

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MODELO PREDITIVO DE SITUAÇÕES COMO APOIO  
À CONSCIÊNCIA SITUACIONAL E AO PROCESSO  
DECISÓRIO EM SISTEMAS DE RESPOSTA À  
EMERGÊNCIA**

**CLAUDIA BEATRIZ BERTI**

**ORIENTADORA: PROFA. DRA. REGINA BORGES DE ARAUJO**

São Carlos - SP  
Agosto/2017

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MODELO PREDITIVO DE SITUAÇÕES COMO APOIO  
À CONSCIÊNCIA SITUACIONAL E AO PROCESSO  
DECISÓRIO EM SISTEMAS DE RESPOSTA À  
EMERGÊNCIA**

**CLAUDIA BEATRIZ BERTI**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutora em Ciência da Computação, área de concentração: Processamento de Imagens e Sinais.

Orientadora: Profa. Dra. Regina Borges de Araujo

São Carlos - SP  
Agosto/2017

---

Folha de Aprovação

---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado da candidata Claudia Beatriz Berti, realizada em 28/08/2017.



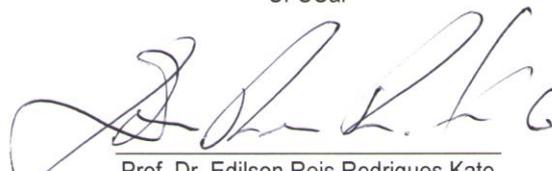
---

Prof.ª Dr.ª Regina Borges de Araújo  
UFSCar



---

Prof. Dr. José Hiroki Saito  
UFSCar



---

Prof. Dr. Edilson Reis Rodrigues Kato  
UFSCar

---

Prof.ª Dr.ª Fátima de Lourdes dos Santos Nunes Marques  
USP

---

Prof. Dr. Celso Massaki Hirata  
ITA

Certifico que a defesa realizou-se com a participação à distância dos membros Fátima de Lourdes dos Santos Nunes Marques e Celso Massaki Hirata, depois das arguições e deliberações realizadas, os participantes à distância estão de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.



---

Prof.ª Dr.ª Regina Borges de Araújo (presidente)

## Dedicatória

*Dedico toda energia salutar que nasceu e se desenvolveu na construção desta tese  
à querida e eterna amiga Rosana Silva.*

# AGRADECIMENTOS

*“O importante é caminhar sempre, com decisão e coragem”. Ir. Tecla Merlo  
Minha eterna gratidão a todos que, mesmo pela mais singela forma, me apoiaram e me incentivaram durante esta caminhada.*

À minha querida orientadora, professora Regina Borges de Araujo, pela oportunidade, por compartilhar seu tempo, seus conhecimentos, por toda dedicação, por ter estado sempre ao meu lado e principalmente por me fazer acreditar.

À minha querida instituição, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), campus Diamantina, por me proporcionar todo o suporte a esta qualificação.

A CAPES, pelo apoio promovido por meio do programa Prodoutoral.

À Polícia Militar do Estado de São Paulo, pela prestimosa colaboração, representada pelo Tenente Murilo Goes de Almeida.

Ao Centro de Processamento de Dados da Polícia Militar do Estado de São Paulo, representado pelo Major Márcio Ribeiro de Campos, Chefe da Seção de Arquitetura de Software, pela colaboração inestimável com o estudo de caso.

Ao primo e policial militar Gustavo Seconde Panagio, pelas importantes informações e orientações durante as análises do estudo de caso.

Aos colegas professores e funcionários do Departamento de Computação da UFVJM pela colaboração e apoio durante meu período de afastamento.

Especial agradecimento a Marcus Vinícius C. Guelpeli pelo prestimoso apoio e colaboração em minhas pesquisas iniciais.

À Universidade Federal de São Carlos, ao Departamento de Computação, professores e funcionários, pelo brilhantismo profissional e dedicação.

Aos colegas do laboratório de pesquisa WINDIS, pelos maravilhosos e inesquecíveis momentos de aprendizagem e também de descontração.

Às amigas queridas, que São Carlos me deu de presente, Jane Elisa Almeida Lacombe, Gabriela Perone Rocha, Ondina Amaral, Janaina Cintra Abib e Claudia Vicci pela amizade, carinho e apoio incondicional. Vibrações de carinho à saudosa e querida amiga Rosana Silva.

