

**UNIVERSIDADE FEDERAL DE SANTA MARIA**  
**CENTRO DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**GERÊNCIA PRÓ-ATIVA DE INCIDENTES DE TI ATRAVÉS DA  
QUANTIFICAÇÃO DE DADOS E DA UTILIZAÇÃO DE MÉTODOS  
ESTATÍSTICOS MULTIVARIADOS**

**Dissertação de Mestrado**

**Érico Marcelo Hoff do Amaral**

**Santa Maria, RS, Brasil**

**2010**

**GERÊNCIA PRÓ-ATIVA DE INCIDENTES DE TI  
ATRAVÉS DA QUANTIFICAÇÃO DE DADOS E DA  
UTILIZAÇÃO DE MÉTODOS ESTATÍSTICOS  
MULTIVARIADOS**

por

**Érico Marcelo Hoff do Amaral**

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Área de Concentração em Qualidade e Produtividade, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Engenharia de Produção.**

**Orientador Prof. Dr. Raul Ceretta Nunes**  
**Co-Orientador Prof. Dr. Adriano Mendonça Souza**

**Santa Maria, RS, Brasil**  
**2010**



**Universidade Federal de Santa Maria  
Centro de Tecnologia  
Programa de Pós-Graduação em Engenharia de Produção**

A Comissão Examinadora, abaixo assinada,  
aprova a Dissertação de Mestrado

**GERÊNCIA PRÓ-ATIVA DE INCIDENTES DE TI ATRAVÉS DA  
QUANTIFICAÇÃO DE DADOS E DA UTILIZAÇÃO DE MÉTODOS  
ESTATÍSTICOS MULTIVARIADOS**

elaborada por  
**Érico Marcelo Hoff do Amaral**

como requisito parcial para obtenção do grau de  
**Mestre em Engenharia de Produção**

**COMISSÃO EXAMINADORA:**

---

**Dr. Raul Ceretta Nunes - UFSM**  
(Presidente/Orientador)

---

**Dr. Adriano Mendonça Souza - UFSM**  
(Co-Orientador)

---

**Dra. Roseclea Duarte Medina - UFSM**

---

**Dr. Adriano Petry - INPE**

**Santa Maria, 2010.**

*“Nada nesta vida  
é por acaso...”*

## **Agradecimentos**

A minha noiva Danuza, pelo apoio durante todo o mestrado, principalmente nas intermináveis noites, finais de semana, feriados e férias dedicados ao estudo e escrita desta dissertação. Pelo carinho e compreensão despendidos nos momentos mais importantes.

A toda minha Família, por sua dedicação, pelas palavras de incentivo e por serem sempre meu porto seguro.

Aos meus colegas, pelo companheirismo e amizade, em especial ao Claudio pela orientação estatística e parceria nos artigos escritos e publicados, ao Breno e Henrique pelo apoio nas atividades do CRS.

Ao professor Raul Ceretta, orientador e amigo, pela confiança depositada em meu trabalho, além da disponibilidade de me orientar nos momentos necessários. Também agradeço ao meu co-orientador professor Adriano Mendonça, por sua boa vontade e pela oportunidade de repassar os conhecimentos necessários sobre o tema desenvolvido neste trabalho.

Aos integrantes do CRS e INPE, em especial ao Koiti, grande mestre e amigo, sempre disposto a auxiliar e repassar seu conhecimento.

## SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO .....</b>	<b>18</b>
1.1 Motivação .....	19
1.2 Problema de Pesquisa .....	19
1.3 Objetivo Geral .....	20
1.4 Objetivos Específicos .....	20
1.5 Principais Contribuições.....	21
1.6 Organização da Dissertação.....	21
<b>CAPÍTULO 2 - REFERENCIAL TEÓRICO: FUNDAMENTAÇÃO CONCEITUAL SOBRE GESTÃO DE INCIDENTES.....</b>	<b>22</b>
<b>2.1 Incidente.....</b>	<b>22</b>
<b>2.2 Gerência de Incidentes .....</b>	<b>25</b>
2.2.1 Descrição do Processo de Gerência de incidentes .....	29
<b>2.3 Níveis de Serviço .....</b>	<b>34</b>
2.3.1 Principais funções do SLA.....	36
2.3.2 Tipos de SLA .....	38
2.3.2.1 SLAs na organização .....	39
2.3.2.2 SLAs Externos .....	39
2.3.2.1 SLAs Internos .....	39
2.3.3 SLM - Gerenciamento de Nível de Serviços .....	40
<b>2.4 Apoio a Decisão .....</b>	<b>42</b>
2.4.1 Ferramentas de Apoio à Decisão .....	43
<b>2.5 Considerações Parciais .....</b>	<b>50</b>
<b>CAPÍTULO 3 - REFERENCIAL TEÓRICO: FUNDAMENTAÇÃO CONCEITUAL SOBRE ANÁLISE ESTATÍSTICA MULTIVARIADA .....</b>	<b>51</b>
<b>3.1 Análise Estatística .....</b>	<b>51</b>

3.1.1	Ferramentas para Análise Estatística Descritiva.....	52
<b>3.2</b>	<b>Análise Multivariada .....</b>	<b>53</b>
3.2.1	Termos e Conceitos Básicos .....	55
3.2.1.1	Dados Multivariados.....	55
3.2.1.2	Padronização, Escalonamento e Medidas de Similaridade .....	56
3.2.1.3	Técnicas de Análise Multivariadas .....	57
3.2.2	Análise de Agrupamentos .....	58
3.2.3	Análise de Fatorial .....	60
3.2.4	Análise de Componentes Principais .....	61
<b>3.3</b>	<b>R – Linguagem para Cálculos Estatísticos .....</b>	<b>62</b>
3.3.1	Ambiente de Desenvolvimento.....	63
<b>3.4</b>	<b>Considerações Parciais .....</b>	<b>66</b>
<b>CAPITULO 4 – PROJETO DO PORTAL CORPORATIVO E DA FERRAMENTA PARA O SUPORTE A DECISÃO .....</b>		<b>67</b>
<b>4.1</b>	<b>Intranet X Portal Corporativo.....</b>	<b>67</b>
<b>4.2</b>	<b>Portal Corporativo.....</b>	<b>69</b>
4.2.1	Tecnologias e funcionalidades do Portal Corporativo .....	69
4.2.1.1	Controle de usuários .....	71
4.2.1.2	Controle de materiais .....	72
4.2.1.3	Controle de rede.....	72
4.2.1.4	Informativos gerais .....	72
4.2.1.5	Controle de usuários online e colaboradores .....	73
4.2.1.6	Base de Conhecimento.....	73
<b>4.3</b>	<b>Projeto do Portal Corporativo.....</b>	<b>73</b>
<b>4.4</b>	<b>Proposta do Service Desk via Portal Corporativo (SDvPC).....</b>	<b>75</b>



4.4.1	Modulos do SDvPC .....	76
4.4.2	Funcionalidades e requisitos do SDvPC .....	77
4.4.3	Quantificação das informações .....	78
4.4.4	Categorização.....	80
<b>4.5</b>	<b>Análise estatística .....</b>	<b>81</b>
<b>4.6</b>	<b>Considerações Parciais .....</b>	<b>83</b>
<b>CAPITULO 5 – IMPLEMENTAÇÃO, TESTES E DISCUSSÕES.....</b>		<b>86</b>
<b>5.1</b>	<b>Laboratório de Pesquisa – CRS/INPE-MCT .....</b>	<b>86</b>
5.1.1	Monitoramento de Incidentes no âmbito do CRS.....	88
<b>5.2</b>	<b>Projeto do Portal Corporativo e do <i>Service Desk</i> .....</b>	<b>89</b>
5.2.1	Análise do Portal Corporativo e do <i>Service Desk</i> .....	90
<b>5.3</b>	<b><i>Framework Django</i> .....</b>	<b>96</b>
5.3.1	Informações Técnicas do projeto com Django .....	97
<b>5.4</b>	<b>Implementação do SDvPC .....</b>	<b>98</b>
5.4.1	Interface do Usuário.....	100
5.4.2	Interface do Administrador .....	101
5.4.3	Integração – Python, Java e R.....	106
<b>5.5</b>	<b>Resultados e Discussões .....</b>	<b>108</b>
5.5.1	Descrição dos Resultados .....	109
5.5.2	Análise e Validação Comprobatória dos Resultados.....	113
5.5.3	Obtenção da Pró-atividade.....	124
<b>5.5</b>	<b>Considerações Parciais .....</b>	<b>125</b>
<b>CAPITULO 6 - CONCLUSÕES.....</b>		<b>126</b>
<b>6.1</b>	<b>Trabalhos Futuros .....</b>	<b>128</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>		<b>130</b>

## LISTA DE TABELAS

Tabela 1- Distribuição de Frequências do Perfil do Usuário.....	109
Tabela 2- Exemplos de Casos e Prioridades.....	112
Tabela 3- Comparação entre o Tempo Previsto de Atendimento e Tempo Real de Atendimento .....	115

## LISTA DE FIGURAS

Figura 1 - Relação erro, problema e incidente.....	24
Figura 2 - A estrutura das Publicações do ITIL.....	27
Figura 3 - Processos de Suporte a Serviços.....	28
Figura 4 - Processos de Entradas e Saídas do ITIL.....	30
Figura 5 - Processos envolvidos na Gerência de Incidentes.....	31
Figura 6 - Hierarquia de níveis de serviços.....	37
Figura 7 - Organograma dos níveis funcionais de um SI.....	44
Figura 8 - Componentes de um SAD.....	46
Figura 9 – Descrição da Técnica de Representação <i>Box-plot</i> .....	53
Figura 10 - Console de programação em Linguagem R.....	65
Figura 11 - Console de saída em Linguagem R.....	65
Figura 12 - Transição do acesso a intranet para os portais corporativos.....	68
Figura 13 - Modelo do Ciclo de Vida Clássico (Cascata).....	74
Figura 14 - Inserção do Portal Corporativo na estrutura do CRS.....	75
Figura 15 - Etapas do funcionamento do SDvPC.....	78
Figura 16 - Matriz de Prioridade – “Impacto vs. Urgência”.....	79
Figura 17 - Tabela de Prioridades.....	79
Figura 18 - nível de profundidade “hierarquia” ou funcionalidade a qual deve atender um determinado incidente.....	80
Figura 19 - O FMEA é o Resultado do Relacionamento entre as Funções: Ferramenta, Procedimento e Diário.....	81
Figura 20 – Amostra de dados disponível para a análise estatística.....	82
Figura 21 – Organização da Inferência estatística.....	83
Figura 22 – Organização do capítulo Implementação, Testes e Discussões.....	86
Figura 23 – Console Equipamentos, Salas de Equipamentos e Armários de Telecomunicações.....	87
Figura 24 – Caso de Uso – acesso ao Portal, cadastro e atendimento ao usuário.....	91
Figura 25 – Caso de Uso – execução das aplicações em Java e R.....	92
Figura 26 – Diagrama de Sequencia – registro de incidente (ticket) no <i>Service Desk</i> via Portal.....	93
Figura 27 – Diagrama de Sequencia – atendimento aos chamados dos usuários.....	93
Figura 28 – Diagrama de Sequencia – execução da análise estatística.....	94

Figura 29 – Diagrama de Sequencia – análise estatística.....	94
Figura 30 – Diagrama Entidade-Relacionamento – estrutura da base de dados do sistema .....	96
Figura 31 – Trecho de código com a estrutura ORM.....	97
Figura 32 – <i>View</i> com os últimos informativos.....	98
Figura 33 – SDvPC – Modulos e composição do sistema.....	99
Figura 34 – Tela de cadastro de usuários - SDvPC .....	100
Figura 35 – Interface para a administração dos incidentes reportados ao SD - SDvPC.....	101
Figura 36 – Janela com informações detalhadas sobre os tickets abertos pelos usuários .....	104
Figura 37 – Área de acesso a ferramenta para a coleta de dados e análise estatística.....	106
Figura 38 – Arquivo com amostra de dados para análise multivariada .....	107
Figura 39 – Distribuição das variáveis no círculo de correlção .....	111
Figura 40 – Gráfico da distribuição da nuvem de pontos.....	112
Figura 41 – Dendograma baseado na mostra de dados do SDvPC .....	114
Figura 42 – Histograma de comparação entre TPA e TA .....	115
Figura 43 – <i>Box-plot</i> das variáveis TPA e TA .....	116
Figura 44 – <i>Box-plot</i> comparativo entre as Categorias de Incidentes .....	117
Figura 45 – <i>Box-plot</i> para análise do índice de reportes por Perfil de Usuário.....	118
Figura 46 – <i>Box-plot</i> para análise do índice de reportes por Prioridade dos Tickets reportados .....	119
Figura 47 – Diagrama de Pareto – Análise das categorias com maior número de incidentes	120
Figura 48 – Diagrama de Pareto – Análise dos perfis de usuários .....	120
Figura 49 – Diagrama de Pareto – Verificação dos Reportes por Prioridade dos Tickets .....	121
Figura 50 – Gráfico de Controle – Comparação entre os tempos de atendimento.....	122
Figura 51 – Gráfico de Controle – Análise dos reportes por categoria dos incidentes .....	122
Figura 52 – Gráfico de Controle – Análise dos reportes considerando o Perfil dos usuários	123
Figura 53 – Gráfico de Controle – Verificação dos reportes considerando a prioridade de atendimento .....	123

## LISTA DE SIGLAS

ACP	-	Análise de Componentes Principais
AW	-	Ataques à Web
AF	-	Análise Fatorial
AGC	-	Algoritmo Genético Construtivo
B2E	-	<i>Business to Employee</i>
BD	-	Banco de Dados
BDGC	-	Base de Dados do Gerenciamento da Configuração
BI	-	<i>Business Intelligence</i>
BRAMS	-	<i>Brazilian Atmospheric Modeling System</i>
CEP	-	Controle Estatístico de Processos
CPTEC	-	Centro de Previsão de Tempo e Clima
CRECTEALC	-	Centro regional de Educação em Ciência e Tecnologia para América Latina e o Caribe
CRS	-	Centro Regional Sul de Pesquisas Espaciais
CSV	-	<i>Comma-separated values</i>
DDM	-	Diálogos, Dados e Modelo
DOS	-	<i>Denial of Service</i>
DRY	-	<i>Don't Repeat Yourself</i>
DSS	-	<i>Decision Support Systems</i>
EAD	-	Ensino à Distância
EAPS	-	<i>Ethernet Automatic Protection Switching</i>
ESRP	-	<i>Extreme Standby Router Protocol</i>
FMEA	-	<i>Failure Mode and Effect Analysis</i>
FINCLAS	-	Classificação Financeira
FIFO	-	<i>First In, First Out</i>
FTP	-	<i>File Transfer Protocol</i>
GED	-	Gerenciamento Eletrônico de Documentos
GRUMA	-	Grupo de Modelagem Atmosférica de Santa Maria
HTTPS	-	<i>HyperText Transfer Protocol Secure</i>
INPE	-	Instituto Nacional de Pesquisas Espaciais
IRIS	-	<i>Interactive Robustness analysis and parameters' Inference for Multicriteria Sorting</i>

ITIL	-	<i>Information Technology Infrastructure Library</i>
KMO	-	Kaiser-Meyer-Olkin
LACESM	-	Laboratório de Ciências Espaciais de Santa Maria
LES	-	<i>Large Eddy Simulation</i>
MANCOVA	-	Análise Multivariada de Covariância
MANOVA	-	Análise Multivariada de Variância
MCT	-	Ministério da Ciência e Tecnologia
OGC	-	<i>Office of Government Commerce</i>
ORM	-	<i>Object-relational mapping</i>
P2P	-	<i>Peer-to-Peer</i>
PC	-	Portal Corporativo
RAM	-	<i>Random Access Memory</i>
RFC	-	<i>Requests For Change</i>
RMD	-	Requisição de Mudança
ROI	-	Retorno sobre Investimento
SD	-	<i>Service Desk</i>
SDvPC	-	<i>Service Desk</i> via Portal Corporativo
SI	-	Segurança da Informação
SLA	-	<i>Service Level Agreement</i>
SLM	-	<i>Service Level Management</i>
SLR	-	Requisitos dos Níveis de Serviço
SPOC	-	<i>Single Point of Contact</i>
SPSS	-	<i>Statistical Package for the Social Sciences</i>
SAD	-	Sistema de Apoio a Decisão
SGBD	-	Sistema Gerenciador de Banco de Dados
SGBM	-	Sistema Gerenciador do Banco de Modelos
SIP	-	<i>Session Initiation Protocol</i>
SQL	-	<i>Structured Query Language</i>
SSL	-	<i>Secure Socket Layer</i>
STP	-	<i>Spanning Tree Protocol</i>
TI	-	Tecnologia da Informação
VIP	-	<i>Variable Interdependent Parameters</i>
VoIP	-	<i>Voice over Internet Protocol</i>

## **RESUMO**

### **Dissertação de Mestrado Programa de Pós-Graduação em Engenharia de Produção Universidade Federal de Santa Maria**

#### **GERÊNCIA PRÓ-ATIVA DE INCIDENTES DE SEGURANÇA ATRAVÉS DA QUANTIFICAÇÃO DE DADOS E DA UTILIZAÇÃO DE MÉTODOS ESTATÍSTICOS MULTIVARIADOS**

AUTOR: ÉRICO MARCELO HOFF DO AMARAL

ORIENTADOR: DR. RAUL CERETTA NUNES

CO-ORIENTADOR: DR. ADRIANO MENDONÇA SOUZA

Data e Local da Defesa: 08 de Março de 2010, Santa Maria

É reconhecido que no cenário atual das organizações a informação tornou-se um ativo de suma importância. Analisando esta tendência, percebe-se que, da mesma forma, a evolução e o dinamismo das ameaças e incidentes de segurança sobre este ativo é um fato incontestável. É fundamental que os responsáveis pela gestão organizacional das empresas envidem esforços para monitorar os incidentes relacionados à área de Tecnologia da Informação (TI), atuando de forma pontual sobre esses problemas, tratando-os de maneira pró-ativa e inteligente, possibilitando assim a tomada de decisões precisas e rápidas, objetivando a garantia de continuidade do negócio.

Esta dissertação apresenta uma ferramenta de gestão de incidentes relacionados a serviços e sistemas de TI, denominada SDvPC (*Service Desk* via Portal Corporativo), que integra um sistema de *Service Desk* a um Portal Corporativo e disponibiliza, de forma centralizada, procedimentos formais de reporte e escalonamento dos problemas identificados pelos usuários em uma organização. O SDvPC auxilia para que as fragilidades da organização sejam notificadas de maneira ágil e simples, tão logo quanto possível, e permite a gestão pró-ativa de incidentes na área de TI ao explorar a quantificação dos dados qualitativos coletados no *Service Desk* e o agrupamento dos incidentes através da análise multivariada. Como resultado, esta ferramenta possibilita o conhecimento tácito da falhas, deficiências e

dificuldades agregadas aos serviços e sistemas de TI em um ambiente organizacional, possibilitando a visão e o planejamento estratégico sobre as atividades da área de suporte.

**Palavras-chave:** Gestão de Incidentes, Estatística Multivariada, Apoio a Decisão, Portal Corporativo.



## **ABSTRACT**

*Master's Dissertation  
Program of Master Degree in Production Engineering  
Federal University of Santa Maria*

### **GERÊNCIA PRÓ-ACTIVA DE INCIDENTES DE SEGURANÇA ATRAVÉS DA QUANTIFICAÇÃO DE DADOS E DA UTILIZAÇÃO DE MÉTODOS ESTATÍSTICOS MULTIVARIADOS**

*AUTHOR: ÉRICO MARCELLO HOFF DO AMARAL*

*ADVISER: RAUL CERETTA NUNES, DR.*

*CO-ADVISER: ADRIANO MENDONÇA SOUZA, DR.*

*Date and Local: March 08, 2010, Santa Maria*

*Its recognized that in the organizations the information has become an asset of paramount importance. Analyzing this trend, one realizes that, in the same way, the evolution and dynamics of the threats and security incidents on this asset is an indisputable fact. Moreover, its essential that those responsible for organizational management in the companies strive to monitor the incidents related to the area of Information Technology (IT), acting in a timely manner about these issues, treating them in a proactive and intelligent, way allowing the accurate and rapid decisions, aimed at ensuring continuity of business.*

*This paper presents a tool for incident management related to it service and systems, called SDvPC (Service Desk via Portal Corporativo), which includes a Service Desk to a corporate portal and provides in a centralized way, formal procedures for reporting and scheduling the problems identified by users in an organization. The tool helps to ensure that the weaknesses of the organization are reported in a quickly and simply, as soon as possible, and allows proactive management of incidents in the area of IT to explore the quantification of qualitative data collected in the Service Desk and the grouping of incidents using multivariate analysis. As a result, SDvPC allows the tacit knowledge of the failures, shortcomings and difficulties attached to the IT services and systems in an organizational environment, providing vision and strategic planning on the activities of the support area.*

***Keynotes: Incident Management, Multivariate Statistics, Decision Support, Corporate Portal.***

## **CAPÍTULO 1**

### **INTRODUÇÃO**

A forte concorrência nas diversas áreas de negócio, aliada às conturbações econômicas enfrentadas pelas empresas nos dias atuais, forçam estas a serem mais criativas na resolução de seus problemas, tornando clara a necessidade de melhorar seus procedimentos de negócio para obter uma maior qualidade agregada aos produtos e a satisfação de seu público-alvo (ZSCHORNACK, 2003). Sob essas circunstâncias a informação passa a ter maior valor para a organização, seja ela privada ou governamental, pois possibilita o conhecimento, comunicação, participação e maior esclarecimento. Segundo Davis (1974), a informação é um dado processado em uma forma significativa para o receptor, com valor real ou percebido em uma decisão atual ou futura. Com base nesta colocação percebe-se que a informação não se restringe apenas a dados meramente coletados, mas sim a dados ordenados, organizados e fundamentados de maneira que se tornem úteis, infinitamente reutilizáveis e não se deteriorem. A importância da informação e o seu valor determinado exclusivamente pelo usuário, a caracteriza como um elemento importante na tomada de decisões pertinentes, de melhor qualidade e no momento adequado (GILKA; ADIEL, 2002). O tratamento correto das informações pode ser de extrema utilidade para as organizações, considerando o grau de valor agregado a elas.

Para se obter uma maior confiabilidade na tomada de decisões, é necessário estabelecer uma maneira de tratar as informações, de forma que o resultado gerado pelo agrupamento e estudo dessas variáveis sirva de subsídio concreto para a definição das estratégias da organização. Segundo Beuren (1998), a informação ocupa um espaço importante no apoio à formulação de estratégias de decisão, como também no acompanhamento das operações empresariais. Inteirando-se da importância da informação para o sucesso das organizações, percebe-se uma lacuna que deve ser revista e priorizada: a segurança e como os incidentes que agregam riscos às informações devem ser tratados. Para garantir a continuidade do negócio com segurança, um dos pontos-chaves é implementar e manter um sistema de gestão de incidentes (ITIL, 2008).

Na gestão de incidentes a organização deve envidar esforços para monitorar os incidentes relacionados à Tecnologia da Informação (TI) e tomar decisões precisas e rápidas. Em um processo de gestão de incidentes, a notificação de fragilidades e eventos de segurança da informação exige procedimentos formais de registro e escalonamento, bem como a notificação dos incidentes com brevidade ao ponto de contato da organização (NBR ISO/IEC

17799, 2005). Apenas com uma coleta de dados eficiente é possível alcançar uma Gestão Pró-ativa, que neste trabalho esta diretamente relacionada à manipulação do conhecimento, extraído das informações sobre incidentes, e a viabilidade da gerência de TI se beneficiar deste conhecimento, para avaliar suas atividades e planejar de forma estratégica seus procedimentos. A adoção desta postura permite uma avaliação eficiente dos processos bem como agendar tarefas futuras e manutenções, evitando assim a degradação dos serviços de TI.

## **1.1 Motivação**

Percebe-se a necessidade de um ambiente computacional capaz de concentrar todas as informações pertinentes a incidentes de segurança reportados no âmbito da organização. Sabe-se que a abrangência destas informações e seu variado número de fontes criam um problema, a dificuldade de interpretação. Existem, segundo a literatura, inúmeras maneiras de se tratar esse tipo de problema. Uma das soluções está calcada na construção de modelos matemáticos que contemplem resultados previsíveis. Porém esta tarefa possui um nível de complexidade diretamente relacionado ao número de variáveis a serem analisadas, que neste caso são muitas. A análise do contexto acima, remete a utilização de métodos estatísticos multivariados, que procuram contemplar as inúmeras variáveis de forma integrada, onde as inter-relações são exploradas em sua máxima profundidade e as soluções para os problemas são mais consistentes e úteis (FERREIRA, 2008).

A relevância científica desta pesquisa recai sobre o estudo e implementação de um Serviço de Atendimento (*Service Desk* ou SD) via Portal Corporativo, que servirá como base de dados para registro de incidentes de segurança, e este, como laboratório para análise estatística multivariada de incidentes, visando obter um método pró-ativo voltado ao tratamento destes problemas, aumentando a disponibilidade e competitividade das organizações.

## **1.2 Problema de Pesquisa**

Segundo o CERT.br, um incidente de segurança pode ser definido como qualquer evento adverso, confirmado ou sob suspeita, relacionado à segurança em sistemas computacionais, de informação ou redes de computadores. A escala desses eventos é muito abrangente dentro das organizações, o que demanda um alto custo de tempo e recursos

humanos necessários para classificá-los, categorizá-los e enumerá-los de forma a alcançar seu efetivo tratamento. Este trabalho busca encontrar uma solução viável e factível para o seguinte problema de pesquisa:

É possível implementar uma solução para gerência de incidentes de segurança pró-ativa, utilizando métodos estatísticos multivariados, e desta maneira restabelecer todos os serviços de TI em uma determinada organização a fim de reduzir o impacto destes sobre o negócio?

### **1.3 Objetivo Geral**

O objetivo principal deste trabalho é estudar e desenvolver uma ferramenta para coletar registros de incidentes de segurança, e a partir do repositório de dados gerado, tratar via técnicas de análise multivariada as informações armazenadas para prever o futuro, criando uma maneira pró-ativa para detectar as ameaças, vulnerabilidades e eventos críticos no ambiente das organizações.

### **1.4 Objetivos Específicos**

Esta dissertação teve como objetivos específicos um conjunto de atividades, que serviram de aporte para alcançar o objetivo geral. Dentre essas atividades destacaram-se:

- Realizar um estudo sobre Incidentes de Segurança e as principais metodologias de Governança de TI para o tratamento destes problemas;
- Identificar técnicas estatísticas multivariadas que permitam analisar incidentes em uma organização;
- Analisar e implementar uma ferramenta para coleta centralizada dos dados relacionados a incidentes de segurança;
- Desenvolver uma técnica para quantificação das informações qualitativas de incidentes coletadas;
- Validar os resultados obtidos, a fim de reconhecer o nível de pró-atividade e as melhorias no processo de Tomada de Decisão.

## 1.5 Principais Contribuições

Com o intuito de reconhecer e tornar pró-ativo o processo decisório na área de TI das organizações, este trabalho se propôs a estudar o tratamento dos incidentes de segurança de uma forma multidisciplinar, apresentando assim como importantes contribuições:

- O desenvolvimento de um mecanismo de coleta centralizada de dados descritivos dos Incidentes de Segurança, o SDvPC (*Service Desk* via Portal Corporativo) ;
- A elaboração de uma metodologia válida e eficaz para a quantificação das informações sobre Incidentes, armazenadas no *Service Desk*;
- O estudo e implementação de uma ferramenta para análise multivariada dos dados quantificados, a qual é capaz de extrair o conhecimento tácito contido nestas informações, viabilizando a Gerência Pró-ativa dos incidentes no âmbito de uma organização, e se tornando peça fundamental para o apoio a decisão.

## 1.6 Organização da Dissertação

Esta dissertação está estruturada como segue. O capítulo 2 descreve o referencial teórico sobre como tratar e gerir incidentes de segurança, além de mostrar como essas atividades podem auxiliar no apoio à decisão no âmbito das organizações. O capítulo 3 apresenta a revisão de literatura sobre os métodos estatísticos, definindo como as técnicas de análise multivariada podem ser aplicadas sobre as amostras de reportes de incidentes, e assim realizar a predição a partir da base de conhecimento gerada. O capítulo 4 descreve o projeto de uma ferramenta de SD sobre um Portal Corporativo e o ambiente para o qual a mesma foi implementada. O capítulo 5 descreve os testes e experimentos realizados, a avaliação e validação dos resultados obtidos. Finalizando, o Capítulo 6 descreve as conclusões e trabalhos futuros.

## CAPÍTULO 2

### REFERENCIAL TEÓRICO: FUNDAMENTAÇÃO CONCEITUAL SOBRE GESTÃO DE INCIDENTES

Este capítulo apresenta na seção 2.1 a fundamentação teórica dos principais aspectos que podem caracterizar um incidente de segurança da informação. Como é feita a gerência desses incidentes é mostrada na seção 2.2.. Na seção 2.3 são apresentados os Acordos de Nível de Serviço (SLA), contratos firmados dentro da organização que garantem a obtenção de níveis adequados, a serem respeitados, no atendimento de cada incidente reportado. Por último, na seção 2.4 é demonstrado como estas práticas podem servir de subsídios para a tomada de decisão em um ambiente organizacional.

#### 2.1 INCIDENTE

Um incidente é qualquer evento que não faz parte do funcionamento normal de um serviço que causa, ou pode causar, a sua interrupção ou uma redução da qualidade (BOM, 2006). Também pode ser considerado um incidente, a exploração de vulnerabilidades em sistemas de informação ou em componentes de aplicações por meio de ataques de especialistas (*hackers*), o que pode levar as organizações a perdas incalculáveis sob o ponto de vista operacional, financeiro e mercadológico (PAIVA, *et al*, 2007). Em um nível gerencial um incidente é qualquer acontecimento que altere os níveis de serviço estipulados nos acordos de nível de serviço (SLA), ou ainda qualquer acontecimento que não faça parte da operação normal dos sistemas e ambientes estabelecidos (OGC, 2001). Outra linha, mais radical, define incidente de segurança como qualquer fato ou evento que afete a segurança de recursos humanos ou de materiais no âmbito da organização. Sob qualquer um dos pontos de vista citados, nota-se que a ocorrência de um problema ou incidente pode ser definido por um conjunto de fatores diferentes que afetam vários setores em uma organização. Porém todas as definições apresentam um ponto em comum: o impacto causado pelo incidente reflete diretamente em aspectos financeiros e orçamentários das empresas.

Ao se tratar de incidentes é conveniente reconhecer os principais termos utilizados para identificar e relacionar os eventos que afetam e colocam em risco a integridade, confidencialidade e disponibilidade dos ativos (ITIL, 2008). Pode-se citar como parte desses:

- Vulnerabilidade: é um ponto fraco em um processo, na gestão ou na tecnologia, que se possa explorar (ou executar) para comprometer a segurança da TI;
- Ameaça: é a possibilidade de que uma fonte de ameaça exerça (acione acidentalmente ou explore intencionalmente) uma vulnerabilidade específica;
- Risco: é uma função da probabilidade de que uma determinada ameaça exerça uma vulnerabilidade específica, e o impacto resultante desse incidente adverso sobre a empresa;
- Problema: é quando a causa raiz de um ou mais incidentes é desconhecida;
- Erro conhecido: é quando a causa raiz de um problema é conhecida e já se tem uma solução de contorno;
- Solução de Retorno (*work-around*): é um método de contornar um incidente a partir de uma reparação temporária;
- Processo: é constituído por um conjunto de atividades interligadas que partem de objetivos estipulados e são executadas a fim de atingir resultados pré-definidos entre os envolvidos em um determinado processo.
- Central de atendimentos ou central de serviços (SD): ponte de contato entre os participantes dos processos da empresa (*stakeholders*), clientes, usuários, Serviços de Tecnologia de Informação e Organizações de suporte subcontratadas (*customers, users, IT services and third-party support organizations*) e o setor de Tecnologia da Informação de uma instituição. É a central que monitora os processos de solução dos incidentes ou das solicitações de serviços registradas. É o primeiro contato dos usuários com a área de TI quando existir qualquer necessidade de contato (MAGALHAES; PINHEIRO, 2007).
- Suporte de Primeiro Nível (*Initial support team*): equipe responsável pelo pronto atendimento as solicitações feitas por usuários através do SD.
- Requisição de Serviço (*Service Request*): trata-se da requisição de um serviço inédito ou alteração de um já existente. Os tipos mais comuns são as Requisições de Alteração, *Requests For Change* (RFCs) e Requisições de informações, *Requests For Information* (RFIs).

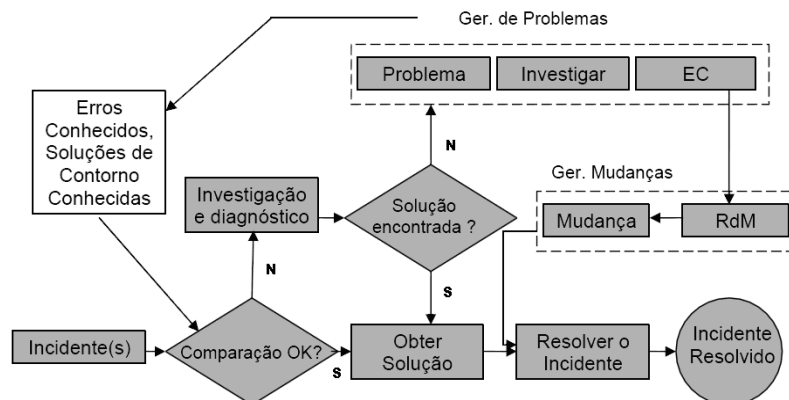


- Solução/correção permanente (*Solution/permanent fix*): ao contrário da solução de contorno (Workaround), que fornece uma solução temporária, a solução permanente (*permanent fix*), denomina a forma possível e definitiva de resolver a causa de um incidente ou problema.

O escopo de um incidente pode ser classificado de acordo com a indisponibilidade de algum dos serviços oferecidos pela área de TI ao usuário, sendo os mais comuns:

- Falha de equipamento (hardware),
- Erro de programa ou aplicativo (software),
- Solicitações de informações,
- Troca de senha,
- Novos funcionários,
- Solicitação para compra de suprimentos (cartucho de tinta, *mouse*),
- Problemas de desempenho.

É importante a distinção entre erro, problema e incidente: problema é definido como a causa com raiz desconhecida de um ou mais incidentes; erro é um problema ou incidente cuja raiz é conhecida; erro conhecido é algo para o qual já foi identificada uma solução de contorno ou permanente; e por fim incidente é a consequência ou resultado de um problema ou erro. A Figura 1 demonstra a relação existente entre problema e incidente.



**Figura 1 – Relação erro, problema e incidente.**

Fonte: *Computer Associates, 2000.*

Devido ao impacto que os diversos tipos de incidentes podem ocasionar, percebe-se a necessidade de um atendimento rápido e eficiente dos mesmos. Partindo deste princípio, com

o intuito de que os incidentes sejam tratados de forma correta, é preciso estabelecer prioridades. Estas prioridades se baseiam na gravidade do erro para o negócio e para o usuário. Mediante consulta ao problema e de acordo com as cláusulas de SLA, as prioridades são determinadas e a ordem de tratamento dos incidentes definida (BOM, 2006). Os critérios que influenciam a ordem de atendimento são os seguintes:

- **Impacto:** extensão do desvio em relação ao nível normal de serviço, medido pelo número de usuários ou de processos de negócio afetados. Os incidentes críticos são aqueles cujo grau de impacto sobre os sistemas, processos e usuários são extremos. Os incidentes para os quais o tempo de ruptura, mesmo sendo a porcentagem de usuários relativamente pequena, são excessivos e devem ser considerados graves.
- **Urgência:** a demora aceitável para o usuário ou processo de negócio, na solução do problema.

A prioridade é baseada, sobretudo a partir da análise da urgência *versus* o impacto. Quando são identificados incidentes com a mesma prioridade, o esforço exigido para a solução dos mesmos pode definir a ordem em que estes incidentes serão tratados. O próximo item Gerência de Incidentes aborda com riqueza de detalhes como são administrados e tratados os incidentes de segurança.

## 2.2 GERÊNCIA DE INCIDENTES

A “Gerência de incidentes” está descrita na biblioteca de infra-estrutura ITIL (*Information Technology Infrastructure Library*) desenvolvida pelo CCTA (*Central Computer and Telecommunications Agency*) atualmente OGC (*Office of Government Commerce*), no final dos anos 80. Esta biblioteca foi criada a partir de uma necessidade do governo britânico, para implementar uma abordagem de melhores práticas, para gerenciar a utilização eficiente e responsável dos recursos de TI, aplicáveis a organizações com necessidades técnicas e de negócio distintas (FERNANDES; ABREU, 2008). Quando se fala em boas práticas refere-se a resultados significativos que atendam a melhoria da TI, isto é, operações de empresas e aprimoramento de níveis de serviços. Entre os principais objetivos da adoção da melhores práticas da ITIL se destacam os seguintes:

- Alinhar os Serviços de TI com as necessidades atuais e futuras do negócio e seus clientes. Todos os processos da ITIL falam que a TI precisa entender os requisitos de

negócio da empresa para poder planejar e prover seus serviços para atender as expectativas.

- Melhorar a qualidade dos serviços de TI. Através de um programa de melhoria contínua deve se buscar a consistência na entrega dos serviços, atendendo as necessidades de negócio.
- Reduzir custos na provisão de serviços. Este é um dos motivos chaves que levam os gestores de TI a adotarem as melhores práticas, já existem vários casos de sucesso onde observa-se grande redução dos custos operacionais e investimentos em TI.
- Processos mais eficientes e eficazes, buscando rapidez e resultados nos processos.
- Adoção das melhores práticas, evitando reinventar a roda.

O Modelo de Processos do ITIL pode ser observado como tendo duas áreas de cobertura principais: área do Suporte ao Serviço mais concentrada nas tarefas de execução diária e no Suporte aos Serviços de TI, enquanto a área da Entrega de Serviço está mais concentrada no planejamento de longo prazo e respectivo melhoramento. As gerências que compõem o ITIL estão a seguir.

A área de Entrega de Serviço possui cinco gerências:

- Gerenciamento de Níveis de Serviço;
- Gerenciamento de Continuidade;
- Gerenciamento Financeiro;
- Gerenciamento de Capacitação/Capacidade;
- Gerenciamento de Disponibilidade.

Na área de Suporte ao Serviço outras cinco gerências são implementadas:

- Gerenciamento de Incidentes;
- Gerenciamento de Problemas;
- Gerenciamento de Configuração;
- Gerenciamento de Mudanças;
- Gerenciamento de Liberação.

A biblioteca também descreve outras gerências como: Gerência de Aplicação, da Segurança, da Comunicação e do Relacionamento. A Figura 2 mostra a coleção atual das publicações do ITIL sobre as melhores práticas. O processo de Gerenciamento de Serviço, no centro da estrutura, está dividido nas duas áreas mais importantes de Suporte e Entrega.

