurou-se analisar parâmetros de qualidades internacionais referendados pela empresa em análise. Na Tabela AIII. 4 mostra-se alguns dos parâmetros analisados.

Tabela AIII. 4 : Parâmetros de qualidade.

Indicador	Média	Melhoria Proposta
TMI	88 Horas	80 Horas
TRP	6 %	5 %
TMR	4 Horas	3 Horas
QRPR	8 %	12 %

Fonte: elaborado pelo autor.

ANEXO IV

Breve Histórico das Comunicações.

O processo de comunicação visual foi desenvolvido no princípio da década de 1790, pelo engenheiro francês Claude Chappe, que inventou a palavra telégrafo (do grego, “escrever à distância”). Consistia em transmitir letras, palavras e frases através de um código visualizado a partir de 3 réguas de madeira articuladas colocadas na parte alta de um poste ou edifício. A primeira linha de semáforos data de 1794 e ligava Paris a Lilly, distantes 225 km. Este sistema teve larga difusão no século XVIII e princípios do século XIX na França e em outros Países. Um destes telégrafos de semáforos esteve instalado no alto das Torres de São Suplicio em Paris e foi usado para transmitir as notícias das campanhas Napoleônicas. Estes processos visuais de comunicação estavam obviamente dependentes das condições naturais de visibilidade.

Só após a criação dos processos elétricos o telegrafo foi impulsionado de forma ímpar a velocidade e o alcance da transmissão de mensagens à distância. Um dos pioneiros foi o médico espanhol Francisco Salvá, de Barcelona. Em 1795 transmitiu mensagens por meio da descarga de um condensador. Em 1804, o mesmo Salvá criou outro tipo de telégrafo elétrico constituído por fios (cada um correspondendo a uma letra), por vasos de água nas extremidades e por uma pilha de Volta. A letra transmitida era detectada pela formação de bolhas gasosas formadas no vaso correspondente à letra. As bolhas eram obtidas por eletrólise da água. Com este sistema conseguiu enviar mensagens até 1 km de distância. Em 1812, o Dr. Samuel Von Sömmering de Munique, na Alemanha, atingiu 3 km de distância com um telégrafo do mesmo tipo. O seu sistema era constituído por 24 fios ligados a 24 voltâmetros que correspondiam às 24 letras do alfabeto e mais 10 para algarismos.

Em 1832, o diplomata e barão russo Pawel Schilling inventou um sistema de telégrafo elétrico com a inovação de necessitar apenas de 6 fios, ao contrário dos anteriores. As letras transmitidas eram detectadas pelo movimento de agulhas magnéticas colocadas sobre bobinas. Ao serem percorridas por corrente, as bobinas produziam campos magnéticos que faziam desviar as agulhas. Também os físicos alemães Gauss e Weber fizeram funcionar um telégrafo em 1833, baseado nos movimentos da agulha magnética.

O inglês William Cook viu uma demonstração do telégrafo de Schilling em 1836 em *Heidelberg* e construiu vários aparelhos semelhantes, associando-se a Charles Wheatstone, professor do *King's College* de Londres. O primeiro telégrafo por eles desenvolvido foi patenteado em 1837. Tinha 6 fios e 5 agulhas magnéticas, donde lhe veio o nome de telégrafo de 5 agulhas. As agulhas eram acionadas por eletroímãs. Eram acionadas duas agulhas de cada vez, pois cada letra era definida por duas agulhas. Por exemplo, a letra H era definida pelas primeira e segunda agulha. Este sistema foi usado em 1839 em Inglaterra, entre *Paddington e West Drayton*, numa distância de 21 km, servindo para informar as posições dos comboios. Tornou-se o primeiro serviço telegráfico comercial e foi também o primeiro uso comercial da eletricidade. Para chegar a este ponto foi necessário o desenvolvimento de várias técnicas, nomeadamente das pilhas elétricas, dos eletroímãs e da fabricação dos fios de cobre em lugar dos fios de ferro. Mais tarde, em 1839, Cook e Wheatstone criaram um telégrafo mais simples, o telégrafo de 2 agulhas. Podia transmitir vinte e duas palavras por minuto.

Nos Estados Unidos, o pintor Samuel Morse inventou um sistema mais prático, com um interruptor, um eletroímã e apenas um fio. Em 1838, Morse registrou uma patente com a descrição do seu telégrafo. Visto usar apenas um fio foi necessário utilizar um código para cada letra constituído por pontos, traços e espaços. Os pontos correspondiam a uma ação breve sobre o eletroímã, o traço a uma ação mais longa e o espaço a uma pausa. Assim era à base do código Morse.

