

UNIVERSIDADE PAULISTA - UNIP
PROGRAMA DE MESTRADO EM ENGENHARIA DE PRODUÇÃO

**CONSTRUÇÃO DE *BUSINESS INTELLIGENCE* PARA
GERAÇÃO DE VALOR A UMA SIDERÚRGICA:
APLICAÇÃO NA MANUTENÇÃO INDUSTRIAL**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Paulista, para obtenção do título de Mestre em Engenharia de Produção.

LEONARDO MATSUMOTA

**São Paulo
2016**

UNIVERSIDADE PAULISTA - UNIP

PROGRAMA DE MESTRADO EM ENGENHARIA DE PRODUÇÃO

**CONSTRUÇÃO DE *BUSINESS INTELLIGENCE* PARA
GERAÇÃO DE VALOR A UMA SIDERÚRGICA:
APLICAÇÃO NA MANUTENÇÃO INDUSTRIAL**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Paulista, para obtenção do título de Mestre em Engenharia de Produção.

Orientador: Prof. Dr. Oduvaldo Vendrametto

Área de Concentração: Gestão de Sistemas de Operação

Linha de Pesquisa: Redes de Empresa e Planejamento da Produção

Projeto de Pesquisa: Aplicação de Ferramentas e Métodos Quantitativos na Solução de Problemas em Engenharia de Produção e Logística

LEONARDO MATSUMOTA

São Paulo

2016

FICHA CATALOGRÁFICA

Matsumota, Leonardo.

Construção de *business intelligence* para geração de valor a uma siderúrgica: aplicação na manutenção industrial / Leonardo Matsumota. - 2016.

78 f. : il. color. + CD-ROM.

Dissertação de Mestrado Apresentada ao Programa de Pós-Graduação em Engenharia da Produção da Universidade Paulista, São Paulo, 2016.

Área de Concentração: Gestão de Sistemas de Operação – Redes de Empresa e Planejamento da Produção.

Orientador: Prof.º Dr. Oduvaldo Vendrametto.

Coorientador: Prof.º Dr. Rodrigo Franco Gonçalves.

1. *Business intelligence*. 2. Manutenção industrial. 3. *Data mining*.

I. Vendrametto, Oduvaldo (orientador). II. Gonçalves, Rodrigo Franco (coorientador). III. Título.

**CONSTRUÇÃO DE *BUSINESS INTELLIGENCE* PARA
GERAÇÃO DE VALOR A UMA SIDERÚRGICA:
APLICAÇÃO NA MANUTENÇÃO INDUSTRIAL**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Paulista, para obtenção do título de Mestre em Engenharia de Produção.

LEONARDO MATSUMOTA

Aprovado em:

BANCA EXAMINADORA

_____/_____/_____
Prof. Dr. Oduvaldo Vendrametto (orientador)
Universidade Paulista - UNIP

_____/_____/_____
Prof. Dr. Mario Mollo Neto
UNESP - Tupã

_____/_____/_____
Prof. Dr. Rodrigo Franco Gonçalves
Universidade Paulista – UNIP

DEDICATÓRIA

A minha família, pelo apoio, compreensão e parceria, que foram fundamentais ao meu desenvolvimento profissional e por terem me direcionado aos estudos desde sempre, valorizando a importância do conhecimento para o crescimento e as realizações.

AGRADECIMENTOS

Agradeço a Deus por esta conquista! Não há maior recompensa na vida que atingir qualquer meta por meio de nosso esforço.

Ao professor orientador Dr. Oduvaldo Vendrametto, por seus valiosos conselhos, conhecimentos e críticas que ajudaram na construção de um trabalho melhor.

Aos professores doutores: Mario Mollo Neto, Rodrigo Franco Gonçalves, José Benedito Sacomano, João Gilberto Mendes dos Reis, Márcia Terra e Irenilza de Alencar Nääs.

Aos companheiros de estudos durante o curso de mestrado, inspiradores de novas metas e desafios.

Aos meus amigos da Usiminas, que contribuíram para o desenvolvimento das técnicas e ferramentas utilizadas no trabalho.

Agradeço também a todos aqueles que tornaram a realização deste trabalho possível a todos que contribuíram, direta ou indiretamente, para o meu desenvolvimento acadêmico e profissional.

RESUMO

Esta dissertação apresenta a construção de um ambiente de *business intelligence* para auxiliar a gestão da área de manutenção na indústria siderúrgica e, dessa forma, reforçar a importância da utilização de indicadores de desempenho, processamento analítico e *data warehouse* no controle das atividades e ocorrências críticas capazes de comprometer o processo de produção da empresa. Para tanto, foi desenvolvido um sistema de monitoramento que verifica o *status* de funcionamento dos principais equipamentos (*hardware*) e aplicações (*software*) instalados nas fábricas, o *Customer Relationship Management* (CRM) como meio de solicitação/abertura de chamados a fim de gerenciar o relacionamento com os clientes das áreas e também a implantação do *business intelligence* que compreende a análise multidimensional destas informações para suporte às decisões gerenciais. Para validar os resultados e objetivos propostos, a pesquisa aplicada utilizou os dados oriundos de uma usina siderúrgica de grande porte e com longo tempo de atividade no país, relacionando a indisponibilidade operacional com seus respectivos motivos, a fim de elaborar planos de manutenção que evitem a quebra de equipamentos críticos para o processo de produção, dessa forma, identificando também problemas recorrentes e a eficiência dos atendimentos realizados, através da definição de indicadores de manutenção. A mineração de dados, com as abordagens de modelagem descritiva (regras de associação) e preditiva (classificação e regressão), possibilitou descobrir padrões e relacionamentos ocultos nas ocorrências e atendimentos realizados na usina.

Palavras-chave: *Business intelligence*; Manutenção industrial; *Data mining*.

ABSTRACT

This study aims to build a business intelligence environment to assist the management of the maintenance area in the steel industry. With this, reinforce the importance of using performance indicators, analytical processing and data warehousing in control of critical activities and occurrences that can compromise the company's production process. Thus, a monitoring system that checks the operating status of the main equipment (hardware) and applications (software) installed in the factories was developed, CRM (Customer Relationship Management) as a means of request to manage the customer relations areas, and also the implementation of business intelligence comprising the multidimensional analysis of information to support management decisions. To validate the results and objectives proposed in this study, the applied research used the data from a large steel mill and with long activity in the country, relating the operational stop with their reasons, in order to develop maintenance plans prevent breakage of critical equipment for the production process. Also identify recurring problems and efficiency of assistance provided through the definition of maintenance indicators. Data mining, with the approaches of descriptive modeling (association rules) and predictive (classification and regression), made it possible to discover patterns and relationships hidden in the occurrences and assistance provided at the plant.

Keywords: Business intelligence; Industrial Maintenance; Data mining.

LISTA DE SIGLAS

ABRAMAN	Associação Brasileira de Manutenção
BI	<i>Business Intelligence</i>
CLP	Controlador Lógico Programável
COM	<i>Component Object Model</i>
CRM	<i>Customer Relationship Management</i>
DBSCAN	<i>Density-Based Spatial Clustering of Applications with Noise</i>
DCOM	<i>Distributed Component Object Model</i>
DENCLUE	<i>Density-based clustering</i>
DW	<i>Data Warehouse</i>
EM	<i>Expectation Maximization</i>
ERP	<i>Enterprise Resources Manager</i>
ETL	<i>Extract Transform Load</i>
FK	<i>Foreign Key</i>
KDD	<i>Knowledge Discovery in Databases</i>
KPI	<i>Key Performance Indicators</i>
MDX	<i>MultiDimensional eXpressions</i>
MES	<i>Manufacturing Execution System</i>
MTBF	<i>Mean Time between Failures</i>
MTTR	<i>Mean Time to Repair</i>
MTW	<i>Mean Time Waiting</i>
OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>
OPC	<i>OLE for Process Control</i>
OPTICS	<i>Ordering Points to Identify the Clustering Structure</i>
PIMS	<i>Plant Information Management System</i>
PK	<i>Primary Key</i>
SAP	<i>Systeme, Anwendungen und Produkte in der Datenverarbeitung</i>
SCADA	<i>Supervisory Control and Data Acquisition</i>
SSAS	<i>SQL Server Analysis Services</i>
SSIS	<i>SQL Server Integration Services</i>
SSRS	<i>SQL Server Reporting Services</i>
SGBD	Sistema Gerenciador de Banco de Dados

STING	<i>Statistical Information Grid Approach to Spatial Data Mining</i>
TI	Tecnologia da Informação
XI	<i>Exchange Infrastructure</i>
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

LISTA DE FIGURAS

Figura 1 – Tendência das empresas na contratação de serviços.....	14
Figura 2 – Indicadores de disponibilidade operacional devido à manutenção	15
Figura 3 – As três gerações de evolução da manutenção	19
Figura 4 – Arquitetura genérica de um sistema de <i>Business Intelligence</i>	22
Figura 5 – Esquema Estrela (modelagem multidimensional).....	24
Figura 6 – Esquema Flocos de Neve (modelagem multidimensional).....	24
Figura 7 – Etapas do processo KDD	26
Figura 8 - Característica multidisciplinar do KDD	28
Figura 9 – Mineração de dados no processo de KDD	28
Figura 10 – Tarefas da mineração de dados	30
Figura 11 – Registros agrupados em três <i>clusters</i>	32
Figura 12 – Arquitetura cliente-servidor do padrão OPC.....	36
Figura 13 – Sistema de aquisição de paradas operacionais	37
Figura 14 – Tela de abertura de chamado no CRM.....	38
Figura 15 – Preenchimento do CRM pelo executante	38
Figura 16 – <i>Control Flow</i> do processo de ETL	39
Figura 17 – <i>Data Flow</i> do processo de ETL	40
Figura 18 – Transferência dos dados: OLTP e OLAP.....	40
Figura 19 – Modelo multidimensional para a manutenção industrial	41
Figura 20 – Cubo OLAP no <i>Analysis Services</i>	42
Figura 21 – Configuração dos KPIs de atendimento e parada operacional.....	42
Figura 22 – Configuração dos <i>calculated members</i> do projeto	43
Figura 23 – Cubo da duração de paradas operacionais por equipe	44
Figura 24 – Tela inicial do <i>Weka</i>	45
Figura 25 – Interface do <i>Explorer Weka</i>	45
Figura 26 – Configuração do ambiente do <i>Reporting Services</i>	47
Figura 27 – Configuração do Excel com o <i>Analysis Services</i>	47
Figura 28 – <i>Dashboard</i> de indicadores de atendimento dos serviços	48
Figura 29 – <i>Dashboard</i> de indicadores de parada por falha e satisfação do cliente.....	49
Figura 30 – Tempo médio de atendimento das ocorrências	49
Figura 31 – Relatório de atendimentos realizados em prazo maior do que três meses	50
Figura 32 – Análise de disponibilidade por fábrica.....	50
Figura 33 – Análise das paradas operacionais por linha de produção.....	51
Figura 34 – Análise das paradas operacionais pela natureza da falha.....	51
Figura 35 – Operação <i>Drill Down</i> nos dados de paradas	52
Figura 36 – Operação <i>Drill Up</i> nos dados de diagnóstico de falha.....	52
Figura 37 – Operação <i>Drill Across</i> relacionada com as tabelas fato.....	52
Figura 38 – Operação <i>Slice and Dice</i> relacionada com as tabelas fato	53
Figura 39 – Parâmetros e saída do modelo de regressão linear	54
Figura 40 – Gráficos com estatísticas dos atributos e instâncias.....	55
Figura 41 – Relatório de classificação - algoritmo J48	55

Figura 42 – Árvore de decisão (algoritmo J48).....	56
--	----

LISTA DE TABELAS

Tabela 1 – Matriz de confusão	34
Tabela 2 – Tarefas e técnicas de mineração de dados	35
Tabela 3 – Regras de associação (<i>Apriori</i>)	57

SUMÁRIO

RESUMO	7
ABSTRACT	8
LISTA DE SIGLAS	9
LISTA DE FIGURAS	11
LISTA DE TABELAS	13
1 CONSIDERAÇÕES INICIAIS	13
1.1 Introdução	13
1.2 Justificativa	14
1.3 Objetivos	15
1.3.1 Objetivo geral	15
1.3.2 Objetivos específicos	16
1.4 Procedimento metodológico	16
1.5 Estrutura do trabalho	17
2 REVISÃO DA LITERATURA	18
2.1 Manutenção industrial	18
2.2 Evolução da manutenção	19
2.3 Tipos de manutenção	19
2.4 <i>Business Intelligence</i>	20
2.5 ETL (<i>Extract, Transform and Load</i>)	25
2.6 Descoberta de conhecimento em banco de dados (KDD)	26
2.7 <i>Data Mining</i>	28
2.7.1 Tarefas da mineração de dados	30
3 MATERIAIS E MÉTODOS	36
3.1 Sistemas de aquisição	36
3.2 Transformação dos dados	39
3.3 Modelagem multidimensional	41
3.4 <i>Analysis services</i>	42
3.5 <i>Weka</i>	44
3.6 Apresentação da informação	46
4 RESULTADOS E DISCUSSÃO	48
4.1 Implementação do <i>Data Warehouse</i>	48
4.2 Mineração de dados na siderúrgica	53
4.2.1 Regressão linear	53
4.2.2 Classificação	54
4.2.3 Associação	56
5 CONCLUSÃO	58
5.1 Recomendações de trabalhos futuros	58
REFERÊNCIAS	60
ANEXO A: ARTIGO DA REVISTA ESPACIOS – Desenvolvimento de um sistema de monitoramento e gestão da manutenção industrial	66

1 CONSIDERAÇÕES INICIAIS

1.1 Introdução

A atual crise econômica vivida pelas indústrias brasileiras sugere que algumas medidas precisam ser adotadas para reduzir os custos de produção e melhorar a produtividade dos funcionários no âmbito organizacional. A inovação dos processos e a eficiência na manutenção, utilizando equipamentos e sistemas inteligentes são outras oportunidades para aprimorar e garantir a continuidade da produção industrial. O enfoque da gestão é de atuar evitando a quebra do equipamento e a perda de produção, ao contrário do paradigma de que a execução de um bom reparo significava sucesso na manutenção (PINTO e RIBEIRO, 2002).

O'Brien (2004) menciona estudos de casos realizados em empresas que investiram na informatização para aumento da competitividade e obtiveram redução dos custos proveniente da desburocratização dos processos, aumento da eficiência e eficácia. Entre eles, o caso da Lucent Technologies, que implantou o sistema *Manufacturing Execution System* (MES) e *Enterprise Resources Manager* (ERP), deixando os clientes mais participativos no projeto para evitar especificações errôneas, atrasos nas entregas entre outros.

A Associação Brasileira de Manutenção (Abraman) afirma a importância da manutenção nas empresas:

[...] precisamos seguir os mesmos requisitos exigidos da produção: desenvolvimento econômico, preservação ambiental e responsabilidade social. Diretamente ligada à redução de custos, a manutenção firmou-se como elemento essencial à garantia de qualidade e da sustentabilidade dos processos produtivos [...] (ABRAMAN, 2010).

Em 2013 a entidade divulgou Documento Nacional (Figura 1), no qual aponta que a contratação de serviços nas empresas não tem aumentado, reforçando a necessidade de efetividade e melhoria nos processos já existentes.

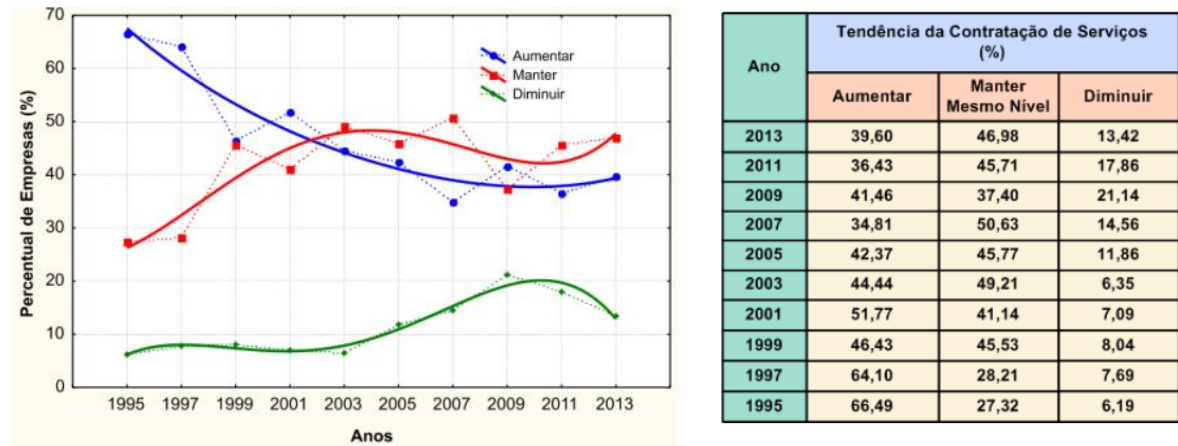


Figura 1 – Tendência das empresas na contratação de serviços

Fonte: Abramam (2013).

Dentro desse contexto, essa dissertação realiza um levantamento dos atendimentos e ocorrências dos sistemas de automação de uma indústria siderúrgica de grande porte localizada em Cubatão, com foco na prevenção das falhas na área de manutenção industrial que ocasionam paradas de produção e trazem riscos ambientais ou de acidente no trabalho. A redução do processo produtivo proveniente de uma parada inesperada em uma determinada linha de produção é um dos principais problemas enfrentados pelas indústrias. Fica evidente que consertar uma máquina após sua quebra não é viável, sendo mais sensato evitar que isto aconteça, fazendo as inspeções preventivas ou preditivas (KARDEC, 2001).

1.2 Justificativa

Devido à complexidade do ambiente siderúrgico, alguns departamentos como automação, elétrica, instrumentação e tecnologia da informação são tratados por diferentes gestores na estrutura organizacional da empresa. Por isso, possuem sistemas de fornecedores distintos, o que dificulta o acesso a seus equipamentos e gestão da informação, por não estarem integrados. Essas dificuldades são, em sua maioria, ocasionadas pela aquisição de soluções completas que englobam desde o equipamento (responsável pela produção de um produto) até o sistema de controle, que muitas vezes utiliza tecnologias diferentes dos padrões já existentes na empresa.

Assim, este estudo torna-se importante na medida em que irá determinar como realizar a integração entre os sistemas, extraindo e transformando os dados de cada um deles para criar uma base de dados única e em condições de análises confiáveis. A mineração dos dados

é capaz de evidenciar padrões e auxiliar na descoberta de conhecimento de muita valia aos gestores da empresa e o processamento analítico traz a visão de múltiplas perspectivas própria para análises comparativas.

Com estas informações a indústria tem condições de atuar na causa dos problemas mais importantes e direcionar seus atendimentos/investimentos com maior precisão, a fim de evitar a escassez dos recursos. Os equipamentos necessários para a operação das fábricas, por exemplo, têm custo alto e complexidade de atualização, o que torna inviável substituí-los com pouco tempo de vida útil. Assim, a empresa precisa trabalhar com planos rigorosos de manutenção para garantir o funcionamento adequado. A Figura 2 demonstra que a partir de 2009 ocorre a redução de disponibilidade e aumento da indisponibilidade devido à manutenção, evidenciando a necessidade de melhoria na gestão dos serviços.

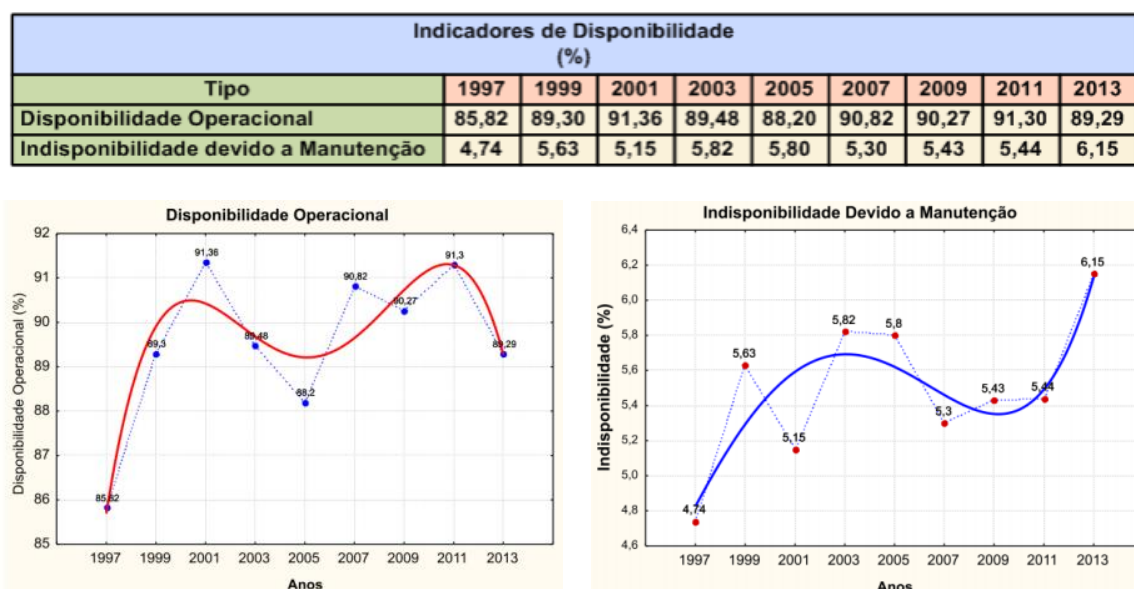


Figura 2 – Indicadores de disponibilidade operacional devido à manutenção

Fonte: Abraman (2013).

1.3 Objetivos

1.3.1 Objetivo geral

Desenvolver um sistema de *business intelligence* com integração e consolidação da base de dados das áreas de automação e tecnologia da informação para determinar indicadores de eficiência dos equipamentos essenciais para continuidade operacional e dos atendimentos realizados.

1.3.2 Objetivos específicos

1. Desenvolver um sistema de monitoração dos equipamentos de automação industrial, classificados como imprescindíveis para a continuidade de produção na empresa.
2. Gerar indicadores de desempenho de manutenção através do sistema de BI (*Business Intelligence*).
3. Aplicar técnicas de mineração de dados para descobrir padrões consistentes.
4. Integração entre o sistema desenvolvido com o ERP (*Enterprise Resources Manager*) da empresa.
5. Criar um sistema de gestão dos serviços de manutenção, que direcione ordens de serviço automaticamente, de acordo com a necessidade de atuação nos equipamentos.

1.4 Procedimento metodológico

A pesquisa aplicada foi realizada para identificar os critérios da modelagem multidimensional, implantação das técnicas de mineração de dados e apoio na tomada de decisão sobre a realização de serviços ou investimentos no setor de manutenção industrial.

O trabalho é uma aplicação prática dos conceitos relevantes apresentados em Kimball e Ross (2002), que possibilita a análise multidimensional com o intuito de organizar e realizar consultas analíticas *Online Analytical Processing* (OLAP) com efetividade, e Fayyad (1996), com o processo *Knowledge Discovery in Databases* (KDD) utilizado para auxiliar a descoberta de conhecimento em bases de dados.

Os estágios do processo de tomada de decisão estão embasados em Simon (1960 *apud* LAUDON e LAUDON, 2004) e os conceitos de manutenção industrial em Kardec (2001), Muassab (2002), Pinto e Ribeiro (2002) e nas informações da Associação Brasileira de Manutenção (Abraman).

A pesquisa quantitativa foi dividida em três fases: (i) pesquisa com base nas referências teóricas; (ii) análise para modelagem dos dados e critérios de elaboração dos indicadores de desempenho e (iii) apresentação dos resultados e respectivos comentários.