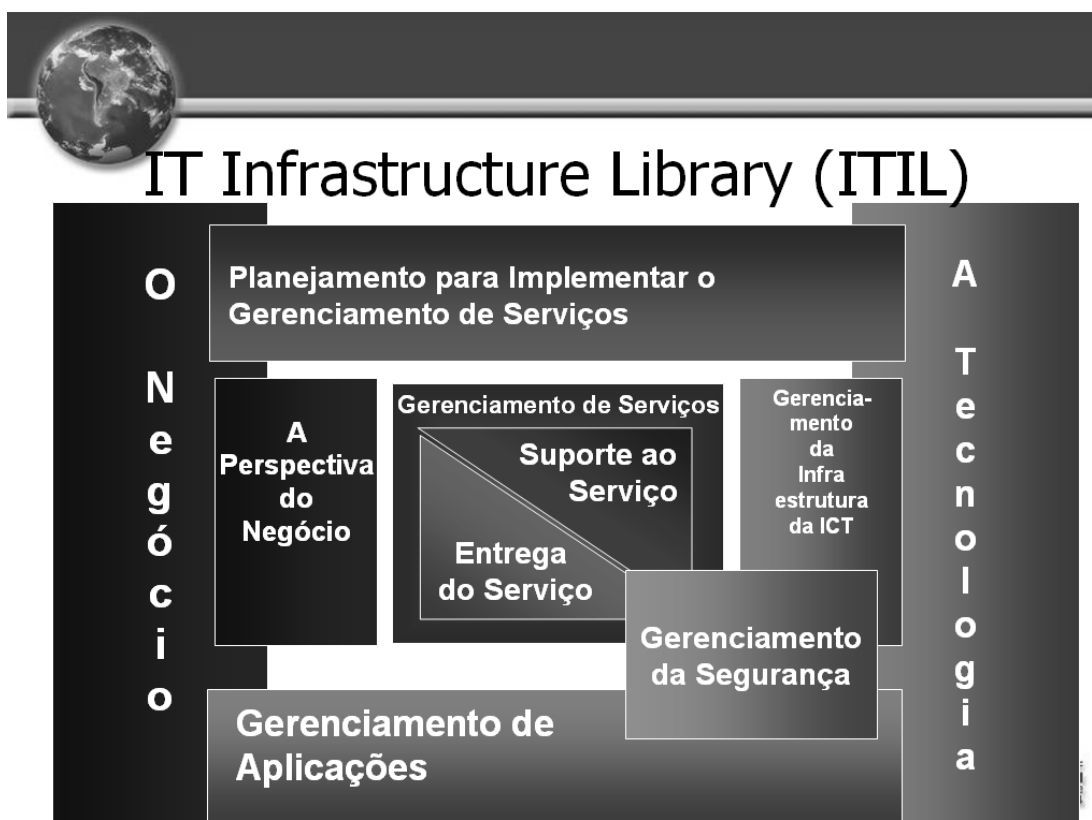
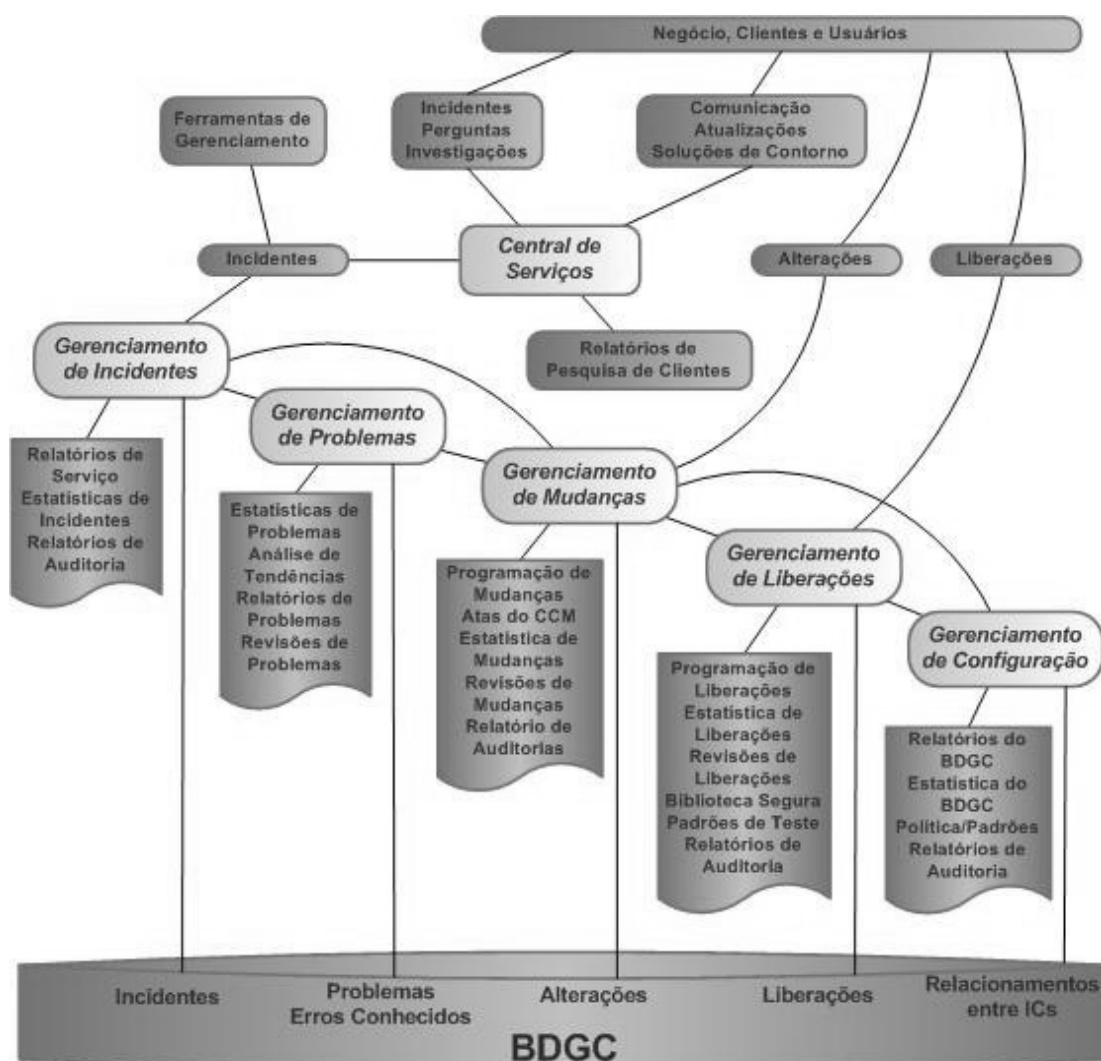


Figura 2 – A estrutura das Publicações do ITIL.

Fonte: OGC, 2005.

O ITIL fornece uma descrição detalhada de um grande número de importantes práticas para o Gerenciamento de Serviços em uma organização de TI, através de processos, tarefas, procedimentos, lista de verificações (*checklists*) e responsabilidades que podem ser adotadas e implantadas em qualquer organização, fazendo com que ela cresça e amadureça seus processos e serviços oferecidos (BON *et al*, 2006). A Gerência de Incidentes está integrada a área de Suporte aos Serviços, que descreve como os clientes e os usuários podem obter acesso aos serviços apropriados de suporte a suas atividades e ao negócio, além de propor como esses serviços podem ser suportados.

A Figura 3 mostra os principais produtos e interfaces dos processos de Suporte aos Serviços:



**Figura 3 – Processos de Suporte a Serviços.**

Fonte: MACFARLANE; RUDD, 2005.

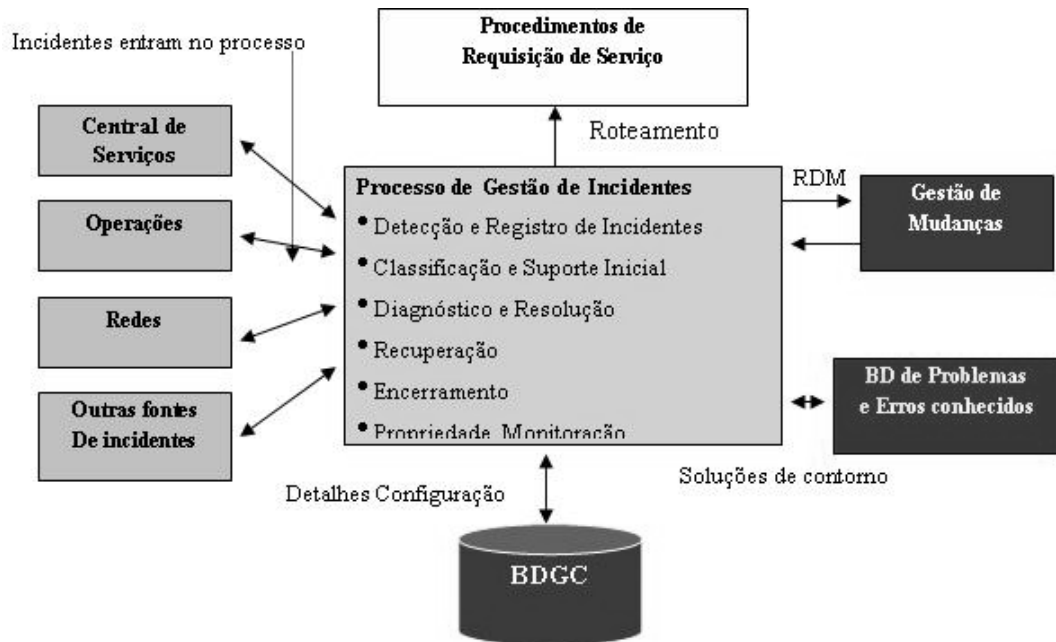
A função do Gerenciamento de Incidentes é a resolução mais rápida possível destes erros, retomando o estado de funcionamento das soluções de TI para os níveis acordados, ou níveis normais de operação (MACFARLANE; RUDD, 2005). Este é um dos processos mais reativos, pois entrará em atuação a partir dos incidentes levantados por usuários ou ferramentas de monitoramento. Entretanto este processo é vital para manter a agilidade dos serviços de TI. É importante considerar também que as informações dos incidentes levantadas neste processo serão de grande importância para o processo de Gerenciamento de Problemas. O Processo de Gerenciamento de Incidentes tem como objetivos:

- Resolver os incidentes o mais rápido possível, restabelecendo o serviço normal dentro do prazo acordado nos SLAs;

- Manter a comunicação dos status dos incidentes aos usuários;
- Escalonar os incidentes para os grupos de atendimento para que seja cumprido o prazo de resolução;
- Fazer avaliação dos incidentes e as possíveis causas informando ao processo de Gerenciamento de Problemas. Este processo não é responsável por fazer o diagnóstico identificando a causa raiz, apenas auxiliará o processo de Gerenciamento de Problemas que tem este foco.

### **2.2.1 Descrição do Processo de Gerência de Incidentes**

Todo processo tem como base entradas e saídas. A entrada principal deste processo são os incidentes. Os incidentes podem vir de muitas fontes, tais como usuários, equipes de operações, redes ou ferramentas de monitoramento que identificam irregularidades nos serviços. Soluções de contornos podem ser buscadas a partir de uma base de erros conhecidos, ajudando a resolver o incidente mais rápido. A Base de Dados do Gerenciamento da Configuração (BDGC) auxiliará na identificação do item de configuração relacionado ao incidente atual, anteriores, mudanças já registradas, problemas abertos e o possível impacto e itens relacionados ao incidente. Determinadas solicitações de usuários podem necessitar de um Registro de Mudança, como por exemplo, uma nova regra de negócio ou instalação de um novo componente. A Figura 4 mostra as principais entradas e saídas que podem ser identificadas na Gerência de Incidentes:



**Figura 4 – Processos de Entradas e Saídas do ITIL.**

Fonte: OGC, 2001.

Na Figura 4 destaca-se a Central de Serviços, como ponto central de contato para os clientes reportarem dificuldades, queixas e questões. Pode servir de interface para outras atividades, tais como, solicitações de mudança, contratos de manutenção, licenças de software, acordos de níveis de serviços e gerenciamento de configurações (OGC, 2001). A proposta do *Service Desk* (Central de Serviços) é separar dentro das operações de TI, quem faz parte do suporte aos usuários e quem vai realizar atividades de resolução de problemas. Ter uma área específica para o suporte traz vantagens para os usuários, propiciando um suporte com maior agilidade e qualidade, e para a equipe de TI mais eficiência, pois o técnico especialista acaba não sendo mais interrompido pelas chamadas diretas dos usuários. Com isso esta unidade passa a ter várias responsabilidades, entre elas:

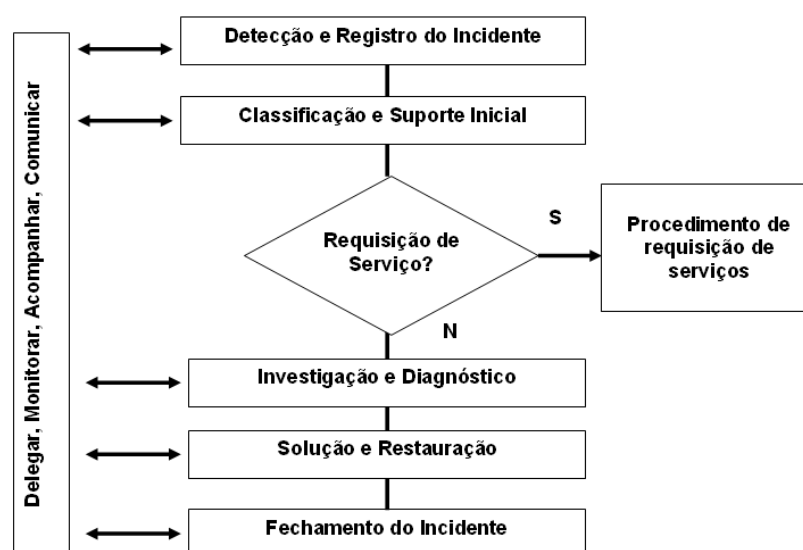
- Receber e gravar TODAS as chamadas dos usuários,
- Gravar e acompanhar incidentes e reclamações,
- Prover uma avaliação inicial dos incidentes,
- Monitorar / escalar incidentes por Acordo de SLA,
- Comunicar mudanças planejadas nos níveis de serviço,
- Encerrar os incidentes com confirmação,
- Manter os usuários informados sobre o progresso de suas requisições,
- Produzir relatórios de gerenciamento,

- Coordenar os grupos de suporte de segundo (2º) e terceiro (3º) nível,
- Prover informações gerenciais,
- Identificar necessidades de treinamento dos usuários,
- Contribuir na identificação de problemas.

A Gerência de Incidentes ocupa uma posição estratégica na prestação de serviços relacionados à atividades de coleta e tratamento de incidentes, destacando-se:

- Detecção de incidentes e registro,
- Classificação e suporte inicial,
- Investigação e diagnóstico,
- Resolução e restauração,
- Fechamento do incidente,
- Responsabilidade pelo incidente, monitoração, acompanhamento e comunicação.

O processo de Gerenciamento de Incidentes está claramente demonstrado na Figura 5. A representação demonstra o incidente desde o momento de sua detecção até o instante de seu encerramento e contabilização.



**Figura 5 – Processos envolvidos na Gerência de Incidentes.**

**Fonte: FERNANDES; ABREU, 2008.**

Cada passo realizado durante a administração do incidente está descrito abaixo segundo (OGC, 2001):



- Detecção de Incidentes e registro: incidentes na maioria das vezes são oriundos de necessidades de suporte dos usuários. O contato com a Central de Serviços poderá acontecer por telefone ou e-mail. Atualmente a maioria das empresas está adotando sistemas *web*, que permitem aos usuários abrirem chamados de suporte diretamente da Intranet ou *website*. Esta postura permite a redução de gargalos no atendimento e facilitando a vida dos analistas de suporte, que tem mais tempo para resolver os incidentes ao invés de gastar o tempo no registro do chamado via telefone. É importante que todos os incidentes sejam registrados, mesmo que resolvidos por telefone. O histórico de incidentes registrados ajudará no processo de identificação de tendências de problemas e também para a extração de informações gerenciais úteis.
- Classificação e suporte inicial: incidentes devem ser classificados de tal forma que permita a identificação de erros conhecidos e gere informações gerenciais que permitam a identificação dos tipos de incidentes mais frequentes. É importante determinar o Impacto e Urgência de cada incidente para determinar a sua prioridade. A prioridade determina qual será a ordem de execução para resolver os incidentes. Para determinar a prioridade utilize como boa prática a combinação entre Impacto e Urgência do incidente. O Impacto será considerado quantas pessoas ou sistemas serão prejudicados pelo incidente. Já a urgência determina a velocidade em que o incidente precisa ser resolvido.
- Investigação e Diagnóstico: uma vez registrado o incidente a atividade de investigação e de diagnóstico será iniciada. Se a Central de Serviços não puder resolver um incidente, ele será encaminhando a outros níveis de suporte que irão investigar o mesmo, usando um conjunto de habilidades e ferramentas disponíveis tais como uma base de conhecimento de Erros Conhecidos. É importante que todas as partes que trabalham com os Incidentes mantenham o registro de suas ações, atualizando o registro do incidente.
- Resolução e Restauração: uma vez que uma solução de contorno ou definitiva para o incidente é encontrada, esta será implementada. Se uma mudança for necessária, uma Requisição de Mudança (RMD) será submetida ao Gerenciamento de Mudanças.
- Fechamento do Incidente: a etapa de fechamento do incidente inclui:
  - Atualização dos detalhes do incidente,
  - Comunicação ao usuário sobre a solução.

- Responsabilidade pelo Incidente, monitoração, acompanhamento e comunicação: é importante que durante todo o ciclo de vida do incidente a Central de Serviços permaneça proprietária do incidente, sendo ela responsável pelo seu fechamento. Desta forma tem-se um comprometimento maior da Central de Serviços para o cumprimento dos prazos, escalando o incidente para o grupo disponível quando necessário. Sendo assim, sempre que o usuário entrar em contato com a Central de Serviços terá uma pronta resposta sobre a situação de suas chamadas. Não é conveniente que os usuários tenham contato direto com os solucionadores finais do incidente, isto fará com que os usuários comecem a manter o contato direto com eles.

Observe que durante o processo de gerenciamento, o incidente assume alguns estados que caracterizam seu ciclo de vida (LIRA; SILVA, 2008). Os estados podem ser denominados como:

- Novo: Estado do incidente ao ser registrado.
- Aceito: O incidente recebe esta denominação após sua primeira análise e classificação quanto sua prioridade.
- Programado: O estado indica que o incidente aguarda atendimento, já foi aceito e espera a definição de um analista para execução de atendimento técnico.
- Atribuído: O incidente é repassado a um técnico responsável.
- Em andamento: O incidente neste estado significa que sua investigação e diagnóstico já foram iniciados.
- Em espera: significa que sua investigação e diagnóstico foram suspensos.
- Resolvido: Implementada uma solução de contorno (*workaround*), permanente ou definitiva (*permanent fix*) e aplicada ao incidente. O serviço de TI afetado é solucionado.
- Encerrado: A equipe de suporte contata o usuário, informa que o serviço foi solucionado e obtêm a confirmação para a restauração do serviço de TI.

A utilização de um padrão para distinção das etapas no ciclo de vida do tratamento de incidentes permite que todos integrantes da equipe de TI acompanhem o andamento dos registros de problemas (*report*), possibilitando então a gerência de incidentes.

## 2.3 NÍVEIS DE SERVIÇOS

O tratamento de incidentes deve ser um serviço oferecido no âmbito da organização, e a produção e entrega desses serviços exige um nível adequado de interação entre o cliente e o provedor. A identificação da qualidade do serviço prestado pode ser estabelecida por parte do cliente, avaliando aspectos técnicos e de interação. A qualidade técnica está calcada basicamente em fatores tecnológicos, enquanto a qualidade de interação é totalmente dependente da percepção do cliente, o que muitas vezes é imprevisível e subjetiva. Para que este processo de prestação de serviços ocorra corretamente, é necessário a definição de um conjunto claro de compromissos entre as duas partes envolvidas. Esta forma de agir é extremamente importante para a área de TI, ao ponto que esta é vista dentro da organização como um provedor de serviços. A tendência é que isso seja baseado em um contrato que descreva explicitamente os produtos (bens ou serviços a serem contratados) e os índices a serem atingidos para o cumprimento do conjunto de compromissos acordados. Esse contrato tem sido representado por um instrumento denominado SLA (STURM *et al.*, 2000). Segundo (TUDE, 2003), SLA é um documento formal, negociado entre as partes, na contratação de um serviço de TI ou telecomunicações.

A negociação de um SLA é o processo no qual há um intercâmbio de informações visando uma conclusão razoável. O grupo usuário precisa estar apto a comunicar suas necessidades claramente. Eles também precisam poder explicar os impactos comerciais dos vários níveis de serviço. Nos casos de SLAs na empresa e nos processos internos, o acordo pode não construir um contrato de obrigação legal, entretanto, a estrutura será a mesma. (GOTTSCHALK, 2005).

O SLA é colocado, geralmente, como anexo do contrato e tem por objetivo especificar os requisitos mínimos aceitáveis para o serviço proposto. O não cumprimento do SLA implica em penalidades, estipuladas no contrato, para o provedor do serviço. A gestão destes acordos com um conjunto de clientes, é possível por meio da implementação de uma Gerência de Nível de Serviços (*Service Level Management, SLM*) que tem como principal finalidade o controle e melhoria contínua na qualidade dos serviços, além de manter a mensuração destes níveis e o equilíbrio entre a demanda e a capacidade (OGC, 2001).

A gestão de serviços de TI tem como filosofia definir, executar e manter os níveis requeridos de serviços de TI para o conjunto de usuários do negócio da empresa, de acordo com as prioridades do negócio e a um custo aceitável. Essa é uma atividade que vem sendo praticada há muito tempo nas áreas responsáveis pela TI nas empresas, tipicamente

relacionadas à gestão e entrega efetiva de aplicações de missão crítica. Uma gestão de serviços efetiva requer que a organização de TI entenda profundamente os seus serviços, incluindo a importância e a prioridade relativas de cada um deles para o negócio. Mas, por ser voltada principalmente para as necessidades do usuário do negócio, se implantada, ela viabiliza um alto retorno do investimento nos gastos de TI e apresenta poucas falhas, por apontar as expectativas apropriadas (MCBRIDE, 1998).

Existe um conjunto de razões para o estabelecimento dos níveis de serviço em uma organização, alguns deles podem ser definidos segundo (STURM *et al.*, 2000):

- Propaganda (*Marketing*) interna dos serviços de TI: caracteriza-se como uma atividade necessária, com o intuito de demonstrar que a área de TI provê bons serviços aos clientes e ao negócio, denotando a sua importância para a organização;
- Controle de custos: Um SLA deve ser capaz de identificar as áreas de serviço que necessitam maior investimento, e em quais áreas os níveis de serviço estão satisfatórios para que os recursos possam ser canalizados de maneira eficiente;
- Estabelecimento de uma estratégia defensiva: Mantendo SLAs precisos entre a área de TI e clientes, e os níveis de serviço documentados, é possível que fatos possam tomar o lugar de percepções, disponibilizando argumentos plausíveis para a terceirização;
- Gestão das expectativas: Inicialmente todos os usuários consideram-se satisfeitos, porém é natural um aumento gradativo de exigência em relação aos serviços oferecidos pela área de TI, a respeito do fato de que mais usuários estão acessando o sistema, mais aplicações devem ser gerenciadas e novas tecnologias estão sendo incorporadas;
- Regulamentação de recursos: A área de TI esta constantemente ligada a manutenção dos níveis de serviço, dessa forma ela se torna atenta ao desenvolvimento de problemas por sobreposição da capacidade e falta de recursos e, então, pode tomar providências antes da degradação dos serviços;
- Satisfação do cliente: O provedor de serviços de TI deve entender o que o cliente percebe como um bom serviço (ex.: prioridade da disponibilidade do e-mail em relação à disponibilidade de outra aplicação) e o cliente deve saber o que é razoável esperar da área de TI;

Neste ponto é importante identificar a principal função dos acordos de serviço (seção 2.3.1), além de reconhecer a caracterização por tipo de cada SLA (seção 2.3.2) e como deve ocorrer o gerenciamento destes acordos (seção 2.3.3).

### 2.3.1 Principais funções dos SLAs

Um dos principais aspectos positivos na criação de um **SLA** está na definição do tempo de resolução de um determinado problema. Inicialmente é definido pelo SLA, quais níveis de serviço são considerados aceitáveis pelos usuários e podem ser fornecidos pelos prestadores de serviços. Existe uma característica básica da natureza humana, querer sempre mais e melhor, independentemente do assunto (STURM *et al.*, 2000). Os requisitos dos níveis de serviço (SLR) definem as métricas e padrões de temporização a serem utilizados.

Tais funções são incorporadas pelo provedor do serviço e pela equipe de TI do cliente, juntamente com os usuários finais. No trabalho conjunto, essas instâncias identificam indicadores chave, focados nas necessidades de negócio do cliente, também definem requisitos dos níveis de serviço e articulam métricas e responsabilidades para o cumprimento dos SLAs.

Os requisitos dizem respeito diretamente à disponibilidade, desempenho, estabilidade e suporte aos serviços de TI que são prestados na organização, desta maneira sendo considerados como a etapa inicial para a implementação dos níveis de serviço. Para (MAURER *et al.*, 2000) a definição das SLAs são criados para garantir a qualidade dos serviços oferecidos aos usuários, e desta forma devem ser validadas a partir de um conjunto de métricas e indicadores, que são formuladas com base nos seguintes pontos de vista:

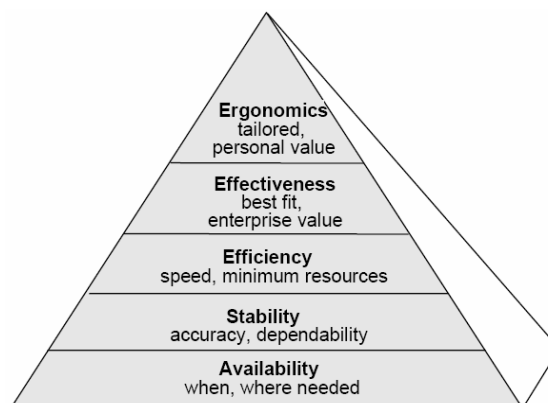
- **Provedor de Serviços:** Esse ator tem a finalidade de focar como os serviços serão entregues. Diante dessa perspectiva, o provedor leva em consideração o desempenho dos sistemas e dos recursos humanos disponíveis, bem como os padrões, políticas, procedimentos e práticas de gerenciamento que devem ser seguidas.
- **Cliente dos serviços:** Esse ator considera a entrega do serviço, gerenciando o provedor com base nos resultados dos serviços entregues. Entre as funções de entrega que o cliente deve considerar incluem-se, o nível operacional e de suporte, o serviço de plano de contingência e o suporte administrativo.
- **Usuário final:** Para esse ator as necessidades e definições de níveis de serviço devem ser significativas, uma vez que refletem diretamente no seu dia-a-dia.

Os indicadores dos níveis de serviço especificam as métricas estatísticas pelas quais os níveis de serviço podem ser medidos. Quando registrar as métricas é importante capturar as métricas objetivas (números), tanto quanto as subjetivas (percepção do usuário) que podem

ser bem diferentes dos valores contabilizados (VAN GREMBERGEN *et al.* 2003, KAMINSKI e PERRY, 2006).

A construção dos SLAs parte de processos delicados, que para alcançar seus objetivos devem estar alinhados de acordo com as perspectivas e métricas de serviço do cliente e do provedor, sendo desta maneira considerada essenciais para o estabelecimento de garantias de qualidade dos serviços e assegurando para ambas as partes os mesmos critérios de avaliação da qualidade (VAHIDOV; NEUMANN, 2008). Com base na afirmação acima, deve ser considerado os seguintes pontos:

- A maturidade do serviço: as definições dos níveis de serviço dependem da maturidade apresentada pelo mesmo. A compreensão dessa maturidade é um fator importante na elaboração dos SLAs, pois referencia um ponto em particular dentro do ciclo de vida do serviço.
- A hierarquia de requisitos do serviço: os níveis de serviço podem ser representados por um conjunto de camadas organizadas de forma hierárquica, conforme mostra a Figura 6. Tais camadas são interdependentes, ou seja, para que se possam atingir os níveis mais altos na hierarquia, os mais baixos, obrigatoriamente, devem ter sido atingidos. Considerando esta estrutura pode-se descrever as camadas da seguinte maneira: a camada de eficiência é responsável pelo tempo de resposta do serviço; a disponibilidade é a primeira camada e serve de base da hierarquia das necessidades; para atender o nível da disponibilidade tem-se a camada da estabilidade que define a precisão e confiabilidade do sistema; a próxima camada trata da eficiência que aborda a velocidade e os requisitos dos recursos; as necessidades dos clientes são atendidas pelo nível da efetividade onde são atendidos os valores do negócio do cliente.



**Figura 6: Hierarquia de níveis de serviços.**

**Fonte: MAURER *et al.*, 2000**

Um SLA oferece, por meio de indicadores, uma linguagem comum para a integração entre a área de TI, que são prestadores de serviço, e seus clientes (usuários). A documentação de um entendimento mútuo obtido durante o processo de negociação do acordo de nível de serviços oferece clareza aos negócios. Existem seis benefícios primários que podem ser esperados dos acordos de nível de serviços (STURM *et al*, 2000):

- Oferecer continuidade,
- Oferecer clareza,
- Servir como veículo de comunicação,
- Proteger contra o fantasma da expectativa,
- Definir padrões mútuos de serviço,
- Definir como será medido o nível de serviços.

Para que os SLA alcancem estes benefícios, são necessários cuidados com alguns fatores críticos como: processo de suporte em geral, disponibilidade, tempo de resposta, horário do serviço, estabilidade, suporte e desempenho (MAURER *et al*, 2000).

### **2.3.2 Tipos de SLA**

Para (STURM *et al*, 2000), existem três tipos de SLAs: o mais comumente utilizado é o SLA na organização, que envolve o prestador de serviços e um cliente interno. Em segundo existem os SLAs externos, e por último os SLAs internos, sendo que o SLA na empresa é aquele que envolve um prestador de serviço ou um departamento fornecedor de serviços de informática e os demais clientes ou departamentos de usuários, dentro da mesma organização. Os acordos de serviços são definidos a partir de um mesmo processo de criação, com isso os conteúdos de acordos diferentes apresentam semelhanças e pontos em comum. As principais diferenças identificadas estão no processo de criação de um SLA que é, fundamentalmente, o mesmo para cada tipo de processo de elaboração do acordo na linguagem utilizada, e nas consequências que resultarão caso não sejam cumpridas as obrigações de nível de serviços. E também podem existir aquelas que integram ao SLA estratégias e planos que garantam a continuidade dos serviços oferecidos (GOTTSCHALK, 2005).

### **2.3.2.1 SLAs na organização**

Neste acordo de serviço, o prestador de serviços e o cliente encontram-se na mesma organização, o que não deve impedir a implementação e o estabelecimento legal e detalhado de um acordo de nível de serviços, o que de modo geral leva a resultados positivos para a empresa. Segundo (STURM *et al*, 2001) o resultado cumulativo da adesão é rigorosa a esses acordos de nível serviço na própria organização, a um nível total de confiança, já está sendo utilizado como estratégia de venda dos serviços fins ou produtos das organizações que estão utilizando esse processo.

### **2.3.2.2 SLAs Externos**

É o tipo de acordo mais rigoroso, geralmente baseado em um contrato legal entre empresas. Muitas empresas ignoram a necessidade de uma revisão legal do acordo externo, com isso acabam com um documento irrelevante e sem valor. Uma característica necessária para esse tipo de acordo é a adesão do mesmo por todos prestadores de serviço, a fim de evitar problemas em suas atividades. A participação de todos interessados (empresa e clientes) permite a manutenção de níveis satisfatórios de serviço. Após assinar um documento sem uma garantia desses níveis, o cliente torna-se vulnerável as atividades realizadas pelos prestadores de serviço e com opções bem limitadas poderá causar efeitos colaterais desagradáveis como encerramento do contrato, níveis de serviços inferiores aos desejados, término prematuro do mesmo, com potencial de aplicação de duras penalidades, ou até mesmo a renegociação dos termos. Segundo (STURM *et al*, 2000), a empresa que se enquadrar em algumas das situações acima poderá buscar amparo legal para o atendimento de suas necessidades.

### **2.3.2.3 SLAs Internos**

Este é um contrato relativamente simples, redigido de maneira informal e que muitas vezes apresenta o texto, que é bastante técnico, inserido em outros documentos formalizados na organização como objetivos e metas individuais e de departamentos, ou como critério no plano de bonificação da empresa.



Uma vez especificado um SLA, independente do tipo e resultado, o SLA precisa ser medido sistematicamente, considerando os aspectos de qualidade, tecnologia e negócio (REGAZZINI, 2001). Dado a conclusão das negociações, a etapa seguinte é documentar o acordo. LEE (1996) explica que o interesse dos contratantes e dos contratados nem sempre é o mesmo e, por isto, o contrato escrito é conseqüentemente o instrumento mais importante para definir as responsabilidades e as expectativas de ambos. São considerados componentes básicos de um acordo de nível de serviço (STURM *et al.*, 2000): Partes do acordo ou contratantes, duração do acordo, escopo, limitações, objetivos, indicadores de desempenho, serviços adicionais, exclusões, relatórios, administração e revisões. É imprescindível que as empresas possam medir os níveis dos serviços prestados, para isso, faz-se uso de uma ferramenta para gerenciar estes níveis, também chamada de Gerencia de Nível de Serviços (*Service Level Management, SLM*).

### **2.3.3 SLM - Gerenciamento dos Níveis de Serviço**

O processo de Gerenciamento do Nível de Serviço é a base para o gerenciamento dos serviços que a área de TI fornece para a organização. O SLM é um conceito aplicado para administrar os serviços que fazem parte dos negócios de uma empresa. É uma metodologia disciplinada e proativa com procedimentos aplicados que visam assegurar que os níveis adequados de serviço sejam prestados a todos os usuários que utilizam tecnologia da informação, de acordo com as prioridades organizacionais e a um custo razoável (LEÃO, 2002). Conforme (IEC, 2005), SLM é o conjunto de pessoas e sistemas que garantem à organização os recursos necessários sendo fornecidos de maneira eficiente, conforme descrito no SLA.

Esta gerência visa alocar adequadamente os recursos disponíveis e gerenciá-los de forma integrada, fazendo com que a qualidade do conjunto seja percebida pelos clientes e usuários, evitando a ocorrência de problemas na entrega e na operação dos serviços de TI. Sendo assim, a ISO/IEC 20000-1 (2005) cita que o SLM deve ser capaz de prover a flexibilidade de adaptar o serviço a essas mudanças. O processo do SLM deve realizar o gerenciamento e a coordenação dos envolvidos no nível de serviço, os seguintes objetivos:

- manter o acordo dos requisitos do serviço e a capacidade esperada do sistema;
- manter o acordo dos objetivos do serviço;
- medir e gerar relatórios do nível de serviço e capacidade obtido, bem

- como a justificativa, caso os objetivos não forem alcançados;
- iniciar rapidamente as ações corretivas, se necessárias; e
- dar informações para melhorar os serviços.

O processo do Gerenciamento do Nível de Serviço pode ser dividido nos seguintes subprocessos: revisão dos serviços disponibilizados, negociação com os clientes, revisão dos contratos de serviços com fornecedores externos, desenvolvimento e monitoração dos acordos de nível de serviço, implementação das políticas e dos processos de melhoria contínua, estabelecimento de prioridades, planejamento do crescimento dos serviços, definição de custo dos serviços em conjunto com o gerenciamento financeiro e da forma de ressarcimento desses custos. Para a implementação de uma solução de SLM, é preciso definir os parâmetros dos níveis de serviços acordados entre a empresa fornecedora e o cliente, além de planejar os processos, firmar compromisso por parte das diferentes gerências relacionadas aos serviços mapeados e checar a necessidade de aquisição de novas ferramentas de gerenciamento. Para (TUDE, 2007), indiretamente, o processo de gerenciamento do nível de serviço traz alguns benefícios, tais como:

- Satisfação dos clientes;
- Melhor entendimento das regras e responsabilidades por ambas as partes;
- Melhor foco da TI nas áreas-chave do negócio;
- Monitoramento dos serviços, que permite identificar problemas não previstos; e
- Identificação da deficiência de treinamento e conhecimento por parte dos clientes ou usuários.

Atualmente para a área de TI gerenciar níveis de serviços pode ser considerado como uma atividade básica de sobrevivência. Pode-se pensar que a TI é uma empresa, com uma gama de serviços ofertados aos seus clientes. É raro encontrar uma empresa bem-sucedida que vende seus produtos como se fossem únicos no mercado. Em vez disso, as empresas se esforçam e investem muito para definir claramente as capacitações e as limitações do produto que estão vendendo. Geralmente, essas definições são apresentadas em um contrato com o comprador. De modo semelhante, uma empresa não pode gerar uma expectativa de êxito se não implementar controles de produção para assegurar que o que é enviado está dentro das especificações (STURM, 2001; GOTTSCHALK, 2005 ).

## 2.4 APOIO A DECISÃO

Atualmente as atividades relacionadas ao apoio à decisão no âmbito das organizações devem ser cuidadosamente elaboradas, pois são fatos que acontecem todo o tempo, em todos os níveis, e influenciam diretamente o desempenho das empresas. Este processo precisa ser bem compreendido para ser levado a bom termo. A “decisão perfeita” pode ser muitas vezes ambígua, pois é influenciada por dificuldades encontradas no momento da escolha além de outros fatores e pontos de vista diferentes. Muitas decisões são complexas, pois envolvem múltiplos objetivos a serem alcançados, diferentes alternativas, conflitos de valores entre grupos de pessoas, além de uma enorme quantidade de informações qualitativas e quantitativas que devem ser levadas em conta no processo decisório (HAMMOND *et al*, 1998).

O ato da tomada de decisão acontece nas mais variadas circunstâncias, idades e posições sociais dos indivíduos, e tem basicamente dois objetos: a ação no momento e a descrição para o futuro (SIMON, 1965). Esta ação no momento possui uma qualidade imperativa, pois seleciona um estado de coisas futuras em detrimento de outro e orienta o comportamento rumo à alternativa escolhida. E a descrição de um estado futuro, num sentido estritamente empírico, pode ser correta ou errada. Em termos conceituais, o processo de tomada de decisão pode ser entendido como o processo de identificar um problema ou uma oportunidade e selecionar uma linha de ação para resolvê-lo. O problema normalmente ocorre porque o estado natural não está de acordo com o desejado pela organização, já a oportunidade, ocorre quando as circunstâncias oferecem uma chance de ultrapassarem objetivos e metas (LACHTERMACHER, 2002). Pode-se dizer então que uma decisão baseia-se em conhecimento ou crenças sobre as relações de causa e efeito das opções disponíveis, e visa alternativa cujas consequências são preferíveis.