Em 1844 este sistema foi experimentado entre *Baltimore e Washington*, numa distância de 64 km, utilizando aquele código. Este sistema de Morse permitiu um grande desenvolvimento do telégrafo. Em 1852 haviam sido instalados 64 000 km de linhas telegráficas no mundo. Em 1848 Wheatstone construiu o primeiro telégrafo ABC, com um só fio e um eletroímã. Possuía um disco que indicava as letras recebidas. Em 1851 Wheatstone modificou o modelo para imprimir as letras numa tira de papel. Verificou-se que os operadores de Morse facilmente decifravam o código por audição dos eletroímãs, sem recorrer à leitura do papel, o que levou à adaptação do aparelho de forma a produzir um estalido. Com este processo a transmissão atingia as trinta palavras por minuto.

Em 1855, o professor inglês Davis Hughes inventou na América o telégrafo impressor, constituído por um teclado no lado emissor, em que cada tecla correspondia a uma letra e por uma máquina impressora no lado receptor. Wheatstone continuou a melhorar os telégrafos,

nomeadamente o de Morse. Um dos aperfeiçoamentos foi à inclusão, em 1858, de uma perfuradora de fita de papel, o que permitiu atingir quase seiscentas palavras por minuto, pois as mensagens eram perfuradas na fita antes de serem enviadas. Por outro lado o aumento do tráfego telegráfico trazia grandes problemas técnicos e econômicos às empresas telegráficas que ansiavam por aumentar a capacidade de tráfego das suas linhas.

As próximas invenções vieram nesse sentido. Em 1872, o americano Joseph B. Stearns inventou o telégrafo duplex que permitia a transmissão de duas mensagens simultâneas pela mesma linha, uma em cada sentido. Em 1872 o francês Jean-Maurice-Émile Baudot, inventou o telégrafo *multiplex*, que veio permitir a transmissão de duas ou mais mensagens simultaneamente, pela mesma linha e no mesmo sentido. Edison também desenvolveu um telégrafo duplex, aperfeiçoando o de Stearns e em 1874 demonstrando que o seu telégrafo quadruplex era capaz de transmitir simultaneamente quatro mensagens pelo mesmo fio, duas num sentido e duas no outro. Em 1876 ano da Exposição Centenária da Philadelphia, Gram Bell decidiu fazer uma exibição de suas invenções a Sir. William Thomson. A Bell Telephone foi criada em 1878, em Nova Iorque, com 5.000 ações.

As primeiras centrais telefônicas automáticas

Pouco tempo após a invenção do telefone e das centrais de comutação, surgiu a idéia de que a ligação entre as várias linhas telefônicas poderia ser feita automaticamente, sem a ajuda de operadores. Em uma central telefônica automática, a própria pessoa que quer telefonar envia sinais elétricos especiais de seu aparelho para certos instrumentos na central telefônica, e esses instrumentos ligam a pessoa com o telefone desejado. Em 1879, os irmãos Thomas e Daniel Connelly, juntamente com Thomas J. McTighe, patentearam o primeiro sistema em que um usuário podia controlar um mecanismo de comutação à distância.

Esquema do sistema de comutação automática de Connelly e McTighe, o aparelho se baseava nos telégrafos ABC de Wheatstone e era bastante primitivo, nunca tendo sido utilizado na prática. A parte principal do sistema era um tipo de roda dentada, semelhante aos dispositivos utilizados em relógios, que era movida por meio de um eletroímã e que só podia girar o espaço de um "dente" de cada vez. Cada vez que o eletroímã recebia um pulso de eletricidade, ele atraía uma barra metálica, e esta fazia a roda dentada girar um espaço. À

medida que a roda girava, ela ia movendo um braço metálico que entrava em contato sucessivamente com os contatos referentes às várias linhas telefônicas. Assim, enviando sucessivos pulsos elétricos, era possível escolher, à distância, a ligação desejada com um pequeno número de linhas.

Em 1884, Ezra Gilliland, da companhia Bell, desenvolveu um sistema de comutação automática que podia trabalhar com um máximo de 15 linhas. Nesse sistema primitivo, semelhante ao de Connely e McTighe, porém um pouco mais simples, havia um contato metálico que se movia passo a passo, pulando de uma posição para outra, quando o usuário apertava um botão. O número de vezes que o botão era apertado determinava a conexão que era produzida. Esse sistema também não foi utilizado, na prática.

O sistema automático Strowger

O avanço realmente importante ocorreu em 1889, quando um agente funerário da cidade de Kansas chamado Almond B. Strowger começou a desenvolver um sistema automático de comutação que realmente funcionou. Conta-se que Strowger desconfiava das telefonistas, e acreditava que elas propositadamente desviavam os chamados de seus clientes para um outro agente funerário. Por isso, ele resolveu inventar um sistema de comutação que dispensasse o uso das telefonistas. Almond Strowger Após vários estudos e tentativas, conseguiu construir, com a ajuda de um relojoeiro, um sistema para 100 linhas telefônicas. O sistema foi patenteado em 1891, e no mesmo ano Strowger estabeleceu a *Automatic Electric Company* para comercializar seu invento.