O sistema de aquisição dos dados foi desenvolvido em linguagem C# (*Dot Net*) com banco de dados MS *SQL Server*, utilizando uma rede TCP/IP (*Transmission Control Protocol - Internet Protocol*) para teste de conectividade dos equipamentos e obtenção dos resultados.

A apresentação de *status* de funcionamento de cada equipamento e serviço de notificação por *e-mail* utiliza o pacote da OSI SOFT (*ProcessBook* e *PI Notification*).

A arquitetura *SQL Server Integration Services* (SSIS) foi utilizada no processo de *Extract Transform Load* (ETL) para buscar os dados no ambiente *Online Transaction Processing* (OLTP) e carregar a base de dados do trabalho no *Online Analytical Processing* (OLAP). A integração deste sistema com o ERP da empresa, o *Systeme, Anwendungen und Produkte in der Datenverarbeitung* (SAP), ocorreu por meio do SAP XI (*Exchange Infrastructure*). O *software Weka* foi utilizado para aplicar técnicas e ferramentas de mineração de dados.

1.5 Estrutura do trabalho

O trabalho é composto por quatro capítulos. No Capítulo 1 apresenta-se as Considerações iniciais, com o tema, problema, objetivos, justificativa, procedimentos metodológicos e embasamento teórico.

No Capítulo 2, Revisão da literatura, estão elucidados os conceitos de manutenção industrial, a importância da prevenção das falhas em equipamentos e sistemas primordiais nas unidades de produção.

Os Materiais e Métodos são descritos no Capítulo 3, explicando as ferramentas e tecnologias utilizadas na aquisição, extração e transformação dos dados para preparação do *Data Warehouse*. A aplicação das técnicas de mineração de dados e a elaboração de relatórios para identificação das principais ocorrências na manutenção industrial, de acordo a literatura apresentada no Capítulo 2.

Os resultados e a integração entre sistemas são apresentados no Capítulo 4 e, em seguida, a Conclusão, as Referências e o Anexo A, com o artigo sobre o desenvolvimento de um sistema de monitoramento e gestão da manutenção industrial.

2 REVISÃO DA LITERATURA

Neste capítulo são abordados os tópicos centrais da dissertação com o objetivo de embasar teoricamente seus conceitos. Subdividem-se em manutenção industrial, *business intelligence* e mineração de dados.

2.1 Manutenção industrial

Dentre as diversas definições sobre o conceito de manutenção, Cabral (2006) trata a manutenção como o conjunto das ações destinadas a assegurar o bom funcionamento das máquinas e das instalações, garantindo que elas são intervencionadas nas oportunidades e com o alcance certo por forma a evitar que avariem ou baixem de rendimento.

Ainda que os autores tenham diferentes visões sobre a manutenção, o objetivo apresentado é transversal. COOKE (2000) e MADU (2000), em consenso, pontuam a manutenção dos equipamentos e a confiabilidade do sistema como fatores importantes que afetam a capacidade das organizações de fornecerem serviços de qualidade aos seus clientes. Além disso, a função manutenção é vital para o desempenho sustentável de qualquer empresa.

A norma EN 13306 (2001), com o propósito de uniformizar internacionalmente estas definições, apresenta que a manutenção é a “combinação de todas as ações técnicas, administrativas e de gestão durante o ciclo de vida de um bem, destinadas a mantê-lo ou repô-lo num estado em que possa desempenhar a função requerida”.

A manutenção industrial, como função estratégica na empresa, não consiste apenas em reparar ou evitar a quebra de equipamentos que comprometem a linha de produção, e sim, contribuir na melhoria da produtividade, segurança, redução de custos, adequação a normas e ao meio ambiente. Logo, melhorias na confiabilidade e disponibilidade dos equipamentos produtivos podem contribuir significativamente para aumentar a competitividade das organizações (SANTOS et al., 2007; MÁRQUEZ et al., 2009; PINTO e XAVIER, 2001; BLOOM, 2006; RAUSAND, 1998).

Neste contexto, muitas empresas já perceberam a importância de uma gestão efetiva da manutenção, principalmente dos recursos de influência direta na disponibilidade produtiva (NGUYEN et al., 2008). A eficácia dos custos empregados e a acurácia das atividades são os dois critérios básicos para uma boa manutenção (NIU et al., 2010).

2.2 Evolução da manutenção

Segundo Pinto (1999), a manutenção e os conceitos associados passaram por três grandes períodos:

- Primeiro período: surge o conceito de manutenção e “reparar a avaria” era o grande trabalho. Compreendeu o período após a Primeira Guerra Mundial até a década de 30.
- Segundo período: surge o conceito de manutenção preventiva. A produção em massa não parava de aumentar e as empresas passaram a “evitar a avaria” devido ao custo inerente às reparações.
- Terceiro período: a partir dos anos 70 do século passado, a manutenção está mais orientada para o controle das avarias (“adivinhar a avaria”). O conceito de manutenção preditiva é reforçado nesta época, com completo domínio das metodologias preventivas.

Kardec e Nascif (2009) definem a manutenção de modo semelhante, dividindo-a em três gerações: 1ª geração - manutenção não planejada, 2ª geração - manutenção preventiva e 3ª geração - manutenção preditiva (Figura 3).

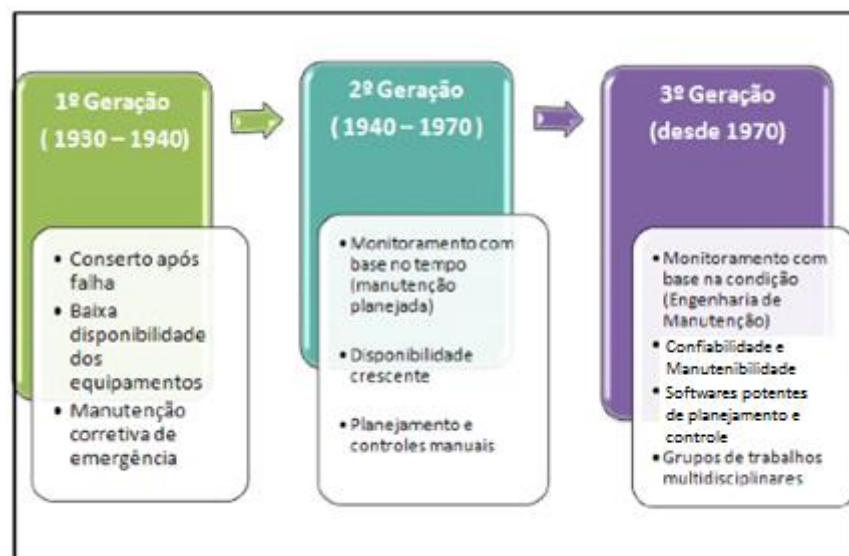


Figura 3 – As três gerações de evolução da manutenção

Fonte: adaptado de Kardec e Nascif (2009).

2.3 Tipos de manutenção

Existem vários tipos de manutenção dos quais se destacam:

- Manutenção preventiva: segundo a NBR 5462 (ABNT, 1994), é "[...] efetuada em intervalos de tempo predeterminados, ou de acordo com critérios prescritos, destinada a reduzir a probabilidade de falha ou degradação do funcionamento de um item".
- Manutenção corretiva: é aquela efetuada após a ocorrência de uma pane destinada a recolocar um item em condições de executar uma função requerida (ABNT, 1994). Bonin (1988) identifica que exige um esforço técnico e administrativo bem mais intenso, sendo que mesmo assim este fato torna-as normalmente de baixa produtividade.
- Manutenção preditiva: com base na aplicação sistemática de técnicas de análise, utiliza-se de meios de supervisão centralizados ou de amostragem, para reduzir ao mínimo a manutenção preventiva e diminuir a manutenção corretiva (ABNT, 1994).
- Manutenção detectiva: segundo Fagundes Neto et. al. (2006), é a estratégia que visa identificar as causas de falha e anomalias, auxiliando a manutenção em seus planos, com o objetivo de atacar a origem do problema e não apenas o sintoma.

2.4 *Business Intelligence*

O conceito de *Business Intelligence* (BI) surgiu na década de 1980, cunhado pela *Gartner Group* e refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte à gestão de negócios (PRIMAK, 2008).

Pode ser entendido, em uma de suas vertentes, como diretamente relacionado ao apoio e subsídio dos processos de tomada de decisão baseado em dados trabalhados especificamente para a busca de vantagem competitiva (COOPER et al., 2000; BARBIERI, 2001).

Com a utilização de uma arquitetura de BI, muitas empresas estão obtendo vantagem competitiva (BARBIERI, 2001), permitindo maior fidelização dos clientes, visão de mercados promissores e melhora na qualidade dos produtos ou serviços. Tudo isso devido à influência direta nos processos estratégicos da organização.

De um ponto de vista mais técnico, segundo Amaral (2011), o BI é uma habilidade que as corporações possuem ao explorar dados e informações que estão contidos em um *Data Warehouse* (DW) / *Data Marts* (DM) e através de ferramentas específicas, podem ser

analisadas e desenvolvidas percepções e entendimentos, permitindo obter informações fundamentadas e mais confiáveis para a tomada de decisão.

Inmon (2005) define que o DW é uma coleção de dados orientados por assuntos, integrados, variáveis com o tempo e não voláteis, para dar suporte ao processo de tomada de decisão. Para Kimball (2002), um conjunto de todos os *Data Marts* da organização, em que a informação é sempre armazenada em um modelo dimensional.

Um *Data Mart* normalmente se concentra em um assunto ou departamento específico, sendo considerado um subconjunto do DW. As principais características do DW, conforme Inmon (2005), são:

- Orientado por assuntos: a modelagem é desenvolvida especificamente sobre os principais assuntos e relativas às áreas estratégicas de uma empresa;
- Integrado: é possível padronizar uma representação única para os dados de todos os sistemas que formarão a base de dados do DW;
- Não voláteis: as informações ao serem carregadas na base analítica do DW se tornam uma carga não volátil, ou seja, a partir daquele momento as informações não serão alteradas, somente consultadas.
- Variante no tempo: um sistema de apoio à decisão deverá possuir um histórico de operações ocorridas de imediato, curto e longo prazo.

Para Kimball (2002), existem quatro componentes separados e distintos a serem considerados em um ambiente de DW – sistemas de suporte ao operacional, uma área de estágio de dados, uma área de representação de dados e uma ferramenta de acesso.

A Figura 4 apresenta uma arquitetura genérica do BI com seus principais componentes:

- Fonte de dados: provenientes de diferentes sistemas dentro da empresa;
- Processo de *Extract, Transform, Loading* (ETL): processo para extrair, transformar e carregar os dados no DW. Utiliza-se para isso uma área intermediária denominada área de estágio.
- *Data Warehouse*: consiste em um banco de dados central e *Data Marts* com dados modelados, de forma a atender as necessidades analíticas de uma determinada área de negócio.
- Apresentação: são as ferramentas para trabalhar os dados do DW e apresentá-los ao usuário final.

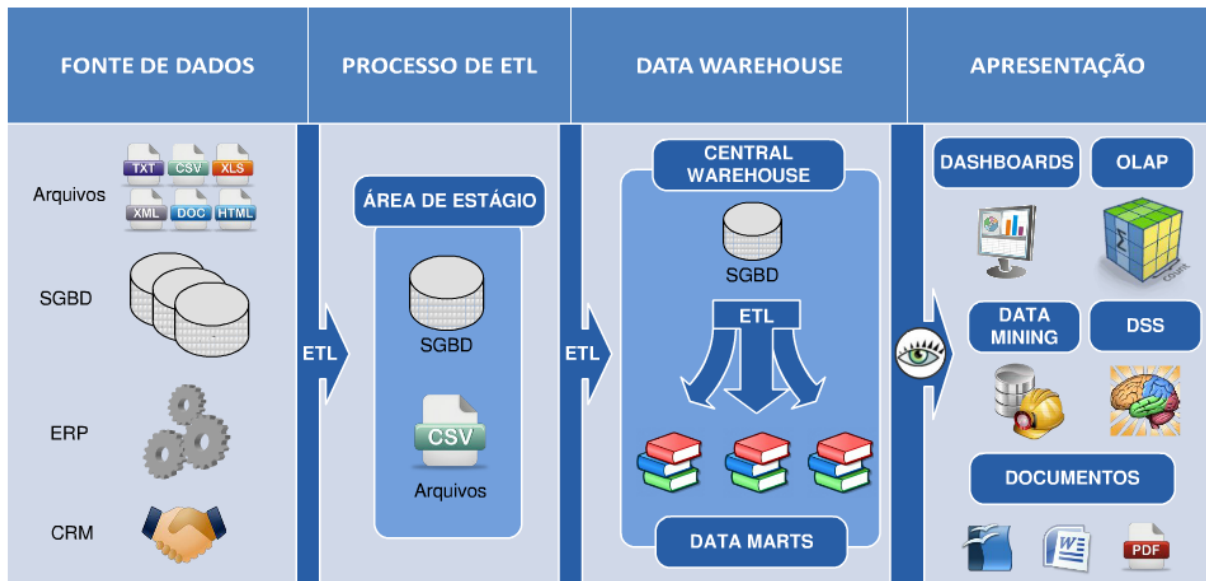


Figura 4 – Arquitetura genérica de um sistema de *Business Intelligence*

Fonte: Lorena (2011).

Conforme Kimball (2002), a modelagem dimensional é uma técnica de projeto lógico que busca apresentar os dados em uma estrutura padronizada mais intuitiva e que permite alto desempenho de acesso, evitando a utilização da técnica de normalização, conhecida em bancos de dados relacionais.

Imhoff et. al. (2003) definem o modelo dimensional como “[...] uma forma de modelagem de dados que acondiciona os dados de acordo com consultas e progressos específicos do negócio [...]”.

Nardi (2007) explica que as análises em bases multidimensionais envolvem o uso de alguns termos como dimensões, medidas, fatos e agregações:

- **Dimensões:** estabelecem a organização dos dados, determinando possíveis consultas/cruzamentos. Por exemplo: região, tempo, canal de venda etc.. Cada dimensão pode ainda ter seus elementos, chamados membros, organizados em diferentes níveis hierárquicos. A dimensão tempo, por exemplo, pode possuir duas hierarquias: calendário gregoriano (com os níveis ano, mês e dia) e calendário fiscal (com os níveis ano, semana e dia);
- **Medidas:** são os valores a serem analisados, como médias, totais e quantidades;
- **Fatos:** são os dados a serem agrupados, contendo os valores de cada medida para cada combinação das dimensões existentes. O tamanho da tabela que contém os fatos merece atenção especial do analista;
- **Agregações:** totalizações calculadas nos diversos níveis hierárquicos.

O modelo dimensional é composto da tabela fato e suas dimensões (tabelas dimensionais). Segundo Gubiani et. al. (2008), a tabela fato armazena a essência dos dados de um processo de negócio, medição numérica do negócio obtida da intersecção de todas as dimensões, continuamente valorado, aditivo (incremental) e as tabelas dimensões armazenam dados descritivos do negócio.

Essa estrutura provê condições para a realização de operações OLAP, que conforme Jacobson e Misner (2007), são um conjunto de ferramentas e técnicas que permite explorar os dados de um DW e visualização de grandes conjuntos de dados, transformando-os em informação.

Kimball (1998) considera as três principais abordagens OLAP:

- **MOLAP** (*Multidimensional Online Analytical Processing*): OLAP multidimensional. Constitui-se basicamente de um banco de dados multidimensional, através de um conjunto de interfaces de usuário, aplicações e banco de dados, com tecnologia proprietária.
- **ROLAP** (*Relational Online Analytical Processing*): OLAP relacional. Constitui-se de um conjunto de interfaces de usuário e aplicações que dá ao banco de dados relacional características dimensionais.
- **HOLAP** (*Hybrid Online Analytical Processing*): OLAP híbrido. É a combinação entre ROLAP e MOLAP, na qual os dados ficam retidos no banco de dados, enquanto as agregações ficam no MOLAP.

O cubo é a estrutura que armazena os dados em formato multidimensional, e as operações mais comuns que podem ser executadas são (HAN e KAMBER, 2001):

- *Slice-dice*: através do *slice*, o valor do atributo de uma dimensão é selecionado e fixado para análise em relação às outras dimensões. O *dice* permite definir um subcubo selecionando atributos em uma ou mais dimensões.
- *Drill Down*: permite navegar através do cubo a partir de dados agregados para dados mais detalhados, descendo um nível na hierarquia de uma dada dimensão, como por exemplo, na dimensão tempo é possível mudar do nível trimestre para ano.
- *Drill Up / Roll Up*: consiste numa agregação de dados do cubo que pode ser obtida subindo para um nível mais elevado da hierarquia de uma dada dimensão, como na dimensão localização, na qual é possível subir do nível país para o nível região.

- *Pivot/Rotate*: permite rodar os eixos dos dados de forma a disponibilizar uma representação alternativa dos mesmos.

Segundo Kimball (2001) e Inmon (2005), existem dois esquemas lógicos utilizados na modelagem multidimensional:

- Esquema Estrela: criado por Kimball (2001), que propõe a presença de dados altamente redundantes. A tabela “fato” fica centralizada com várias dimensões que não têm outro relacionamento nas suas pontas (Figura 5).

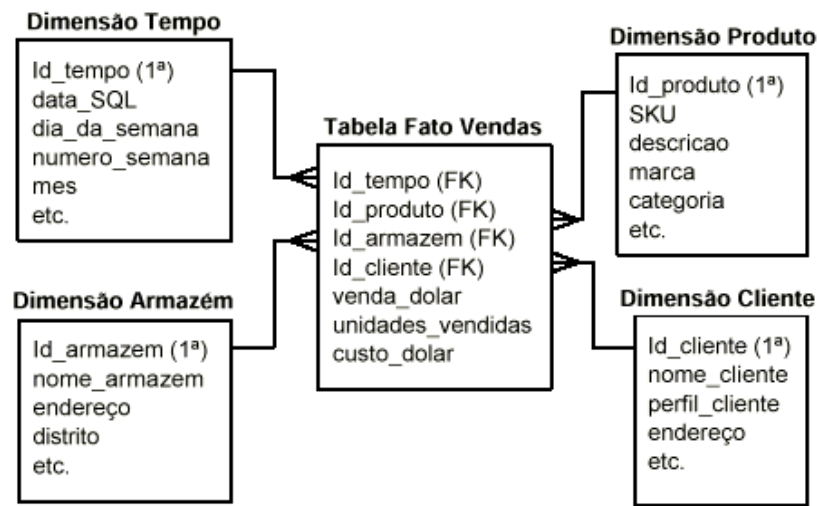


Figura 5 – Esquema Estrela (modelagem multidimensional)

Fonte: Inmon (2005).

- Floco de Neve: as dimensões relacionam-se com a tabela “fato” e com outras dimensões, conforme Figura 6. Este esquema objetiva a normalização das tabelas dimensionais para diminuir o espaço ocupado por elas (INMON, 2005).

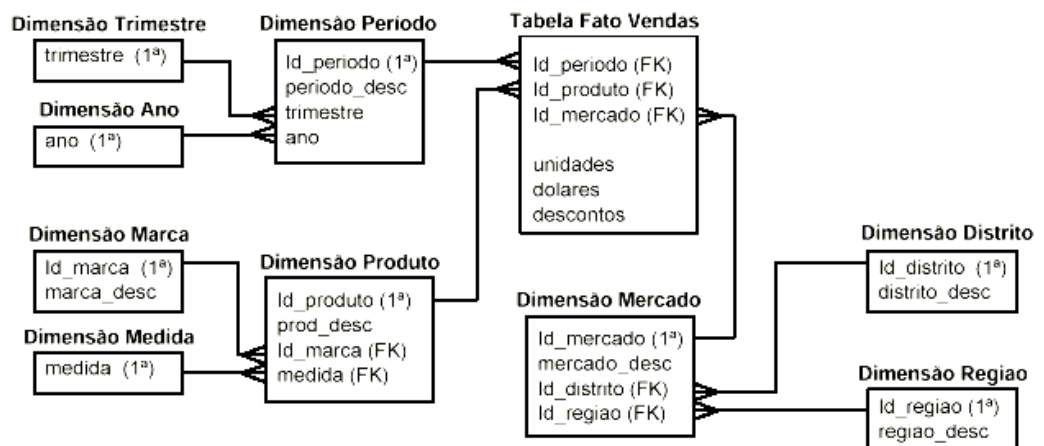


Figura 6 – Esquema Floco de Neve (modelagem multidimensional)

Fonte: Inmon (2005).

Machado (2010) explica que o formato de estrela origina-se por ter a tabela fato centralizada no modelo e as dimensões arranjadas ao redor dessa unidade central. Amaral (2011) acrescenta que as tabelas dimensões possuem chave estrangeira de relacionamento *Foreign Key* (FK) e assim, por meio dos atributos, se relacionam diretamente com a tabela fato, que contém a chave primária *Primary Key* (PK). A tabela fato contém os valores numéricos, ou medidas dos fatos consultados.

2.5 ETL (*Extract, Transform and Load*)

A etapa de *Extract, Transform and Load* (ETL) lida com questões de qualidade dos dados, tratando erros e inconsistências oriundos do ambiente de transações OLTP. A remoção de dados incompletos, repetidos, de tipagem e integração de dados heterogêneos são as principais adequações necessárias ao *Data Warehouse*. As três tarefas realizadas são:

- Extração: aquisição dos dados dos sistemas de origem. A maioria dos projetos consolida registros que existem em muitos formatos e localizações diferentes. Tal busca de dados pode ser obstruída por problemas como a distribuição das origens dos dados, que podem estar em bases distintas com plataformas diferentes gerando a demanda de utilização de formas de extração diferentes para cada local (ALMEIDA, 2006 *apud* ABREU, 2007).
- Transformação: aplica uma série de regras para determinar a padronização dos dados (homogeneização). O objetivo é garantir a integridade dos dados e sua transformação, que exige muita ou pouca manipulação, dependendo da fonte de dados. A correção de erros de digitação, a descoberta de violações de integridade, a substituição de caracteres desconhecidos, a padronização de abreviações podem ser exemplos de limpeza desta etapa (GONÇALVES, 2003 *apud* ABREU, 2007).
- Carga: os dados são efetivados no DW, que devido a sua utilização contínua, não pode ter um alto custo de processamento e tempo de carga. Para Almeida (2006), basicamente são carregadas as dimensões estáticas, de modificação lenta ou remanescente e fatos integrantes ao modelo do DW. Alguns cuidados como o desligamento de índices, referência de integridade e utilização de comandos que não geram armazenamento de informações para recuperação (exemplo: *truncate* ao invés de *delete*) podem ser adotados antes do início da carga de dados.

Os dados armazenados no DW precisam ser confiáveis, pois a partir destas informações são tomadas decisões que afetam os negócios da organização. Assim, o ETL é um processo considerado crítico e a maior parte do esforço exigido no desenvolvimento de um DW é consumida neste momento e não é incomum que 80% de todo esforço seja empregado no processo de ETL (INMON, 1997 *apud* ABREU, 2007).

2.6 Descoberta de conhecimento em banco de dados (KDD)

Com a evolução da tecnologia de armazenamento em banco de dados, torna-se imprescindível para as organizações, o desenvolvimento de ferramentas computacionais que auxiliem a analisar e interpretar as informações de forma automática, e assim, transformá-las em conhecimento que possa ser utilizado pela empresa.

Uma metodologia emergente, que tenta solucionar o problema da análise de grandes quantidades de dados e ultrapassa a habilidade e a capacidade humanas, é o *Knowledge Discovery in Databases* (KDD) (CARDOSO e MACHADO, 2008). Sem essa capacidade de interpretação correta dos dados, as organizações passam a analisar e tomar decisões seguindo suas próprias intuições, o que pode ocasionar erros graves na estratégia da empresa.