Na tarefa de decidir um fator importante a ser considerado é o tempo, pois segundo (ANSOFF, 1997) as exigências em relação ao tempo do responsável pela tomada de decisões sempre parecem ser superiores ao tempo total de que dispõe; decisões de grande importância misturam-se a exigências triviais, embora demoradas; a natureza das decisões possui faceta e modifica-se continuamente. Esta diversidade tende a aumentar com o nível de responsabilidade e torna-se particularmente pronunciada no caso do mais alto executivo da empresa. Também se consideram variáveis importantes no processo de tomada de decisão (FREITAS *et al.*, 1997): os objetivos da organização, os critérios de racionalidade e de eficácia, as informações, raciocínio, valores, crenças, recursos, etc. Estas variáveis servem

de apoio ao decisor, principalmente quando o conteúdo das informações possibilita formar uma base de conhecimento e ajuda no raciocínio, contribuindo para uma decisão acertada. Com o objetivo de realmente sustentar e apoiar as decisões nas organizações atuais, a informação precisa ter como suporte uma adequada TI, a fim de disponibilizar as respostas rápidas e eficientes que a competitividade está constantemente exigindo (FREITAS *et al.*, 1997). Para (BEUREN, 1998), a informação ocupa um espaço importante no apoio à formulação de estratégias, como também no acompanhamento das operações empresariais. Ela coloca que “(...) a informação é fundamental no apoio às estratégias e processos de tomadas de decisão”, devendo, assim, ser disponibilizada aos responsáveis pela elaboração da estratégia empresarial. Neste mesmo contexto, (DRUCKER, 1992) afirma que “(...) as organizações modernas são fundamentadas na informação e no conhecimento”, recomendando, assim, que as organizações trabalhem a informação e o conhecimento como capital estratégico. A abrangência das informações e seu número variado de fontes levam, quase sempre, a uma dificuldade de interpretação o que revela a importância de um sistema de informação que responda a inúmeras questões. Foi para atender a essas necessidades que as organizações, a partir dos anos 1990, passaram a desenvolver ferramentas de tecnologias de informação (TI) que estão permitindo a sistematização do saber e do decidir em todos os níveis gerenciais (JAMIL, 2001). Desse modo, o sucesso empresarial passa a depender, fundamentalmente, da capacidade da organização em termos de administrar a base informacional da empresa e aproveitar as oportunidades de diferenciação que as novas TI oferecem (TORRES, 1995). Considerando que os sistemas de informação como apoio à tomada de decisão têm sido utilizados cada vez mais ao longo das últimas décadas. Neste sentido, os sistemas de apoio à decisão são sistemas informatizados e interativos, utilizados no processo decisório, que proporcionam ao decisor, acesso fácil ao banco de dados e modelos, apoiando a tomada de decisão semi-estruturada ou não-estruturada (SPRAGUE; WATSON, 1989).

#### **2.4.1 Ferramentas de Apoio à Decisão**

Conforme (STÁBILE, 2001) é quase impossível para uma organização, no atual ambiente competitivo, gerenciar suas informações sem o auxílio da tecnologia da informação. As rotinas operacionais obrigatórias, por si mesmas, geram grandes quantidades de informações, tornando sua manipulação difícil e demorada. Ao mesmo tempo, estas informações devem ser passadas aos

administradores para que estes possam desempenhar bem seu trabalho, decidindo os rumos da organização.

A necessidade de informações eficazes faz com que decisores busquem, cada vez mais, ferramentas que os auxiliem no processo de tomada de decisão. Assim, as organizações utilizam os Sistemas de Informação (SI) para manipular suas informações e lutar com mais chances pelos seus objetivos. Com o intuito de alavancar a produtividade e capacidade competitiva das empresas foram desenvolvidas ferramentas de apoio a decisão, relacionadas com Sistemas de Informação. Os SI tornaram-se elementos de primeira necessidade, eles vêm sendo gradativamente implantados nas empresas com base em métodos voltados para identificar dados estratégicos e aspectos críticos dos negócios. Em decorrência do próprio objetivo, vêm crescendo em importância na medida em que agilizam o diagnóstico de problemas e a descoberta de oportunidades. Um SI pode ser definido como um conjunto de elementos ou componentes inter-relacionados que coletam, armazenam, processam e distribuem dados e informações com a finalidade de dar suporte às atividades de planejamento, direção, execução e controle em uma organização (LAUDON e LAUDON, 2001).

Basicamente, pode-se estabelecer os sistemas de informações em 4 níveis funcionais, como estão ilustrados na Figura 7.



**Figura 7: Organograma dos níveis funcionais de um SI.**

**Fonte: BOUZON e CORRÊA, 2007**

Na base da Figura 7 estão os Sistemas Transacionais, que representam a base das outras operações, de onde são retiradas as informações das atividades de planejamento e coordenação. No nível seguinte estão os Controles Gerenciais com a finalidade de buscar as informações no sistema transacional para poder gerenciar as atividades. O Apoio à Decisão utiliza aplicativos (software) como ferramenta decisória para as atividades operacionais e estratégicas complexas, para que estas não sejam praticadas com embasamento somente no

sentimento (*feeling*). No topo Planejamento Estratégico utiliza-se as informações obtidas dos três níveis abaixo, que servem como suporte para o desenvolvimento e melhoria contínua da estratégia logística.

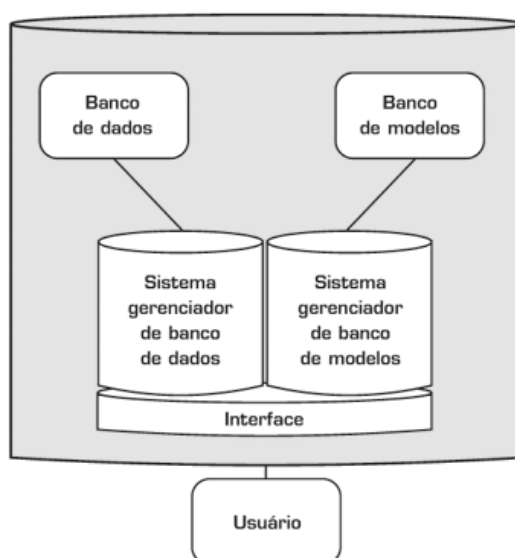
Os SI desenvolvidos como ferramentas para o aporte à decisão são chamados de Sistema de Apoio à Decisão (SAD ou DSS, *Decision Support Systems*) e atuam nos níveis gerenciais das empresas, servindo de suporte às gerências médias da organização, que desenvolvem as atividades relacionadas à monitoração e controle em nível operacional. Os SAD apresentam maior capacidade analítica, o que permite empregar vários modelos para análise da informação, considerando informações geradas por banco de dados internos e de fontes externas. A finalidade é auxiliar seus usuários na tomada de decisão, nos diversos tipos de problemas decisórios, quer sejam de natureza econômica, industrial, política e, até mesmo, social. Dificilmente existem situações a serem tratadas sob um único enfoque, normalmente vários aspectos, ou critérios, devem ser simultaneamente considerados, objetivando a identificação das opções mais satisfatórias (VINCKE, 1992).

Esses sistemas não são relevantes apenas para os gerentes e decisões estratégicas, são apropriados para todas as decisões semi ou não-estruturadas, onde o julgamento e as fontes de informação são ingredientes chaves para tomada de decisão (BARBOSA, 2003). Os SAD dão suporte a decisões menos rotineiras e estruturadas, dificilmente conhecidas de antemão, apoiando os processos de tomada de decisão, fornecendo ao decisor mais detalhes sobre o problema e melhorando seu entendimento sobre as consequências das possíveis decisões e mudanças no ambiente (FUGLSETH; GRONHAUG, 2003). Eles incluem ferramentas analíticas mais avançadas, tais como simulação de cenários e a possibilidade de incluir filtros e reordenar as informações apresentadas. No nível estratégico as decisões são bem menos estruturadas e referem-se ao posicionamento da organização frente às mudanças em seu ambiente e ao planejamento das consequências internas deste posicionamento (LAUDON e LAUDON, 2001). Estes não são sistemas simples, pois apresentam complicados modelos matemáticos, que dependem de parâmetros subjetivos ou da realização de complicadas rotinas matemáticas (GUGLIELMETTI; MARINS; SALOMON, 2003). O resultado do processamento das informações deve servir de subsídios para a tomada de decisão, para (ROBBINS, 2000) a qualidade de toda decisão depende muito da qualidade de informação à disposição. Na complexidade do mercado atual, em que a habilidade para tomar decisões rápidas e inteligentes é uma necessidade absoluta para a sobrevivência, o controle da informação tem adquirido importância significativamente maior. Uma das principais características dos SAD refere-se à estruturação do problema a ser tratado. Conforme ressaltado por (ALMEIDA *et al.*

2002), o nível de estruturação do problema cuja solução é apoiada pelos sistemas de informação é fundamental para identificar a abordagem adequada a esses sistemas, principalmente, para diferenciar sistemas de informação gerencial (SIG) e SAD, uma vez que ambos se prestam a apoiar na tomada de decisões. Outras características técnicas podem ser ressaltadas em um SAD:

- No uso dos SAD não são comuns comportamentos padrões entre o usuário e o aplicativo (o usuário pode incluir fatores e analisar cenários variados, por exemplo);
- A repetição de uma decisão é freqüentemente baixa (o número de vezes de uso do aplicativo é limitado, por exemplo);
- Um aplicativo específico pode nunca ser concluído, estando em constante evolução (em geral tratam de decisões não-estruturadas sobre as quais vai se adquirindo conhecimento aos poucos);
- Contam com a necessidade de dados não formais (intuição, experiências e heurística em adição à lógica).

O modelo conceitual de um SAD proposto por (SPRAGUE e WATSON, 1989), chamado de paradigma DDM (Diálogos, Dados e Modelos), é composto por dois bancos: o banco de dados (BD) e o banco de modelos (BM); e três subsistemas: o sistema gerenciador de banco de dados (SGBD), o sistema gerenciador do banco de modelos (SGBM) e uma interface amigável. Para estes autores, este paradigma é a própria arquitetura de um SAD. A Figura 8 mostra os componentes de um SAD.



**Figura 8: Componentes de um SAD.**

Fonte: (SPRAGUE E WATSON,1989).

Entre os muitos sistemas que podem ser utilizados como apoio à decisão, (SOUZA *et al.*, 2007) destacam os seguintes:

- Sistemas de GED (Gerenciamento Eletrônico de Documentos): ferramenta utilizada para a gestão do conhecimento da empresa (*Knowledge Management*), no sentido de organizar, preservar e compartilhar todo o conhecimento de qualquer natureza existente dentro da empresa, seja ele relacionado com tecnologia, produtos, serviços, Gestão da Qualidade, processos produtivos e de gestão da organização ou qualquer outro que possa agregar valor e diferenciar a empresa.
- Sistemas de workflow: um conjunto de técnicas, sistemas e tecnologias para a utilização de computadores com a finalidade de prover suporte ao trabalho em grupo de pessoas que possuem um objetivo comum de negócio. Entre suas aplicações temos: Sistema de apoio à decisão em grupo; Sistemas de gerenciamento de documentos; Gerenciadores de fluxo de trabalho.
- *Business Intelligence*: a inteligência empresarial surgiu na década de 80 e descreve as habilidades das corporações para acessar dados e explorar as informações, analisando-as e desenvolvendo percepções e entendimentos a seu respeito, o que permite incrementar e tornar mais pautada em informações a tomada de decisão.
- Sistemas de mapa de conhecimento: Para (DAVENPORT e PRUSAK, 1998), um mapa de conhecimento ou um sistema de páginas amarelas indica o conhecimento, mas não o contém. Trata-se, portanto, de um guia, não de um repositório.
- Ferramentas de inteligência competitiva: processo informacional pró-ativo que conduz à melhor tomada de decisão seja ela estratégica ou operacional. É um processo sistemático que visa descobrir as forças que regem os negócios, reduzir o risco e conduzir o tomador de decisão a agir antecipadamente, bem como proteger o conhecimento gerado.
- *Data Warehouse*: um depósito de informações integrado, disponíveis para busca e análise. Estas informações são provenientes de diferentes fontes operacionais da empresa (marketing, finanças, etc.) e armazenadas em um só local, o banco de dados central, onde ocorre um compartilhamento visando a maior integração das atividades da empresa. Além dos dados da própria empresa, o *Data Warehouse* também pode



comportar informações de fontes externas, tais como dados demográficos de consumidores, informações pessoais de cada cliente, entre outras.

- *Data Mining*: é uma metodologia que procura uma descrição lógica ou matemática, normalmente complexa, de possíveis padrões existentes em um conjunto de dados. Em outras palavras, é uma ferramenta capaz de lidar com grandes massas de dados de uma maneira mais eficiente que a estatística, definindo padrões para fenômenos complexos dependentes de muitos parâmetros.
- Portais corporativos: uma evolução das atuais intranets é uma plataforma que integra pessoas, sistemas, serviços e conteúdo corporativo em um ambiente único e personalizado, sob a ótica da Gestão do Conhecimento. Entre suas principais características destacam-se: Totalmente baseado na tecnologia web, ou seja, acessado via navegador; Publicação descentralizada de conteúdo; Permite fácil integração de sistemas web e acesso a bancos de dados; Acesso personalizado a conteúdo e serviços; Interface dinâmica e personalizável.

De forma geral (CLERICUZI *et al.*, 2006) lista algumas aplicações de SAD encontradas na literatura ou apresentadas pelo EURO Working Group Multicriteria Decision Aiding (EWG-MCDA):

- *Interactive Robustness analysis and parameters' Inference for Multicriteria Sorting (IRIS)*: programa concebido para apoiar as decisões de ações em projetos, candidatos ou alternativas, caracterizadas pelas suas avaliações, segundo múltiplos critérios, a um conjunto de classes ordenadas e pré-definidas, segundo uma variante pessimista do método *ELECTRE TRI*. O *IRIS* inclui um módulo para ajudar a ultrapassar eventuais inconsistências quando não há nenhuma forma de satisfazer todas as restrições em simultâneo. Por outro lado, quando estas restrições não são mutuamente incompatíveis, o *IRIS* permite obter conclusões robustas, ao indicar qual a gama de categorias a que uma ação pode ser afetada sem contradizer nenhuma restrição (DIAS; MOUSSEAU, 2002).
- *Variable Interdependent Parameters (VIP Analysis)*: dedica-se à agregação de desempenhos em múltiplos critérios, através da função valor aditivo com informação imprecisa, sendo por isso adequado a situações de escolha nas quais os decisores não consigam fixar valores precisos para os parâmetros que refletem a importância dos critérios de avaliação. Estes parâmetros são vistos como variáveis interdependentes

sujeitas a intervalos e restrições lineares. O *VIP Analysis* incorpora múltiplas ferramentas de análise adequadas a estas situações. Propõe uma metodologia de redução progressiva do número de alternativas e da imprecisão dos parâmetros, introduzindo um conceito de tolerância que permite aos decisores utilizarem algumas destas ferramentas de uma forma muito flexível (DIAS e CLÍMACO, 2000).

- *TRI 2.0a.*: desenvolvido na linguagem de programação de C++ e funciona no ambiente *Microsoft Windows*. O aplicativo visa categorizar alternativas de preferência predefinidas. Resulta numa comparação de todos os critérios, permite ao usuário incorporar os dados requeridos (critérios, alternativas, pesos, perfis e pontos iniciais) e/ou usar as funcionalidades assistentes e os resultados são vistos pelo usuário com rapidez, comparando alternativa e visualizando gráficos (MOUSSEAU *et al.*, 2000).
- Sistema de Apoio à Decisão em Escalada: O objetivo do sistema é apoiar um líder de escalada nas suas decisões, onde existe um grande risco associado. Essas decisões referem-se à configuração de recursos e otimização de rota alpina (HOFFMANN, 2001).
- *Equity*: ajuda a gerentes em todas as organizações a enfrentar o problema comum de como melhor alocar recursos limitados. Minimizando custo, maximizando valores e aceitando um nível tolerado de risco (PHILIPPS, 2003).
- FINCLAS (Classificação Financeira): é o resultado de uma tentativa de integrar metodologias de decisão multicritério, (CORREÇÃO) a fim de fornecer aos analistas financeiros uma ferramenta amigável para resolução de problemas financeiros em tempo real (ZOPOUNIDIS, 2001).
- Algoritmo Genético Construtivo (AGC): um algoritmo para roteamento e localização de recursos; juntamente com um Sistema de Informação Geográfica, forma um sistema de apoio à decisão para problemas de roteamento e localização. Ele pode resolver problemas tanto no domínio rural quanto no domínio urbano. Localização de silos, postos de saúde, rotas de ônibus, caminhões para escoamento da produção, localização de escolas, hospitais, etc. (NARCISO; LORENA, 2002).

Os decisores dentro da organização atuam de forma diferenciada e suas decisões implicam em realizar ponderações sobre vários aspectos, onde a definição de uma ferramenta adequada pode ser um fator crucial para chegar a requisitos mínimos e desta forma encontrar alternativas qualificadas de solução. Tendo em vista a complexidade de

tais problemas e a necessidade de soluções de qualidade, o decisor necessita contar com métodos, técnicas e tecnologias de suporte eficientes (SHIM *et al.*, 2002).

## 2.5 CONSIDERAÇÕES PARCIAIS

Este capítulo demonstrou que a gestão das empresas no cenário organizacional atual necessita de uma administração eficiente dos recursos de informação, onde problemas e ameaças sobre esses ativos podem afetar diretamente a competitividade destas organizações. A identificação, análise e o correto tratamento destes incidentes, assim como a criação de uma base de conhecimento, podem ser diferenciais positivos para a definição de soluções proativas que visem a excelência dos processos destas empresas. Para alcançar esses objetivos podem ser adotados alguns ensaios (*frameworks em inglês*), que tratam a gerência dos serviços de TI. Nesta pesquisa foram estudados os preceitos definidos pela biblioteca ITIL, que agrega um conjunto de gerências e tem por finalidade obter um alto grau de eficiência dos serviços de TI contratados pelas empresas. Pode-se observar também que para enfrentar o novo cenário competitivo exige-se a transformação da TI, na qual se insere os sistemas de informação e de Apoio a Decisão (SAD), mudando também a percepção dos profissionais sobre estes sistemas. Estes estão diretamente relacionados à gestão estratégica da informação e com isso desempenham papel fundamental no âmbito das organizações, pois agilizam o diagnóstico de problemas e a descoberta de oportunidades.

## **CAPÍTULO 3**

### **REFERENCIAL TEÓRICO: FUNDAMENTAÇÃO CONCEITUAL SOBRE ANÁLISE ESTATÍSTICA MULTIVARIADA**

A estatística é uma ciência utilizada desde a antiguidade para fins econômico-administrativos, coletando dados populacionais, registro de colheitas, e outros fatos ligados a fenômenos naturais, caracterizando a estatística como auxiliar da administração pública (CALVO, 2004). Ela pode ser considerada segundo (TRIOLA, 1998) uma coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões (TRIOLA, 1998). A seção 3.1 tem por finalidade mostrar os princípios básicos da estatística. A seção 3.2 descreve os métodos multivariados para análise científica, peça chave para alcançar os objetivos desta dissertação. A última seção deste capítulo, seção 3.3, apresenta a linguagem R, principal ferramenta utilizada para análise estatística neste trabalho.

#### **3.1 ANÁLISE ESTATÍSTICA**

Em sua essência, a estatística é a ciência que apresenta processos próprios para coletar, apresentar e interpretar adequadamente conjuntos de dados, sejam eles numéricos ou não. Para (KACHIGAN, 1986) a estatística é a disciplina científica cujo fim consiste na “recolha, organização e interpretação de dados de acordo com procedimentos bem definidos”.

Alguns conceitos que devem ser reconhecidos nesta área do conhecimento são:

- População: conjunto de elementos que tem pelo menos uma característica em comum.
- Amostra: metodologia tratada pela estatística inferencial, refere-se ao subconjunto de elementos de uma população. Deve possuir dimensão menor que o da população e seus elementos devem ser representativos da população. A seleção dos elementos que irão compor a amostra pode ser feita de várias maneiras e irá depender do conhecimento que se tem da população e da quantidade de recursos disponíveis.
- Variável: é a característica que vai ser observada, medida ou contada nos elementos da população ou da amostra e que pode variar, ou seja, assumir um valor diferente de elemento para elemento. Podem ser qualitativa dividindo-se em qualitativa ordinal e

nominal. Ou quantitativas, que por sua vez, classificam-se em quantitativa discreta e quantitativa contínua.

Método de coleta de dados em um levantamento estatístico:

- Método qualitativo: amplamente utilizado em pesquisas de natureza qualitativa, com objetivo exploratório. Este método consiste em uma abordagem não-estruturada que visa promover uma compreensão do contexto do problema em estudo.
- Método de observação: Este método consiste no registro, de forma sistemática, dos padrões de comportamento das pessoas/objetos, para obter informações sobre a situação em estudo.
- Método *survey*: É considerado o método mais comum de coleta de dados primários em uma pesquisa. Trata-se de uma forma estruturada e padronizada, através de um questionário, para a obtenção das informações necessárias para o estudo. Os dados podem ser obtidos pelo preenchimento de questionários através de entrevistas

A interpretação das principais características dos dados por meio de métodos gráficos gerados a partir de listas numéricas é bem mais simples. O método mais apropriado dependerá da natureza dos dados, que podem ser classificados da seguinte forma (EHLERS, 2006):

- Variáveis qualitativas ou categóricas que podem ser:
  - nominais, por exemplo sexo (masculino, feminino), classificação de defeitos em uma máquina.
  - ordinais, com categorias ordenadas, por exemplo salinidade (baixa, média e alta).
- Variáveis quantitativas que podem ser:
  - discretos, contagens ou número inteiros, por exemplo número de ataques de asma no ano passado.
  - contínuos, medidas numa escala contínua, tais como volume, área ou peso.

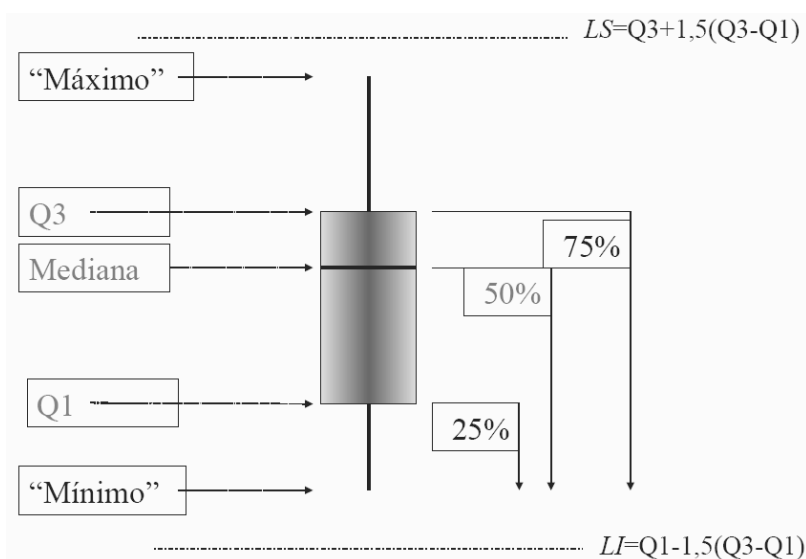
### 3.1.1 Ferramentas para Análise Estatística Descritiva

Após a coleta de uma amostra de dados, a análise descritiva sobre estas informações permite ao pesquisador familiarizar-se com os dados, organizá-los e sintetizá-los de forma a obter as informações necessárias para responder as questões que estão sendo investigadas. Para auxiliar nesta análise, várias ferramentas para representação gráfica dos resultados podem ser encontradas na literatura. Entre elas destacam-se:

- Gráficos em Barras

- Histogramas
- Diagrama Circular
- Medidas de Dispersão
- Gráfico de Sequência
- *Box-plot*

Dentre todas as representações citadas, para esta pesquisa serão usados os Boxplots, um tipo de gráfico que representa os dados através de um retângulo construído com os quartis, conforme a Figura 9, utilizando cinco medidas estatísticas: valor mínimo, valor máximo, mediana, primeiro e terceiro quartil da variável quantitativa. As medidas do Box-plot, oferecem a idéia da posição, dispersão, assimetria, caudas e dados discrepantes. A posição central é dada pela mediana e a dispersão pelo desvio interquartil  $dq = Q3 - Q2$ . As posições relativas de Q1, Q2 e Q3 dão uma noção da assimetria da distribuição. Os comprimentos das caudas são dados pelas linhas que vão do retângulo aos valores atípicos e nas extremidades tem-se LS como limite superior, valores máximos, e LI limite inferior ou valores mínimos.



**Figura 9 – Descrição da técnica de representação Box-plot.**

### 3.2 ANÁLISE MULTIVARIADA

As várias áreas do conhecimento pesquisadas atualmente apresentam uma quantidade de dados e informações que refletem a complexidade dos inúmeros fenômenos que as compõem, o que remete obrigatoriamente a análise de muitas variáveis diferentes. Para a

explicação desses fenômenos, o pesquisador deve coletar e analisar todos os dados possíveis de acordo com uma hipótese. A avaliação realizada a partir da amostragem ou experimentação dessas hipóteses muitas vezes pode modificar a explicação do fenômeno. Neste contexto, a inferência estatística é realizada de acordo com o paradigma hipotético-dedutivo (BOCK, 1975). Além disso, o avanço da tecnologia computacional tem proporcionado avanços e facilidades no que diz respeito à análise destes dados em diversas áreas. Assim, percebe-se uma necessidade de procedimentos de análise e interpretação adequados para a extração das informações, principalmente para tomada correta de decisões de forma a possibilitar um gerenciamento mais eficiente e um crescimento na produtividade (BAKKE *et al.*, 2008).

De acordo com (HAIR, 2005) e (MALHOTRA, 2001), a Análise Multivariada é uma ferramenta estatística que processa informações de modo a simplificar a estrutura dos dados, quando o número de variáveis envolvidas é muito grande, facilitando o entendimento da relação existente entre as variáveis do processo. Segundo (DAVIS, 1986), os métodos multivariados são poderosos permitindo ao pesquisador manipular diversas variáveis simultaneamente. Esses métodos de análise têm comprovado amplamente sua eficácia no estudo de grandes massas de informação. Trata-se de métodos chamados de multidimensionais que permitem a confrontação entre duas ou mais variáveis. Pode-se, então, extrair-se as tendências mais sobressalentes e hierarquizar-las, eliminando os efeitos que perturbam a percepção global (ESCOFIER e PAGES, 1992).

A necessidade de compreensão das relações entre as diversas variáveis faz com que as análises multivariadas sejam complexas ou até mesmo difíceis, porém o uso destas técnicas é justificado pelo fato de que a utilização de apenas uma variável é geralmente insuficiente para a explicação de um experimento, ou incapaz de discriminar indivíduos com relação a qualquer critério que seja empregado. Quanto maior o número de variáveis, maior será a relação entre as mesmas, o que torna complexa a análise por métodos comuns de estatística univariada.

Independente de possuir um modelo teórico estruturado, a análise multivariada pode ser aplicada com diversas finalidades. A finalidade de sua aplicação pode ser de reduzir dados ou de simplificação estrutural, de classificar e agrupar, de investigar a dependência entre variáveis, de predição e de elaborar hipóteses e testá-las (JOHNSON; WICHERN, 1992). Para (PISANI, 1969), as questões fundamentais a ser enfrentadas pela análise multivariada são:

- Problemas de dependência, quando utilizadas análises de regressão múltipla e correlação canônica;
- Problemas de inferência, quando utilizados em extensões multivariadas de métodos univariados, como análise de variância;

- Problemas de redução e escala, quando são utilizadas a análise de componentes principais, análise dos fatores, distância generalizada “D” de *Mahalanobis* e etc.;
- Problemas de discriminação e classificação, quando são utilizadas funções discriminantes e as diversas técnicas de análise de agrupamentos.

Os modelos multivariados possuem, em geral, um propósito através do qual o pesquisador pode testar ou inferir a respeito de uma hipótese sobre um determinado fenômeno. No entanto a sua utilização adequada depende do bom conhecimento das técnicas e das suas limitações. Essas técnicas têm sido amplamente usadas em diversas áreas: na agronomia, para avaliação de questões relacionadas à produção e até mesmo genótipo dos frutos, entre outros estudos conforme (BARROSO; ARTES, 2003); na biologia, em pesquisas para o melhoramento de plantas onde os dados são avaliados por meio de variadas análises (LEDO *et al.*, 2003); nas ciências médicas para avaliação de diagnósticos, classificação de pacientes entre outras atividades; na área de economia, para consultas de perfis e liberação de créditos e também no agrupamento de clientes de acordo com características e comportamento (BARROSO; ARTES, 2003) e (JOHNSON; WICHERIN, 2002); na engenharia civil com a finalidade de identificar correspondências entre diferentes variáveis e fatores comuns em classes e tipos de imóveis (TRIVELLONI; HOCHHEIM, 1968); e em todas as demais áreas do conhecimento como psicologia, educação, geologia, ciências sociais, ergonomia, tecnologia da informação e etc.

### 3.2.1 Termos e Conceitos Básicos

Neste item serão apresentados alguns termos e conceitos importantes para o entendimento das técnicas de análise estatística multivariada desenvolvida neste trabalho.

#### 3.2.1.1 Dados Multivariados

MINGOTTI (2005) define que análise multivariada consiste em um conjunto de métodos aplicados em situações, onde inúmeras variáveis são medidas simultaneamente em cada elemento amostral. Uma observação multivariada de dimensão  $p$ , ou  $p$ -variada, é um vetor:

$$X = (X_1, X_2, \dots, X_p)^t$$



onde  $t=$  transposto e cujas coordenadas  $X_1$  a  $X_p$  são variáveis aleatórias oriundas de várias medidas de um mesmo elemento amostral (BAKKE *et al.*, 2008). A representação dos dados em uma análise multivariada pode adotar a notação  $y_{ij}$  para mostrar um valor particular da  $i$ -ésima unidade amostral ou experimental e da  $j$ -ésima variável mensurada. A matriz resultante deste processo apresenta as seguintes características:

$$Y_{(n \times p)} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{bmatrix} = [y_1 \ y_2 \ \cdots \ y_p] \quad (2)$$

onde  $Y_i$  é um vetor ( $p \times 1$ ) de observações das  $p$  variáveis no indivíduo  $i$  ( $i=1,2,\dots, n$ );  $y_j$  é um vetor ( $n \times 1$ ) de observações da variável  $j$  nos  $n$  indivíduos. Assim, cada indivíduo é representado no espaço  $p$ -dimensional por um ponto, onde suas coordenadas são dadas por  $Y_i=(y_{i1}, y_{i2},\dots,y_{ip})$ , para  $i=(1,2,\dots,n)$  (SARTORIO, 2008).

Para (HAIR *et al.*, 2006) os dados multivariados devem ser devidamente definidos de acordo com seu tipo, para que a aplicação correta da técnica estatísticas seja viabilizada. Este autor define os dados em dois tipos:

- Dados métricos: conhecidos como dados quantitativos, são medidas intervalares ou proporcionais que identificam, descrevem indivíduos ou objetos a partir de um grau e atribuição de valor.
- Dados não-métricos: são dados qualitativos, definem propriedades e atributos dos indivíduos ou objetos a partir de suas características, também chamados de dados ordinais ou nominais.

### 3.2.1.2 Padronização, Escalonamento e Medidas de Similaridade

A finalidade da padronização e escalonamento dos dados originais é expressar cada observação em termos de variações inerentes ao sistema (autoescalonamento).

Sobre Medidas de Similaridade pode-se dizer que cada objeto é representado por um ponto no espaço  $n$ -dimensional e, portanto, pode ser agrupado com outros que estejam próximos e mais se assemelham a ele. Dois critérios de melhor associação podem ser utilizados:

- Covariância e Correlação: Partindo da matriz de dados  $Y$  ( $m \times n$ ), obtém-se a matriz de covariância  $C$ , onde seus elementos são dados por:

$$c_{kl} = \frac{1}{m-1} \sum_{i=1}^m (d_{ik} - \bar{d}_k)(d_{il} - \bar{d}_l) \quad \text{onde} \quad \bar{d}_k = \frac{1}{m} \sum_{i=1}^m d_{ik} \quad (3)$$

$c_{kl}$  é grande e positivo quando, para a maior parte das amostras, os valores das variáveis  $k$  e  $l$  desviam da média na mesma direção. Portanto, a covariância de duas variáveis é uma medida de sua associação. Para cada elemento da matriz de covariância pode ser calculado o coeficiente de correlação, conseqüentemente a matriz de covariância pode ser transformada numa matriz de correlação  $R$ , onde seus elementos são dados por:

$$r_{kl} = \frac{c_{kl}}{s_k \cdot s_l} \quad (4)$$

onde  $s_k$  e  $s_l$  são os desvios padrões das variáveis  $K$  e  $l$

Os valores de  $r_{kl}$  são uma covariância padronizada entre  $-1$  e  $+1$ .

- Medidas de distâncias: Na análise de agrupamentos (*cluster analysis*) a similaridade entre duas amostras pode ser expressa como uma função da distância entre os dois pontos representativos destas amostras no espaço  $n$ -dimensional. A maneira mais usual de calcular a distância entre dois pontos  $a$  e  $b$  no espaço  $n$ -dimensional é conhecida por distância euclidiana ( $x_{ab}$ ) e é dada por:

$$x_{ab}^2 = \sum_{j=1}^n (d_{aj} - d_{bj})^2 \quad (5)$$

Existem outras maneiras de calcular distâncias, como a distância de Mahalanobis, mas estas não são apresentadas neste estudo científico.

### 3.2.1.3 Técnicas de Análise Multivariadas

Para uma análise multivariada eficaz, pode-se utilizar um conjunto de técnicas diferentes, de forma individual ou em conjunto. Algumas destas técnicas estão descritas a seguir (BAKKE, 2008):

- **Análise fatorial:** inclui a análise de componentes principais e análise dos fatores comuns. É aplicada quando há um número grande de variáveis correlacionadas entre si, com o objetivo de identificar um número menor de novas variáveis alternativas, não correlacionadas e que, de algum modo, sumarizem as informações principais dos dados originais, encontrando os fatores ou variáveis latentes (MINGOTI, 2005);
- **Análise de agrupamento:** trata-se de uma análise que identifica grupos em objetos de dados multivariados. O objetivo é formar grupos com propriedades homogêneas de amostras heterogêneas grandes. Devem-se buscar os grupos mais homogêneos possíveis, onde as diferenças sejam as maiores possíveis (HÄRDLE e SIMAR, 2007);

### **3.2.2 Análise de Agrupamentos**

A Análise de agrupamentos tem por objetivo agrupar as variáveis, conforme sua proximidade ou suas características comuns, buscando mostrar a homogeneidade dentro do grupo e a heterogeneidade entre os grupos. Conforme (HAIR, 2005) esta análise pode ser definida como “um grupo de técnicas multivariadas cuja finalidade primária é agregar objetos com base nas características que eles possuem”. Da mesma forma, (MALHOTRA, 2001) define a análise de conglomerados como uma técnica usada para classificar objetos ou casos em grupos relativamente homogêneos chamados conglomerados.

A formação dos diversos grupos homogêneos pode ter como objetivo um estudo exploratório com a formação de classes de objetos, uma simplificação das informações; ou ainda a identificação de relacionamentos entre as observações. Os grupos, que são obtidos através de uma ou mais técnicas de Análise de Agrupamentos, devem apresentar tanto uma grande homogeneidade interna (dentro de cada grupo), como uma grande heterogeneidade externa (entre grupos). Portanto, se a classificação for bem sucedida, quando representados em um gráfico, os objetos dentro dos clusters (grupos) estarão muito próximos, e os grupos diferentes ficarão afastados.

A Análise de Agrupamentos tem aplicabilidade em diversas áreas do conhecimento, com destaque para marketing, por exemplo, na segmentação de mercado e mais recentemente nas áreas de investimentos, economia e finanças. Efetuar uma análise de cluster consiste inicialmente em selecionar as variáveis ou características que serão determinantes na

formação dos agrupamentos. Em seguida é feito um tratamento das variáveis (transformações e padronização), e a escolha da medida de similaridade a ser adotada.

Podem ser empregadas várias técnicas aglomerativas para a análise, as quais segundo (DAVIS, 1986) dividem-se em quatro tipos básicos:

- Métodos de partição: classificam regiões no espaço, em função de variáveis, avaliando sua ocupação em termos de observações.
- Métodos com origem arbitrária: procuram classificar as observações segundo “k” pontos arbitrários, os quais servirão como centróides iniciais, e as observações irão se agrupar por similaridade, em torno desses centróides e formar agrupamentos.
- Métodos por similaridade mútua: procuram agrupar observações que tenham uma similaridade em comum. Inicialmente uma matriz  $n \times n$  de similaridades entre colunas são repetidamente recalculadas. Colunas representando membros de um único agrupamento tenderão apresentar intercorrelações próximas a 1 e valores menores que dos não membros.
- Métodos por agrupamentos hierárquicos: são as técnicas mais comumente usadas em Geologia. A partir da matriz inicial de dados é gerada uma matriz simétrica de similaridades e inicia-se a detecção de pares de casos com a mais alta similaridade, ou a mais baixa distância. Para a combinação, segundo níveis hierárquicos de similaridade, escolhe-se entre os diversos procedimentos aglomerativos, de tal modo que cada ciclo de agrupamento obedeça a uma ordem sucessiva no sentido do decréscimo de similaridade.

Com a aplicação da análise e de posse dos resultados sobre qualquer amostra de dados, é possível escolher o método mais adequado para o objetivo do estudo, tanto no que diz respeito à técnica empregada, quanto ao número de clusters. Com isso os agrupamentos são caracterizados, interpretados e validados, com base no perfil das suas observações. Existem muitas maneiras de procurar agrupamentos no espaço n-dimensional. Segundo (NETO e MOITA, 1997) a maneira matematicamente mais simples consiste em agrupar os pares de pontos que estão mais próximos, usando a distância euclidiana, e substituí-los por um novo ponto localizado na metade da distância entre eles. Este procedimento, quando repetido até que todos os pontos sejam agrupados em um só ponto, leva a construção do dendrograma, onde, no eixo horizontal são colocadas as amostras e, no eixo vertical, o índice de similaridade,  $s_{ij}$ , entre os pontos  $i$  e  $j$ , calculado segundo a seguinte expressão:

$$s_{ij} = 1 - \frac{d_{ij}}{d_{\text{máx}}} \quad (6)$$

Os dendrogramas, portanto, consistem em diagramas que representam a similaridade entre pares ou grupos de amostras, em uma escala que vai de um (identidade) a zero (nenhuma similaridade). Para a construção dessas estruturas representativas podem ser usadas várias técnicas para o cálculo das distâncias que ligam os eixos horizontais no gráfico. Neste estudo se optou pelo uso de medidas correlacionais (Correlação de Pearson), a qual representa a similaridade pela correspondência de padrões ao longo dos atributos, descartando assim a magnitude dos valores e considerando apenas o padrão global dos mesmos. A partir da definição destas distâncias, é importante observar como os objetos são agrupados, o que denota o método para o cálculo da distância entre os novos clusters. Para estes casos, várias metodologias de ligação podem ser utilizadas, entre elas: Ligação Simples, Ligação Completa, Médias dos Grupos, Centróide dos Grupos e o Método de Ward. Este último método, por utilizar a análise de variância para calcular as distâncias entre os clusters, foi o que apresentou melhores resultados sobre a amostra de dados disponível neste trabalho, sendo assim adotado como padrão para este estudo.

Por fim, é importante salientar a questão relacionada a padronização dos dados, visto que os dados obrigatoriamente devem ser padronizados antes que as similaridades sejam calculadas. Este fato se comprova devido a característica das medidas de distância serem demasiadamente sensíveis a diferentes escalas ou magnitudes entre as variáveis. A construção dos dendrogramas pode ser feita diretamente por todos os programas estatísticos que fazem classificação de dados através de agrupamento hierárquico (*Hierarchical Analysis* ou *Cluster Analysis*).

### 3.2.3 Análise Fatorial

A Análise Fatorial (AF) é uma técnica estatística que envolve um processo composto de vários métodos estatísticos multivariados, com o propósito de definir a estrutura subjacente em uma matriz de dados. Para (MALHOTRA, 2001), “a Análise Fatorial é um nome genérico que denota uma classe de processos utilizados essencialmente para redução e sumarização de dados”. Conforme (Pereira, 2004), “esta é uma análise multivariada que se aplica à busca de identificação de fatores num conjunto de medidas realizadas”.

A AF tem como princípio básico, a redução do número original de variáveis a um conjunto menor de “fatores” independentes, de forma que esses fatores possam explicar as variáveis originais (FERNANDES e LIMA, 2005). Na composição desses fatores, tem-se que:

- as variáveis mais correlacionadas se combinam dentro do mesmo fator;
- as que compõem determinado fator são praticamente independentes das que constituem outro fator;
- a derivação dos fatores se processa visando a maximizar o percentual de variância total relativa a cada fator consecutivo;
- os fatores não são correlacionados entre si.