A primeira central telefônica automática, utilizando o sistema de Strowger, foi aberta em *La Porte*, Indiana, em 1892. Nos dez primeiros anos após o invento de Strowger, foram instaladas mais de 70 centrais automáticas nos Estados Unidos. O sistema de Strowger era um aperfeiçoamento dos aparelhos anteriores. Ele também tinha um dispositivo com um contato metálico principal, móvel, que se deslocava passo a passo, acionado por eletroímãs, "varrendo" diversos contatos fixos, cada um deles conectado a uma linha telefônica. Mas havia uma diferença importante: o sistema se movia dentro de um cilindro e podia tanto girar em torno do eixo do cilindro como também se mover para cima e para baixo. O cilindro tinha, em sua parte interna, 10 fileiras com 10 contatos metálicos cada uma, totalizando 100 contatos. A

"vassoura" ou "escova" metálica central podia assim se deslocar facilmente e escolher um dos 100 contatos, cada um dos quais representava uma linha telefônica.

O sistema de comutação automática desenvolvido por Strowger tem duas partes básicas. Uma é o dispositivo existente na própria central telefônica, que realiza as conexões entre as linhas telefônicas. A outra é um dispositivo colocado nos aparelhos dos usuários, que permite enviar sinais à central telefônica informando o número do telefone com o qual se quer fazer conexão. O dispositivo precisa conectar os fios do telefone que está chamando com qualquer uma das conexões dos outros aparelhos. Inicialmente, os usuários não dispunham de nenhum mecanismo especial para enviar os sinais à central. Eles tinham dois botões na caixa dos seus telefones, e deviam apertar cada um deles um determinado número de vezes. Por exemplo, para se conectar com o telefone número 34, era necessário apertar o primeiro botão 3 vezes e o segundo botão 4 vezes. Cada vez que um botão era apertado, ele enviava um pequeno pulso elétrico para a central e fazia o dispositivo se mover um passo. O sistema funcionava, portanto, da seguinte forma. Quando uma pessoa queria telefonar, ela tirava o telefone do gancho e então apertava os botões do seu aparelho. Na central telefônica, o dispositivo ligado a este telefone ia se movendo, primeiro na direção vertical, depois girando, até fazer a ligação com o número desejado. Então, a pessoa girava a manivela do magneto, para produzir um sinal na campainha do número chamado. Durante toda a conversa, o dispositivo Strowger se mantinha na mesma posição, ligando as duas linhas. Quando a conversa terminava, era necessário apertar um botão que fazia o dispositivo Strowger da central telefônica voltar para a sua posição inicial. Se a pessoa não apertasse esse botão, seu telefone continuaria conectado à linha que havia sido chamada antes. Com um dispositivo do tipo Strowger original, era possível escolher apenas uma de 100 linhas telefônicas.

Seria possível construir cilindros com maior número de contatos, mas isso era complicado, sob o ponto de vista técnico. Havia outros problemas com o sistema. Cada telefone precisava estar ligado à central telefônica através de vários fios: os fios que enviavam a voz, e outros fios que enviavam os sinais elétricos para mover o dispositivo automático. Isso encarecia bastante o sistema porque, o custo dos fios telefônicos era bastante alto. No caso do primeiro tipo de sistema, com botões, era necessário um fio para cada botão, aumentando dois fios para uma rede de 99 telefones. Um segundo problema era que cada telefone da rede precisava dispor do seu próprio dispositivo na central telefônica. Esses dispositivos eram caros

e ficava quase todo o tempo parado, pois cada usuário utilizava o telefone apenas durante uma pequena parte do dia. Outro problema era que uma pessoa podia se conectar a um telefone que já estava sendo usado. Nas centrais com telefonistas, isso não acontecia, porque a telefonista sempre verificava se a linha já estava ocupada. Nas primeiras centrais automáticas, não havia nada que impedisse uma pessoa de se ligar a uma linha ocupada, e nesse caso ela podia ouvir a conversa de outras pessoas ou intrometer-se na conversa. Por fim, cada pessoa precisava se lembrar de apertar um último botão ao terminar a conversa, para fazer o dispositivo do tipo Strowger da central voltar à posição inicial.

Aperfeiçoamentos do sistema Strowger

Strowger empregou algumas pessoas que ajudaram a aperfeiçoar o seu sistema automático. Em 1892, ele contratou Anthony E. Keith e, em 1894, Frank A. Lundquist e os irmãos Erickson (John e Charles). Foram essas pessoas, e não o próprio inventor, que aperfeiçoaram esse sistema.