O KDD é o processo não trivial de identificar padrões de dados válidos, novos, potencialmente úteis e desconhecidos em grandes bancos de dados (FAYYAD et al., 1996). Este processo é composto de várias etapas que são executadas de forma interativa e iterativa. Brachnad e Anand (1996) definem as etapas como interativas, pois envolvem a cooperação do responsável pela análise de dados, cujo conhecimento sobre o domínio orientará a execução do processo. A iteração é devido ao processo não ser executado de forma sequencial, mas com repetidas seleções de parâmetros e conjuntos de dados, aplicações das técnicas de mineração de dados a fim de refinar os conhecimentos extraídos dos resultados. A Figura 7 demonstra as etapas do processo KDD.

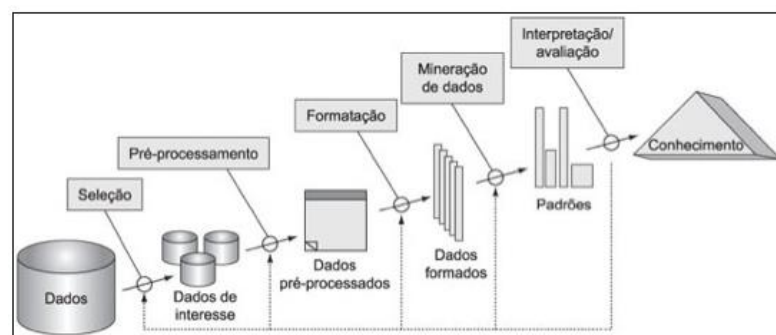


Figura 7 – Etapas do processo KDD

Fonte: Fayyad et al. (1996).

Damasceno (2010) explica as etapas do processo KDD da seguinte forma:

- Seleção: os dados são selecionados de acordo com as necessidades do usuário. Após o levantamento dos requisitos, cria-se um conjunto de dados com as informações necessárias aos algoritmos de mineração.
- Pré-processamento: é responsável pela qualidade e formatação dos dados e pela remoção de erros e exemplos fora do padrão essencial para a obtenção de resultados confiáveis em ferramentas de mineração.
- Transformação (formatação): os dados pré-processados passam por outra transformação, que localiza características úteis, melhores exemplos e atributos presentes no conjunto de dados, visando facilitar o uso das técnicas de mineração.
- Mineração de dados: aplicação de técnicas de mineração para alcançar os objetivos definidos na primeira etapa, o que envolve ajuste de modelos e determinação de características nos dados. As técnicas são escolhidas de acordo com os requisitos dos usuários e as características dos dados.
- Interpretação/Avaliação: é a etapa final do processo de mineração, que inclui visualizar e interpretar os padrões descobertos, utilizando medidas técnicas (precisão, erro médio, erro quadrático e taxas de falsos positivos e falsos negativos) e subjetivas (utilidade, entendimento ou complexidade) para avaliar os padrões extraídos.

A criação de um *Data Warehouse* é considerada como um dos primeiros passos para viabilizar a análise em grandes massas de dados (REZENDE, 2005). Tarapanoff (2001) acrescenta que o *Data Warehouse* é um aliado ao tornar a mineração de dados mais eficiente, pois além da análise de transações individualizadas é possível tirar proveito da agregação e sumarização de coleções de dados-alvo.

Lemos (2003) define o processo KDD como interdisciplinar, pois envolve diversas áreas do conhecimento que completam o processo de transformação dos dados: o aprendizado de máquina, a estatística, banco de dados, sistemas especialistas e visualização de dados, conforme Figura 8.



Figura 8 - Característica multidisciplinar do KDD

Fonte: adaptado de Lemos (2003).

2.7 Data Mining

Segundo Fayyad et al. (1996), o KDD é um processo não trivial de identificação de padrões, que tem por objetivo a descoberta de conhecimento em banco de dados, a fim de extrair informações implícitas e potencialmente úteis. Para Tan et al. (2009), este processo consiste de uma série de passos de transformação, do pré-processamento dos dados até o pós-processamento dos resultados da mineração, conforme mostrado na Figura 9.

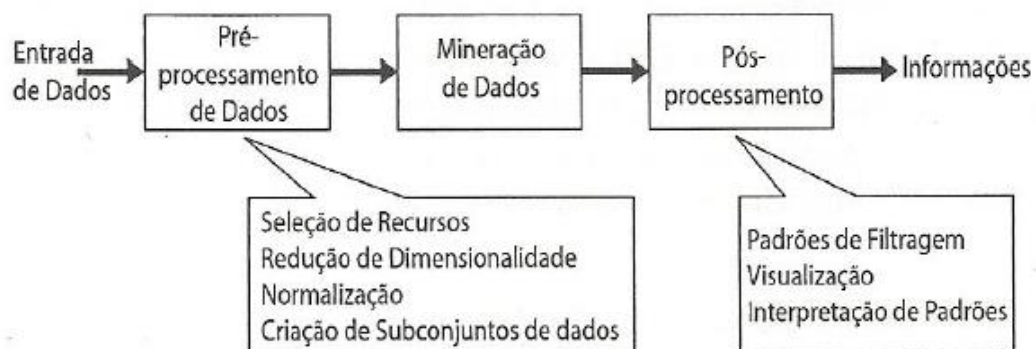


Figura 9 – Mineração de dados no processo de KDD

Fonte: Tan et al. (2009).

Dunkel et al. (1997) afirmam que a identificação de dados inapropriados dentro do conjunto selecionado é problemática e dificulta a automatização desta fase. Definir um dado como ruim dentro do conjunto depende da estrutura do mesmo e também de que aplicação é dada a ele. Tan et al. (2009) explicam que ao integrar os resultados da mineração de dados

com os sistemas de apoio a decisões, um passo de pós-processamento é requerido para assegurar que apenas resultados válidos e úteis sejam incorporados a estes sistemas.

Fayyad et al. (1996) reforçam que a mineração de dados é uma tecnologia poderosa na descoberta de informações ocultas nos bancos de dados, mas que não elimina a necessidade de conhecimento do negócio e o entendimento dos dados. Também é preciso conhecer muito bem o algoritmo e a ferramenta escolhida na busca dos dados.

A aplicação do processo de KDD à indústria de processos é apenas uma das diversas áreas em que podem ser aplicadas técnicas de mineração de dados. Wang (1999) apresenta várias técnicas de mineração de dados e uma série de considerações sobre o processo de KDD nas indústrias de processos:

- Análise estatística multivariada: utiliza técnicas como o *Principal Component Analysis* (PCA) para a identificação de variáveis mais importantes nos processos industriais;
- Aprendizado supervisionado: utiliza redes neurais para definição de modelos de identificação de falhas em processos. As classes de dados usadas para treino são sempre conhecidas e há um histórico de dados que permite prever dados futuros. Um conjunto de exemplos $E = \{E_1, E_2, \dots, E_N\}$ é fornecido ao sistema de aprendizado, no qual cada exemplo $E_i \in E$ possui um rótulo associado, que define a classe a qual o exemplo pertence. Pode-se dizer que cada exemplo $E_i \in E$ corresponde a uma tupla $E_i = (\vec{x}_i, y_i)$, na qual \vec{x}_i é um vetor de valores que representam as características (atributos) do exemplo E_i e y_i o valor da classe deste exemplo. Como o propósito do aprendizado supervisionado é induzir um mapeamento geral dos vetores \vec{x} para valores y , o sistema de aprendizado deve construir um modelo, tal que $y = f(\vec{x})$, onde f é uma função desconhecida (função conceito) que permite predizer valores y .
- Aprendizado não supervisionado: para identificar comportamentos semelhantes entre variáveis com a implementação de um classificador Bayesiano. O rótulo da classe das amostras de treino não é conhecido e, a princípio, o número de classes a ser treinada também pode ser desconhecido. O sistema de aprendizado recebe um conjunto de exemplos E , e cada exemplo consiste apenas de vetores \vec{x} , não contendo a informação da classe y . O objetivo é construir um modelo que procura por regularidades nos exemplos, formando agrupamentos ou *clusters* de exemplos com características similares. Utiliza algoritmos descritivos e assim

trabalham com conjuntos de dados que não possuem uma classe determinada. O objetivo é identificar padrões de comportamento semelhantes nestes dados. As tarefas descritivas podem ser: associação, agrupamento e generalização.

Mesmo sendo um dos principais meios de criar conhecimento e prever eventos futuros, a inferência indutiva deve ser utilizada com muita atenção e cuidado, pois segundo Batista (2003), um argumento indutivo e correto pode, perfeitamente, admitir uma conclusão falsa, ainda que suas premissas sejam verdadeiras. Se as premissas de um argumento indutivo são verdadeiras, o melhor que pode ser dito é que a sua conclusão é provavelmente verdadeira.

2.7.1 Tarefas da mineração de dados

Tan et al. (2009) afirmam que as tarefas de mineração de dados são geralmente divididas em duas categorias principais:

- Tarefas preditivas: o objetivo é prever o valor de um determinado atributo (“variável dependente” ou “alvo”) baseado nos valores de outros atributos (“variáveis independentes” ou “explicativas”).
- Tarefas descritivas: o objetivo é derivar padrões de comportamento comum nos dados. A validação e explicação do resultado frequentemente requerem técnicas de pós-processamento, pois muitas vezes são de natureza exploratória.

A Figura 10 ilustra as tarefas de mineração organizadas em atividades preditivas e descritivas:

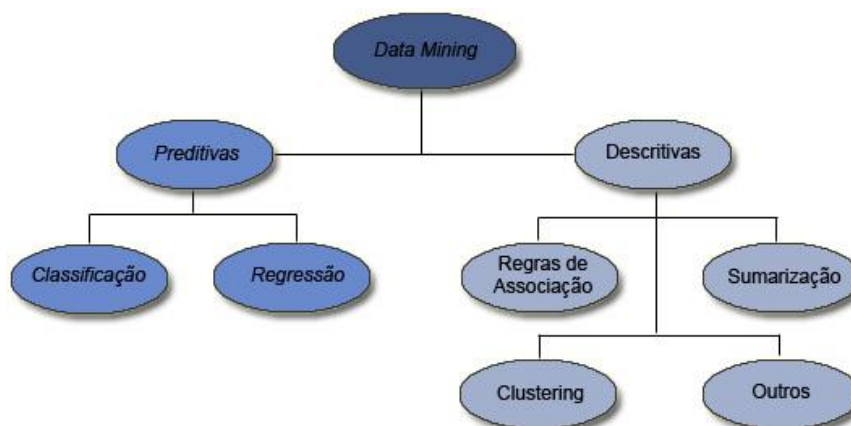


Figura 10 – Tarefas da mineração de dados

Fonte: Rezende (2003).

2.7.1.1 Associação

A associação é uma técnica para descobrir relacionamentos interessantes em um determinado conjunto de dados. O uso deste método foi empregado por Silva et al. (2009) para descobrir fraudes no uso dos cartões de pagamento do Governo Federal.

As regras podem mostrar relacionamento forte existente entre itens comprados juntos, tipos de defeitos, entre outros. Uma regra de associação é um padrão da forma $X \rightarrow Y$ (X implica em Y), em que X e Y são conjuntos disjuntos de valores ($X \cap Y \neq \emptyset$). A força da regra pode ser medida em (TAN et al., 2009):

- Suporte: determina a frequência da aplicação de uma regra a um determinado conjunto de dados.

$$\text{Suporte } (X \rightarrow Y) = \frac{\text{Transações que contém X e Y}}{\text{Quantidade total de transações}}$$

- Confiança: determina a frequência que os itens Y aparecem nas transações que contenham X .

$$\text{Confiança } (X \rightarrow Y) = \frac{\text{Transações que contém X e Y}}{\text{Transações que contém X}}$$

Os algoritmos *Apriori* e *FP-growth* são muito utilizados para descobrir as regras de associação. De acordo com Agrawal e Srikant (1994), o *Apriori* utiliza a estratégia de itens frequentes, ou seja, se um *itemset* (itens que ocorrem em conjunto) é frequente, então todos os seus subconjuntos também são frequentes. Este princípio é conhecido como a propriedade antimonotônica, onde o suporte de um *itemset* nunca é maior que o suporte de seus subconjuntos.

O algoritmo *FP-growth* (HAN et al., 1999) realiza a busca por regras de associação sem a necessidade de geração de um conjunto de *itemsets* candidatos, o que faz dele um algoritmo menos custoso que o *Apriori*. Isso porque a estrutura de dados *FP-Tree* armazena informações de forma compacta sobre os *itemsets* frequentes, evitando a realização de varreduras constantes no banco de dados.

2.7.1.2 Agrupamento (*clustering*)

Segundo Tan et al. (2009), o agrupamento (*clustering*) é uma abordagem descritiva que agrupa objetos baseado apenas em informações encontradas nos dados que descrevem os

objetos e seus relacionamentos. O objetivo é que os objetos dentro de um grupo, ou *cluster*, sejam semelhantes (ou relacionados) entre si e diferentes de (ou não relacionados aos) outros objetos de outros grupos. Os elementos com pouca ou nenhuma semelhança (dissimilares) são distanciados e inseridos em outros *clusters*. A Figura 11 mostra a identificação de grupos de dados similares.

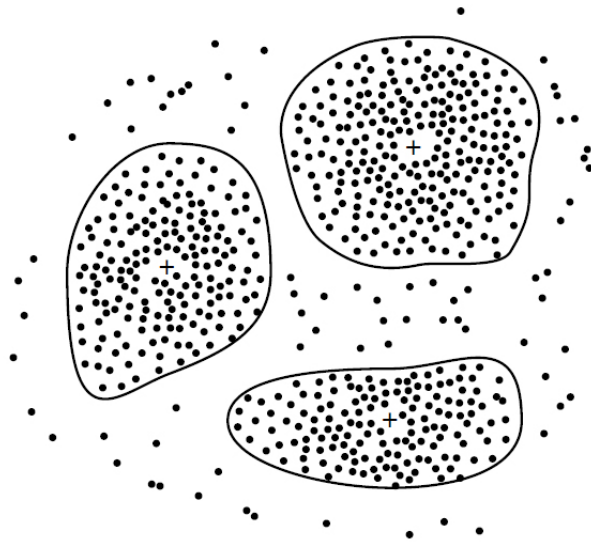


Figura 11 – Registros agrupados em três *clusters*

Fonte: Han e Kamber (2006).

Diferente da tarefa de classificação que tem rótulos predefinidos, a clusterização precisa automaticamente identificar os grupos de dados que o usuário deverá atribuir rótulos, como por exemplo, uma empresa de telecomunicações pode realizar um processo de clusterização para obter grupos de clientes que compartilhem o mesmo perfil de compra de serviços (GOLDSCHMIDT e PASSOS, 2005).

Os algoritmos de agrupamento podem ser classificados em:

- Métodos de particionamento: agrupa elementos em K grupos, onde K é a quantidade de grupos desejados. O *k-Means* e o *k-Medoids* são os algoritmos mais comuns de agrupamento.
- Métodos hierárquicos: agrupa por meio da divisão ou da aglomeração dos elementos do conjunto, nos quais os métodos aglomerativos e divisivos são utilizados.
- Métodos baseados na densidade: trabalham com a distribuição dos valores dos dados mais densa, onde cada ponto tem muitos pontos em sua vizinhança. Os algoritmos *DBSCAN*, *OPTICS* e *DENCLUE* são exemplos.

- Métodos baseados em grade: dividem os registros em um número finito de células de uma grade multidimensional, onde as operações de agrupamento são realizadas. O *STING* e *WaveCluster* são os principais algoritmos.
- Métodos baseados em modelos: criam um modelo para cada agrupamento e buscam identificar o melhor modelo para cada objeto. O *EM*, *COBWEB* e *CLASSIT* são os algoritmos utilizados.

2.7.1.3 Sumarização

Sumarização é, para Dunham (2002), um método descritivo que mapeia dados em subconjuntos com associações descritivas simples, utilizando técnicas básicas de estatística (média, moda, mediana, desvio-padrão e variância) para resumir os dados. Ou seja, procura identificar e indicar características comuns entre conjuntos de dados. É comum aplicar a tarefa de sumarização a cada um dos agrupamentos obtidos pela tarefa de clusterização (GOLDSCHMIDT e PASSOS, 2005).

Segundo Fayyad (1996), a tarefa de sumarização envolve métodos para encontrar uma descrição compacta para um subconjunto de dados. E como exemplo desta tarefa, a tabulação do significado e desvios-padrão para todos os itens de dados.

2.7.1.4 Classificação

De acordo com Romão (2002), a classificação é uma das tarefas mais referenciadas na literatura de MD. Seu objetivo é descobrir um relacionamento entre um atributo e um conjunto de atributos previsores. Paula (2004) diz que para tal, é necessária a descrição prévia das classes, expressando suas características por meio de definições, fórmulas e/ou atributos.

A construção do modelo utiliza dados de entrada com rótulos conhecidos, chamados de dados de treino. Os dados de teste utilizam os dados gerados por este modelo e rotulam os dados corretamente. Para avaliar o desempenho do modelo, é feita uma contagem destes dados de teste previstos corretamente e incorretamente. A Tabela 1 é uma matriz de confusão que faz a tabulação destas contagens:

Tabela 1 – Matriz de confusão

		Classe prevista	
		Classe = 1	Classe = 0
Classe real	Classe = 1	f_{11}	f_{10}
	Classe = 0	f_{01}	f_{00}

Fonte: Ohana (2012).

As entradas $f_{11} + f_{00}$ correspondem ao número de exemplos classificados com previsões corretas e as entradas $f_{10} + f_{01}$, às previsões incorretas. As métricas de precisão e erro utilizadas na avaliação de desempenho são definidas nas seguintes equações:

$$\text{Precisão} = \frac{\text{previsões corretas}}{\text{total de previsões}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

$$\text{Erro} = \frac{\text{previsões incorretas}}{\text{total de previsões}} = \frac{f_{01} + f_{10}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

Alguns exemplos do uso da tarefa de classificação são citados por Larose (2006):

- Sistemas bancários: identificar fraude em uma transação de cartão de crédito;
- Educação: encontrar necessidades especiais para novos alunos;
- Medicina: contribuição no diagnóstico de doenças específicas;
- Leis: determinar a veracidade de um testamento;
- Segurança: identificar ameaça terrorista por meio do comportamento pessoal.

2.7.1.5 Regressão

Para Fayyad et al. (1996), a regressão compreende a busca por uma função que mapeie um item de dado para uma variável de predição real estimada. A regressão é uma técnica de modelagem preditiva na qual a variável alvo a ser avaliada é contínua, e não discreta, como na tarefa de classificação. Seu objetivo é encontrar uma função-alvo que possa ajustar os dados de entrada com um erro mínimo (TAN et al., 2009).

No campo da análise estatística, os métodos de estimativa são largamente utilizados para determinar pontos estatísticos, intervalos de confiança, regressão linear simples, correlações e regressão múltipla (LAROSE, 2006). As técnicas matemáticas de regressão utilizadas são: regressão linear, não linear, logística, múltipla e de Poisson. A Tabela 2 apresenta um resumo das tarefas com as técnicas associadas:

Tabela 2 – Tarefas e técnicas de mineração de dados

Tarefas	Técnicas
Associação	<i>Apriori, FP-Growth, DHP, DIC, GSP, SPIRIT.</i>
Agrupamento (<i>Clustering</i>)	<i>K-means</i> , agrupamentos hierárquicos, DBSCAN, baseados em protótipos, baseados em densidade, baseados em grafos, agrupamentos escaláveis.
Detecção de anomalias	Técnicas estáticas baseadas em proximidade, em densidade e em agrupamento.
Classificação	Árvores de decisão, Bayesianos, rede neural artificial (RNA), máquinas de vetor de suporte (SVM), classificadores baseado em regras e de vizinho mais próximo.
Regressão	Regressão linear, regressão não linear, regressão logística, regressão múltipla e regressão de Poisson.

Fonte: Adaptado de Tan et al. (2009).

3 MATERIAIS E MÉTODOS

3.1 Sistemas de aquisição

No ambiente de produção da empresa, os sistemas de aquisição utilizam o padrão OPC, baseado nas tecnologias *Component Object Model* (OLE COM) e *Distributed Component Object Model* (DCOM). Segundo Fonseca (2002), eles estabelecem as regras para que sejam desenvolvidos sistemas com interfaces padrões para comunicação dos dispositivos de campo (CLPs, SDCDs etc.) com sistemas de monitoração, supervisão e gerenciamento (SCADA, MES, ERP etc.).

Puda (2008) explica que o padrão OPC trabalha com um ou mais servidores fornecendo dados para as aplicações-cliente, arquitetura conhecida como cliente-servidor (Figura 12). O servidor OPC é o responsável pela tradução do protocolo proprietário do aplicativo para o padrão OPC, que permite acesso aos dados nas especificações (SOUZA et al. 1998):

- Leitura e escrita síncrona: executadas imediatamente pelo servidor e só retornam para o cliente após completar a operação.
- Leitura e escrita assíncrona: mais eficientes, pois o cliente é imediatamente liberado após fazer a requisição.
- Atualização enviada pelo servidor: permite ao cliente requisitar ao servidor de forma periódica ou por exceções, atualizando um determinado conjunto de valores.



Figura 12 – Arquitetura cliente-servidor do padrão OPC

Fonte: Puda (2008).

O *OPC Data Access (DA)* é a funcionalidade mais utilizada do padrão OPC no desenvolvimento deste trabalho. Segundo a *OPC Foundation* (2006), o OPC DA fornece a funcionalidade de transferência de dados de tempo real e contínua de CLPs, SDCDs e outros, para IHMs, sistemas supervisórios e similares.

O sistema coletor de paradas (Figura 13) foi desenvolvido em ambiente *Dot Net* (linguagem C#) e utiliza o padrão OPC, explicado anteriormente, para fazer a conexão com os CLPs. Neste caso, o *driver* da *GE Ethernet* foi instalado no servidor e conectado no CLP GE 9030 para verificar o *status* de funcionamento da fábrica. O evento de mudança no sinal de funcionamento é indicado pelo evento “disparaParada_DataChange” que utiliza a referência *OPCAutomation* (OPCDAuto.dll) para trabalhar com os métodos, interfaces e propriedades de comunicação do padrão OPC (*OPC Foundation*).