Aos amigos queridos, irmãozinhos de pesquisa, Leonardo Botega, Fernando V. Duarte, Allan Oliveira, Guilherme Cruz pelo companheirismo sempre.

À amiga Cinthia Brigante, sua mãe Iara e seu pai Walter pela afetuosa acolhida na chegada a São Carlos e pela amizade maravilhosa que se formou.

Ao querido casal Dona Maria e Senhor Benedito Simões que me acolheram como filha enquanto estive em São Carlos. Meu carinho e agradecimento a toda sua linda família. Vibrações de carinho à saudosa e querida Simone Simões.

À Márcia Maria Cavalcanti Gonçalves pelo profissionalismo, carinho, atenção, dedicação e força transformadora.

Aos meus amigos fofos Luciana Pereira de Assis e Alessandro Vivas de Andrade pelas irrefragáveis mostras de amizade durante os períodos mais difíceis da minha vida.

À Marlúcia Aguiar pelo apoio, pelo profissionalismo e pela amizade.

Às amigas de Diamantina Thais Magalhães, Débora Vitorino, Caroline Queiroz e Marlene Souza pelo companheirismo e incentivo nos momentos difíceis.

Ao meu amado pai Otávio, minha amada e saudosa mãe Dinorah e minha amada irmã Marisa, alicerces da minha vida.

À minha família que amo muito. Estamos sempre juntos em coração e recebo suas energias de luz: tia Dalila, tia Darziza, primas Diva, Clarissa, Alessandra e Letícia e a amiga-irmã Claudia A. Alegreti.

Gratidão a Deus, à vida, a todas as conquistas, problemas e obstáculos durante os estudos deste trabalho, pois foram eles que me transformaram no que hoje sou.

*O primeiro degrau para a sabedoria é a humildade.*

*Tomás de Aquino*

# RESUMO

Consciência da situação ou consciência situacional (*Situation Awareness – SAW*) é um conceito amplamente utilizado em áreas que requerem tomada de decisão crítica, e se refere à habilidade de um indivíduo ou equipe de percepção, compreensão e antecipação de estado futuro de uma situação corrente, que é influenciada pela dinamicidade e natureza crítica de eventos. SAW é considerada como principal precursora do processo decisório. Em domínios, por exemplo, de resposta à emergência, obter e manter SAW requer do operador humano grande esforço, pela sobrecarga cognitiva exigida na atividade, alto nível de estresse que envolve o atendimento, turnos exaustivos que podem refletir negativamente no processo de atendimento e conseqüentemente no processo decisório como um todo. Sistemas de apoio à tomada de decisão que contemplam aspectos da SAW podem contribuir no enriquecimento e manutenção da SAW do operador e no processo decisório. Diante desse contexto, este trabalho apresenta um Modelo Preditivo de Situações para sistematizar o desenvolvimento de módulos de apoio a SAW de operadores humanos em sistemas de resposta à emergência, que prevê a utilização de modelos de atendimento e protocolos das instituições atuando como situações prototípicas. Objetivamente o modelo propõe a previsão e ou a identificação prematura da situação em tempo real ao atendimento da emergência. Conjuntamente foi desenvolvido um Modelo Conceitual que norteou a construção do Modelo Preditivo e servirá como base a outros desenvolvimentos. Atualmente os denominados sensores humanos e sensores sociais, especialmente de redes sociais, estão sendo utilizados, de forma crescente, como importantes fontes de informação para a melhor compreensão de situações em diferentes áreas de aplicação. No domínio de resposta à emergência, objeto de estudo desta tese, os sensores humanos são a principal fonte de informação, sobre a qual técnicas de aprendizagem de máquina como classificadores de texto foram aplicadas com resultados muito positivos. Para ser validado, o Modelo Preditivo de Situações foi implementado com a criação de um vocabulário baseado nos modelos decisórios reais da Polícia Militar do Estado de São Paulo (PMESP) e com o desenvolvimento de algoritmos de dois métodos classificadores (*Bag of Words* e *Naïve Bayes*). Testes foram realizados com quatro tipos diferentes de instâncias de entrada (frases). Para todas as métricas analisadas (precisão, acurácia e cobertura) os testes demonstraram superioridade do algoritmo *Naïve Bayes*. A diferença entre a taxa de acerto em relação ao algoritmo *Bag of Word* para a classe de instâncias com maior grau de dificuldade de identificação foi superior a 37%. Tais resultados demonstraram bom potencial do Modelo Preditivo de Situações de colaborar com os sistemas já existentes de atendimento emergencial, possibilitando maior efetividade no atendimento e diminuição da sobrecarga cognitiva a que são submetidos os atendentes cotidianamente.