Em 1896, Strowger teve problemas de saúde e se afastou da companhia que havia criado, morrendo em 1902. Keith resolveu em 1893 um problema simples: a pessoa que telefonava não precisava mais apertar um botão quando terminava a conversa. A própria colocação do telefone no gancho enviava um sinal à central telefônica, que fazia o dispositivo Strowger voltar à posição inicial. Além disso, em 1894, Keith, Lundquist e os irmãos Erickson eliminaram a possibilidade de uma pessoa se conectar a linhas já ocupadas.

Em 1896, Keith e os irmãos Erickson desenvolveram um processo que eliminava a necessidade de que os usuários ficassem apertando vários botões. Foi colocado no aparelho telefônico um sistema que enviava seqüências de pulsos do aparelho do usuário para a central. Esses dispositivos empregavam discos que eram girados e que, ao voltarem para sua posição, iam fazendo essencialmente à mesma coisa que a pessoa fazia antes apertando o botão várias vezes. É exatamente por causa desses discos, utilizados durante quase um século, que até hoje utilizamos a expressão "discar um número", embora utilizemos telefones que não possuem mais discos, e sim teclas. No sistema de discagem utilizado durante muitas décadas, o funcionamento sempre foi essencialmente o mesmo. A pessoa girava o disco até uma certa posição (um número) e largava o disco. Então, uma mola fazia o disco voltar para a posição

inicial, e ao mesmo tempo um mecanismo atrás dele fazia uma sucessão de contatos elétricos, enviando uma série de "clicks" para a central telefônica. Esses sinais produziam o mesmo efeito que apertar o primeiro botão várias vezes. Depois, quando a pessoa novamente girava o disco e enviava novos sinais, o efeito produzido na central telefônica era o mesmo de apertar o segundo botão várias vezes. Esse sistema reduziu também o número de fios ligando cada aparelho à central telefônica. No mesmo ano em que inventaram o sistema de discagem, Keith e os irmãos Erickson começaram a desenvolver um modo de utilizar o dispositivo de Strowger em uma rede com 1.000 linhas telefônicas. Em vez de construir um dispositivo maior, eles resolveram utilizar dois dispositivos. A idéia é simples. As linhas são divididas em "troncos", cada uma com 100 linhas. Cada tronco pode, portanto, utilizar um dispositivo Strowger comum, com 100 posições. Utilizando 10 dispositivos Strowger, é possível fazer conexões com 10 troncos de 100 linhas, atingindo assim 1.000 linhas.

Seletor de linha (1905) desenvolvido por Keith esse sistema, quando um usuário quer se conectar a uma das linhas, ele disca primeiramente um número, que corresponde à escolha do tronco; então, na central telefônica, um aparelho simples (um seletor com 10 posições) estabelece sua ligação com o tronco. Ao ser conectado a esse tronco, a sua linha é ligada a um segundo aparelho, do tipo Strowger, e os dois números seguintes enviados pelo usuário vão escolher a linha exata dentro do tronco. Não há limites para o sistema, pois é possível formar 10 grupos de 10 troncos (um total de 10.000 linhas), por exemplo, introduzindo mais uma etapa da discagem. Sistema automático tipo Strowger "passo-a-passo" com várias etapas, um telefone é ligado, primeiramente, a um dispositivo que tem 10 opções, cada uma delas é ligada a um aparelho tipo Strowger com 100 ligações, cada uma dessas 100 ligações pode levar a um outro aparelho Strowger, e assim por diante. Uma central desse tipo (com dois estágios) foi instalada em 1897, para 400 linhas, cada telefone tinha, na central, um seletor simples de apenas 4 posições, e 4 dispositivos do tipo Strowger de 100 posições. Portanto, a central telefônica tinha, ao todo, 400 seletores simples e 1.600 dispositivos do tipo Strowger. O sistema funcionou, mas era extremamente caro. Para redes com número ainda maior de linhas, o sistema se torna proibitivamente dispendioso. Portanto, embora o sistema pudesse ser aumentado de forma ilimitada, seu custo aumentava tanto que o tornava inviável para redes grandes.

