

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MÁRCIO JOSÉ DA SILVA

**MODELAGEM SEMÂNTICA DE CONTEXTO APLICADA EM
UM HISTÓRICO DE ALARMES DE PROCESSO**

Porto Alegre

2016

MÁRCIO JOSÉ DA SILVA

**MODELAGEM SEMÂNTICA DE CONTEXTO APLICADA EM
UM HISTÓRICO DE ALARMES DE PROCESSO**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Controle e Automação

ORIENTADOR: Prof. Dr. Carlos Eduardo Pereira

Porto Alegre

2016

MÁRCIO JOSÉ DA SILVA

MODELAGEM SEMÂNTICA DE CONTEXTO APLICADA EM UM HISTÓRICO DE ALARMES DE PROCESSO

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Carlos Eduardo Pereira, UFRGS

Doutor pela Universidade de Stuttgart, Stuttgart, Alemanha

Banca Examinadora:

Prof. Dr. José Palazzo Moreira de Oliveira, PPGC UFRGS

Doutor pelo Instituto Politécnico de Grenoble – Grenoble, França

Prof. Dr. João César Netto, PPGC UFRGS

Doutor pela Universidade Católica de Louvain – Louvain, Bélgica

Prof. Dr. Edison Pignaton de Freitas, PPGEE UFRGS

Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Coordenador do PPGEE: _____

Prof. Dr. Valner João Brusamarello

Porto Alegre, novembro de 2016

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Aos Professores Carlos Eduardo Pereira e Marcelo Götz pela orientação e ensinamentos durante este trabalho.

Aos familiares, amigos e colegas que acompanharam o desenvolvimento deste trabalho.

RESUMO

Atualmente, os avanços tecnológicos, principalmente nas áreas de controle e automação, facilitam a inclusão de alarmes em sistemas de supervisão de plantas industriais. É possível incluir um número quase que ilimitado de alarmes com variação de tipos para cada ponto de medição de um processo. Consequentemente, o volume de informações cresce significativamente e isso pode ser prejudicial, uma vez que limita a habilidade do operador no gerenciamento de anomalias e pode exceder sua capacidade de realizar ações eficazes durante o funcionamento do processo. Este trabalho apresenta um estudo sobre modelagem semântica de contexto e utiliza uma base histórica de informações de eventos para análise de padrões. Dessa forma, o intuito é, por meio dos dados de contexto, obter conhecimento útil para inferência e determinação da situação. Uma aplicação real onde são investigados eventos ocorridos em uma planta de uma usina térmica de geração de energia elétrica é usado como estudo de caso para aplicar as ideias desenvolvidas bem como para validar a proposta. Decorrente desse estudo, é proposta uma ontologia de domínio específico implementada a partir de um modelo semântico de contexto. Por fim, é apresentada uma implementação de regras semânticas.

Palavras-chave: Alarmes, Ontologia, Padrões, Regras Semânticas.

ABSTRACT

Nowadays, technological advance, especially in the areas of control and automation, make it easy alarm inclusion in supervision of industrial plant systems. You can include a number almost unlimited of alarms with different types for each measurement point of a process. Consequently, the volume of information grows significantly and this can be harmful since it limits the ability of the operator in managing anomalies and may exceed its ability to carry out effective actions during operation of the process. This work presents a study of semantic modeling of context and uses historical bases event information to identify patterns. Thus, the intention is use this context data to obtain useful knowledge for inference and define the situation. A real application where events of a thermal power plant for electricity generation are investigated is used as a case study to apply the ideas developed and to validate the proposal. As a Result of this study, it is proposed a specific domain ontology implemented from a semantic model of context. Finally, it is presented an implementation of semantic rules.

Keywords: Alarms, Ontology, Standards, Semantic Rules.

SUMÁRIO

1	INTRODUÇÃO	11
1.1	MOTIVAÇÃO	16
1.2	OBJETIVO	16
1.3	ORGANIZAÇÃO DO TEXTO	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	SISTEMAS DE ALARMES	18
2.2	GUIAS E NORMAS	20
2.3	MINERAÇÃO DE DADOS	25
2.3.1	Dados, informação e conhecimento	25
2.3.2	O Processo de Descoberta de conhecimento	27
2.4	ANÁLISE DE CONTEXTO	30
2.4.1	Dados de Contexto	32
2.4.2	Geração de Modelos de Contexto	33
2.5	CONCEITOS DE ONTOLOGIA	34
3	TRABALHOS RELACIONADOS	38
4	METODOLOGIA PROPOSTA	49
4.1	DESCRIÇÃO DA PROPOSTA	49
4.2	CONTEXTOS DE INTERESSE	55
4.3	EVENTOS AVALIADOS	55
5	MODELAGEM CONCEITUAL DE CONTEXTO	58
5.1	DESENVOLVIMENTO DO MODELO CONCEITUAL	58
5.2	ANÁLISE DO MODELO CONCEITUAL	60
5.3	AGREGAÇÃO DO CONTEXTO	61
6	IMPLEMENTAÇÃO DA PROPOSTA	65
6.1	DESCRIÇÃO DO ESTUDO DE CASO	65
6.2	FERRAMENTAS PARA IMPLEMENTAÇÃO	66
6.3	MINERAÇÃO DE PROCESSOS	68
6.4	IMPLEMENTAÇÃO DO MODELO	70
7	APLICAÇÃO DA METODOLOGIA E ESTUDO DE CASO	76
7.1	VALIDAÇÃO EXPERIMENTAL	76
7.2	SITUAÇÕES DE INTERESSE	83
7.2.1	Estudo de Caso 1	84
7.2.2	Estudo de Caso 2	89
7.2.3	Estudo de Caso 3	92
7.3	CONCLUSÕES	93
8	CONSIDERAÇÕES FINAIS	95
8.1	TRABALHOS FUTUROS	96
8.2	TRABALHO COMPLETO PUBLICADO EM CONFERÊNCIA	96
	REFERÊNCIAS	98

LISTA DE ILUSTRAÇÕES

Figura 1 Aumento do número de alarmes controlados por operador ao longo dos anos. Fonte: Traduzido de (HABIBI e HOLLIFIELD, 2006).....	11
Figura 2 Distribuição dos motivos de falhas nas indústrias. Adaptado de (HONEYWELL, 2004).....	12
Figura 3 Níveis de um sistema de automação. Fonte: Adaptado de (SOUZA <i>et al.</i> , 2005).....	13
Figura 4 Visão geral de um sistema de alarmes de processo genérico. Fonte: o autor.	19
Figura 5 Ciclo de vida do gerenciamento de alarmes. Traduzido de (ANSI/ISA-18.2-2016). 23	
Figura 6 Representação do processo KDD. Traduzido de (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).	28
Figura 7 Ontologias para diagnóstico de falha em um sistema de energia. Adaptado de (BERNARAS <i>et al.</i> , 1996).	38
Figura 8 Ontologias para metodologia proposta. Fonte (AIZPURÚA, GALÁN e JIMÉNEZ, 2008).....	40
Figura 9 Ciclo de vida dos alarmes de processo. Fonte (LEITÃO, GUEDES e ARAUJO, 2012).	43
Figura 10 Classes principais da ontologia para planta industrial. Fonte (LIMA <i>et al.</i> , 2013). 44	
Figura 11 Representação do ambiente, situação, contexto, dados de contexto, sensor e relações semânticas. Fonte (MACHADO <i>et al.</i> , 2013).	45
Figura 12 Visão geral das classes do modelo. Fonte (MACHADO e OLIVEIRA, 2014).....	46
Figura 13 Visão geral da abordagem desenvolvida. Fonte: o autor.	50
Figura 14 Estrutura do processo. Fonte: O autor.....	51
Figura 15 Representação dos diferentes contextos e situações presentes na planta. Fonte: O autor.	56
Figura 16 Visão parcial do modelo ontológico conceitual desenvolvido. Fonte: O autor.	58
Figura 17 Visão parcial do modelo desenvolvido. Fonte: O autor.....	62
Figura 18 Visão parcial do modelo desenvolvido. Fonte: O autor.....	62
Figura 19 Visão geral do modelo desenvolvido. Fonte: O autor.....	63
Figura 20 Credenciais para acesso à base de dados. Fonte: Autor.....	67
Figura 21 Mapeamento e importação das instâncias. Fonte: Autor.....	68
Figura 22 Registro dos eventos. Fonte: O autor.....	69
Figura 23 Registro dos eventos. Fonte: O autor.....	69
Figura 24 Visão geral da ontologia elaborada pelo plug-in OWL Viz do Protégé.....	71
Figura 25 Tela principal do ProM6.....	77
Figura 26 Tela principal do ProM6.....	77
Figura 27 Seleção dos compostos referentes a unidade de geração 4. Fonte: O autor.....	78
Figura 28 Visão geral do filtro de dados ProM. Fonte: O autor.....	79
Figura 29 Visão geral do filtro de dados ProM. Fonte: O autor.....	80
Figura 30 Visão geral do filtro de dados ProM. Fonte: O autor.....	81
Figura 31 Visão geral do <i>plug-in Dotted Chart</i> ProM. Fonte: O autor.....	81
Figura 32 Visão geral do <i>plug-in Dotted Chart</i> ProM. Fonte: O autor.....	82

Figura 33	Visão parcial do <i>plug-in Show Sequence and Patterns</i> no ProM. Fonte: O autor. ...	83
Figura 34	Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.	84
Figura 35	Visão parcial do <i>plug-in Show Sequence and Patterns</i> . Fonte: O autor.	85
Figura 36	Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação atual. Fonte: O autor.	86
Figura 37	Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.	87
Figura 38	Visão geral do <i>plug-in Dotted Chart</i> ProM. Fonte: O autor.	88
Figura 39	Visão parcial do <i>plug-in Show Sequence and Patterns</i> . Fonte: O autor.	89
Figura 40	Visão parcial do <i>plug-in Show Sequence and Patterns</i> . Fonte: O autor.	90
Figura 41	Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.	91
Figura 42	Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido. Fonte: O autor.	93

LISTA DE TABELAS

Tabela 1 índice de desempenho EEMUA 191.....	22
Tabela 2 índice de desempenho ISA 18.2.	25
Tabela 3 Análise das técnicas para modelagem de contexto.....	32
Tabela 4 Resumo dos Trabalhos Relacionados.	47
Tabela 5 Interpretação das condições.....	73
Tabela 6 Tabela das regras definidas no estudo de caso 3.	92

LISTA DE ABREVIATURAS

CML: Context Modelling Language

CSV: Comma-Separated Value

DCS: Distributed Control System

EEMUA: Engineering Equipment and Materials Users Association

ESD: Emergency Shutdown

HSE: Health and Safety Executive

ISA: International Society of Automation

JESS: Java Expert System Shell

KDD: Knowledge Discovered in Database

MEBN: Multi-Entity Bayesian Network

OWL: Ontology Web Language

OWL-DL: Ontology Web Language Description Language

PLC: Programming Logic Controllable -Controller

PR-OWL: Probabilistic OWL

SCADA: Supervisory Control and Data Acquisition

SWRL: Semantic Web Rule Language

UFRGS: Universidade Federal do Rio Grande do Sul

W3C: World Wide Web Consortium

1 INTRODUÇÃO

Os avanços na automatização de processos industriais possibilitaram inúmeros ganhos em diversos setores da economia. Diversas soluções que permitem monitorar de forma mais eficiente a produção acompanharam esse avanço. Contudo, segundo (HABIBI e HOLLIFIELD, 2006), a indicação das condições anormais do processo e as formas de acesso às informações oferecidas por esses sistemas não são ideais.

A tecnologia tornou possível coletar, produzir e analisar em um ambiente digital, um volume significativo de dados e informações. Porém, esse volume dificulta a localização e a seleção de informações relevantes que possam auxiliar na tomada de decisão (KORDIC *et al.*, 2007).

A Figura 1 mostra o aumento do número de alarmes usualmente controlados por um operador em uma planta industrial nos últimos anos. Um aumento significativo no número de alarmes controlados pelo operador é observado a partir da década de 70. Segundo o trabalho de (HABIBI e HOLLIFIELD, 2006), isso se deve ao surgimento dos sistemas de controle distribuído (do inglês *Distributed Control System* (DCS) presentes nas plantas industriais.

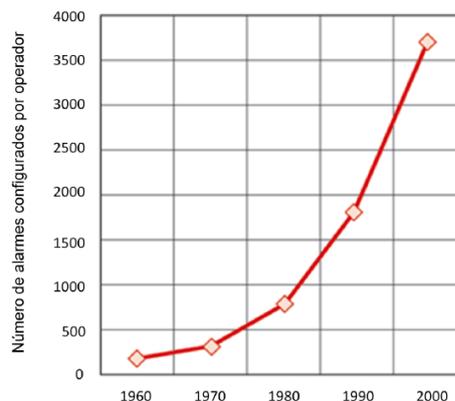


Figura 1 Aumento do número de alarmes controlados por operador ao longo dos anos. Fonte: Traduzido de (HABIBI e HOLLIFIELD, 2006).

Diagnosticar um alarme pode tornar-se uma tarefa difícil em função da grande quantidade de variáveis envolvidas, aliado ao fato que muitas vezes as informações disponíveis são incompletas e insuficientes. Alarmes são configurados e habilitados com o intuito de tornar mais simples a identificação de anomalias nos processos e equipamentos, embora nem sempre passem por uma análise mais aprofundada.

Os atuais sistemas eletrônicos permitem programar com grande facilidade novos alarmes. A consequência disso é um acréscimo de inúmeras informações desnecessárias para análise pelo operador.

De acordo com (LEITÃO *et al*, 2008), um olhar mais atento para a segurança e efetividade dos sistemas de alarmes de processo somente ocorreu após acidentes marcantes como o ocorrido na refinaria Milford Haven da Texaco em 1994, quando surgiram guias e normas específicos para esse tema. Entretanto, a observação de normas por si só não garante que um sistema esteja seguro. Durante situações de distúrbios na planta industrial uma grande quantidade de alarmes é gerada e muitas vezes os responsáveis pelo processo realizam ações que tornam o problema ainda maior. A Figura 2 apresenta a distribuição dos motivos de falhas nas indústrias (HONEYWELL, 2004).

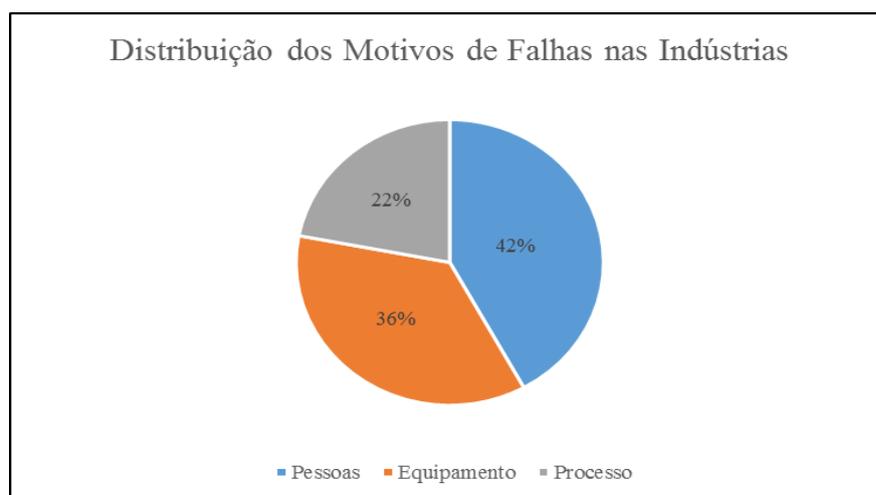


Figura 2 Distribuição dos motivos de falhas nas indústrias. Adaptado de (HONEYWELL, 2004).

Segundo (DUNN, 2003), as falhas ocorrem por fatores como o despreparo da operação e também a incompreensão do responsável técnico sobre esse fator. O autor destaca os métodos de análise de riscos que podem auxiliar na identificação e tratamento de falhas, conhecidos por *Failure Modes and Effect Analysis* (FMEA) e *Fault Tree Analysis* (FTA), considerados métodos qualitativos que possibilitam a análise de eventos de falha, em contraste com o método *Risk Analysis* (RA) que é quantitativo, permitindo analisar a probabilidade de falha em um hardware, um software ou em outro componente (ARUNRAIJ, N. S.; MAITI, J, 2007).

Atualmente os sistemas de gerenciamento de alarmes obtêm informações de dados de processo por meio de computadores, dispositivos e aplicativos, e encontram-se no nível de gerenciamento de um sistema de automação, conforme apresentado na Figura 3.

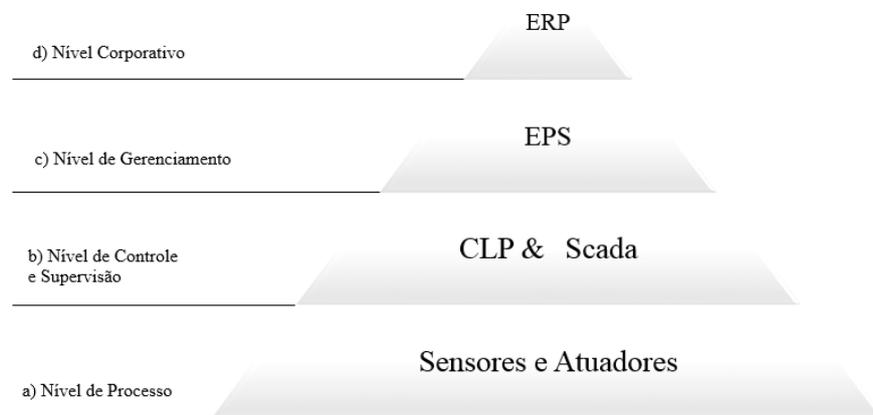


Figura 3 Níveis de um sistema de automação. Fonte: Adaptado de (SOUZA *et al.*, 2005)

Para melhor compreensão desse sistema, são apresentados os 4 níveis segundo (SOUZA *et al.*, 2005). São eles:

- a) Nível de Processo: Onde se encontram os sensores e atuadores responsáveis pela medição e intervenção no processo.

- b) Nível de Controle e Supervisão: Onde se encontram os controladores lógicos programáveis (CLP) e os *Supervisory Control and Data Acquisition* (SCADA) que são responsáveis controle e supervisão do processo.
- c) Nível de Gerenciamento: Onde ocorre a coleta de dados da camada de controle e supervisão. Esta camada é composta pelo *Enterprise Production System* (EPS).
- d) Nível Corporativo: Onde se encontram os sistemas *Enterprise Resource Planning* (ERP), responsáveis por analisar os dados da empresa nos seus diversos setores e transformar em informação de negócio.

Segundo (LEITÃO *et al.*, 2008), os processos de manipulação de informações estão cada vez mais presentes na indústria e o volume de informações que uma pessoa dispõe cresce diariamente. Muitas indústrias adotam algum sistema de gerenciamento de alarmes independentes dos supervisórios SCADA permitindo a implementação de funcionalidades para análise dos alarmes gerados e armazenados em uma base de dados. Dessa forma, transformar a grande quantidade de dados gerados nesses processos em conhecimento torna-se um desafio.

Um estudo realizado pela *Health and Safety Executive* (HSE)¹ concluiu que durante a operação normal de um processo de determinado porte, um operador recebe em média um alarme a cada dois minutos. Após um distúrbio na planta, ou seja, em uma condição anormal, ocorrem cerca de 90 alarmes no primeiro minuto, seguidos de 70 alarmes ao longo dos próximos 10 minutos, tornando a operação não gerenciável (BRANSBY e JENKINSON, 1998). Diante disso, o grande volume de eventos gerados a partir dos processos industriais torna os sistemas de gerenciamento ineficazes.

¹ HSE: <http://www.hse.go.uk>

Segundo (ETZION e NIBLETT, 2010), um evento é uma ocorrência dentro de um sistema, sendo algo que já aconteceu. Os eventos podem alterar o estado do ambiente produzindo novas situações. Contudo, os alarmes devem ser configurados para informar somente os eventos mais necessários, seguindo uma priorização que pode ser realizada por meio da racionalização, ou seja, uma investigação sistemática da base de dados onde estão armazenados (O'BREIN e WOLL, 2004).

Diante da grande quantidade de informação, torna-se necessário possuir meios que facilitem a obtenção das informações necessárias, extraindo conhecimento para auxiliar na tomada de decisão. Nesse sentido, análises estatísticas podem identificar os alarmes e eventos mais frequentes.

Segundo (MIRANDA, 1999), é possível estabelecer uma hierarquia entre dado, informação e por último conhecimento, de forma que os dados são os itens mais elementares e que uma vez coletados, organizados e ordenados tornam-se dentro de um contexto uma informação. Já o conhecimento encontra-se na camada superior às informações e nada mais é do que um padrão que envolve dados e informações. Obter conhecimento na análise manual de uma grande quantidade de dados não é uma tarefa trivial devido às limitações humanas que tornam o processo lento e com alto custo. Devido a esta dificuldade, a utilização de técnicas de mineração de dados de processos para a melhoria no tratamento e organização destes dados apresenta-se como uma possibilidade de aumento da eficiência no tratamento de alarmes, apoiando operadores em sua tomada de decisão.

A mineração de processos é uma técnica que permite identificar e classificar eventos do processo e com isso identificar determinado comportamento presente em um histórico de contexto (IEEE, 2011).

Algumas metodologias podem ser utilizadas para facilitar esse processo. Uma delas trata da descoberta de conhecimento em base de dados chamada KDD (*Knowledge Discovery in Databases*) (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).

Segundo (ALMEIDA e BAX, 2003), técnicas de tratamento e organização de dados podem utilizar ontologia. Essa técnica permite a representação formal do conhecimento e gera uma especificação formal e explícita de um conceito. Neste sentido, ontologias são desenvolvidas para o compartilhamento e entendimento comum de algum domínio do conhecimento, bem como a reutilização de informações.

1.1 MOTIVAÇÃO

Diante dos problemas citados anteriormente, esse trabalho aborda a detecção de desvios de comportamento normal da planta industrial. A partir da identificação de informações relevantes que possibilitem a formalização do conceito de situação, é proposto o uso de inferência para determinação da situação e a avaliação dos padrões comportamentais dos eventos registrados em uma base de dados do processo.

Assim, a questão que motiva esse trabalho é: uma abordagem suportada por técnicas de mineração e ontologias pode ser utilizada para extrair conhecimento de uma base de dados de alarmes de processo?

1.2 OBJETIVO

Nesta dissertação, o objetivo principal é apresentar uma metodologia de acesso e manipulação das informações de alarmes e eventos armazenadas em uma base de dados de uma planta industrial, através de modelagem semântica de contexto de um domínio específico. Esta dissertação apresenta um modelo de contexto, implementa uma ontologia e busca gerar regras para avaliar padrões para determinadas situações de interesse na planta.

Nesta dissertação é apresentado um modelo de contexto por meio de ontologia codificada em linguagem *Ontology Web Language (OWL)*. Os objetivos específicos referem-se a compreender o processo de descoberta de conhecimento, mineração de processos e modelagem semântica de contexto.

1.3 ORGANIZAÇÃO DO TEXTO

O texto é estruturado da seguinte maneira: No Capítulo 2 são apresentadas discussões de fundamentação, bem como, os conceitos utilizados no trabalho. No Capítulo 3 são apresentados os trabalhos relacionados ao sistema proposto. No Capítulo 4 é apresentada uma descrição da proposta desta dissertação. No Capítulo 5 são apresentados aspectos da modelagem de contexto. No Capítulo 6 a definição da abordagem para implementação do modelo é apresentada. No Capítulo 7 são apresentadas as regras de inferência definidas para esse estudo e a validação experimental. No Capítulo 8 são apresentadas as considerações finais, bem como os trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta as áreas relacionadas com o trabalho e está dividido em cinco seções. Na seção 2.1 é apresentado o conceito de sistemas de gerenciamento de alarmes. Na seção 2.2 são tratados aspectos das normas relacionadas. Na seção 2.3 são apresentados os conceitos que abrangem a área de mineração de dados de processo. Na seção 2.4 são tratados os conceitos de análise de contexto. Por fim, na seção 2.5 são apresentados os conceitos de ontologia.

2.1 SISTEMAS DE ALARMES

Os sistemas de automação onde estão presentes os sistemas de gerenciamento de alarmes apresentam uma estrutura em pirâmide, conforme apresentado na Figura 3 do capítulo anterior, onde os diferentes níveis possuem características e propósitos distintos. A integração vertical entre os diferentes níveis permite a melhora e integração dos processos, sendo que os sistemas SCADA criados na década de 70, juntamente com as interfaces *Man Machine Interface* (MMI) (PEREIRA e JUNIOR, 2003), assumiram um papel importante no controle de processos de automação, apresentando de forma amigável os dados do processo. A medida que esses sistemas foram evoluindo as plantas industriais passaram a instalar um número cada vez maior de equipamentos de campo, como por exemplo sensores, atuadores, entre outros.

Com o uso de sistemas supervisórios é possível gerar sinalizações ou “alarmes” para os comportamentos da planta, que auxiliem na operação. Contudo, a complexidade das operações, em diferentes processos, como uma usina de geração térmica de energia elétrica, que dispõe de inúmeros equipamentos, necessita basicamente seguir guias e normas que serão apresentadas na seção seguinte. Sabe-se que a instalação e configuração desses equipamentos, em geral, não segue uma metodologia formal, sendo, muitas vezes baseada na intuição e

experiência das pessoas envolvidas no processo. A consequência surge durante a ocorrência de um distúrbio na planta, provocando a geração de um volume considerável de alarmes na tela do supervisor, dificultando a análise e correta tomada de decisão do operador, e que pode levar a danos ambientais, dos equipamentos e da própria segurança no processo.

Segundo a *Engineering Equipment and Materials Users Association* (EEMUA, 1999), os sistemas de gerenciamento de alarmes ajudam a manter as condições de operação da planta e a identificar situações de emergência ESD (*Emergency Shutdown*), evitando ou diminuindo os danos descritos anteriormente. A Figura 4 apresenta um exemplo genérico de um sistema de gerenciamento de alarmes.