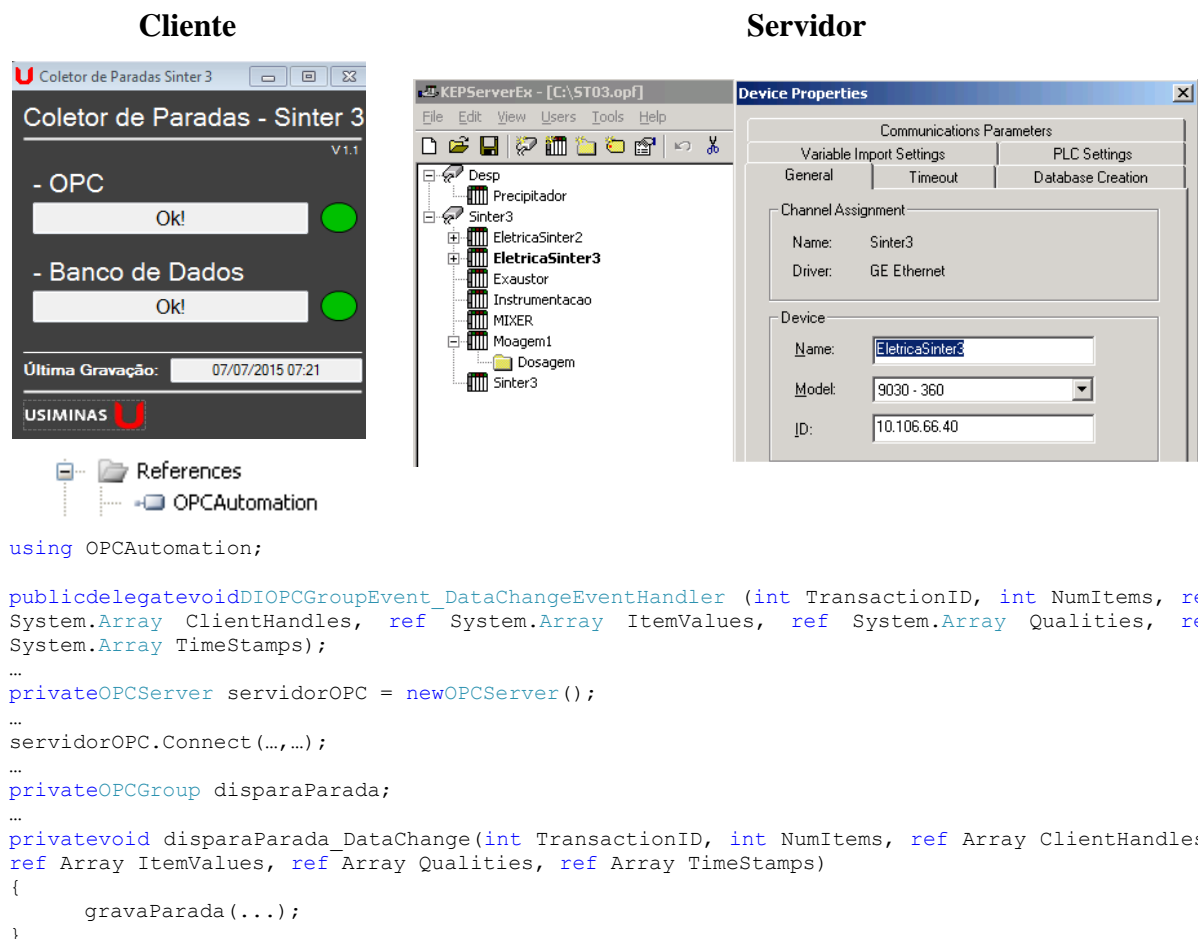


Figura 13 – Sistema de aquisição de paradas operacionais

O sistema CRM é utilizado na empresa para gerenciar as solicitações realizadas pelos clientes na abertura, acompanhamento e execução de serviços. A Figura 14 exhibe como o cliente realiza a abertura de uma solicitação preenchendo: o título do chamado, descrição,

responsável por solicitar (contraparte), responsável pelo chamado (proprietário), sistema e unidade de trabalho da linha de produção a ser realizada.

Figura 14 – Tela de abertura de chamado no CRM

O executante da solicitação preenche as informações de início e fim do atendimento, qual o motivo da ocorrência, se utilizou uma solução de contorno, reincidência da ocorrência e consequência operacional (em caso de indisponibilidade de produção), conforme Figura 15.

Figura 15 – Preenchimento do CRM pelo executante

Os sistemas de aquisição, explicados anteriormente, registram as informações em banco de dados relacional no ambiente OLTP e o processo ETL é o responsável pela extração, transformação e carga da informação no *Data Warehouse*, possibilitando o uso posterior no ambiente OLAP.

3.2 Transformação dos dados

A arquitetura *SQL Server Integration Services* (SSIS) permite o desenvolvimento dos pacotes (*packages*) em projetos de *Extract, Transform and Load* (ETL), sendo necessária a configuração do fluxo de controle (*Control Flow*) e seus respectivos itens: os *containers* e as tarefas (*Tasks*). O *Sequence Container* permitiu a organização das tarefas dos fluxos de dados (*data flow task*) responsáveis pela inserção dos dados no DW, conforme Figura 16.

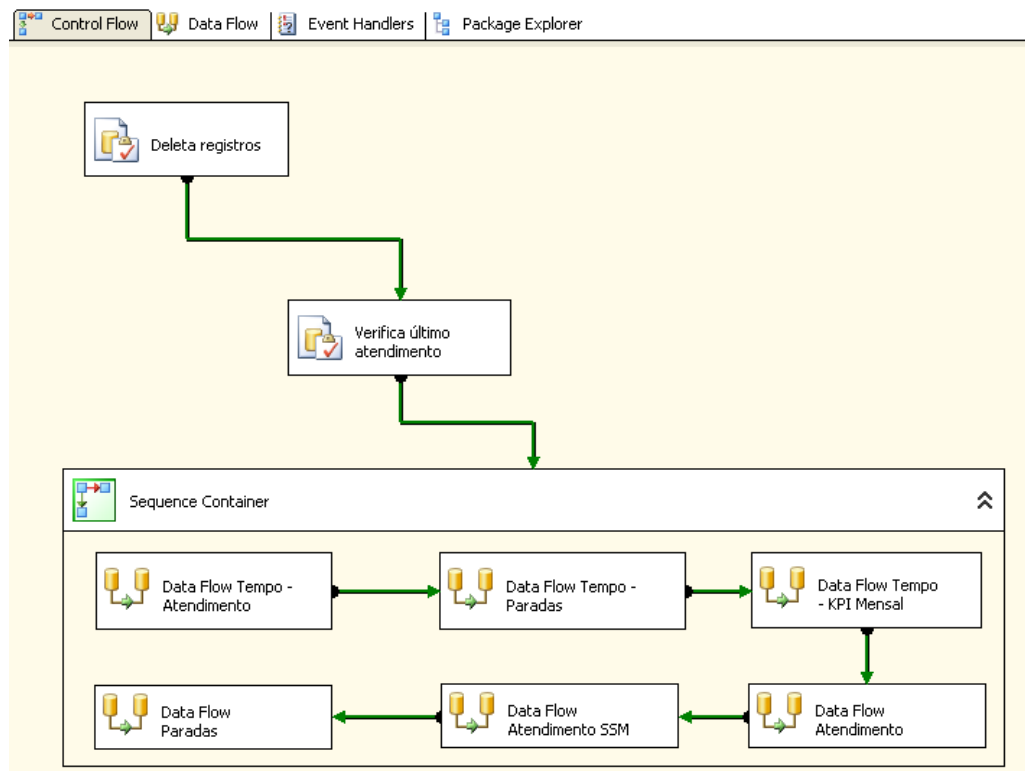


Figura 16 – *Control Flow* do processo de ETL

No processo de *Extract, Transform and Load* (ETL), a etapa de extração é a responsável por coletar os dados armazenados no sistema OLTP, que utilizam o banco de dados *SQL Server 2008*. A etapa seguinte trata da qualidade dos dados (limpeza e transformação) com o objetivo de padronizá-los.

Algumas colunas consideradas chaves primárias do sistema transacional CRM são do tipo *varchar* e foram convertidas para o formato *int*. Também precisaram de tratamento alguns valores nulos (por exemplo: atendimentos que não foram finalizados; pesquisa de satisfação não respondida; forma de atendimento não preenchida) para não prejudicar a projeção dos cálculos finais. A última etapa, considerada de carga dos dados no DW, tratou de executar a carga dos dados entre origem (OLTP) e destino (DW), conforme Figura 17.

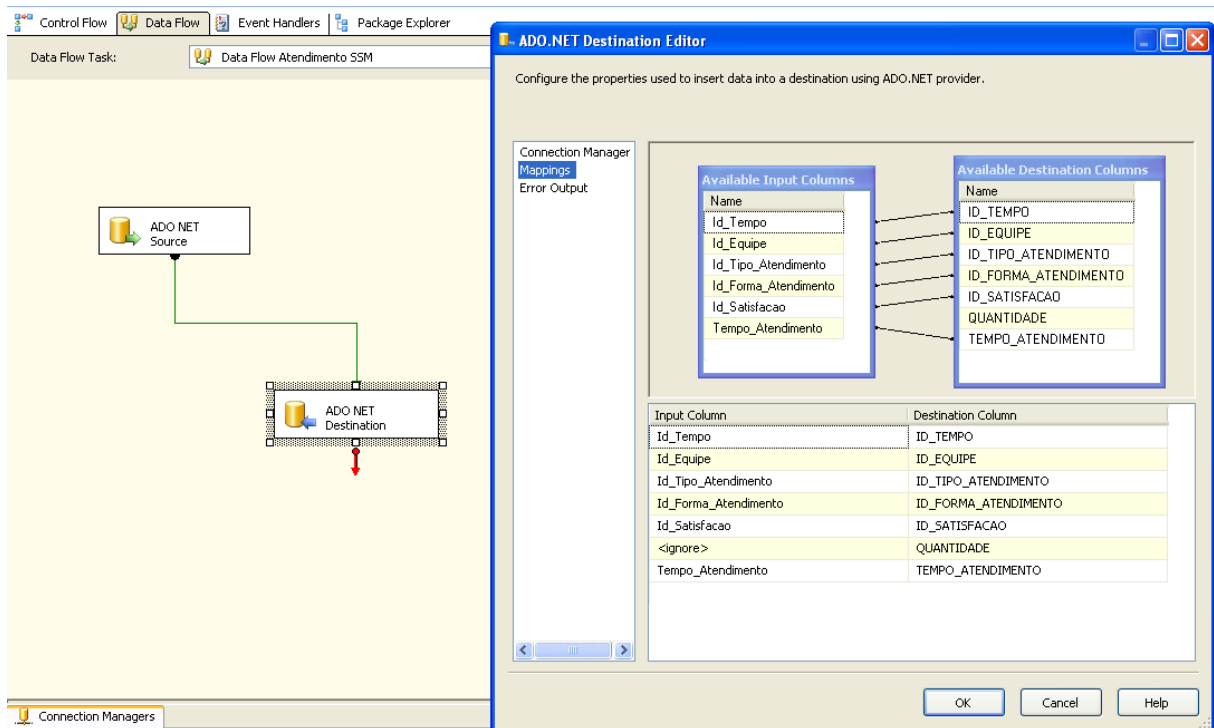


Figura 17 – Data Flow do processo de ETL

Assim, os dados armazenados no ambiente OLTP puderam ser extraídos, transformados e armazenados no DW, permitindo o processamento analítico OLAP das informações, muito importantes na construção dos cubos, conforme Figura 18.

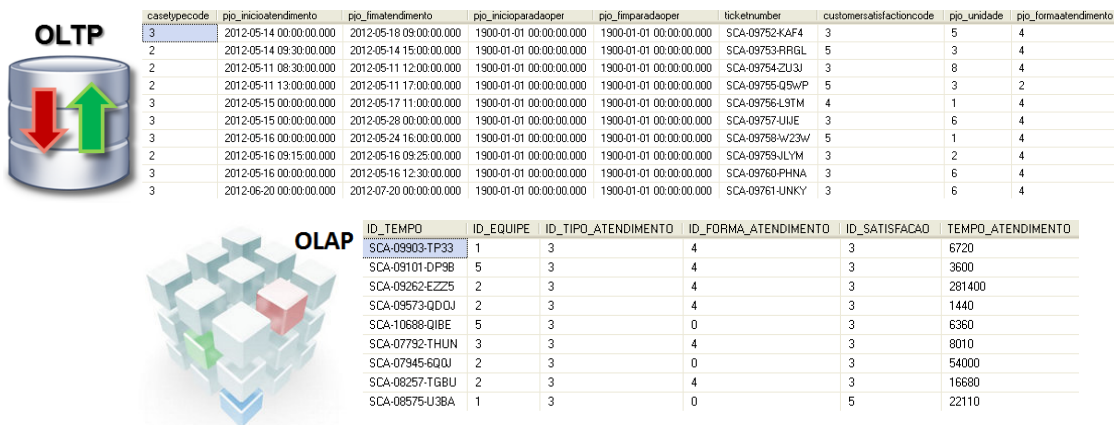


Figura 18 – Transferência dos dados: OLTP e OLAP

3.3 Modelagem multidimensional

A modelagem multidimensional do Data Warehouse, seguindo o esquema estrela, foi considerada neste trabalho. A tabela fato fica localizada no centro do modelo e suas tabelas de dimensões ficam arranjadas ao redor dessa unidade central, originando um formato semelhante ao de uma estrela (MACHADO, 2010). A Figura 19 demonstra o relacionamento das tabelas fato FATO_PARADA_OPERACIONAL, FATO_ATENDIMENTO_SSM, FATO_ATENDIMENTO e as tabelas de dimensão DIM_TIPO_ATENDIMENTO, DIM_SATISFACAO, DIM_EQUIPE, DIM_TEMPO, DIM_FORMA_ATENDIMENTO, DIM_DIAGNOSTICO, DIM_LINHA_PRODUCAO e DIM_FABRICA. O SGBD utilizado foi o SQL SERVER 2008 da Microsoft.

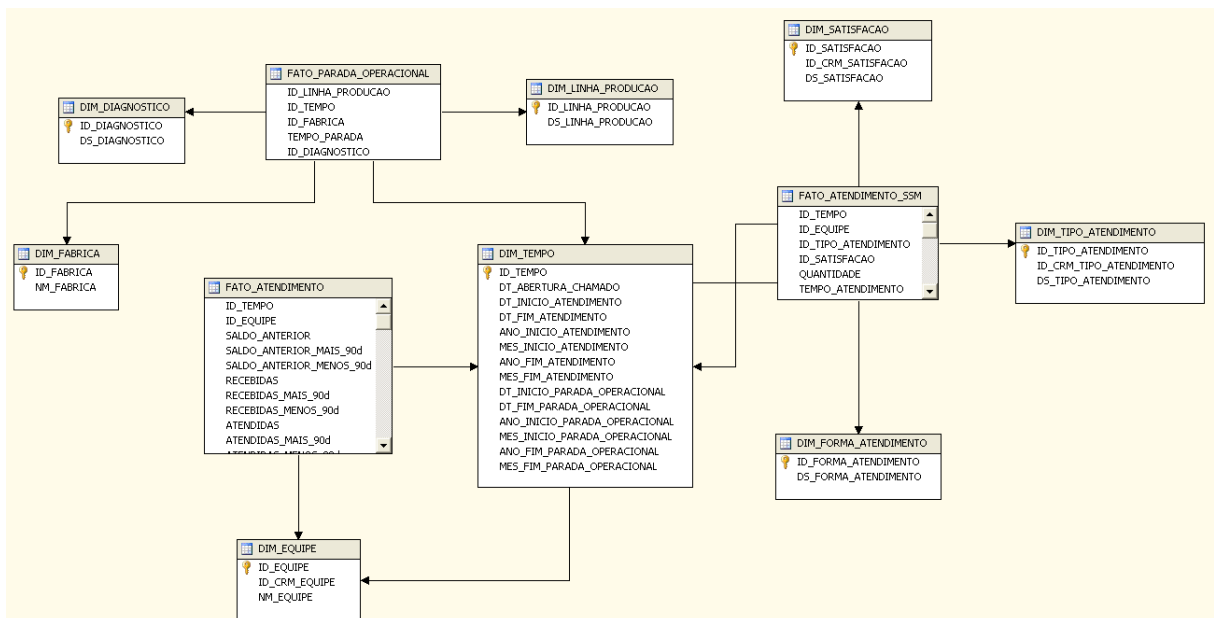


Figura 19 – Modelo multidimensional para a manutenção industrial

As principais métricas de atendimento dos serviços solicitados à manutenção estão nas tabelas FATO_ATENDIMENTO e FATO_ATENDIMENTO_SSM: quantidade e tempo de atendimento. O tempo e quantidade de paradas operacionais estão na tabela FATO_PARADA_OPERACIONAL e as dimensões foram projetadas com os dados de negócio necessários à manutenção, importantes no momento do cálculo dos indicadores das tabelas fato.

Outras métricas podem ser calculadas utilizando a aba *Calculations* do SSAS e assim obter valores calculados (*calculated members*) através da combinação dos dados do cubo, operações aritméticas, números e funções. A Figura 22 mostra a criação de cálculos de taxa de avaria, MTTR, MTBF e taxa de disponibilidade com o uso desta funcionalidade.

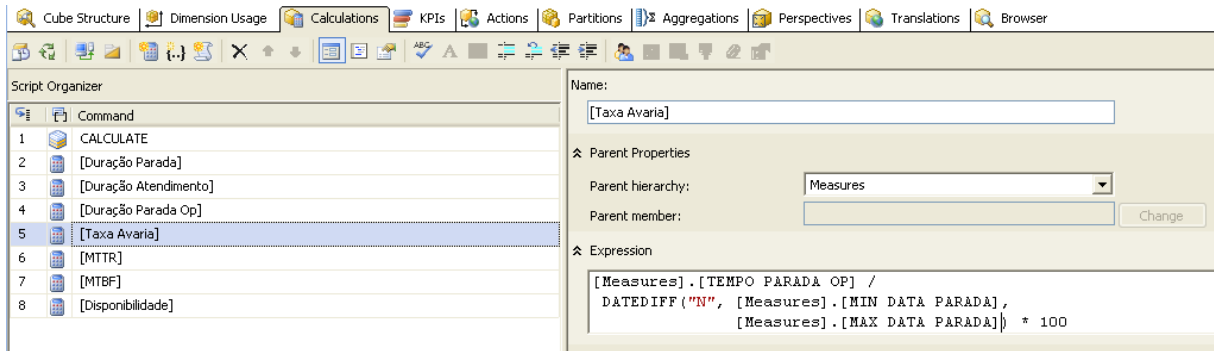


Figura 22 – Configuração dos *calculated members* do projeto

De acordo com Assis (2004), os principais indicadores de manutenção são:

- A taxa de avarias (λ) resulta o número de avarias por unidade de utilização, ou seja, a quantidade de avarias do equipamento em um determinado tempo.

$$\text{Taxa de Avarias} = \frac{\text{número de avarias}}{\text{tempo total de funcionamento no período}} * 1000$$

- O MTTR indica o tempo médio para recuperação do sistema em caso de avaria. A unidade do MTTR é expressa em horas.

$$MTTR = \sum \frac{\text{tempos utilizados nas reparações no período}}{\text{número de avarias}}$$

- O MTBF é um termo de fiabilidade que representa o tempo médio entre avarias. A unidade do MTBF costuma ser expressa em horas.

$$MTBF = \sum \frac{\text{tempos de funcionamento no período}}{\text{número de avarias}}$$

- O MTW indica a média dos tempos de espera de atendimento de solicitações de reparos de avarias num determinado período.

$$MTW = \sum \frac{\text{tempos de espera no período}}{\text{número de avarias}}$$

- A disponibilidade representa a fração do tempo em que o sistema encontra-se operacional e acessível.

$$Disponibilidade = \frac{MTBF}{MTBF + MTTR + MTW}$$

Por fim, a ferramenta do *Cube Browser* permite demonstrar as métricas de desempenho e as dimensões relacionadas com elas, utilizando filtros (se necessário) para visualizar as perspectivas presentes na análise dos dados. A Figura 23 demonstra a soma de tempo das ocorrências em sistemas ou equipamentos que ocasionaram parada de produção entre os anos de 2008 e 2012. No lado esquerdo encontram-se as métricas de desempenho das tabelas fato e as dimensões a serem relacionadas; no lado direito, acima, podem ser incluídos filtros que contribuem para uma análise mais específica dos dados.

Measure Group:	Dimension	Hierarchy	Operator	Filter Expression
<All>	DIM TIPO ATENDIMENTO	DS TIPO ATENDIMENTO	Equal	{ Ocorrência }
	<Select dimension>			

ANO INICIO PARADA OPERACIONAL						
	2008	2009	2010	2011	2012	Grand Total
NM EQUIPE	Duração Parada	Duração Parada	Duração Parada	Duração Parada	Duração Parada	Duração Parada
Aciaaria	38:31	4:55	37:13	1:16		1:21:55
Chapas Grossa	2:12	1:37	1:37			5:26
Energia e Transporte	1:00					1:00
Laminação a frio	21:37	30:08	4:28	2:46	3:30	1:02:29
Laminação a quente	:32	2:04	32:08	1:19	1:45	37:48
Grand Total	1:03:52	38:44	1:15:26	5:21	5:15	3:08:38

Figura 23 – Cubo da duração de paradas operacionais por equipe

3.5 Weka

O software livre *Waikato Enviroment for Knowledge Analysis (Weka)* é desenvolvido em Java e formado por um conjunto de algoritmos para diversas técnicas de mineração de dados (UNIVERSITY OF WAIKATO, 2010). Está sob domínio da licença GPL e contém ferramentas para classificação (árvore de decisão, *Naive Bayes* entre outras), predição numérica (regressão linear, *perceptron* multicamadas e outras), agrupamento (*SimpleKMeans*, *DBScan* e outras) e associação (*Apriori*, *FPGrowth* e outras).

Como descrito anteriormente, um dos primeiros passos no KDD é a seleção dos dados, armazenados no SGBD *Microsoft SQL Server* para limpeza, preparação e processamento

deles em busca de padrões e regras de interesse. O *Weka* trabalha com o formato ARFF para organizar os dados, que utiliza marcações como *@relation*, *@attribute* (definição do tipo e valores que ele pode representar) e *@data* (instâncias presentes nos dados) nas atividades de mineração de dados. A Figura 24 representa a tela inicial do *Weka*, que permite operacionalizar por linhas de comando ou interface gráfica.



Figura 24 – Tela inicial do *Weka*

A opção *Explorer* é utilizada neste trabalho, pois é a principal interface para análise de dados do *software* de mineração de dados. Dentre as outras aplicações, a *Experimenter* realiza comparações de desempenho entre algoritmos, a *KnowledgeFlow* concede uma alternativa para a *Explorer* baseada em componentes e a *SimpleCLI* é um *shell* de comandos simples. A Figura 25 exhibe a *interface* com as seis guias para explorar os dados: *Preprocess*, *Classify*, *Cluster*, *Associate*, *Select attributes* e *Visualize*.

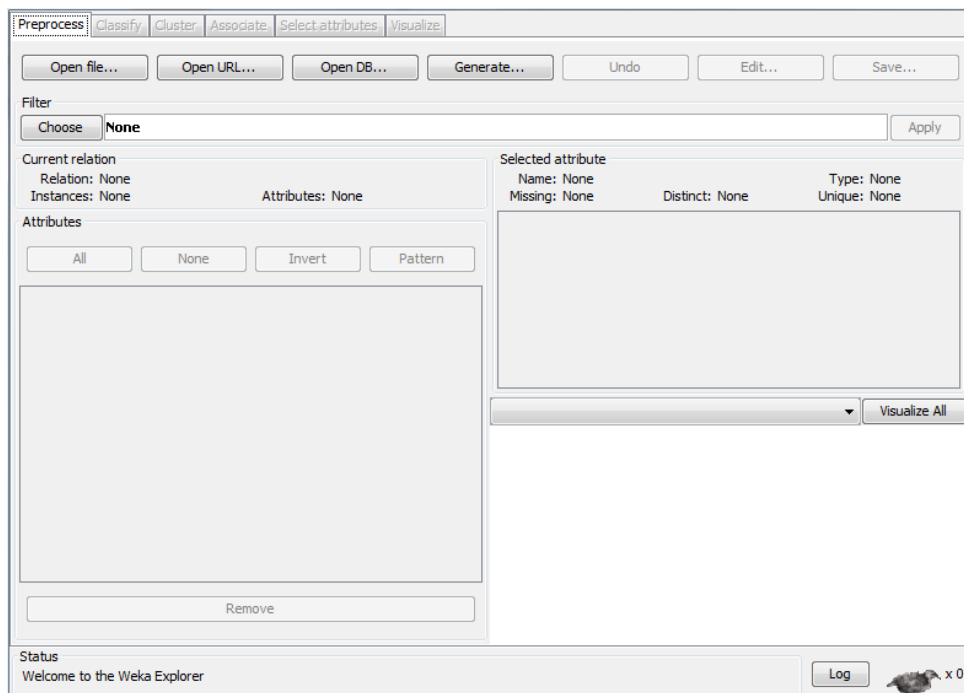


Figura 25 – Interface do *Explorer Weka*

Os algoritmos de regressão linear, classificação e associação foram utilizados nas análises da pesquisa. A classe *weka.classifiers.LinearRegression* para o uso da regressão linear, adota o critério *Akaike* para a seleção do modelo, capaz de lidar com as instâncias ponderadas. O objetivo é descobrir o custo gerado para a empresa a partir das ocorrências nas fábricas, ocasionadas pelas falhas dos sistemas de automação industrial.