Inicialmente o pesquisador tem várias medidas (variáveis) e não percebe que elas podem ser resumidas em um ou mais fatores que condensem esse volume de informações. A análise fatorial como técnica estatística multivariada, obtém com êxito através de correlações a redução das variáveis pelos fatores resultantes.

(RODRIGUES, 2002), explica que a AF ou análise do fator comum, tem como objetivo a identificação de fatores subjacentes às variáveis observacionais, o que contribui para facilitar a interpretação dos dados. Um exemplo prático seria ao invés de buscar entender o comportamento de 15 ou 20 variáveis observacionais, o pesquisador deverá procurar entender o comportamento de 3 ou 4 fatores latentes através do comportamento dos seus “scores” fatoriais. Complementando esta afirmativa, (JOHNSON e WICHERN, 1992) explicam que na análise fatorial as variáveis são agrupadas em função de suas correlações. Ou seja, as variáveis que compõem um determinado fator devem ser altamente correlacionadas entre si e fracamente correlacionadas com as variáveis que entram na composição do outro fator.

A AF faz uma abordagem estatística, com o propósito de analisar as variáveis e explicá-las através de suas dimensões comuns, que são os fatores calculados. É possível encontrar um número de fatores tantos quanto forem às variáveis, porém há uma perda de objetividade ao se utilizar um grande número de fatores. Para a execução da AF, alguns passos são necessários como: a formulação do problema, determinar a matriz de correlação, calcular os autovalores, que fornecerão elementos para encontrar a variância total explicada por cada um dos fatores encontrados.

### **3.2.4 Análise de Componentes Principais**

Análise de Componentes Principais é utilizada para obter a participação de cada variável na construção de um índice ou no alcance de resultados. Segundo Souza (2000), a

idéia matemática do método de análise de componentes principais é conhecida há muito tempo, apesar do cálculo das matrizes dos autovalores e autovetores não ter sido possível até o advento dos computadores.

Esta análise é utilizada na redução das variáveis e identificação das que possuem maior influência no sistema. Ela consiste essencialmente em reescrever as coordenadas das amostras em outro sistema de eixo mais conveniente para a análise dos dados. Em outras palavras, as  $n$ -variáveis originais são escritas em combinações lineares por meio dos autovalores que geram os autovetores dos dados originais e, esta combinação é denominada de componente principal (CP). A principal característica da CP, além da ortogonalidade, é obter em ordem decrescente de máxima variância, ou seja, a primeira componente principal detém mais informação estatística que a segunda componente principal, e assim por diante. Este método permite a redução da dimensionalidade dos pontos representativos das amostras, sem contudo desprezar os dados originais. Pois, embora a informação estatística presente nas  $n$ -variáveis originais seja a mesma das  $n$ -componentes principais, é comum obter em apenas 2 ou 3 das primeiras componentes principais mais que 70% desta informação (NETO e MOITA, 1997).

Para (FERREIRA, 2008) componentes principais são uma técnica de análise intermediária e, portanto não se constituem em um método final e conclusivo. Esse tipo de análise se presta fundamentalmente como um passo intermediário em grandes investigações científicas, podendo ainda ser aplicada na análise de regressão múltipla, principalmente, nos casos de colinearidade ou de multicolinearidade além de aplicar-se à análise de agrupamento e como estimadores de fatores nas técnicas multivariadas denominadas de análises fatoriais.

### **3.3 R – LINGUAGEM PARA CÁLCULOS ESTATÍSTICOS**

A linguagem R é um ambiente para computação estatística e geração de gráficos, que disponibiliza uma grande variedade de métodos estatísticos, como modelagem linear, não linear, testes estatísticos, séries temporais, classificação, métodos multivariados, geoestatística, além de técnicas gráficas. Um dos seus pontos fortes é a facilidade na construção de gráficos em alta qualidade (R PROJECT, 2009). O R além de ser uma linguagem gratuita, apresenta código fonte aberto, podendo ser modificado ou implementado com novos procedimentos desenvolvidos pelo usuário a qualquer momento. Esta é uma importante ferramenta na análise e manipulação de dados, com testes paramétricos e não paramétricos, modelagem linear e não linear, análise de séries temporais, análise de

sobrevivência, simulação e estatística espacial, além de apresentar facilidade na elaboração de diversos tipos de gráficos. Para (PRADO, 2009), a formação de pesquisadores exige treinamento na implementação de análises estatísticas, necessárias ao processamento dos dados de pesquisa e à elaboração e teste de hipóteses, por este motivo o uso da linguagem “R” tem crescido em diversas áreas da ciência.

Apesar do seu caráter gratuito o R é uma ferramenta bastante ampla e com boas capacidades ao nível da programação, com um conjunto bastante vasto de pacotes que acrescentam potencialidades a versão base do R. Algumas atribuições do time de desenvolvimento em relação a Linguagem R (R DEVELOPMENT CORE TEAM, 2003):

- prover suporte ao projeto, considerando o R uma ferramenta madura e disponível para o desenvolvimento continuado, permitindo futuras inovações em programas para estatística e computação científica;
- prover um ponto de referência para pessoas, instituições ou empresas que querem interagir com a comunidade que desenvolve o R;
- manter e administrar os direitos do R e sua documentação

Por ser uma linguagem orientada a objetos o usuário pode criar suas próprias funções, e sua própria rotina na análise de dados. Outro atributo do R é sua capacidade de interagir com outros programas estatísticos, bem como de banco de dados, isto pelo fato do mesmo possuir inúmeros pacotes já prontos, chamados *standard packages*, os quais possibilitam criar novas rotinas e funções. O R apresenta-se como uma ferramenta promissora, podendo adaptar-se a diversas áreas de interesse, como na educação, no desenvolvimento de soluções comerciais, etc. Como vantagens da linguagem R, pode-se citar pequeno consumo de memória, velocidade de processamento e código fonte aberto (RIBEIRO JUNIOR, 2009). Por esses, e por outros motivos esta linguagem vem sendo utilizada cada vez mais por diferentes usuários em diversas áreas do conhecimento (SILVA JUNIOR, 2005).

As execuções de operações no sistema “R” são efetuadas por meio de ordens em linha de comando, processo esse de uso bastante conhecido por profissionais da área de informática, mas não muito didático ou apreciado por profissionais de outras áreas do conhecimento (SILVA, 2009).

### 3.3.1 Ambiente de Desenvolvimento

O ambiente R é formado por um conjunto integrado de programas utilizados para manipulação de dados, cálculos e geração de gráficos. Entre outras características dispõe de:



- armazenamento e manipulação efetiva de dados;
- operadores para cálculos sobre variáveis indexadas e cálculos matriciais;
- uma ampla, coerente e integrada coleção de ferramentas para estatísticas e análise de dados;
- capacidade gráfica para análise exploratória de dados, que permitem a visualização diretamente sobre o vídeo de uma estação de trabalho ou obter cópias em discos rígidos;
- uma linguagem de programação bem desenvolvida, simples e efetiva que inclui: estruturas condicionais, cíclicas, funções recursivas e possibilidades de entradas e saídas de dados.

O R é semelhante a outras linguagens de programação, como C, Java e Perl. Assim como estas, o R é capaz de executar uma ampla variedade de tarefas de computação, por meio da interação e manipulação de diversos comandos. Esta ferramenta contém diversos mecanismos incorporados para a organização de dados, execução de cálculos e criação de representações gráficas. Atualmente o R está disponível para a família UNIX (incluindo LINUX), a maior parte dos Mac OS e ainda Windows 95, 98, NT, 2000, Me, XP.

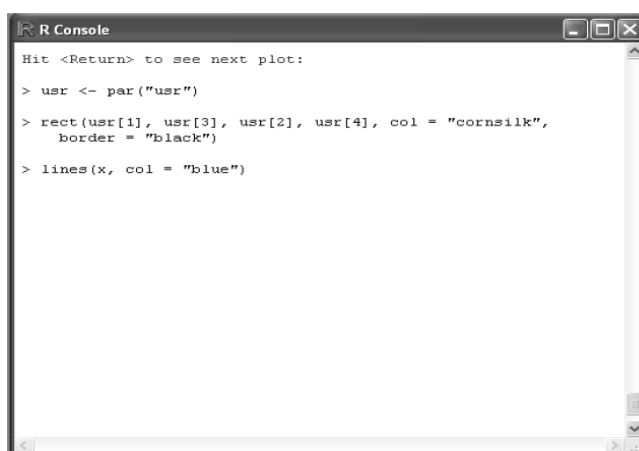
A linguagem permite a manipulação de arranjos de dados em várias dimensões, sistemas univariados, bivariados e multivariados. Além disso, o R possui um conjunto de pacotes que auxiliam os pesquisadores a desenvolver cálculos estatísticos precisos, interagindo com a linguagem por meio de interfaces gráficas ou através de linhas de comando. Os comandos R devem ser preferencialmente escritos em um editor de texto como o bloco de notas, possibilitando editar tais arquivos, denominados de scripts, nos quais podem executar os códigos digitados. No ambiente de desenvolvimento, cada linha deve conter um comando, uma função, uma atribuição ou um comando.

Exemplo:

```
# programa R demonstrando o uso de vetores lógicos
x = c (2.2, 1.3, 4.7, 5.1, 3.2) # cria um vetor de tamanho 5
x                               # imprime o vetor
```

O ambiente de programação do R está baseado em expressões, que são comandos. Estas expressões podem ser organizadas em grupos com uso de chaves  $\{expr_1; expr_2; \dots; expr_m\}$ . Os comandos são separados por ponto e vírgula ou por uma nova linha. O ponto e

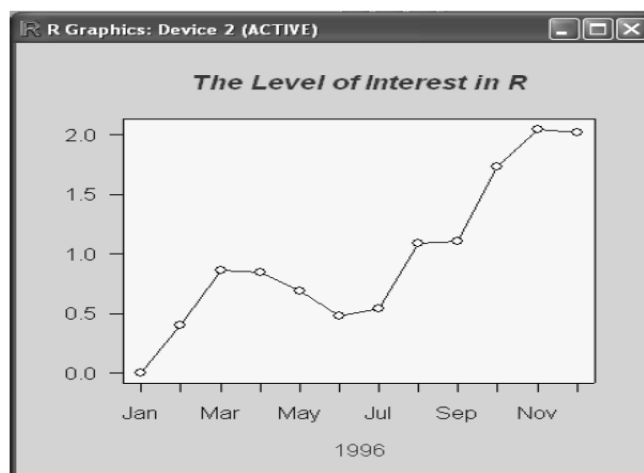
vírgula sempre indica o fim do comando, enquanto uma nova linha pode indicar ou não o fim do mesmo. O comando pode ser associado a um objeto ou não. Quando não estiver associado a um objeto, o valor do comando é retornado imediatamente (FERREIRA, 2008). O desenvolvimento pode ser realizado através de uma janela console e o resultado apresentado em uma janela de saída gráfica, apresentados na Figura 10 e Figura 11, respectivamente. Na janela console podem ser desenvolvidos programas ou utilizados os pacotes disponibilizados no *site* da linguagem para a análise de dados. Os resultados do tipo tabela são apresentados na própria janela console, as saídas gráficas são apresentadas na janela gráfica.



```
R R Console
Hit <Return> to see next plot:
> usr <- par("usr")
> rect(usr[1], usr[3], usr[2], usr[4], col = "cornsilk",
      border = "black")
> lines(x, col = "blue")
```

**Figura 10 – Console de programação em Linguagem R.**

Fonte: (PACHECO, 2005)



**Figura 11 – Console de saída em Linguagem R.**

Fonte: (PACHECO, 2005)

O ambiente de programação R disponibiliza o que existe de mais atual em técnicas de análises estatísticas, visto que disponibiliza um conjunto integrado de facilidades computacionais para manipulação de dados, cálculos e apresentação gráfica.

### **3.4 CONSIDERAÇÕES PARCIAIS**

O desenvolvimento dos métodos científicos para o tratamento de dados e informações permite aos cientistas evoluírem em suas pesquisas de forma dinâmica. A intenção de solucionar problemas com inúmeras variáveis leva a necessidade da adoção de métodos estatísticos, pois segundo (FREUND, SIMON, 2000), a estatística “compreende o manejo dos dados para resumí-los ou descrevê-los, sem ir além, isto é, sem procurar inferir qualquer coisa que ultrapasse os próprios dados”. Partindo desse conceito, é comum encontrar problemas, como incidentes de segurança, que apresentam um conjunto significativo de variáveis a serem tratadas. Nestes casos a estatística multivariada demonstra-se um instrumento útil e adequado para o tratamento eficiente de tais problemas. Esta técnica permite ao pesquisador manipular diversas variáveis simultaneamente com o auxílio de um conjunto de métodos como análise de agrupamentos, fatorial, componentes principais entre outros. Porém é uma técnica complexa que exige uma ferramenta computacional de apoio, além de um conhecimento avançado em estatística.

Como software de apoio para a análise de dados, pode-se adotar a Linguagem R, uma ferramenta poderosa no que diz respeito à manipulação, visualização e análise dos dados. As técnicas existentes atualmente para manipulação e estudo de informações quantitativas, e também as que estão em desenvolvimento, na sua maioria são disponibilizadas pelo R, além do fato que esta linguagem possui muitos recursos gráficos, que permitem gerar e exportar os resultados em diversos formatos.

## CAPITULO 4

### PROJETO DO PORTAL CORPORATIVO E DA FERRAMENTA PARA O SUPORTE A DECISÃO

Para aprender com os incidentes de segurança e alcançar decisões mais eficazes é necessária uma análise estatística do universo das informações geradas pela organização. Este trabalho propôs e desenvolveu um Portal Corporativo, com diversas funcionalidades voltadas para coleta de dados e apoio a decisão, a fim de prover uma ferramenta pró-ativa para a gerência dos incidentes de segurança. O estudo explorou a quantificação dos dados coletados, com a utilização de uma sequência de processos desde a definição dos perfis de usuários, categorização dos eventos e utilização de uma matriz de priorização para possibilitar o estudo estatístico multivariado da amostra dos dados reportados ao sistema. Desta forma, o SDvPC (*Service Desk* via Portal Corporativo) demonstrou-se uma ferramenta modelada para o apoio e tomada de decisões pró-ativas, no que diz respeito à detecção e tratamento dos riscos, ameaças e vulnerabilidades que possam ser impactantes à organização.

Neste capítulo, a seção 4.1 mostra uma breve comparação entre Intranet e Portal Corporativo, sendo este último detalhado na seção 4.2, onde estão descritas as principais funcionalidades agregadas a esta ferramenta. O projeto e a proposta do Portal estão descritos nas seções 4.3 e 4.4, consecutivamente. A seção 4.5 apresenta como as técnicas estatísticas são aplicadas no contexto deste trabalho. As considerações parciais encerram o capítulo com a seção 4.6.

#### 4.1 Intranet X Portal Corporativo

Não existe um padrão, entre os autores, sobre critérios de classificação das intranets e portais corporativos. Alguns classificam essas aplicações de acordo com um padrão evolutivo (1ª, 2ª e 3ª geração) (CHADRAM, 2003), outros classificam pelo uso, funcionalidades e pela ausência ou presença de uma determinada tecnologia. Entretanto, podemos considerar as Intranets como ambientes digitais corporativos, que funcionam como um canal de comunicação entre a empresa e o colaborador, classificadas pela sigla B2E (*business to employee*) onde a informação tem apenas uma direção. Neste sistema a informação está

disponível na intranet e o usuário deve buscá-la, interagindo por meio de acesso a arquivos, lista de ramais, calendário de eventos e etc.

Um portal corporativo é apresentado geralmente como uma evolução da Intranet. Uma definição mais abrangente adotada por (COLLINS, 2003): é “O portal corporativo é uma interface personalizada de recursos on-line que permite aos trabalhadores do conhecimento acessarem informações, tomarem decisões e realizarem ações independentemente de sua localização física, do formato da informação e do local em que ela está armazenada”. Segundo estas afirmações, um portal corporativo é uma aplicação que agrega conteúdo, colaboração, conhecimento e aplicativos transacionais, todos numa interface única de forma a estar em sintonia com os objetivos do negócio da empresa. Existem algumas funcionalidades básicas que precisam estar presentes em um portal corporativo, das quais destacam-se a integração de seus módulos, gestão de conteúdo, segurança e recursos de administração dos sistemas. Dentre essas a integração é vista como uma funcionalidade padrão, que vai servir de base para os outros componentes, sem ele o portal se comportaria apenas como uma fachada bonita para uma estrutura de informação desorganizada (FIRESTONE, 2003). Os portais corporativos podem ser vistos como uma evolução do uso das intranets, conforme mostra a Figura 12, onde estão descritas as 3 gerações de Intranets. Na primeira tem-se como principal característica o uso de conteúdos estáticos. A segunda geração é caracterizada pela grande quantidade de páginas web, pouco padronizadas, porém com recursos dinâmicos. O Portal Corporativo se destaca como a terceira geração, com um conjunto de atributos relacionados a interação homem-sistema.

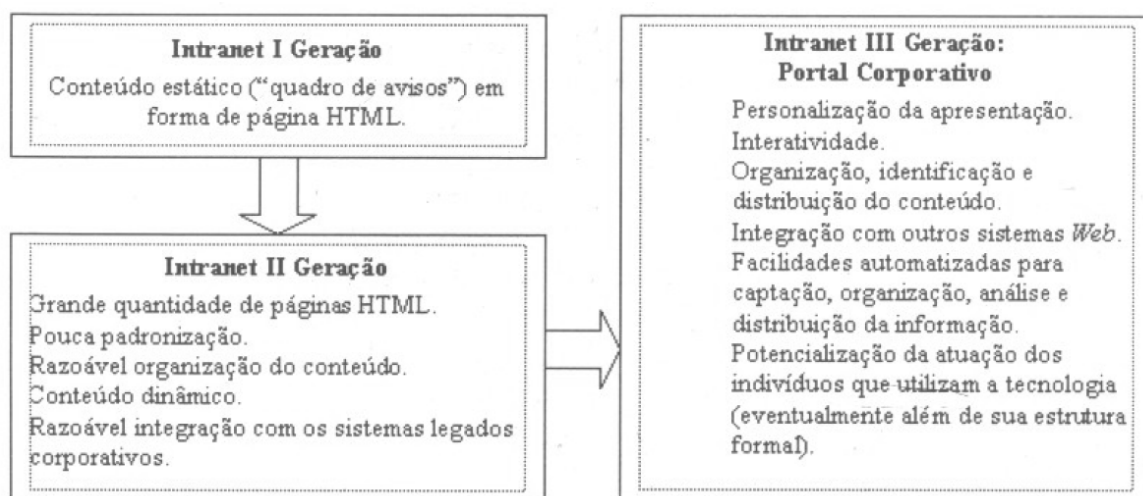


Figura 12 – Transição do acesso a intranet para os portais corporativos.

Fonte: (FERRO;CAMILA, 2002)

Com base nos estudos sobre intranets e Portais Corporativos, e considerando as características das ferramentas e o problema de pesquisa a ser resolvido, optou-se pela adoção da tecnologia de Portal para servir de base para a implementação do sistema de SD. No capítulo 4.2 são apresentados alguns aspectos importantes sobre Portais Corporativos, as quais fundamentam a sua escolha.

## **4.2 Portal Corporativo**

Na literatura especializada encontram-se vários termos similares para este tipo de ferramenta, como: para a definição de um mesmo sistema, como portal de informações corporativas, portal de negócios, portal de informações empresariais e portal corporativo. Para (FIRESTONE, 2000) o processo de definição do portal corporativo, como qualquer outra estratégia de negócios, é um processo político, isto é, uma tentativa de persuadir os usuários e os investidores da área de tecnologia da informação que uma definição é mais adequada do que outra, favorecendo os interesses de um ou outro fornecedor/consultor. Esta ferramenta de comunicação aplicada sobre os recursos de uma intranet permite à disponibilização de um ponto de acesso único as informações e recursos no âmbito interno de uma organização, se tornando desta maneira o mais importante projeto de gestão da informação na empresa (COLLINS, 1999). Percebe-se importância estratégica dos portais corporativos vinculados a área de TI (CAMPOS, 2001), pois estes são instrumentos práticos e abrangentes para a disponibilização de informações e serviços disponibilizando uma infraestrutura voltada pontualmente para as necessidades dos usuários. Segundo (REYNOLDS; KOULOPOULOS, 2000) um Portal Corporativo pode ser visto como um sistema de informações centrado no usuário, integrando e divulgando conhecimentos e experiências de indivíduos e equipes, atendendo, assim, aos padrões atuais de instituições baseadas no conhecimento.

### **4.2.1 Tecnologias e Funcionalidades do Portal Corporativo**

O portal corporativo não é uma única tecnologia ou sistema, mas um conjunto de tecnologias, que se corretamente integradas, provêem uma única interface ao usuário para acessar qualquer recurso de informação e de processos de negócio. As funcionalidades do portal e sua utilidade são moldadas em função de sua arquitetura e dos serviços disponibilizados por essa arquitetura. É, portanto, crítico à construção de portais corporativos,

que se entenda a natureza distinta de cada componente do modelo de arquitetura de portal, bem como o modo no qual cada componente se relaciona aos outros (TOLEDO, 2002). A necessidade de uma ferramenta desse porte nas organizações atuais é eminente, considerando que um portal corporativo é capaz de aliar o conhecimento explícito contido em arquivos, bases de dados, correspondências, páginas *web* e aplicativos empresariais ao conhecimento tácito dos times de projeto, das heurísticas profissionais e das comunidades institucionais (REYNOLDS; KOULOPOULOS, 2000).

Considerando a base de conhecimento gerada a partir da utilização de portais corporativos, deduz-se que é totalmente viável o armazenamento de informações referentes à área de segurança da informação nestes bancos de dados. Alguns *frameworks* para a governança de TI, como ITIL e COBIT, defendem a utilização de ferramentas como portais corporativos para a coleta de incidentes, devido à característica de serem pontos únicos de contato entre as diversas equipes de trabalho em uma organização. A partir dessa visão, um novo conceito para portal corporativo é implementado, onde além de servir como meio de comunicação, o mesmo pode armazenar informações pertinentes ao apoio a decisão pelos níveis gerenciais das empresas. A utilização de portais pode aumentar as oportunidades de mercado para as áreas de segurança, gerência de conteúdo, serviços, banco de dados e *Business Intelligence* (LYNCH, 1998).

O ITIL defende a utilização de ferramentas como portais para servirem de SD nas organizações. Estas aplicações recebem todas as informações pertinentes a reportes de incidentes de segurança, internos e externos, que possam de alguma maneira afetar o bom andamento dos serviços de TI. Contudo a criação dessa base de dados insere um problema básico gerado pelo fato do conjunto de informações armazenadas induzirem a necessidade de formas especiais de tratamento, visto o alto número de variáveis agregadas aos relatos de incidentes. A análise dos requisitos descritos até este momento, são fundamentais para a definição de quais serviços devem ser oferecidos pelo portal corporativo, assim como a definição das ferramentas de desenvolvimento que devem ser usadas para a sua implementação, e principalmente deve ficar claro qual a melhor estratégia científica utilizada para o tratamento dessa base multivariada de informações, de forma que a mesma se torne uma fonte rica em conhecimento agregando valor a tomada de decisão. Além disso (ECKERSON, 1999) define mais alguns requisitos que devem ser respeitados na implementação de um portal corporativo:

- Fácil para usuários eventuais: as informações devem ser de fácil acesso, intuitivas, com o mínimo de treinamento.

- Classificação e pesquisa intuitiva: as informações devem ser armazenadas de forma lógica, organizada, indexada, contendo mecanismos de busca eficientes.
- Compartilhamento cooperativo: os usuários devem poder compartilhar informações com todos integrantes da organização. Deve existir uma ampla interação entre usuários e grupos.
- Conectividade universal aos recursos de informações: o portal deve permitir amplo acesso a toda informação armazenada no mesmo, garantindo a compatibilidade com sistemas heterogêneos.
- Arquitetura baseada em servidor: o portal corporativo deve ser capaz de suportar um número elevado de usuários, assim como informações armazenadas em seus repositórios, permitindo sessões concorrentes de acesso, para isso baseando-se em uma estrutura cliente-servidor.
- Segurança: o sistema deve estar provido de padrões mínimos de segurança, isto significa a adoção de técnicas básicas como criptografia, autenticação, firewall entre outras.
- Fácil administração: o sistema deve possuir a capacidade de gerenciar todas as informações armazenadas, além de monitorar e administrar seus recursos de forma dinâmica e centralizada.

Com base nos requisitos descritos, o projeto do portal corporativo deve ser capaz de suprir todas as necessidades identificadas, com o objetivo de servir como uma ferramenta de gestão e apoio a decisão sobre incidentes de segurança.

#### **4.2.1.1 Controle de usuários**

Para utilizar o portal corporativo é necessário um sistema de controle e autorização de acesso integrado a aplicação. A finalidade deste requisito é manter a confidencialidade das informações pertinentes a organização, e que devem ficar expostas no ambiente. É necessário que o Portal disponibilize um formulário onde o usuário possa acessar o ambiente e inserir seus dados. As informações obrigatórias para o cadastro devem ser obtidas por meio de documentos que normatizem os procedimentos na instituição.

Existem informações que são inseridas de forma automática no sistema, como: permissão (por padrão todos os usuários cadastrados tem o nível de acesso mais baixo) e data de cadastro. Cada usuário recebe um nível de prioridade que varia de 1 a 5, onde 1 significa



que tem a prioridade mais alta e 5, que tem a prioridade mais baixa. Essa prioridade vai depender do tipo de função desempenhada pelo usuário, por exemplo, se ele for da direção da instituição recebe prioridade 1, se for estudante recebe prioridade 5. Esse dado é utilizado pela aplicação de SD. Quando um usuário faz seu cadastro, ele precisa aguardar o mesmo ser aceito pelo administrador do sistema, pois é na hora de validar o cadastro que o administrador atribui a prioridade e nível de acesso do usuário.

#### **4.2.1.2 Controle de materiais**

O controle de materiais é um requisito de suma importância a análise e implementação do portal. Os materiais devem ser categorizados em itens de consumo (placa de rede, periféricos, memória RAM e etc.) e etiquetados de acordo com a padronização de patrimônio da organização.

A movimentação de material de consumo ou de patrimônio é realizada apenas sobre itens cadastrados no sistema. Este cadastro deve ocorrer durante o atendimento de uma chamada ao Serviço de Atendimento. Cada patrimônio e material de consumo cadastrado devem possuir um estado (*status*) associado, que é utilizado para saber o estado em que se encontra o item (danificado ou não). É necessário que a aplicação seja capaz de controlar a quantidade de materiais de consumo existentes no Centro e sua localização, além de viabilizar o rastreamento da mesma.

#### **4.2.1.3 Controle de rede**

O Portal deve manter um registro (*log*) de todos os endereços IP dos computadores da rede que acessarem o sistema. Também deve armazenar todos os pontos de rede e suas respectivas salas. Essa ferramenta deve permitir um controle de todas as máquinas que estão em funcionamento no prédio, explicitando os endereços lógicos dos hosts na rede, assim como a qual ponto físico este pertence. O intuito disso é a identificação, por exemplo, de um usuário e equipamento que estão causando ou tendo problema.

#### **4.2.1.4 Informativos gerais**

A página principal do Portal (*home Page*) deve disponibilizar informações de interesse dos usuários da instituição, tais como: notícias, eventos, comunicados, lista de ramais, álbum de fotos, avisos de viagens da instituição disponíveis para transporte, informações sobre o

clima da região, manuais de normas, entre outras. Essa ferramenta deve servir como um canal para difusão de informações entre os colaboradores.

#### **4.2.1.5 Controle de usuários online e colaboradores**

O Portal também deve permitir um controle dos usuários que estão autenticados no sistema, e com isso manter registros de acesso e gerar dados estatísticos que permitam a verificação do uso deste pelos colaboradores, provendo assim uma comunicação mais eficiente entre todos. Uma solução a ser agregada as funcionalidades do sistema é a capacidade de informar quais usuários encontram-se nas instalações da organização. A integração do Portal Corporativo a um sistema de controle de acesso, o qual esta em desenvolvimento, possibilitará esta funcionalidade tendo em vista que o sistema de controle de acesso irá utilizar a base de usuários do portal para o controle de acesso ao prédio.

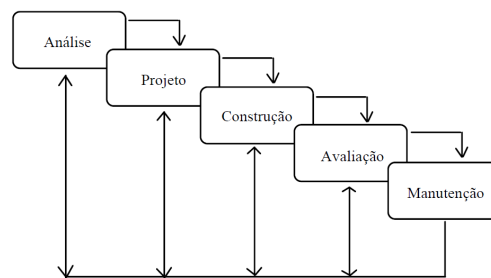
#### **4.2.1.6 Base de conhecimento**

A partir da base de conhecimento gerada pelos dados armazenados no repositório do Portal, é possível obter o controle total de todos os processos organizacionais. Isto permitirá um controle total de todos os materiais usados dentro da instituição, emissão de relatórios com as atividades do grupo de suporte em um determinado período de tempo, além de gerar estatísticas relacionadas aos incidentes registrados.

### **4.3 Projeto do Portal Corporativo**

O processo de desenvolvimento de software contempla um conjunto de atividades realizadas para construir a aplicação, levando em consideração os produtos sendo construídos, as pessoas envolvidas e as ferramentas com as quais trabalham. Com esse intuito várias tecnologias vêm sendo experimentadas dentro do contexto de apoiar o ciclo de vida para implementação de software. Para que o projeto seja bem sucedido é necessária a definição de uma metodologia voltada a disciplinar o processo de desenvolvimento através do estabelecimento de etapas bem definidas, proporcionando, desta forma, um mecanismo de controle para este processo. Para (THIVES, 2000), processo “*é um conjunto de atividades estruturadas e de medidas destinadas a resultar num produto especificado para um determinado cliente ou mercado*”.

Baseado na afirmação de (COSTA, 1999) um modelo de processo é uma coleção de asserções, estratégias, atividades, métodos e tarefas que estão organizados para atingir um conjunto de metas e objetivos. O modelo é importante porque fornece a base para a resposta às questões críticas de planejamento, autoridade, predição e rastreamento de como um resultado é obtido. Neste contexto o desenvolvimento do projeto para o portal corporativo, considerando as peculiaridades da integração com um SD e os demais aspectos gerenciais necessários a uma organização, deve seguir um conjunto de atividades planejadas. Para alcançar estes objetivos, foi seguido um modelo baseado no ciclo de vida clássico, reconhecido por (PRESSMANN, 1995; SOMMERVILLE, 1992; GHEZZI, 1991; YOURDON, 1992). A Figura 13 demonstra as fases que compõem esse ciclo de desenvolvimento, seguindo um modelo em cascata, onde é requerida uma abordagem seqüencial e sistemática, iniciando pela Análise e avançando pelo Projeto, Construção, Avaliação e Manutenção (GHEZZI, 1991; PRESSMAN, 1992).



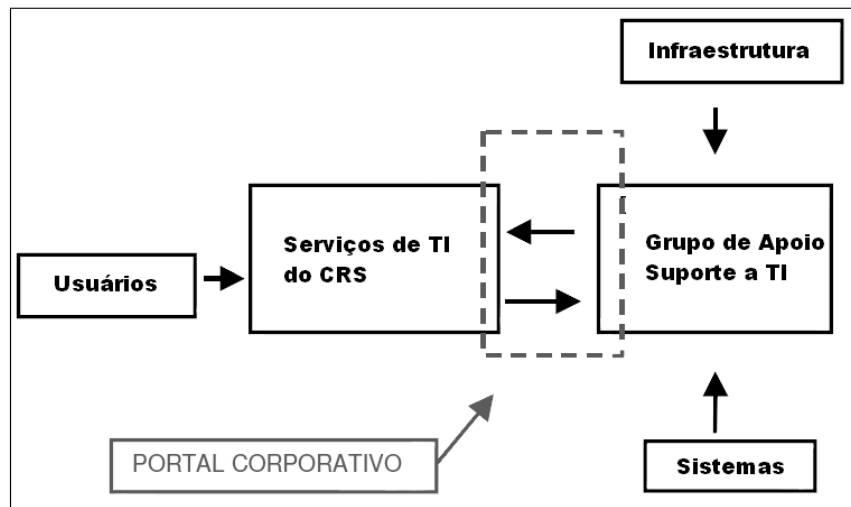
**Figura 13 – Modelo do Ciclo de Vida Clássico (Cascata)**

**Fonte: (PRESSMANN, 1995)**

As fases que compõem o modelo em cascata podem ser descritas da seguinte forma:

- **Modelagem do Sistema:** onde são estabelecidos os requisitos do sistema ao qual o software está sendo realizado, incluindo os requisitos de informação e de negócios. Nesta fase deve ocorrer o alinhamento dos serviços de TI com as necessidades atuais e futuras do negócio. O fato da grande dependência gerada pela área de TI reflete diretamente sobre os custos e riscos nas organizações, com isso torna-se muito importante a justificativa para o retorno sobre investimento (ROI) em TI. Para alcançar um ROI satisfatório é necessário que todo e qualquer incidente de segurança no âmbito da organização seja conhecido e tratado de forma pró-ativa ou no mínimo dentro dos SLAs estabelecidos. Um sistema de apoio a decisão, que auxilie a gerencia de TI em suas atividades e que também sirva de ponto de contato único entre os usuários deve ser modelado. Com base no exposto, a adoção da tecnologia de Portal

Corporativo integrado a uma ferramenta de apoio a decisão, utilizando métodos estatísticos multivariados, se demonstra uma solução inteligente. O SDvPC deve ser modelado de forma que todos integrantes da organização tenham acesso ao mesmo, com uma interface simples e acessível permita os usuários reportarem todos incidentes identificados. A Figura 14 mostra dentro da estrutura de negócio, onde deve atuar o Portal Corporativo, identificando a sua posição de interligação entre serviços de TI e as equipes que apóiam essa área.



**Figura 14 – Inserção do Portal Corporativo na estrutura do CRS.**

Fonte: Adaptação de (LEME, 2005)

- Análise de requisitos: onde são modelados os requisitos de informação, funcionais, comportamentais, de desempenho e de interface do software. O levantar requisitos significa fazer cortes no espaço real de solução para delimitar soluções aceitáveis, buscando um nível satisfatório de qualidade. Esta atividade deve permitir uma visão geral do sistema, dos processos de negócio e do conjunto de requisitos coletados a partir do conhecimento dos usuários e dos resultados de suas ações.
- Projeto: onde são planejadas as estruturas de dados, a arquitetura do sistema e os comportamentos são mapeados em procedimentos;
- Codificação: onde o projeto é transformado em uma linguagem compreendida pelo computador;
- Testes: onde verificamos e validamos o software;
- Manutenção: onde garantimos a usabilidade do software

#### 4.4 Proposta do *Service Desk* via Portal Corporativo (SDvPC)

A excessiva demanda das empresas em manter seus sistemas funcionais, acarreta a criação de inúmeros chamados (incidentes) que devem ser tratados pelas equipes de TI. Atualmente as empresas utilizam sistemas chamados de *Service Desk*, os quais são softwares que tem por objetivo criar um ponto único para abertura e registro de chamados de serviço. Também chamado de ponto único de contato (SPOC – *Single Point of Contact*), o SD auxilia as equipes de suporte a coordenar e sanar problemas ocorridos no âmbito da empresa, pois constitui um mecanismo computacional facilitador da informação organizacional (CAVALARI e COSTA, 2005).

Várias metodologias podem ser adotadas para a coleta de chamados dos usuários, dentre as quais, e-mail, telefone, mensagens instantâneas, VoIP, fax, atendimento pessoal e fóruns. Porém é importante salientar que para todos esses métodos é necessária a intervenção humana de um operador para o lançamento do chamado de serviço, o que demanda custos para a organização além de aumentar a probabilidade de erros devido à imparcialidade do operador e problemas de comunicação com o usuário.

Para minimizar a interação com o operador do SD, no momento dos registros de incidentes identificados por usuários, este trabalho propõe uma ferramenta que integra o sistema de SD a um Portal Corporativo, o qual é a principal infra-estrutura tecnológica para a gestão do conhecimento em uma empresa segundo (CARVALHO, FERREIRA e CHOO, 2007).

O Portal Corporativo é uma interface personalizada de recursos *on-line* que permite que os trabalhadores do conhecimento acessem e compartilhem informações, tomem decisões e realizem ações independentes da sua localização física, do formato da informação e do local que ela está armazenada (COLLINS, 2003). Deste modo, a integração do SD com o portal corporativo permite a todos integrantes do sistema lançar seus reportes diretamente, sem a intervenção de terceiros, o que garante mais integridade nas informações repassadas.

##### 4.4.1 Módulos do SDvPC

O SDvPC é uma de aplicação que agrega vários módulos independentes, integrados a partir da necessidade de atendimento às solicitações dos usuários que utilizam os recursos de TI na organização. A principal característica do SDvPC é realizar a gestão dos usuários e dos

seus direitos (autenticação, autorização, auditoria) definindo níveis de prioridade para cada um. Para que esses níveis sejam definidos de forma clara e precisa, alguns dados devem ser informados durante o cadastramento dos usuários no SDvPC como, por exemplo, atividade, grupo de trabalho, área de atuação e responsabilidades. Esse conjunto de informações possibilita identificar o grau de relevância do usuário através de uma Escala *Likert*, técnica flexível que permite a inferência estatística de dados e não interfere na interpretação de médias baseadas em intervalos variáveis (GÜNTER, 2003). Com uma escala de cinco pontos, 1 é o valor com maior relevância, e 5 o de menor relevância.

Para analisar de uma maneira ampla onde está concentrado o maior número de problemas atendidos pelo sistema, além da classificação por perfil, os usuários podem ser agrupados de acordo com alguns critérios predefinidos como, por exemplo, grupo de trabalho, setor e ocupação. O foco do SDvPC é manter um acordo sobre o nível de serviço - SLA, que é um acordo entre o provedor de serviços e seus clientes para estabelecer a qualidade mínima de serviço que a empresa necessita (HILES, 1994). No caso da gestão de incidentes, o SLA corresponde aos níveis de serviço combinados para, por exemplo, tempo de resposta, disponibilidade, continuidade e interação do usuário. Deste modo, ao definir pontos de controle garante-se que o departamento de TI terá o compromisso de prover um serviço orientado ao cliente, e para tal o gerenciamento do nível de serviços é essencial (ACADEMY, 2003).

#### **4.4.2 Funcionalidades e Requisitos do SDvPC**

Para que sejam alcançados os objetivos desta pesquisa, o SDvPC deve ser capaz de armazenar e tratar um quantidade substancial de informações sobre incidentes de segurança reportados na instituição, além de manipular de forma eficiente as amostras de dados utilizadas durante a análise estatística destes incidentes. Com base nisso, (LEEDY, 1989) afirma que “a natureza dos dados governa o método que é apropriado para interpretar os dados e a ferramenta de pesquisa que é necessária para processar aqueles dados”. Portanto para que a análise dos dados coletados possa sofrer a inferência estatística desejada, é necessária a conversão destas informações em dados quantitativos. Esta seção explica como esta conversão é realizada no SDvPC.

#### 4.4.3 Quantificação das Informações

Os reportes de incidentes devem ser realizados pelos próprios usuários sem a intervenção do suporte de TI, e a informação repassada normalmente possui um aspecto descritivo, para que possa auxiliar na identificação e solução rápida do problema. Porém o número elevado de situações heterogêneas reportadas impõe a necessidade da coleta de uma gama razoável de informações descritivas para compreender e explorar os dados de forma qualitativa (KRUSKAL e WISH, 1978). Uma abordagem quantitativa é mais científica e recomendada para ser utilizada como um procedimento de confirmação de hipóteses, de forma a ganhar força de argumento e qualidade nas conclusões elaboradas (FREITAS e MUNIZ, 2002). Deste modo, a abordagem descritiva adotada na interface do SDvPC visa obter na íntegra a percepção e cognição dos usuários sobre todos os aspectos pertinentes aos problemas identificados, mas sua quantificação e tratamento por meio da inferência estatística visa desprender idéias e ações, com o intuito de produzir uma informação mais rica para o tratamento de incidentes.