O sistema “passo-a-passo”

Esse tipo de problema foi resolvido por Frank Lundquist, que havia saído da companhia de Strowger em 1896 e desenvolvido um novo processo, que foi patenteado em 1897. A idéia consistia essencialmente em utilizar um menor número de dispositivos na central telefônica automática, levando em conta que, da forma como o sistema funcionava, quase todos os dispositivos ficavam a maior parte do tempo sem funcionar. Na prática, como apenas cerca de 10% dos telefones era utilizado ao mesmo tempo, o número de dispositivos realmente utilizados a cada momento era ainda menor. Portanto, devia haver um modo de, em vez de ter um conjunto de dispositivos Strowger para cada linha telefônica, utilizar um número menor, partilhado por todos. Como fazer isso? Lundquist imaginou que bastaria introduzir um novo seletor que procurasse um dispositivo desocupado para usar. Vamos imaginar, primeiramente, uma rede com apenas 100 linhas, que exigia (pelo sistema antigo) 100 dispositivos do tipo Strowger. Pode-se agora construir uma central telefônica que tenha apenas 10 desses dispositivos. Cada um deles serve para fazer conexões com todas as 100 linhas, mas eles não "pertencem" a nenhum assinante em particular. Eles ficam disponíveis, esperando que alguém precise deles. Para que possam ser partilhados, é necessário haver um pré-seletor, ou seletor de dispositivo desocupado. Cada linha telefônica, ao chegar à central, é ligada a um pré-seletor desse tipo. No exemplo que estamos considerando, esse pré-seletor tem 10 posições diferentes, podendo ser ligado a cada um dos 10 dispositivos Strowger. Quando uma pessoa quer fazer uma ligação e tira o telefone do gancho, esse pré-seletor é colocado em movimento na central e vai girando de posição em posição e testando, até encontra um dispositivo do tipo Strowger desocupado. Então, um som especial é ouvido pela pessoa que está querendo fazer a ligação, e ela pode discar o número desejado, enviando os sinais elétricos para o seletor Strowger que foi selecionado. Quando termina a ligação, esse dispositivo Strowger fica desocupado, e pode ser utilizado por outra pessoa. Esse novo tipo de sistema, embora introduza uma etapa a mais, utiliza um menor número de dispositivos seletores, e por isso é muito mais barato. Assim, o sistema de ligação de uma linha a outra linha telefônica utiliza uma seqüência de etapas, passando por vários dispositivos encadeados. Esse sistema é chamado “passo-a-passo”.

Uma central para 4.000 linhas foi instalada em *New Bedford* (Estado de Massachusetts), e depois em outros locais (principalmente fora dos Estados Unidos). Esse sistema foi aperfeiçoado e adotado no mundo todo, até 1926, quando foi introduzido o sistema *crossbar* na Suécia. O serviço de comutação automática se expandiu mais fortemente na Europa (especialmente na França), na primeira década do século XX. Na década de 1910, as centrais automáticas se espalharam nos Estados Unidos, mas apenas entre os sistemas telefônicos "independentes" (isto é, os que não pertenciam ao sistema Bell). A Bell passou a utilizar no início do século XX algumas pequenas centrais automáticas em vilas e regiões rurais em que era difícil treinar e empregar telefonistas, mas apenas na década de 1910 começou a se preocupar com a necessidade de centrais automáticas exigidas por grandes cidades. Somente na década de 1920 a própria Bell adotou o sistema automático, utilizando telefones com discos, que passou a ser chamado de "sistema francês", embora tivesse sido inventado nos Estados Unidos. A substituição das centrais manuais foi lenta: até 1936, 52% das centrais dos Estados Unidos utilizavam telefonistas.

Histórico das Telecomunicações no Brasil.

Observa-se o universo das telecomunicações ao longo de sua existência, em três importantes períodos evolutivos:

1. A invenção e universalização do telefone,
2. a digitalização do sistema de telefonia e
3. a convergência total das plataformas de telecomunicações.

O primeiro ciclo teve como precedente a invenção do telégrafo, em 1837, iniciando-se efetivamente, em 1876, com a criação do telefone. Com duração de aproximadamente 90 anos, o primeiro grande ciclo das telecomunicações foi marcado, fortemente, por grandes descobertas e invenções, encerrando-se com as primeiras ligações DDD (Discagem Direta a Distância), em 1960.

A partir da década de 60, um novo ciclo teve início, estendendo-se até os dias atuais. Essencialmente voltadas à digitalização dos sistemas de telecomunicações, suas características

principais vêm sendo: a compactação constante do hardware, implicando em equipamentos cada vez menores, e o aumento crescente do software, acarretando em capacidades de processamento de informações cada vez maiores. Nesta fase, além dos serviços de voz, a comunicação de dados passou, gradativamente, ocupar maior espaço nas telecomunicações. O processo de digitalização iniciou, com maior abrangência, sobre os sistemas de transmissão, utilizando a técnica de digitalização de sinais denominada PCM (*Pulse Code Modulation*). Com menor intensidade nos primeiros anos, a digitalização foi também, aos poucos, sendo introduzida no ambiente das centrais telefônicas, através de equipamentos com tecnologia CPA (Controle por Programa Armazenado). Finalmente, as redes de acesso de última milha, que permaneciam analógicas quase que em toda extensão, começaram a ser equipadas, parcialmente, com sistemas digitais. As soluções mais utilizadas são compostas por dispositivos digitais compactos (ópticos e elétricos) PDH e SDH, *Modems* ópticos e enlaces de rádio digital.