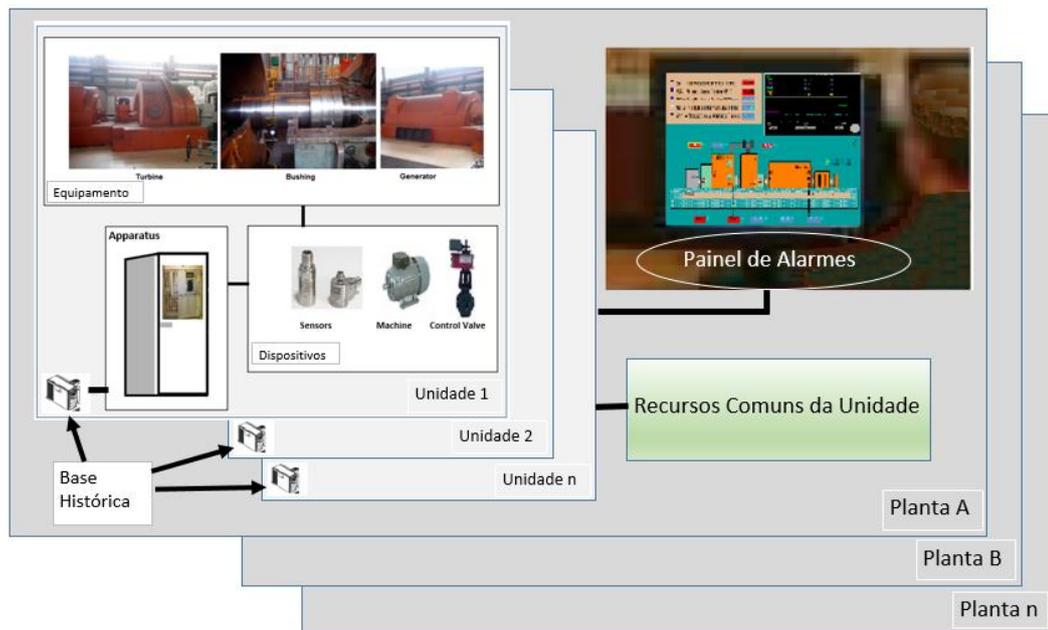


Figura 4 Visão geral de um sistema de alarmes de processo genérico. Fonte: o autor.

No exemplo genérico apresentado na Figura 4, os alarmes e eventos que ocorrem nas diferentes plantas são configurados para serem armazenados em banco de dados histórico (*Historical Database*) e apresentados em painéis supervisórios, conforme definido durante o comissionamento ou posteriormente a partir de intervenções ao longo do processo produtivo. Recursos comuns (*Unit Common Resources*), presentes na planta, como por exemplo os sistemas de abastecimento de matéria prima, torres de resfriamento de água do processo ou

estações de tratamento de água e efluentes, em geral, geram sinalizações com as mesmas características do restante da planta, ou seja, são apresentadas na tela do supervisor e armazenadas na base dados, de acordo com uma configuração prévia realizada no sistema de automação.

Segundo (QUINTÃO e ROSÁRIO, 2008), um alarme é efetivo quando indica o ponto onde o operador deve agir. A qualidade da informação que os sistemas de alarme transmitem e a sua capacidade na orientação do operador é um indicativo de eficiência. Como técnica para o desenvolvimento da proposta apresentada pelos autores (QUINTÃO e ROSÁRIO, 2008), funções necessárias para gerenciamento de condições críticas estão conectadas a uma máquina de inferência para monitoramento em *background*. As condições críticas ocorrem a partir de uma combinação de eventos que podem causar acidentes ou perdas para o processo (QUINTÃO e ROSÁRIO, 2008).

Os sistemas de alarmes têm por característica durante uma perturbação da planta gerar uma quantidade de alarmes que excede a capacidade de análise pelo operador. Um estudo realizado em uma refinaria da Suécia constatou que, durante uma parada de máquina, no primeiro minuto desde o primeiro evento relacionado, uma taxa de um alarme a cada dois segundos foi gerada, tornando inevitável a parada dos equipamentos (MATTIASSON, 1999).

Segundo (LEITÃO *et al.*, 2008), são necessários a partir de sistemas supervisórios um gerenciamento efetivo dos alarmes gerados na planta. Para isso existem guias e normas que estabelecem princípios e que acompanham esse gerenciamento. A seção a seguir apresenta os principais documentos relacionados a esse tema.

2.2 GUIAS E NORMAS

Conhecida como a melhor prática na área de gerenciamento de alarmes a *Engineering Equipment and Materials Users Association* (EEMUA), em sua publicação número 191,

fornece recomendações validadas por profissionais a partir de suas experiências, assim como estudos de fatores humanos (HATCH, 2005). Esse guia apresenta os principais fundamentos que norteiam um sistema de alarmes. Entre eles estão:

- a) Usabilidade: Visa garantir que o projeto de um sistema de alarmes absorva as necessidades do usuário, operando dentro de suas limitações. Deve indicar claramente qual resposta é requerida, apresentar-se em uma taxa na qual o usuário pode gerenciar e ser amigável.
- b) Segurança: Visa garantir a identificação de que o sistema de alarmes preserva a segurança das pessoas, do meio ambiente e dos equipamentos.
- c) Monitoramento de desempenho: Determina que a performance de um sistema de alarmes deve ser avaliada durante o projeto e comissionamento para garantir sua usabilidade e eficiência durante todas as condições de operação.
- d) Investimento em engenharia: define que o sistema de alarmes segue uma metodologia de projeto que o justifica.

Em um sistema de gerenciamento de alarmes é fundamental a priorização dos alarmes para que se tornem mais evidentes para os operadores, facilitando a tomada de decisão. Para isso, o guia propõe três métodos para determinar a prioridade de alarmes. O primeiro deles é baseado na soma das consequências, que podem ser econômicas, de segurança e meio ambiente. O segundo determina a prioridade de acordo com a consequência máxima do alarme. O último é baseado em um diagrama de fluxo. Nesse guia, também são definidos índices de desempenho para o funcionamento adequado de um sistema de alarmes. O primeiro deles define que o operador deve ter pelo menos 10 minutos para avaliar cada alarme que surge num total de 144 alarmes por dia. Em seguida, determina que a lista de alarmes não contenha mais do que 9 alarmes contínuos, evitando assim que não haja dificuldade na

detecção de alarmes pendentes. É sugerida uma distribuição de prioridade de 80% para alarmes de baixa prioridade, 15% para alarmes de prioridade média e somente 5% para alarmes de alta prioridade. Contudo, uma tabela contendo valores encontrados na indústria, em sua segunda edição da EEMUA publicada em 2007, mostra um comparativo destes índices de desempenho.

Tabela 1 índice de desempenho EEMUA 191.

Indicador	EEMUA	Óleo e Gás	Petroquímica	Energia	Outras
Alarmes/dia	144	1200	1500	2000	900
Alarmes constantes	9	50	100	65	35
Alarmes por 10 minutos (picos)	10	220	180	350	180
Alarmes por 10 minutos (média)	1	6	9	8	5
Distribuição de prioridade (Baixa/Média/Alta)	80/15/5	25/40/35	25/40/35	25/40/35	25/40/35

Fonte: Traduzido de EEMUA, 2007.

A partir da Tabela 1 observa-se que os sistemas de alarmes em geral estão muito distantes do que é considerado aceitável.

Um comitê da *International Society of Automation* (ISA) denominado *Instrument Signals and Alarm* é responsável por definir terminologias e práticas para sistemas de alarmes de processo utilizados em computadores. No padrão *Management of Alarm Systems for Process Industry* ISA 18.2 (ANSI/ISA-18.2-2016) que trata especificamente de gerenciamento de alarmes, é estabelecido um modelo representado pelo ciclo de vida. Ele é composto por uma definição de filosofia de alarmes, gerenciamento de mudanças, auditoria e avaliação, conforme mostra a Figura 5.

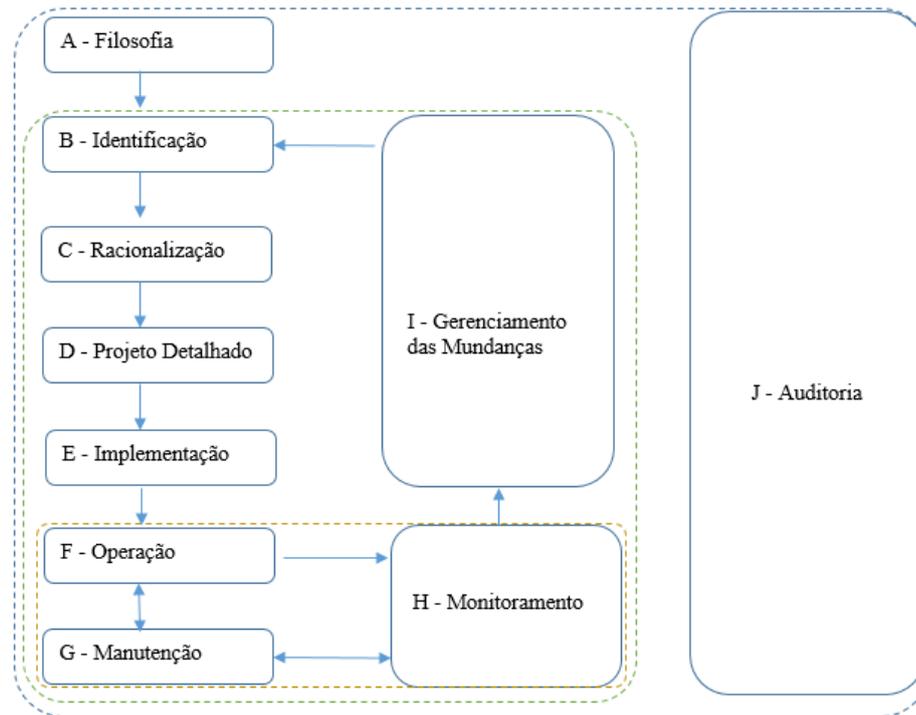


Figura 5 Ciclo de vida do gerenciamento de alarmes. Traduzido de (ANSI/ISA-18.2-2016).

A seguir são as fases do ciclo representado na figura 4, que são:

- a) **Filosofia:** Define as normas de como a empresa realizará o gerenciamento de alarmes por meio de todas as fases do ciclo de vida. Nessa etapa, devem estar presentes as regras para a classificação e priorização de alarmes, estabelecendo parâmetros de desempenho.
- b) **Identificação:** Visa determinar o conjunto mínimo de alarmes necessários para manter o processo seguro e sob controle e poderá ser realizada por um método de análise de riscos.
- c) **Racionalização:** A partir dos alarmes identificados na etapa anterior e com base nos princípios e requisitos definidos na primeira etapa dessa filosofia, ocorre uma revisão e justificativa para que eles cumpram os requisitos de um alarme. Também é determinado nessa fase a documentação de

prioridade de cada alarme, classificação, limites, causas e consequências, além de ações corretivas no banco de dados onde ficam armazenados.

- d) Projeto: Define a interface de visualização e acesso aos alarmes e métodos de gerenciamento. Nessa etapa, cada procedimento deve ser documentado, indicando os procedimentos realizados.
- e) Implementação: Visa colocar em prática a etapa anterior. Prevê o treinamento dos operadores no sistema, a fim de esclarecer as funcionalidades de cada um dos alarmes.
- f) Operação: Etapa de execução do sistema de alarmes e armazenamento na base de dados.
- g) Manutenção: Visa descrever os procedimentos que devem ser seguidos para retirar um alarme de serviço, incluindo a documentação de justificativa para remoção e os testes necessários para colocar novamente em serviço.
- h) Monitoramento: Visa analisar os dados gerados pelo sistema durante a etapa de operação para identificar problemas com os parâmetros adotados. Esta etapa prevê o monitoramento de desempenho e comparação com métricas da norma.
- i) Gerenciamento das Mudanças: Visa controlar da forma estruturada os processos de configuração, alteração, remoção ou adição de alarmes. Prevê a utilização de ferramentas e procedimentos que garantam que as alterações passem por revisões e sejam aprovadas antes da implementação.
- j) Auditoria: Visa a revisão periódica dos processos de trabalho e desempenho dos sistemas de alarme. Tem por objetivo manter sua integridade em todo ciclo de vida para identificar melhorias.

Para cada uma das etapas estão previstos procedimentos e recomendações, bem como índices de desempenho semelhantes aos encontrados na EEMUA 191, conforme apresentado na Tabela 2.

Tabela 2 índice de desempenho ISA 18.2.

Indicador	Valor	
	Satisfatório	Máximo gerenciável
Alarmes por tempo		
Hora	~6 (média)	~12 (média)
10 minutos	~1 (média)	~2 (média)
% de períodos de 10 minutos com mais de 10 alarmes	~<1%	
Máximo de alarmes no período de 10 minutos	<=10	
% tempo com enxurrada de alarmes	~<1%	
% contribuição dos 10 alarmes mais frequentes	~<1 até 5	
Quantidade de alarmes incômodos	Zero	
Alarmes contínuos	<5/dia	
Distribuição de prioridade (Baixa/Média/Alta) %	80/15/5	
Distribuição de prioridade (Baixa/Média/Alta/Muito Alta) %	80/15/5/<1	

Fonte: Traduzido de ANSI/ISA-18.2-2016

A partir das métricas apresentadas é possível determinar a integridade ou a qualidade dos sistemas de alarmes.

2.3 MINERAÇÃO DE DADOS

Antes de tratar os conceitos relacionados à mineração de dados, torna-se necessário compreender o significado conceitual de dado, pois auxiliará no entendimento de outros conceitos como informação e conhecimento.

2.3.1 Dados, informação e conhecimento

A mineração de dados é uma das etapas de um processo conhecido como extração de conhecimento. Do ponto de vista de um processo computacional, (REZENDE, 2005) define dado como um conjunto de registros que é fornecido como entrada de um processo, que

agrupado, categorizado e padronizado adequadamente, transforma-se em informação, ou seja, aquilo que processo fornece como saída.

Segundo (MIRANDA, 1999), os termos dados, informação e conhecimento muitas vezes são utilizados de forma indiferente e tratados como sinônimos. Segundo (SETZER, 2004), dado é puramente sintático e informação contém necessariamente semântica. Computadores são máquinas sintáticas. Dessa forma, processam dados e dessa forma a informação é armazenada por meio de uma representação em forma de dados. Nesta dissertação, as informações estão contidas em uma base de dados de alarmes de processo. A informação armazenada pode gerar conhecimento que auxilie na operação, por meio da análise de padrões históricos de eventos.

Segundo (SETZER, 2001), o conhecimento é uma abstração interior e pessoal de algo que foi experimentado ou vivenciado, estando relacionado a vivência pessoal. Assim, um conjunto de dados ou uma informação se transformará em conhecimento quando estiver em um contexto que foi vivenciado pelo receptor. Apesar da informação ser capaz de alterar o conhecimento atual de uma pessoa, não é de toda a informação que uma pessoa necessita.

Informação relevante é aquela que o usuário necessita para satisfazer sua necessidade em um determinado momento e deve estar em um contexto de que o usuário precisa. Além disso, dados são objetivos e envolvem sintaxe. Informações são objetivas e subjetivas e envolve semântica, podendo ser transmitidas através de dados ou não, e o significado depende de cada receptor. Conhecimento é compreendido como sendo totalmente subjetivo e depende da forma com que a pessoa percebe o mundo (SETZER, 2001).

O conhecimento é utilizado pela pessoa para resolver problemas e executar tarefas que surgem em diversas situações. Com o aumento do volume de dados, é ineficiente a extração de conhecimento por analistas de forma manual. Segundo (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996), a forma manual de sondagem de um conjunto de dados é lenta,

cara e altamente subjetiva. Portanto há a necessidade de utilizar metodologias que envolvem conceitos computacionais, para extrair conhecimento de uma base de dados.

Para esta dissertação, a situação de interesse é a informação que se deseja obter a partir dos dados e análise de padrões obtidos a partir da etapa de mineração de processo. O contexto de interesse são as informações que definem a ação a ser tomada. Os dados referentes aos valores de vibração da unidade de geração de energia elétrica, por exemplo, são agregados com entidades de contexto para gerar informação de mais alto nível, caracterizando a situação da planta. Com essa caracterização é possível a tomada de decisão. A utilização de axiomas e regras permite raciocinar sobre o contexto, o que gera conhecimento por meio do ambiente modelado.

Segundo (CHEN e KOTZ, 2000), entidade pode ser um lugar, sensor, pessoa, entre outros, onde os dados são capturados e onde são gerados os contextos, podendo raciocinar sobre um domínio de interesse.

2.3.2 O Processo de Descoberta de conhecimento

Como foi descrito anteriormente, a mineração de dados é uma das etapas de um processo conhecido como extração de conhecimento. Nesse sentido, descobrir conhecimento é identificar, receber informações, computá-las e adicioná-las a um conhecimento prévio. Assim, para que um problema possa ser resolvido, ocorre um processo de mudança no estado de um conhecimento atual, tornando-se um processo amplamente complexo e subjetivo dependendo do problema a ser solucionado e da forma como uma situação é percebida pela pessoa, passando inclusive pela forma com que a pessoa processa essas informações para solucionar um determinado problema, e assim, tomar decisões.

Um dos maiores problemas na tomada de decisões diz respeito a quantidade de informações. Somente com apoio computacional é possível manipular a informação de forma

eficiente a fim de descobrir novos conhecimentos. Nesse sentido a descoberta de conhecimento com auxílio computacional é um processo de análise de dados. Os métodos e ferramenta utilizados para realizar esse processo surgiram com base em métodos estatísticos, métodos de inteligência artificial e de recuperação de informação. Um dos métodos mais conhecidos é o KDD que permite selecionar dados presentes em um banco de dados. O processo é dividido em cinco etapas, tendo como entrada um conjunto de dados e como saída o conhecimento (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996). A Figura 6 apresenta as etapas previstas nesse processo e indica que da mineração de dados obtém-se padrões que são avaliados posteriormente.

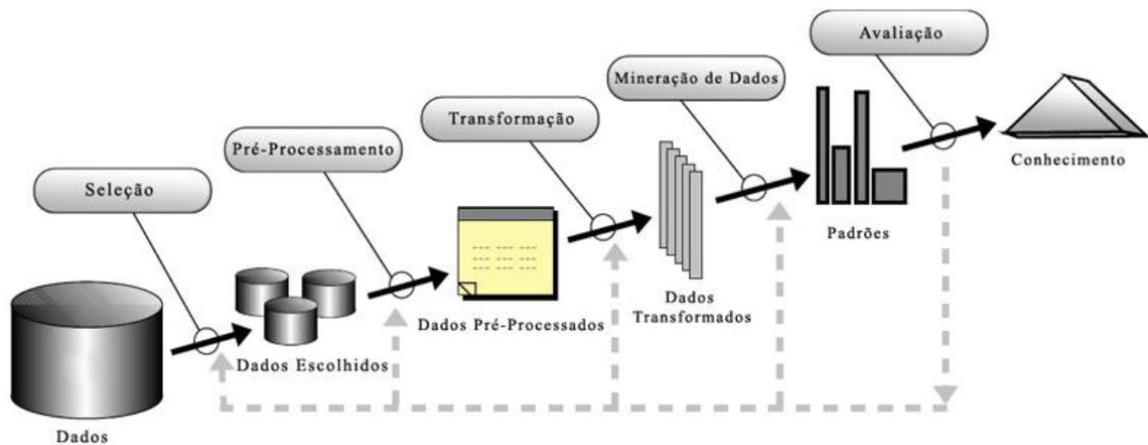


Figura 6 Representação do processo KDD. Traduzido de (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).

A primeira etapa apresentada na Figura 6 resulta na seleção dos dados que serão utilizados, os quais podem ser derivados de diversas fontes. A escolha dos dados relevantes ao contexto influencia na qualidade dos padrões que serão encontrados nas etapas finais. Essa seleção de dados pode ser realizada por um especialista no domínio. A etapa seguinte irá eliminar dados discrepantes ou que estejam incompletos. Essa etapa pode tomar um tempo expressivo. Na etapa de transformação o conjunto de dados será formatado e armazenado em um formato que permita a etapa seguinte. Por fim, ocorre a avaliação dos padrões

encontrados, sendo possível retornar para etapas anteriores do processo (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).

Pela Figura 6, a extração de conhecimento é obtida pela análise dos padrões obtidos da mineração de dados por meio de técnicas e algoritmos específicos, sendo considerada uma área multidisciplinar, envolvendo estatística, inteligência artificial, aprendizagem de máquina entre outras (LIU, 2011). A escolha da técnica de mineração de dados está intimamente ligada à tarefa que será utilizada, podendo ser empregadas várias técnicas dependendo do objetivo desejado. Como exemplo a aprendizagem de máquina, que é um método bastante comum de classificação de dados (WITTEN, EIBE e HALL, 2011).

Conforme apresentado na Figura 6, somente depois de um pré-processamento é que as técnicas de mineração de dados são utilizadas para gerar padrões. O trabalho de (KOFI, LUO e QUIN, 2013), descreve que métodos de mineração de dados são usados por pesquisadores para construir modelos de predição, como análise de regressão, redes neurais, teorias Bayesianas entre outras. Os autores apresentam um método disponível na ferramenta *Waikato Environment for Knowledge Analysis* (WEKA) (HALL *et al.*, 2009) que classifica informações comparando com exemplos disponíveis na própria base de dados.

Segundo (KLEMETTINEN, M.; MANNILA, H.; TOIVONEN, H., 1997), o método KDD pode ser aplicado para construir um sistema de correlação a partir de características de alarmes em uma base de dados de alarmes de telecomunicações. O método proposto pelos autores facilita a correlação e a identificação de falhas, gerando regras de associação a partir de predicados analisados das mensagens de alarmes que ocorrem frequentemente.

Segundo (DEY e ABOWD, 1999), as pessoas podem utilizar informações sobre tudo que acontece em um contexto. No entanto, os computadores não estão habilitados a analisar um determinado contexto durante a interação com seres humanos, o que torna a definição de contexto no campo computacional uma tarefa delicada. Ainda segundo os autores, contexto é

qualquer informação que possa ser usada para caracterizar a situação de uma entidade. Entidade entende-se por qualquer pessoa, lugar ou objeto considerado relevante na interação entre o usuário e a aplicação. Na próxima seção serão tratados os conceitos de dados e modelagem de contexto.

2.4 ANÁLISE DE CONTEXTO

Segundo (DEY e ABOWD, 1999), contexto é qualquer informação que possa ser usada para caracterizar a situação de uma pessoa, lugar ou objeto, ou seja, uma entidade. O termo contexto trata de tudo que influencia na forma com que o usuário interage com o ambiente físico, ou que ocorre ao redor deste usuário.

Segundo (CHEN e KOTZ, 2000), contexto é um conjunto de estados do ambiente que determina o comportamento da aplicação e que seja de interesse do usuário. Já (SANTOS, 2008), apresenta o conceito de elemento contextual como qualquer dado, informação ou conhecimento que permite caracterizar uma entidade em um domínio. O contexto da interação entre agente e aplicação para a execução de uma tarefa, segundo a autora, é um conjunto de elementos contextuais que são necessários para apoiar uma tarefa.

Segundo (DEY, 1999), o projetista de um sistema utiliza informações de um contexto para determinar porque ocorre uma determinada situação e determina ações que serão executadas pela aplicação. A questão é que, segundo o autor, histórico de contexto explora as informações do estado atual de funcionamento do sistema e retorna conhecimento.

Considerando as definições descritas, será adotado o conceito de contexto como sendo todos os dados que influenciam o funcionamento do sistema, bem como o conceito definido por (SANTOS, 2008), que é um conjunto de elementos contextuais que são necessários para apoiar uma tarefa e um conjunto de estados do ambiente que determinam a situação da planta.

Nesta dissertação, é utilizado ontologia para estruturação do conhecimento no domínio da planta, de modo a dar suporte para área de automação. Segundo (QUINTÃO e ROSÁRIO, 2008), os sistemas de automação concentram toda supervisão e controle da planta industrial nas salas de controle, onde os operadores recebem uma quantidade muito grande relacionada a alarmes e eventos. Conforme apresentado por (DEY e ABOWD, 1999), uma aplicação não pode determinar sozinha porque uma situação está ocorrendo, mas o desenvolvedor da aplicação consegue. Dessa forma, o desenvolvedor utiliza informações de contexto para determinar por que uma situação ocorre e utiliza isso para determinar ações necessárias diante da situação.

Uma situação por exemplo, é definida como uma interpretação semântica externa dos dados do sensor. Essa interpretação significa atribuir significado aos dados dos sensores. O fato de ser externa diz respeito a metodologia ou aplicação, enquanto semântica significa que a interpretação atribui significado em dados de sensores baseados em estruturas e relações dentro de um mesmo tipo de dado do sensor (YE, STEVENSON e DOBSON, 2011).

Uma situação pode ser definida a partir da identificação de contextos relevantes, descobrindo correlações significativas entre eles, e assim rotulá-los com um nome descritivo. Esse nome pode ser a definição descritiva de uma situação, que é definido pelo ser humano sobre um estado das coisas. A expressão lógica de predicados de contexto correlacionados é chamada de uma especificação de uma situação lógica (YE, STEVENSON e DOBSON, 2011). O predicado de contexto também pode ser entendido como uma ação, sendo associado com relacionamentos do tipo propriedade de objeto entre os conceitos. Ainda, sujeito e objeto são relacionados com conceitos ontológicos ou classes, que são um dos elementos da ontologia. Modelos definidos que representam contexto utilizam predicados de primeira ordem, possuem em sua estrutura um sujeito, um predicado e um verbo (WANG *et al*, 2004).

Por exemplo, um dado contexto de estado pode ser derivado através de informações produzidas por eventos, tais como um alarme, sendo possível identificar a situação da planta.