Para realizar a quantificação das informações, o SDvPC segue as etapas apresentadas na Figura 15. Um usuário previamente cadastrado e classificado num perfil de 1 a 5 realiza um reporte de incidente. Este reporte é rotulado como um TICKET, o qual é classificado e armazenado em grupos por afinidade (tipo de incidente).

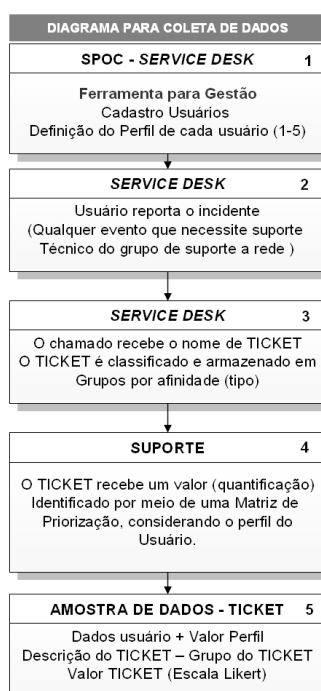


Figura 15 – Etapas do funcionamento do SDvPC.

Para realizar a quantificação do reporte é adotada uma estratégia baseada em uma Matriz de Prioridades que, segundo (BRASSAND, 1989), objetiva determinar as características prioritárias, sendo, acima de tudo, uma técnica de decisão que toma por base critérios com pesos predefinidos. As etapas de construção da Matriz de Prioridades variam fortemente em função do grau de complexidade do problema, e do tempo disponível. Para este estudo a Matriz de Prioridade esta descrita na Figura 16. As colunas que formam o item “impacto” definem o grau de criticidade para o negócio, enquanto as linhas definidas por “urgência” sugerem a velocidade que o incidente deve ser resolvido.

		Impacto		
		Alto	Médio	Baixo
Urgência	Alta	1	2	3
	Média	2	3	4
	Baixa	3	4	5

**Figura 16 – Matriz de Prioridade – “Impacto vs. Urgência”.**

**Fonte: ITIL v.2**

De acordo com o valor obtido com a quantificação de cada reporte pela Matriz de Prioridade “Impacto vs. Urgência”, resultante de uma escala de cinco pontos, a prioridade de resposta da equipe de TI e o tempo de atendimento ao usuário podem ser quantificados de acordo com a tabela de prioridade do ITIL versão 2, Figura 17. Esta tabela mostra o SLA desejável para o atendimento dos incidentes relatados, pois a equipe de TI deve trabalhar de forma organizada e orientada a prioridade, que por sua vez define a ordem de trabalho e a agilidade nos serviços de suporte.

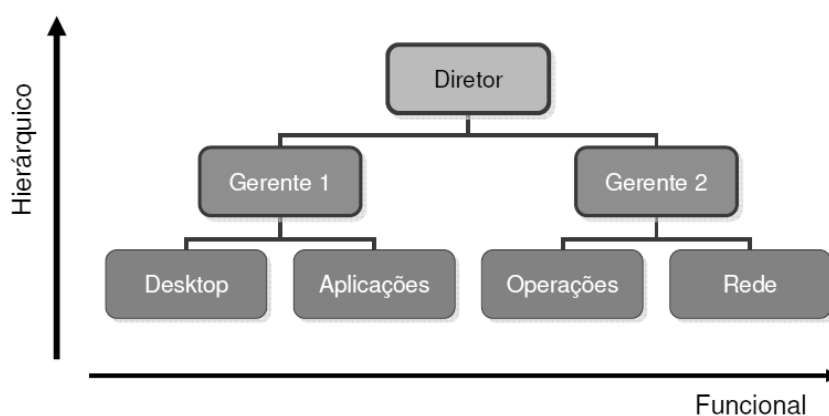
Prioridade	Descrição	Tempo para atendimento
1	Crítica	1 hora
2	Alta	4 horas
3	Média	24 horas
4	Baixa	48 horas
5	Planejada	-

**Figura 17 – Tabela de Prioridades.**

**Fonte: ITIL v.2**



No momento da identificação dos incidentes reportados, o responsável pelo suporte técnico deve considerar as informações recebidas e também avaliar o perfil do usuário que realizou o reporte. O perfil do usuário nunca pode ser maior que o valor atribuído pela escala ao seu ticket. A Figura 19 apresenta um exemplo de escalonamento em profundidade para um incidente, onde se define perfis e respeita-se a hierarquia e funcionalidade dos integrantes da organização. De acordo com a Figura 18, um Diretor sempre terá prioridade maior do que os Gerentes, e estes terão prioridade maior que o pessoal da área de operações e assim sucessivamente.



**Figura 18 – nível de profundidade “hierarquia” ou funcionalidade a qual deve atender um determinado incidente.**

Fonte: ITIL v.2

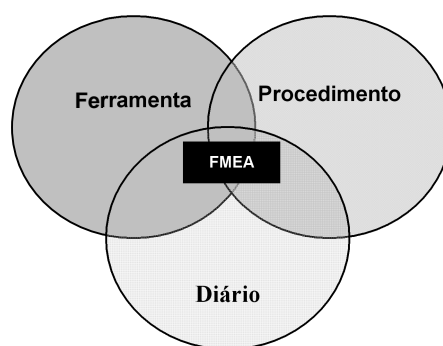
#### 4.4.4 Categorização

O objetivo primordial da categorização é obter informações gerenciais para o negócio a partir de um determinado conjunto de dados que podem ser chamados de dimensões. Estes dados devem fornecer informações relevantes para distinguir as categorias dos principais problemas, identificar gargalos, encontrar ações corretivas, tudo isso com o intuito de buscar melhoria contínua, reduzir custos e aumentar a qualidade dos serviços.

Para a atividade de categorização foi utilizada uma metodologia de Análise do Tipo Falha e Efeito, conhecida como FMEA (*Failure Mode and Effect Analysis*) (Chang *et al.* 2001). A utilização do FMEA é sugerida para a identificação dos processos que estão ocasionando os incidentes e as atividades relacionadas. Esta técnica utiliza um método sistemático com três

funções simples, conforme a Figura 19: a ferramenta, técnica de baixo risco com o intuito de prevenir problemas e identificar soluções eficazes; o Procedimento, que oferece uma abordagem estruturada para a condução e avaliação das atividades, podendo ser utilizado também para associar e manter os documentos de uma organização; e Diário, iniciado na concepção do projeto e persiste durante o ciclo de vida do sistema, garantindo assim que qualquer alteração que venha a ser feita, deve inicialmente ser avaliada e documentada pelo FMEA.

As informações retiradas da ferramenta dão origem às categorias que são utilizadas pelo SDvPC. Posteriormente são avaliados os riscos de cada reporte relatado por meio de índices e, com base nestes valores, são tomadas as ações necessárias para diminuir os riscos e reavaliar as categorias implementadas.



**Figura 19 – O FMEA é o Resultado do Relacionamento entre as Funções: Ferramenta, Procedimento e Diário.**

**Fonte: CASTILHO, 2007**

Como exemplo da utilização do FMEA pode-se definir uma situação pontual: a grande demanda de um incidente específico, que está sendo reportado dentro da categoria “outros incidentes”. O fato deste problema não ser relevante até o momento o descaracterizou, porém a sua elevada ocorrência tende a modificar a mostra de dados. Neste caso a análise dos reportes com base em procedimentos diários permite a identificação deste tipo de problema e a sua retificação, de maneira que o mesmo não afete a amostra utilizada na análise estatística.

#### **4.5 Análise Estatística**

Segundo Mingoti (2005) a análise estatística multivariada é dividida em dois grupos: um primeiro, basicamente aplicado neste estudo, que consiste em técnicas exploratórias de

simplificação da estrutura da variabilidade dos dados, possuindo um apelo prático muito interessante, pois, independem em sua maioria do conhecimento da forma matemática da distribuição de probabilidades geradora dos dados amostrais. Em linhas gerais o uso da multivariada neste trabalho busca simplificar ou facilitar a interpretação de fenômenos específicos, através de métodos chamados de multidimensionais que permitem a confrontação entre duas ou mais variáveis. O intuito é extrair as tendências mais sobressalentes e hierarquizá-las, eliminando os efeitos que perturbam a percepção global.

Com base no contexto apresentado, as informações utilizadas neste trabalho são compostas por um conjunto de reportes coletados através do SDvPC durante a sua implantação. Estes reportes constituem solicitações reais, informados por integrantes do CRS no período de março a abril de 2009, as quais foram atendidas pela equipe de suporte. Após a quantificação das informações expostas pelos usuários, estas são armazenadas em uma base de dados e estratificadas como amostra para análise estatística, conforme a Figura 20. As variáveis apresentadas abaixo são: código e nome do usuário, perfil do usuário com grau de importância, grupo de incidentes (categorização dos incidentes reportados), valor do ticket que define a prioridade de atendimento ao chamado de acordo com a SLA definida na organização, tempo previsto e efetivo do atendimento.

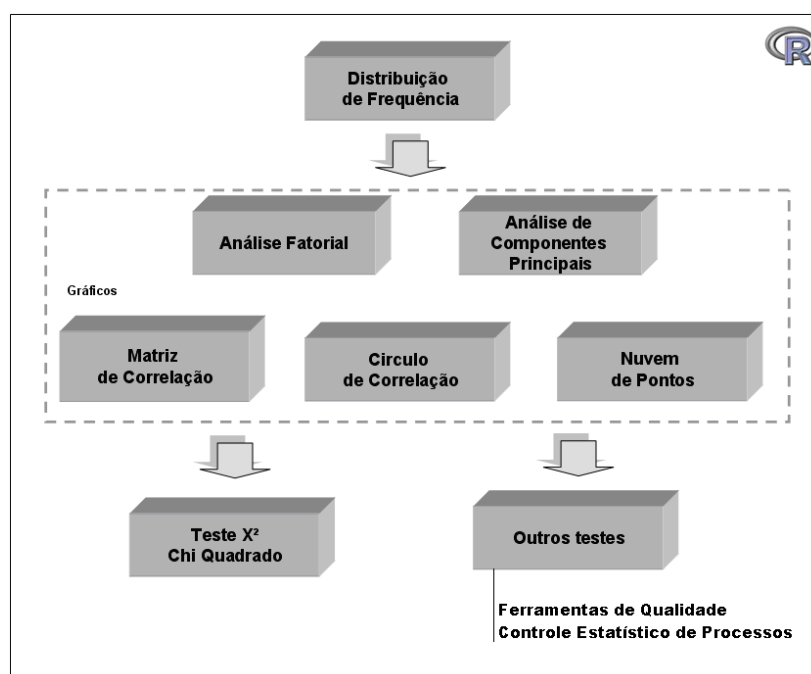
Código usuário	Nome Usuário	Perfil usuário	Grupo incidente	Valor TICKET	Tempo Previsto Minutos	Tempo Atendimento Minutos
156	João Silva	3	6	2	240	120
248	Marcos Hubert	5	1	5	4320	80
87	Marcia Santos	1	5	1	60	20

**Figura 20 – Amostra de dados disponível para análise estatística.**

Para a adequação e validação dos dados importados da base, inicialmente deve ser realizada uma análise estatística descritiva para identificar a incidência e o comportamento destes registros. Conjuntamente com a análise descritiva é interessante a elaboração de um gráfico *box-pot*, que é uma representação gráfica de valores amostrais, com o objetivo de evidenciar a distribuição das ocorrências e suas amplitudes.

Na seqüência devem ser aplicadas técnicas de Análise Fatorial e Análise de Componentes Principais, pois a análise conjunta das variáveis através de procedimentos multivariados permite inferir o inter-relacionamento das variáveis em estudo. Como resultados destes processos são apresentados os gráficos de Matriz de Correlação, Circulo de Correlação e Nuvem de Pontos, por meio dos quais é possível verificar visualmente a relação das variáveis da amostra estudada.

O processamento dos dados, a geração dos resultados e os gráficos obtidos nestas análises foram efetuados inicialmente com o auxílio do software de estatística, STATISTICA 7.0 e da planilha Excel 2003 que de maneira simples permitiram validar a inferência, posteriormente foram utilizadas funções em R, visando à automação destes processos. A Figura 21 mostra graficamente os passos utilizados pelo SDvPC para a inferência estatística.



**Figura 21 – Organização da inferência estatística.**

Ao findar a análise podem ser aplicados alguns testes de validação. Neste trabalho foram adotados o teste do qui-quadrado, que segundo Downing (2006), verifica se há diferença significativa entre duas ou mais variáveis, ou inferir se as variáveis são independentes. Ele auxilia na verificação da multinormalidade e na identificação de possíveis *outliers*.

#### 4.6 CONSIDERAÇÕES PARCIAIS

O conjunto de funcionalidades agregadas ao Portal Corporativo o torna uma eficiente ferramenta para a comunicação interna e interação de processos de controle e de usuários no ambiente da rede. Com base nesse conjunto de possibilidades a disponibilização de um *Service Desk* integrado as ferramentas do Portal, torna-se uma opção viável para a difusão de um método eficaz de reporte dos incidentes de segurança identificados pelos usuários. O *Service Desk* serve também como um canal único de comunicação entre a equipe de suporte e os integrantes da organização.

Esta aplicação como um todo permite o sistema coletar, processar, quantificar e armazenar um conjunto de informações relevantes sobre os problemas identificados e reportados na área de TI da organização, os quais compõem uma base de conhecimento sobre a qual é possível proceder um conjunto de análises estatísticas a fim de obter informações relevantes para o apoio a decisão na área de TI. A automação destas análises possibilita a área de gerência levantar um conjunto de dados importantes, os quais servem de alicerce para toda decisão relacionada a administração dos sistemas computacionais no âmbito da organização.

## CAPITULO 5

### IMPLEMENTAÇÃO, TESTES E DISCUSSÕES

Este capítulo apresenta em uma primeira etapa o Centro Regional Sul de Pesquisas Espaciais, o qual serviu como laboratório para este trabalho. Nas próximas etapas estão descritas a seção 5.1 com o projeto do Portal Corporativo e os processos de análise do SDvPC. A seção 5.2 justifica a utilização do *framework* Django como solução e ferramenta de desenvolvimento, na seção 5.3 é apresentado a implementação do *Service Desk*, enquanto a seção 5.4 demonstra uma discussão sobre os principais resultados obtidos a partir da análise multivariada utilizando os softwares R, Statística, SPSS e Excel. A seção 5.5 descreve um conjunto de análises a fim de fundamentar e comprovar os resultados obtidos através da análise multivariada. E ao final do capítulo é exposto uma conclusão parcial.

Com o intuito de alcançar um nível de pró-atividade na resolução dos incidentes de segurança, identificados em uma organização, este trabalho estudou, avaliou e implementou uma solução em software para tais problemas, afim de tratá-los de forma pontual. A Figura 22 mostra, graficamente, os passos realizados durante a análise da amostra dos *tickets* no *Service Desk* via Portal Corporativo, os quais estão descritos neste capítulo. É possível identificar a ordem cronológica das ações do SDvPC, desde a fase de projeto do sistema até as considerações parciais obtidas. De maneira clara é demonstrado o ponto onde as técnicas multivariadas são aplicadas, as discussões sobre os resultados e a realização de outras análises, as quais baseiam-se na reorganização das informações, por meio de tabelas de frequências, somatórios e interligação dos dados amostrais. Esta fase tem como finalidade comprovar a validade dos métodos descritos e utilizados no Portal.

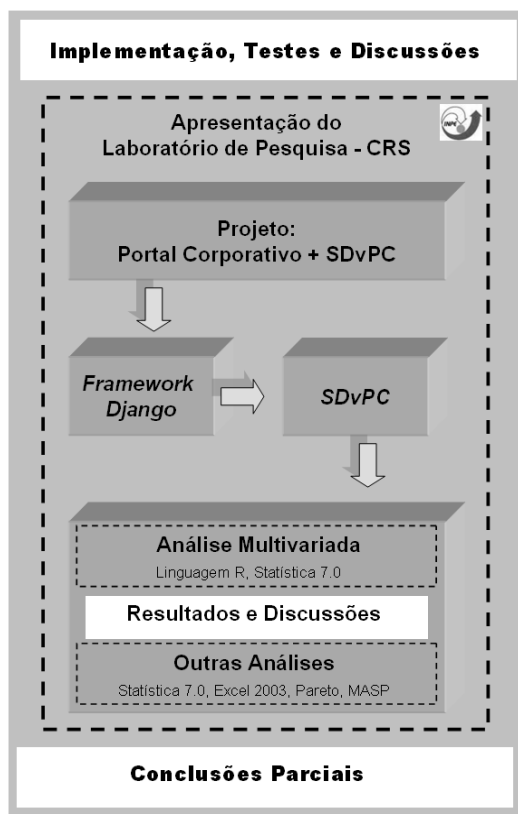


Figura 22 – Organização do capítulo Implementação, Testes e Discussões.

É importante salientar que o uso dos métodos estatísticos confrontados com outras formas de análise possibilita uma correta interpretação dos dados utilizados, permitindo aos analistas de suporte identificar os principais pontos de falha na estrutura computacional da organização e no atendimento aos usuários. Os SLAs definidos pela organização são utilizados como ferramentas métricas para avaliação e definição de tais falhas.

### 5.1 Laboratório de Pesquisa – CRS/INPE - MCT

O Centro Regional Sul de Pesquisas Espaciais (CRS/INPE-MCT) é uma instituição de pesquisa que possui mais de 500 colaboradores entre estudantes, pesquisadores e funcionários. Em suas instalações atuam vários grupos de pesquisa, entre os quais se destacam:

- Projeto Antártico do INPE (PAN);
- NANOSAT - Projeto para construção de nanosatélites;
- CRECTEALC – Centro regional de Educação em Ciência e Tecnologia para América Latina e o Caribe;
- Grupo de Modelagem Atmosférica de Santa Maria (GruMA);

- Grupo Geodesastres Sul;
- Grupo de Gestão e Tecnologia em Segurança da Informação (GTSeg/UFSM);
- Grupo de suporte à Tecnologia de Informação

O Centro ainda conta com um sistema de cabeamento estruturado que permite alcançar alto desempenho sustentado e imune a mudança de tecnologia, onde os meios de transmissão integram infra-estrutura única de fluxo de dados, interconectando à rede de telecomunicações com diferentes tipos de aplicações, tais como: servidores, impressoras, VoIP e uma central telefônica SIP, integrando as centrais telefônicas instaladas em Santa Maria, São Martinho da Serra, São José dos Campos e a Universidade Federal de Santa Maria. O *backbone* da rede de dados e voz opera sob o padrão *Gigabit ethernet*, através de fibra óptica monomodo, com taxas de transmissão ao usuário final de 10/100/1000 Mbps.

O CRS utiliza uma solução Extreme Networks, com 2 equipamentos *Alpine*, modelo 3808 no *backbone*, 14 switches Summit200-48 e um Summit300-48, como equipamentos de borda. O *Summit300-48* é responsável pelo gerenciamento da rede sem fio (*wireless*), que é constituída por 8 estações Altitude, Figura 23, que atendem todo perímetro do Centro.

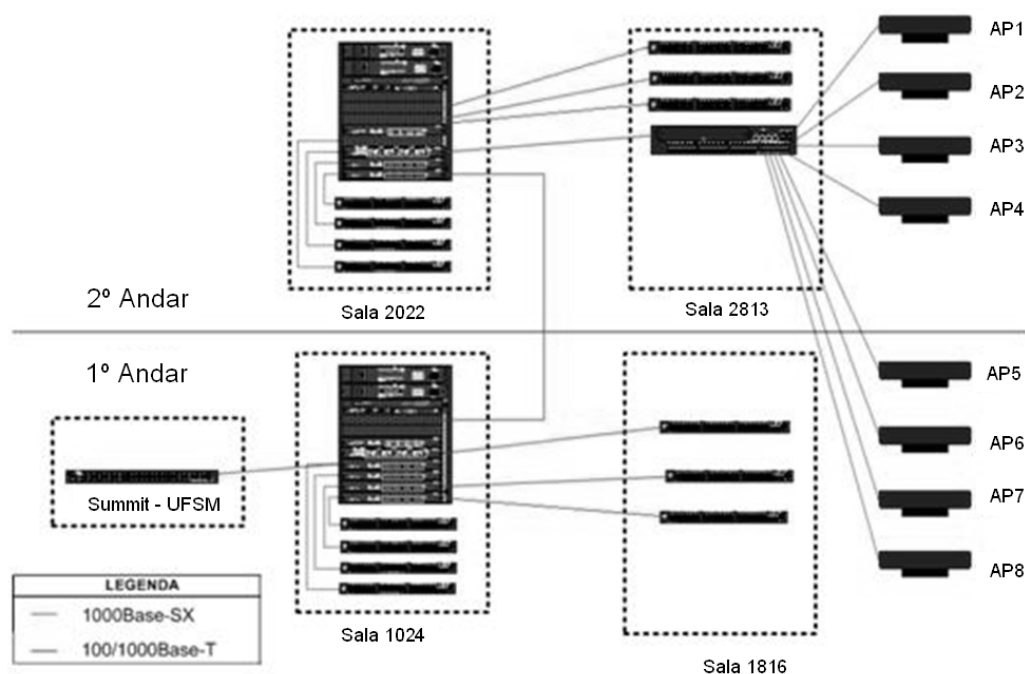


Figura 23 – Console Equipamentos, salas de equipamentos e armários de telecomunicações.

Fonte: (KOITI, 2008)

Os ativos de rede *Alpine* estão dispostos em duas salas de equipamentos e possuem capacidade para rodar protocolos que auxiliam na tolerância à falhas: os protocolos *Extreme*



*Standby Router Protocol* (ESRP), *Ethernet Automatic Protection Switching* (EAPS), *Spanning Tree Protocol* (STP) e extensões do protocolo STP da Extreme Networks, e agregação de links para redundância de Camada 3 e Camada 2 dentro dos uplinks da rede, para garantir que falhas em outros equipamentos não resultem em perda da conectividade para os usuários.

Com este sistema a rede do CRS incorpora alguns serviços de TI, tais como: o serviço de e-mail, que atende o Laboratório de Ciências Espaciais de Santa Maria (LACESM); um servidor de FTP, para repositório de dados, que atende as necessidades de troca de informações entre os orientadores do INPE/UFSM e os alunos orientados no CRS; outro servidor FTP que auxilia as atividades do grupo de suporte; dois clusters utilizados pelo CPTEC (Centro de Previsão de Tempo e Clima) para execução do modelo numérico de previsão de tempo e clima (BRAMS) e do modelo de dispersão de poluentes (L.E.S); e por último um laboratório de computação para a Coordenadoria de Educação a Distância da UFSM, que atende 32 cidades pólos.

### **5.1.1 Monitoramento de Incidentes no âmbito do CRS**

O CRS apresentou características pontuais, as quais o demonstravam como um laboratório ideal para a realização desta pesquisa, tendo em vista o grande fluxo de informações que envolvem dados, voz, imagens geradas no âmbito do Centro, além do expressivo aumento do número de usuários e pesquisadores locados no CRS. Com base nessas informações denotou-se de forma clara a necessidade da implantação de um sistema de gerência e monitoramento, capaz de identificar o elevado número de incidentes computacionais reportados por seus colaboradores.

Dentre os incidentes ocorridos no CRS, com mais frequência, alguns foram definidos como comuns, mas de grande influência no desempenho e funcionamento dos sistemas. Cita-se como exemplo: a não atualização de antivírus e dos sistemas operacionais, que aumentam as vulnerabilidades do sistema e a propagação de *worms*, *trojans* e vírus; a duplicação de IPs e a formação de *bootnet*, que permitem o ataque em massa a servidores na Internet; e tentativas de quebras de senha da rede sem fio, que podem indisponibilizar este meio de acesso. Todos os fatos descritos proporcionam conflitos e inacessibilidade de serviços na rede, o que pode levar ao conseqüente bloqueio de toda comunicação do Centro. Com menos intensidade, tem-se ainda os incidentes ocasionados pela utilização de programas P2P, para o compartilhamento de software, os quais demandam grande taxa de tráfego e causam

sobrecarga e lentidão da rede, além de disponibilizar portas de comunicação abertas, que podem ser exploradas.

## 5.2 Projeto do Portal Corporativo e do *Service Desk*

Para o desenvolvimento do projeto optou-se pelo paradigma de programação orientado a objetos, pois ela proporciona um nível de abstração na resolução de problemas, reaproveitamento de código e a utilização de um padrão conceitual durante todo o desenvolvimento do projeto (SILVA, 1996).

Então para implementar o sistema optou-se pela utilização de um *framework* para desenvolvimento web, o qual permite concentrar-se com mais abstração nos problemas análise tornando mais eficiente a resolução de problemas e por consequência com mais produtividade. Dentre os vários *frameworks* existentes foram escolhidos dois para análise: *Ruby on Rails* e Django.

*Ruby on Rails* (URUBATAN, 2009) é um *framework* que tem como base a linguagem Ruby, ele utiliza o conceito de DRY (*Don't Repeat Yourself*) que é o conceito de definir nomes, propriedades e código em um único local. Ele utiliza mapeamento objeto-relacional (ou ORM) onde as tabelas do banco de dados são representadas por classes e os campos por atributos, assim o programador não precisa trabalhar diretamente com os comandos da linguagem SQL, utilizando uma interface orientada a objetos que faz todo o trabalho de persistência (*Ruby Project*). Ele possui um bom conjunto de ferramentas, dentre elas a capacidade de evolução do banco de dados onde é possível fazer mudanças nas tabelas do banco e migrá-las para a nova versão ou para outra plataforma. Também tem grande quantidade de documentação. Porém o *Ruby on Rails* apresenta como aspectos negativos a não sincronia com bancos de dados legados existentes além de não gerar interfaces administrativas automaticamente.

O *framework* Django tem como base a linguagem de programação Python, que possui uma grande versatilidade, tem uma biblioteca padrão que inclui módulos de processamento de texto, expressões regulares, protocolos de rede, criptografia, interface gráfica (ANDERSON, 2007). Além de ser uma linguagem fácil de aprender, pois possui uma sintaxe simples o que ajuda também na manutenção do código. O Django trabalha com DRY e possui um ORM próprio. Um dos grandes diferenciais desse *framework* é a possibilidade de gerar uma interface administrativa automática que possibilita maior produtividade e agilidade na fase de desenvolvimento.

Tendo como base esta análise, optou-se por utilizar o Django, principalmente pela versatilidade e fácil aprendizado da linguagem Python e pela geração da administração automática.

### 5.2.1 Análise do Portal Corporativo e do *Service Desk*

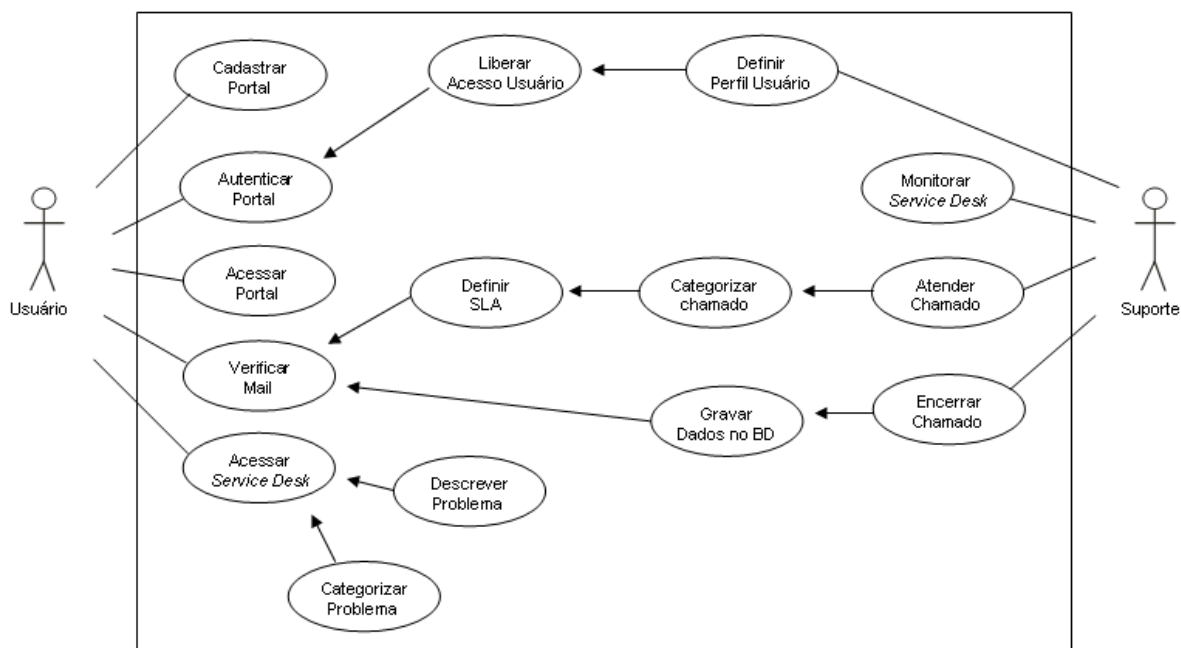
Para a especificação, construção e visualização do projeto SDvPC foi adotada uma linguagem de modelagem livre que permite que desenvolvedores visualizem os produtos de seu trabalho em diagramas padronizados, a *Unified Modeling Language* (UML). A UML auxilia a visualizar o desenho e a comunicação entre objetos, permitindo a visualização dos projetos através de diagramas padronizados (GUEDES, 2004). Junto com uma notação gráfica, a UML também especifica significados, isto é, semântica. É uma notação independente de processos, embora o RUP (*Rational Unified Process*) tenha sido especificamente desenvolvido utilizando a UML. Esta linguagem tem origem na compilação das "melhores práticas de engenharia" que provaram ter sucesso na modelagem de sistemas grandes e complexos. Sucedeu aos conceitos de Booch, OMT (Rumbaugh) e OOSE (Jacobson) fundindo-os numa única linguagem de modelagem comum e largamente utilizada.

O UML define como diagrama uma representação gráfica de um sistema sob determinada visão, e nesta seção serão apresentados basicamente dois tipos de diagramas comportamentais: os diagramas de caso de uso e diagramas de seqüência, os quais demonstram de forma clara as principais atividades do Portal Corporativo e do *Service Desk*.

Um diagrama de casos de uso descreve uma funcionalidade provida pelo sistema ou um grupo de atividades que produz um resultado concreto e tangível. São descrições de interações típicas entre os usuários de um sistema e o sistema propriamente dito. Eles representam a interface externa do sistema e especificam um conjunto de exigências do que o sistema deve fazer. O diagrama de casos de uso ilustra os comportamentos que o programa terá quando estiver pronto. Ele descreve relacionamentos e dependências entre um grupo de *Caso de Uso* e os Atores participantes no processo.

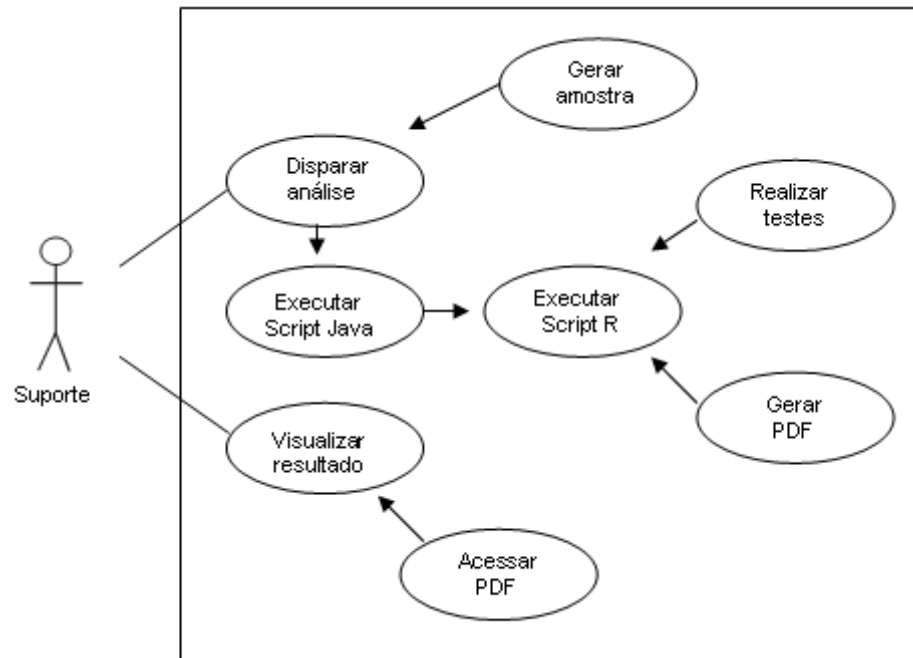
O diagrama apresentado na Figura 24, mostra o caso de uso para o acesso ao Portal e atendimento ao usuário. O diagrama apresenta os módulos do sistema que os usuários, representados pelo ator Usuário, e o Grupo de suporte, representados pelo ator Suporte, poderão interagir de forma direta e a interrelação entre esses módulos. Este caso de uso indica que os usuários poderão acessar a área de cadastro do Portal, realizar sua autenticação e utilizar todas as funcionalidades disponíveis no ambiente, como o *Service Desk*. O Grupo de

suporte irá definir o perfil de acesso do usuário, monitorar e atender todos os chamados reportados ao sistema, definindo a categoria e urgência de atendimento para cada um, de acordo com o SLA definido.



**Figura 24 – Caso de Uso – acesso ao Portal, cadastro e atendimento ao usuário.**

A Figura 25 representa o caso de uso com a sequência de execução das aplicações em Java e R. Neste diagrama o ator Suporte, dispara a análise das informações armazenadas na base de conhecimento, as quais dão origem a amostra de dados a ser utilizada durante a análise estatística. Em seguida é disparada a aplicação desenvolvida em Java, a qual permite a interação entre o Portal Corporativo e a linguagem R. Por fim, é representada, através do módulo Visualizar resultado, o acesso ao arquivo de saída do SD, no formato PDF.



**Figura 25 – Caso de Uso – execução das aplicações em Java e R.**

Para a representação do comportamento das operações, métodos (procedimentos ou funções) entre objetos no cenário do SDvPC foram utilizados diagramas de seqüência, que mostram a seqüência dos processos (mais especificamente de mensagens passadas entre os objetos) do sistema. O cenário é uma forma de ocorrência de um caso de uso, definindo como o software realizará seu papel (ordem de operações). A adoção deste tipo de representação gráfica, se justifica por ser um diagrama simples e lógico, tornando óbvios a seqüência e o fluxo de controle.

O diagrama de seqüência da Figura 26 descreve o cenário do registro de chamados pelo usuário, partindo da autenticação no Portal, acesso ao *Service Desk* e registro da solicitação de suporte. Um detalhe importante nesta representação é a confirmação do registro enviada pelo Portal diretamente ao usuário, via mail, o qual contém dados sobre o SLA, hora de abertura do ticket e tempo previsto de atendimento.

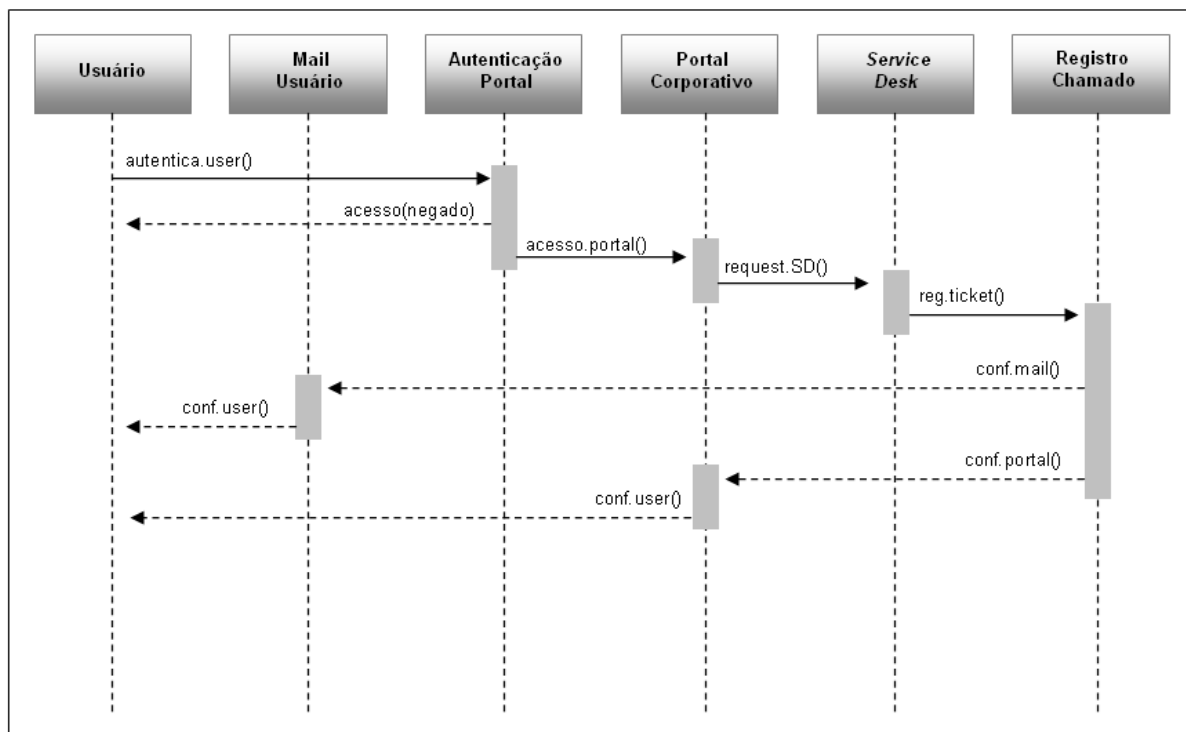


Figura 26 – Diagrama de seqüência – registro de incidente (ticket) no Service Desk via Portal.

A Figura 27 mostra o diagrama implementado para o atendimento aos *tickets* de usuários pela equipe de suporte. O destaque desta representação está na definição do SLA de atendimento ao chamado, por meio da quantificação do reporte e análise do perfil do usuário e por fim o armazenamento das informações na base de dados do sistema.