Estas soluções têm possibilitado a digitalização dos meios de acesso para grandes e médios usuários. Outra forma, que permitiu aumentar o índice de digitalização da rede de acesso de última milha, foi introduzida através de terminais RDSI (Rede Digital de Serviços Integrados). Na década de 80 as linhas ADSL começaram a ser desenvolvida na Europa e nos Estados Unidos, esta tecnologia emergente vinha para suportar as aplicações de vídeo conferencia entre dois pontos, o que convencionamos chamar na literatura técnica de circuitos ponto-a-ponto.

No ano 2000 as empresas telefônicas brasileiras iniciaram um processo de vendas baseadas em um produto de grande valor agregado à telefonia fixa, chamado ADLS, o sistema era baseado em uma tecnologia avançada, e de algoritmos criativos, com o intuito de se obter a compressão das informações sobre o par de linhas metálicas.

Em julho 2002, o ITU finalizou as recomendações G.992.3 (G.dmt.bis) e G.992.4, estes dois novos padrões para a tecnologia ADSL, foram denominados pelos seus criadores de ADSL2, como alusão a uma versão melhorada da tecnologia ADSL. Em janeiro 2003, ao mesmo tempo em que o número de usuários da primeira geração do padrão ADSL passou da marca de 30 milhões, a recomendação G.992.5 foi oficialmente adicionada à família ADSL2 com o nome ADSL2plus, ou simplesmente ADSL2+ como é usualmente conhecido. Diversas

outras características e melhorias foram incorporadas também na forma de novos anexos e recomendações.

As prestadoras de serviço de acesso banda larga e os usuários desempenharam um papel importante na conclusão das recomendações do padrão ADSL2, já que os resultados obtidos a partir do uso prático do ADSL puderam ser incorporados pelo ITU nessas recomendações na forma de novas funcionalidades ou de melhorias de desempenho. Como consequência, o ADSL2 proporciona um uso mais amigável para os assinantes, e torna-se mais rentável para os prestadores de serviço, mantendo e ampliando o grande sucesso do ADSL.

O ADSL2 (ITU G.992.3 e G.992.4) apresenta novas características e funcionalidades que têm o objetivo de melhorar o desempenho e a interoperabilidade, e adiciona suporte para novas aplicações e serviços, e para novas formas de implementação dos serviços. Entre as novas características oferecidas podem-se destacar as melhorias na taxa de bits e na distância do enlace, o ajuste adaptativo de taxa de bits, as novas facilidades de diagnóstico, e a nova modalidade *stand-by* para o controle do uso de energia. O ADSL2+ (ITU G.992.5) duplica a largura da banda usada para a transmissão de dados *downstream*, duplicando também as taxas de bits *downstream*, e conseguindo taxas de até 20 Mbit/s em linhas telefônicas com distâncias de até 1,5 km. As soluções que usam o ADSL2+ geralmente são do tipo multimodo, ou seja, permitem operação conjunta, na mesma infra-estrutura, do ADSL, do ADSL2, e também dos *Chipsets* ADSL2+.

O ADSL2+ permite que os prestadores de serviço possam evoluir suas redes para suportar serviços avançados, tais como as aplicações de Vídeo, de forma flexível, permitindo que uma única solução possa atender tanto os serviços em enlaces curtos (mais próximos da estação telefônica) como os serviços em enlaces longos (mais distantes da estação telefônica). Além disso, o ADSL2+ inclui todos os benefícios das características e do desempenho de ADSL2, ao mesmo tempo em que pode operar com os equipamentos legados da versão ADSL.

Desta forma, os prestadores de serviço podem oferecer a nova tecnologia ADSL2+ nas suas redes sem ter que atualizar os equipamentos existentes, permitindo uma transição gradual para serviços avançados. Sendo assim esta tecnologia se mostra atualmente como o futuro das linhas banda larga no mundo.

ANEXO V

Regressão Linear Múltipla Utilizando o *Software SPSS*

O *package* estatístico SPSS para *windows* é um sistema de análises estatísticas e manuseamento de dados, num ambiente gráfico, em que a utilização mais freqüente, para a maioria das análises, se resume a seleção das respectivas opções de menus e caixas de diálogos.

Editor de Dados

O editor de dados do SPSS é um programa do tipo planilha de calculo que permite criar ou editar dados. Muitas das características do editor são similares aos das planilhas *excel*. Cada célula contém um valor ou uma observação de um indivíduo, em relação a uma determinada característica ou variável. As células podem apenas conter valores numéricos ou alfanuméricos, não sendo permitido definir formulas em células.