**Palavras-chave:** Consciência da Situação, Tomada de Decisão, Resposta à Emergência, Aprendizado de Máquina, Classificação de Texto.

# ABSTRACT

Situation Awareness (SAW) is a concept widely used in areas that require critical decision making, and refers to the ability of an individual or team to perceive, understand and anticipate the future state of a current situation, which is influenced by the dynamicity and critical nature of events. SAW is considered as the main precursor of the decision-making process. In the emergency response area, obtaining and maintaining SAW requires a great effort from the human operator, the cognitive overload required in the activity, high level of stress involving the care, exhaustive shifts that may negatively reflect the care process and consequently the decision process as one all. Decision support systems that address aspects of the SAW can contribute to the enrichment and maintenance of the operator's SAW and in the decision-making process. Given this context, this work presents a Situational Predictive Model to systematize the development of modules to support the human operator's SAW in emergency response systems, which provides for the use of service models and protocols of institutions acting as prototypical situations. Objectively the model proposes the prediction and or the premature identification of the situation while the applicant has emergency assistance. A Conceptual Model was developed that guided the construction of the Predictive Model and will serve as basis for other developments. So-called human sensors and social sensors have become important sources of information especially in social networks. For the treatment of this data, text classifier methods are used with satisfactory results that cover the areas of education, security, entertainment, commercial, among others. For the emergency responses domain, object of this thesis, human sensors are the main source of information and machine learning techniques as text classifiers show important alternatives. In order to be validated, the Predictive Situations Model was implemented with the creation of a vocabulary based on the actual decision-making models of the Military Police of the State of São Paulo (PMESP) and the development of algorithms two classifying methods (Bag of Words and Naïve Bayes). Tests were performed with four different types of input instances (sentences). For all the metrics analyzed (accuracy, accuracy and coverage) the tests demonstrated superiority of the Naïve Bayes algorithm. The difference between the hit rates in relation to the Bag of Word algorithm for the class of instances with the highest degree of identification difficulty was over 37%. These results demonstrated good potential the Predictive Situations Model to collaborate with the existing systems of emergency services, allowing more attendance effectiveness and reduction of the cognitive overload that the attendants are routinely subjected to.

**Keywords:** Situation Awareness, Decision Making, Emergency Response, Machine Learning, Text Classification.

# LISTA DE FIGURAS

Figura 2.1 – Situações e percepções. Fonte: traduzido e adaptado de (KOKAR et al., 2009). .....	31
Figura 2.2 – Automaticidade no Processo Cognitivo. Fonte: traduzido e adaptado de (ENDSLEY e GARLAND, 2000). .....	31
Figura 2.3 - Modelo Ciclo Perceptual de Niesser. Fonte: traduzido e adaptado de (SMITH e HANCOCK, 1994). .....	34
Figura 2.4 - Modelo de SAW de Bedny e Meister. Fonte: traduzido e adaptado de (BEDNY e MEISTER, 1999). .....	35
Figura 2.5 - Modelo de Consciência de Situação em tomada de decisão dinâmica. Fonte: traduzido e adaptado de (ENDSLEY, 1995). .....	36
Figura 2.6 – Percepção dos elementos da situação. Fonte: adaptado de (ENDSLEY e JONES 2012), pág. 16. ....	37
Figura 2.7 - Círculo e triângulo. Mera ilustração representativa da Teoria do Modelo de (JOHNSON-LAIRD, 2001). .....	39
Figura 2.8 - Esquema, modelos mentais e SAW. Fonte: traduzido e adaptado de (ENDSLEY e JONES, 2012). ....	41
Figura 2.9 - O nível superior de SAW pode ser usado para direcionar a busca de dados e fornecer valores padrão quando a informação não está disponível. Fonte: traduzido e adaptado de (ENDSLEY, 2015). ....	45
Figura 3.1 – Inferência de fatos situacionais por meio de raciocínio automático. Fonte: traduzido e adaptado de (ROY e DAVENPORT, 2010). ....	56
Figura 3.2 - Valores propriedade delineados por notificações de evento. Fonte: traduzido e adaptado de (MATHEUS et al., 2003) .....	59
Figura 3.3 - Imagem da base de dados VIRAT com (A) representação trajetórias dos objetos de interesse e (B) instantâneo de vídeo do conjunto de dados, exibindo o mesmo cenário ao mesmo tempo. Fonte: adaptado de (FISCHER e BEYERER, 2012). .....	62
Figura 3.4 – Ilustração de cenário de consciência situacional marítima ideal. Fonte: traduzido e adaptado de (Snidaro et al., 2013). ....	63
Figura 3.5 - Ciclo para construir a rede situacional usando redes bayesianas. Fonte: traduzido e adaptado de (NADERPOUR et al., 2014) .....	65
Figura 3.6 – Movimentos normais no canal. Fonte: traduzido e adaptado de (CASTALDO et al., 2016). .....	66