O contexto pode ser definido como um conceito estruturado que descreve propriedades do ambiente, como dados dos sensores abstraídos em um conjunto de conceitos do domínio. Dessa forma, contextos em um domínio podem ter uma estrutura diferente, que é distinta dos contextos de outros domínios. Também podem ser classificados em termos das propriedades que os descrevem. Por exemplo, no domínio de localização, o contexto pode ser identificado com três valores numéricos medidos em metros, enquanto no domínio de temperatura, o contexto é identificado como um valor numérico que possui como unidades de medida Celsius ou Fahrenheit (YE, STEVENSON e DOBSON, 2011).

2.4.1 Dados de Contexto

Segundo (PREKOP e BRUNETT, 2003), o contexto pode ser dividido em externo e interno, sendo os dados de sensores pertencentes à dimensão externa e as atividades executadas e as informações relativas ao usuário pertencentes à dimensão interna. Este estudo considera os dados de contexto da dimensão externa para modelagem.

A complexidade dos sistemas torna necessária a geração de modelos, que aumentem a abstração que auxiliem no entendimento do ambiente. Abordagens relacionadas a modelagem de contexto são apresentadas nos trabalhos de (STRANG, LINNHOFF-POPIEN, 2004; MOORE *et al.*, 2007). As pesquisas apresentam as exigências relacionadas a composição distribuída (CD), validação parcial (VP), qualidade da informação (QI), dados incompletos (DI), Formalidade (NF) e aplicabilidade (AP), conforme apresentado na Tabela 3.

Tabela 3 Análise das técnicas para modelagem de contexto.

	Métricas					
Modelos	CD	VP	QI	DI	NF	AP

Valor-Chave	-	-	-	-	-	+
Esquemas de Marcação	+	++	-	-	+	++
Gráficos	-	-	+	-	+	+
Orientados a Objetos	++	+	+	+	+	+
Baseados em Lógica	++	-	-	-	++	+
Baseados em Ontologias	++	++	+	+	++	+
Aprendizagem de Máquina	+	+	-	++	++	-

Fonte: Traduzido de (STRANG, LINNHOFF-POPIEN, 2004; MOORE et al., 2007).

Na Tabela 3 é apresentado o comparativo dos modelos definidos para representação de contexto. O sinal de (-) representa limitação do modelo. Nesse sentido, um modelo baseado em lógica é apresentado como fraco com relação à validação parcial devido à dificuldade existente em reduzir erros e com relação ao desenvolvimento de *Reasoners* lógicos, que permitem inferir consequências lógicas, ou seja, raciocínio. Os sinais de (+) e (++) representam uma estrutura mais adequada para modelagem, onde segundo os autores a modelagem baseada em ontologias contemplou de forma mais adequada o modelo definido para representação de contexto.

2.4.2 Geração de Modelos de Contexto

A geração de modelos ocorre a partir da tarefa de pensar sobre um domínio, que é a representação da realidade destinada a um propósito definido. Esses modelos devem extrair a essência dos fenômenos do mundo. A conceituação é a representação abstrata de fenômenos que estão presentes no mundo e no imaginário das pessoas, que são representadas por símbolos com significado dado pelas pessoas.

A modelagem de contexto é feita através de ontologia. A mais comum utilizada atualmente é a OWL. Uma variação dessa ontologia é a *Description Logic* (DL) que permite

incluir restrições de propriedades que fazem com que valores de determinada propriedade pertençam a uma determinada classe e dessa forma estruturas e expressões OWL-DL servirão como base para motores de inferência inferir entidades de contexto. Segundo (POLI e OBRST, 2010), ontologias podem ser divididas em dois pesos. As de peso leve relacionadas a hierarquias formais e as de peso pesado que adicionam as ontologias de peso leve uma interpretação semântica compreensível pela máquina. As ontologias são utilizadas para dotar sistemas computacionais de meta-conhecimento, permitindo capacidade de raciocínio, ou seja, permitindo um maior entendimento sobre o domínio no qual devem interagir.

Aplicações sensíveis ao contexto utilizam ontologias para representar o raciocínio lógico. Segundo (YE, STEVENSON e DOBSON, 2011), técnicas de raciocínio permitem a descoberta de conhecimento e assistência sensível à situação. A sensibilidade a um determinado contexto depende do conhecimento do estado atual de funcionamento de um sistema, sendo possível a classificação de situações.

2.5 CONCEITOS DE ONTOLOGIA

Ontologias agrupam conceitos de um domínio de conhecimento. Seu desenvolvimento segundo (NOY e MCGUINNES, 2001) envolve a definição de um conjunto de elementos como entidade, classe, propriedade, restrição, predicados e funções e pode ser escrita em uma linguagem formal que permite a descrição do conhecimento, sendo utilizada para realizar inferência sobre os objetos do domínio.

Segundo (GRUBER, 1993), em uma ontologia os conceitos e restrições de uso são explicitamente definidos, podendo ser entendida por máquina, captura um conhecimento e se refere a um modelo abstrato do mundo real. Conforme o autor, uma ontologia pode ser descrita como um conjunto de termos de uma representação, existindo a associação entre classes, funções e objetos.

A *Web Ontology Language* (OWL) é a linguagem mais utilizada para construção de ontologia. O *World Wide Web Consortium* (W3C) possui como padrão a linguagem OWL que possui como ferramenta de desenvolvimento o Protégé² (STANFORD, 2015).

Segundo (MCGUINNESS e HAMERLEN, 2004), três sublinguagens definem o nível de expressividade, sendo OWL-Lite a que fornece o menor grau, mantendo restrições simples e uma classificação hierárquica. Em seguida OWL-DL que permite maior expressividade e todas as conclusões são validadas para serem computáveis. Por fim, OWL-Full apresenta a máxima expressividade. As versões DL e Lite permitem a utilização de motores de inferência que permitem validar o modelo desenvolvido, descobrir conhecimento e inferir relacionamentos existentes entre as diferentes classes. Uma ontologia de contexto permite compartilhar conhecimento.

Segundo (BERNERS-LEE, T.; HENDLER, J.; LASSILA, 2001), a ontologia possui como base uma taxonomia, que define classes de objetos e suas relações, além de um conjunto de regras de inferência que possibilita o raciocínio.

A metodologia para construção da ontologia definida por (NOY, N. F.; MCGUINNES, D. L., 2001) define regras que facilitam a tomada de decisão em um projeto, sendo: (1) não há uma maneira correta de modelar um domínio, mas sim alternativas viáveis, sendo que a melhor opção muitas vezes depende da aplicação que se deseja construir; (2) o desenvolvimento de ontologias é necessariamente um processo iterativo; (3) conceitos em ontologia devem ser próximos a objetos (físicos ou lógicos) e relacionamentos no seu domínio de interesse. Estes podem ser substantivos (objetos) ou verbos (relacionamentos) em sentenças que descrevem seu domínio.

O método 101 (NOY, N. F.; MCGUINNES, D. L., 2001) possui como apoio uma estrutura de ações que define etapas para as atividades durante a construção da ontologia. No

² Protégé: <http://protege.stanford.edu>

primeiro passo, são analisados questionamentos tais como: Qual é o domínio em questão? Pelo que está se utilizando ontologia? No passo seguinte, verifica-se a possibilidade de reutilização, que é útil para possibilitar a comunicação com outras aplicações, para o aperfeiçoamento e manutenção do formalismo. Em seguida são listados termos que auxiliam na declaração e nos relacionamentos. Em seguida, são definidas as classes e hierarquias, que ajudam a definir de que tipo são as instâncias. As propriedades das classes são definidas para permitir a implementação do domínio da ontologia. Essa definição de propriedades é necessária pois elas podem possuir diferentes facetas que definem o tipo de valor, valores permitidos, cardinalidade de valores, entre outras características que a propriedade pode assumir.

Regras podem ser utilizadas juntamente com ontologias, como *Semantic Web Rule Language* (SWRL), permitindo a identificação de uma situação que está ocorrendo no presente. É possível determinar a probabilidade através de *Probabilistic OWL* (PR-OWL) de uma situação acontecer no futuro por meio de *Multi-Entity Bayesian Network* (MEBN) (LASKEY, 2008). Nessa dissertação o foco é a identificação de situações atuais por meio de regras SWRL e motores de inferência.

O uso de motores de inferência permite o raciocínio a partir de regras e ontologias, sendo programas que inferem consequências lógicas, tendo como base fatos ou axiomas. A capacidade computacional disponível atualmente aliada ao conhecimento de apoio expresso por meio de redes semânticas, tesouros ou ontologias que podem ser usadas para modelagem, bem como raciocínio sobre contexto. De acordo com (O'BREIN, 2009), sistemas reativos reagem automaticamente à mudanças no ambiente. Para detectar situações, algumas informações são descritas na definição da situação, como o contexto relevante durante a situação detectada, os eventos que irão participar da situação detectada, as condições de consumo dos eventos e as condições semânticas das situações.

O tratamento de situação é entendido como um conceito de mais alto nível, indicando um conjunto de características de contexto que são invariáveis dentro de um espaço de tempo (HERVÁS, BRAVO e FONTECHA, 2010).

No próximo capítulo serão apresentados trabalhos que propõem abordagens baseadas em ontologia, em formas de representação de contexto e em sistemas de alarmes presentes na indústria.

3 TRABALHOS RELACIONADOS

Este capítulo apresenta trabalhos relacionados com os temas da dissertação: modelagem semântica de contexto e aplicação de ontologia em sistemas de alarmes para indústria.

Segundo o trabalho de (BERNARAS *et al.*, 1996), em sistemas de energia o conhecimento envolvido é relativamente complexo, sendo necessário o desenvolvimento de diversas ontologias. A Figura 7 apresenta as diferentes ontologias propostas pelo autor.

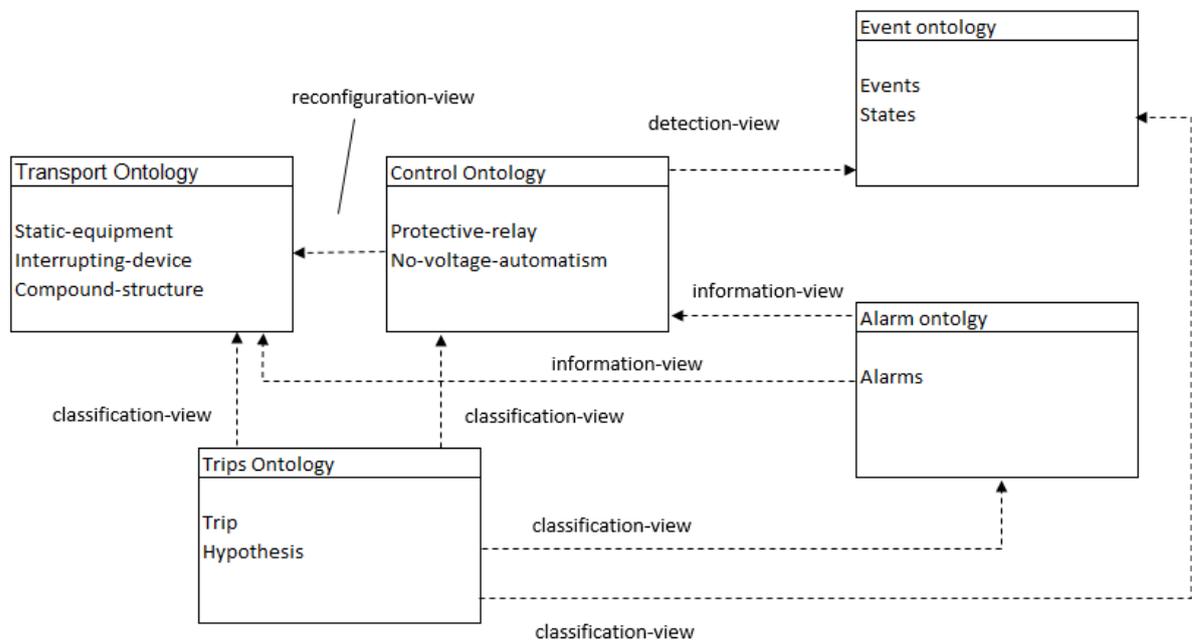


Figura 7 Ontologias para diagnóstico de falha em um sistema de energia. Adaptado de (BERNARAS *et al.*, 1996).

Um evento é uma mudança discreta de um estado do sistema ou dispositivo. A ontologia de evento armazena o conhecimento necessário para identificar um distúrbio quando ocorrer, notificando o engenheiro de controle (BERNARAS *et al.*, 1996).

A empresa Honeywell³ possui um sistema colaborativo denominado *Abnormal Event Guidance and Information System* (AEGIS) aplicado ao setor petroquímico que possui um estimador que fornece o status de uma planta (COCHRAN, MILLER e BULLEMER, 1996). O sistema proposto permite o acesso à infraestrutura de aplicação e informação para auxiliar

operadores em situações anormais. Os autores apresentam problemas existentes em diferentes domínios, apresentando situações presentes em uma sala de controle.

No trabalho de (LAALLAM e SELLAMI, 2007), apresenta-se uma abordagem ontológica para capturar o conhecimento no domínio da manutenção industrial em uma estação de compressão de gás presente em uma turbina que transforma energia térmica em energia elétrica e que constitui uma situação dinâmica. O objetivo é o desenvolvimento de um sistema inteligente que permita tomada de decisão no domínio da manutenção industrial. Este trabalho não apresenta mecanismos de *Reasoning* bem como uma modelagem de contexto que apresente recomendações ou permita inferir sobre o processo auxiliando a operação.

O trabalho de (AIZPURÚA, GALÁN e JIMÉNEZ, 2008), apresenta uma ontologia de domínio que modela hipótese de falhas de uma planta de uma usina hidroelétrica localizada na República do Panamá. A partir de regras observadas com base em 156 alarmes digitais são geradas recomendações aplicando uma rede neural em um histórico de uma base de dados. Segundo apresentado nesse trabalho, a próxima etapa em mineração de dados é implementar um software de detecção de falhas, predição e análise. Assim, um conjunto de características pode ser reconhecido quando ocorrerem.

O modelo proposto é apresentado na Figura 8. O objetivo é apresentar uma árvore de alarmes por meio de informações obtidas através de dados de alarmes, histórico de falhas, base de dados de alarmes e técnicas de inteligência artificial por meio de sistemas especialista e redes neurais (AIZPURÚA, GALÁN e JIMÉNEZ, 2008).

³ Honeywell: <http://www.honeywell.com>

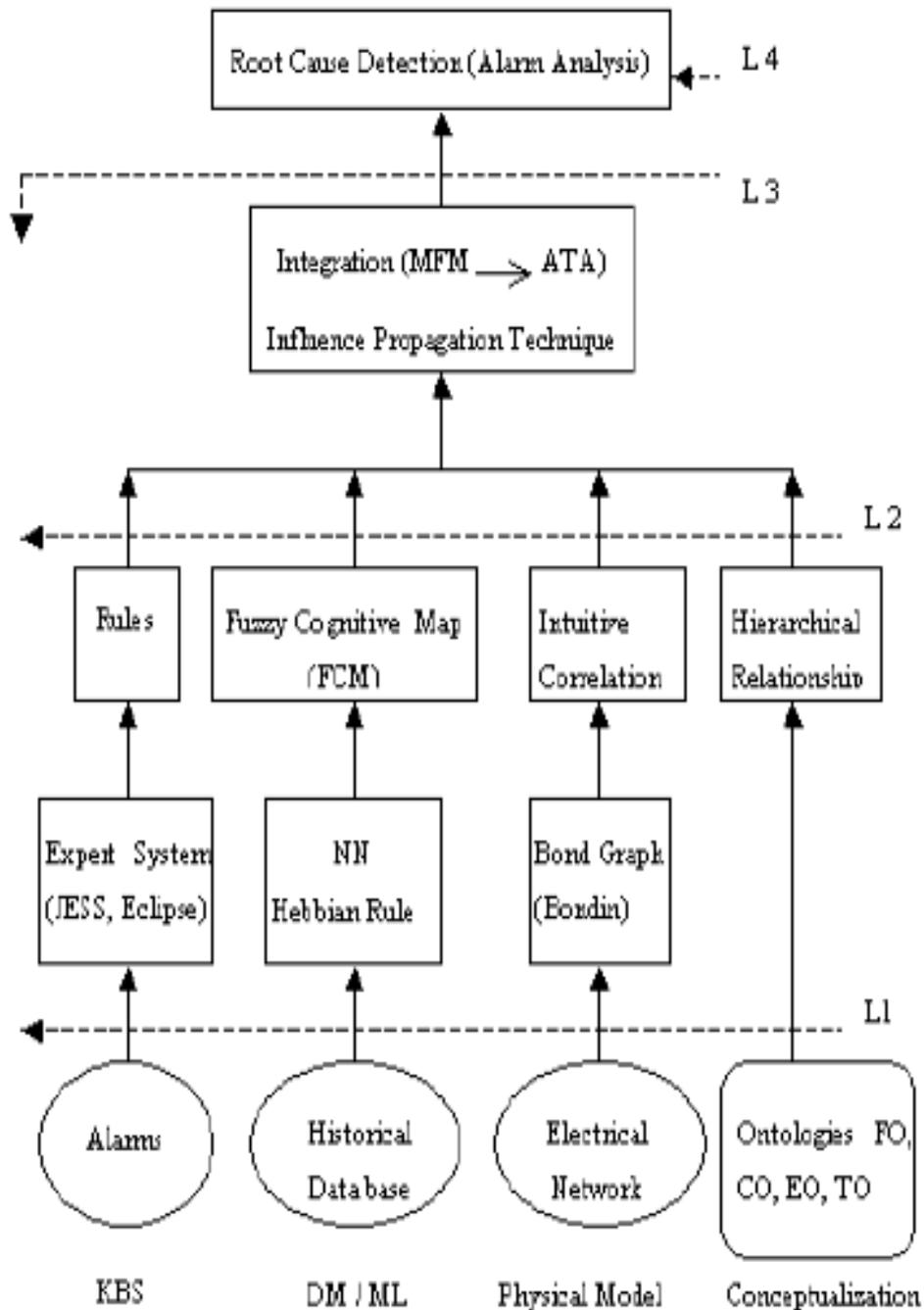


Figura 8 Ontologias para metodologia proposta. Fonte (AIZPURÚA, GALÁN e JIMÉNEZ, 2008).

O modelo apresentado na Figura 8 apresenta diferentes níveis. No nível L1 somente informação é encontrada oriunda do sistema SCADA. Já o nível L2 é apresentado como um nível especialista que permite prover regras a partir da base de alarmes. Nesse caso o Java

Expert System Shell (Jess)⁴ é utilizado como máquina de inferência lógica. No nível seguinte L3 ocorre a integração do algoritmo por meio da aplicação da técnica como *Multilevel Flow Modelling* (MFM). Essa técnica prove a árvore de falhas, derivando situações anômalas da planta. Por fim, no nível L4 os resultados são integrados em uma interface com a ferramenta Protégé, permitindo a análise dos alarmes para detecção de causa.

No trabalho apresentado por (IZADI *et al.*, 2009), uma visão geral sobre sistemas de alarmes é apresentada, abordando os falsos alarmes. Não foi identificado um modelo implementado de contexto ou ontológico. Entretanto os autores apresentam uma análise sobre o volume de falsos alarmes que são apresentados para os operadores e uma proposta para diminuição de alarmes ruidosos. Segundo os autores, a sobrecarga de informações ocorre devido ao volume de sensores e atuadores monitorados. Para análise de base de dados e erro predição são citados métodos estatísticos e modelos dinâmicos. O uso de filtros também é apresentado pelos autores como solução para redução de ruído e de falsos alarmes.

No trabalho apresentado por (SIHEM, TAREK e RINGWOOD., 2009), o Jess é integrado a ontologia para inferir conhecimento e resolver possíveis casos de manutenção. Uma *mono-ontology* é obtida por múltiplos CCOs (*Canonical Conceptual Ontologies*) que são associados à base de dados constituindo uma base de conhecimento. Também são definidos NCCO (*Non Canonical Conceptual Ontology*) que representam um mapa de inter-relação entre os CCOs, expressando seus relacionamentos, consistindo basicamente em representar a correspondência entre os esquemas de dados integrados. Nesse trabalho é apresentada uma ontologia voltada para integração com a base de dados a fim de inferir conhecimento e propor diagnóstico, com base em uma ontologia de domínio, para uma turbina a vapor. Com base em características dos equipamentos e casos de manutenção, são definidos os sintomas e defeitos. Contudo, não foram apresentados claramente os dados dos

⁴ Jess é uma marca registrada de *Sandia National Laboratories*.

equipamentos ou mesmo uma estrutura que represente classes, objetos ou propriedades de objetos. Apenas uma estrutura simples de representação das instâncias e um mapa do *Reasoning* utilizando ferramenta de software Jess. Também é apresentado uma interface *Graphical User Interface* (GUI) descrita para diagnóstico e manutenção que se conecta ao modelo proposto, desenvolvida em *Java 2 Enterprise Edition* (J2EE), mas não são apresentadas recomendações de ações ou mesmo sinalização sobre o estado da turbina.

No trabalho apresentado por (AGUIAR, ALMEIDA e MEIRA, 2010), uma abordagem voltada para mineração de regras de associação e análise de correlação cruzada a fim de detectar alarmes redundantes e padrões de alarmes aplicada a etapa de racionalização de alarmes. Os autores propõem uma combinação de algoritmos para detecção de alarmes redundantes que permite detectar relações temporais adequadas a diferentes tipos de processos.

No trabalho de (LEITÃO, GUEDES e ARAUJO, 2012), é apresentada uma análise que permite a racionalização de alarmes prevista na ISA SP 18.2 e no guia EEMUA 191. Dados de alarmes de uma planta petroquímica são analisados por meio de técnicas de filtragem e cálculos de correlação, a partir da análise do ciclo de vida dos alarmes. Os autores citam diferentes incidentes provocados por sistemas de alarmes ineficientes, como o caso da plataforma de petróleo da Petrobrás⁵ em maio de 2001.

Uma representação do ciclo de vida de alarmes é apresentada conforme a Figura 9.

⁵ Petrobrás: <http://www.petrobras.com.br>

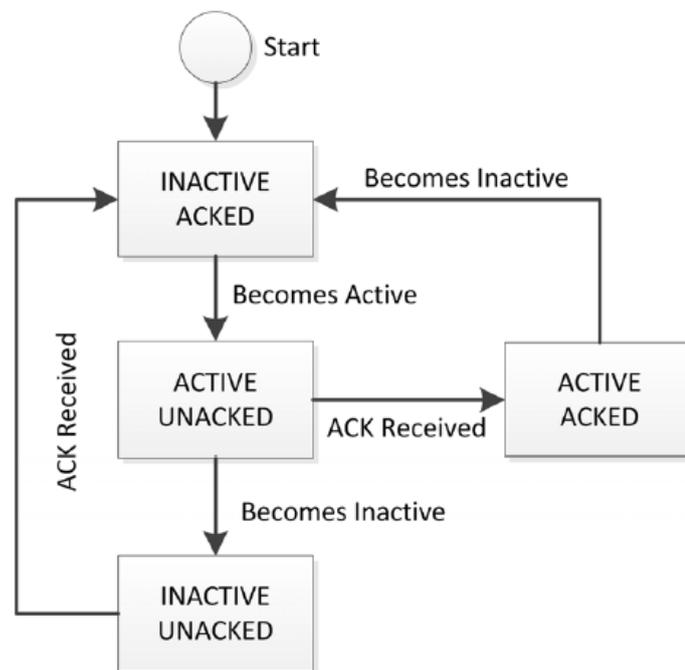


Figura 9 Ciclo de vida dos alarmes de processo. Fonte (LEITÃO, GUEDES e ARAUJO, 2012).

A Figura 9 apresenta em seu estado inicial ativação e reconhecimento do alarme. Uma notificação é enviada para o operador sobre a ocorrência de um alarme. Em seguida já é possível que o operador reconheça esse alarme. No caso de uma variável de processo monitorada por alarme estiver em um estado normal, o alarme permanece em um estado inativo (LEITÃO, GUEDES e ARAUJO, 2012). Um exemplo da variável temperatura que é monitorada, na qual, para cada um dos tipos de alarmes HH (*High High*), H (*High*), LL (*Low low*) e L (*Low*) um evento indicativo é gerado. No trabalho apresentado o cálculo de correlação de eventos é realizado por meio de janelas de tempo, sendo que duas janelas são utilizadas para cada ocorrência de alarme: uma posicionada na ativação e outra na normalização do evento. Esse trabalho, segundo os autores, possibilitou identificar alarmes redundantes e conhecer a ativação dinâmica de alarmes. Como desafio, o trabalho aponta o fato de ser necessário um conhecimento significativo do processo para identificar e estimar o tempo das janelas propostas, e propõe para trabalhos futuros identificar algoritmos que calculem as janelas de tempo automaticamente.

Em (LIMA *et al.*, 2013), é descrita a estruturação de uma ontologia aplicada a uma planta de tratamento de dietanolamina (DEA) a partir de dados de alarmes gerados na planta, buscando inferir sobre a correlação dos alarmes gerados. A ontologia proposta tem como objetivo modelar o domínio de plantas e processos industriais, a fim de dar suporte a aplicações de automação industrial, conforme mostra a Figura 10.

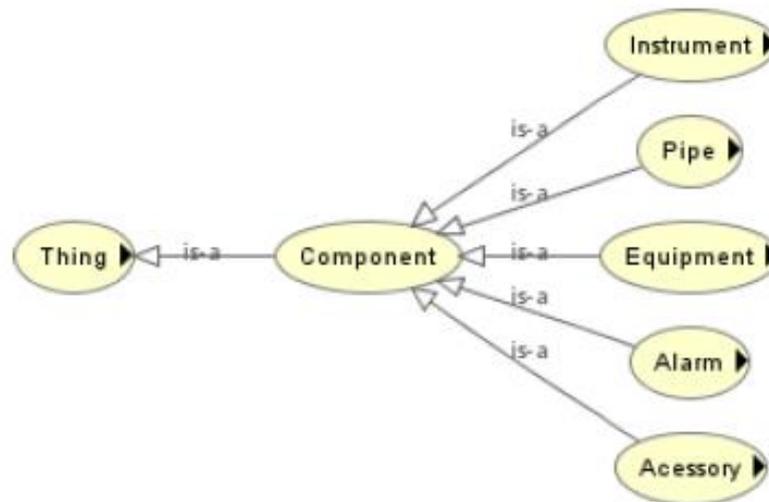


Figura 10 Classes principais da ontologia para planta industrial. Fonte (LIMA *et al.*, 2013).