Para as tarefas de classificação, o *Weka* dispõe do algoritmo J48 (atualização do C4.5), desenvolvido por Ross Quinlan, que permite a criação de modelos de decisão em árvore, utilizando uma tecnologia na qual todas as decisões que ele toma, visa àquela com a mais imediata e óbvia vantagem (MARKEY, 2011; QUINLAN e KOHAVI, 1999). As árvores de decisão são construídas de forma hierárquica (do topo para a base), por meio da seleção do atributo mais apropriado para cada situação (BOUCKAERT et al., 2008). Basgalupp (2010) reforça que ainda acontece o processo de poda no final do algoritmo para reduzir o excesso de ajustes (*overfitting*) aos dados de treinamento.

O algoritmo *Apriori* trabalha para identificar as regras de associação entre os atributos selecionados, configurando um suporte e confiança mínima nos parâmetros do *Weka*. Como ele tem limitações para trabalhar com valores numéricos, o valor de duração das atividades foi discretizado em < 8 horas e ≥ 8 horas durante a geração de conjuntos de itens frequentes. Lembrando que o Suporte de $X \rightarrow Y$ considera o $\frac{\text{número de casos que contém } X \text{ e } Y}{\text{número total de registros}}$ e a Confiança de $X \rightarrow Y$ é o $\frac{\text{número de registros que contém } X \text{ e } Y}{\text{número de registros que contém } X}$, que são os parâmetros *lowerBoundMinSupport* (suporte mínimo) e *minMetric* (confiança mínima) no *Weka*.

3.6 Apresentação da informação

O *SQL Server Reporting Services* (SSRS) foi utilizado para apresentar as informações disponíveis nos ambientes OLTP e OLAP, pois permite criar relatórios e gráficos baseados em banco de dados relacionais e multidimensionais. A Figura 26 demonstra a configuração de um *report* realizando uma consulta MDX na tabela fato e usando as dimensões ano e mês do atendimento como parâmetros de filtragem na leitura dos resultados.

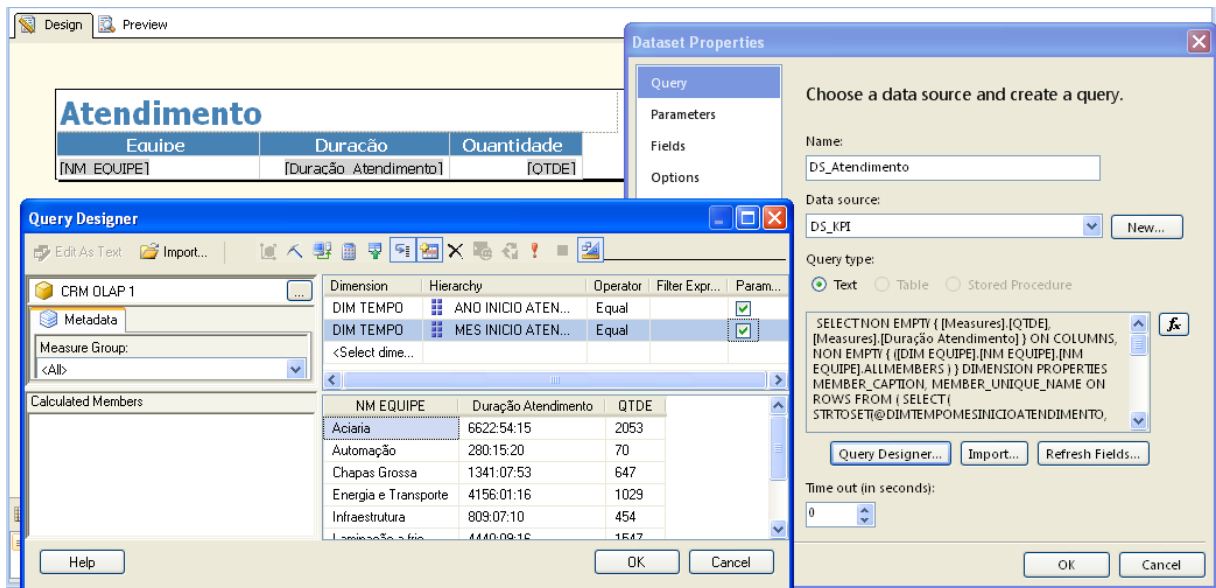


Figura 26 – Configuração do ambiente do Reporting Services

O Excel possui *drivers* de conexão com banco de dados e com os cubos criados no *Analysis Services*, permitindo também criar relatórios e gráficos por meio de tabelas dinâmicas (Figura 27). Este recurso de análise adicional é usado com filtros, classificações e ordenações para resumir grandes quantidades de dados, importados de fontes externas.

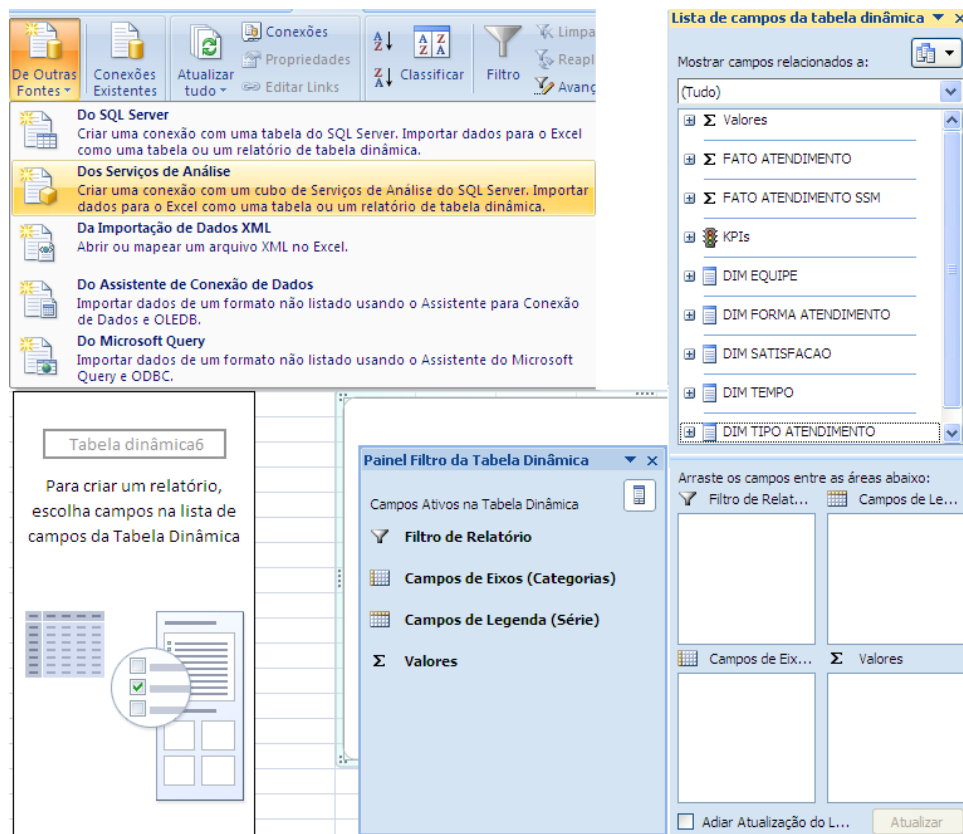


Figura 27 – Configuração do Excel com o Analysis Services

4 RESULTADOS E DISCUSSÃO

4.1 Implementação do *Data Warehouse*

Os resultados do trabalho apresentaram os benefícios do uso do BI no âmbito industrial, pois como se pôde observar na fundamentação teórica, a informação e o conhecimento dos processos da empresa são essenciais para a tomada de decisões estratégicas.

O KPI de “solicitações atendidas” mensura os chamados atendidos em relação ao saldo anterior (em carteira) com os chamados recebidos, desconsiderando os chamados cancelados, que pode ter sido aberto indevidamente. O ideal é que mais de 97% dos chamados tenham sido atendidos.

$$\text{IndicadorAtend.} = \frac{\text{Atendidas}}{\text{Saldoanterior} + \text{Recebidas} - \text{Canceladas}} * 100$$

O *dashboard* criado no Excel (Figura 28) é usado para acompanhamento destes indicadores. Os rótulos de linha utilizados para as tabelas e gráficos dinâmicos foram: ano de atendimento – 2012 e nome da equipe.

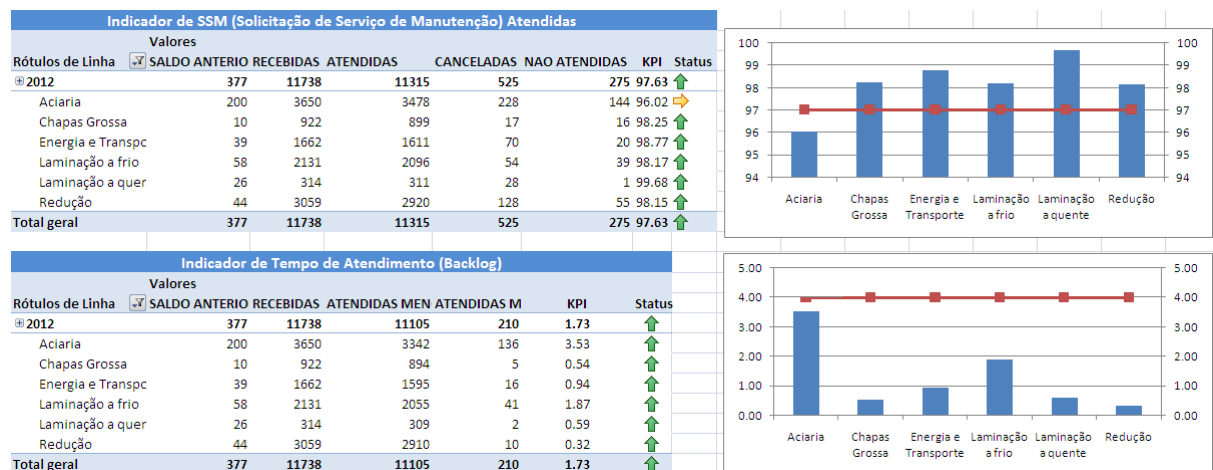


Figura 28 – *Dashboard* de indicadores de atendimento dos serviços

No *dashboard* da Figura 29 também foram criados indicadores de paradas da linha de produção por falha em equipamentos ou falha em sistemas da gerência em questão. O indicador considerado máximo pela gerência é de 01h35m/mês de parada na linha de produção por falha de manutenção. Os outros gráficos representam a satisfação dos clientes

em relação às solicitações de serviços e à quantidade de atendimentos com mais de 90 dias em relação aos que tiveram menos de 90 dias.

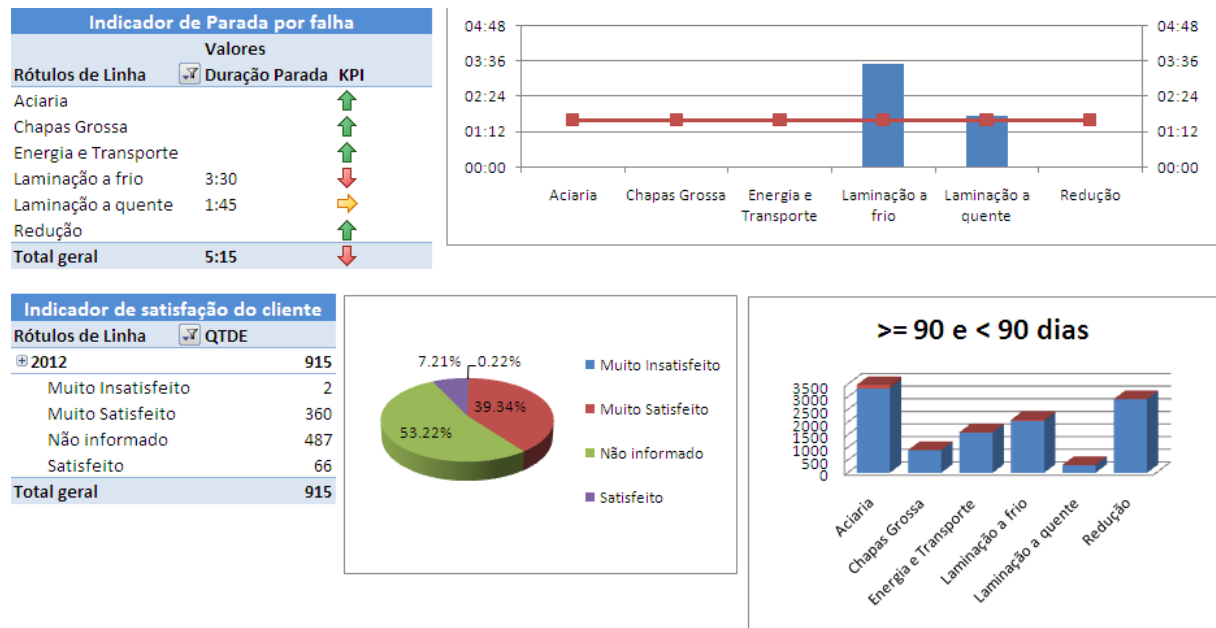


Figura 29 – Dashboard de indicadores de parada por falha e satisfação do cliente

Outro importante KPI é o tempo de atendimento das ocorrências (*backlog*), ou seja, em quanto tempo um problema considerado como emergência levou para ser resolvido. É o que representa os dados da Figura 30: quantidade de atendimentos em ocorrências e tempo médio para resolução entre os anos de 2008 e 2012.

	ANO INICIO ATENDIMENTO ▾											
	2008		2009		2010		2011		2012		Grand Total	
NM EQUIPE ▾	QTDE	Tempo médio	QTDE	Tempo médio	QTDE	Tempo médio	QTDE	Tempo médio	QTDE	Tempo médio	QTDE	Tempo médio
Aciaria	314	5:54	176	4:27	162	8:31	161	4:52	205	6:93	1018	5:03
Chapas Grossa	46	9:38	50	6:88	50	3:99	57	8:67	116	6:55	319	6:86
Energia e Transporte	73	4:67	69	10:96	67	5:63	44	2:18	110	3:45	363	5:43
Laminação a frio	224	3:57	151	1:94	124	4:28	90	5:67	114	3:93	703	3:54
Laminação a quente	99	3:59	72	8:78	71	3:51	84	5:95	121	1:44	447	4:63
Redução	182	6:33	173	15:12	102	10:39	79	9:21	122	8:74	658	10:26
Grand Total	938	4:21	691	7:06	576	6:89	515	5:31	788	5:97	3508	5:47

Figura 30 – Tempo médio de atendimento das ocorrências

Os relatórios criados no ambiente do SSRS contêm os dados necessários para análises mais detalhadas sobre o andamento de qualquer indicador mencionado anteriormente. A Figura 31 exibe todos os atendimentos com duração superior a três meses, que deveriam ser tratados como um projeto e não como um atendimento, de acordo com definição interna da empresa.

Relatório - Atendimentos com duração > 90 dias

ID	Título	Início	Término	Duração - dias
SCA-07720-QP1W	IL02 - Auxiliar a configuração de uma saída analógica no CLP da TRFA2 e configurar nos Supervisórios da Ilha2.	07/06/2005 00:06	28/06/2011 10:06	2212
SCA-01135-NXMV	LTF1 - Tempo de Parada no LTF	21/06/2007 00:06	21/06/2009 16:06	731
SCA-02856-URT2	IMPLANTAÇÃO DO SISTEMA DE REGISTRO DE VARIÁVEIS AMBIENTAIS	14/01/2008 00:01	21/05/2009 00:05	493
SCA-02021-NDG7	FOP1 /S: Projeto de modernização da automação do Forno Panela	05/03/2008 00:03	30/06/2009 08:06	482
SCA-04203-ER1L	DEP6 /S: Implantação do novo Supervisório de Despeiramento do CV5/6 na plataforma Windows XP x Ellipse E3	10/03/2008 00:03	22/04/2009 00:04	408
SCA-01864-3VCA	LI01- Transferir controle de emissão de etiquetas para a tabela "TB_etiqueta_in"	31/03/2008 00:03	15/06/2009 00:06	441
SCA-02891-M8HS	Implantação das telas da turbina de topo do AF#2 na máquina Principal do SSE.	12/06/2008 00:06	18/06/2009 16:06	371
SCA-02864-YNOL	CETT - IMPLEMENTAR COMUNICAÇÃO ENTRE A MÁQUINA DE TRAÇÃO SHIMADZU DE 100t E O NÍVEL 3	11/07/2008 00:07	21/07/2009 09:07	375
SCA-04162-2KOE	TF01 - Implantação do sistema de leitor óptico.	28/07/2008 00:07	27/03/2009 00:03	242
SCA-02883-K94D	SMRU - Controle das Mesas e Transferencias dos Pátios do Acabamento	18/08/2008 00:08	27/06/2009 00:06	313
SCA-03286-7YCG	LTF1 - Verificar um método de atualização automática do Symantec End Point Manager.	22/08/2008 00:08	02/03/2009 00:03	192
SCA-03407-CXD3	LTF1 - Melhoria da Adaptação do Modelo Matemático de Otimização do Laminador de Tiras a Frio	16/09/2008 00:09	29/06/2009 15:06	286

Figura 31 – Relatório de atendimentos realizados em prazo maior do que três meses

Na utilização do SSAS e suas ferramentas de BI foram realizadas diversas análises, e entre elas, cubos que ajudaram a identificar os equipamentos com maior incidência de quebra e o prejuízo na produção que ocasionaram. A Figura 32 demonstra a quantidade de incidências em duas fábricas da usina e as métricas de confiabilidade de cada linha de produção.

Dimension	Hierarchy	Operator	Filter Expression
DIM FABRICA	■ NM FABRICA	Equal	{ Sinterização 2, Sinterização 3 }
DIM TEMPO	■ ANO INICIO PARADA OPERACI...	Equal	{ 2014 }

NM FABRICA ▼	DS LINHA PRODUCAO ▼	QTDE PARADA	Duração Parada Op	MTBF	MTTR	Taxa Avaria	Disponibilidade
■ Sinterização 2		2028	25:53:04	196.39	45.95	18.96	81.04
■ Sinterização 3		1945	26:20:28	203.93	48.75	19.29	80.71
Grand Total		3973	52:13:32	76.38	47.32	38.26	61.74

Figura 32 – Análise de disponibilidade por fábrica

A partir das informações obtidas foi possível notar que em 2014, a unidade Sinterização 2 teve mais ocorrências de paradas operacionais que a Sinterização 3, porém com uma duração total inferior (tempo menor) o que justifica o índice de disponibilidade melhor. O tempo médio de reparo foi de 46 minutos (45,95) e a cada três horas (196,39 min.), em média, ocorreu uma parada operacional.

Ao incluir a dimensão DIM LINHA PRODUCAO nesta análise foi possível examinar quais linhas de produção tiveram mais influência nas paradas da fábrica (Figura 33).

NM FABRICA	DS LINHA PRODUCAO	QTDE PARADA	Duração Parada Op	MTBF	MTTR	Taxa Avaria	Disponibilidade
Sinterização 2	Outros Sistemas/Equipamentos	310	2:58:59	1,550.73	34.64	2.19	97.81
	Sistema Abastec. Blendado	219	1:57:07	2,212.05	32.09	1.43	98.57
	Resfriador	194	36:54	2,521.92	11.41	0.45	99.55
	Sistema Envio Sinter (abastecimento)	185	1:35:50	2,625.49	31.08	1.17	98.83
	Sistema Sinter Produto (Preparação)	151	1:27:45	3,219.87	34.87	1.07	98.93
	Sistema Mist. Intensivo	111	35:18	4,408.54	19.08	0.43	99.57
	Manutenção Programada (Orçamento)	89	6:55:38	5,241.89	280.20	5.07	94.93
	Outros Serviços Operacionais	78	11:55	6,291.68	9.17	0.15	99.85
	Peneiras Sinter	67	1:08:09	7,274.28	61.03	0.83	99.17
	Máquina de Sinter (tremonha a penei	60	27:56	8,163.17	27.93	0.34	99.66
	Sistema Carregamento MS	54	12:38	9,087.19	14.04	0.15	99.85
	Limpeza Emerg. Calhas	52	18:29	9,429.94	21.33	0.23	99.77
	Sistema Abastec. Diversos	44	25:27	11,134.98	34.70	0.31	99.69
	Sistema Retorno	41	25:06	11,950.24	36.73	0.31	99.69
NM FABRICA	DS LINHA PRODUCAO	QTDE PARADA	Duração Parada Op	MTBF	MTTR	Taxa Avaria	Disponibilidade
Sinterização 3	Outros Serviços Operacionais	407	1:52:36	1,190.93	16.60	1.37	98.63
	Sistema Envio Sinter (abastecimento)	195	1:23:22	2,494.69	25.65	1.02	98.98
	Manutenção Programada (Orçamento)	165	11:17:29	2,732.22	246.36	8.27	91.73
	Sistema Abastec. Blendado	152	51:59	3,212.81	20.52	0.63	99.37
	Sistema Mist. Intensivo	135	47:41	3,619.30	21.19	0.58	99.42
	Outros Sistemas/Equipamentos	95	36:38	5,150.19	23.14	0.45	99.55
	Sistema Sinter Produto (Preparação)	73	23:40	6,712.96	19.45	0.29	99.71
	Sistema Carregamento MS	61	31:17	8,026.05	30.77	0.38	99.62
	Outras Limpezas Emergenciais	54	13:04	9,086.70	14.52	0.16	99.84
	Limpeza Emerg. Correias (retorno, c	48	7:33	10,229.44	9.44	0.09	99.91
	Grande Reparo	45	4:06:00	10,593.47	328.00	3.00	97.00
	Instrumentação Analítica	45	20:41	10,893.89	27.58	0.25	99.75
	Limpeza Emerg. Calhas	45	12:45	10,904.47	17.00	0.16	99.84
	Máquina de Sinter (tremonha a penei	45	8:39	10,909.93	11.53	0.11	99.89
	Peneiras Sinter	42	17:21	11,676.79	24.79	0.21	99.79
	Resfriador	38	13:50	12,911.47	21.84	0.17	99.83


Figura 33 – Análise das paradas operacionais por linha de produção

Também acrescentou-se a dimensão DIM DIAGNOSTICO nesta análise para relacionar a natureza da falha com a linha de produção e assim conhecer quais os problemas recorrentes e danosos à operação (Figura 34).