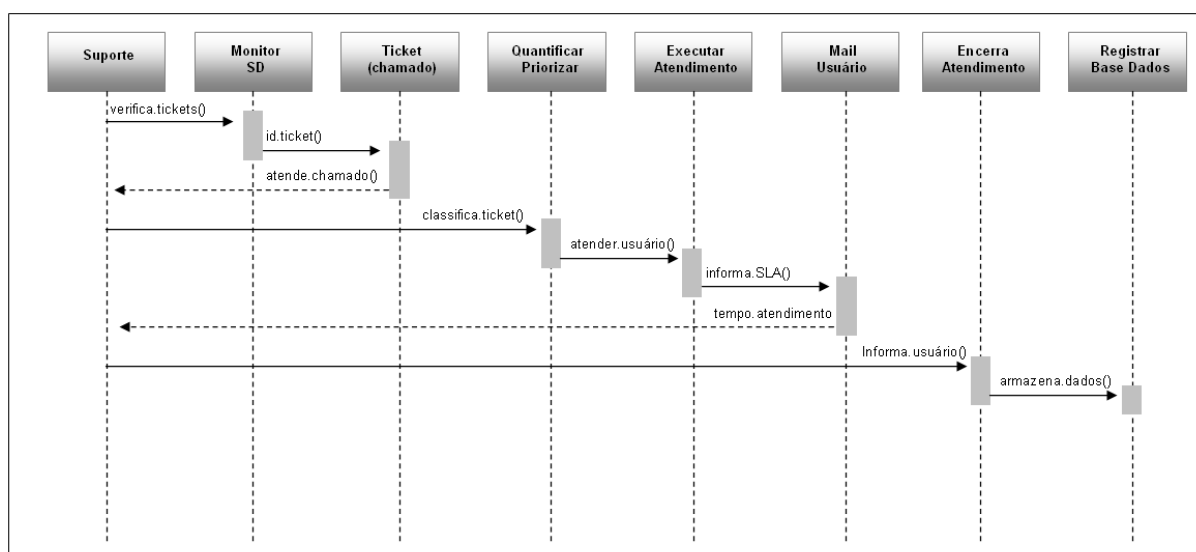
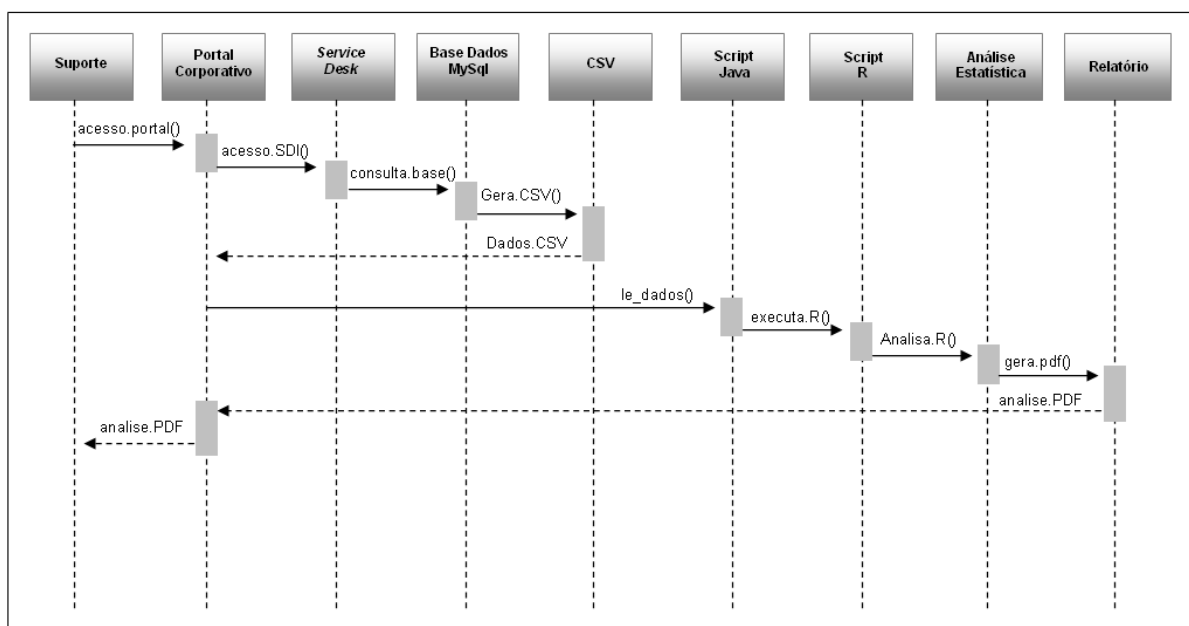
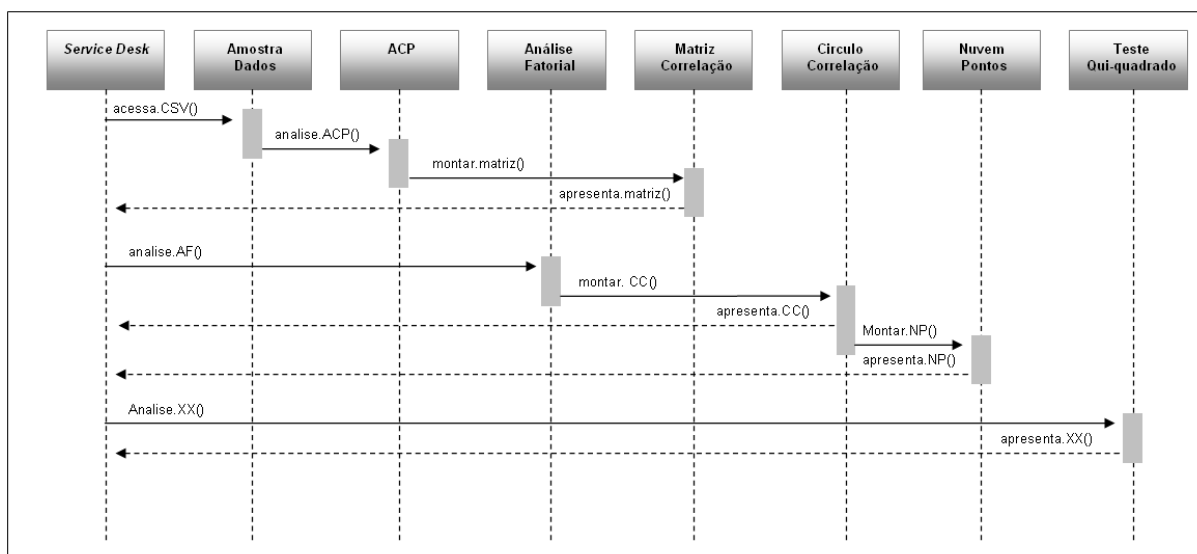


Figura 27 – Diagrama de seqüência – atendimento aos chamados dos usuários.

A amostra de dados para a análise estatística esta descrita no diagrama da Figura 28, a qual apresenta os principais passos realizados pelo sistema para organização das informações, desde o acesso ao *Service Desk*, geração do arquivo CSV e execução dos scripts em Java (leitura do CSV) e na linguagem R para a execução da própria análise e emissão do relatório utilizado para o apoio a decisão.



**Figura 28 – Diagrama de sequência – execução da análise estatística.**



**Figura 29 – Diagrama de sequência – análise estatística.**

Para o desenvolvimento da aplicação em R, o diagrama da Figura 29 mostra a sequência de análises realizadas e os resultados emitidos. Iniciando pela análise de componentes principais, por meio da matriz de correlação, seguido da análise fatorial tendo como resultado o círculo de correção e nuvem de pontos. Como último passo da análise, pode ser realizado o teste do qui-quadrado, a fim de verificar a real correlação entre as variáveis da amostra.

Para o entendimento do sistema SDvPC por meio de um modelo conceitual, é utilizado o diagrama de entidade-relacionamento (E-R), uma técnica de representação gráfica que facilita a visualização das relações entre todas as entidades componentes da aplicação. Este método se aproxima de um modelo teórico relacional, sendo usado na análise e desenvolvimento da aplicação em uma fase anterior a criação da base de dados. A Figura 30 apresenta a modelagem da base de dados do Portal, identificando os objetos relevantes, suas características e propriedades. A organização desta representação é feita pela utilização de entidades, conjunto de objetos de mesmo tipo, atributos, que são as características de cada objeto e por fim as relações, que mostram de maneira clara o inter-relacionamento entre as entidades por meio do uso de cardinalidades. A Figura 30, é composta por um conjunto de 25 entidades, que representam as tabelas da base de dados da aplicação, as quais armazenam todas informações relevantes ao Portal Corporativo, como: tabelas de usuários, salas, chamados, categorias de chamados, informações sobre a rede de dados e telefonia, funções e perfis de usuários, grupos entre outras.

A relação entre estas entidades permite ao sistema compor uma base de conhecimento simples de ser utilizada, e útil especificamente para a ferramenta de apoio à decisão baseada em análise multivariada. O dados armazenados nesta base podem ser coletados através de diferentes amostras e diferentes formas, o que garante uma multidisciplinaridade na escolha dos métodos de avaliação e estudo dos incidentes obtidos pelo sistema



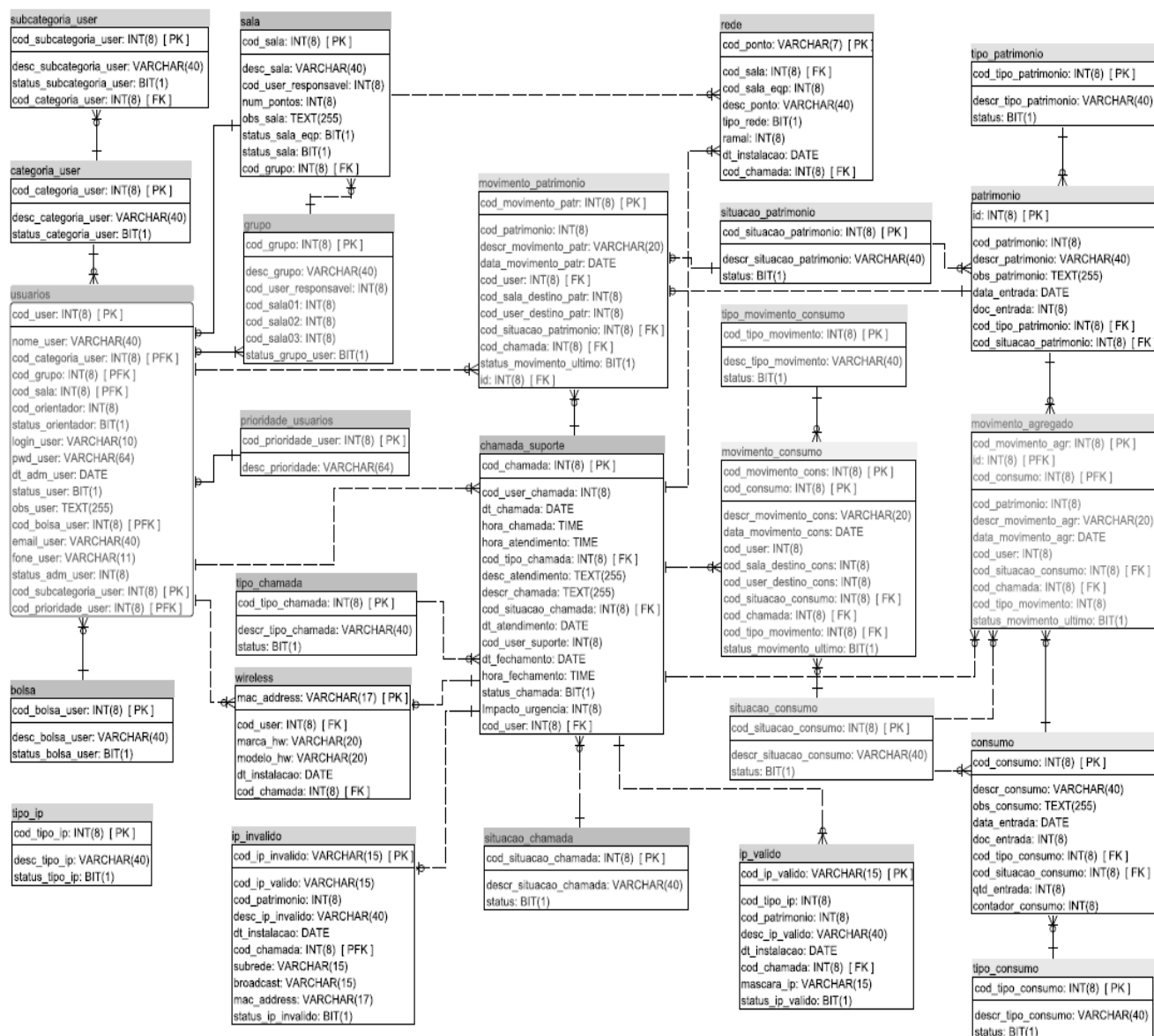


Figura 30 – Diagrama Entidade-Relacionamento – estrutura da base de dados do sistema.

### 5.3 Framework Django

O Django trabalha com 3 elementos principais (BRANDÃO, 2009): o Modelo, onde vão as definições dos dados, como eles devem ser armazenados e tratados; Visão, que são as funções que recebem requisições e retornam respostas; e Controle que é formado por um conjunto de *middlewares* e verificadores de endereço e caminho de requisições. Esta ferramenta possui um sistema para a criação de formulários automáticos de acordo com a definição dos modelos de dados criados por meio do ORM o que proporciona um ganho de produtividade muito grande. O Django usa uma linguagem baseada em modelos (*templates*), que permite separar design, conteúdo e código Python.

A interface administrativa é uma das vantagens mais fortes em relação a outros *frameworks*. No Django é possível gerar uma interface administrativa totalmente automática para os modelos criados por meio do ORM. Nela, é possível alterar, adicionar deletar informações, além de fazer buscas à objetos diversos.

### 5.3.1 Informações Técnicas do Projeto com Django

Um projeto desenvolvido em Django é dividido em aplicações. Essas aplicações podem se comunicar entre si e podem ser plugadas em outros projetos. No portal corporativo foram desenvolvidas até agora 5 aplicações: informativos, documentos, transportes, imagens e suporte.

Para gerir todas as informações criadas pelas aplicações, utilizou-se o mapeamento objeto-relacional (ORM). Segue a Figura 31, que apresenta a estrutura ORM simplificada da base de dados usada para publicar notícias no portal:

```
class Noticia(models.Model):
    class Meta:
        ordering = ('-publicacao',)
    def __unicode__(self):
        return self.titulo
    titulo = models.CharField(max_length=350)
    conteudo = models.TextField()
    publicacao = models.DateTimeField(default=datetime.now,blank=True,
                                     editable=False,unique=True)
    imagem = models.ImageField(upload_to='imagens/noticias/original',
                              blank=True,default='imagens/noticias/logo_grande_inpe.jpg')
```

Figura 31 – Trecho de código com a Estrutura ORM.

Cada modelo é representado por uma classe derivada da classe *django.db.models.Model* e tem um número de variáveis de classe, as quais por sua vez representam os campos de banco de dados no modelo. Cada campo é representado por uma instância de uma classe *Field*. *CharField* para campos do tipo caractere, *TextField* para campos de texto, *DateTimeField* para campos de data e *ImageField* para campos de imagens. Algumas instâncias exigem que sejam especificados parâmetros como tamanho máximo do campo (*max\_length*), outras permitem receber valores padrão. No modelo Meta pode-se definir a ordenação dos dados, nome da tabela etc. A estrutura ORM também permite

sobrescrever métodos. Um uso comum disso é quando for necessário fazer uma verificação antes de persistir os dados no banco, para isso sobrescreve-se o método *save*.

Todas as requisições enviadas para os aplicativos são processadas na *View* antes de serem exibidas no arquivo html. Na Figura 32 uma *View* responsável por exibir os últimos informativos na tela.

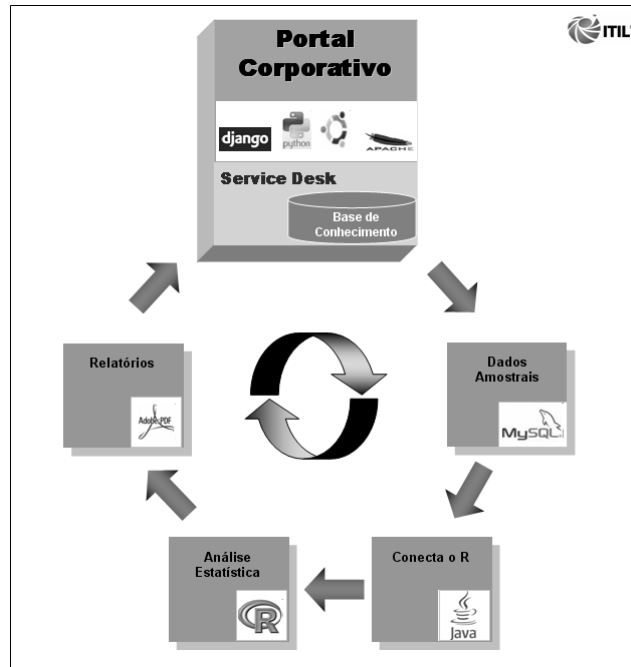
```
def index(request):
    ultimas_noticias = Noticia.objects.all().order_by('-publicacao')[:3]
    ultimos_comunicados = Comunicado.objects.all().order_by('-publicacao')[:4]
    ultimos_eventos = Evento.objects.all().order_by('-publicacao')[:4]
    return render_to_response('portal/conteudo_principal.html',
                              locals(),
                              context_instance=RequestContext(request))
```

**Figura 32 – View com os últimos informativos.**

As *views* recebem pelo menos 1 argumento (*request*) que é uma classe responsável por informar, por exemplo, que tipo de método na requisição foi usado (POST ou GET), usuário logado na sessão etc. Depois são feitas consultas ao banco de dados utilizando métodos da classe *django.db.models.Model*. Essas consultas retornam uma instância contendo os dados da pesquisa. Depois esses objetos instanciados são passados para o arquivo html usando a função *render\_to\_response*.

#### 5.4 Implementação do SDvPC

Para que o SDvPC possa cumprir sua finalidade, é necessário a utilização de um conjunto de ferramentas que permitam integrar a tecnologia de Portal Corporativo e a base de conhecimento gerada, resultando numa ferramenta de apoio a decisão. A Figura 33 apresenta o conjunto de módulos e tecnologias que compõem o sistema.



**Figura 33– SDvPC – Módulos e Composição da Ferramenta de análise multivariada de Incidentes de Segurança.**

Fonte: CASTILHO, 2007

Para o entendimento do sistema, faz-se necessário compreender a integração de todas as tecnologias empregadas para a implementação do Portal e da ferramenta de análise dos Incidentes e Apoio à Decisão:

- O Portal Corporativo, já na sua segunda versão, foi inteiramente desenvolvido com o *framework* Django, e projetado para a plataforma Ubuntu Linux. A integridade e confidencialidade das informações que trafegam no ambiente são garantidas através do uso de tecnologias como HTTPS e SSL, que são recursos computacionais que utilizam técnicas criptográficas para a troca segura de informações.
- A tecnologia MySQL é utilizada como base para o armazenamento de todo dado e informação manipulado pelo sistema. Para que possa ser realizada a análise estatística multivariada é necessário obter uma amostra correta dos incidentes de segurança armazenados. As rotinas de pesquisa e busca em banco de dados são realizadas por meio de consultas SQL, as quais podem ter sua saída direcionada para tela, impressora e também arquivos de texto simples.
- A partir do momento em que a amostra de dados está disponível, o portal corporativo dispara a ferramenta estatística R, a qual tem sua execução iniciada através de uma classe de interconexão desenvolvida em linguagem Java. O R reconhece e interpreta

todos os dados armazenados, e dispara um conjunto de técnicas estatísticas com a finalidade de traduzir e coletar todas as informações disponíveis na amostra.

- Ao término da análise, são gerados um conjunto de gráficos, que apresentam as informações coletadas, as quais são armazenadas em arquivos do tipo pdf. Neste ponto o portal corporativo identifica todos os arquivos gerados pela análise, e os disponibiliza novamente aos usuários do sistema via uma interface WEB.

#### 5.4.1 Interface do Usuário

Para que os usuários tenham acesso ao sistema de SD é necessário que os mesmos estejam cadastrados no Portal Corporativo para que possam realizar a autenticação no sistema. A Figura 34 apresenta a tela de cadastro dos usuários do CRS. Existe um conjunto de campos chaves que devem ser obrigatoriamente preenchidos, estes estão marcados com asterisco. Ao se cadastrar, o processo de quantificação dos incidentes mais tarde reportados é iniciado. Dentre os dados informados ao sistema, está o perfil do usuário, o qual define o grau de prioridade em seu atendimento, além de informações adicionais, também relevantes, como: categoria, grupo, sala, status entre outros.

**Figura 34– Tela de cadastro de usuários - SDvPC.**

Uma vez autenticado no Portal o usuário tem acesso a interface de gerenciamento de seus dados pessoais e também, mais importante, as funções de SD. No SD podem ser reportados os chamados (tickets) sobre incidentes de segurança vivenciados pelos usuários. Nesta mesma área é possível acessar o histórico de todos os chamados realizados pelo usuário, identificar o status de atendimento e as SLAs relacionadas. Para a abertura de um

novo ticket, Figura 35, é necessário categorizar e descrever o problema, preferencialmente de forma clara e precisa para que o mesmo seja identificado e atendido de forma rápida e eficiente. Todo ticket de usuário é armazenado em uma base de dados e enviado a uma fila de atendimento do tipo FIFO (*first in, first out*) em um servidor, onde os primeiros chamados submetidos são os primeiros a serem atendidos, com exceção dos que forem classificados com prioridade 1, ou seja, as solicitações classificadas como críticas terão preferência de atendimento.

#### 5.4.2 Interface do Administrador

O sistema de administração do SD é acessado a partir do Portal Corporativo, porém o mesmo é totalmente independente da área onde os usuários reportam seus *tickets*. A escolha deste padrão é apoiado pela necessidade de se manter a integridade e confidencialidade das informações reportadas ao sistema.

Ao efetuar o *login* na área administrativa, o administrador dispõe de uma interface completa para a gerência dos incidentes reportados. Todas as informações relacionadas ao problema e ao usuário estão disponíveis de forma clara e de fácil interação, conforme é mostrado na Figura 35.

The screenshot displays the 'Administração do Portal Corporativo do CRS' interface. At the top right, it shows the user 'Breno' and options to 'Alterar senha' or 'Encerrar sessão'. A navigation menu includes 'Auth', 'Flatpages', 'Informativos', 'Materiais', 'Rede', 'Sites', 'Subdados', 'Suporte', and 'Transporte'. The main content area is divided into several sections:

- IPs Bloqueados:** A table listing blocked IP addresses.
 

Usuário	Sala	IP
Marta Helena	2048	10.10.10.115
Alessandro Vasconcelos	2020	10.10.10.124
- Chamadas Não Atendidas:** A table showing pending calls.
 

Usuário	Sala	Data/Chamada	Prioridade/Atendimento	Suporte Destinatário	Problema
Gislaine Gabriele	2023	05/10/2009 09:17:39	.....	None	<a href="#">Ver</a> <a href="#">Atender</a>
Tatiana Mora	1010	04/10/2009 14:20:18	.....	None	<a href="#">Ver</a> <a href="#">Atender</a>
- Chamadas em Atendimento:** A table showing calls currently being handled.
 

Usuário	Sala	Data/Chamada	Data/Atendimento	Prioridade/Atendimento	Suporte	Problema	Ausências
Afranio Silva	1030	01/12/2009 10:31:16	01/12/2009 10:31:21	Média	Breno Simonetti	Ver	0
- Últimas Chamadas Finalizadas:** A table showing recently completed calls.
 

Usuário	Sala	Data/Chamada	Data/Atendimento	Prioridade/Atendimento	Suporte	Problema
Alessandro Vasconcelos	2020	05/10/2009 11:19:42	05/10/2009 11:55:07	Baixa	Breno Simonetti	<a href="#">Ver</a>
Marta Helena	2048	05/10/2009 09:19:59	05/10/2009 09:41:12	Alta	Breno Simonetti	<a href="#">Ver</a>
Marta Helena	2048	05/10/2009 08:20:30	05/10/2009 09:21:02	Média	Breno Simonetti	<a href="#">Ver</a>
- Usuários Esperando Liberação:** A table listing users waiting for release.
 

Nome	Sala	Categoria
Afranio Silva	1030	Servidor
Luiz Antônio	2057	Tercerizado

Figura 35– Interface para administração dos incidentes reportados ao SD - SDvPC.

A principal página do SD disponibiliza cinco tipos de informações:

- Endereços Ips bloqueados, computadores com problemas de segurança identificados pela equipe de suporte a TI tem seu acesso a rede bloqueado até a resolução dos mesmos;

- Usuários cadastrados no sistema e esperando liberação de acesso. Embora o cadastro no Portal Corporativo seja uma atividade automatizada, a liberação de acesso e confirmação do perfil do usuário é emitido de forma manual pela área de TI;
- Status dos *tickets*, toda chamada ao ser recebida pelo SD é identificada de acordo com seu status, que pode ser: aberto, em andamento ou finalizado.

Para que a interface de administração atendesse todos requisitos do sistema, houve a necessidade de customização do *framework* Django. A forma mais simples para alcançar o resultado esperado foi através da utilização do “*template tags*”, um método disponível para a reutilização de código em outros projetos e aplicações. Por exemplo: para gerar a lista de ips bloqueados foi desenvolvido o seguinte *template tag*:

```
class IpsBloqueados(Node):
    def __init__(self, varname):
        self.varname = varname
    def render(self, context):
        context[self.varname] = Ip.objects.filter(status=Ip.BLOQUEADO)
        return ''
    def get_ip_bloqueado(parser, token):
        bits = token.contents.split()
        if len(bits) != 3:
            raise TemplateSyntaxError, "get_ip_bloqueado tem que ter exatamente 3 argumentos"
        if bits[1] != 'as':
            raise TemplateSyntaxError, "primeiro argumento tem que ter 'as'"
        return IpsBloqueados(bits[2])
```

O método `get_ip_bloqueado` recebe uma string de argumentos vindos do template html da página de administração. Depois de verificar a quantidade de argumentos passados, é invocado a classe `IpsBloqueados` que vai fazer a consulta ao banco de dados através do mapeamento objeto relacional (`Ip.objects.filter(status=Ip.BLOQUEADO)`). E finalmente o template invoca o método da seguinte forma:

```
{% get_ip_bloqueado as ips %}
{% for ip in ips %}
    <tr class="{% cycle 'row1' 'row2' %}">
        <td>{{ip.chamada.usuario_solicitante}}</td>
        <td>{{ip.chamada.usuario_solicitante.sala}}</td>
        <td><a href="/admin/rede/ip/{{ip.id}}">{{ip.endereco}}</a></td>
    </tr>
{% endfor %}
```

O uso de *templates tags*, permite uma listagem interativa na tela, conforme Figura 35, do sistema dos usuários que se cadastraram no Portal, mas ainda não tiveram seu acesso liberado. Esse controle é realizado para que cada usuário receba a classificação de prioridade, a qual será utilizada no controle de incidentes. Ainda na Figura 36, ao centro da página o tem-se a listagem dos tickets abertos, em atendimento e finalizadas.

Todo *ticket* ao ser atendido também deve ser priorizado, a classificação do chamado é baseada em 3 premissas: o nível de prioridade do usuário, a categoria e a descrição do incidente. O resultado da priorização é utilizado para definição do tempo de atendimento a ser respeitado. Ao encerrar o processo de identificação e quantificação do incidente, é necessário armazenar em uma base de conhecimento, todas informações do reporte, pois estas informações serão utilizadas como apoio a decisão. Além disso, segundo a biblioteca ITIL, o usuário deve permanecer informado sobre a situação de seu atendimento, para alcançar esse objetivo o SD envia um e-mail descritivo com dados do suporte em períodos específicos. A seguir são apresentados os trechos de códigos que executam o armazenamento das informações do chamado a *view* “atender(self,request)” e também o método “email\_to\_user(chamada.ATENDIMENTO)” que passa os argumentos, como tipo de mensagem, a ser enviado ao usuário.

Ainda na tela de Administração do Portal, Figura 37, pode-se obter informações detalhadas de cada reporte de maneira simples, clicando sobre “Detalhes/Mudar Atendimento” é apresentada uma janela, Figura 33, contendo informações relacionadas ao atendimento.

```
def atender(self,request):
    if (request.method=='POST') and
    (request.POST['prioridade_chamada']!=''):
        chamada =
        Chamada.objects.get(id=request.POST['chamada_id'])
        chamada.usuario_suporte =
        Usuario.objects.get(id=request.user.id)
        chamada.data_atendimento =
        datetime.now().isoformat(' ')
        chamada.impacto_urgencia =
        int(request.POST['prioridade_chamada'])

        chamada.status = Chamada.EM_ATENDIMENTO
        chamada.save()
        chamada.email_to_user(chamada.ATENDIMENTO)
        return HttpResponseRedirect('/admin/')
    else:
        return HttpResponseRedirect('/admin/');
    return
    render_to_response('admin/index.html',locals(),context_instan
    ce=RequestContext(request),)
```



Usuario solicitante:	Afranio Silva
Usuario suporte:	Breno Simonetti
Destino suporte:	----- ▾ +
<b>Tipo chamada:</b>	Solicitação de Hardware/Software ▾ +
<b>Atendimento:</b>	<div style="border: 1px solid gray; height: 80px;"></div>
Descricao chamada:	Solicito a instalação de um impressora HP na minha sala.
Impacto urgencia:	Média ▾
<b>Ips</b>	
Ip: #1	
Adicionar IPs ( Mostrar )	
<b>Ponto Redes</b>	
Ponto Rede: #1	
Adicionar Pontos de Rede ( Mostrar )	
<b>Wireless</b>	
Wireless: #1	
Adicionar Wireless ( Mostrar )	
<b>Movimento Patrimonios</b>	
Movimento Patrimonio: #1	
Movimento de Patrimônio ( Mostrar )	
<b>Movimento Consumos</b>	
Movimento Consumo: #1	
Movimento de Consumo ( Mostrar )	

**Figura 36– Janela com informações detalhadas sobre os *tickets* abertos pelos usuários.**

Na interação com a tela de atendimento, Figura 36, o responsável pelo suporte descreve a solução encontrada para o problema. As abas localizadas abaixo, são utilizadas para o cadastro de informações relevantes para a administração da rede e movimentação de material envolvidos no chamado. Uma especificidade no atendimento dos incidentes, é o fato do usuário não estar presente em sua sala no instante do suporte, impossibilitando assim a resolução do problema. Neste casos, para que não haja uma distorção nos dados armazenados na base de dados sobre incidentes, o próprio sistema reporta ao usuário por meio de mail a sua ausência, fato repetido por 3 vezes. Caso não haja retorno do usuário o *ticket* é imediatamente descartado, conforme mostra o código da *view* “notificar\_ausencia(self,request,chamada\_id)”. A tela principal, Figura 38 apresenta o número de notificações enviadas.

Abaixo, esta o método responsável por enviar todos os email para os usuários. Este função recebe como argumento o tipo de mensagem a ser transmitida. Uma thread foi

implementada para que o envio dos mails pudessem ser feitos em segundo plano, sem a interrupção do sistema.

```
def email_to_user(self, TIPO_MSG):
    if TIPO_MSG == self.ATENDIMENTO:
        titulo = "Atendimento de chamado"
        texto = u""" Caro usuário, seu chamado recebeu prioridade %s , então de acordo
com a escala GLA seu chamado será atendido %s .
Att,
GTI
""" %(self.get_prioridade(),self.get_prioridade_horas())
    pass
    if TIPO_MSG == self.NOTIFICACAO:
        titulo = "Notificação de Ausência"
        texto = u""" Caro usuário, o integrante %s do grupo GTI foi em sua sala em %s,
mas você não estava.
São toleradas no máximo 3 ausências, mais que isso o seu chamado será cancelado e
você terá que realizar um novo chamado para ser atendido.
Att,
GTI
""" %(self.usuario_suporte.get_full_name(),datetime.now().strftime("%d/%m/%Y as
%H:%M"))
    pass
    if TIPO_MSG == self.CANCELAMENTO:
        titulo = "Cancelamento de chamado"
        texto = u""" Caro usuário, os integrantes do grupo GTI foi em sua sala 3 vezes
mas você não estava.
São toleradas no máximo 3 ausências, por isso seu chamadado será cancelado. Caso
queira ser atendido favor submeter novamente uma solicitação.
Att,
GTI
"""
    pass
    t = threading.Thread(target=send_mail,
                        args=[titulo,texto,EMAIL_HOST_USER,[self.get_email_user()]])
    t.setDaemon(True)
    t.start()
```

Com o encerramento do processo de atendimento ao usuário, o sistema de SD já reúne uma base de informações relevantes, as quais serão utilizadas como amostra para a análise estatística multivariada. Neste ponto o sistema deve dispor dos mecanismos que permitam a equipe de suporte interagir com a ferramenta estatística e também coletar a partir da base de dados as informações necessárias. A Figura 37 mostra a área de acesso à ferramenta para coleta de dados e execução da aplicação em R, que fornece como resultado as informações da análise multivariada.



Figura 37– Área de acesso a ferramenta para coleta de dados e análise estatística.

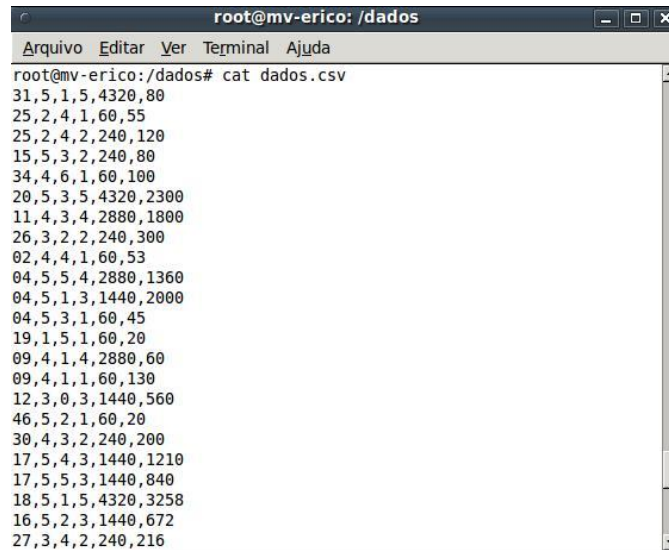
### 5.4.3 Integração – Python, Java e R

Para alcançar os resultados almejados, foi necessário integrar ao Portal 3 linguagens de programação: a linguagem Python, Java e R. A linguagem Python utilizada no desenvolvimento do próprio Portal, também é responsável pela obtenção da amostra de dados. O trecho de código abaixo, em Python, apresenta a organização, filtro e coleta das informações na base de dados do sistema, informações estas que serão utilizadas para análise estatística.

```
def model_to_csv(request):
    chamadas = Chamada.objects.filter(status=Chamada.FINALIZADA)
    arq = csv.writer(open("media/csv_e_estatisticas/%s.csv", "w", datetime.now))
    for chamada in chamadas:
        previsao_tempo_atendimento = chamada.get_minutos_prev_atend()
        tempo_atendimento = chamada.data_fechamento - chamada.data_atendimento
        tempo_total_minutos = tempo_atendimento.days*24*60 + tempo_atendimento.seconds/60
        arq.writerow([chamada.usuario_solicitante.id,
                     chamada.usuario_solicitante.prioridade,
                     chamada.tipo_chamada.id,
                     chamada.impacto_urgencia,
                     previsao_tempo_atendimento,
                     tempo_total_minutos,
                     ])

    return render_to_response('admin/', locals(),
                              context_instance=RequestContext(request))
```

O resultado da coleta anteriormente, é um arquivo de texto, no formato CSV, o qual possui uma tabela de dados referentes aos incidentes de segurança registrados no período. A Figura 38 mostra este arquivo e seu conteúdo. A tabela é formada pelas seguintes colunas da esquerda para direita: código do usuário, perfil do usuário, categorização do incidente, valor do *ticket*, tempo de atendimento previsto pela SLA (minutos) e tempo real do atendimento (minutos).



```

root@mv-erico: /dados
Arquivo  Editar  Ver  Terminal  Ajuda
root@mv-erico:/dados# cat dados.csv
31,5,1,5,4320,80
25,2,4,1,60,55
25,2,4,2,240,120
15,5,3,2,240,80
34,4,6,1,60,100
20,5,3,5,4320,2300
11,4,3,4,2880,1800
26,3,2,2,240,300
02,4,4,1,60,53
04,5,5,4,2880,1360
04,5,1,3,1440,2000
04,5,3,1,60,45
19,1,5,1,60,20
09,4,1,4,2880,60
09,4,1,1,60,130
12,3,0,3,1440,560
46,5,2,1,60,20
30,4,3,2,240,200
17,5,4,3,1440,1210
17,5,5,3,1440,840
18,5,1,5,4320,3258
16,5,2,3,1440,672
27,3,4,2,240,216

```

**Figura 38– Arquivo com amostra de dados para análise multivariada.**

Com os dados prontos, é necessário disparar um programa escrito em linguagem R, para realizar a análise multivariada. Neste ponto foi necessário implementar um código em linguagem Java, pois a linguagem Python não possui bibliotecas disponíveis para invocar e executar aplicações em R. Abaixo é apresentado o código em Python que ativa o servidor da linguagem R e executa a aplicação `analisa.class`.

```

ativaR="R CMD Rserve"
os.system(ativaR)
executaJava="java analisa.class"
os.system(executaJava)

```

A seguir é apresentado o pacote implementado em linguagem Java, que faz a leitura do arquivo `dados.csv` e direciona seu conteúdo, para que a aplicação `analisa.R` realize a análise multivariada dos dados. Para o programa em Java interagir com a linguagem R, é necessário que o mesmo utilize o pacote `JRClient`, que possui um conjunto de classes especialistas com a finalidade de permitir a integração entre as linguagens.

```

package le_dados;
import org.rosuda.JRclient.*;
public class Main {
public static void main(String[] args){
    try{
        Rconnection c = new Rconnection();
        //caso queira mudar o diretorio de trabalho
        //c.voidEval("setwd('~~/fazendo')");
        //dae se o fonte do R estiver no mesmo diretorio
        //c.voidEval("source('análise.R')");
        //Lê o arquivo pro objeto de nome dados
        c.voidEval("dados<-
read.table('~~/NetBeansProjects/le_dados/src/le_dados/dados.csv', sep=',')");
        //executa o que tiver no arquivo fonte exemplo.R e no diretório de trabalho jái
        //vai ter o obj dados
        c.voidEval("source('~~/NetBeansProjects/le_dados/src/le_dados/análise.R')");
        //Ãtil pra debug: guarda em x a saÃda da expressÃo
        REXP x=c.eval("linha(dados, 1)");
        //Ãtil pra debug: lista o que tinha na expr dizendo o tipo
        System.out.println(x.toString());

        //Do the toxic waltz
    }catch(RSrvException ex){
        System.out.println("problema");
        ex.printStackTrace();
    }
}
}
}

```

Após o tratamento correto das informações geradas pelo SD através dos processos de quantificação, coleta e disponibilização dos dados, a aplicação desenvolvida na linguagem R (análise.R) implementa as funções necessárias para a análise estatística da amostra disponibilizada.

O conjunto de passos realizados pelo programa análise.R para a disponibilização das informações de apoio a decisão é apresentado no item 5.5. Inicialmente é realizada uma análise descritiva dos dados e mostrada através de uma tabela de distribuição de frequências. Os próximos passos constituem-se na redução do número de fatores e da aplicação da análise fatorial, com o propósito de definir a estrutura subjacente em uma matriz de dados a fim de identificar um conjunto de fatores relevantes. Para isso é utilizada a Matriz de correlação, identificando-se assim a relação e o nível de importância das variáveis na amostra. As representações do Círculo de Correlação e Nuvem de Pontos mostram graficamente estas relações. Para comprovar as informações obtidas, a aplicação realiza o teste do qui-quadrado para comprovar a veracidade da relação entre os dados.

## 5.5 Resultados e Discussões

Nesta seção serão discutidos, no item 5.5.1, os principais resultados obtidos com a aplicação da análise estatística multivariada sobre a amostra de dados coletas a partir do

SDvPC. O item 5.5.2 apresenta um estudo comprobatório dos resultados obtidos, este item utiliza ferramentas estatísticas e da qualidade para definir a validade do conhecimento obtido com as técnicas multivariadas. A última parte desta seção, item 5.5.3. apresenta como pode-se alcançar uma pró-atividade na tomada de decisões a partir dos fornecidos pelo SDvPC.