No trabalho apresentado por (LIMA *et al.*, 2013), a extração de conhecimento é realizada por meio de um algoritmo de ordenação dos alarmes instanciados para planta DEA, onde, segundo os autores, é possível inferir o ordenamento dos alarmes por meio de correlação semântica e filas de ordenação, descrevendo a dinâmica do processo.

No trabalho de (VENCESLAU *et al.*, 2014), é descrito a aplicação de árvore de falhas e ontologia para detecção de situações de falhas nos componentes presentes em um tanque de água que representa um processo industrial. Uma ontologia de domínio é apresentada para representar o conhecimento sobre esse processo semelhante ao apresentado por (LIMA *et al.*, 2013).

Sistemas que detectam uma situação são abordados por (MACHADO *et al.*, 2013). Os autores apresentam uma proposta de desenvolvimento de um sistema inteligente que permita

o gerenciamento de aplicações pervasivas de monitoramento e adaptação, aplicado ao monitoramento de pacientes com necessidades especiais ou limitações físicas, cognitivas, que reduzam a capacidade de realizações de tarefas diárias. Os autores propõem um modelo conceitual para determinar a visão do mundo e acionar ações proativas de acordo com situações detectadas. É proposto um middleware chamado *Situations as a Service* (SIaaS) para gerenciamento de contexto e situações de interesse.

O contexto de interesse é apresentado como a contextualização de entidades e dados, bem como suas relações. A Figura 11 apresenta o modelo apresentado pelos autores.

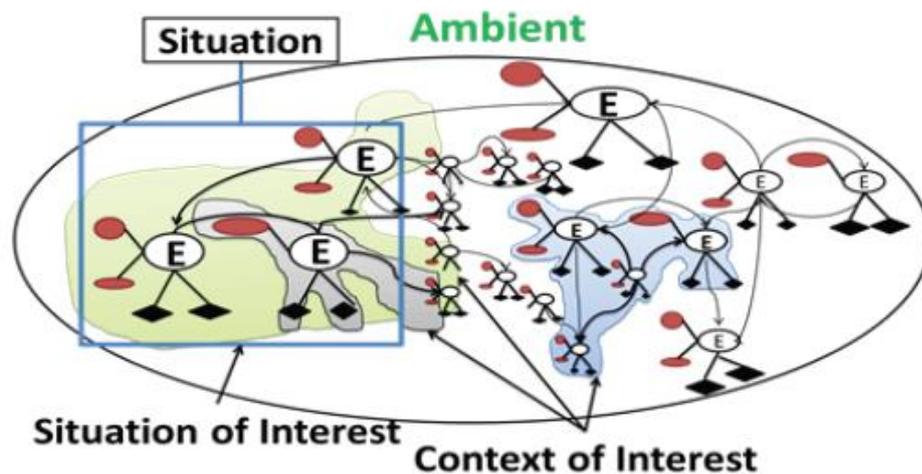


Figura 11 Representação do ambiente, situação, contexto, dados de contexto, sensor e relações semânticas. Fonte (MACHADO *et al.*, 2013).

A proposta dos autores apresenta uma ontologia para descrever o modelo de contexto. As diferentes situações são implementadas por meio de regras SWRL. O trabalho prevê um subsistema de gerenciamento de ações proativas. Para informar o contexto de interesse, foram desenvolvidas instâncias do tipo triplas (Paciente, Sensor, *hasSensor*). Segundo os autores, as principais contribuições são a consciência da situação para ações proativas, sendo essas implementadas por meio de uma abordagem probabilística em Redes Bayesianas, o desenvolvimento de uma arquitetura que provê recursos aplicáveis em uma residência, o uso

de um modelo conceitual para interoperabilidade semântica e a construção de um método que permitiu construir uma aplicação pervasiva genérica para um domínio específico.

No trabalho apresentado por (MACHADO e OLIVEIRA, 2014), é apresentada uma abordagem voltada para recomendação de objetos de aprendizagem. A seleção de recursos em ambientes pervasivos onde se dispõe de um volume significativo de recursos possui o usuário como elemento central, além dos elementos de domínio e de contexto. A união do modelo instanciado com as regras semânticas e um motor de inferência lógicas, implementa a base de conhecimento utilizada para realizar a recomendação sensível ao contexto.

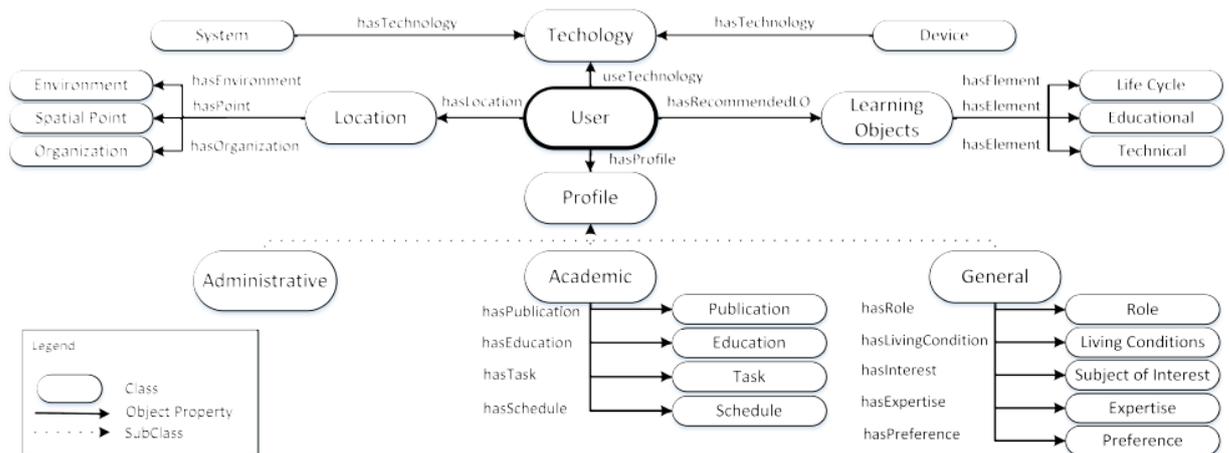


Figura 12 Visão geral das classes do modelo. Fonte (MACHADO e OLIVEIRA, 2014).

Conforme apresentado na Figura 12, os autores propõem um modelo no qual o usuário é elemento central que se relaciona com as quatro informações contextuais identificadas. Um filtro semântico com regras SWRL é desenvolvido para tornar o modelo apto a realizar recomendações, fazendo uma correspondência entre as características de contexto.

Os trabalhos apresentados propõem, em geral, modelos baseados em cenários fictícios e não a um exemplo real de aplicação. Além disso, em sua maioria, os trabalhos mais recentes não objetivam a aplicação industrial. Alguns descrevem contextos para ambientes educacionais ou relacionado ao comportamento do usuário e não de um processo. Alguns

trabalhos apresentam modelos de contexto baseados em ontologias, porém não buscam o uso de técnicas de mineração de dados ou de processos para identificação de padrões de eventos para serem avaliados.

Assim, observa-se que existe uma lacuna com relação à inclusão de um modelo de contexto que defina um conjunto de estados do ambiente e os modelos apresentados nos trabalhos relacionados, que determine o comportamento de uma planta industrial e que permita inferir situações a partir da análise de padrões e que será apresentado por meio da metodologia proposta nesta dissertação.

A Tabela 4 a seguir apresenta um resumo dos trabalhos relacionados comparando as principais dimensões de informação e formalismos semânticos observados para esta dissertação.

Tabela 4 Resumo dos Trabalhos Relacionados.

Trabalhos Relacionados	Dimensão de Informação				Formalismo Semântico		
	Evento	Situação	Estado	Ação	OWL	SWRL	Outros
BERNARAS et al., 1996	X	X	X	X			
COCHRAN, MILLER e BULLEMER, 1996	X	X	X	X			
LAALLAM e SELLAMI, 2007		X			X		
AIZPURÚA, GALÁN e JIMÉNEZ, 2008	X	X	X	X			X
IZADI et al., 2009	X	X	X	X			
SIHEM, TAREK e RINGWOOD, 2009					X		
AGUIAR, ALMEIDA e MEIRA, 2010	X	X		X			
LEITÃO, GUEDES e ARAUJO, 2012	X	X	X	X			
LIMA et al., 2013	X	X	X	X	X		
MACHADO et al., 2013	X	X	X	X	X	X	

VENCESLAU et al., 2014	X	X	X		X		
MACHADO e OLIVEIRA, 2014	X	X			X	X	

Fonte: Autor.

Observa-se a partir da Tabela 4 que a maioria dos trabalhos relacionados citam as principais dimensões de informação de interesse para esta dissertação. Com relação ao formalismo semântico, somente os trabalhos mais recentes apresentaram implementações aplicáveis a este estudo.

4 METODOLOGIA PROPOSTA

Este trabalho está relacionado à proposta de (BERNARAS *et al.*, 1996) que modela uma ontologia para formar uma base de conhecimento, mas é baseado em formalismos mais atuais e que possibilitam o reuso, conforme apresentado nos trabalhos de (LIMA *et al.*, 2013) e (MACHADO e OLIVEIRA, 2014). A mineração de dados do processo será implementada a partir das técnicas apresentadas no Capítulo 2. A proposta para esta dissertação é modelar uma ontologia e determinar um conjunto de regras semânticas que possam ser aplicadas para determinar uma situação.

4.1 DESCRIÇÃO DA PROPOSTA

Esta seção descreve a metodologia desenvolvida para implementação da proposta de elaboração de um modelo conceitual de contexto que possibilite inferir diferentes situações a partir de instâncias obtidas de uma base de dados de alarmes do processo. A metodologia é representada na Figura 13.

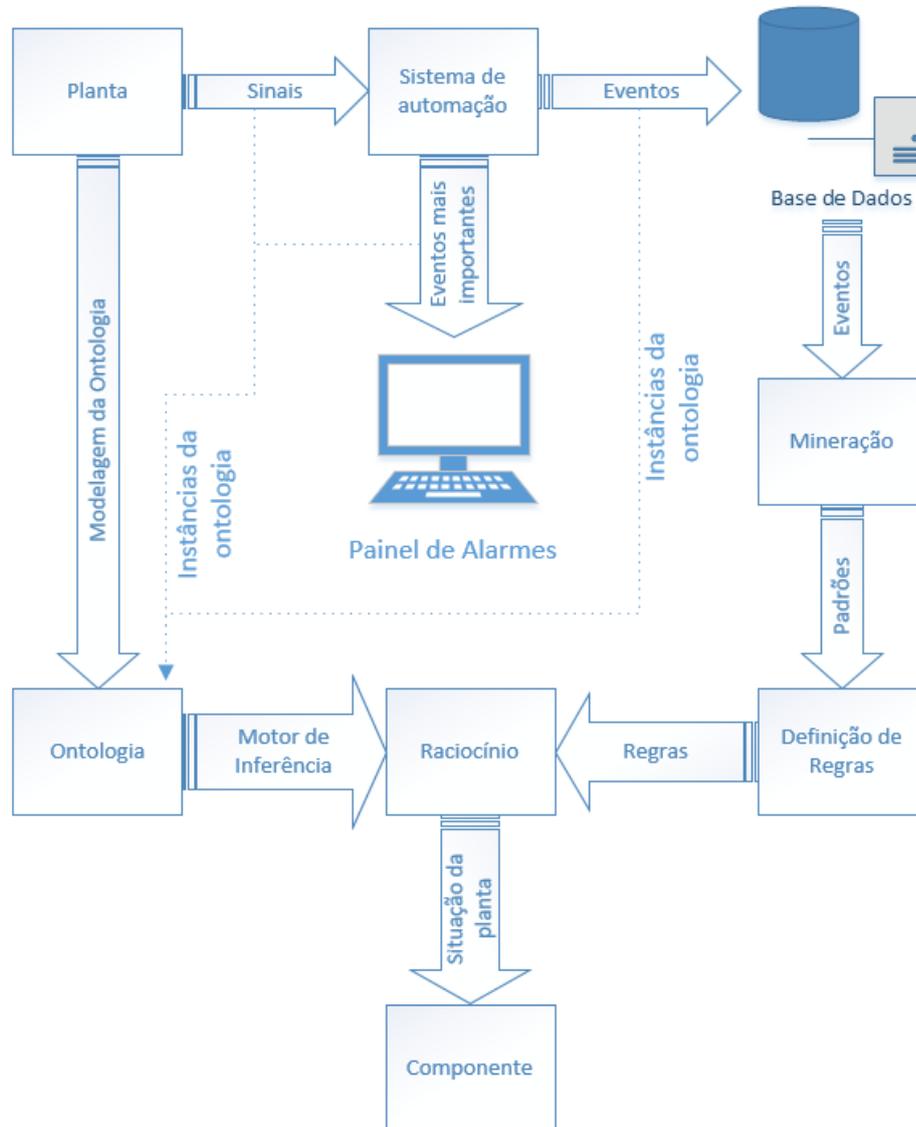


Figura 13 Visão geral da abordagem desenvolvida. Fonte: o autor.

A mineração dos eventos presentes na base de dados permite a identificação de padrões e análises estatísticas onde, a partir da planta industrial são analisados os eventos ocorridos no processo. Esses padrões de eventos possuem as informações necessárias para, a partir de um modelo conceitual implementado por meio de uma ontologia, caracterizar uma situação da planta.

O estudo da base de dados inicia-se na etapa da seleção de dados. É necessário uma etapa de filtragem pois nem todos os dados são necessários. Assim, ferramentas de mineração

são aplicadas para filtragem e análise dos dados da base. Existem diversas técnicas que implementam esta etapa. Nesta dissertação é aplicada a técnica de agrupamento de sequências de eventos em função de um identificador. Considerando os conceitos apresentados no Capítulo 2, a primeira parte deste estudo consistiu em identificar os eventos de acordo com suas descrições, prioridades e tipos. Além disso, esses eventos foram classificados considerando um identificador, que agrupa um conjunto de dispositivos. A partir da seleção dos identificadores é possível classificar os eventos ocorridos no processo.

A configuração da planta analisada, conforme mostra a Figura 14 a seguir, apresenta camadas onde os blocos representam os dispositivos de campo instalados e os compostos representam múltiplos blocos de um mesmo equipamento. Os blocos e os compostos são ativados conforme definido na configuração realizada durante o comissionamento da planta ou nas intervenções de manutenção, de acordo com o funcionamento do processo que pode estar ligado a um ou mais blocos.

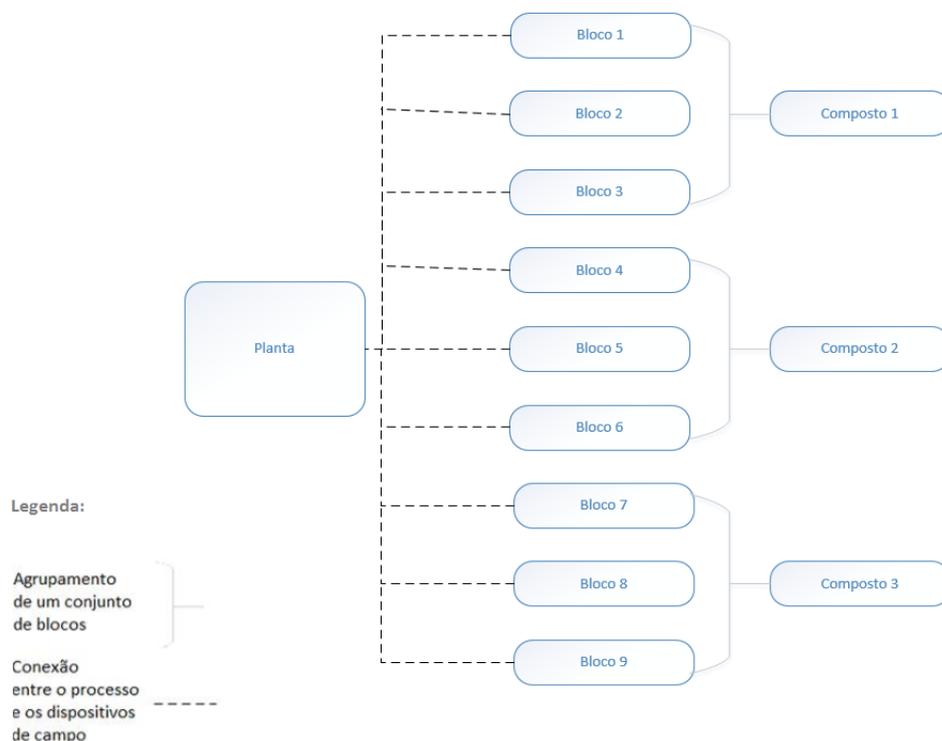


Figura 14 Estrutura do processo. Fonte: O autor.

A utilização da estrutura em blocos e compostos apresentados na Figura 14 permitirá a mineração dos dados da planta a partir das ferramentas acadêmicas disponíveis na literatura. O composto representa um agrupamento de um ou mais blocos e pode ser entendido como um equipamento, enquanto que os blocos representam os dispositivos ligados a esse equipamento. Dessa forma, os eventos que ocorrem na planta estão primeiramente associados aos dispositivos. Contudo, a tarefa de analisar dispositivos individualmente, embora seja possível, torna o trabalho mais demorado. Assim, agrupasse os eventos ocorridos na planta, relacionados aos equipamentos de interesse, permitindo otimizar essa tarefa.

Os sistemas de alarmes, em geral, são configurados para apresentar eventos individualmente em forma de lista e armazenados em uma base de dados. Este é um dos pontos observados neste estudo. Diferentemente de uma situação comum, com consulta manual à base de dados, que geralmente consome um tempo significativo e nem sempre retorna uma resposta adequada quando se busca, por exemplo, pelas causas do problema ocorrido no processo, a proposta desta dissertação estabelece como critério a classificação dos eventos a partir de características como prioridade e tipo. Em seguida, uma sequência de eventos classificados são apresentados em padrões a partir da aplicação de técnicas de mineração de dados do processo. A partir da identificação de eventos em um mesmo local semântico, tem-se uma correlação por meio de um padrão que é identificado a partir de mineração de processo, e isso forma uma especificação lógica de uma situação.

Também, a partir da base de dados, é possível importar instâncias da base de dados para ontologia, analisar padrões e extrair dados estatísticos, que podem ser comparados às métricas apresentadas pela norma ISA 18.2 e o guia EEMUA 191.

A base de dados de alarmes e eventos de processo possui um histórico de informações pouco explorado devido as características dos atuais sistemas de automação. O alinhamento entre os dados de informação de contexto e de domínio específico escolhido para esta

dissertação, permite criar regras para as situações que serão reconhecidas. O desenvolvimento do modelo por meio de ontologia foi definido a partir da fundamentação teórica apresentada na Seção 2.5. A ontologia é implementada a partir do modelo conceitual de contexto desenvolvido como proposta para esta dissertação.

O processo de construção de uma base de conhecimento necessita primeiramente ter uma visão completa do domínio. A identificação de conceitos chave que descrevem formalmente um domínio é formada por uma lista finita de termos e relacionamentos. O modelo conceitual começa com entidades que são conceitos utilizados para raciocinar sobre o domínio de interesse e implementado por meio de ontologia e de regras de inferência. A planta é representada por um conjunto de entidades e suas relações semânticas que caracterizam o contexto. O contexto de interesse é um subconjunto de instâncias com as suas relações semânticas. Quando estas relações são avaliadas é possível caracterizar uma situação por meio de raciocínio conforme mostra a Figura 13, ou seja, ao invés de analisar todos os eventos que são gerados em uma planta em função, por exemplo, da vibração de uma unidade de geração de energia, o sistema pode inferir relações entre predicados de contexto com o objetivo de gerar informações úteis, ou seja, conhecimento sobre um domínio de interesse ligado a um componente.

As informações são interpretadas em um nível conceitual mais alto, como por exemplo o caso em que surge um alerta para bloqueio da turbina, que caracteriza uma situação em um determinado contexto. O contexto de interesse é identificado a partir de informações produzidas pelos próprios eventos, como por exemplo, o evento de unidade sincronizada.

O formalismo definido até esta etapa não permite a realização de raciocínio sobre as instâncias definidas por (NOY, N. F.; MCGUINNES, D. L., 2001), obtidas da base de dados. Para isso, torna-se necessário a implementação de uma linguagem de regras semânticas para determinação da situação da planta. Esse raciocínio é permitido através da inter-relação entre

o formalismo da ontologia e de uma linguagem em um formato específico (HORROCKS *et al.*, 2005). Assim, o modelo instanciado de contexto é combinado com regras semânticas, aplicando um motor de inferências lógicas, formando uma base de conhecimento.

Na avaliação de padrões ocorre a caracterização da situação onde as regras de inferência são processadas a partir do acesso integrado das informações da ontologia. Para a determinação da situação da planta é utilizada uma linguagem baseada em regras que permite a realização de raciocínio sobre as instâncias da ontologia definidas na base de conhecimento, as quais constituem, neste trabalho, dos predicados de contexto necessários para avaliação da situação da planta. Com isso, tem-se um nível mais alto, que indica a situação da planta.

Dessa forma, para estruturar um conhecimento são analisados padrões obtidos a partir de uma etapa de mineração de dados de processo em uma base de dados real. A ontologia proposta é implementada por meio de uma linguagem formal que representa a informação semântica e permite gerenciar as situações da planta por meio de uma representação padrão.

As relações semânticas entre os contextos de um domínio podem ser exploradas a partir de uma base de conhecimento. Nesta dissertação, o conhecimento de domínio é considerado como específico para os contextos existentes. A partir dos diferentes estados da planta, é possível especificar entre os contextos de interesse, um determinado estado que será disjunto a outro.

Para possibilitar o raciocínio, é necessário um formalismo que irá permitir gerar conhecimento através do ambiente modelado. A situação pode ser entendida como um estado particular da planta. Ela pode ser captada a partir de dados de sensores e é interessante para os sistemas de gerenciamento de alarmes, pois podem ser úteis para permitir a indicação de ações a serem tomadas.

O mecanismo de raciocínio possui entre seus processos a identificação de situação, que deriva uma situação por meio de interpretação ou fusão de várias partes do contexto.

Assim, é possível identificar a situação da planta o que permite a partir das classes e propriedades modeladas, inferir novas informações.

4.2 CONTEXTOS DE INTERESSE

Os contextos de interesse definidos para essa dissertação estão divididos em três estados na planta. O estado normal é o que exige o menor número de ajustes de acordo com o guia EEMUA 191. O estado *Upset* exige maior atenção e possui como característica a região onde ocorre o maior número de ações humanas, que não serão analisadas neste trabalho. Por fim, o estado Shutdown caracteriza uma situação atual reativa da planta onde ações automatizadas de segurança são necessárias (EEMUA, 1999).

4.3 EVENTOS AVALIADOS

Os eventos avaliados a partir da base de dados foram classificados de acordo com os diferentes contextos apresentados a seguir. A partir dos objetivos para sistemas de alarmes definidos pelo guia EEMUA 191 (EEMUA, 1999) e da proposta apresentada por (MACHADO *et al.*, 2013) é apresentada a representação proposta para esta dissertação. Na planta surge a situação atual que é iniciada e terminada por eventos (E), conforme representado na figura abaixo. Situações preditivas podem ser identificadas a partir de uma etapa de predição, embora não apresentadas neste estudo.

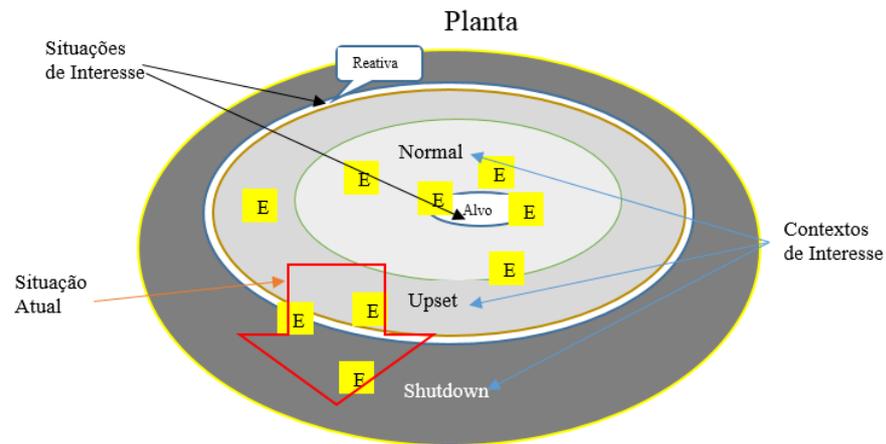


Figura 15 Representação dos diferentes contextos e situações presentes na planta. Fonte: O autor.

Conforme mostra a Figura 15, a situação de interesse reativa pode ser vista como um limite entre dois contextos de interesse no sentido *Upset* para *Shutdown*, sendo necessárias ações de emergência (ESD).

Existem diferentes casos que são identificados a partir de eventos ocorridos em uma unidade de geração térmica de energia elétrica, como por exemplo: sobre velocidade do conjunto turbina-gerador, temperatura nos mancais de escora, vibrações, pressão no condensador, entre outros. Com base nessas informações o operador deve bloquear a turbina caso identifique falha no automatismo.

A situação de interesse alvo pode ser vista como uma região central ao contexto *Normal*. Nesta situação não são necessárias ações sobre a planta.