NM FABRICA	DS DIAGNOSTICO	QTDE PARADA	Duração Parada Op	NM FABRICA	DS DIAGNOSTICO	QTDE PARADA	Duração Parada Op
Sinterização 2	resfriador c. queda de velocidade	101	3:02	Sinterização 3	PARADA PROGRAMADA	104	8:32:24
	PARADA PROGRAMADA	54	4:25:48		PREVENTIVA	41	1:56:49
	CONTROLE DE ESTOQUE DE SINTER.	40	3:36:55		GRANDE REPARO.	35	2:46:00
	Retirando pedra da comporta do rolo de mistura	27	1:50		NÍVEL BAIXO DA TREMONHA DE MISTURA	31	2:02
	QUEDA DE VELOCIDADE NO RESFRIADOR	19	:37		Ajustando nível da tremonha de mistura.	30	1:51
	PARADA PREVENTIVA	18	1:27:12		TROCA DAS PENEIRAS DO 3º PENEIRAMENTO	30	9:37
	CT 130 - DETETOR DE METAL.	17	:27		PICO DE QUEIMA NA MAQUINA DE SINTER	22	4:46
	Parada programada.	13	58:03		av 14 com nível super alto	19	6:55
	CT. 130- DETETOR DE METAL.	10	:18		TREMONHA DE MISTURA ENGATOLADA	19	2:17
	ENGATOLAMENTO DA TREMONHA DE MISTURA	9	1:10		resfriador c. queda de velocidade	17	9:36
	CT 115 - SOBRE CARGA (DESALINHAMENTO).	7	2:52		SOBRECARGA NA MAQUINA DE SINTER	15	1:15
	CT 168V31 - DESALINHAMENTO	7	1:23		CFW 07 C.DESVIO DE PESO	12	2:20
	FALTA DE MISTURA	7	7:31		ESTOQUE DE SINTER ALTO	12	33:28
	Calha chela na ct. 123.	6	:41		NÍVEL BAIXO DA TREMONHA DE FALSA GRELHA	11	:14
	CT 120 - OBSTRUÇÃO NA CALHA.	6	1:53		DESARME NO SISTEMA S	10	1:28
	CT M301 C/ LIMITE DE DESALINHAMENTO	6	:37		GRANDE REPARO	10	1:20:00
	FALTA DE MISTURA.	6	2:29		RESFRIADOR - QUEDA DE VELOCIDADE.	10	:49
	NÍVEL BAIXO DA TREMONHA DE MISTURA	6	:13		ENGATOLAMENTO DA TREMONHA DE MISTURA	9	:51
	RESFRIADOR - QUEDA DE VELOCIDADE	6	:06		PARADA PREVENTIVA	9	21:16
	Aferição nas balanças da dosagem.	5	:55		CT168V11 C. LIMITE DE EMERGENCIA	8	1:36
	CT 124 - OBSTRUÇÃO NA CALHA (PEDRA GRAN)	5	:34		DRENANDO MATERIAL DO RESFRIADOR - MATE	8	:40
	CT. 120- OBSTRUÇÃO DA CALHA (MATERIAL UN	5	4:21		FALTA DE GCO.	8	16:32
	ESTOQUE DE SINTER ALTO	5	22:53		LIMPEZA NAS MALHAS DA PENEIRA	8	2:45
	FALTA DE MISTURA - SILO 02 E 04 VAZIO.	5	2:38		máquina de sinter com sobrecarga	8	2:38
	PARADA P/ PIQUE DE QUEIMA.	5	1:58		CARRO SEPARADOR DE PEDRAS OBSTRUÍDO	7	2:14
	PREPARAÇÃO PARA TROCA DE SETOR	5	:16		CONTROLE DE ESTOQUE DE SINTER.	7	27:32
	Queda de velocidade no misturador Primário (04	5	:33		AV 14 - OBSTRUÇÃO, PEDRA GRANDE	6	:45
	SOBRE-CARGA MAQUINA DE SINTER MOTOR LA	5	2:54		CT 168V11 DESALINHADA	6	:57
	CT - 125 - QUEDA DE VELOCIDADE (PEDRA GRA	4	1:22		CT 168V21 - CALHA CHEIA.	6	1:13
	CT 121 - QUEDA DE VELOCIDADE (CENTRÍFUGO	4	:52		CT 168V31 - DESALINHAMENTO	6	:44
	CT 133 - QUEDA DE VELOCIDADE.	4	:17		CT M305 - DESALINHAMENTO.	6	:57
	CT. 124- QUEDA DE VELOCIDADE.	4	1:00		CT. 168V21- LIMITE DE CALHA CHEIA.	6	1:29
	CT. 133- LIMITE DE CALHA CHEIA.	4	:16		CT168V21 C.LIMITE DE EMERGENCIA	6	:33

Figura 34 – Análise das paradas operacionais pela natureza da falha

A organização dos dados, seguindo o modelo multidimensional, permite a realização de operações OLAP, na qual o usuário pode analisar os dados de diferentes perspectivas e variados níveis de detalhe. Kimball (1998) define a operação de *Drill Down* como a apresentação de dados num nível de abstração mais específico, aumentando o nível de detalhe da informação. A Figura 35 mostra um maior detalhamento da informação como quantidade e total de horas paradas no ano de 2014, usando a hierarquia de ano-mês da dimensão tempo, para exibir as ocorrências específicas do mês de maio.




ANO ▼		
2014		
NM FABRICA ▼	QTDE PARADA	Duração Parada
Sinterização 2	2028	25:53:04
Sinterização 3	1945	26:20:28
Grand Total	3973	52:13:32

NM FABRICA ▼	ANO ▼	MES	QTDE PARADA	Duração Parada
<input type="checkbox"/> Sinterização 2	<input type="checkbox"/> 2014	5	326	2:49:03
<input type="checkbox"/> Sinterização 3	<input type="checkbox"/> 2014	5	79	1:04:55
Grand Total			405	3:53:58

Figura 35 – Operação *Drill Down* nos dados de paradas

A operação de *Drill Up* (*Roll Up*) apresenta os dados em um nível de abstração superior e segundo Machado (2010), o usuário aumenta o nível de granularidade, diminuindo o nível de detalhamento da informação. Na Figura 36, a visualização do diagnóstico da falha que estava no mês de setembro de 2014 passou a ser sumarizada pelo ano de 2014, sem especificação de mês de parada.



ANO ▼ MES		
<input type="checkbox"/> 2014		
9		
DS DIAGNOSTICO ▼	QTDE PARADA	Duração Parada
TROCA DAS PENEIRAS DO 3º PENEIRAMENTO	15	4:08
NIVEL BAIXO DA TREMONHA DE MISTURA	13	:42
Grand Total	28	4:50

ANO ▼		
2014		
DS DIAGNOSTICO ▼	QTDE PARADA	Duração Parada
NIVEL BAIXO DA TREMONHA DE MISTURA	31	2:02
TROCA DAS PENEIRAS DO 3º PENEIRAMENTO	30	9:37
Grand Total	61	11:39

Figura 36 – Operação *Drill Up* nos dados de diagnóstico de falha

Com a operação *Drill Across*, ao menos dois cubos precisam de uma das dimensões em comum. A Figura 37 demonstra um registro de falha no sistema de pesagem, relacionando as tabelas fato FATO ATENDIMENTO e FATO PARADA OPERACIONAL.

DS DIAGNOSTICO ▼	Duração Atendimento	Duração Parada Op
Sem indicação de peso na balança B-1.	1:05	2:10
Grand Total	1:05	2:10

Figura 37 – Operação *Drill Across* relacionada com as tabelas fato

E por fim, a operação *Slice and Dice*, definida por Machado (2010) como uma forma simples de reduzir o escopo dos dados em análise, mudando a ordem das dimensões e respectiva orientação na qual os dados são visualizados. A Figura 38 mostra a operação *Slice* que seleciona o nível de satisfação com relação aos atendimentos realizados em 2013 e a operação *Dice* que envolve três dimensões (ano de atendimento = 2013 || tipo de atendimento = “Ocorrência” || equipe de atendimento = “Chapas Grossas” ou “Energia e Transporte” ou “Laminação a quente” ou “Laminação a frio” ou “Redução” ou “Aciaria”).

Dimension	Hierarchy	Operator	Filter Expression
DIM TEMPO	Ano-mes Atendimento	Equal	{ 2013 }
DIM TIPO ATENDIMENTO	DS TIPO ATENDIMENTO	Equal	{ Ocorrência }
DIM EQUIPE	NM EQUIPE	Equal	{ Chapas Grossa, Energia e Transporte, Laminação a quente, Laminação a frio, Redução, Aciaria }

DS SATISFACAO ▼					
Muito Satisfeito Não informado Satisfeito Grand Total					
DS FORMA ATENDIMENTO ▼	QTDE	QTDE	QTDE	QTDE	
Local (fora do ADM)	3	1	1	5	
Não Acionado	42	29		71	
Remoto (fora do ADM)	45	115	2	162	
Remoto (no ADM)	33	26		59	
Grand Total	123	171	3	297	

Figura 38 – Operação *Slice and Dice* relacionada com as tabelas fato

4.2 Mineração de dados na siderúrgica

4.2.1 Regressão linear

Foi utilizado o valor salarial com a duração do atendimento, também o custo da unidade parada (em caso de interrupção no processo de produção) e a duração como “variáveis dependentes”; o cargo ocupado e a fábrica de produção como “variáveis independentes”. A Figura 39 exhibe os parâmetros (arquivo ARFF) e a equação gerada pelo algoritmo *LinearRegression* para descobrir o custo das ocorrências nas fábricas, ocasionada pelos sistemas de automação industrial.


```

@relation Manutencao

@attribute Cargo numeric
@attribute DuracaoAtendimento numeric
@attribute HomemHora numeric
@attribute Unidade numeric
@attribute DuracaoParada numeric
@attribute CustoUnidadeParada numeric
@attribute Total numeric

@data
1, 240, 0.28, 3, 0, 17.05, 68.18
2, 960, 0.66, 1, 0, 7.58, 636.36
2, 10, 0.68, 2, 0, 11.36, 6.82
1, 330, 0.36, 3, 0, 17.05, 118.75
1, 80, 0.46, 4, 0, 21.78, 36.74
2, 180, 0.66, 2, 0, 11.36, 119.32
2, 90, 0.68, 2, 0, 11.36, 61.36
2, 1020, 0.62, 2, 0, 11.36, 627.84
1, 30, 0.36, 2, 0, 11.36, 10.8
1, 480, 0.33, 1, 0, 7.58, 159.09
3, 180, 1.25, 2, 0, 11.36, 225
1, 9630, 0.33, 1, 0, 7.58, 3191.76

```

```

Classifier output

Linear Regression Model

Total =

-117.7218 * Cargo +
  0.6387 * DuracaoAtendimento +
 671.4651 * HomemHora +
1107.0974 * Unidade +
 10.602 * DuracaoParada +
-227.5713 * CustoUnidadeParada +
 269.8213

Time taken to build model: 0.25 seconds

=== Evaluation on test split ===
=== Summary ===

Correlation coefficient           0.9524
Mean absolute error              145.9213
Root mean squared error          546.4271
Relative absolute error          28.8243 %
Root relative squared error      32.4409 %
Total Number of Instances       1388

```

Figura 39 – Parâmetros e saída do modelo de regressão linear

Por meio da equação, o gestor poderá justificar a contratação de um turno ou terceiros para atender chamados de emergência e dimensionar o valor do homem/hora com relação à continuidade do negócio. O coeficiente indica que o cargo e o custo da unidade parada influem negativamente no valor final e a unidade (fábrica) tem o maior peso da equação, pois uma fábrica de produto final (laminador, por exemplo) tem o valor agregado muito maior para a empresa que uma fábrica de matéria-prima (sinterização, por exemplo).

4.2.2 Classificação

Ao selecionar o arquivo ARFF, o *Weka* permite explorar os oito atributos e informações estatísticas sobre suas instâncias. A Figura 40 demonstra os gráficos de visualização destas informações: o tipo de chamado (3.872 ocorrências e 2.337 solicitações de serviços); o tempo de atendimento (3.770 com duração inferior a oito horas e 2.444 chamados com mais de oito horas para resolver); o turno de trabalho (4.761 atenderam no horário administrativo e 1.453 fora do horário convencional); se teve consequência ou parada (interrupção) operacional, se reincidente; se utilizou uma solução de contorno para resolver o problema e a área de atendimento.

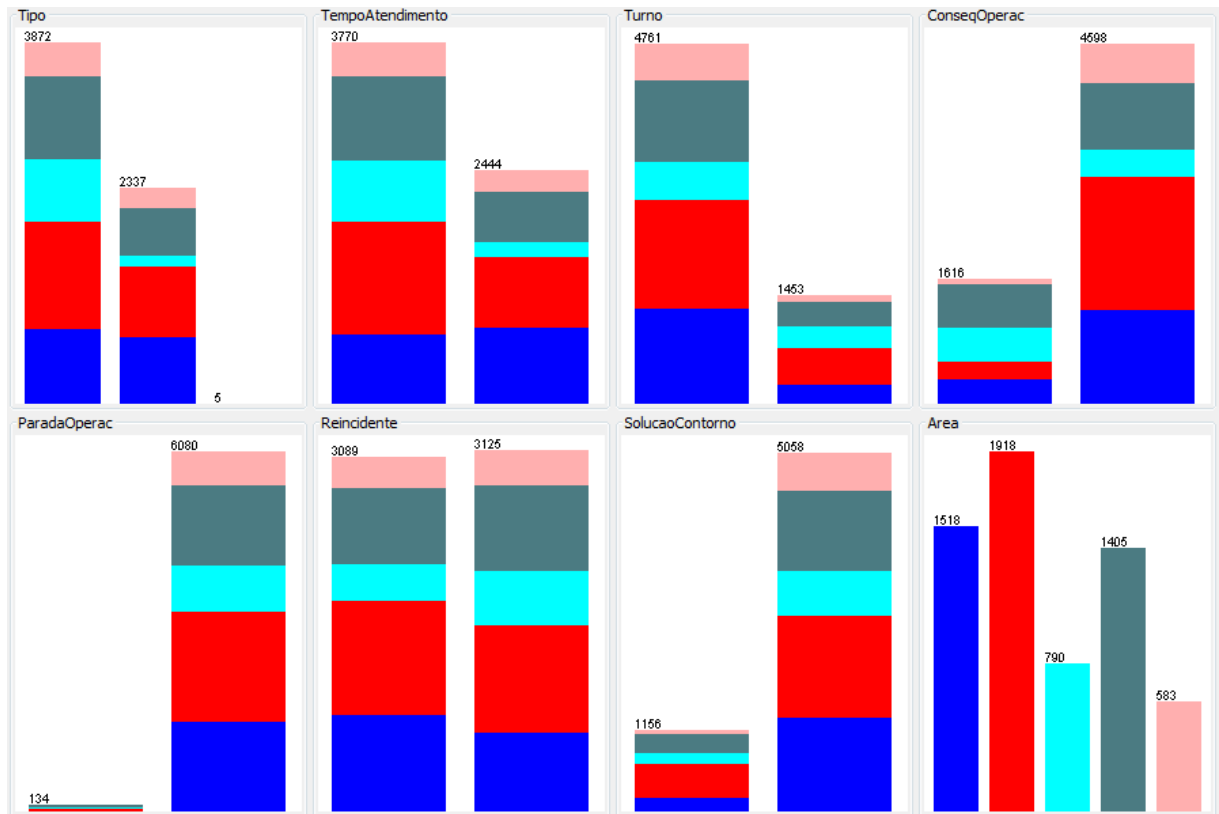


Figura 40 – Gráficos com estatísticas dos atributos e instâncias

Na guia *Classify*, após executar o algoritmo J48 com treinamento de 70% do conjunto de dados, cujos resultados são apresentados na Figura 41, e ainda a construção da árvore, seguida do número de folhas, tamanho da árvore e o tempo de construção do modelo. O relatório informa as instâncias classificadas corretamente e incorretamente, o índice *kappa* – medida de qualidade do modelo - que apresentou 0,834 de concordância, gerando uma matriz de confusão com poucas amostras classificadas erroneamente.

```

J48 pruned tree
-----

TempoAtendimento = <8h: Ocorrencia (3770.0/246.0)
TempoAtendimento = >=8h
|   ConseqOperac = S: Ocorrencia (142.0/8.0)
|   ConseqOperac = N
|   |   Reincidente = Sim
|   |   |   SolucaoContorno = Sim: Ocorrencia (9.0)
|   |   |   SolucaoContorno = Nao
|   |   |   |   Turno = ADM: SSA (2005.0/58.0)
|   |   |   |   Turno = FORA: Ocorrencia (4.0)
|   |   |   Reincidente = Nao
|   |   |   |   Turno = ADM
|   |   |   |   |   Area = RED
|   |   |   |   |   |   SolucaoContorno = Sim: Ocorrencia (8.0/2.0)
|   |   |   |   |   |   SolucaoContorno = Nao: SSA (84.0/30.0)
|   |   |   |   |   |   Area = ACI: SSA (80.0/37.0)
|   |   |   |   |   |   Area = LIQ: Ocorrencia (28.0/8.0)
|   |   |   |   |   |   Area = LIF: SSA (30.0/12.0)
|   |   |   |   |   |   Area = CHG: Ocorrencia (29.0/12.0)
|   |   |   |   Turno = FORA: Ocorrencia (25.0/4.0)

Number of Leaves :    12
Size of the tree :    20
Time taken to build model: 0.02 seconds

=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances      1718      92.1674 %
Incorrectly Classified Instances    146       7.8326 %
Kappa statistic                    0.834
Mean absolute error                 0.0802
Root mean squared error             0.2089
Relative absolute error             25.4422 %
Root relative squared error         52.2537 %
Total Number of Instances          1864

=== Detailed Accuracy By Class ===
                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                -----  -----  -
                0.959    0.136    0.915     0.959    0.937     0.925    Ocorrencia
                0.865    0.041    0.932     0.865    0.898     0.926    SSA
                0         0         0         0         0         0.695    Suporte_Tecnico
Weighted Avg.   0.922    0.098    0.922     0.922    0.921     0.926

=== Confusion Matrix ===
      a  b  c  <-- classified as
1083  46  0  |  a = Ocorrencia
 99   635  0  |  b = SSA
 1     0  0  |  c = Suporte_Tecnico

```

Figura 41 – Relatório de classificação - algoritmo J48

O *visualize tree* do *Weka* apresenta a árvore de decisão (Figura 42), fornecendo os descritores e os limites considerados na execução do modelo. O resultado indica que a maioria das ocorrências é resolvida em menos de oito horas. Com prazo superior a oito horas de trabalho, é classificada como ocorrência caso haja consequência operacional (qualquer atividade que se não realizada prejudique a operação da fábrica) ou solução de contorno (quando o colaborador não teve recurso, conhecimento ou tempo suficiente para resolver o problema em definitivo). Nos casos de problemas reincidentes, existe variação na definição (ocorrência ou solicitação) do chamado de acordo com o horário e equipe de atendimento.

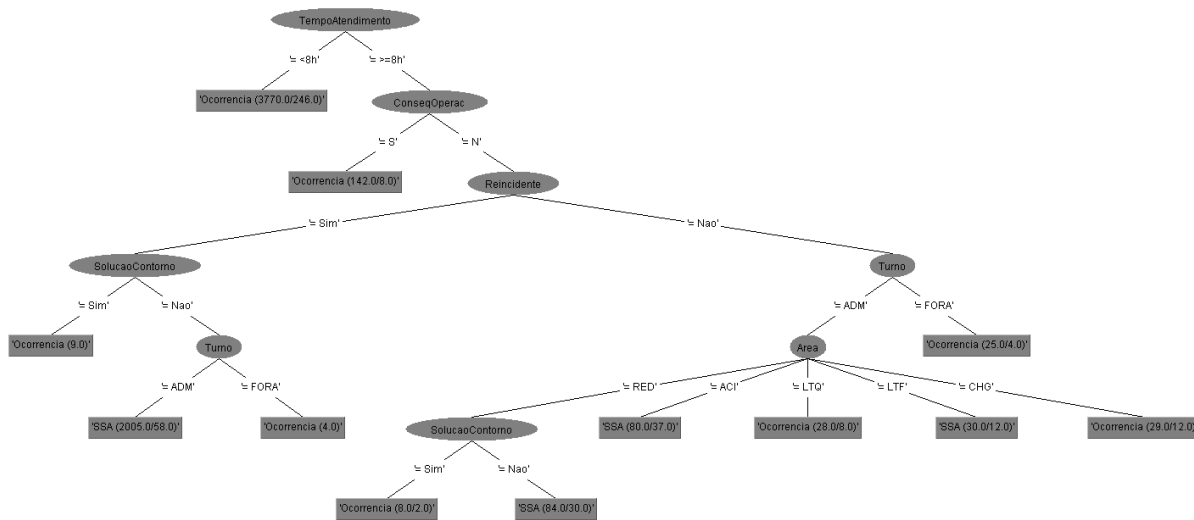


Figura 42 – Árvore de decisão (algoritmo J48)

4.2.3 Associação

Ao configurar o suporte mínimo com 40% e a confiança mínima de 70%, o *Weka* apresentou as melhores regras de associação, que também podem ser configuradas no *software*. A Tabela 3 demonstra as regras de associação detectadas pelo algoritmo *Apriori*, a partir do arquivo ARFF com os atributos “tipo do atendimento”, “duração” ($> 8h$ ou $\leq 8h$), “parada” (se o problema gerou parada na produção da fábrica), “reincidente” (se o problema voltou a acontecer) e “contorno” (utilizou uma solução imediata para resolver o problema e não a melhor solução).

Tabela 3 – Regras de associação (*Apriori*)

	Regra de associação	Confiança
1	Tipo=0 Parada=0 3052 ==> Contorno=0 3017	conf:(0.99)
2	Tipo=1 Contorno=0 3165 ==> Parada=0 3103	conf:(0.98)
3	Reincidente=1 3907 ==> Parada=0 Contorno=0 3733	conf:(0.96)
4	Tipo=1 Parada=0 4324 ==> Duracao_menor_8h=1 3921	conf:(0.91)
5	Reincidente=0 3610 ==> Duracao_menor_8h=1 Parada=0 3016	conf:(0.84)
6	Duracao_menor_8h=1 Parada=0 4252 ==> Reincidente=0 3016	conf:(0.71)

- Regra 1: em solicitações de atendimento (tipo = 0 é uma solicitação de serviço planejada) sem parada operacional, não foram utilizadas solução de contorno no atendimento com 99% de acerto, o que significa que foi usada a melhor solução técnica na maioria dos casos.
- Regra 2: em ocorrências (tipo =1 é um atendimento não planejado, de caráter emergencial) que não foram utilizadas solução de contorno, com 99% de acerto, nas unidades de produção.
- Regra 3: em problemas reincidentes, com 96% de acerto, não ocasionaram parada de produção, nem resolvidas com solução de contorno.
- Regra 4: as ocorrências sem parada operacional, com 91% de acerto, foram atendidas em menos de oito horas.
- Regra 5: em problemas não reincidentes, com 84% de acerto, foram resolvidos em menos de 8 horas com a planta em operação.
- Regra 6: em atendimentos resolvidos em menos de 8 horas, sem parada de produção, com 71% de acerto, não foram reincidentes (problemas que voltaram a acontecer).

5 CONCLUSÃO

Apesar de ter sido aplicado na manutenção industrial, o trabalho demonstrou como os processos de mineração de dados e BI auxiliam a gestão das informações, com possibilidade de aplicação em outras áreas da siderúrgica, como na gestão de processos e da produção.

As técnicas de mineração de dados foram aplicadas aos dados existentes para descobrir padrões e relacionamentos acerca dos atendimentos realizados. Com isso, foi determinado o impacto das ocorrências de sistemas industriais nas fábricas, que auxiliam os gestores a direcionarem investimentos em equipamentos críticos e contratação de turno para atendimentos emergenciais (fora do horário administrativo).

A criação do *Data Warehouse* possibilitou a análise multidimensional e definição de KPIs para controle e monitoramento dos serviços realizados pela gerência de automação e contribuiu com a gestão da manutenção industrial, muito importante no processo de produção da indústria. Com os indicadores de MTBF e MTTR relacionados ao motivo de cada parada operacional, permitiu a elaboração de planos de manutenção dos equipamentos, com o objetivo de melhorar a vida útil e os serviços de inspeção dos mesmos.

Por fim, pôde-se perceber que a consistência e a confiabilidade das informações dependem do uso correto do sistema de abertura de chamados (título, descrição, unidade, sistema, data de início, data de término, solução de contorno, reincidência, motivo da ocorrência e consequência operacional) e do preenchimento da natureza da falha, pois o *status* de funcionamento é uma aquisição automática. Para isso, é importante que os solicitantes dos chamados e os responsáveis pela manutenção dos equipamentos sejam treinados para utilizar corretamente o sistema proposto.