Figura 3.7 – Fluxograma dos dados dos navios e representação da rede bayesiana. Fonte: traduzido e adaptado de (Castaldo et al., 2016). .....	67
Figura 3.8 – Gráfico em <i>word cloud</i> dos <i>tweets</i> para as categorias Alimento e Esporte. Fonte: (KOTEVSKA et al., 2016) .....	69
Figura 3.9 – Detecção de eventos pelo Twitter. Fonte: traduzido e adaptado de (DEIVA RAGAVI e USHARANI, 2014) .....	70
Figura 4.1 – Modelo conceitual para o desenvolvimento de ferramentas orientadas à consciência da situação. Fonte: elaborado pela autora. ....	83
Figura 4.2 – Descrição do problema. Fonte: elaborado pela autora.....	87
Figura 4.3 – Modelo Preditivo de Situações, considerando os sensores situacionais. Fonte: elaborado pela autora. ....	87
Figura 4.4 – Modelo Preditivo de Situações prevendo a não existência de Modelos Decisórios. Fonte: elaborado pela autora.....	91
Figura 5.1 – Sistema e Informações Operacionais de Polícia Militar do estado de São Paulo para atendimento às chamadas 190. Fonte: PMESP (ALMEIDA, 2013). .....	96
Figura 5.2 – Parte da árvore de decisão para situação de Roubo. Fonte: PMESP...	97
Figura 5.3 - Descrição do problema. Fonte: elaborada pela autora. ....	99
Figura 5.4 - Frequência de palavras visualizadas no <i>Word Cloud</i> . Fonte: elaborada pela autora. ....	101
Figura 5.5 - Histograma com palavras com frequência maior que 100 ocorrências. Fonte: elaborada pela autora. ....	102
Figura 5.6 - Parte da matriz fluxograma X vocabulário com a frequência de cada palavra no respectivo fluxograma (árvore de decisão). Fonte: elaborado pela autora. ....	104
Figura 5.7 - Algoritmo Bag of Words. Fonte: elaborado pela autora. ....	105
Figura 5.8 - Exemplo de tabela de frequência das palavras do vocabulário nas árvores de decisão (fluxogramas). Fonte: elaborado pela autora. ....	106
Figura 5.9 - Resultado da classificação Naïve Bayes. Fonte: elaborado pela autora. ....	108
Figura 5.10 - Algoritmo Naïve Bayes. Fonte elaborado pela autora.....	108
Figura 6.1 - Boxplot e p-valores para as métricas Precisão, Acurácia e Cobertura para o grupo de instâncias Teste_10. Fonte: elaborado pela autora. ..	120
Figura 6.2 – Recorte de parte da tabela de resultados para todos os grupos de instâncias analisados, com retorno dos cinco melhores em total de acertos. Fonte: elaborado pela autora.....	122