A análise de eventos e a caracterização das situações a partir de dados de contexto permite formar uma base de conhecimento para inferir situações futuras, como por exemplo na detecção de fraudes ou no congestionamento do trânsito, conforme apresentado por (ETZION e NIBLETT, 2010), onde a fase de detecção de padrões, que antecede a predição, pode ser confrontada com dados históricos.

No próximo capítulo será descrito o modelo conceitual de contexto proposto para essa dissertação.

5 MODELAGEM CONCEITUAL DE CONTEXTO

Esta seção apresenta o desenvolvimento do modelo conceitual de contexto. Neste trabalho, o modelo de contexto possibilitará por meio de regras semânticas determinar se uma situação está acontecendo no tempo atual com base em dados de contexto de dimensão externa. Assim, o modelo instanciado juntamente com um motor de inferências lógicas e as regras semânticas geram a base de conhecimento.

5.1 DESENVOLVIMENTO DO MODELO CONCEITUAL

Neste capítulo é apresentada a visão específica do trabalho a respeito da modelagem conceitual desenvolvida, buscando caracterizar o contexto e as situações de interesse na planta industrial. As figuras que apresentam o modelo conceitual estão em inglês pois descrevem a ontologia que é implementada, sendo essa a forma escolhida para apresentação.

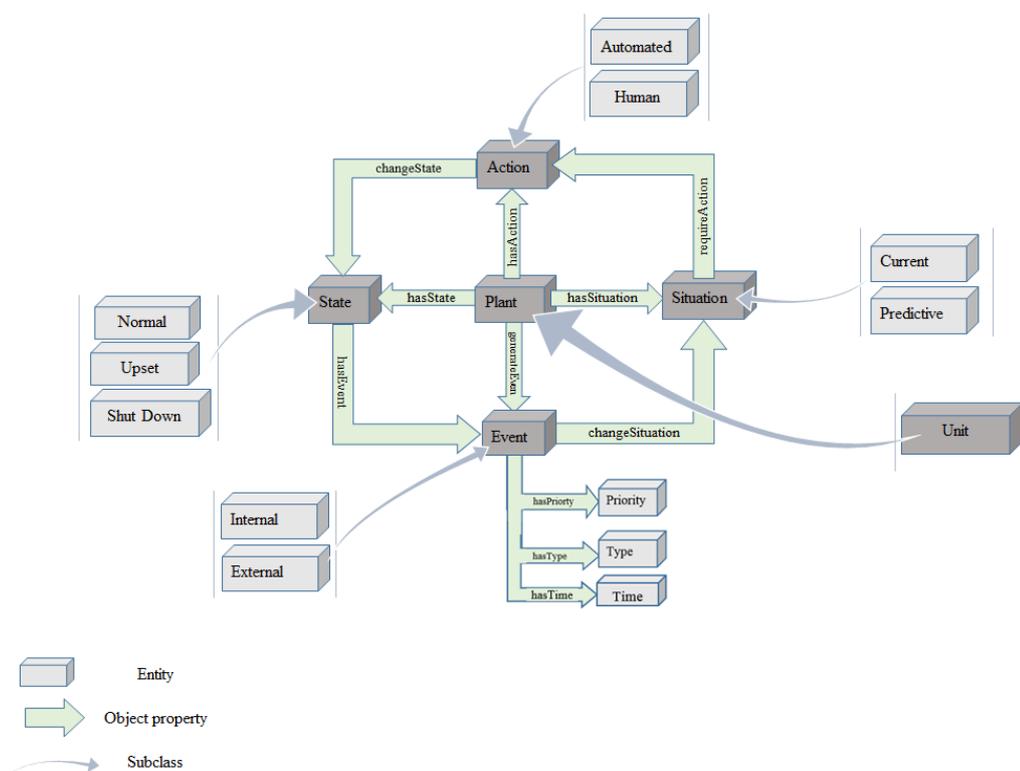


Figura 16 Visão parcial do modelo ontológico conceitual desenvolvido. Fonte: O autor.

O modelo ontológico conceitual proposto tem como elemento central a planta conforme mostra a Figura 16. As dimensões são descritas como segue. A dimensão da planta (*Plant*) é composta por uma ou mais unidades, conforme apresentado na Figura 4 e tem por objetivo a realização do processo para o qual foi projetado. A unidade modela as informações necessárias sobre a coleção de equipamentos em que um ou mais passos do processo podem ser realizados, bem como instrumentação, alarmes entre outros. A dimensão ação (*Action*) modela informações sobre as ações requeridas pela planta, representando os conceitos de ações humanas ou automatizadas. Na dimensão estado (*State*) os conceitos que compõem essa dimensão representam os estados descritos na EEMUA 191 (EEMUA, 1999). A sub dimensão *normal* está prevista quando mínimos ajustes ou intervenções na planta por parte da operação são necessários. A sub dimensão *upset* está prevista para os casos em que o operador deverá fazer intervenções que permitam o retorno da planta para o estado *normal*. Por fim, a sub dimensão *shutdown* ocorre quando uma ação de segurança prevista na planta é necessária. Na dimensão evento (*Event*) os eventos estão divididos em duas sub dimensões: eventos internos representam as ações automatizadas previstas para sistema de alarmes; eventos externos que representam ações humanas que são externas ao sistema. A dimensão prioridade (*Priority*) indica a prioridade definida para os diferentes eventos e auxilia na operação do processo. A dimensão tipo (*Type*) é necessária para definir os diferentes tipos de eventos que ocorrem na planta. A dimensão tempo (*Time*) permite analisar a ordem de ocorrência dos eventos durante o processo. A dimensão situação (*Situation*) é o mais alto nível de abstração do modelo, que está dividida em duas sub dimensões, atual (a partir de mecanismos com regras de inferência) e preditiva (definida no modelo de contexto, mas que não será apresentada neste trabalho). Assim, a ontologia desenvolvida acessa os elementos de contexto e raciocina sobre estes dados para definir a situação corrente.

Para facilitar a disseminação do modelo e das regras associadas, o modelo e a implementação estão em inglês. Essa modelagem de contexto considera o reuso, a flexibilidade, a interoperabilidade e o compartilhamento, sendo os requisitos para representação do conhecimento. Uma ontologia analisada para reuso durante esse desenvolvimento, descrita no projeto OntoSafe (NATARAJAN, GHOSH e SRINIVASAN, 2012), apresenta uma aplicação para processos de supervisão para indústria química, propondo um gerenciamento semântico de situações anormais para indicar o estado do processo. Contudo, trata-se de uma ontologia genérica e optou-se por outras ontologias.

O modelo conceitual foi construído com base nos conceitos apresentados nas ontologias de (MACHADO *et al.*, 2013; MACHADO *et al.*, 2014; LIMA *et al.*, 2013). Além disso, para definição das classes foram considerados os termos presentes na norma (ANSI/ISA-5.1-2009), que define classes de instrumentos, sistemas de instrumentação e funções utilizadas para monitoramento, medição e controle.

5.2 ANÁLISE DO MODELO CONCEITUAL

O uso de máquinas de inferência permite a verificação da consistência do modelo. Essa análise é realizada logo após a implementação da ontologia. Somente após essa verificação é que são criadas as instâncias da ontologia. A modelagem semântica é utilizada para produzir uma semântica formal para as entidades de contexto e descrever entidades complexas de contexto que não poderiam ser representadas por outras linguagens (BETTINI *et al.*, 2010).

5.3 AGREGAÇÃO DO CONTEXTO

Relações semânticas $\{R\}$ formam um contexto, conforme apresentado por (MACHADO *et al.*, 2013; YE *et al.*, 2015), sendo representadas por triplas do tipo $\langle Es, p, Eo \rangle$. Tanto sujeito (Es) quanto objeto (Eo), presentes na tripla e representadas por entidades instanciadas presentes no ambiente, são encapsuladas por um predicado de contexto, de forma que sujeitos e objetos podem ser representados por regras de inferência.

Por exemplo, o contexto de interesse, conforme apresentado por (MACHADO *et al.*, 2013), pode ser definido como um grupo de relações semânticas representadas como triplas:

$$\{R\} = \begin{aligned} & \langle \text{Unit}, \text{hasDevice}, \text{Sensor_Nível} \rangle \\ & \langle \text{Alarme_Sensor_Nível}, \text{hasPriority}, \text{Priority_1} \rangle \\ & \langle \text{Alarme_Sensor_Nível}, \text{hasType}, \text{LLABS} \rangle \end{aligned} \quad (1)$$

Assim, o conjunto de instâncias com relações semânticas define o contexto que é a informação de interesse.

Na Figura 17, a planta como elemento central se relaciona com quatro dimensões de informação contextuais identificadas. Assim, uma planta tem estado (*hasState*), tem situação (*hasSituation*), tem ação (*hasAction*) e gera evento (*generateEvent*).

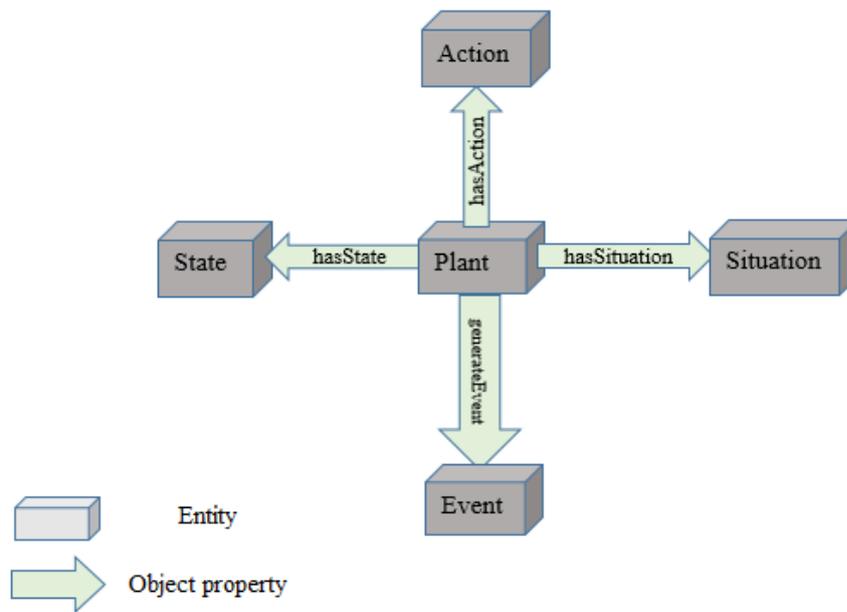


Figura 17 Visão parcial do modelo desenvolvido. Fonte: O autor.

A dimensão situação (*Situation*) possui como propriedade requer ação (*requireAction*). A dimensão ação (*Action*) muda estado (*changeState*). A dimensão estado (*State*) tem evento (*hasEvent*). A dimensão evento (*Event*) muda situação (*changeSituation*), pois toda situação é iniciada e finalizada por um evento, tem prioridade (*hasPriority*), está localizado no tempo (*hasTime*) e tem tipo (*hasType*). Na Figura 18 são apresentadas dimensões descritas acima.

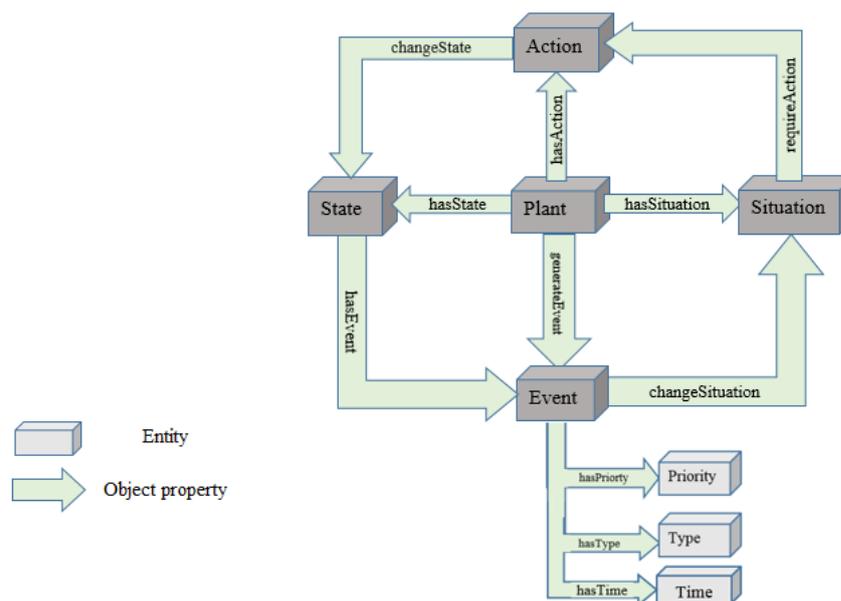


Figura 18 Visão parcial do modelo desenvolvido. Fonte: O autor.

A subdimensão unidade (*Unit*) tem alarme (*hasAlarm*), tem aparatos (*hasApparatus*), tem equipamento (*hasEquipment*), tem tubulação (*hasPipe*) e tem dispositivos (*hasDevice*). Assim, na Figura 19 é apresentado o modelo conceitual proposto para esta dissertação.

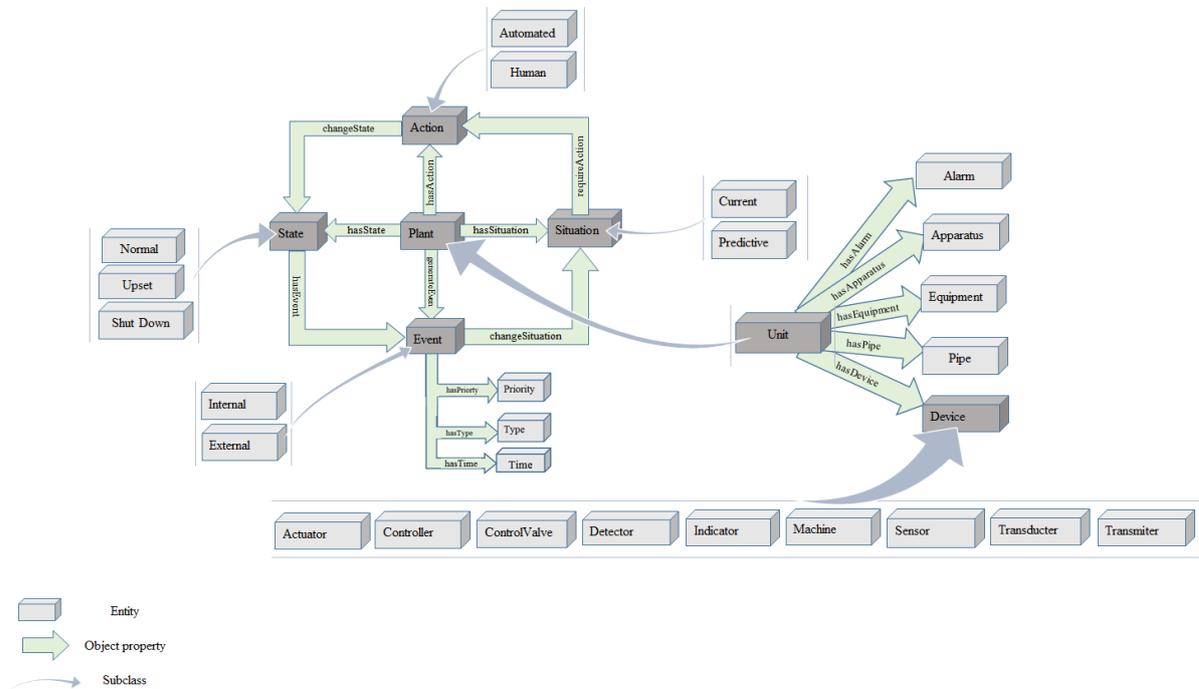


Figura 19 Visão geral do modelo desenvolvido. Fonte: O autor.

As Ações Humanas e Automatizadas são representadas por uma relação semântica {R} e observada a partir de eventos registrados na base de dados. Essas ações são aplicadas à planta e modificam o seu estado.

Os eventos são representados segundo (MACHADO *et al.*, 2014), pela seguinte declaração:

$$\text{Evento} : (\text{nome}, \text{tipo}, \text{tempo}, \{R\}, p) \quad (2)$$

Os eventos podem ser do tipo externo ou interno, estão situados no tempo (*hasTime*) utilizado como conceito discreto, possuem um conjunto de relações semânticas {R} e devem conter um padrão (*p*), no caso de eventos complexos. Dados de sensores são considerados instâncias de entidades de contexto, ou seja, eventos simples.

A partir de relações semânticas, é possível observar uma parte do contexto, descrito como contexto de interesse. Dessa forma, um conjunto de instâncias, que são entidades contextualizadas ligadas por suas relações semânticas, forma o contexto de interesse.

Nessa dissertação foram consideradas para análise do estudo de caso as situações a partir dos conceitos analisados no trabalho de (MACHADO *et al.*, 2014). Segundo o trabalho dos autores, eventos podem determinar a evidência do início e do fim da situação atual, bem como influenciar uma situação preditiva. A declaração a seguir representa o modelo de situação atual, definida nesta dissertação como a situação corrente.

$$Sa : (\textit{nome}, Ie, Fe) \quad (3)$$

A partir da definição apresentada em (3), onde situação atual (Sa) possui um nome, além de dois eventos. Esses eventos são observados a partir do início (Ie) e fim (Fe).

Nesse capítulo, foi apresentado o modelo conceitual proposto e as definições adotadas a partir dos conceitos analisados nos capítulos anteriores. No próximo capítulo será apresentada a implementação da proposta desta dissertação.

6 IMPLEMENTAÇÃO DA PROPOSTA

Este capítulo descreve as etapas desenvolvidas a partir dos conceitos apresentados nos capítulos anteriores para posterior aplicação no estudo de caso proposto. Para facilitar o entendimento, na sequência o estudo de caso será brevemente descrito.

6.1 DESCRIÇÃO DO ESTUDO DE CASO

Para esta dissertação são avaliadas informações relacionadas aos eventos armazenados em uma base de dados industrial de uma usina térmica de geração de energia elétrica. Este estudo de caso concentra-se na análise dos eventos relacionados ao conjunto turbina-gerador a partir de incidentes relacionados ao bloqueio da turbina e as devidas ações de emergência definidas a partir de documentos de operação, durante o condicionamento, acendimento e sincronismo das unidades de geração de energia elétrica. A utilização deste sistema como estudo de caso permitirá a aplicação e a avaliação da proposta em uma situação real. Vale também salientar que para o estudo da planta e aplicação da proposta tem-se acesso aos dados da base e o conhecimento sobre o processo.

Este sistema é composto por uma base de dados presente em uma planta de geração térmica de energia elétrica (além de dezenas de componentes, os quais não interessam para esta dissertação) que geram cada uma um arquivo de eventos e alarmes. O volume de dados registrados é muito grande, e isso dificulta a análise manual, principalmente quando ocorrem distúrbios na planta, onde um volume significativo de dados é gerado em um curto espaço de tempo.

A proposta deste trabalho pode ser representada a partir da Figura 4, do Capítulo 2, onde a partir de uma base histórica de dados instalada na planta, é extraído um arquivo de eventos e alarmes, que em seguida será realizado o processo de mineração, conforme descrito na Figura 6, cujas atividades envolvem a representação de padrões conforme os dados do

domínio. A mineração dos dados do processo se baseia em instâncias onde um identificador denominado composto representa um conjunto de blocos.

Neste estudo, os eventos do processo armazenados em uma base de dados do sistema supervisorio representam o objeto de pesquisa.

6.2 FERRAMENTAS PARA IMPLEMENTAÇÃO

Os dados disponíveis na base de dados foram analisados por meio de *plug-ins* disponíveis no ProM⁶ (VERBECK *et al.*, 2010; AALST, 2010) e permitiram gerar regras para o modelo desenvolvido no Protégé, permitindo formar uma base de conhecimento a partir do modelo conceitual instanciado.

Diante da grande quantidade de informação, torna-se necessário possuir meios que facilitem a obtenção das informações necessárias. Por meio da ferramenta de mineração de processos ProM é possível realizar as etapas iniciais de seleção, pré-processamento, transformação e mineração de dados do processo apresentado na Figura 6. Além disso, a ferramenta disponibiliza um conjunto de *plug-ins* que permitem a filtragem, clusterização, análises estatísticas e a identificação de padrões dos eventos ocorridos.

A ontologia desenvolvida foi codificada em OWL-DL utilizando o software Protégé *desktop* (STANFORD, 2015). Esse software desenvolvido em Java possui uma versão *web* que juntamente com a versão *desktop* são mantidas por uma comunidade de desenvolvedores, sendo *open-source* e com compatibilidade para os formatos RDF(S) e XML.

O *Reasoning* utilizado para validação de consistência do modelo foi o pellet (SIRIN *et al.*, 2007; DENTLER *et al.*, 2011). Esse motor de inferência, segundo os autores acima, é a mais indicada devido a solidez, ser compatível ao Protégé e ser *open source*. Além disso, apresenta maior expressividade em lógica descritiva.

⁶ ProM é uma marca registrada do Grupo *Process Mining*, Eindhoven Technical University.

Segundo (NOY, N. F.; MCGUINNES, D. L., 2001), os diferentes *frameworks* utilizados na construção de uma ontologia possuem abordagens distintas. No último passo são criadas as instanciadas as classes. Nesta etapa, é utilizado além da inclusão manual o paradigma *Ontology Based Data Access* (OBDA) (CALVANESE, D. *et al.*, 2009) para importar instâncias por meio do *plug-in* Ontop⁷ (KONTCHAKOV, R. *et al.*, 2014). Esse *plug-in* permite o acesso aos dados da base de dados por meio de mapas definidos a partir do modelo conceitual e da ontologia (HARRIS, S.; SEABORNE, A., 2013).

A tecnologia OBDA provê uma visão conceitual sobre uma base de dados por meio de uma ontologia. O *plug-in* Ontop é um sistema de inferência que permite esse acesso por meio regras de mapeamento. Mapeamentos são axiomas que se relacionam com os dados da base e atende as recomendações do W3C. As entradas são a ontologia e o mapeamento e o *reasoner* retorna o resultado da consulta estando conectado a base.

A conexão na base necessita a definição de alguns parâmetros, conforme mostra a Figura 20.



Figura 20 Credenciais para acesso à base de dados. Fonte: Autor.

⁷ Ontop é um projeto *open source* desenvolvido pela Free University of Bozen-Bolzano. <http://ontop.inf.unibz.it/>

A importação das instâncias ocorre por meio do *plug in* Ontop, conforme apresentado na Figura 21.

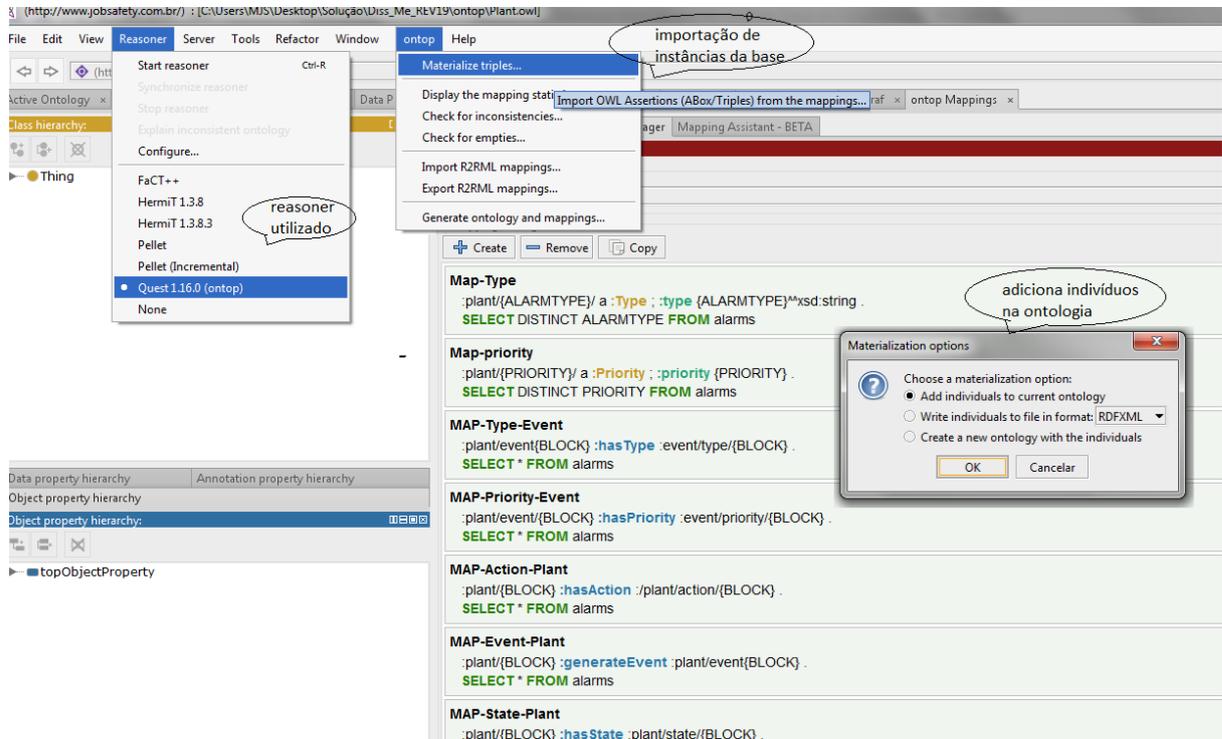


Figura 21 Mapeamento e importação das instâncias. Fonte: Autor.

6.3 MINERAÇÃO DE PROCESSOS

Esta seção descreve o processo de mineração para posterior aplicação no estudo de caso proposto. A mineração de processos (*Process Mining*), está inserida entre mineração de dados e aprendizagem de máquina, que consiste na aplicação de técnicas de análise de dados em uma base de dados, com o objetivo de descobrir padrões que não são possíveis de serem descobertos manualmente. Nessa etapa, foram analisados registros de eventos ocorridos em uma planta de geração térmica de energia elétrica.

O intervalo de geração de energia elétrica apresentado neste exemplo, compreende o dia 23/03/2015 13h00min até o dia 24/03/2015 06h59min.