5.1 Recomendações de trabalhos futuros

Após ter a consolidação dos dados de atendimento, paradas operacionais, causa e consequência da falha, pretende-se desenvolver, em trabalhos futuros, um modelo preditivo que atue na descoberta de uma condição operacional (desvio) que possa acarretar em parada de produção da linha, com o objetivo de garantir a manutenção adequada dos equipamentos e segurança dos funcionários da empresa.

Pretende-se ainda aplicar as técnicas de mineração de dados e construção do BI em dados relacionados aos processos de produção das fábricas, com o objetivo de identificar e

executar melhorias nos processos, podendo colaborar com a redução de custos e eliminação de gargalos na produção.

REFERÊNCIAS

- ABRAMAN, editorial. **Documento Nacional 2013**. Disponível em: <http://www.abraman.org.br/Arquivos/403/403.pdf>. Acesso em: 19 jun. 2015.
- ABRAMAN, editorial. **Revista da Associação Brasileira de Manutenção**, São Paulo: n.126, Set, 2010, p.06.
- ABREU, F. S. **Estudo de usabilidade do software *Talend Open Studio* como ferramenta padrão para ETL dos sistemas clientes da aplicação PostGeoOlap**. Monografia (Graduação em Sistemas de Informação) – Faculdade Salesiana Maria Auxiliadora, Macaé, 2007.
- AGRAWAL, R; SRIKANT, R. **Fast algorithms for mining association rules**. 20th International Conference on Very Large Data Bases, p. 487–499, 1994.
- ALMEIDA, A. M. **Proposição de indicadores para avaliação técnica de projetos de Data Warehouse: um estudo de caso no Data Warehouse da plataforma Lattes**. Monografia (Pós- Graduação em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2006.
- AMARAL, C. R. **O uso do *Business Intelligence* como gerador de indicadores de desempenho e investimentos no setor de manutenção industrial**. Indaiatuba: Centro Estadual de Ensino Tecnológico Paula Souza, Faculdade de Tecnologia de Indaiatuba (FATEC), Curso de Análise e Desenvolvimento de Sistemas, 2011.
- ASSIS, R. **Apoio à Decisão em Gestão da Manutenção – Fiabilidade e Manutenibilidade**. Editora Lidel, 1ª ed., 2004.
- Associação Brasileira de Normas Técnicas. **NBR 5462: confiabilidade e manutenibilidade - terminologia**. Rio de Janeiro, 1994.
- BASGALUPP, M. P., **LEGAL-Tree: Um algoritmo genético multi-objetivo lexicográfico para indução de árvores de decisão**. Tese de D.Sc., ICMC/USP SãoCarlos, São Carlos, SP, Brasil. 2010.
- BARBIERI, C. **Business Intelligence: Modelagem e Tecnologia**. Rio de Janeiro: Axcel Books, 2001.
- BATISTA, G. E. A. P. A. **Pré-Processamento de Dados em Aprendizado de Máquina Supervisionado**. Tese de Doutorado, Ciências da Computação e Matemática Computacional, Instituto de Ciência Matemática e de Computação – ICMC-USP, São Carlos – SP. 2003.
- BLOOM, N. B. **Reliability Centered Maintenance: implementation made simple**. New York, McGraw-Hill, 2006.
- BOUCKAERT, R. R.; FRANK, E.; HALL, M.; KIRKBY, R.; REUTEMANN, P.; SEEWALD, A.; SCUSE, D. **WEKA Manual for Version 3-6-0**. 2008.

BOUMAN, R.; DONGEN, J. van; **Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL**. 1ª ed. Indianapolis: Wiley Publishing, 2010. ISBN: 978-0-470-48432-6.

BRACHNAD, R. J.; ANAND, T. The process of knowledge discovery in databases. In FAYYAD, U. M. et al. **Advances in Knowledge Discovery in Data Mining**. Menlo Park: AAAI Press. 1996.

CABRAL J. **Organização e Gestão da Manutenção, dos conceitos à prática**. Editora Lidel, 6ª edição, 2006.

CARDOSO, O. N. P.; MACHADO, R. T. M. Gestão do Conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. **Revista de Administração Pública**, vol. 42, nº 3, p. 495-528, 2008. Disponível em: <<http://www.scielo.br/pdf/rap/v42n3/a04v42n3.pdf>>. Acesso em: 27 mai. 2015.

COOKE, F. L. Implementing TPM in plant maintenance: some organizational barriers. **International Journal of Quality & Reliability Management**, v.17 n. 9, p.1003–1016. 2000.

COOPER, B. L.; WATSON, H. J.; WIXOM B. H.; GOODHUE, D. L. Data Warehousing Supports Corporate Strategy at First American Corporation. **MIS Quarterly**, v. 24, n. 4, p. 547-567, December 2000.

DAMASCENO, M. **Introdução a Mineração de Dados utilizando o WEKA**. Instituto Federal do Rio Grande do Norte – IFRN. Anais... V CONNEPI - Congresso Norte-Nordeste de Pesquisa e Inovação, Macau, RN, Brasil, 2010. Disponível em: <<http://connepi.ifal.edu.br/ocs/anais/conteudo/anais/files/conferences/1/schedConfs/1/papers/258/public/258-4653-1-PB.pdf>>. Acesso em: 08 jun. 2014.

DUNHAM M.H. **Data Mining, Introductory and Advanced Topics**, Editora Prentice Hall, 2002.

DUNKEL, B.; SOPARKAR, N.; SZARO, J.; UTHURUSAMY, R. Systems for KDD: From concepts to practice. **Future Generation Computer Systems**, n. 13, p. 231-242. 1997.

EUROPEAN STANDARD EN 13306 - **Maintenance terminology**. Brussels: European committee for standardization. CEN, 2001.

FAYYAD, U. M.; PIATETSKY-SHAPIO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AAAI Press/The MIT Press**. 1996.

FONSECA, M. de O. **Comunicação OPC – Uma abordagem prática**. In: SEMINÁRIO DE AUTOMAÇÃO DE PROCESSOS DA ABM, 6., 2002, Vitória, Brasil. Anais... Vitória, 2002.

GOLDSCHMIDT, R.; PASSOS, E. **Data Mining: um guia prático**. Rio de Janeiro:Elsevier, 2005.

GUBIANI, J. S.; BRAGA, M. de M.; MIRANDA, J. B.; TEDESCO, J. L. **Inteligência de negócios como um recurso para o processo decisório**. XV SIMPEP – Simpósio de Engenharia de Produção, Bauru, Anais... 2008.

HAN, J; KAMBER, M. **Data Mining: Concepts and Techniques**. Morgan Kaufmann Publishers, 2001.

HAN, J; KAMBER, M. **Data Mining: Concepts and Techniques**. Elsevier, 2006.

HAN, J.; PEI, J.; YIN, Y. Mining partial periodicity using frequent pattern trees. **In GSTech**, Rep, 99-10, Simon Fraser University, July 1999.

IMHOFF, C.; GALEMMO, N.; GEIGER, J. **G. Mastering data warehouse design: relational and dimensional techniques**. Indianapolis: Wiley Publishing, Inc., 2003.

INMON, W. H. **Building the data warehouse**. Fourth Edition. Indianapolis: Published by Wiley Publishing, Inc., 2005.

JACOBSON, R.; MISNER, S. **Microsoft SQL Server 2005: Analysis Services**, Microsoft Press, 2007. Disponível em:
http://www.uniritter.edu.br/graduacao/informatica/sistemas/downloads/tcc2k9/TCCII_Andre_2009_2.pdf. Acesso em 18 mai.2015.

JOHN, V. M.; BONIN, L. C. **Princípios de um Sistema de Manutenção**. In: Seminário sobre manutenção de edifícios. Anais. Porto Alegre: UFRGS, 1988.

KARDEC, A; NASCIF, J. **Manutenção: Função Estratégica**. Rio de Janeiro: Editora Qualitymark, 2.^a Edição. 2001.

KARDEC, A; NASCIF, J. **Manutenção: Função estratégica**. 3.^a ed. Rio de Janeiro: Qualitymark: Petrobras. 2009.

KIMBALL, R. et al. **The data warehouse lifecycle toolkit: expert methods for designing, developing and developing data warehouse**. New York: John Wiley & Sons. 1998.

KIMBALL, R.; ROSS, M. **The data Warehouse Toolkit: The Complete Guide to Dimensional Modeling**, United States of America, John Wiley and sons, inc, 2nd ed. 2002.

LAROSE, D. T. **Discovering Knowledge in Data**. An Introducing to DATA MINING. New Jersey, John Wiley & Sons, 2006. Disponível em:
 <<http://books.google.com.br/books?id=JbPMdPWQIOwC&printsec=frontcover&hl=pt-BR#v=onepage&q=medicine&f=false>> Acesso em: 02 jun. 2015.

LAUDON, K. C.; LAUDON, J. P. **Sistemas de informação gerenciais: administrando a empresa digital**. São Paulo: Pearson Prentice Hall, 2004.

LEMOES, E. P. **Análise de crédito bancário com o uso de Data Mining: Redes Neurais e Árvores de Decisão**. Dissertação de M.Sc, Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba, PR, Brasil, 2003. Disponível em:
 <<http://www.ppgmne.ufpr.br/arquivos/diss/70.pdf>>. Acesso em: 28 fev. 2015.

LORENA H. N. L. **Utilização de técnicas de Business Intelligence para avaliar o processo seletivo Fatec-SJC**. Monografia (Tecnólogo em Banco de Dados) – Faculdade de Tecnologia de São José dos Campos, 2011.

MADU, C. Competing through maintenance strategies. **International Journal of Quality & Reliability Management**, v.17, n. 9, p.937–948. 2000.

MACHADO, F. N. R. **Tecnologia e projeto de data warehouse**. 5ª ed. São Paulo: Érica, 2010.

MARKEY, J. **Using Decision Tree Analysis for Intrusion Detection: A How-To Guide**. SANS Institute Reading Room, Security, 503. 2011.

MÁRQUEZ, A. C; LÉON, P. M; FERNÁNDEZ, J. F. G; MÁRQUEZ, C. P; CAMPOS, M. L. The maintenance management framework: A practical view to maintenance management. **Journal of Quality in Maintenance Engineering**, v. 15, n. 2, p. 167-178, 2009.

MUASSAB, J. R. **Gerenciamento da Manutenção na Indústria Automobilística**, Taubaté. Monografia – Universidade de Taubaté. 2002.

MOUBRAY, J. **Manutenção Centrada em Confiabilidade** (*Reliability-centred Maintenance*). United Kingdom, Biddles Ltd. 2000.

NARDI, A. R. **Fundamentos e modelagem de banco de dados multidimensionais**, MSDN. 2007. Disponível em: <<http://msdn.microsoft.com/pt-br/library/cc518031.aspx>>. Acesso em: 14 mai. 2015.

NGUYEN, D. Q; BRAMMER, C; BAGAJEWICZ, M. New tool for the evaluation of the scheduling of preventive maintenance for chemical process plants. **Industrial & Engineering Chemistry Research**, v. 47, n. 6, p. 110-124, 2008.

NIU, G; YANG, B; PECHT, M. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. **Reliability Engineering and System Safety**. v. 95, p. 786-796, 2010.

O'BRIEN, J. A. **Sistemas de informação e as decisões gerenciais na era da internet**. 2ª edição. São Paulo: Editora Saraiva. 2004.

OHANA, I. **Experimentos de Mineração de Dados Aplicados a Sistemas SCADA de Usinas Hidrelétricas**. Tese (Doutorado), Universidade Federal do Pará. 2012.

OPC FOUNDATION. **OPC Introduction 2012**. Disponível em: <<http://www.opcconnect.com/opcintro.php>>. Acesso em: 08 jul 2015.

PAULA, M. V. **Explorando o Potencial da Plataforma Lattes como fonte de conhecimento organizacional em Ciência e Tecnologia**. Dissertação de M. Sc, Programa de Pós-Graduação em Gestão do Conhecimento e da Tecnologia da Informação, Universidade Católica de Brasília, Brasília, DF, Brasil, 2004. Disponível em: <http://www.bdtd.ucb.br/tede/tde_arquivos/3/TDE-2004-12-03T093632Z-156/Publico/Dissertacao%20Marcelo.pdf>. Acesso em: 01 jun. 2015.

PINTO, A. K.; XAVIER, J. A. N. **Manutenção: função estratégica**. 2.^a Edição. Rio de Janeiro: Qualitymark, 2001.

PINTO, A. K.; RIBEIRO, H. **Gestão Estratégica e Manutenção Autônoma**. Rio de Janeiro: Qualitymark. 2002.

PINTO, V. **Gestão da Manutenção**. Editora IAPMEI, 1a edição, 1999.

PRIMAK, F. V. **Decisões com BI (*Business Intelligence*)**. Rio de Janeiro: Ciência Moderna, 2008.

PUDA, A. P. **Padronização da comunicação através da tecnologia OPC**. 2008. Disponível em: <http://www.isarj.org.br/pdf/artigos/Padronizacao-da-Comunicacao-atraves-da-Tecnologia-OPC.pdf>. Acesso em 07 jul. 2015.

QUINLAN, J. R.; KOHAVI, R. **Decision Tree Discovery**, 3 (Hunt 1962). 1999.

RAUSAND, M. **Reliability Centered Maintenance**. Reliability Engineering and System Safety, v. 60, p. 121-132, 1998.

REZENDE, S. O. **Sistemas Inteligentes: Fundamentos e Aplicações**. Editora Manole Ltda, Barueri, SP, Brasil, 2005.

ROMÃO, W. **Descoberta de Conhecimento relevante em Banco de Dados sobre Ciência e Tecnologia**. Tese de D. Sc., Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina – UFSC, Florianópolis, SC, Brasil, 2002. Disponível em: http://www.din.uem.br/~intersul/intersul_arquivos/documentos/Tese%20Wesley.pdf. Acesso em: 14 dez. 2015.

SANTOS, W. B; COLOSIMO, E. A; MOTTA, S. B. Tempo ótimo entre manutenções preventivas para sistemas sujeitos a mais de um tipo de evento aleatório. **Revista Gestão e Produção**, v. 14, n. 1, p. 193-202, 2007.

SILVA, F. M.; STOPANOVSKI, M.; ROCHA, H.; COSAC, D. **Cartão de Pagamento do Governo Federal: uma Análise de Regras de Associação**. In: Workshop Franco-Brasileiro sobre Mineração de Dados, 2009, Recife/PE. Caderno de Resumos. Recife/PE, 2009. p. 40.

SOUZA, L. C. A.; SEIXAS FILHO, C.; PENA, R. T. **Padrão de Acesso a Dados OPC e sua Implementação em um driver OPC-Modbus**. In: II Congresso Mineiro de Automação, V Simpósio Regional de Instrumentação da ISA-BH / GRINST-MG, 1998, Belo Horizonte. Livro de Anais - II Congresso Mineiro de Automação, V Simpósio Regional de Instrumentação da ISA-BH / GRINST-MG, 1998. p. 157-164.

TAN, P. N.; STEINBACH, M.; KUMAR, V. **Introdução ao Data Mining - Mineração de Dados**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2009.

TARAPANOFF, K. **Inteligência Organizacional e Competitiva**. Editora Universidade de Brasília, Brasília, DF, Brasil, 2001. 344 p.

TAVARES, L. A. **Administração Moderna da Manutenção**. Rio de Janeiro, Editora NAT. 1999.

UNIVERSITY OF WAIKATO. **Weka3** – Machine Learning Software in Java. 2010. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka2>>. Acesso em 25 ago. 2015.

WANG, X. Z. **Data Mining and Knowledge Discovery for Process Monitoring and Control**. Advances in Industrial Control, Springer, London, 1999.

ANEXO A: ARTIGO DA REVISTA ESPACIOS – Desenvolvimento de um sistema de monitoramento e gestão da manutenção industrial

1 INTRODUÇÃO

Nos últimos anos vivemos diante de um mercado repleto de incertezas econômicas, escassez financeira entre outras dificuldades que obrigaram as empresas a adotarem uma política mais rigorosa com relação à redução de custos, qualidade no processo e principalmente competitividade.

Os meios de produção e os recursos disponíveis precisam de reformulação com o objetivo de reduzir ou eliminar os desperdícios existentes na empresa. A adequação ambiental e o retorno financeiro exigem a otimização dos processos existentes e a gestão aprimorada da manutenção dos equipamentos e sistemas atuais.

A manutenção industrial, muito importante na Siderurgia, não deve ser dissociada da gestão de operação ou da qualidade. Segundo a ABRAMAN – Associação Brasileira de Manutenção: [...] precisamos seguir os mesmos requisitos exigidos da produção: desenvolvimento econômico, preservação ambiental e responsabilidade social. Diretamente ligada à redução de custos, a Manutenção firmou-se como elemento essencial à garantia de qualidade e da sustentabilidade dos processos produtivos [...]. (Revista ABRAMAN. São Paulo: n.126, Set. 2010).

Este trabalho faz um levantamento das principais ocorrências da gerência de manutenção de sistemas industriais (instrumentação e automação) com foco na prevenção das falhas que ocasionam paradas de produção, questões ambientais ou risco de acidente no ambiente de trabalho.

A redução do processo produtivo proveniente de uma parada inesperada em uma determinada linha de produção é um dos principais problemas enfrentados por empresas dos mais diversos setores da indústria. Fica evidente que consertar uma máquina após sua quebra não é viável, é mais sensato evitar que isto aconteça fazendo as inspeções preventivas ou preditivas (KARDEC 2001).

Com as técnicas e ferramentas presentes na empresa, o sistema desenvolvido faz a monitoração dos principais equipamentos de controle e automação, essenciais para a operação da planta. Estas informações são armazenadas em banco de dados e viram inteligência empresarial por ajudar na estatística das principais ocorrências. Muassab (2002) afirma que a Engenharia e a Gestão da Manutenção ocupam posições importantes no plano estratégico,

tendo inerente participação nas melhorias no processo produtivo, por manter os equipamentos em perfeita condição de funcionamento.

Neste trabalho é desenvolvido um sistema de monitoração dos equipamentos de automação industrial, classificados como imprescindíveis para a continuidade de produção na empresa. A monitoração é realizada através de uma rede TCP/IP com uma aplicação Dot Net que testa a conectividade de cada um. A interface utiliza o *ProcessBook* que cria as animações de funcionamento e dispara mensagens em caso de perda de comunicação de algum equipamento na rede. Com o uso do BI (*Business Intelligence*), serão definidos indicadores de desempenho de manutenção, e apresentados de forma aplicada, como resultados obtidos, a partir de um estudo de caso.

Por fim, será discutida a integração deste sistema de automação com o SAP, sistema de gestão da empresa, através do SAP XI (*Exchange Infrastructure*) com a finalidade de abrir notas de inspeção e manutenção de modo automático e alocar as mãos de obra para execução do serviço preventivo ou corretivo adequadamente. E ainda notificar as pessoas responsáveis pelas ocorrências de quebra ou instabilidade, usando o serviço de e-mail ou mensagem de texto para aparelhos celulares.

2 METODOLOGIA

Além da apresentação deste sistema, a gestão da manutenção em sistemas industriais recorre a indicadores específicos de manutenção que traçam as metas de desempenho, o que permite comparar resultados, identificar tendências e definir um plano de ação para as ocorrências.

Segundo Moubray (2000), a RCM (*Reliability Centred Maintenance*) trata a manutenção por meio de um estudo de confiabilidade de cada sistema, trazendo para esta função, um tratamento mais científico e baseado em indicadores. Estes indicadores geralmente estão relacionados com tempo (de funcionamento, reparações, etc.) e custos (de manutenção preventiva, corretiva, melhoria, etc.).

Os principais indicadores teóricos de manutenção são:

Nome	Fórmula	Dados
Taxa de Avarias	$TA = \frac{n^{\circ} \text{avarias}}{TF} * 1000$ $= \frac{n^{\circ} \text{avarias}}{Km} * 10000$	TF – Tempo total de funcionamento no período
MTBF	$MTBF = \sum \frac{TFi}{n^{\circ} \text{avarias}}$	TFi – Tempos de funcionamento no período
MTTR	$MTTR = \sum \frac{TRi}{n^{\circ} \text{avarias}}$	TRi – Tempos utilizados nas reparações no período
MTW	$MTW = \sum \frac{TEi}{n^{\circ} \text{avarias}}$	TEi – Tempo de espera no período
Disponibilidade	$D = \frac{MTBF}{MTBF + MTTR + MWT}$	MTBF – Tempo médio entre avarias MTTR – Tempo médio entre avarias MWT – Tempo médio de espera

Tabela 1: Indicadores teóricos da manutenção e suas respectivas fórmulas.

Fonte: Cabral J. (2006)

A Taxa de Avarias (λ) representa o número de avarias por unidade de utilização, ou seja, quantas quebras um equipamento sofreu em um determinado tempo.

O MTBF (*Mean Time Between Failures*) indica o tempo médio de bom funcionamento, ou ainda, o tempo médio entre avarias. É um parâmetro de fiabilidade, cujo objetivo principal é que o equipamento dure o máximo de tempo possível sem falhas (ASSIS 2004).

O MTTR (*Mean Time to Repair*) indica o tempo necessário de recuperação do sistema em caso de avaria. Coloca ênfase num equipamento que em caso de avaria, seja reparado o mais rapidamente possível (ASSIS 2004).

O indicador MTW (*Mean Time Waiting*) representa a eficiência de resposta de um departamento às solicitações de manutenção. A fórmula efetua uma média dos tempos de espera de atendimento de pedidos de reparação de avarias num determinado período.

Já o indicador disponibilidade retrata a percentagem de bom funcionamento do equipamento em relação ao tempo total. A norma europeia (European Standard EN 13306, 2001), define “disponibilidade como a aptidão de um bem para estar em estado de cumprir uma função requerida em condições determinadas, num dado instante ou em determinado intervalo de tempo, assumindo que é assegurado o fornecimento dos necessários meios externos”.

Os indicadores de confiabilidade podem colaborar no planejamento da programação das paradas e evitar que determinados equipamentos fiquem indisponíveis durante ciclos produtivos. Entre os principais, estão: Taxa de Avarias, MTBF – *Mean Time Between Failure* (Tempo Médio Entre Falhas), MTTR – *Mean Time To Recovery* (Tempo Médio Para Reparo) e Disponibilidade (TAVARES, 1999). A Figura 1 representa a tela de monitoramento dos principais equipamentos da gerência de manutenção de sistemas industriais.