### 5.5.1 Descrição dos Resultados

O resultado das consultas à base de conhecimento do Portal, retorna uma amostra completa das informações sobre os incidentes de TI reportados ao sistema e atendidos pela equipe de suporte. Esta amostra pode ser utilizada de várias maneiras e com diversas finalidades, de acordo com a necessidade identificada. Neste trabalho o foco está na utilização de técnicas multivariadas para elucidar as informações disponíveis nesses conjuntos de dados. O intuito é fornecer subsídios para o apoio a decisão. Desta forma após a coleta dos dados foram elaboradas tabelas de distribuição de frequências para verificar o comportamento das variáveis “Perfil do Usuário”, “Grupo Incidente” e “Valor do *Ticket*” (prioridade para o atendimento do seu chamado), com o uso, principalmente, da Linguagem R, e apoio de alguns outros softwares disponíveis no mercado como: *Statistica*, *SPSS* e da planilha Excel.

O Perfil do Usuário varia de acordo com seu grau de importância dentro da empresa em uma escala de 1 a 5, onde um determinado cliente com perfil igual a 1 tem prioridade de atendimento em relação aos clientes classificados nos perfis dos números seguintes. A Tabela 01 mostra a tabela de distribuição de frequências dos reportes ocorridos para os cinco tipos de perfis de usuários existentes, demonstrando que os clientes com perfil 1 possuem menor índice de chamados, em oposição aos clientes com perfil 5, responsáveis pelo maior número de solicitações ao suporte, 30%. Esta tabela de distribuição de frequências permite verificar que quanto menor a importância do cargo do cliente, maior o seu perfil de usuário. Este grupo é formado pelo maior número de usuários, com isso há também uma maior necessidade de auxílio do suporte a esses clientes.

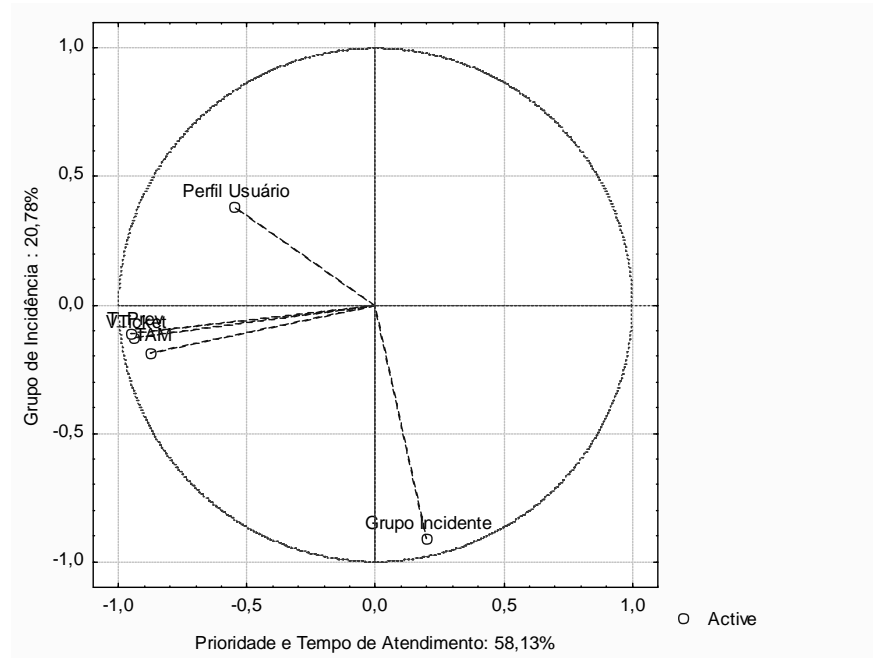
**Tabela 1 – Distribuição de frequências do Perfil do Usuário**

Perfil	Frequência de Chamados	Frequência Acumulada	Frequência Relativa %	Frequência Relativa Acumulada %
1	3	3	6	6
2	8	11	16	22
3	10	21	20	42
4	14	35	28	70
5	15	50	30	100

Além das variáveis indicadas acima, foram utilizadas as variáveis Tempo Previsto para o Atendimento e Tempo de Atendimento Efetivo, ambas quantificadas em minutos. Com este número de variáveis é inviável a utilização de técnicas estatísticas univariadas, onde se faz necessário aplicar então, métodos estatísticos multivariados de modo a simplificar a estrutura dos dados e a sintetizar informações, MALHOTRA (2001).

Através da Análise Fatorial as variáveis são agrupadas em função de suas correlações e tem-se o propósito de explicá-las através de suas dimensões comuns, os fatores calculados. Desta forma aplicou-se o teste de Kaiser-Meyer-Olkin, conhecido como KMO, para análise da adequabilidade da amostra o qual apresentou valor de 0,733, indicando que a amostra é passível de ser analisada pelas técnicas da análise fatorial. Com base nesta premissa procedeu-se a AF, calculando-se a matriz de correlação e a determinação dos autovalores e percentual de variância explicada por cada um. Para fins de análise foram considerados apenas dois fatores por resultarem num percentual de variância explicada de 78,91%; isto indica que a partir dessa análise, no lugar de cinco variáveis, passam a ser utilizados apenas 2 fatores.

Analisando os fatores encontrados, pode-se concluir que o fator 1 é o mais importante para o estudo, pois é derivado do maior autovalor e possui uma variância explicada de 58,13%, sendo um fator representativo da prioridade de atendimento, tempo previsto para este atendimento e do tempo efetivo para a solução do problema pelo suporte técnico; uma vez que as variáveis que mais contribuem para este fator são “Valor do *Ticket*”, “Tempo Previsto para o Atendimento” e “Tempo de Atendimento”. O fator 2 é derivado do segundo autovalor e fornece uma explicação de variância de 20,78%, sendo representado pela variável Grupo Incidente. Observa-se na Figura 39 a distribuição das variáveis no círculo de correlação, onde as variáveis mais próximas ao círculo de correlação são altamente representativas para o plano fatorial traçado. Uma das principais utilizações do círculo unitário de correlação é realizar a sua sobreposição sobre o plano fatorial; possibilitando dessa forma, identificar visualmente, quais variáveis estão relacionadas com os casos em estudo.

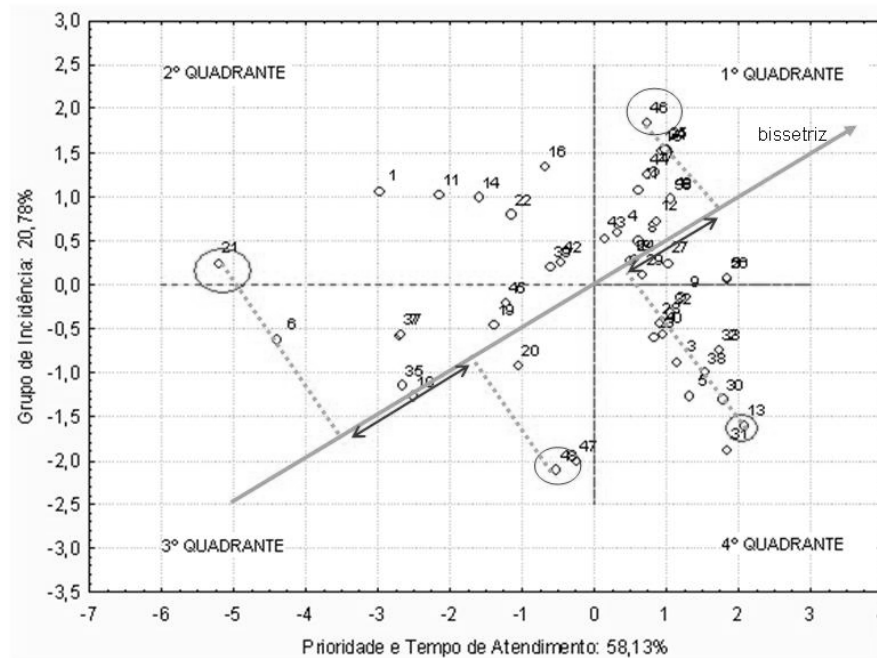


**Figura 39 - Distribuição das variáveis no círculo de correlação**

A análise simultânea das Figuras 39 e 40 remetem a verificar que o atendimento de nº 21 possui Perfil de Usuário com valor 5, nível hierárquico mais baixo, valor do *ticket* igual a 5 e grupo de incidência identificado como sendo de nº 1. Da mesma forma os chamados de números 1, 11, 14 e 16 são todos pertencentes a ocorrências do grupo de incidência de número 1 e estão neste quadrante em função do poder de explicação do fator 2. Em tempo, é importante salientar que na verificação das variáveis perfil de usuário e grupo incidente, na Figura 39, é possível identificar um ângulo de aproximadamente 180 graus, indicando que o perfil de usuário possui pouca relação com o grupo de incidente, por serem pontos opostos.

A análise do poder de explicação dos fatores permite inferir que os casos grafados no primeiro e quarto quadrantes têm maior prioridade de atendimento e devem ser atendidos no menor tempo possível em função do nível hierárquico e da natureza do atendimento de quem os solicitou. Da mesma forma, os chamados constantes no segundo e terceiro quadrantes têm menor prioridade e, portanto um tempo maior para serem solucionados pelos técnicos de suporte.





**Figura 40-** Gráfico da distribuição da nuvem de pontos.

Na tabela 2 foi elaborado um exemplo com a finalidade de demonstrar como a interpretação da metodologia utilizada pode ajudar nas ações a serem tomadas para auxiliar o gerenciamento do suporte técnico. Foram selecionados 4 chamados (casos) para o *Service Desk* constantes no gráfico da Figura 40. O caso de número 13, no 4º quadrante é de um usuário com perfil de alta gerência (Perfil do Usuário = 1) e prioridade de atendimento 1 (máxima) e teve seu problema solucionado em 20 minutos. O chamado número 46 embora sendo de um usuário de nível hierárquico baixo, teve sua solicitação definida como de alta prioridade e alocada neste quadrante em função disto e de seu grupo de incidente. O usuário representado pelo caso número 48 tem perfil e prioridades médios e, se enquadra no grupo de incidência 6, tendo um tempo longo para ter seu problema sanado. O chamado de número 21 corresponde a um defeito do grupo 1, mas é de um usuário de baixo nível hierárquico e, portanto com pouca ou nenhuma prioridade para ser atendido e teve de esperar 54 horas para ter seu caso resolvido.

**Tabela 2:** Exemplos de casos e prioridades

Caso	Prioridade (Valor Ticket)	Perfil Usuário	Grupo Incidente	Tempo Real Atendimento
13	1 - Alta	1	5	20 min
46	1 - Alta	5	1	50 min
48	3 - Média	3	6	15h 24min
21	5 - Baixa	5	1	54h

Na fase final da análise utilizou-se o teste qui-quadrado para verificar de forma pontual se há ou não uma relação entre o cargo ocupado (perfil do usuário) e o grupo de incidência dos tipos de atendimento reportados. Para a aplicação desta técnica, foi definido como hipótese de nulidade o fato de não haver relação entre o perfil do usuário (cargo ocupado) e o tipo de ocorrência (grupo de incidente), e a hipótese alternativa é o fato de que há uma relação entre o cargo ocupado e a necessidade de atendimento. O resultado do teste demonstrou-se significativo com um valor de 147,21 e 49 graus de liberdade, com nível de significância onde o *p-value* foi  $p < 0,0000$ , portanto, rejeitando-se a hipótese  $H_0$ , indicando que há forte relação entre as variáveis estudadas.

Os usuários de menor nível hierárquico respondem por 30% dos chamados distribuídos em todos os grupos de incidência, em oposição ao grupo de maior nível hierárquico, grupo 1, responsável por apenas 6% do total de chamadas. A partir destas observações fica evidenciada a necessidade da adoção de um conjunto de ações visando diminuir o número de solicitações, com ênfase principalmente para os grupos de menor nível hierárquico, responsáveis pela maior parte dos chamados.

Os resultados obtidos possibilitaram a definição de uma política de atendimento para o *Service Desk*, permitindo assim avaliar as necessidades de atendimento, os tipos de problemas ocorridos e suas variações, as prioridades de atendimento e o tempo gasto pelos técnicos para executarem os serviços de suporte ao usuário e sistemas.

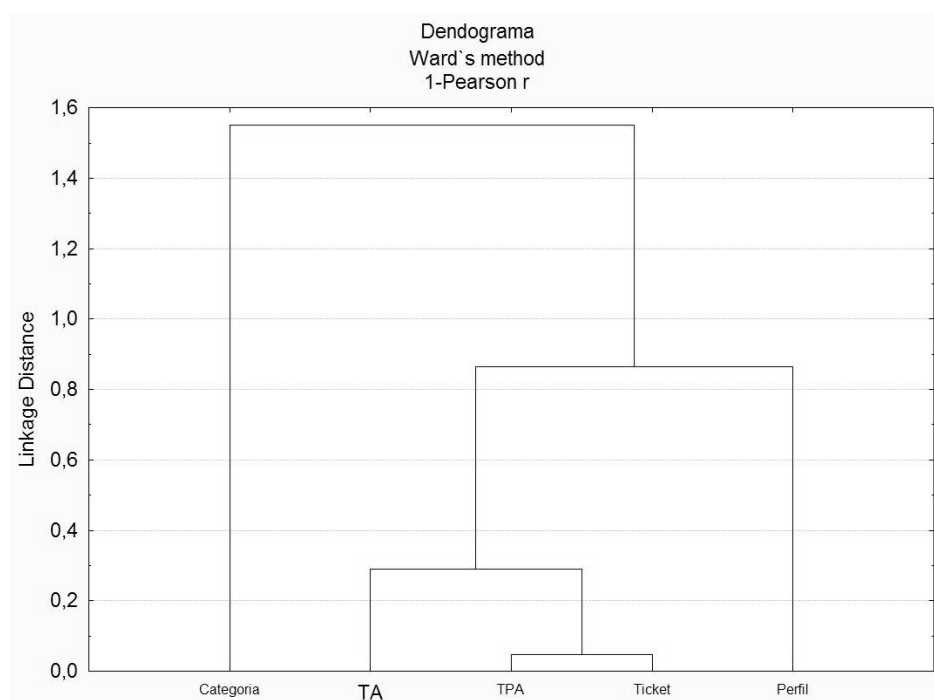
### **5.5.2 Análise e Validação Comprobatória dos Resultados**

Considerando que todo processo está sujeito a variação, por mais que seja perfeito, e baseado no fato de que todo resultado está propenso a variabilidade, a execução dos testes apresentados nesta seção tem por objetivo validar, de forma crítica, as saídas fornecidas pelo SDvPC. Para alcançar esta validação e confirmar o cenário dos incidentes projetados pelo Portal, foram implementadas análises baseadas em controle estatístico de processos e ferramentas da qualidade, a fim de comprovar e facilitar o entendimento das informações armazenadas na base de conhecimento do sistema.

Em um primeiro momento foi utilizada uma técnica de representação denominada dendograma ou *cluster*, um gráfico de agrupamento que apresenta o nível de relação entre as variáveis amostrais. A Figura 41 mostra o dendograma formado a partir da amostra inicial dos dados, mediante a técnica de Análise de Agrupamentos utilizando como parâmetros os

métodos de Ward e a distância 1-Pearson. Por meio desta representação é possível identificar quatro grupos distintos no dendograma:

- O primeiro grupo é formado pelas variáveis “*Ticket*”, que representa o valor quantificado dos reportes de incidentes, e pela variável TPA, tempo previsto para o atendimento ao reporte, demonstrando a relação direta entre o valor do *ticket* e o tempo previsto para o atendimento e evidenciando a característica de ambas variáveis estarem diretamente relacionadas aos SLAs definidos pelo sistema.
- O segundo grupo, formado pela relação entre TA, tempo real de atendimento, e o primeiro grupo (TPA, *Ticket*), mostrando que o atendimento aos reportes está dentro do tempo previsto respeitando o SLA.
- O terceiro grupo apresenta as variáveis Perfil do usuário e TA comprovando que o atendimento aos reportes de incidentes respeita o perfil estabelecido para cada usuário.
- Por último, com uma relação menor tem-se a categoria dos incidentes, que apenas identificam os grupamentos através dos quais o incidentes estão organizados.



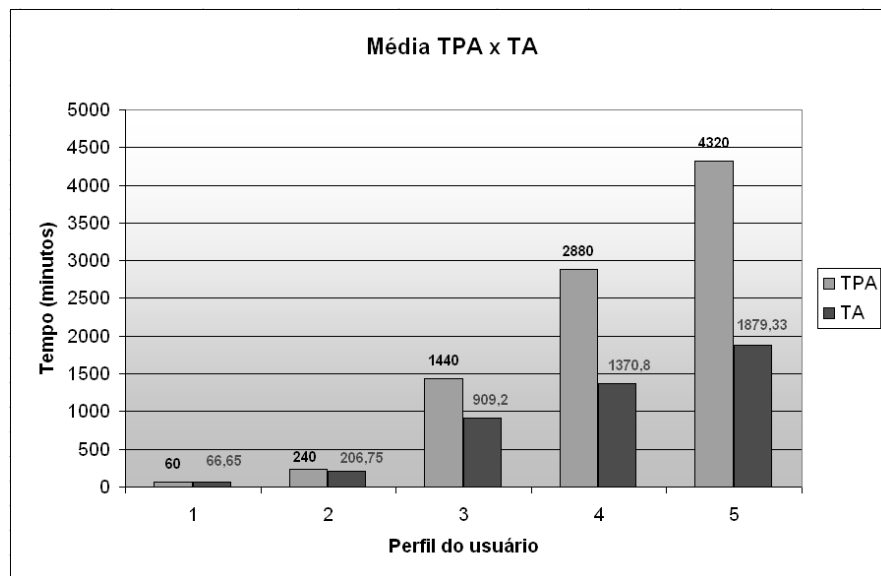
**Figura 41- Dendrograma baseado na amostra de dados do SDvPC.**

A utilização da análise de agrupamentos deixou claro que o grupo de suporte a TI está atuando de forma eficaz no tratamento dos incidentes reportados no âmbito do CRS,

respeitando os níveis de serviço descritos pela biblioteca ITIL. A comprovação do resultado obtido com o dendograma é feita a partir da construção de um histograma, Tabela 3 e Figura 42 , que apresenta a comparação entre os tempos médios, em minutos, de atendimento e o tempo previsto pelo SLA para o atendimento ao usuário, definido dentro de cada perfil.

**Tabela 3: Comparação entre o Tempo Previsto de Atendimento e Tempo Real de Atendimento**

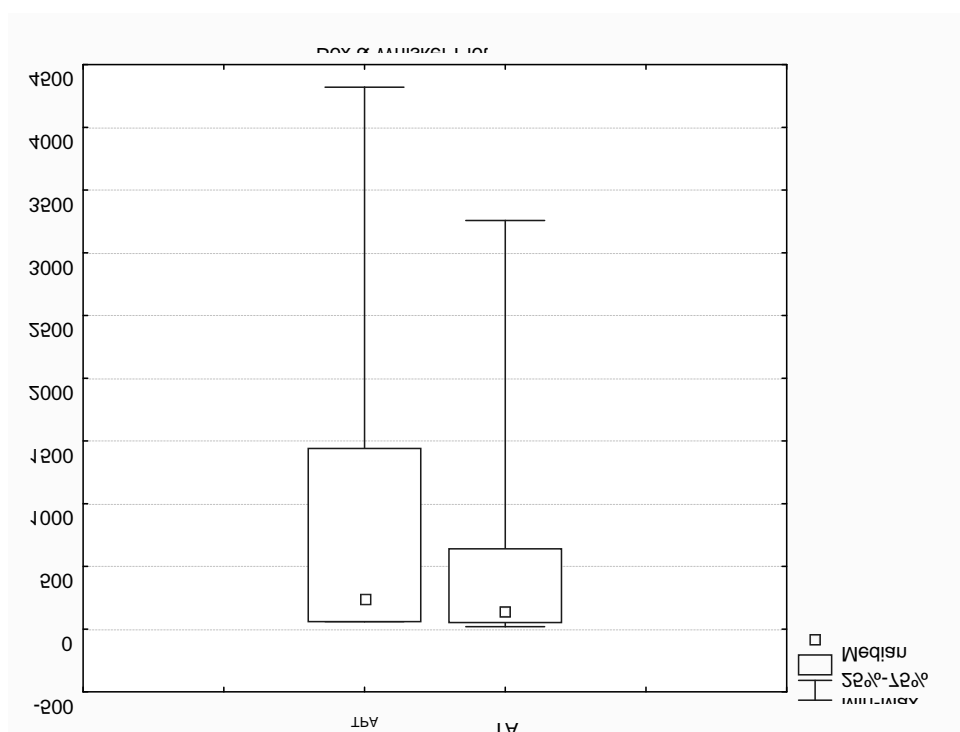
Perfil	TPA	TA
1	60	66,65
2	240	206,75
3	1440	909,2
4	2880	1370,8
5	4320	1879,33



**Figura 42- Histograma da comparação ente TPA e TA.**

Neste ponto um questionamento ainda pode ser apresentado: até onde a proposta do SDvPC pode auxiliar na tomada de decisões pró-ativas e, se os subsídios fornecidos por esta ferramenta são verdadeiramente válidos? Para esclarecer esta questão, foram realizadas algumas reestruturações da amostra de dados utilizadas pelo sistema, e a partir destes novos modelos de dados, foram aplicados outros métodos comumente usados no Controle Estatístico de Processos (CEP), permitindo desta maneira inferir no comportamento no SDvPC e validar seus resultados.

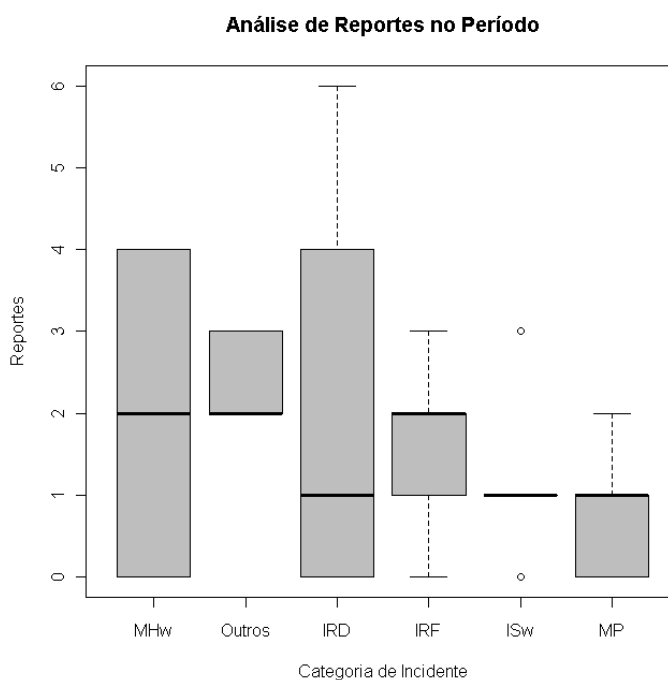
Seguindo na análise dos tempos de SLA, temos a Figura 43, que apresenta o *Box-plot* de comparação entre TPA e TA. A mediana em ambas as variáveis se equivale, porém existe uma discrepância óbvia no limite máximo e no terceiro quartil da variável TPA. Essas observações são explicadas pelo fato dos tempos previstos para o suporte ao usuário serem bem mais elevados que o tempo atendimento real. Esta conclusão pode sugerir a necessidade de realizar um recálculo nos tempos definidos para os níveis de serviço da organização, a fim de estipular metas de tempo menores para a solução dos reportes.



**Figura 43- Box-plot das variáveis TPA e TA.**

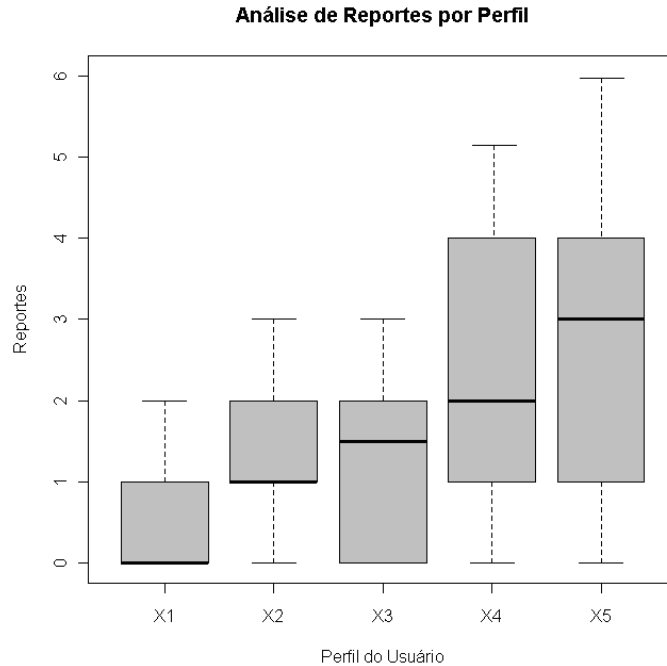
É importante identificar neste ponto o comportamento dos incidentes, para o correto entendimento do que ocorre com os tempos de atendimento aos usuários, e também para que seja possível estipular uma estratégia de redução dos problemas no ambiente computacional da instituição. A análise das categorias ou grupos de incidentes informados no SDvPC, pelos usuários, está representada pelo *Box-plot* da figura 44. O gráfico está constituído pelo conjunto de 6 grupos, através dos quais os incidentes são categorizados: MHw, manutenção de hardware; Outros incidentes não catalogados; IRD, instalação de redes de dados; IRF, instalação de redes de telefonia; ISw, instalação de software; e MP, movimentação de patrimônio. Os grupos MHw e IRD destacam-se por apresentar uma maior dispersão de reportes, caracterizado pelos seus quartís, sendo que a mediana da MHw apresenta-se em um

nível acima da IRD. Em contrapartida o limite superior da categoria IRD é maior que as demais, denotando um número maior de atendimentos para este grupo. Desta maneira percebe-se quais os pontos de falha que devem ser atendidos de maneira precisa. As demais variáveis apresentam um comportamento padrão, o qual deve ser avaliado constantemente, para evitar que anomalias indesejadas possam prejudicar as atividades do CRS.



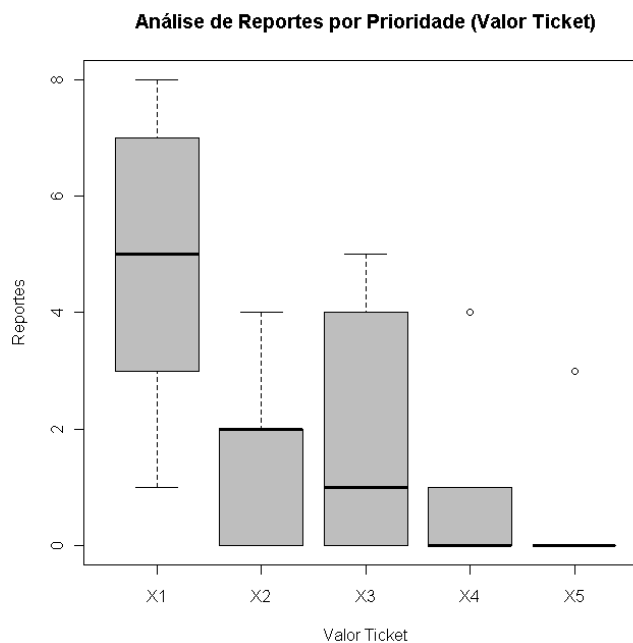
**Figura 44- Box-plot comparativo entre as categorias de incidentes.**

O reconhecimento dos pontos que levam a falha no sistema permite a área de TI prevêr anomalias de forma ágil, e desta maneira alcançar uma pro-atividade intrínseca ao viabilizar planos de prevenção e conscientização dos recursos humanos na empresa. Percebe-se então que os usuários são peças chave para o correto funcionamento dos recursos de TI, com isso é necessário reconhecer como ocorre a interação destes com o SDvPC. A Figura 45 caracteriza esta relação, nela são apresentados os reportes de incidentes organizados por perfil de usuário. Denota-se que integrantes do CRS com o perfil de nível 1 (X1), mais alto, tem uma baixa relação com o sistema, devido ao baixo número de reportes caracterizados por uma mediana próxima a zero. Os usuário de perfil 4 e 5 (X4 e X5), possuem o maior índice de reportes submetidos ao Portal, tendo um comportamento semelhante, porém os usuários com o nível mais baixo (X5) apresentam uma mediana e limite superior bem mais elevados.



**Figura 45- Box-plot para análise do índice de reportes por Perfil de Usuário.**

Neste ponto já são conhecidos os tempos de atendimento aos usuários, as categorias com maior volume de reportes e quais são os perfis com responsabilidade sobre estes índices. Agora resta apenas identificar a quantidade de reportes por prioridade de *ticket*, que estão sendo atendidos pelo sistema. Tendo em vista que esta prioridade é definida com base em uma relação entre o perfil do usuário e a descrição do problema, é possível enxergar pontualmente quais são as anomalias que devem ser estudadas e tratadas de forma preventiva. O *Box-plot* da Figura 46 mostra que o maior índice de incidentes está sendo classificado com perfil mais alto (X1), que possui a SLA com menor tempo de resposta (60 minutos), em segundo tem-se reportes de prioridade média (X3) com uma baixa mediana, porém um número razoável de *tickets*. Já em contrapartida, têm-se os reportes (X5) de menor prioridade e com tempo de atendimento planejado, que não se destacam no gráfico, pois possuem uma mediana em zero.



**Figura 46- Box-plot para análise do índice de reportes por Prioridade dos *Tickets* reportados.**

Como parte final da análise comprobatória dos resultados obtidos pelo SDvPC, foram utilizadas 2 ferramentas da qualidade (Pareto e Gráficos de Controle), a fim de fundamentar as análises estatísticas realizadas. Originalmente estas ferramentas têm por objetivo mensurar, analisar e propor soluções para problemas que possam afetar o funcionamento de determinados processos. Neste trabalho, as ferramentas de qualidade são usadas como técnicas para visualizar de forma analítica as informações disponíveis na base de conhecimento do Portal, e em um nível mais alto validar todas as análises realizadas até este momento.

A primeira ferramenta a ser aplicada foi o Pareto, que é uma representação gráfica utilizada para dividir um problema maior em partes menores, e assim mostrar a contribuição de cada item para o efeito total, e desta forma identificar oportunidades de melhoria para cada um deles. Este tipo de gráfico foca nos pontos de maior importância separando os muitos triviais dos poucos essenciais. Basicamente o princípio de Pareto defende que “um número pequeno de causas é responsável pela maior parte dos problemas”, sendo assim, o tratamento de um pequeno número de causas, resolve uma grande parte dos problemas.

Seguindo a ordem das análises estatísticas realizadas, o primeiro gráfico de Pareto implementado verifica as categorias de incidentes reportadas, Figura 47, onde são apresentados os 6 grupos categorizados pelo SDvPC. Destacam-se nesta análise as 3 primeiras



categorias (Outros, MHw e IRD), sendo que o grupo outros possui o maior número de chamados. O tratamento deste problema deve ser realizado com o uso do FMEA para a verificação da incidência e dos tipos de chamados que estão sendo agrupados na categoria outros.

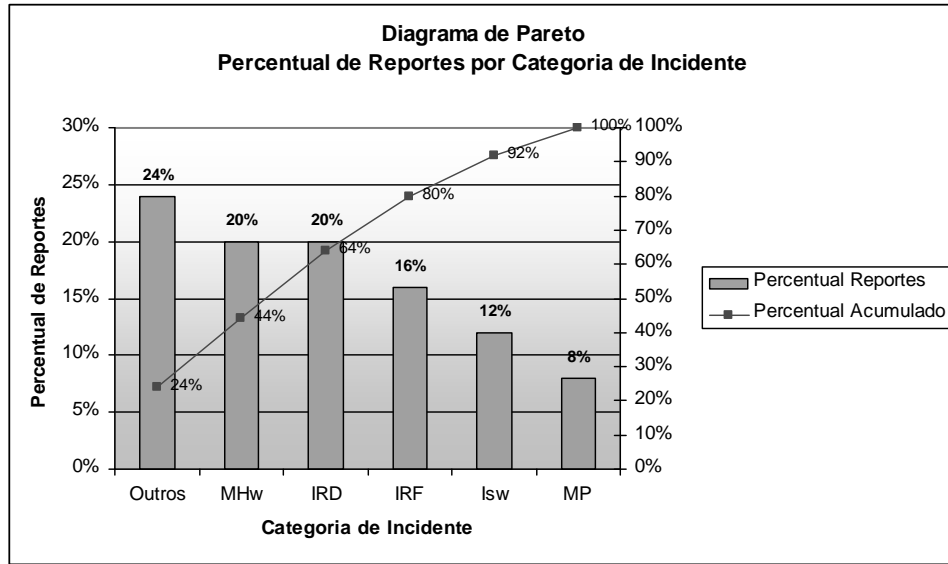


Figura 47- Diagrama de Pareto – Análise das categorias com maior número de incidentes.

A análise do reportes por perfil de usuário na Figura 48, comprova os fatos identificados com as técnicas estatísticas, os usuários com perfis 5 e 4 apresentam-se como os poucos essenciais, desta forma a maior atenção as necessidade destes integrantes reduzirá consideravelmente o índice de problemas no âmbito do CRS.

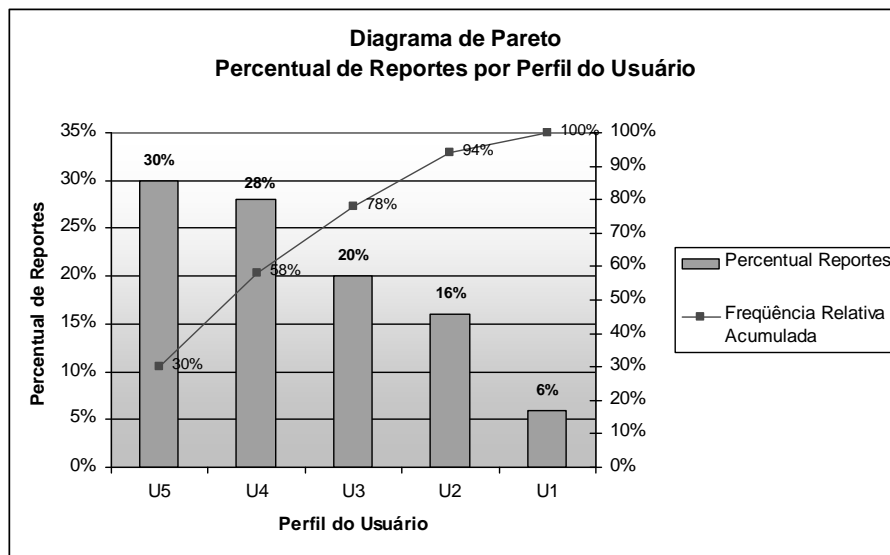
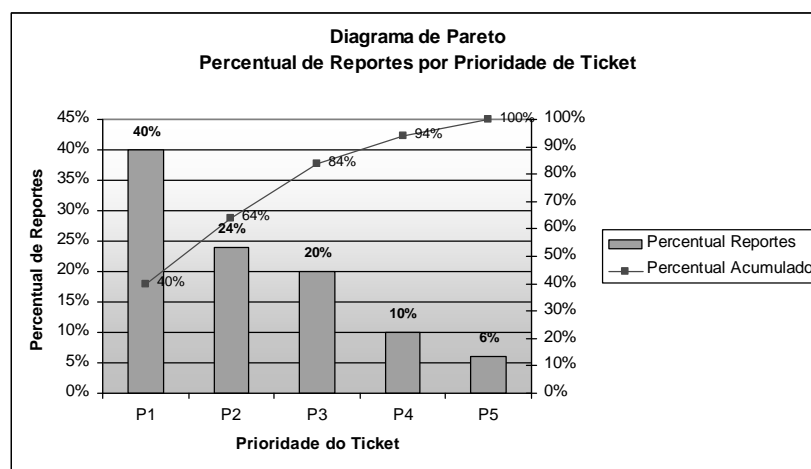


Figura 48- Diagrama de Pareto – Análise dos perfis de usuários.

A verificação sobre as prioridades dos *tickets* lançados no SDvPC, conforme a Figura 49, mostra que os reportes com maior prioridade de atendimento (P1) caracterizam-se como os poucos triviais, isto é, concentram 40% do número de chamados do sistema. Este resultado confirma as análises anteriores, onde mostram que a grande maioria dos problemas deve ser resolvido em um curto prazo de tempo (60 minutos), de acordo com a SLA prevista. Nesta situação a proposição de métodos mais eficazes e ágeis no atendimento ao usuário podem amenizar este quadro.

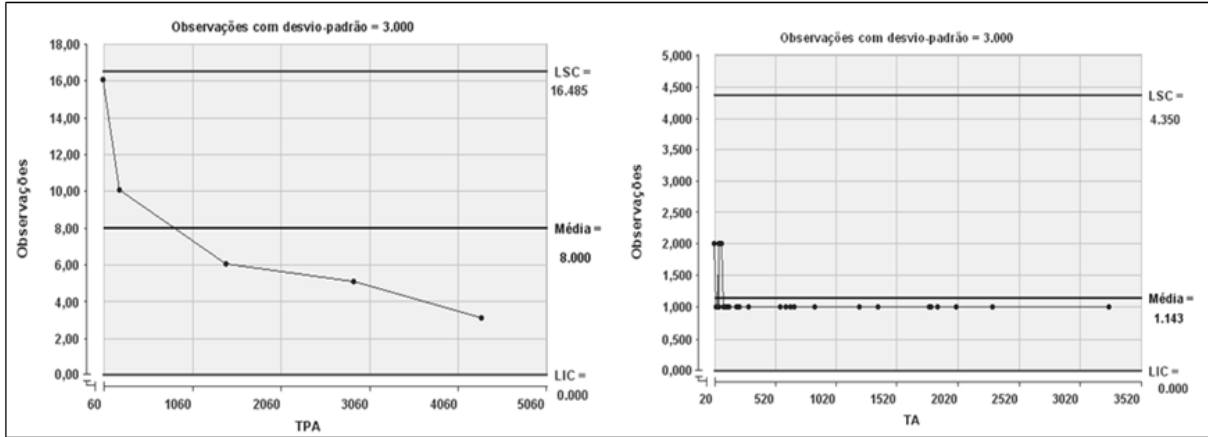


**Figura 49- Diagrama de Pareto – Verificação dos reportes por Prioridade dos *Tickets*.**

E por fim, a última ferramenta da qualidade utilizada para análise das informações foi o gráfico de controle. Este tipo de gráfico é indicado para avaliar a estabilidade de um processo ou variável, distinguindo as suas variações. De forma simples, sintetiza um conjunto de dados, por meio de métodos estatísticos, para observar as mudanças dentro da amostragem, verificando características fundamentais: centralização, determinada pela média, e dispersão, determinada pelo desvio-padrão. Desta maneira é viável investigar e identificar os fatores que afetam determinado processo. O gráfico é formado por uma linha média, que representa o valor médio da amostra, e pelo limite inferior de controle (LIC) e limite superior de controle (LSC), os quais determinam os limites para que os pontos traçados no gráfico mostrem se o processo está ou não sob controle.