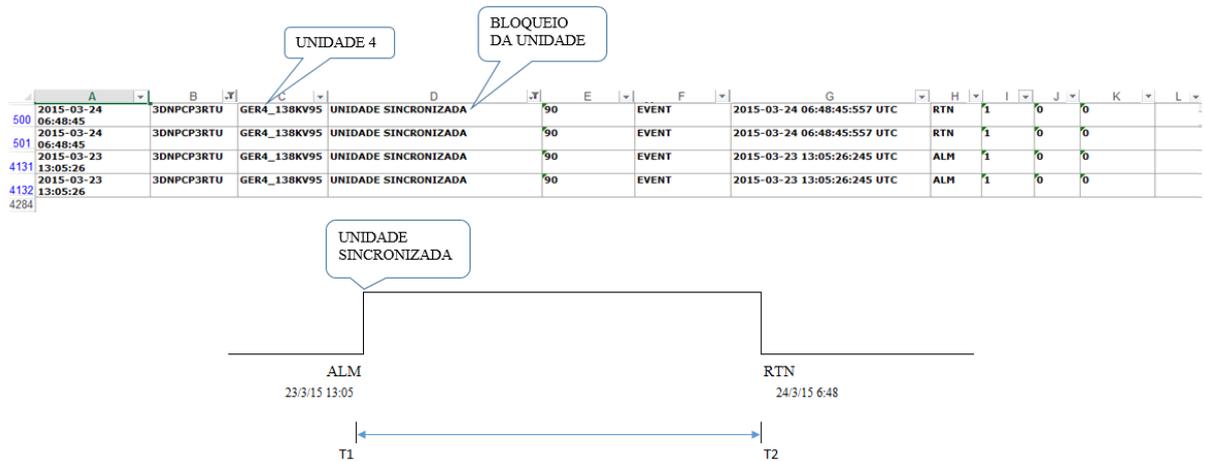


Figura 22 Registro dos eventos. Fonte: O autor.

Cada linha apresentada na Figura 22 corresponde a um evento, e esse evento corresponde a uma atividade que ocorreu no processo. As colunas J e K contém o registro dos valores que ocorreram no processo, enquanto a coluna H indica a transição entre alarme ALM e retorno RTN. O registro é resultante, por exemplo, da passagem por um valor de alarme em um sensor.

A Figura 23 apresenta um exemplo de registro de eventos da mesma base em outro período onde houve vibração no mancal do gerador e a descrição de alguns dados de entrada que podem ser utilizados na etapa de mineração de processo.

	A	B	C	D	E	F	G	H	I	J	K	L
4869	02/04/2015 09:31	3GRECP5TURB	3GREDL415	VIBRAT NOT CORRECT	3GREDL41 STATE	NORMAL		RTN		5		
4870	02/04/2015 09:31	3GSECP4TURB	3GSEMV221	VIBR MANCAL ALTER FRONTAL	HHABS	BLOQUEIO VIB M ALTER FRONTAL		RTN		1	126.594	180 UM
4871	02/04/2015 09:31	3GRECP5TURB	3GSE500UA17	TURBOALT VIBR M ALTO	3GSE500U STATE	NORMAL		RTN		3		
4872	02/04/2015 09:31	3GRECP5TURB	3GSE500UA17	TURBOALT VIBR M ALTO	3GSE500U STATE	ALARME		ALM		3		
4873	02/04/2015 09:31	3GRECP5TURB	3GREDL415	VIBRAT NOT CORRECT	3GREDL41 STATE	ALARME		ALM		5		
4874	02/04/2015 09:31	3GSECP4TURB	3GSEMV221	VIBR MANCAL ALTER FRONTAL	HHABS	BLOQUEIO VIB M ALTER FRONTAL		ALM		1	183.203	180 UM
4875	02/04/2015 09:31	3GSECP4TURB	3GSEMV221	VIBR MANCAL ALTER FRONTAL	HIABS	ADVERTENCIA VIB M ALTER FRONTAL		ALM		2	150.723	120 UM

Figura 23 Registro dos eventos. Fonte: O autor.

Essa amostra de eventos apresentados na Figura 23, ocorridos no processo e relacionados a interrupção da geração de energia, devido a vibração em uma unidade de geração presente na planta, possibilita extrair um conjunto de informações sobre essa situação

típica. Neste caso, pode-se observar que grande parte dos eventos ocorreram no mesmo minuto. Isso pode dificultar a priorização segundo apresentado nos guias e normas da EEMUA e ISA 18.2 detalhadas no Capítulo 2 desta dissertação. Relacionando somente a situação reativa identificada, foram gerados pelo menos três alarmes durante a transição destes eventos, conforme indicam as linhas 4870, 4874 e 4875. Destaca-se a prioridade e tipo dos eventos nas colunas I e F. Observa-se que para analisar dessa forma um volume maior de dados seria necessário um tempo considerável. A seguir será apresentado o resultado do uso de técnicas de mineração aplicada para esse mesmo tipo de situação.

A partir do registro de eventos foram analisadas para a etapa de mineração as seguintes informações: a coluna B apresentada na Figura 23 e é representada pelo composto que será definido como um *Case*. As colunas D e G determinam o nível de detalhe dos passos do processo. As colunas F e I descrevem propriedades específicas que são relevantes para responder questões que se tem sobre o processo, nesse caso definidos como as colunas prioridade e tipo de evento. A coluna A indica o *Timestamp* e permite analisar cada evento na ordem em que ocorreram no processo.

Na seção a seguir, será apresentada a implementação do modelo de contexto instanciado.

6.4 IMPLEMENTAÇÃO DO MODELO

Segundo (MCGUINES e HARMELEN, 2004) que recomenda OWL para o desenvolvimento de ontologias e com base nos dados apresentados na Tabela 3, foram definidos os conceitos relacionados ao domínio. Um dos pontos analisados diz respeito ao reuso de ontologias. Entre as ontologias estudadas, as ontologias definidas em (MACHADO *et al.*, 2013) e em (MACHADO e OLIVEIRA, 2014) foram utilizadas para modelagem de contexto e situação neste trabalho.

A ontologia desenvolvida foi codificada em OWL-DL utilizando o software Protégé.

A Figura 24 a seguir apresenta a estrutura da ontologia implementada.

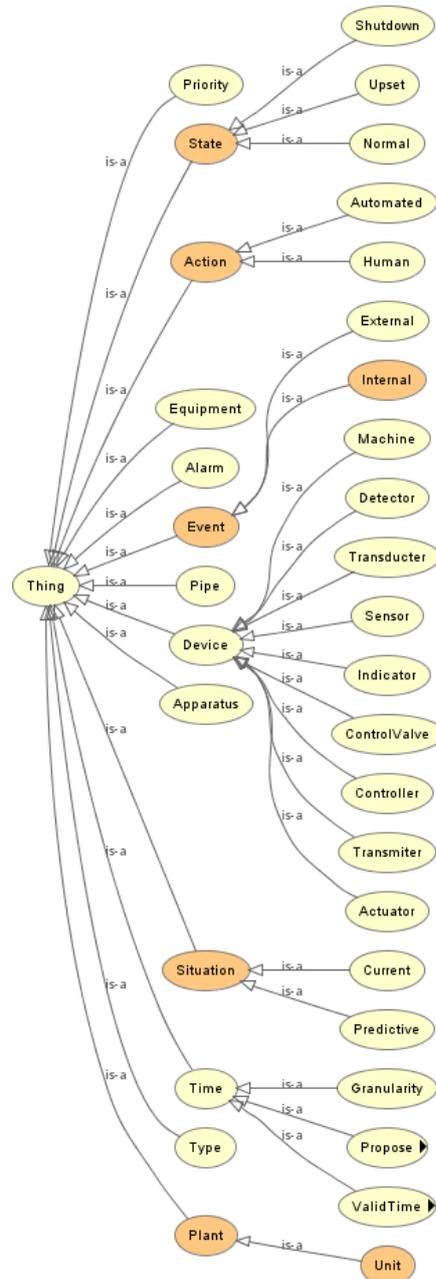


Figura 24 Visão geral da ontologia elaborada pelo plug-in OWL Viz do Protégé.

A ontologia implementada é responsável por acessar os elementos de contexto para definir a situação corrente por meio de raciocínio, sendo utilizado a linguagem OWL que permite descrever classes, propriedades e instâncias, cuja lógica descritiva estrutura um domínio a partir dos conceitos, dos predicados complexos que são representados pelos

indivíduos, e as propriedades que são os relacionamentos existentes entre os conceitos (BERNERS-LEE, T.; HENDLER, J.; LASSILA, 2001).

Para a representação de ontologias utilizou-se como ferramenta o software livre Protégé (STANFORD, 2015), que permite vários formatos de representação, executando um conjunto de estruturas de modelagem de conhecimento. Para inferências de relações de conhecimento baseada em conhecimentos pré-existentes utilizou-se ferramentas de *Reasoning* (SIRIN *et al.*, 2007).

Nessa dissertação, foi escolhido o SWRL para representar a detecção da situação por possuir o seu formalismo fundamentado em OWL-DL, permitindo a modelagem de entidades e restrição de propriedades, permitindo inclusive a análise de possíveis inconsistências no modelo implementado. As regras definidas e uma vez inferidas indicam a situação atual da planta. Tais regras são integradas ao modelo conceitual apresentado, que após ser instanciado é utilizado junto a um motor de inferências lógicas que raciocine sobre regras SWRL para extrair recomendações adequadas ao contexto de interesse.

No Capítulo 2 foram apresentadas as informações para situação segundo (O'BREIN, 2009). O objetivo é focar no modelo conceitual para situações ocorridas em uma usina térmica de energia elétrica. A visão do ambiente é obtida por meio de dados armazenados em uma base histórica de alarmes, aplicando mineração de processo e agregando com entidades do contexto para gerar informação útil (conhecimento).

Nessa dissertação, o raciocínio sobre ontologia ocorre por meio de regras SWRL, pois somente OWL-DL utilizado para implementar a ontologia não permite esse raciocínio. Assim, as regras são adicionadas como axiomas na linguagem OWL-DL e quando são inferidas, indicam uma situação atual (HORROCKS *et al.*, 2005).

Segundo (HORROCKS *et al.*, 2005), as regras em SWRL são definidas na forma de implicação entre um antecedente (corpo) e um conseqüente (cabeça). Uma vez que as

condições especificadas na primeira parte (corpo) forem atendidas, então as condições consequentes (cabeça) devem ser atendidas. Ambas as partes são formadas por átomos que estão presentes nas regras SWRL. A forma descrita por (HORROCKS *et al.*, 2005) $C(x)$, $P(x,y)$, $Q(x,y)$, $sameAs(x,y)$ ou $differentFrom(x,y)$, sendo: C uma classe, P uma propriedade de indivíduos, Q uma propriedade de dados, x e y podem ser indivíduos ou variáveis em OWL e z uma variável ou um valor de dado OWL. A partir da notação definida pelo autor, o modelo é representado por uma tupla I em (4).

$$I = \langle R, EC, ER, L, S, LV \rangle \quad (4)$$

Sendo:

- R é um conjunto de recursos;
- LV um conjunto de valores literais ($LV \subseteq R$);
- EC é o mapeamento das classes e dos tipos de dados para subconjuntos em R e LV ;
- ER é o mapeamento das propriedades para relações binárias em R ;
- L é o mapeamento dos tipos literais para elementos de LV ;
- S é o mapeamento dos nomes de indivíduos para elementos em $EC(owl:Thing)$;

A partir da interpretação OWL descrita em I , uma ligação $B(I)$ é uma interpretação abstrata da OWL que estende I de forma que S mapeia i-variáveis para elementos $EC(owl:Thing)$ e L mapeia d-variáveis para elementos de LV . Assim, um átomo irá satisfazer a ligação $B(I)$ caso atender as condições da Tabela 5.

Tabela 5 Interpretação das condições.

Átomo	Interpretação da condição
$C(x)$	$S(x) \in EC(C)$
$P(x,y)$	$\langle S(x), S(y) \rangle \in ER(P)$

$Q(x,z)$	$\langle S(x), S(z) \rangle \in ER(Q)$
$sameAs(x,y)$	$S(x) = S(y)$
$differentFrom(x,y)$	$S(x) \neq S(y)$

Fonte: Adaptado de (HORROCKS *et al.*, 2005).

A partir da interpretação representada em I e a extensão de I em $B(I)$, é possível definir, seguindo a notação lógica de primeira ordem apresentada, uma regra que determina por exemplo que a planta “x”, está usando sensor de vibração na turbina, está no estado normal, tem um valor que é comparado com um valor de set point, pode ser representado pela regra em (4).

$$\begin{aligned}
& (\forall x, y, z, s, v) \in EC(OWL : Thing) \\
& ((x) \in EC(Plant) \\
& \wedge (y) \in EC(Sensor) \wedge (z) \in EC(State) \\
& \wedge (x, z) \in ER(hasState) \\
& \wedge (y, v) \in ER(hasValue) \\
& \wedge (s, v) \in ER(greaterThanOrEqual) \\
& \rightarrow (x, z) \in ER(isSituation))
\end{aligned} \tag{4}$$

Realizando a tradução da regra em (4) para linguagem SWRL, tem-se a regra:

$$\begin{aligned}
& Plant(?x) \\
& \wedge Sensor(?y) \wedge State(?z) \\
& \wedge hasState(?x, ?z) \\
& \wedge hasValue(?y, ?v) \\
& \wedge swrlb : greaterThanOrEqual(?v, ?s) \\
& \rightarrow isSituation(?x, Reactive)
\end{aligned} \tag{5}$$

As regras desenvolvidas são utilizadas para implementar o raciocínio a partir dos predicados de contexto instanciado determinado a situação corrente da planta. Um predicado de contexto pode ser representado através de variáveis. Assim, no caso do exemplo: $\langle x, hasState, z \rangle$, onde x representa qualquer valor instanciado dentro do domínio da *Plant* e z

qualquer valor instanciado dentro do domínio *State*. Dessa forma, um par de valores relacionados por *hasState* permite validar o predicado.

No próximo capítulo é apresentada a validação experimental e o estudo de caso.

7 APLICAÇÃO DA METODOLOGIA E ESTUDO DE CASO

Este capítulo descreve como o desenvolvimento de uma modelagem conceitual pode ser aplicado em um sistema de alarmes de processo a partir da avaliação dos eventos registrados em uma base histórica de dados e analisados por meio de mineração de processo.

7.1 VALIDAÇÃO EXPERIMENTAL

Múltiplos eventos podem ser analisados a partir do composto (*Case*⁸) apresentado na Figura 14 do Capítulo 4. A primeira etapa de seleção dos dados, conforme descrito por (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996) é implementada por meio dos filtros disponíveis no ProM, permitindo selecionar somente os compostos referentes a uma unidade de geração ou a um equipamento específico.

A plataforma escolhida para pré-processamento e classificação dos dados ProM, apresentado na Figura 25, apresenta uma série de *plug-ins* que podem ser definidos como algoritmos implementados para mineração de processos, sendo uma plataforma aberta e de fácil reutilização de código. Para esta dissertação foram selecionados os *plug-ins* que permitem a identificação do comportamento do processo e a classificação dos eventos.

A etapa inicial de seleção permite a extração de informações com maior facilidade a partir dos filtros disponíveis na ferramenta. Neste exemplo que será apresentado, são avaliados os *plug-ins* disponíveis na ferramenta e que são aplicáveis para esta proposta.

⁸ *Case* é um identificador do processo que permite classificar os eventos (AALST, 2010).



Figura 25 Tela principal do ProM6.

Para a mineração, um arquivo CSV (*Comma-separated Values*) com os dados da base é convertido para um formato específico do ProM denominado *XML-Based standard for event logs* (XES) permitindo o pré-processamento e classificação dos dados da base. O resultado desse processo pode ser analisado pelos *plug-ins* disponíveis na ferramenta. Os dados coletados na planta e armazenados em um historial foram extraídos para um arquivo CSV.

A Figura 26 a seguir apresenta a seleção do composto (*Compound*) que representa um grupo de eventos conforme definido na proposta para esta dissertação.

Figura 26 Tela principal do ProM6.

Para filtrar somente os *Cases* referentes a uma unidade de geração, aplicou-se o filtro *Filter Log on Trace Attribute Values*.

Na Figura 27 é apresentado seleção de todos os compostos referentes à unidade 4, que é caracterizada pela numeração inicial (4) em cada composto, para um período identificado entre a entrada e a saída de sincronismo da unidade de geração de energia.

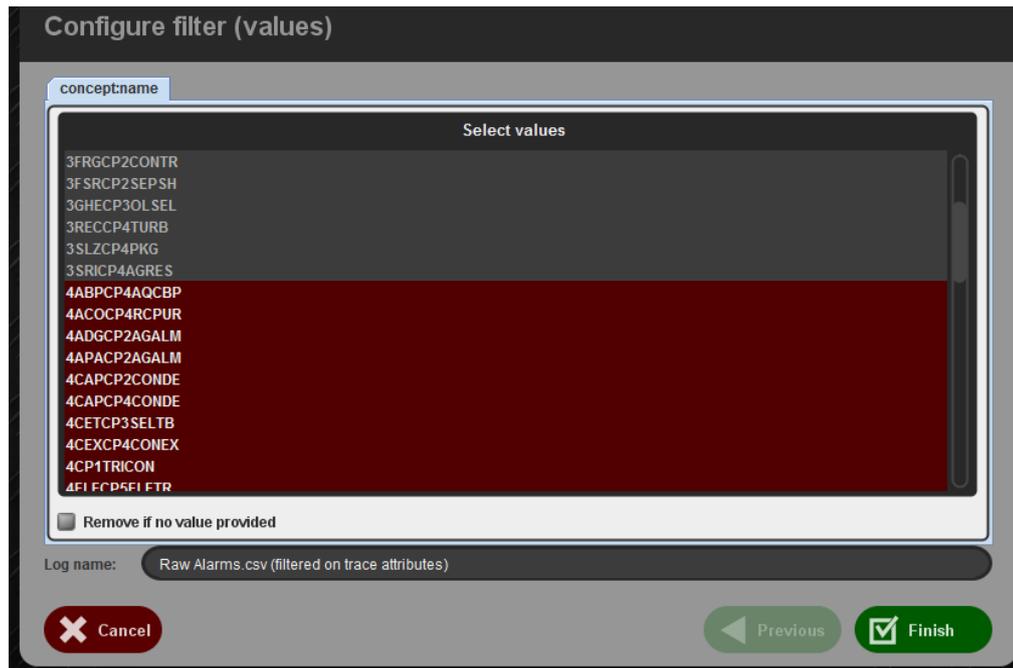


Figura 27 Seleção dos compostos referentes a unidade de geração 4. Fonte: O autor.

Todos os itens selecionados (cor de fundo bordo) representam compostos referentes a uma unidade de geração de energia. O resultado da filtragem apresentada na Figura 27 pode ser observado na Figura 28.



Figura 28 Visão geral do filtro de dados ProM. Fonte: O autor.

Conforme apresentado na Figura 28, o resultado da filtragem realizada para o período descrito no canto superior direito na janela da ferramenta (*Log info*), corresponde ao intervalo de tempo em que a unidade 4 permaneceu em produção. Além desta informação, é possível analisar o total de eventos gerados no período, bem como o número de *cases* (compostos) ativados. Para filtrar somente os *Cases* referentes aos alarmes, aplicou-se o filtro *Filter Log on Event Attribute Values* conforme mostra a Figura 29.

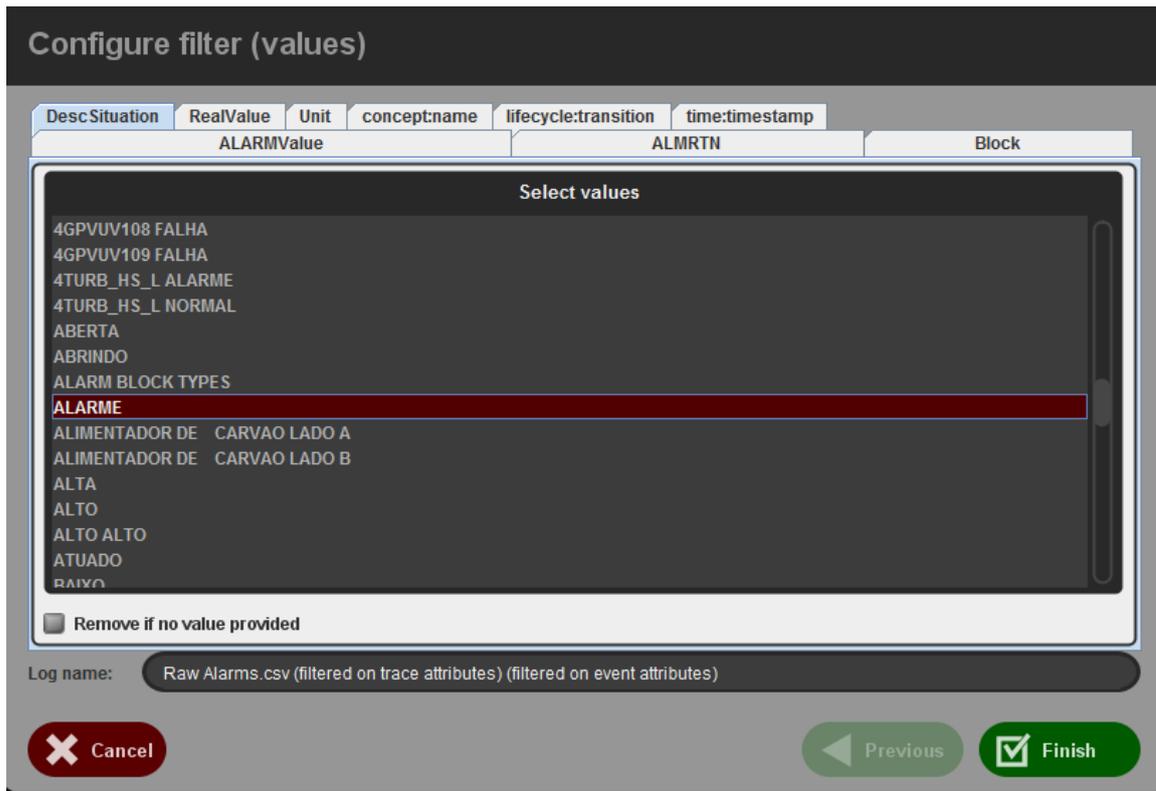


Figura 29 Visão geral do filtro de dados ProM. Fonte: O autor.

A versão 6 do ProM é baseada em *plug-ins* que são algoritmos que podem ser aplicados para mineração de processo proposta para essa dissertação, embora tenham sido concebidos para processos de negócios (AALST, 2010).

Avaliando somente os eventos descritos como alarme, foi possível identificar para o período selecionado um total de 258 alarmes anunciados, conforme resultado do filtro apresentado na Figura 30.



Figura 30 Visão geral do filtro de dados ProM. Fonte: O autor.

Avaliando a partir do *plug-in Dotted Chart* pode-se verificar a indicação temporal relacionada à sequência dos acontecimentos dos alarmes, conforme mostra a Figura 31.



Figura 31 Visão geral do *plug-in Dotted Chart* ProM. Fonte: O autor.

Observa-se na Figura 31 um volume significativo de alarmes nos períodos de início de geração, ocorrido após as 13 horas e no momento em que houve o distúrbio na planta devido a

vibração do conjunto turbina-gerador, ocorrido minutos antes das 7 da manhã e que representam uma avalanche de alarmes.

O *plug-in Dotted Chart* disponível no ProM permite agrupar os eventos ocorridos. Cada evento é apresentado como um ponto em um plano bidimensional (AALST, 2010). Para analisar o registro dos eventos e a identificação de padrões pode-se utilizar um classificador (*Case*), conforme apresentado acima. A Figura 32 apresenta a prioridade e tipo dos eventos.



Figura 32 Visão geral do *plug-in Dotted Chart* ProM. Fonte: O autor.

A partir do *plug-in Dotted Chart* disponível no ProM, é possível observar para o intervalo descrito analisando somente as diferentes prioridades e tipos de eventos ocorridos.

O algoritmo *Show Sequences and Patterns* utilizado neste estudo permite identificar características do processo a partir de execuções que possuem o mesmo comportamento e assim definindo um padrão apresentado por um gráfico sequencial, apresentando a transição e a ligação entre os eventos.

A partir da análise da base de dados é possível identificar a sequência de eventos que antecede ao sincronismo da unidade de geração, conforme apresentado na Figura 33.

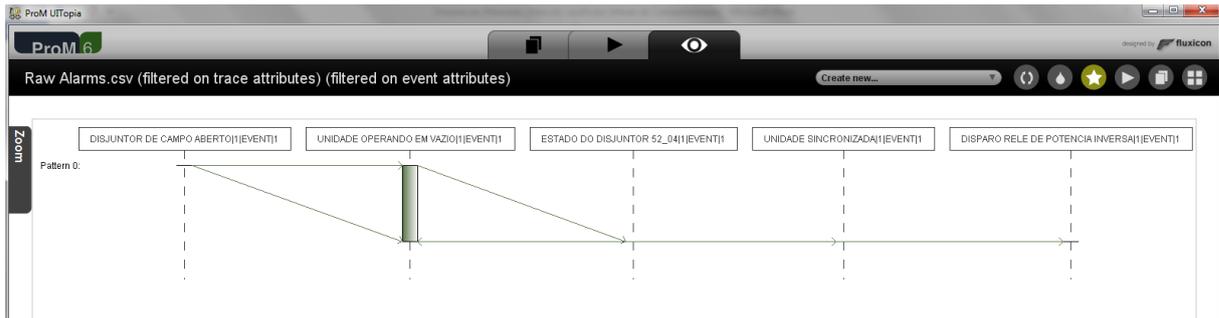


Figura 33 Visão parcial do *plug-in Show Sequence and Patterns* no ProM. Fonte: O autor.