Figura A.1 - Listagem de todas as palavras de todos os fluxogramas. Fonte: elaborado pela autora. ....	141
Figura A.2 - Ordenação e eliminação de palavras repetidas. Fonte: elaborado pela autora. ....	142
Figura A.3 - Conversão para letras minúsculas. Fonte: elaborado pela autora. ....	142
Figura A.4 - Eliminação de caracteres especiais e acentos. Fonte: elaborado pela autora. ....	142
Figura A.5 - Remoção de espaços em branco. Fonte: elaborado pela autora. ....	143
Figura A.6 - Conversão de caracteres ISO8591-1 para UFT-8. Fonte: elaborado pela autora. ....	143
Figura A.7 – Vocabulário completo filtrado com 467 palavras. Fonte: elaborado pela autora. ....	144
Figura A.8 – Palavras extraídas da Árvore de Decisão “Animais/Aves”. Fonte: PMESP. ....	145
Figura A.1 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação envolvendo animais. Fonte: PMESP. ....	147
Figura A.2 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação envolvendo bebê. Fonte: PMESP. ....	148
Figura A.3 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação de homicídio. Fonte: PMESP. ....	149

# LISTA DE TABELAS

Tabela 1.1 - Principais procedimentos metodológicos utilizados nesta pesquisa. Fonte: elaborado pela autora. ....	21
Tabela 2.1 – Parte 1 - Definições de SAW. Fonte: traduzido e adaptado de (DOMINGUEZ, 1994). ....	29
Tabela 2.1 – Parte 2 - Definições de SAW. Fonte: traduzido e adaptado de (DOMINGUEZ, 1994). ....	30
Tabela 3.1 – Parte 1- Resumo dos trabalhos estudados: domínio e métodos utilizados para formalização da situação. Fonte: elaborado pela autora. ....	73
Tabela 3.1 – Parte 2- Resumo dos trabalhos estudados: domínio e métodos utilizados para formalização da situação. Fonte: elaborado pela autora. ....	74
Tabela 6.1 - Exemplo de Matriz de Confusão binária da classe “Furto”. Fonte: elaborada pela autora. ....	113
Tabela 6.2 – Grupos de instâncias de acordo com a composição das frases de entrada dos métodos de classificação. Fonte: elaborada pela autora..	115
Tabela 6.3 - P-valores obtidos após teste de hipótese. Fonte: elaborado pela autora. ....	119
Tabela 6.4 - Média de taxa de acerto dos algoritmos BW e NB para todos os grupos de instâncias analisados. Fonte: elaborado pela autora. ....	121

# SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO.....</b>	<b>13</b>
1.1 Contexto.....	13
1.2 Motivação e Questões de Pesquisa.....	15
1.3 Objetivos.....	19
1.4 Proposições de Pesquisa.....	20
1.5 Metodologia de Desenvolvimento.....	21
<b>1.5.1 Abordagem e Classificação da Pesquisa.....</b>	<b>21</b>
<b>1.5.2 Método de Pesquisa para Coleta de Dados.....</b>	<b>22</b>
<b>1.5.3 Análise dos Dados.....</b>	<b>22</b>
<b>1.5.4 Descrição das Atividades Metodológicas.....</b>	<b>22</b>
1.6 Organização do Texto.....	24
<b>CAPÍTULO 2 - FUNDAMENTOS CONCEITUAIS: CONSCIÊNCIA DA SITUAÇÃO, PROCESSAMENTO DE LINGUAGEM NATURAL E CLASSIFICADORES.....</b>	<b>26</b>
2.1 Considerações Iniciais.....	26
2.2 Consciência Situacional.....	27
<b>2.2.1 Modelos.....</b>	<b>33</b>
<b>2.2.2 Modelos Mentais e Consciência da Situação.....</b>	<b>39</b>
2.3 Elementos do Processamento de Linguagem Natural.....	46
<b>2.3.1 Classificadores.....</b>	<b>47</b>
2.4 Considerações Finais.....	50
<b>CAPÍTULO 3 - ESTADO DA ARTE EM MÉTODOS PARA FORMALIZAÇÃO DA SITUAÇÃO EM SISTEMAS DE APOIO À TOMADA DE DECISÃO.....</b>	<b>52</b>
3.1 Considerações Iniciais.....	53
3.2 Ontologias.....	55
3.3 Métodos Probabilísticos.....	61
3.4 - Classificadores.....	67
3.5 Discussões sobre Métodos de Formalização da Situação.....	75
3.6 Considerações Finais.....	78