Importação de Base de Dados Externos

O programa dispõe de uma rotina de importação de Base de Dados criados em outras aplicações, nomeadamente planilhas de calculo, banco de dados, bases ASCII. Nessa pesquisa importaram-se os dados dos indicadores contidos em uma planilha *Excel*, conforme mostrado na figura AV.1 a seguir:

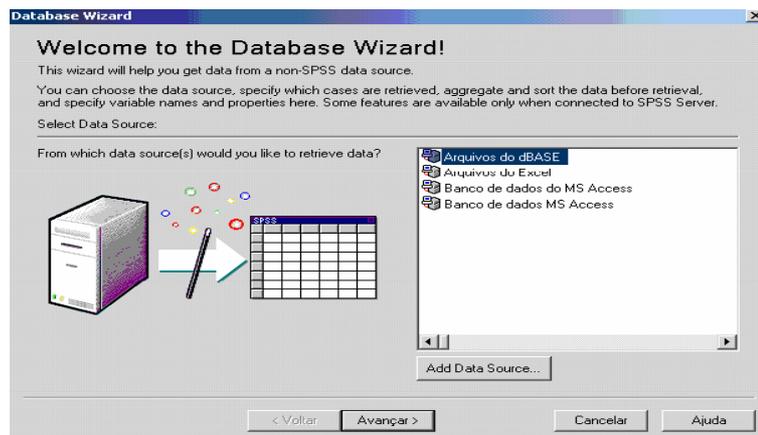


Figura AV.1 Importação de Dados.

Fonte: SPSS adaptado pelo autor.

Os arquivos em *Excel* são constituídos por diversas planilhas, a rotina de importação é assegurada pelo protocolo *Open Database Connectivity (ODBC)*, que é um método padrão de importação de partilha de dados entre bases de dados e outros programa. Os controladores *ODBC* utilizam a linguagem *SQL (Structured Query Language)* padrão para acessar dados de origens exteriores.

Após a importação concluída utiliza-se dos passos conforme mostrada na figura AV.2, para estimar uma equação linear múltipla:

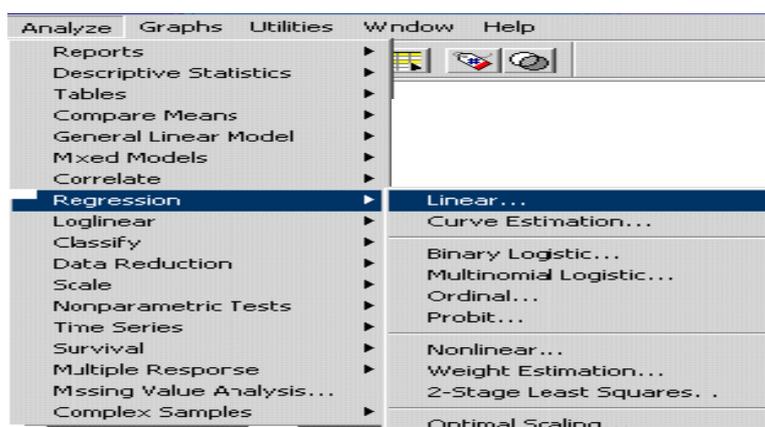


Figura AV.2: Regressão linear múltipla por meio do *SPSS*.

Fonte: *SPSS* adaptado pelo autor.

A seguir se seleciona a opção de regressão linear conforme mostrado na caixa de dialogo da figura AV.3:

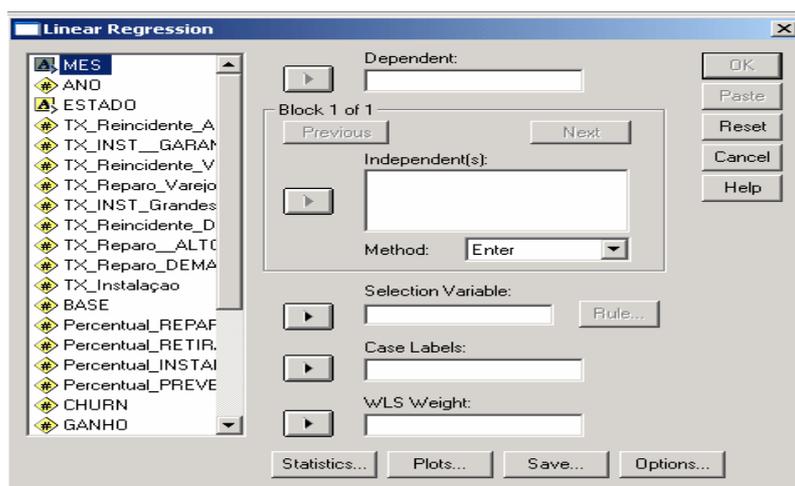


Figura AV.3: Representação da seleção das variáveis.