A partir do padrão apresentado na Figura 33 é possível identificar que o evento de sincronismo ocorreu após o relacionamento entre uma sequência de eventos.

7.2 SITUAÇÕES DE INTERESSE

As situações de interesse instanciadas são descritas na Figura 15 como reativa (*Reactive*) e alvo (*Target*). Para este estudo de caso, a regra definida em (6) representa a situação reativa (*Reactive*). Esses dois conceitos são apresentados a partir do agrupamento de eventos ocorridos e analisados por meio dos *plug ins* apresentados.

Assim, foram definidas as situações de interesse *Target* e *Reactive*, suportadas pela ontologia proposta juntamente com as características reativas avaliadas utilizando os padrões que são compostos por dados e informações encontrados após aplicação das técnicas de mineração de processos. Utilizou-se o motor de inferência Jess para processar as regras semânticas. A Figura 34 apresenta uma visão geral da interface de desenvolvimento no software Protégé com o resultado da inferência obtida.

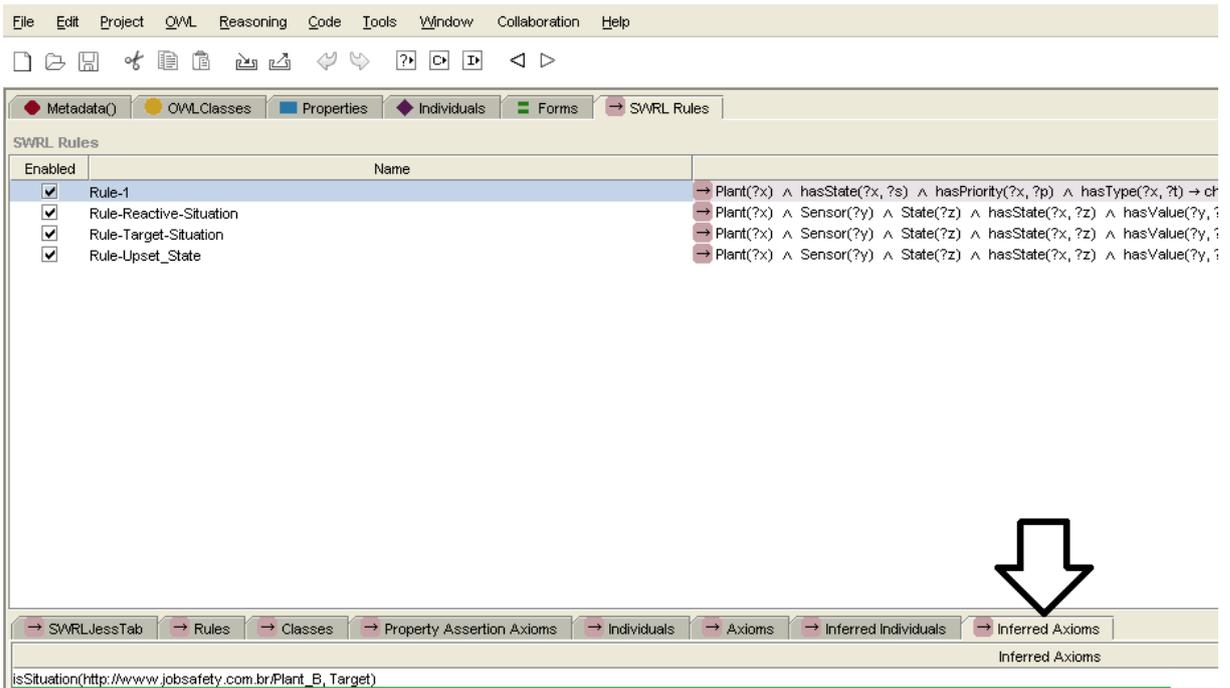


Figura 34 Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.

7.2.1 Estudo de Caso 1

Os padrões permitem identificar as causas das situações indesejadas ocorridas no processo. A Figura 35 apresenta o resultado da aplicação do *plug-in* somente para o composto “4GSECP4TURB”. Este composto (*Case*), concentra entre outros dispositivos (*Devices*) do conjunto turbina-gerador os relacionados ao monitoramento da vibração.

O conjunto turbina-gerador é uma instalação estratégica em um sistema de geração de energia elétrica, que dispõe de um monitoramento por meio de sensores que detectam as manifestações de vibração.

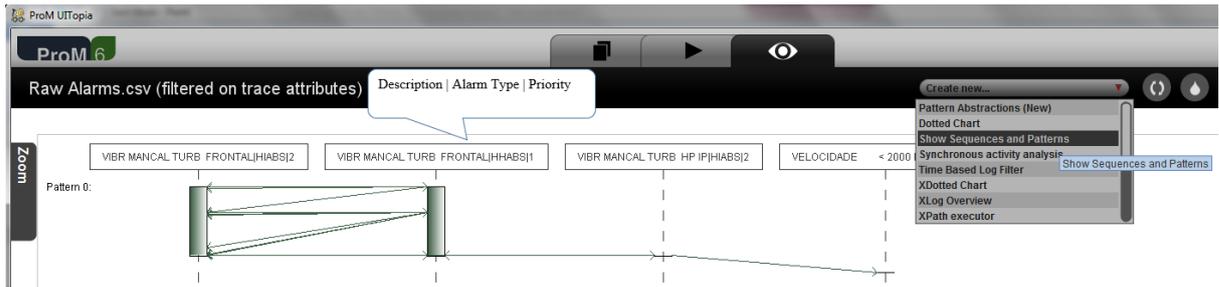


Figura 35 Visão parcial do *plug-in Show Sequence and Patterns*. Fonte: O autor.

A partir da análise da Figura 35 é possível observar a descrição, o tipo e a prioridade dos alarmes de vibração ocorridos no período selecionado. A partir da técnica de mineração apresentada foi possível extrair um padrão da base de dados, permitindo a elaboração de regras de inferência.

A situação atual que surge neste ambiente pode ser vista como uma sequência de eventos, tendo pelo menos um evento ultrapassando a zona limite entre os contextos *Upset* e *Shutdown* e que valida a situação de interesse reativa. Para este estudo de caso, foram analisadas situações em que ocorreram vibrações nas diferentes unidades de geração, suficientes para interromper a geração. Assim, considerando o valor dos diferentes sensores de vibração instanciados como eventos simples, suficientes para a detecção da situação atual, por meio da inferência do modelo conceitual.

A partir da sequência de eventos apresentada e da definição em (5) é possível determinar a situação atual reativa.

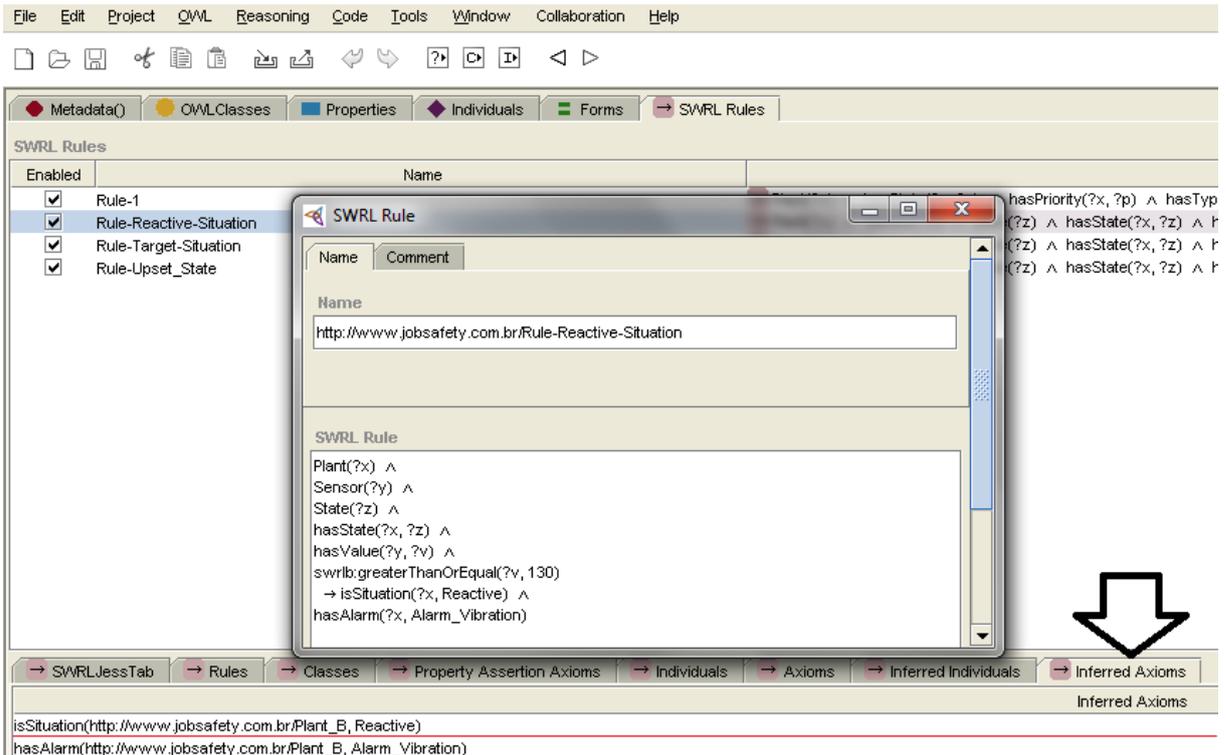


Figura 36 Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação atual. Fonte: O autor.

Conforme apresentado na Figura 36, a regra SWRL determina que se a planta estiver com valor do sensor de vibração maior ou igual a 130 micrômetros, é inferido que a planta “B” está em uma situação atual, e isso determina a situação corrente que é o *Trip* por vibração. Neste caso é considerada a ação automatizada executada diretamente pelos dispositivos de segurança que levam a planta para *Shutdown*. Neste caso, os eventos que são gerados como registro do acontecimento são representados como assertivas que são dadas por um axioma.

O contexto *Shutdown* representado na Figura 15 do Capítulo 4 é um estado indesejado na planta que interrompe a geração de energia elétrica representando uma situação atual. Neste estudo de caso, considerou-se o valor do sensor de vibração definido como instância na ontologia modelada como condição suficiente para indicação da situação corrente, apresentada na Figura 37.

De acordo com o tipo de situação detectada, alertas específicos podem ser apresentados respeitando características contextuais. Na Figura 37 é apresentado um alarme de vibração logo após a detecção da situação.

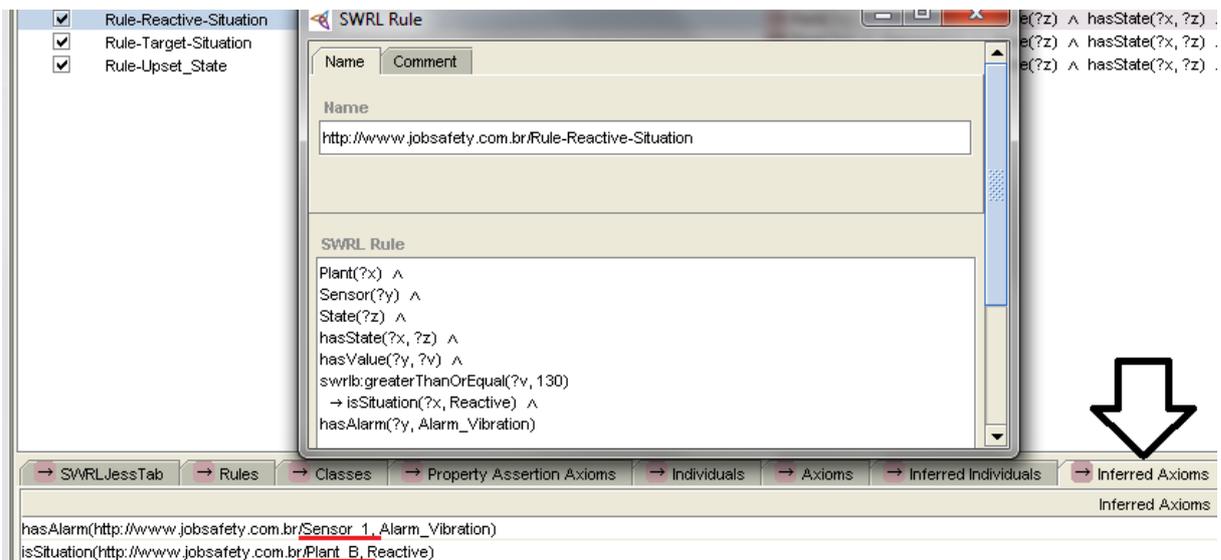


Figura 37 Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.

O estudo de caso apresenta a avaliação de um caso típico que exige uma ação, pois trata-se de uma situação atual de emergência que exige o bloqueio da turbina. Isso deve ser apresentado ao operador como recomendação para que possa se certificar do correto funcionamento do automatismo ou efetuar o bloqueio manualmente.

A Figura 38 apresenta a correlação dos blocos correspondentes aos compostos “3DNPCP3RTU” e “4GSECP4TURB” e a seleção dos principais eventos ocorridos momentos antes do bloqueio da unidade de geração.

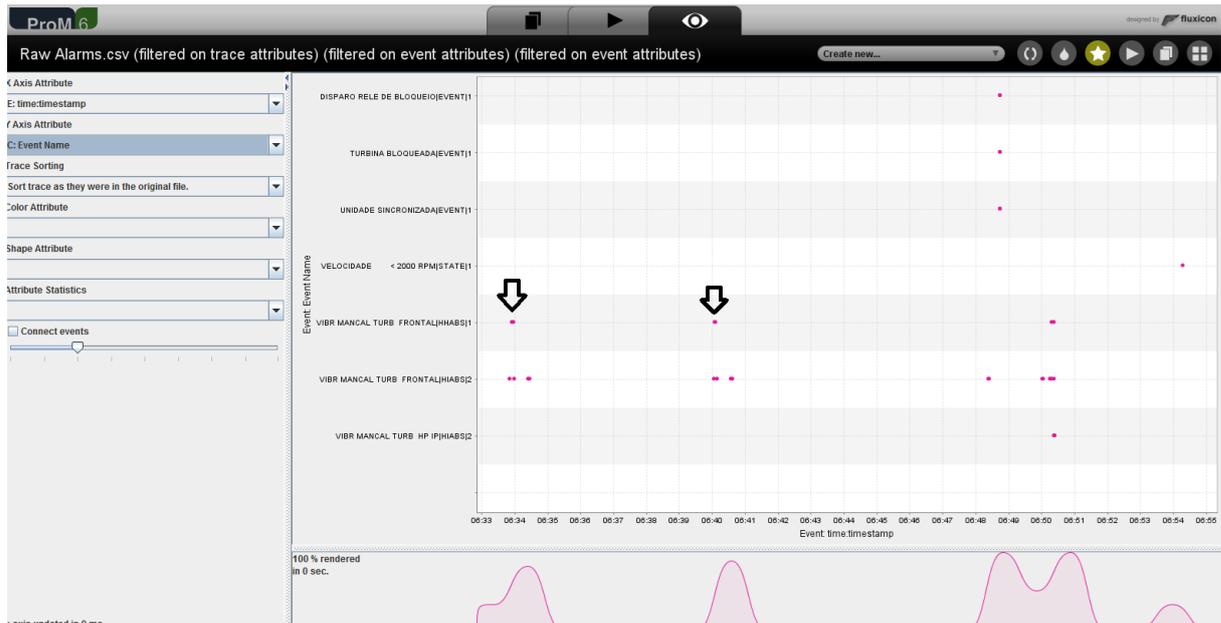


Figura 38 Visão geral do *plug-in Dotted Chart* ProM. Fonte: O autor.

Observa-se que ocorreram alarmes relacionados a vibração do mancal por volta das 06:34 e novamente após as 06:40 que levariam ao bloqueio da unidade, conforme indicado na Figura 40, que ocorre somente após as 06:48. Isso indica uma possível falha no automatismo.

A determinação da relação que define a situação de interesse necessita dos elementos de contexto que, permite o relacionamento dos elementos de contexto definindo um conjunto padrão de relações semânticas que definem uma situação. Assim, o conjunto de instâncias com relações semânticas define o contexto que é a informação de interesse.

A instância do modelo conceitual envolvida na situação de interesse representa dados dos dispositivos da planta que geram os eventos que ficam armazenados na base de dados. A análise dos eventos diretamente na base de dados nem sempre está disponível para a operação.

O acesso aos eventos que ocorreram no processo por meio de telas de supervisão é, em geral, o meio disponível para os envolvidos no processo. Nesta tela, são apresentados os eventos configurados que permitem acompanhar a produção. Contudo, características descritas nos capítulos anteriores indicam que uma série de eventos que definem uma mesma

situação, quando apresentados em uma tela de supervisão, pode gerar um problema ainda maior dependendo da forma como for interpretado. A análise da informação contida em uma sequência de eventos anunciados necessita além de conhecimento sobre o processo, meios de acesso que facilitem a investigação das causas dos eventos e algum tipo de raciocínio que apoie a tomada de decisão, principalmente em casos de falha no automatismo.

O resultado da mineração de dados do processo apresentado nesta dissertação permite conhecer os padrões dos eventos e como eles se relacionam. Neste estudo, os alarmes gerados na planta são ativados pelos blocos que fazem parte dos compostos. Isso facilita a filtragem dos eventos e a aplicação dos *plug-ins* apresentados para detecção de padrões. Na Figura 39 é apresentado um extrato dos eventos que ocorrem logo após o bloqueio da caldeira. A caldeira é o equipamento responsável por gerar o vapor que movimenta a turbina.



Figura 39 Visão parcial do *plug-in Show Sequence and Patterns*. Fonte: O autor.

A partir da Figura 39 é possível identificar os principais eventos que ocorrem logo após o bloqueio da caldeira que ficam registrados na base de dados. Assim, o responsável pelo processo deverá buscar de outra forma a informação que necessita caso essas informações não sejam apresentadas na tela do supervisão.

7.2.2 Estudo de Caso 2

Neste segundo estudo de caso observou-se os eventos relacionados a marcha particular do conjunto turbina-gerador da unidade de geração após o sincronismo, relacionado a pressão do condensador. A Figura 40 o resultado da aplicação do *plug-in* somente para o composto “3GRECP5TURB”.



Figura 40 Visão parcial do *plug-in Show Sequence and Patterns*. Fonte: O autor.

A situação analisada a partir de uma sequência de eventos valida a situação de interesse relativa quando o valor do sensor atinge o seu limite superior de processo, semelhante ao primeiro estudo de caso, onde foram avaliados os valores de vibração. Para este estudo de caso, foram analisados eventos em que os valores de pressão do condensador foram suficientes para a detecção da situação atual e o bloqueio da unidade conforme apresentado na Figura 40, por meio da inferência do modelo conceitual.

Também se considerou os intervalos de pressão acima de 95 e inferiores a 140 mbar e seus respectivos tempos permitidos para funcionamento dentro do contexto de interesse *Upset*.

$$\begin{aligned}
 & (\forall x, y, z, e, s, v, t) \in EC(OWL : Thing) \\
 & ((x) \in EC(Plant) \\
 & \wedge (y) \in EC(Sensor) \wedge (z) \in EC(State) \\
 & \wedge (x, z) \in ER(hasState) \\
 & \wedge (y, v) \in ER(hasValue) \\
 & \wedge (x, e) \in ER(generateEvent) \\
 & \wedge (e, t) \in ER(hasTime) \\
 & \wedge (s, v) \in ER(greaterThanOrEqual) \\
 & \wedge (e, t) \in ER(greaterThanOrEqual) \\
 & \rightarrow (x, z) \in ER(isSituation))
 \end{aligned} \tag{7}$$

Realizando a tradução da regra em (7) para linguagem SWRL, tem-se a regra:

$$\begin{aligned}
 &Plant(?x) \\
 &\wedge Sensor(?y) \wedge State(?z) \\
 &\wedge hasState(?x,?z) \\
 &\wedge hasValue(?y,?v) \\
 &\wedge generateEvent(?x,?e) \\
 &\wedge hasTime(?e,?t) \\
 &\wedge swrlb:greaterThanOrEqualTo(?v,?s) \\
 &\wedge swrlb:greaterThanOrEqualTo(?t,?e) \\
 &\rightarrow isSituation(?x,Prediction)
 \end{aligned}
 \tag{8}$$

O contexto *Upset* é um estado de atenção na planta que, em geral, não interrompe a geração de energia elétrica. Neste estudo de caso, considerou-se o valor do sensor da pressão do condensador definido como instância na ontologia modelada como condição suficiente para indicação da situação corrente, apresentada na Figura 41.

Assim, avaliou-se uma nova situação que se encontra no contexto de interesse *Upset*, o que mostra a importância deste trabalho como base para implementação de um caso preditivo.

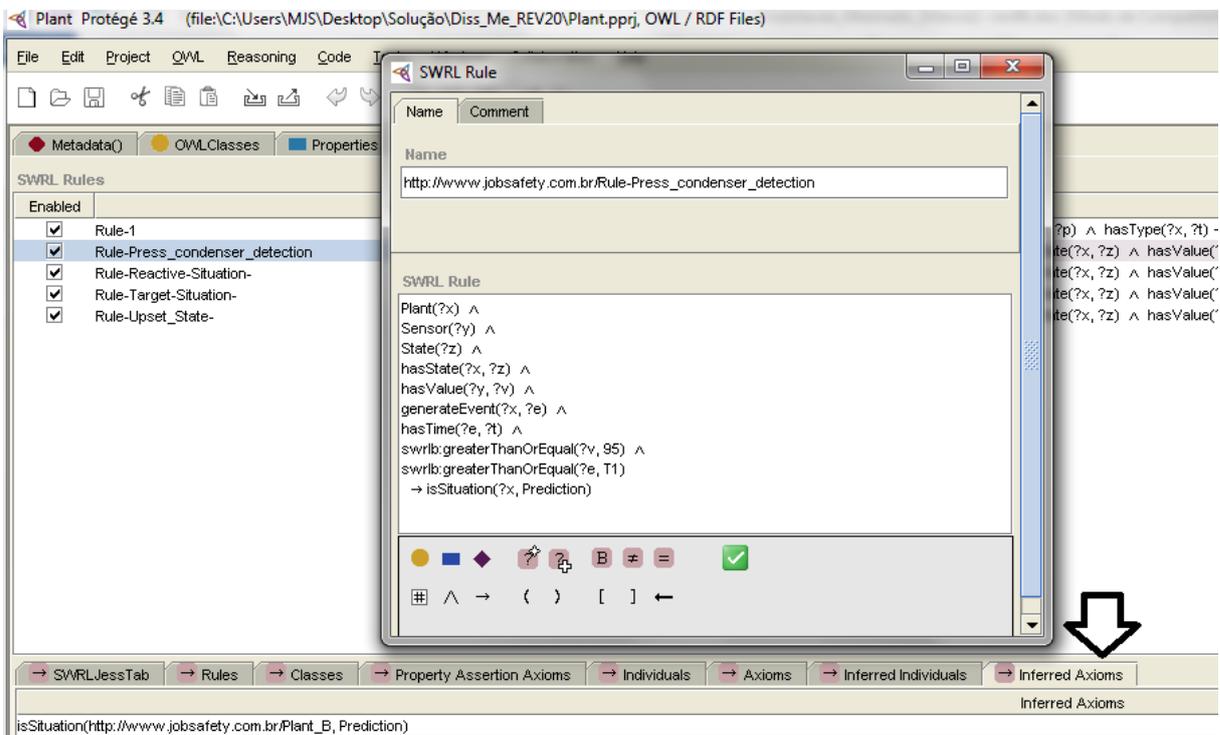


Figura 41 Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido para situação corrente. Fonte: O autor.

7.2.3 Estudo de Caso 3

A execução das regras SWRL requer a análise por meio de um motor de inferência que possibilite o raciocínio utilizando um conjunto de regras e fatos como entrada. O motor de inferência JESS possibilita o raciocínio semântico e habilita as regras SWRL que descrevem a relação semântica no software *Protégé* e retorna a situação de interesse.

A partir da análise dos padrões apresentados nas Figuras 33, 39 e 40 resultantes do processo de mineração de dados, foram definidas as regras semânticas para situações de interesse, conforme mostra a Tabela 6.

Tabela 6 Tabela das regras definidas no estudo de caso 3.

Situação de interesse	Regra representada em SWRL
<i>Target</i> – Figura 33	Plant(?x) \wedge hasEvent(?e, Disjuntor_campo_aberto) \wedge hasEvent(?e, Unidade_operando_vazio) \wedge hasEvent(?e, Estado_disjuntor_52_04) \wedge hasEvent(?e, Unidade_sincronizada) \wedge hasEvent(?e, Disparo_rele_potencia_inversa) \rightarrow isSituation(?x, Target)
<i>Reactive</i> – Figura 39	Plant(?x) \wedge hasEvent(?e, Bloqueio_da_caldeira) \wedge hasEvent(?e, Operacao_manual_disjuntor_52) \wedge hasEvent(?e, Estado_disjuntor_52_03) \wedge hasEvent(?e, Unidade_sincronizada) \wedge hasEvent(?e, Turbina_em_giro) \wedge hasEvent(?e, Disjuntor_campo_aberto) \wedge hasEvent(?e, Turbina_bloqueada) \rightarrow isSituation(?x, Reactive)
<i>Reactive</i> – Figura 40	Plant(?x) \wedge hasEvent(?e, Turbine_load_maior_20_percent) \wedge hasEvent(?e, High_condens_presslimit_signal) \wedge hasEvent(?e, Tripping_press_limitation) \wedge hasEvent(?e, Condenser_press_high) \wedge hasEvent(?e, Turbine_speed_menor_2000_RPM) \wedge hasEvent(?e, Turbine_speed_menor_1000_RPM) \rightarrow isSituation(?x, Reactive)

Fonte: Autor

Conforme apresentado na Tabela 6, as situações de interesse *Target* e *Reactive* são definidas pelos eventos que compõe as relações semânticas obtidas a partir do processo de mineração de dados.