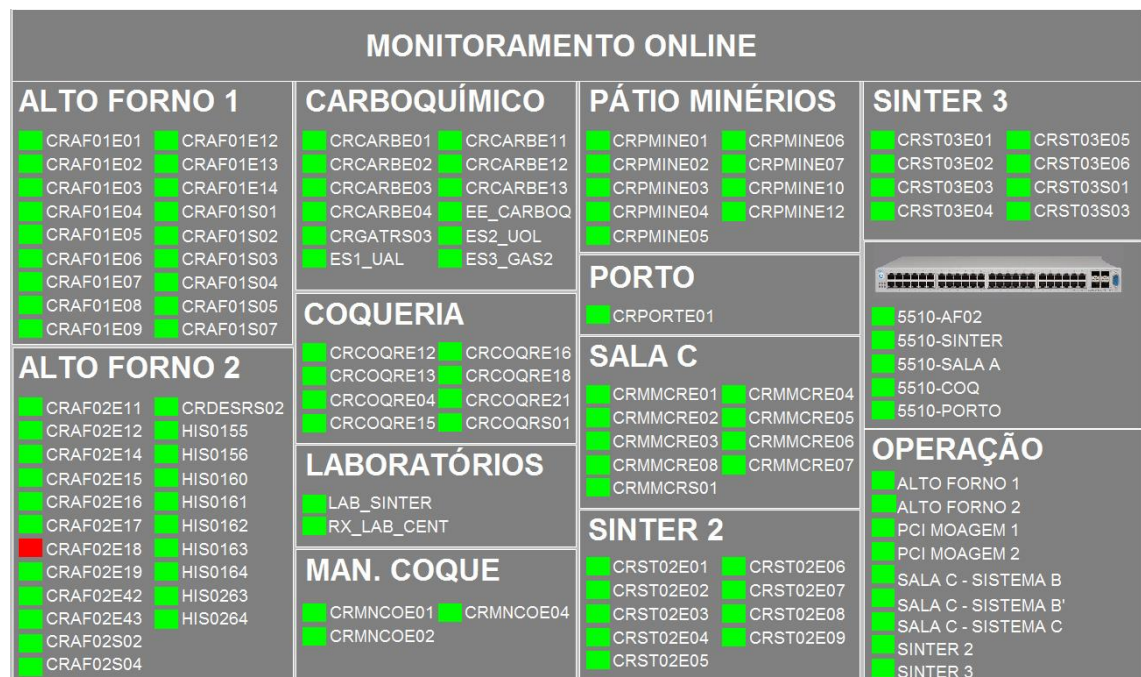


Figura 1: Tela de monitoramento online dos equipamentos de automação industrial

3 GESTÃO DE MANUTENÇÃO

O sistema de monitoramento responsável pelo teste de conectividade com os equipamentos de automação industrial foi desenvolvido em plataforma Dot Net, usando a linguagem de programação C#. Ele utiliza a classe PING entre outras que são consumidas do *Framework* .Net 4.5 e os resultados deste teste de conectividade são armazenadas em um banco de dados MS SQL Server 2008.

Desta forma, criamos uma lista de equipamentos considerados essenciais que podem ocasionar parada operacional ou quebra do ritmo de produção. Quando algum equipamento - servidor ou estação de operação - apresentar falha de comunicação na rede, o sistema modifica o status para “falso” no banco de dados, indicando que aquele equipamento não está disponível na rede, ou seja, sem comunicação.

Assim possuímos um histórico das paradas de equipamentos com data/hora das ocorrências que podem servir de base na criação dos índices de manutenção apresentados anteriormente: número de avarias, tempo médio de bom funcionamento e tempo necessário de recuperação do sistema. Também auxiliam na análise e relação com outras variáveis como, por exemplo, quais equipamentos tiveram maior influência na quebra do ritmo de produção ou contribuíram diretamente em parada operacional.

Sistema de monitoramento

A classe *ping* permite determinar se um computador ou equipamento remoto está acessível pela rede. A Figura 2 exibe o método *PingCompletedEventHandler* que manipulará o evento *PingCompleted*. Assim é enviado um ICMP *echo request* para o host remoto e espera por um ICMP *echo reply*. O resultado *Success* indica que há conectividade normal com o equipamento remoto.

```
using System.Net.NetworkInformation;

public void Ping()
{
    AutoResetEvent waiter;

    byte[] buffer;
    int timeout;
    PingOptions options;
    PingReply reply;
    Ping pingSender = new Ping();

    pingSender.PingCompleted += new
        PingCompletedEventHandler(PingCompletedCallback);

    timeout = 300;
    options = new PingOptions(64, true);

    pingSender.SendAsync(who, timeout, buffer, options, waiter);
}

private void PingCompletedCallback(object sender, PingCompletedEventArgs e)
{
    Int16 Status;
    string Address;

    reply = e.Reply;
    Status = reply.Status.ToString();
    Address = reply.Address.ToString();
}
}
```

Figura 2: Uso das classes de teste de conectividade dos equipamentos

A interface da aplicação, conforme Figura 3, está construída em *Windows Forms* (Visual Studio 2008) e apresenta um resumo dos testes de conectividade na rede. Lembrando que existe uma tela de monitoramento para exibir todos os equipamentos que estão sendo supervisionados.

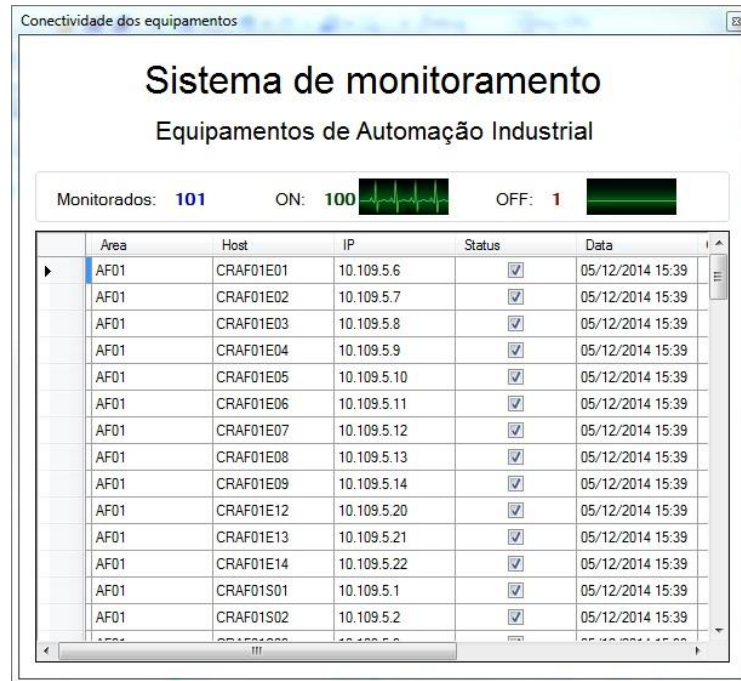


Figura 3: Tela da aplicação que efetua os testes de conectividade de rede

Banco de dados

O SGBD (Sistema Gerenciador de Banco de Dados) utilizado foi o MS SQL Server 2008 com as tabelas listadas na Figura 4. A tabela principal do sistema de monitoramento é a *tbl_monitora Equipamento* que armazena informações de início/fim da perda de comunicação e os códigos do tipo da falha, natureza (responsável pela falha), equipamento e área de instalação (local).

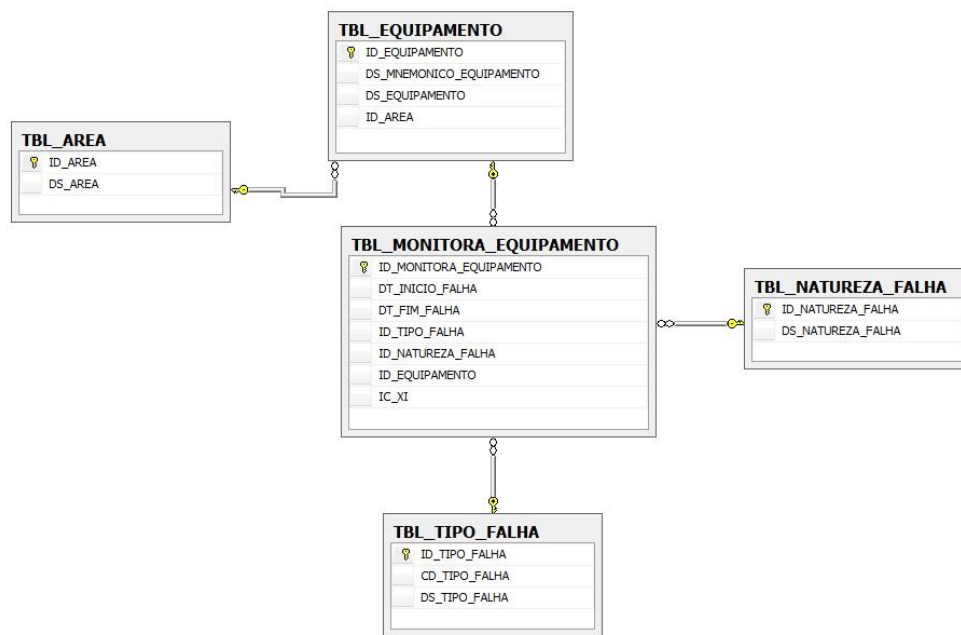


Figura 4: Diagrama das tabelas utilizadas pelo sistema de monitoramento

Interface

As ferramentas da OSI - *PI ProcessBook* e *PI System Explorer* foram imprescindíveis na construção da tela gráfica de monitoramento *online*. No *PI System Explorer* (Figura 5) pode ser criado os elementos de acesso ao banco de dados com a periodicidade de leitura configurável, que para este sistema, a cada segundo está verificando se houve mudanças nos *bits* de *status* de funcionamento dos equipamentos.

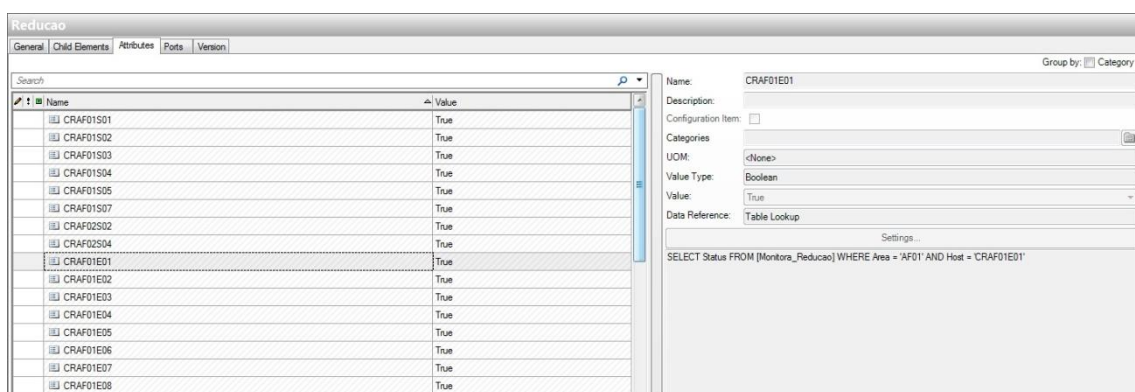


Figura 5: Criação dos elementos de verificação do status de funcionamento

Com os elementos criados, o *PI ProcessBook* (Figura 6) - aplicação que permite criar e visualizar telas de processos para acompanhamento em tempo real, possui um recurso de animação de objetos e assim determinamos as cores mediante o *bit* de *status* de funcionamento dos equipamentos:

- Falso (*bit* sinalizado em zero) = cor vermelha;

- Verdadeiro (*bit* sinalizado em um) = cor verde.

Server: 10.109.1.221 Tag Search...

Tag: E.CRAF01E01 Custom Placeholders

State Info

Number of States: 2

Color for Bad Data: [Yellow] [Blink]

State: 1 Values: >= 0 <= 0 Color: [Red] [Blink]

[OK] [Cancel] [Help]

Convert To Static

Figura 6: Configuração dos objetos com animação

Notificação de falha

Outra funcionalidade importante neste sistema é a notificação de falhas a um grupo de usuários da empresa por *e-mail*. A ferramenta da OSI - PI *Notifications* realiza este envio baseado na configuração de elementos que podem ser o *status* de funcionamento dos equipamentos, condições de processo da área, etc.

Os elementos criados anteriormente verificam o *status* de funcionamento dos equipamentos como condição para determinar o disparo de um *e-mail* informando a avaria. Com a criação dos grupos e a atribuição dos usuários, podemos notificar o técnico, engenheiro e até os níveis gerenciais sobre uma ocorrência em seu departamento. A Figura 7 demonstra a ferramenta de notificação.

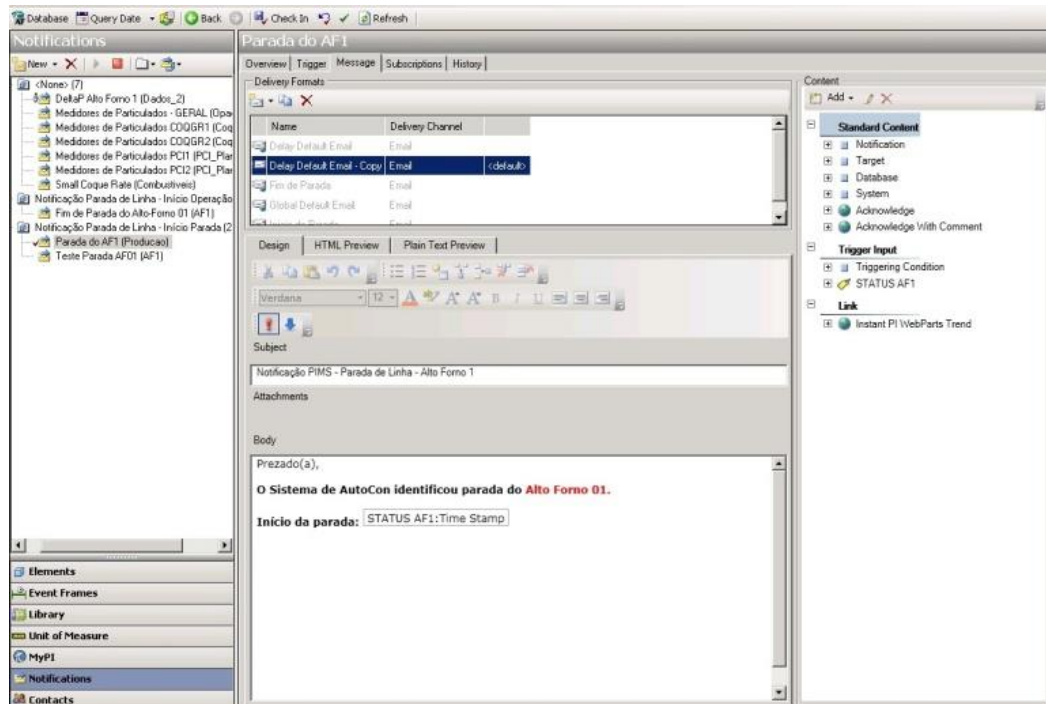


Figura 7: Ferramenta de notificação por e-mail

Abertura de Notas de Manutenção

O sistema de monitoramento dos equipamentos está integrado com o SAP, sistema de gestão empresarial utilizado na organização, através do SAP XI (*Exchange Infrastructure*), que realiza a troca de informações entre o sistema interno e o *software* terceiro. A interação entre os diferentes sistemas operacionais dos aplicativos é facilitada por conta do SAP *Integration Server*, componente central do SAP XI. A família SAP *Business Suite* é baseada na plataforma SAP *NetWeaver* que oferece tecnologias como adaptadores de aplicativos e protocolos para integração de aplicação SAP e não SAP.

A figura 8 representa a estrutura do SAP *NetWeaver* XI e seus principais componentes. Nota-se que destacado ao lado direito, o SAP *NetWeaver* XI faz parte do processo de integração “*Integration of SAP and non-SAP components*”.

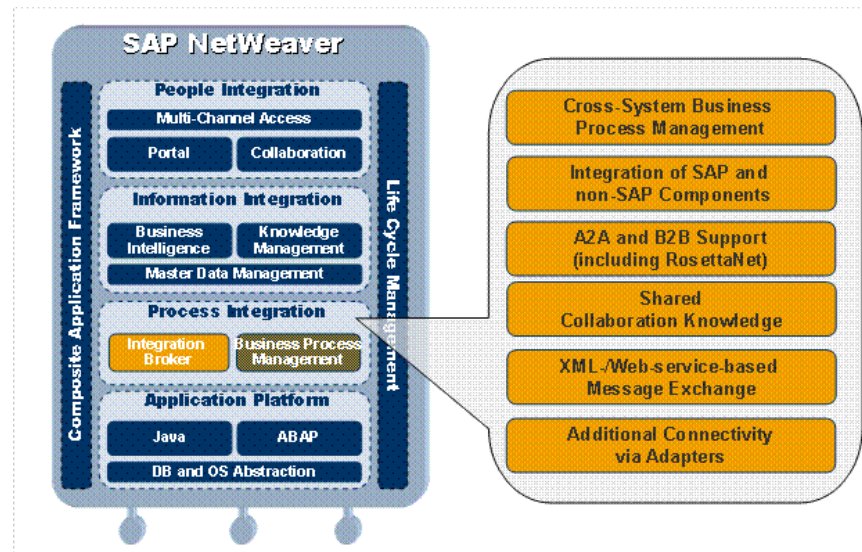


Figura 8: Arquitetura SAP NetWeaver.
Fonte: SAP (2014)

Esta integração torna possível a abertura das notas M1 (Manutenção Planejada), M2 (Nota de Avaria) e M3 (Manutenção Preventiva) por meio do parâmetro NOTIFICATION_TYPE no SAP. Em caso de avaria nos equipamentos, a abertura da nota de manutenção M2 é feita automaticamente, atribuindo uma ordem de serviço ao responsável para execução da mesma.

A realização da inspeção dos equipamentos em datas específicas, com o intuito de manutenção preventiva, é feita mediante a abertura de uma nota M3 ao respectivo “dono do equipamento” que deseja prevenir a avaria do ativo, e executa o serviço de acordo com o plano de inspeção.

Quando o serviço depende de áreas terceiras, como por exemplo, equipamentos que não podem ser desligados com a área de produção em operação, pois os operadores dependem daquele sistema para controlar a fábrica, a abertura das notas está vinculada com o plano de produção da empresa. Ou seja, o serviço é realizado em uma parada programada da área operacional em que o equipamento está instalado.

4 RESULTADOS OBTIDOS

No trabalho apresentado foram utilizadas ferramentas de *software*, serviços e junto ao desenvolvimento de lógicas de programação e manipulação de banco de dados, propiciaram um *dashboard* de manutenção com uma indicação de funcionamento dos principais equipamentos de automação industrial na área de Redução.

Ao utilizar a comunicação com o SAP e a notificação de falhas por *e-mail*, utilizando o serviço *PI Notifications*, pode-se programar a abertura de notas de manutenção do tipo M1, M2 e M3 de modo automático, sempre que houver uma avaria de equipamento ou em caso de manutenção programada. Assim, o responsável pelo equipamento recebe um *e-mail* de notificação de falha e uma nota de manutenção já é atribuída em seu nome para execução do serviço. A Figura 9 demonstra a nota de manutenção no SAP.

Exibir nota PM: Nota de Ação

Nota: 17395270 M3 ADM-RD-AF02-NIVEL2-REMOTO

Status da nota: MSPN

Objeto de referência:

Loc. instalação: CS01-RD-AF002-AAE... Sistema de Automação do AF02

Equipamento:

Conjunto:

Responsabilidades:

Grp.plnj.PM: M19 / CS01

CenTrab respon.: I22I-LMA / CS01 Leonardo Matsumota

Departam.respon: 60408813 Manutenção Sistemas Industriais de

Usuário respon:

Notificador: LEONARDO Data da nota: 18.01.2014 16:00:00

Situação:

Codificação: PMMANUT AUTH MANUT EM EQUIPAMENTO AUTOMAÇÃO PROCE...

Descrição: ADM-RD-AF02-NIVEL2-REMOTO

29.01.2014 10:21:23 Leonardo Matsumota (CC67585)

Descrição: Problema na estação de operação da Granulação CRAF02E18
 Problema: A estação perdeu o HD e foi instalado o Supervisório, Drivers e Programas de Nível 2 na estação CRAF02E15
 Envolvidos: Leonardo Matsumota.
 Horário: 18/01/2014 das 16:00 as 18:00

Figura 9: Nota de manutenção no SAP

Com a aquisição dos dados de início e término das falhas de cada equipamento, obtivemos um histórico de informações que capacitaram à formação dos indicadores de manutenção: Taxa de avarias, MTTR e MTBF. Assim, foram identificadas as maiores reincidências, o tempo de atendimento para resolução da falha em horários administrativos e de turno, quanto tempo determinados componentes ou equipamentos permanecem em funcionamento, etc. A Figura 10 representa os indicadores de manutenção por equipamento.

ID_EQUIPAMENTO	DS_MNEMONICO_EQUIPAMENTO	Taxa_Avaria	MTTR	MTBF
1	CRAF01E01	0,33	7,78	2359,35
2	CRAF01E02	0,46	10,97	2356,16
3	CRAF01E03	0,3	7,16	2359,97
4	CRAF01E04	0,45	10,75	2356,38
5	CRAF01E05	0,68	16,17	2350,95
6	CRAF01E06	0,21	5,01	2362,12
7	CRAF01E07	0,04	0,95	2366,17
8	CRAF01E08	0,5	11,76	2355,37
9	CRAF01E09	0,76	17,93	2349,2
10	CRAF01E12	0,09	2,22	2364,91
11	CRAF01E13	0,37	8,69	2358,44
12	CRAF01E14	0,18	4,16	2362,97
13	CRAF01S01	0,35	8,25	2358,88
14	CRAF01S02	0,26	6,06	2361,07
15	CRAF01S03	0,53	12,47	2354,66
16	CRAF01S04	0,52	12,25	2354,88
17	CRAF01S05	0,19	4,5	2362,62
18	CRAF01S07	0,26	6,12	2361,01

Figura 10: Indicadores de manutenção por Equipamento

Para atender as fronteiras de serviço determinadas na empresa, onde cada equipe de manutenção é responsável por atender equipamentos de áreas de produção diferentes, as consultas também fazem o agrupamento dos índices de manutenção por Área de Produção, conforme Figura 11.

Area_Producao	Taxa_Avaria	MTTR	MTBF
Alto Forno 1	6,47	153,19	2213,94
Alto Forno 2	7,71	182,46	2184,67

Figura 11: Indicadores de manutenção por Área de Produção

5 CONSIDERAÇÕES FINAIS

A principal contribuição deste estudo está relacionada à necessidade de alta disponibilidade dos equipamentos de automação industrial na usina siderúrgica. A utilização de ferramentas de diferentes fornecedores e a integridade dos dados foram grandes desafios para integrar o sistema de monitoramento como um todo.

Como resultado, o sistema disponibiliza o *status* de funcionamento de cada equipamento e prevê uma falha de *hardware* ou *software* do sistema de automação que possa causar uma parada operacional. Logo após detectar a falha, o sistema identificação responsável pelo equipamento e abre uma ordem de serviço para atuação imediata.

A integração com o ERP permitiu a identificação e carteira de serviços dos funcionários da área de manutenção da empresa. Além disso, os supervisores, gerentes e

diretores podem ser notificados por *e-mail* ou mensagem de texto quando a responsabilidade de reparo do equipamento pertence ao seu departamento.

Os indicadores de MTBF, MTTR e a Taxa de Avaria foram adotadas para gerenciar as ocorrências em cada área de produção e auxiliar na tomada de decisão com relação à necessidade de investimentos e aquisição de recursos.

REFERÊNCIAS

ABRAMAN, editorial. Revista da Associação Brasileira de Manutenção, São Paulo: n.126, Set, 2010, p.06.

ASSIS R. Apoio à Decisão em Gestão da Manutenção – Fiabilidade e Manutenibilidade. Editora Lidel, 1ª edição, 2004.

CABRAL J. Organização e Gestão da Manutenção, dos conceitos à prática. Editora Lidel, 6ª edição, 2006.

EUROPEAN STANDARD EN 13306 - Maintenance terminology. Brussels: European committee for standardization. CEN, 2001.

KARDEC, A; NASCIF, J. (2001), Manutenção: Função Estratégica. Rio de Janeiro: Editora Qualitymark, 2.^a Edição.

MUASSAB, J. R. (2002), Gerenciamento da Manutenção na Indústria Automobilística, Taubaté. Monografia – Universidade de Taubaté.

MOUBRAY, J. (2000), Manutenção Centrada em Confiabilidade (*Reliability-centred Maintenance*). United Kingdom, Biddles Ltd.

SAP NetWeaver Exchange Infrastructure. SAP. Disponível em: <http://help.sap.com/saphelp_nw70/helpdata/EN/0f/80243b4a66ae0ce10000000a11402f/frameet.htm>. Acesso em: 16 dez. 2014.

TAVARES, L. A. (1999), Administração Moderna da Manutenção. Rio de Janeiro, Editora NAT.