Fonte: *SPSS* adaptado pelo autor.

No campo de variáveis independentes estão incluídas todas as variáveis explanatórias. O método de cálculo pode ser selecionado entre os métodos sequenciais (*Stepwise, Backward, Forward*) ou combinacionais (*Enter*).

Ao selecionar o botão *Statistics*, se faz as escolhas das opções de descrições das variáveis bem como os testes estatísticos pertinentes a análise conforme mostra a figura AV.4:

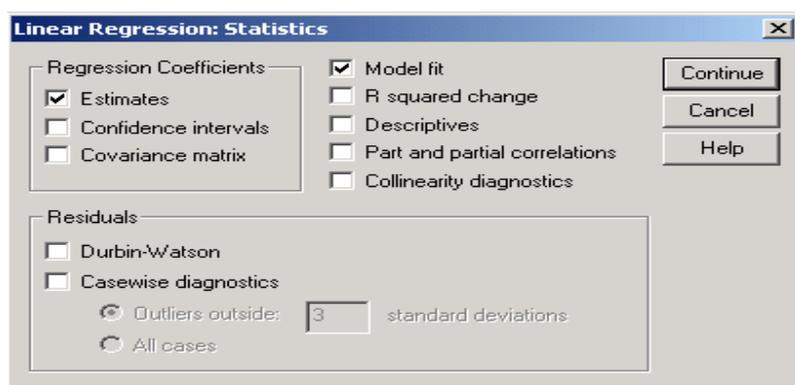


Figura AV.4: Seleção dos dados estatísticos.

Fonte: SPSS adaptado pelo autor.

Por meio da seleção do botão *Plots*, se faz a escolha dos gráficos que devem ser gerados pelo *software SPSS*, conforme mostrado a seguir:

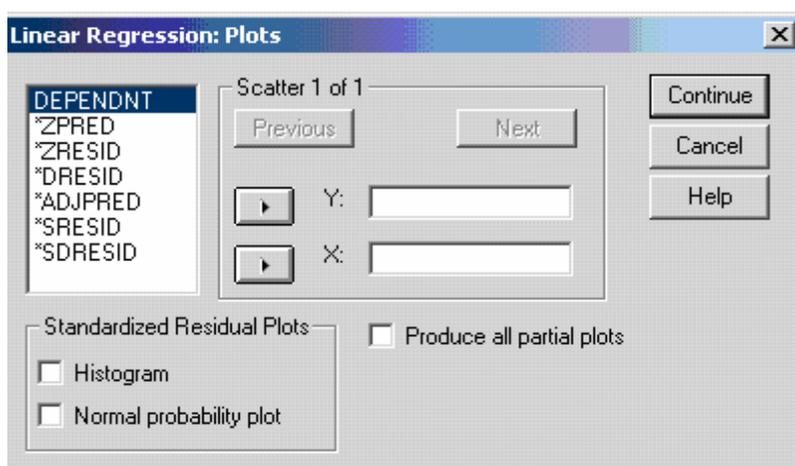


Figura AV.5: Seleção dos gráficos estatísticos.

Fonte: SPSS adaptado pelo autor.

Ao se selecionar o botão *Options*, se faz a escolha da probabilidade “F” que deve ser usado pelo *software SPSS*.

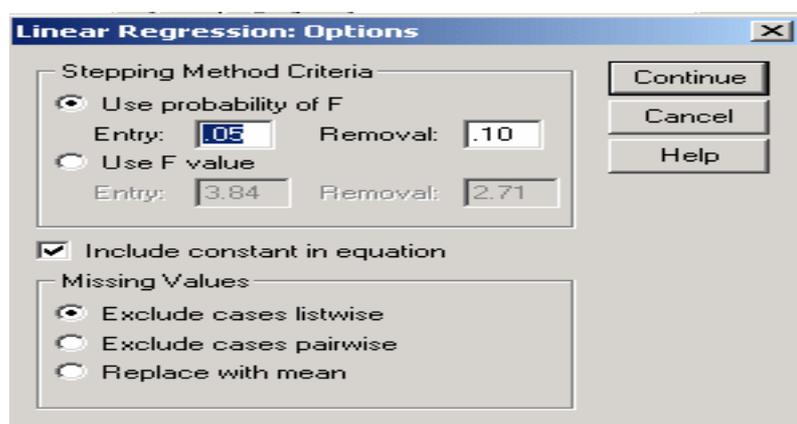


Figura AV.6: Escolha da probabilidade “F”.

Fonte: *SPSS* adaptado pelo autor.

Por fim tecla-se o botão OK conforme mostrada na figura AV.3 para se encerrar a pesquisa encontrando-se as tabelas e gráficos apresentados nesta dissertação.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)