A partir dos padrões analisados foram geradas regras que permitem inferir a situação da planta. Dessa forma, é possível avaliar os padrões de eventos e inferir a situação da planta, sendo possível gerar recomendações a partir da análise de padrões operacionais. A Figura 42 apresenta a detecção da situação reativa referente a Figura 40.

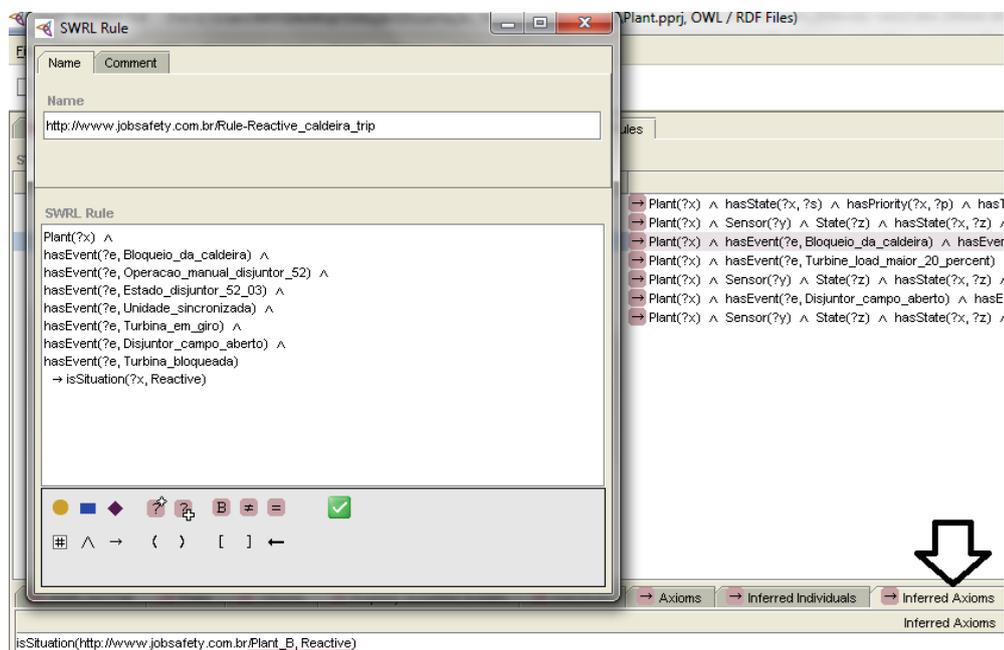


Figura 42 Visão parcial da interface de desenvolvimento de regras SWRL e do axioma inferido.
Fonte: O autor.

7.3 CONCLUSÕES

Nestes estudos de caso foram apresentadas as situações referentes ao contexto atual. A partir das regras semânticas, combinada para formar um padrão ou para eventos simples de um sensor, é possível inferir e recomendar a situação corrente na planta.

O modelo ontológico de contexto comparado a um sem a informação de contexto possibilita, além de compartilhar o conhecimento por meio de uma especificação formal de dados de contexto, o uso de motores de inferência que utilizam lógica descritiva, permitindo a verificação da consistência do modelo e a inferência da situação corrente. A contextualização por meio da abordagem ontológica também permite a ampliação e o reuso do vocabulário de termos apresentado.

Conforme apresentado, os sistemas de gerenciamento de alarmes atuais possuem painéis onde são listados os alarmes para os operadores. A partir de uma configuração prévia, que nem sempre segue uma metodologia formal ou critérios normativos, são definidas as prioridades e tipos dos eventos que serão apresentados. A análise dos alarmes a partir deste modo de sinalização exige do operador um conhecimento amplo, principalmente durante os casos onde as taxas de alarmes ultrapassam os limites aceitáveis, segundo as métricas descritas no Capítulo 2.

Neste estudo, observou-se que é possível analisar um volume de eventos que não seria prático de forma manual. Além disso, estabelecer regras para as diferentes situações e gerar conhecimento a partir de um mecanismo de raciocínio lógico que avalia os padrões detectados.

8 CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma proposta de modelo conceitual aplicado a um sistema de alarmes de processo e sua base de dados, modelando a planta como elemento central. Entre os principais benefícios apresentados a partir deste estudo, destaca-se a base semântica e a metodologia aplicada que possibilitam a continuação para um modelo preditivo.

Destaca-se a abordagem de temas de diferentes áreas, tais como Ontologia, Mineração de Processos e Ambientes Inteligentes. No Capítulo 2, foram aprofundados os conceitos sobre dados, informação e conhecimento, bem como contexto, alarmes e eventos. No Capítulo 3 foram apresentados os trabalhos relacionados a este estudo e as possíveis lacunas existentes. No Capítulo 4 é apresentada a descrição da proposta. No Capítulo 5 foi definido um modelo de contexto utilizando ontologia, onde diferentes domínios foram identificados, tais como Planta, Ação, Situação e Evento. A partir dessa ontologia desenvolvida em OWL-DL foi possível validar as características reativas da planta. O Capítulo 6 apresentou a aplicação de métodos de mineração de processos para avaliação dos padrões, e assim implementar regras que alimentaram a máquina de inferências lógicas Jess. Por fim, o Capítulo 7 apresentou o estudo de caso onde foram analisadas situações de interesse presentes em um caso real industrial.

Os estudos de caso tiveram como objetivo aplicar o modelo desenvolvido em um ambiente real, considerando alguns casos de emergência previstos em procedimentos operacionais para o conjunto turbina-gerador e que necessitam ações automatizadas e humanas, principalmente quando o automatismo apresenta falhas. A partir destes estudos, conclui-se que para a correta intervenção operacional torna-se necessário o acesso ao conjunto de informações válidas presentes na planta industrial a partir da avaliação dos padrões de eventos ocorridos. Considerando uma base de conhecimento, é possível gerar recomendações

a partir das situações detectadas. O aspecto inovador deste trabalho está na forma como os eventos são avaliados.

Dessa forma, a questão de pesquisa foi respondida para os aspectos verificados. A mineração de processo foi utilizada para avaliar padrões. A ontologia foi aplicada como um integrador semântico que possibilitou a implementação de regras de inferência que permitiram identificar situações de interesse formando uma base de conhecimento. Isso possibilita ter um conhecimento sobre o modelo da situação, caracterizada pelos eventos.

8.1 TRABALHOS FUTUROS

Como trabalhos futuros, pode-se destacar:

- Aumentar a abrangência do modelo de contexto apresentado a partir da identificação de novos domínios, como domínio do usuário e tecnológico.
- Explorar os conceitos de redes de ontologia que permita cobrir diferentes domínios, permitindo analisar todos os contextos e situações.
- Analisar algoritmos de mineração de processos e aprendizagem de máquina que possam ser vinculados a tecnologias da Web Semântica.
- Estudo e implementação de *Multy-Entity Bayseian Network* (MEBN) a fim de descrever a influência dos padrões identificados em uma situação preditiva.

8.2 TRABALHO COMPLETO PUBLICADO EM CONFERÊNCIA

Esta dissertação apresentou conceitos aplicados em diferentes áreas de estudo. Inicialmente foi realizado uma análise dos conceitos de situação e contexto. Em seguida foram analisadas as ferramentas para mineração de dados de processo e que permitiram analisar além de dados estatísticos, padrões de dados de eventos ocorridos no processo. Na

sequência do trabalho, foi desenvolvido o modelo conceitual de contexto. Com a definição da modelagem, o conceito de situação foi formalizado e proposto o uso de inferência para determinação da situação da planta industrial. A partir dos testes realizados, resultou na seguinte publicação:

- SILVA, M. J.; PEREIRA, C. E.; GÖTZ, M. A Dynamic Approach for Industrial Alarm Systems. *2016 International Conference on Computer, Information and Telecommunications Systems (CITS 2016)*, Kuming, China, 2016.

A partir deste estudo, resultou na publicação do artigo:

- SILVA, M. J.; PEREIRA, C. E.; GÖTZ, M. Knowledge awareness in Alarm Systems Database. *3rd IFAC Workshop on Advanced Maintenance Engineering, Service and Tecnology (AMEST'16)*, Biarritz, France, 2016.

REFERÊNCIAS

- AALST, W. M. P. **Process Mining-Discovery, conformance and enhancement of business processes**. Chicago: Springer, 2010.
- AGUIAR, L. P.; ALMEIDA, V. A. F.; MEIRA, W. Descoberta de Padrões de Alarmes Redundantes com Técnicas de Mineração de Dados e Redes Complexas. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 18., 2010, Bonito MS. **Anais . . .** Campinas: SBA, 2010. p. 1020-1027.
- AIZPURÚA, O.; GALÁN, R.; JIMÉNEZ, A. A New Cognitive-Based Massive Alarm Management System in Electrical Power Administration, México. In: CARIBBEAN CONFERENCE ON DEVICES, CIRCUITS AND SYSTEMS, 7., 2008, México. **Proceedings...** México: IEEE, 2008, p. 1-6.
- ALMEIDA, M.; BAX, M.; Uma Visão Geral sobre Ontologias: Pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Revista Ciência da Informação**, Brasília, v. 32, n. 3, p. 7-20, dez. 2003.
- ANSI/ISA-5.1-2009 Instrumentation Symbols and Identification. **The International Society of Automation**. Disponível em: <<http://www.isa.org>>. Acesso em: 09 ago. 2016.
- ANSI/ISA-18.2-2016. Management of Alarm Systems for the Process Industries. **The International Society of Automation**. Disponível em: <<http://www.isa.org>>. Acesso em: 09 ago. 2016.
- ARUNRAJ, N. S.; MAITI, J. Risk-based Maintenance –Techniques and applications. **Journal of Hazardous Materials**, New York, v. 142, p. 653-661, 2007.
- BERNARAS, A. *et al.* An Ontology for Fault Diagnosis in Electrical Networks. In: INTELLIGENT SYSTEMS AND APPLICATIONS TO POWER SYSTEMS, 1996, Orlando. **Proceedings...** Orlando: IEEE, 1996, p. 199-203.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American Magazine**, New York, v. 284, n. 5, p. 28-37, May. 2001.
- BETTINI, C. *et al.* A Survey of context modelling and reasoning techniques. **Pervasive and Mobile Computing**, Amsterdam, v. 6, n. 2, p. 161-180, Apr. 2010.
- BRANSBY, M. L.; JENKINSON, J. **The Management of Alarm Systems**. A review of current practice in the procurement, design and management of alarm systems in the chemical and power industries. London: HSE, 1998. Disponível em:<http://www.hse.gov.uk/research/crr_pdf/1998/crr98166.pdf>. Acesso em: 15 Dec. 2015.

CALVANESE, D. *et al.* Ontologies and databases: The DL-Lite approach. In: 5th Int. Reasoning Web Summer School Tutorial Lectures. New York: Springer , 2009.

CHEN, G.; KOTZ, D. **A Survey of Context-Aware Mobile Computing Research.** Technical Report. Hanover: Department of Computer Science, Dartmouth College, 2000.

COCHRAN, E. L.; MILLER, C.; BULLEMER, P. Abnormal Situation Management in Petrochemical Plants: Can a Pilot's Associate Crack Crude?. In: ABNORMAL SITUATION MANAGEMENT JOINT REASERCH AND DEVELOPMENT CONSORTIUM (ASM), 1996, Dayton. **Proceedings...**New York: IEEE, 1996, p. 806-813.

DENTLER, K. *et al.* Comparison of reasoners for large ontologies in the OWL 2 EL profile. **Semantic Web**, v. 2, n. 2, p. 71-87, Apr. 2011.

DEY, A. K., ABOWD, G. D. **Towards a Better Understanding of Context and Context-Awareness.** Technical Report, Georgia: College of Computing. Institute of Technology, Atlanta, 1999.

DEY, A. K. **Providing Architectural Support for Building Context-Aware Applications.** 1999, 34 p. Tese (Doctor of Philosophy in Computer Science) - Georgia Institute of Technology, Atlanta, 1999.

DUNN, W. R. Designing safety-critical computer systems. **IEEE Computer Society**, New York, v. 2, n. 36, p. 40 - 46. Nov. 2003.

EEMUA. **Alarm Systems: A Guide to Design, Management and Procurement.** The Engineering Equipment and Materials Users Association, London, 1999. n. 191. Disponível em: < <http://www.eemua.org/Products/Publications/Print/EEMUA-Publication-191.aspx>>. Acesso em: 10 Nov. 2015.

EEMUA. **Alarm Systems: A Guide to Design, Management and Procurement.** The Engineering Equipment and Materials Users Association, second edition, London, 2007. n. 191. Disponível em: < <http://www.eemua.org/Products/Publications/Print/EEMUA-Publication-191.aspx>>. Acesso em: 10 Nov. 2015.

ETZION, O.; NIBLETT, P. **Event Processing in Action.** Greenwich: Manning Publications Co, 2010.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, Palo Alto, v. 17, n. 3, p. 37-57, 1996.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, London, v. 5, n. 2, p. 199-220, 1993.

HABERLIN, R. **UnBBayes PR-OWL 2.0 Tutorial**, 2013. Disponível em: <<http://sourceforge.net/projects/unbbayes/files/UnBBayes%20Plugin%20Framework/Plugins/Probabilistic%20Networks/MEBN/PR-OWL2/>>. Acesso em: 20 dez. 2015.

HABIBI, E.; HOLLIFIELD, B. Alarm Systems greatly affect offshore facilities amid high oil process. **World Oil Magazine**, Houston, v. 227, n. 9, p. 1-4, Sept. 2006.

- HARRIS, S.; SEABORNE, A. SPARQL 1.1 Query Language. W3C Recommendation. World Wide Web Consortium, Mar. 2013. Disponível em: <<http://www.w3.org/TR/sparql11-query>>. Acesso em: 05/02/2016.
- HATCH, D. **Alarms**: Prevention is better than cure. TCE. Pp 40-42. July, 2005. Disponível em:< <http://www2.emersonprocess.com/pt-BR/Pages/Home.aspx>>. Acesso em: 18 Jul. 2015.
- HALL, M. *et al.* The WEKA Data Mining Software: An Update. **SIKDD Explorations**, New York, v. 11, n. 1, p. 10-18, 2009.
- HERVÁS, R.; BRAVO, J.; FONTECHA, J. A Context Model Based on Ontological Languages: A Proposal for Information Visualization. **Journal of Computer Science**, Sydney, v. 16, n. 12, p. 1539-1555, 2010.
- HONEYWELL. **Operation Management Pro**. White Paper. USA, 2004. 16 p. Disponível em:<<https://www.honeywellprocess.com/library/marketing/whitepapers/OperationsManagementPro.pdf>>. Acesso em: 04 Nov. 2015.
- HORROCKS, I. *et al.* OWL Rules: A proposal and Prototype Implementation. **Journal of Computer Science**, Amsterdam, v. 3, n. 1, p. 23-40, 2005.
- IEEE, Process Mining Manifesto. 2011. In: BUSINESS PROCESS MANAGEMENT WORKSHOPS, 2011, **Lecture Notes in Business Information Processing**, Berlin, v. 99, Springer, 2011, p. 1-15.
- IZADI, I. *et al.* An Introduction to Alarm Analysis and Design. In: SYMPOSIUM ON FAULT DETECTION, SUPERVISION AND SAFETY OF TECHNICAL PROCESSES, 7., 2009, Barcelona. **Proceedings...** Venue: IFAC, 2009. p. 645-650.
- KLEMETTINEN, M.; MANNILA, H.; TOIVONEN, H. A Data Mining Methodology and Its Application to Semi-Automatic knowledge Acquisition. In: INTERNATIONAL WORKSHOP ON DATABASE AND EXPERT SYSTEM APPLICATIONS, 8., 1997, Toulouse. **Proceedings...** New York: IEEE, 1997. p. 670-677.
- KOFI, G. A.; LUO, G.; QIN, K. Multi_Level Data Pre_Processing for Software Defect Prediction. In: INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, INNOVATION MANAGEMENT AND INDUSTRIAL ENGINEERING, 6., 2013, Shannxi. **Proceedings...** New York: IEEE, 2013. p. 170-174.
- KONTCHAKOV, R. *et al.* Answering SPARQL queries over database under OWL 2 QL entailment regime. In: 13th Int. Semantic Web Conference, New York: Springer, 2014.
- KORDIC, S. *et al.* Associative Data Mining for Alarm Groupings in Chemical Processes. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM AND KNOWLEDGE ENGINEERING, 2007, Chengdu. **Proceedings...** Chengdu: Atlantis, 2007. p. 191-198.
- LAALLAM, F. Z.; SELLAMI, M. Gas Turbine Ontology for Industrial Process. **Journal of Computer Science**, Amsterdam, v. 3, n. 2, p. 113-118, 2007.
- LASKEY, K. B. MEBN: A language for first-order Bayesian knowledge bases. **Artificial Intelligence**, Essex, v.172, n.2, p. 140-178, 2008.

LEITÃO, G. B. P. *et al.* Sistema para Análise e Otimização de Alarmes em Plantas Petroquímicas. In: RIO OIL & GAS EXPO AND CONFERENCE, 2008, Rio de Janeiro. **Anais . . .** Rio de Janeiro: IBP, 2008. p. 1-8.

LEITÃO, G. B. P.; GUEDES, L. A.; ARAÚJO, J. R. S. A Correlation-based Approach to Determining Related Alarms in Industrial Processes. In: CONFERENCE ON EMERGING TECHNOLOGIES & FACTORY AUTOMATION, 17., 2012, Krakow. **Proceedings...** New York: IEEE, 2012. p. 1-7.

LI, M.; DU, X.; WANG, S. A Semi-automatic Ontology Acquisition Method for the Semantic Web. In: WAIM'S PROCEEDINGS OF THE 6th INTERNATIONAL CONFERENCE ON ADVANCES IN WEB-AGE INFORMATION MANAGEMENT, 6., 2005, Hangzhou. **Proceedings...** Heidelberg: Springer-Verlag, 2005. p. 209-220.

LIMA, G. F. L. *et al.* Correlação Semântica de Alarmes Utilizando Ontologia. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 2013, Fortaleza. **Anais . . .** Campinas: SBAi, 2013. p. 1-6.

LIU, B. **Web Data Mining.** Exploring Hyperlinks, Contents, and Usage Data. 2nd ed. Springer: New York, 2011. 622 p.

MACHADO, A. *et al.* Situation-Awareness as a Key for Proactive Actions in Ambient Assisted Living. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 15., 2013, France. **Proceedings...** Setúbal: SciTpress, 2013. p. 418-426.

MACHADO, A. *et al.* A Reactive and Proactive Approach for Ambient Intelligence. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 16., 2014, Lisbon. **Proceedings...** Madrid: Springer, 2014. p. 501-512.

MACHADO, G. M.; OLIVEIRA, J. P. M. Context-Aware Adaptive Recommendation of Resources for Mobile Users in a University Campus. In: IEEE INTERNATIONAL CONFERENCE ON WIRELESS AND MOBILE COMPUTING, NETWORKING AND COMMUNICATIONS, 10., 2014, Chipre. **Proceedings...** Larnaca: IEEE, 2014, p. 427 – 433.

MATTIASSON, C. T. The Alarm System from the Operator's Perspective. In: HUMAN INTERFACES IN CONTROL ROOMS, COCKPITS AND COMMAND CENTERS, 1999, Bath. **Proceedings...** New York: IEEE, 1999. p. 1-5.

MCGUINNESS, D. L.; VAN HAMERLEN, F. OWL Web Ontology Language Overview. **W3C recommendation**, Cambridge, v. 10, n. 3, p. 10, 2004. Disponível em: <<https://www.w3.org/TR/owl-features/>>. Acesso em: 22 set. 2015.

MIRANDA, R. C. R. O uso da Informação na formulação de ações estratégicas pelas empresas. **Revista Ciência da Informação**, Brasília, v.28, n.3, p. 286-292, 1999.

MOORE, P. *et al.* A Survey of Context Modelling for Pervasive Cooperative Learning. In: INTERNATIONAL SYMPOSIUM ON INFORMATION TECHNOLOGIES AND APPLICATIONS IN EDUCATION, 7., 2007, Kunming. **Proceeding...** Beijing: IEEE Computer Society, 2007. p. K5-1-K5-6.

- NATARAJAN, S.; GHOSH, K.; SRINIVASAN, R. An Ontology for distributed process supervision of larg-scale chemical plants. **Computer and Chemical Engineering**, Singapore, v.46, p. 124-140, Dec. 2012.
- NOY, N. F.; MCGUINNES, D. L. **Ontology development 101: A Guide to creating your first ontology**. In: STANFORD KNOWLEDGE SYSTEMS LABORATORY, Stanford, 2001. p. 234-252.
- O'BREIN, P. An Ontology for Mobile Situation Aware Systems. **Australian Journal of Information System**, Geelong, v. 15, n. 2, p. 1-28, 2009.
- O'BREIN, P.; WOLL, D. **Alarm Management Strategies**. ARC STRATEGIES, 2004. Disponível em: <<https://www.yokogawa.com/be-fr/cp/pdf/ARC-AlarmStrat.pdf>>. Acesso em: 30 Dec. 2015.
- PEREIRA, C. E.; JUNIOR, W. P. A Supervisory Tool for Real-Time Industrial Automation Systems. In: IEEE INTERNATIONAL SYMPOSIUM ON OBJECT-ORIENTED REAL-TIME DISTRIBUTED COMPUTING, 6., 2003, Washington. **Proceeding...** New York: IEEE Computer Society. 2003. p. 1-8.
- POLI, R.; OBRST, L. **Theory and applications of ontology: Computer applications**. New York: Springer, 2010.
- PREKOP, P.; BURNETT, M. Activities, Context and Ubiquitous Computing. **Computer Communications**, Special Issue on Ubiquitous Computing, Amsterdam v.26, n. 11, p.1168-1176, 2003.
- QUINTÃO, H.; ROSÁRIO, G. Especificação de Requisitos de um Sistema de Gerenciamento de Alarmes baseado em Recomendação de Ações. In: MERORIAS DE LA CONFERENCE IBEROAMERICANA DE SOFTWARE ENGINEERING, 11., 2008. Recife, PE. **Anais . . . Recife: CIbSE**, 2008, p 1-15.
- REZENDE, S. O. Mineração de Dados. In: CONGRESSO SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. 25., 2005, São Leopoldo. **Anais . . . São Leopoldo: SBC**, 2005. p. 397-433. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/enia/2005/0102.pdf>>. Acesso em: 05 out. 2015.
- SANTOS, V. V. **CEManTIKA: A Domain-Independent Framework for Designing Context-Sensitive System**. 2008, 187 p. Tese (Doutorado em Ciência da Computação), Universidade Federal do Pernambuco, Recife, 2008.
- SETZER, V. W. **Dado, Informação, Conhecimento e Competência**. 2001. Disponível em: <<http://www.ime.usp.br/~vwsetzer/dado-inf.html>>. Acesso em: 05 dez. 2015.
- SETZER, V. W. **Dado, Informação, Conhecimento e Competência**. 2004. In: Folha Educação. p. 6-7. Disponível em: <<http://www.ime.usp.br/~vwsetzer/dado-inf.html>>. Acesso em: 05 Dec. 2015.
- SIHEM, K.; TAREK, K. M.; RINGWOOD, J. V. Steam Turbine Diagnostic System based on a Domain Ontology Implemented using J2EE Technology. In: INTERNATIONAL CONFERENCE ON SIGNALS, CIRCUITS AND SYSTEMS, 3., 2009, Medenine. **Proceeding...** Medenine: IEEE, 2009. p. 1-6.

SIRIN, E. *et al.* Pellet: A Practical OWL-DL Reasoner. **Journal Web Semantics: Science, Services and Agents on the World Wide Web**, v. 5, n. 2, p. 51-53. Jun, 2007.

SOUZA, A. J. *et al.* Gerência de Informação de Processos Industriais: Um Estudo de Caso na Produção de Petróleo e Gás. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 7., 2005, São Luís. **Anais . . .** Campinas: IEEE, 2005. p. 1-7.

STANFORD. **Protégé**. Stanford Center for Biomedical Informatics Research (BMIR) at the Stanford University School of Medicine. Disponível em: <<http://protege.stanford.edu>>. Acesso em: 01 ago. 2015.

STRANG, T.; LINNHOFF-POPIEN, C. A. Context Modeling Survey. In: WORKSHOP ON ADVANCED CONTEXT MODELING, REASONING AND MANAGEMENT, 6., 2004, Nottingham. International Conference on Ubiquitous Computing. **Proceedings...** Nottingham: University of Southampton, 2004. p. 1-8.

VENCESLAU, A. *et al.* Ontology for computer-aided fault tree synthesis. In: IEEE EMERGING TECHNOLOGY AND FACTORY AUTOMATION, 2014, Barcelona. **Proceedings...** Barcelona: IEEE, 2014. p. 1-4.

VERBECK, H. M.W. *et al.* ProM6: The Process Mining Toolkit. In: INTERNATIONAL CONFERENCE ON BUSINESS PROCESS MANAGEMENT, 8., 2010, Hoboken, New Jersey. **Proceedings...** Hoboken: CEUR-WS, 2010, p. 34-39.

YE, J. *et al.* Semantic Web Technologies in Pervasive Computing: A Survey and Research Roadmap. **Pervasive and Mobile Computing**, Amsterdam, v. 23, p. 1-25, 2015.

YE, J.; STEVENSON, G.; DOBSON, S. A top-level ontology for smart environments. In: **Pervasive and Mobile Computing**, Amsterdam, v. 7, n.3, p. 359-378, 2011.

WANG, X. H. *et al.* Ontology Based Context Modeling and Reasoning using OWL. In: IEEE Annual Conf. on Pervasive Computing and Communications Workshops, 2., 2004, Singapore. **Proceedings...** Singapore: IEEE, 2004, p. 1-5.

WITTEN, I. H.; EIBE, F.; HALL, M. A. **Data Mining: Practical Machine Learning tools and Techniques**. Burlington, Elsevier, 2011.