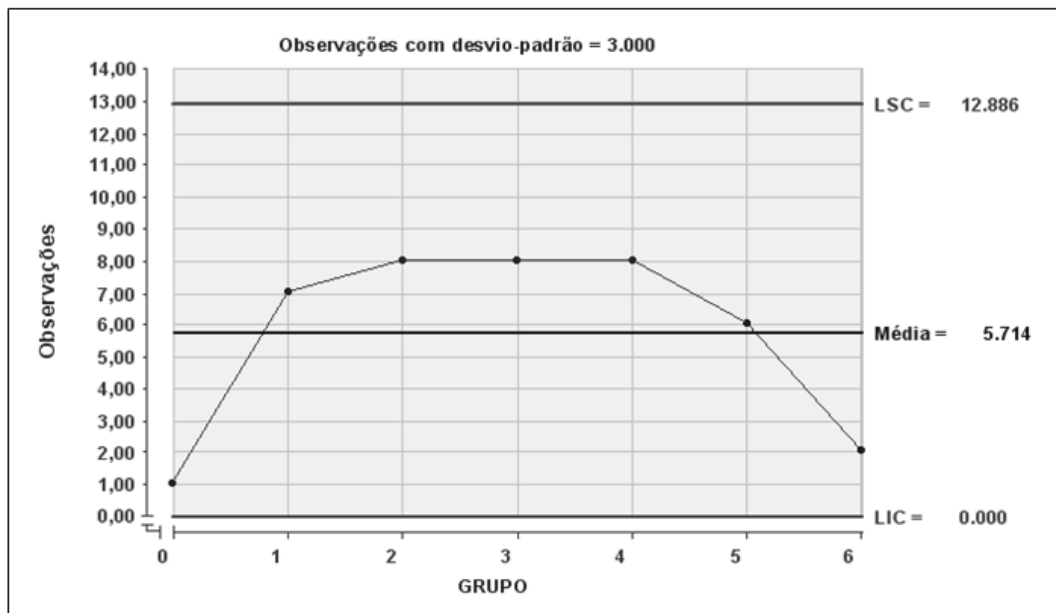
A Figura 50 destaca uma comparação através de Gráficos de Controle das variáveis, tempo previsto de atendimento e tempo real de atendimento. Esta representação confirma os resultados das Figuras 42 e 43, onde os pontos que caracterizam o tempo real de atendimento

ao usuário estão abaixo da média, ou seja, o atendimento aos chamados esta dentro dos níveis de serviços determinados.



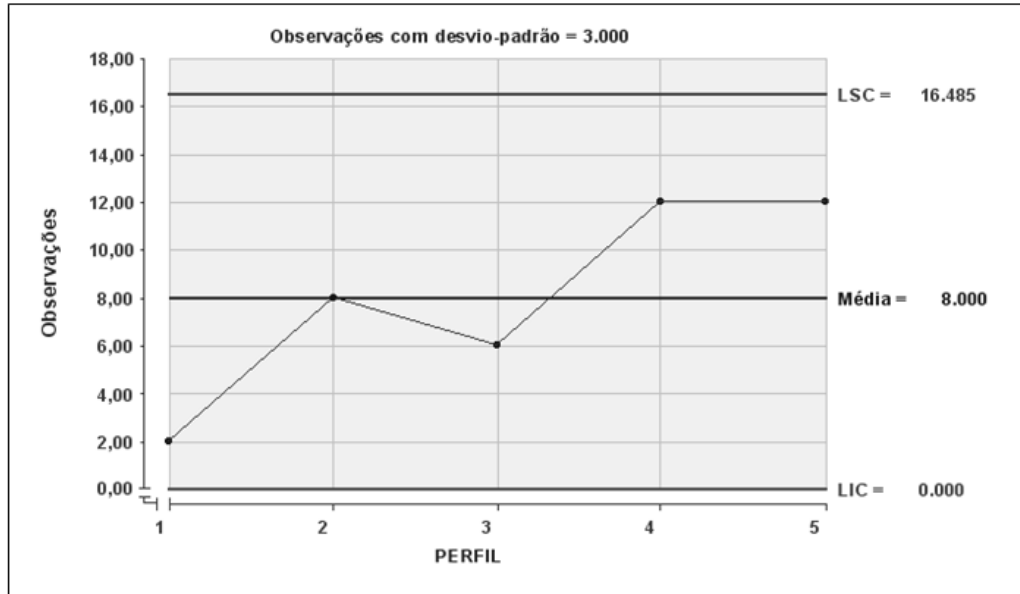
**Figura 50- Gráfico de Controle – Comparação entre os tempos de atendimento.**

Na análise dos incidentes por categoria, Figura 51, os pontos de controle mostram certa estabilidade, com valores acima da média, para as variáveis 2, 3 e 4 (Outros, IRD, IRF). Este fato é plausível, considerando a proximidade do número de reportes em cada grupo.



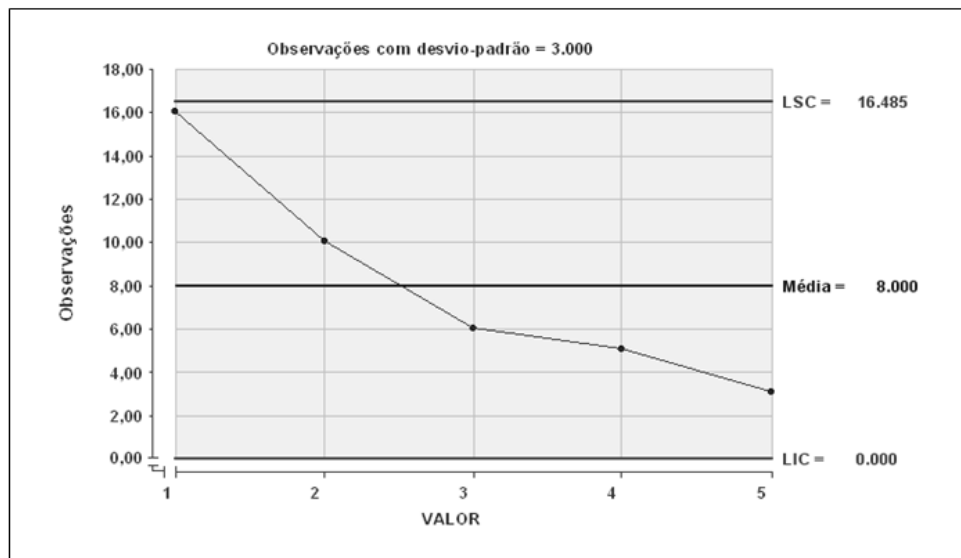
**Figura 51- Gráfico de Controle – Análise dos reportes por categoria dos incidentes.**

Como nas observações anteriores, os pontos de controle da Figura 52 comprovaram que o maior número de reportes é registrado pelos usuários com perfil 4 e 5.



**Figura 52- Gráfico de Controle – Análise dos reportes considerando o Perfil dos usuários.**

A observação da Figura 53 mostra os reportes com prioridades 1 e 2, níveis mais altos de prioridade, como os mais frequentes, ou seja, a grande maioria dos reportes registrados no SDvPC, são de incidentes de alta prioridade e com a necessidade de baixo tempo de resposta pela equipe de TI, comprovando assim, as análises anteriores.



**Figura 53- Gráfico de Controle – Verificação dos reportes considerando a prioridade de atendimento.**

### 5.5.3 Obtenção da Pró-atividade

Para alcançar o alinhamento dos serviços de TI com as necessidades da organização é necessária além de uma postura pró-ativa, uma correta interpretação da importância dos mesmos para o negócio. A partir desta percepção é comum as empresas buscarem a otimização de seus processos e a redução de custos e riscos, garantindo assim um nível de maturidade da área de TI. Uma medida eficiente para a redução dos riscos e custos em uma empresa é a extração do conhecimento, intrínseco, nas informações de incidentes reportadas pelos usuários, e o uso deste conhecimento aliado a boas práticas de governança. Este posicionamento permite aos responsáveis pela área de TI atuarem de forma pró-ativa no âmbito da organização, com a identificação de oportunidades e a implementação de estratégias eficazes de gestão e suporte a decisão.

No contexto deste trabalho, o alcance da pró-atividade nas ações do grupo de TI, está calcada sobre a utilização racional e eficaz dos resultados fornecidos pelo SDvPC, como aporte para algumas atividades:

- manutenção do sistema de gestão de incidentes;
- controle das atividades desenvolvidas pela equipe de TI;
- identificação e estudo sobre a eficiência dos processos de TI;
- capacidade de sugerir e implementar melhorias relacionadas a área de TI .

Este conjunto de ações permite que sejam avaliados e tratados, com a devida atenção, alguns fatores críticos de sucesso, os quais incidem diretamente sobre a capacidade e eficiência da equipe de TI em atender as necessidades da organização. Os principais benefícios que podem ser descritos a partir desta postura não reativa são:

- a busca constante de aperfeiçoamento da equipe de suporte, afim de qualificar e agilizar o atendimento aos incidentes;
- a diminuição de incidentes mais simples, através da prevenção e eliminação de erros conhecidos;
- adequação dos níveis de serviço definidos para os processos de TI;
- organização e otimização nas atividades de suporte aos usuários e sistemas;
- satisfação do usuário, através do reconhecimento e atendimento eficaz de suas necessidades;
- diminuição do tempo previsto para o retorno sobre o investimento (ROI);
- disponibilidade da equipe de TI para atuar em atividades técnicas específicas;
- redução nos custos relacionados a gestão de incidentes.

Em resumo, a pró-atividade almejada durante o desenvolvimento desta dissertação pode ser descrita, de forma simples, como a capacidade técnica de coletar informações que traduzidas em conhecimento específico, viabilizam uma maneira eficaz de otimizar e qualificar as ações relacionadas a Tecnologia da Informação em diferentes ambientes, utilizando diferentes metodologias ou *frameworks* para isso.

## 5.6 CONSIDERAÇÕES PARCIAIS

A integração dos recursos computacionais disponíveis pelas linguagens de programação Python, Java e R, permitiram a modelagem e implementação de uma ferramenta especialista, o SDvPC. O conjunto de funcionalidades agregadas ao Portal Corporativo o torna uma eficiente ferramenta para a comunicação interna, interação de processos de controle e de usuários em um ambiente de rede. Com base nesse conjunto de possibilidades a disponibilização de um SD integrado as ferramentas do Portal, torna-se uma opção viável para a difusão de um método eficaz de reporte dos incidentes de segurança identificados pelos usuários. O SD serve também como um canal único de comunicação entre a equipe de suporte e os integrantes da organização.

Esta aplicação como um todo permite o sistema armazenar um conjunto de informações relevantes sobre os problemas identificados e reportados na área de TI. Estes fatos compõem uma base de conhecimento, sobre a qual é possível desenvolver um conjunto de análises estatísticas, a fim de obter informações relevantes para o apoio a decisão. Em primeiro momento foram realizados testes com métodos multivariados sobre a amostra de dados retirada do SDvPC, os quais apresentaram um panorama real dos incidentes no âmbito do CRS, mostrando suas principais características. Os resultados obtidos a partir do sistema foram comprovados com o uso de algumas ferramentas para o Controle Estatístico de Processos e ferramentas da qualidade. A questão da pró-atividade do *Service Desk* está calcada na capacidade do mesmo em extrair informações relevantes e que sirvam de subsídios reais e válidos para tomada de decisões. Os dados de saída do SD podem ser considerados recursos necessários para o estudo, avaliação e desenvolvimento de processos e procedimentos ativos no tratamento das anomalias reportadas ao sistema.

## CAPITULO 6

### CONCLUSÕES

Os resultados obtidos com este trabalho mostram que o reconhecimento e tratamento correto dos incidentes de TI consistem em uma valiosa solução para a otimização do gerenciamento de todos os aspectos de serviços, recursos e suporte para área de TI na organização. A grande quantidade de dados, relacionados com os incidentes e repassados pelos usuários, denota a importância da informação como objeto para a formulação de estratégias e apoio na tomada de decisão. Os *frameworks* de Governança atuais caracterizam a gestão de incidentes como uma prática fundamental para o alinhamento da TI com as regras de negócio, porém destacam duas peculiaridades nesta gerência que devem ser observadas: a coleta dos reportes de incidentes e o problema da validação científica das informações coletadas.

Como solução ao problema da coleta dos reportes de incidentes, foi proposto, implementado e testado um Portal Corporativo (SDvPC). Este consiste em uma ferramenta capaz de manipular dados institucionais e permitir o acesso as informações por meio de uma aplicação disponível na rede interna da organização. Este software apresentou um excelente desempenho, pois teve agregado ao seu conjunto de funcionalidades, um sistema de *Service Desk*, o qual serviu como interface e ponto único de contato entre os usuários, que descrevem os problemas e solicitam apoio técnico em informática, e a equipe de suporte a TI, que atende esses chamados. O resultado da integração entre a tecnologia de Portal Corporativo e *Service Desk* foi uma aplicação capaz de receber e armazenar, de forma centralizada, todos os reportes de incidentes no âmbito da instituição, dando origem a uma grande base de conhecimento.

Para a validação científica dos dados qualitativos armazenados na base de conhecimento do Portal, foi necessário o desenvolvimento de uma técnica de quantificação destas informações. A implementação da metodologia de conversão justificou-se pelo fato de que o conjunto das técnicas interpretativas disponíveis, que podem ser aplicadas sobre os dados qualitativos do Portal, objetiva apenas descrever ou expressar o sentido de um fenômeno, mas a solução esperada nesta pesquisa deve obrigatoriamente passar por um processo de inferência estatística, a fim de apresentar e identificar dados, indicadores e

tendências observáveis. A partir da quantificação e dos resultados obtidos pela análise estatística sobre os dados, denotou-se a eficácia da solução para conversão das informações, a qual foi implementada usando como requisitos a pontuação (escala likert) das características individuais de cada usuário, avaliando seu perfil, e também considerando suas atividades e posição no organograma da organização.

A avaliação da base de conhecimento gerada, a partir dos reportes de incidentes e da quantificação dos mesmos, expôs a inviabilidade da aplicação de uma análise estatística simples para a obtenção do conhecimento agregado a esses dados. O complexo número de variáveis contidas no SDvPC levou a utilização de técnicas multivariadas para o reconhecimento da sua estrutura e avaliação da situação atual dos processos e serviços de TI no CRS. De forma a alcançar o melhor resultado da análise sobre as informações, foram aplicadas diferentes técnicas de origem multivariada, das quais se destacaram as seguintes: dendograma, usado a fim de reconhecer o grau de agrupamento das variáveis; análise fatorial, utilizada com intuito de reduzir o número de fatores; o teste do qui-quadrado, a fim de comprovar a interação entre as variáveis e por fim aplicados métodos de representação gráfica que facilitaram a visualização da relação entre as variáveis e os resultados obtidos. Uma importante constatação é a existência de uma relação causa-efeito nos incidentes estudados, o que permite um grau de previsão dos fenômenos de suporte identificados. Neste ponto é possível afirmar que o uso de técnicas estatísticas multivariadas possibilitou a detecção e interpretação, de forma precisa, dos grupos de pessoas que realmente necessitam de prioridade no atendimento de suas solicitações. A correta interpretação deste conhecimento possibilita que a equipe de suporte a TI seja capaz de traçar metas e prioridades de atendimento, com base na aplicação das ferramentas descritas, o que permite o planejamento do processo de apoio a decisão e satisfaz as necessidades de entendimento das informações armazenadas na base de dados do Portal.

Para a validação dos resultados obtidos com o SDvPC foi aplicado um conjunto de ferramentas de qualidade e também técnicas estatísticas, a fim de identificar o grau de precisão das informações geradas pelo Portal. Os dados desta segunda análise serviram de fundamentação para as afirmações geradas pelo *Service Desk*, além de garantir as estratégias visionadas para área de TI. A comprovação dos resultados do SD foi baseada no uso das seguintes ferramentas: histogramas, que mostraram a comparação entre os tempos de atendimento; gráficos *box-plot*, que apresentaram as relações entre quantidades de reportes, perfis de usuários e prioridades dos *tickets*; os gráficos de Pareto mostraram a relevância dos



*tickets*, da categorização dos chamados e avaliação dos perfis de usuários; e os gráficos de controle que avaliaram as observações e desvio padrão de todas as variáveis do sistema. Ao final, a interpretação dos dados da análise comprobatória resultou em informações já conhecidas, ou seja, comprovaram as instâncias observadas com as técnicas multivariadas. Com base nestes dados e sabendo que a tomada de decisão pró-ativa está diretamente relacionada ao conhecimento dos processos e de seus resultados, é possível afirmar que o uso adequado das informações obtidas por meio do Portal, aliadas as boas práticas de governança em TI, permitem o tratamento não reativo dos incidentes de segurança, dentro da organização. Essa postura evita a repetição na correção de problemas simples, e a avaliação de soluções mais eficientes para problemas complexos, permitindo a equipe de suporte a TI atentar e disponibilizar mais tempo para o tratamento dos poucos essenciais, sendo que os muitos triviais desta maneira são resolvidos de forma mais ágil.

Ao término deste trabalho é possível identificar que todos os objetivos traçados, no início das pesquisas, foram alcançados e comprovados com êxito. A proposta do Portal Corporativo integrando uma ferramenta de *Service Desk*, de acordo com as orientações da biblioteca ITIL, mostrou-se eficaz na resolução dos problemas de comunicação interna no âmbito da organização, além de centralizar os reportes de incidentes de segurança, com o intuito de observá-los e avaliá-los de forma pontual. A entrada em produção do SDvPC fundamentou as conjecturas apresentadas por este trabalho, firmando a veracidade nos resultados obtidos a partir da coleta centralizada e análise, por meio de técnicas estatísticas, das informações referentes aos incidentes no CRS e a disponibilização de uma base de conhecimento realmente válida. Além dessas observações, podem-se avaliar inúmeras contribuições do SD para o grupo de suporte a TI, entre elas: contabilização da carga de trabalho da equipe, cálculo dos tempos de atendimento (por técnico, grupo de incidente, etc.), identificação dos usuários com maiores dificuldades, implementação de *ranking* de reportes e usuários, definição de programas de treinamento para a equipe de TI, publicação dos resultados do SDvPC em eventos científicos.

## **6.1 TRABALHOS FUTUROS**

Como trabalhos a serem realizados em complemento a esta dissertação, podem ser avaliados e desenvolvidos outros módulos a serem integrados ao SDvPC, tais como:

- Módulo de comunicação entre o SDvPC com o NOC (*Network Operation Center*) do CRS, afim de automatizar o lançamento de reportes relacionados ao funcionamento da rede e de servidores de dados;
- Módulo de interação do SDvPC e a ferramenta de controle de acesso pessoal as instalações do Centro, com o intuito de disponibilizar informações sobre recursos humanos a alta administração.
- Módulo de gerencia de configuração, baseado em regras do ITIL, se caracteriza por uma aplicação destinada a avaliar e autorizar mudanças de configuração estrutural e de sistemas na área de TI.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, A.; COSTA A. P.; MIRANDA C. **Informação e gestão. In: ALMEIDA, A.; RAMOS, F. Gestão da informação na competitividade das organizações.** 2. ed, Recife: Editora da UFPE, 2002.
- AMARAL E. M. H.; BASTOS, C.; FIGUEIREDO, MARIA A.; SOUZA, ADRIANO M.; NUNES, RAUL C. **Análise Estatística Multivariada de Incidentes de Segurança em Redes de Computadores** In: Encontro Nacional de Engenharia da Produção, RJ, 2008.
- ANDERSON, CHRISTIANO. *The definitive Guide to Django*. Ed. Apress, 2007.
- ANSOFF, H. Igor **Estratégia Empresarial**. São Paulo: McGraw-Hill do Brasil, 1977.
- Associates, Computer. “**Documentação on-line**”. *Computer Associates International*, 2000.
- BAKKE, HANNE A.; LEITE, ALEXANDRE S. M.; SILVA, LUIZ B. **Multivariate Statistics: Factorial Analysis Application In Production Engineering** In: Revista Gestão Industrial, Universidade Tecnológica Federal do Paraná – UTFPR, v. 04, 2008.
- BARBOSA, G. R. **Sistemas de Apoio a Decisão sob o enfoque de Profissionais de Tecnologia da Informação e Decisores**. Dissertação de Mestrado, PPGEP, UFPE, Recife, 2003.
- BARDIN, L. **Análise de conteúdo**. 3. Ed. Lisboa: Edições 70, 2004.
- BARROSO, L. P.; ARTES, R. **Análise multivariada**. In: SEAGRO – Simpósio de Estatística Aplicada a experimentação agrônômica, UFLA, Lavras, 2003.
- BEUREN, Ilse Maria. **Gerenciamento da Informação: um recurso estratégico no Processo de gestão empresarial**. São Paulo: Atlas, 1998.
- BOCK, R.D. *Multivariate statistical methods in behavioral research*. McGraw Hill, 1975.
- BOM, J. V. **Fundamento do gerenciamento de Serviços em TI Baseado no ITIL**. ITSMF, Van Haren Publishing, 2006.
- BOUZON, M. e CORRÊA, Kleber C. **Tecnologia da Informação Aplicada a Logística**, Grupo De Estudos Logísticos, Universidade Federal De Santa Catarina, 2007.
- BRANDÃO, Marinho; (2009) **Aprendendo Django no Planeta Terra**, pg 39-43
- CALVO, M. C. M. **Estatística descritiva**. Florianópolis: UFSC, 2004.
- CAMPOS, A. **Portais Acessos dinâmicos à economia digital**. *eManager*, SãoPaulo, v. 2, n. 20, p. 6-8, 2001.

- CASTILHO, K. Adriano, **O Gerenciamento dos Riscos através da Metodologia FMEA – Análise de Modo e Efeitos de Falha**, Monografia - Pós-Graduação em Gerência de Projetos com Ênfase nas Práticas do PMI, Universidade São Judas Tadeu, 2007.
- CHANG, Ching-Liang; LIU, Ping-Hung; WEI, Chiu-Chi. **Failure mode and effects analysis using grey theory**. *Integrated Manufacturing Systems*, v.12, n.3, pp. 211- 216, 2001.
- CLERICUZI, Adriana Z.; ALMEIDA, Adiel T.; COSTA, Ana P. C. S. **Aspectos relevantes dos SAD nas organizações: um estudo exploratório** In: Scielo Brasil, São Paulo, vol.16, 2006.
- COLLIS, J.; HUSSEY, R. **Pesquisa em administração: um guia prático para alunos de graduação e pós-graduação**. 2. ed. Porto Alegre: Bookman, 2005.
- COLLINS, DANIEL. *Data warehouses, enterprise information portal, and the SmartMart meta directory*. *Information Builders Systems Journal*, v. 12, n. 2, p. 53-61, 1999.
- COSTA, Geraldo M., MIRANDA, João Paulo de. **Implantando Processos de Desenvolvimento de Software**. *Developers*, p. 28-30, set. 1999.
- DAVIS, C. B. ; Olson M. *Management information systems: Conceptual foundations, structure and development*. New York: McGraw-Hill, 1974.
- DAVENPORT, T., PRUSAK L.. **Conhecimento empresarial: como as organizações gerenciam seu capital intelectual**. Rio de Janeiro: Campus, 1998.
- DAVIS, J. C. *Statistics and Data Analysis in Geology: 2th Ed., Jhon Wiley and Sons, Inc.*, 1986.
- DIAS, CLÁUDIA A. **Portal Corporativo: conceitos e características**. *Ci. Inf.* vol.30 no.1, UnB, Brasília, 2001.
- DIAS L.; MOUSSEAU V. *IRIS - Interactive Robustness analysis and parameter's Inference for multiple criteria Sorting problems - User Manual*, Document of INESC Coimbra, 2003.
- DIAS, L. C.; J. N. CLÍMACO. *Additive Aggregation with Variable Interdependent Parameters: the VIP Analysis Software*, *Journal of the Operational Research Society*, 2000.
- DOWNING, Douglas; CLARK, J. **Estatística Aplicada**. 2. ed. São Paulo: Saraiva, 2006.
- DRUCKER, Peter F. **Administração, Tarefas, Responsabilidades e Práticas**. São Paulo: Pioneira, 1975.
- ECKERSON, Wayne. **Plumtree blossoms: new version fullfills enterprise portal requirements**. Boston, MA: Patricia Seybold Group, June 1999. [online], abril 2000. [<http://www.plumtree.com/moreinfo/specialoffer.htm>].
- EHLERS, RICARDO S. **Introdução a Estatística**. Departamento de Estatística Universidade Federal do Paraná, 3ª Edição, 2006.

- FENANDES, A. A. e ABREU, V. F. **Implantando a Governança de TI da Estratégia à gestão dos Processos e Serviços**. 2ª Edição, Brasport Editora, 2008.
- FERNANDES, TALIZE A. G.; LIMA, JOÃO E. **Uso de Análise Multivariada para a Identificação de Sistemas de Produção**. Pesquisa Agropecuária Brasileira 26(10), 1991.
- FERREIRA, DANIEL F. **Estatística Computacional Utilizando R**. UFLA, Lavras-MG, 2008.
- FERREIRA, DANIEL F. **Estatística Multivariada**. Editora UFLA, Lavras-MG, 2008.
- FERRO, CAMILA; VANTI, ADOLFO A. **Tecnologia de Portal de Conhecimento Corporativo: um aporte teórico como fundamento à proposta de implementação na área de recursos humanos de uma universidade**. Revista Sistemas de informação, nº16 (19-30), 2002.
- FIRESTONE, JOSEPH M. (2000) *DKMS Brief No. Eight : Enterprise information portals and enterprise knowledge portal*. Disponível em: <http://www.dkms.com/EKPandEIP.html>. Acessado em 07/07/2009.
- FREITAS, H., BECKER, J. L., KLADIS, C. M. e HOPPEN, N. **Informação e Decisão: sistemas de apoio e seu impacto**. Porto Alegre: Ortiz, 1997.
- FREITAS, Henrique; MUNIZ, Raquel Janissek. *Análise quali ou quantitativa de dados textuais*. Quanti & Quali Revista, 2002.
- FREUND, J. E.; SIMON, G. A. **Estatística aplicada**. 9. ed. Porto Alegre: Bookman, 2000.
- FUGLSETH, M.; GRONHAUG, K. *Can computerised market models improve strategic decision-making ? An exploratory study*. *Journal of Socio-Economics*, 2003.
- GHEZZI, C., JAZAYERI, M. and MANDRIOLI, D. *Fundamentals of Software Engineering*. New Jersey, Prentice-Hall International Editions, 1991.
- GILKA ROCHA BARBOSA ; ADIEL TEIXEIRA DE ALMEIDA **Sistemas de Apoio a Decisão sob o Enfoque de Profissionais de TI e de Decisores**. Universidade Federal de Pernambuco - XXII Encontro Nacional de Engenharia de Produção Curitiba – PR, 2002.
- GOTTSCHALK, ALEXSSANDRO A. **Proposta de uma solução EM TI para o gerenciamento dos níveis de serviços de rede alinhados ao planejamento estratégico**, Novo Hamburgo, 2005.
- GUEDES, GILLEANES T. A. **UML: Uma Abordagem Prática**, Edição 1, Novatec, 2004.
- GUGLIELMETTI, Fernando Ribeiro; MARINS, Fernando Augusto Silva; SALOMON, Valério Antonio Pamplona. **Comparação teórica entre métodos de auxílio à tomada de decisão por múltiplos critérios**. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 23., 2003, Ouro Preto. *Anais...* Minas Gerais, 2003.
- GUNTER, Barrie. *News and the Net*. New Jersey: Lawrence Erlbaum Associates, Inc, 2003.
- HAIR, J. et al. **Análise Multivariada de Dados**, 5ª ed. Porto Alegre: Bookman, 2005.

- HAMMOND, J. S.; KEENEY, R. L.; RAIFFA, H. **Decisões Inteligentes**. São Paulo: Campus, 1998.
- HÄRDLE, W.; SIMAR, L. *Applied multivariate statistical analysis*. 2. ed. Berlin: Springer, 2007.
- HOFFMANN, L. T. **Sistema de Apoio à Decisão em Escalada Alpina**. Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, 2001.
- IEC: *Service-Level Management. Service-Level Management*. Tutorial publicado no site do *International Engineering Consortium (IEC)*, sem data de publicação. Disponível em [http://www.iec.org/online/tutorials/service\\_level/](http://www.iec.org/online/tutorials/service_level/).
- ITIL v.2 *Information Systems Audit and Control Foundation, V. 2.0* (ISACF), ISACA, 2008.
- JAMIL, George Leal. **Repensando a TI na Empresa Moderna**. Rio de Janeiro: Axcel Books, 2001.
- JOHNSON, R.A.; WICHERN, D.W. *Applied multivariate statistical analysis*. 4th edition. Prentice Hall, New Jersey, 1998.
- KAMINSKI, H. e PERRY, M. *SLA Automated Negotiation Manager for Computing Services*. In: *IEEE INTERNATIONAL CONFERENCE ON E-COMMERCE TECHNOLOGY AND IEEE INTERNATIONAL CONFERENCE ON ENTERPRISE COMPUTING, E-COMMERCE, AND E-SERVICES (CECEEE)*, 2006.
- KRUSKAL, J. B.; WISH, M. *Multidimensional Scaling*. Sage Publications, Newbury Park, CA, 1978.
- LACHTERMACHER, Gerson. **Pesquisa Operacional na tomada de decisões: uma modelagem em excel**. Rio de Janeiro: Campus, 2002.
- LAUDON, Kenneth C.e LAUDON, Jane P *Management Information Systems* (7<sup>a</sup> edição). Upper Saddle River: Prentice Hall, 2001.
- LEÃO, MAXIMILLIAM M. **Um Estudo Sobre a Concepção e Desenvolvimento de Acordos de Nível de Serviço para Ambientes de Redes de Computadores**. 2002.
- LEDO, C. A. da S.; FERREIRA, D. F.; RAMALHO, M. A. P. **Análise da Variância Multivariada para os Cruzamentos Dialéticos**. Ciências Agortécnicas, Lavras, 2003. **Desenvolvimento do Acordo de Nível de Serviço para Ambientes de Redes de Computadores**, Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, 2002.
- LEEDY, P.D. *Practical research: planning and design*. New York: MacMillan, 1989.
- LEE, Matthew K. O. IT terceirização contracts: *practical issues for management*. **Industrial Management & Data Systems**, 1996.
- LEME, MURILO O. **Portal Corporativo de Gestão do Conhecimento em Projetos: Estudo de Caso em Empresa de Telecomunicações**. Programa de Pós-Graduação em

- Engenharia de Produção, Centro Federal de Educação Tecnológica do Paraná, Ponta Grossa-PR, 2005.
- LIRA, Larissa R.; SILVA, A. **Gerenciamento de Incidentes, segundo a ITIL** MBA Governança de Tecnologia da Informação, Cuiabá, MT, 2008.
- LYNCH, M. *Enterprise Information Portals. Merrill Lynch & Co* – In-depth Report, 1998. Disponível em: [HTTP://genesis.nce.ufrj/ged](http://genesis.nce.ufrj/ged). Acessado em 07/07/2009.
- MCBRIDE, D. “*Successful Deployment of IT Service Management in the Distributed Enterprise*”, <http://www2.hp.com/openview/rpm/pepers/slm.htm>, 1998.
- MACFARLANE, Ivor e RUDD, Colin *IT Service Management*. New Millenium Editora, São Paulo, 2005.
- MAGALHAES, Ivan Luizio. ; PINHEIRO, Walfrido Brito. **Gerenciamento de Serviços de TI na Prática: Uma abordagem com base na ITIL®**. São Paulo. Novatec, 2007.
- MALHOTRA, N.K. **Pesquisa de Marketing**. *Uma orientação aplicada*. Trad. Nivaldo M. Jr. e Alfredo A. de Farias. 3. ed. – Porto Alegre: Bookman, 2001.
- MAURER, W; MATLUS, R; FREY, N. *A Guide to Successful SLA development and Management, Strategic Analysis Report*. GartnerGroup, 2000.
- MINGOTI, SUELI A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Ed. UFMG, 2005.
- MORITA, Hideyuki; SHIMIZU, Tamio; LAURINDO, Fernando J. B. **Modelos para estruturar e avaliar alternativas de decisão em tecnologia de informação**. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 19., 1999, Rio de Janeiro. **Anais...** Rio de Janeiro, 1999.
- MOUSSEAU V.; SLOWINSKI R. e ZIELNIEWICZ P.; *ELECTRE TRI is a Multiple Criteria Sorting Method*. In Newsletter, Series 3, 2000.
- NARCISO, M. G.; LORENA, L. A. N. **Uso de algoritmos genéticos em sistema de apoio à decisão para alocação de recursos no campo e na cidade**. *Revista Brasileira de Agroinformática*, v. 4, 2002.
- NETO, JOSÉ M. M. e MOITA, GRAZIELLA C. **Uma Introdução à Análise Exploratória de Dados Multivariados**, Departamento de Química, Universidade Federal do Piauí, PI, 1997.
- OGC *Office Of Government Commerce. ITIL for Service Delivery*, 1º Edição Stationery Office BO, 2001.
- OGC, Office Government of Commerce – OGC. *IT Infrastructure Library: Service delivery*. London, 2001.
- OGC, Office Government of Commerce – OGC. *IT Infrastructure Library: Service Support*. London, 2001.

- PACHECO, JULIANO A. **Modelos Estatísticos Espaciais no Planejamento da Prestação de Serviços**. Dissertação submetida à Universidade Federal de Santa Catarina. Programa E Pós-Graduação em Ciência da Computação, UFSC, Florianópolis-SC, 2005.
- PAIVA, A. T.; MEDEIROS, I. B.; PEDRO, M. R.; MACHADO, U. L. **Ataque Incidente: Uma resposta ao Desenvolvimento**. SERPRO, Brasília-DF, 2007.
- PHILIPPS L. *Prioritisation and resource allocation with multi-criteria decision analysis*, Newsletter, Series 3, 2003.
- PISANI, J. F. **Análise Estatística Multidimensional em Biologia**. Ciência e Cultura, 1969.
- PRADO, P. I. K. L. “**Uso da Linguagem R para Análise de Dados em Ecologia**”. Disponível em: <<https://sistemas.usp.br/fenixweb/fexDisciplina?sgldis=BIE5782>>, acessado em 07/05/2009.
- PRESSMAN, Roger S. **Engenharia de Software**. São Paulo: Makron Books, 1995.
- PRESSMAN, ROGER S.. "*Software Engineering a practioner's aproach*". *Third Edition*. Singapore: McGraw-Hill, 1992.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. *R Foundation for Statistical Computing*, Vienna, Austria. 2003. Disponível em: <http://www.R-project.org>. Acessado em 07/05/2009.
- R PROJECT. “*The R Project for Statistical Computing*”, <http://www.r-project.org/>, 2009.
- REYNOLDS, HADLEY ; KOULOPOULOS, TOM (2000) *Enterprise knowledge has a face*. *Intelligent Enterprise*, v. 2, n. 5, p. 29-34, Disponível em: [www.intelligententerprise.com/993003/feat1.shtml](http://www.intelligententerprise.com/993003/feat1.shtml). Acessado em 07/07/2009.
- RIBEIRO JUNIOR, P. J. “**O Sistema R e computação Estatística**”. Disponível em: <<http://leg.ufpr.br/~paulojus/embrapa/Rembrapa/estcompR.pdf>>, acessado em 25/04/2009.
- ROBBINS, S. P. **Administração: mudanças e perspectivas**. São Paulo: Saraiva, 2000.
- RODRIGUES, Carlos A. P. **Estudo da Adoção das Melhores Práticas em TI – ITIL e Integração com a Metodologia de Gestão e Avaliação de Desempenho BSC**. Dissertação de Mestrado, Engenharia de Produção, Universidade Federal Fluminense, Niterói – RJ, 2006.
- SARTORIO, SIMONE D. **Aplicações de técnicas de análise multivariada em experimentos agropecuários usando o software R**. Dissertação de Mestrado. Escola Superior de Agricultura “Luiz de Queiroz”, USP, 2008.
- SHIM, J. P.; WARKENTIN, M.; COURTNEY, J.; POWER, D. J.; SHARDA, R.; CARLSSON, C. *Past, present, and future of decision support technology*. *Decision Support System*, v. 33, 2002.
- SHIMAKURA, S. (2007) “**Introdução ao R, Version 1.6.1**”. Disponível em: <<http://www.leg.ufpr.br/Rtutorial/intro.html>>, acessado em 25/04/2009.



- SILVESTRE, ANTÓNIO LUÍS **Análise de Dados e Estatística Descritiva**. Editora: Escolar Editora, Portugal, 2007.
- SILVA JUNIOR, VILSON V. **“Perspectivas para uso da linguagem R”**. Publicado em 19 de agosto de 2005.
- SILVA, M. F **“Apresentando o R”**. <<http://marcosfs2006.googlepages.com/apresentandoor>>, acessado em 27/04/2009.
- SILVA, Ricardo Pereira; **Avaliação de Metodologias de análise e projeto orientadas a objetos voltadas ao desenvolvimento de aplicações, sob a ótica de sua utilização no desenvolvimento de frameworks orientados a objetos**, 1996.
- SIMON, Herbert A. **Comportamento Administrativo**. Rio de Janeiro: USAID, 1965.
- SOMMERVILLE, IAN; SAWYER, P. **Requirements Engineering: A Good Practice Guide**. New Jersey: Prentice Hall. Englewood Cliffs, 1992.
- SOUZA, Sidnei F.; MARTINS, Walter J.; FRANCISO, Antonio C.; SCANDELARI, L. **Ferramentas da Tecnologia da Informação (TI) que auxiliam no levantamento de competências nas empresas** In: 3º Encontro de Engenharia e tecnologia dos Campos gerais, Universidade Estadual de Ponta Grossa – PR, 2007.
- SOUZA, ADRIANO MENDONÇA. **Monitoração e ajuste de realinhamento em processos produtivos multivariados**. Florianópolis: UFSC. 2000, 166p. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, 2000.
- SPRAGUE, Jr.; WATSON, H. **Decision support systems: putting theory into practice**. USA: Prentice-Hall, 1989.
- STÁBILE, S. **A comunicação do profissional de informática em uma empresa fornecedora de sistemas integrados de gestão: um estudo de caso**. In ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, Porto Alegre, 2005.
- STATDLOBER, Juliano. **Help-Desk e SAC com Qualidade**. Brasport Editora, Rio de Janeiro, 2006.
- STURM, R.;MORRIS, W.; JANDER, M. **“Foundations of Service Level management”**. Estados Unidos: Sams, 2000.
- THIVES, Juarez Jonas. **Workflow - uma tecnologia para transformação do conhecimento nas organizações**. Florianópolis: Insular, 2000.
- TOLEDO, Aline M. **Portais Corporativos: Uma Ferramenta Estratégica de Apoio à Gestão do Conhecimento**, UFRJ, Rio de Janeiro, 2002.
- TORRES, N. A. **Competitividade Empresariais com a Tecnologia de Informação**. São Paulo: Makron Books, 1995.
- TRIOLA, M. F., **Introdução à Estatística**. 7º Ed., LTC, Rio de Janeiro, RJ, 1998.

- TRIVELLONI, C. A. P.; HOCHHEN, N. **Avaliação de Imóveis em Técnicas de Análise Multivariada**. In COBRAC – Congresso Brasileiro de Cadastro Técnico Multifinalitário. UFSC, Florianópolis, 1998.
- TUDE, Eduardo. **Service Level Agreement (SLA)**. Tutorial publicado no site *telec.com.br* em 07/07/2003. Disponível em <http://www.teleco.com.br/tutoriais/tutorials/defaul.asp>.
- URUBATAN, RODRIGO. **Desenvolvimento Fácil e Rápido de Aplicações WEB – Ruby on Rails**, Ed. Novatec, 2009.
- VAHIDOV, R. e NEUMANN, D. **Situated Decision Support for Managing Service Level Agreement Negotiations**. In: INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 41, Hawaii, 2008.
- VAN GREMBERGEN, W. e HAES, S. **de. Measuring and Improving IT Governace through the Balanced Scorecard**. *Information Systems Control Journal*, 2005.
- VINCKE, P. **Multicriteria Decision-Aid**. John Wiley & Sons, 1992
- ZOPOUNIDIS C.. **Multicriteria Decision Support for Financial Classification Problems: The FINCLAS system**, *Newsletter, Series 3*, 2001.
- YOURDON, EDWARD - “**Análise Estruturada Moderna**”. Editora Campus, 1992.