<b>CAPÍTULO 4 - MODELO PARA AUTOMAÇÃO E PREVISÃO DA SITUAÇÃO COMO APOIO À CONSCIÊNCIA SITUACIONAL E AO PROCESSO DECISÓRIO EM SISTEMAS DE RESPOSTA À EMERGÊNCIA .....</b>	<b>80</b>
4.1 Considerações Iniciais.....	81
4.2 Modelo Conceitual.....	82
<b>4.2.1 Fontes de dados .....</b>	<b>83</b>
<b>4.2.2 Pré-Processamento de Dados e Informações .....</b>	<b>84</b>
<b>4.2.3 Métodos e Técnicas .....</b>	<b>84</b>
<b>4.2.4 Modelos Decisórios.....</b>	<b>85</b>
<b>4.2.5 Consciência Situacional .....</b>	<b>86</b>
4.3 Modelo Preditivo de Situações .....	86
<b>4.3.1 Entrada .....</b>	<b>88</b>
<b>4.3.2 Submódulo Indexador.....</b>	<b>89</b>
<b>4.3.3 Sentenças .....</b>	<b>89</b>
<b>4.3.4 Modelos Decisórios.....</b>	<b>90</b>
<b>4.3.5 Módulo Filtro.....</b>	<b>91</b>
<b>4.3.6 Vocabulário.....</b>	<b>92</b>
<b>4.3.7 Submódulo Classificador .....</b>	<b>92</b>
<b>4.3.8 <i>Ranking</i> de Situações .....</b>	<b>93</b>
4.4 Considerações Finais .....	93
<b>CAPÍTULO 5 - MODELO PREDITIVO DE SITUAÇÕES: APLICAÇÃO E MÉTODO .....</b>	<b>95</b>
5.1 Contextualização do Problema e Características dos Dados.....	95
5.2 Método de Trabalho .....	99
<b>5.2.1 Filtragem dos Dados e Criação do Vocabulário .....</b>	<b>100</b>
5.3 Métodos de Classificação.....	102
<b>5.3.1 Bag of Words .....</b>	<b>103</b>
<b>5.3.2 Naïve Bayes .....</b>	<b>106</b>
5.4 Considerações Finais .....	110
<b>CAPÍTULO 6 - RESULTADOS E DISCUSSÕES .....</b>	<b>111</b>
6.1 Considerações Iniciais.....	111
6.2 Matriz de Confusão e Métricas para Avaliação de Desempenho .....	112

6.3 Instâncias de Teste .....	114
6.4 Planejamento Experimental.....	116
6.5 Avaliações .....	118
6.6 Considerações Finais .....	123
<b>CAPÍTULO 7 - CONSIDERAÇÕES FINAIS.....</b>	<b>124</b>
7.1 Atendimento à questão de pesquisa, aos objetivos e às proposições .....	124
7.2 Limitações da Pesquisa.....	127
7.3 Trabalhos Futuros .....	127
<b>REFERÊNCIAS.....</b>	<b>129</b>
<b>APÊNDICE A - CÓDIGOS .....</b>	<b>129</b>
<b>APÊNDICE B - PUBLICAÇÕES .....</b>	<b>146</b>
<b>ANEXO A - EXEMPLOS DE ÁRVORES DE DECISÃO DA PMESP .....</b>	<b>147</b>

# Capítulo 1

## INTRODUÇÃO

---

*Entender qual é o nível de consciência situacional do operador dos sistemas de apoio a chamadas emergenciais é um fato determinante, pois o processo decisório será pautado no resultado da compreensão sobre os fatos em curso. Este capítulo inicia com a caracterização de uma situação crítica e a descrição dos aspectos que envolvem o atendimento de uma chamada emergencial. Posteriormente à contextualização do domínio do problema, serão descritos a Motivação e Questões de Pesquisa, Objetivos, Proposições de Pesquisa, Metodologia e Organização do texto.*

### 1.1 Contexto

Uma situação crítica que se caracteriza por eventos que envolvem risco à vida ou ao patrimônio e que ocorre em ambientes complexos e dinâmicos, quando necessita de ações de intervenção, exige que as decisões sejam fundamentadas na aquisição e na manutenção do entendimento desta situação durante todo o seu curso (ROY et al., 2007). Gerenciamento de catástrofes e desastres, operações de centros de comando e controle, atendimentos médicos emergenciais e atendimentos de chamadas 190 emergenciais são alguns exemplos de situações críticas cujo entendimento dependerá para que decisões apropriadas sejam tomadas em relação aos eventos em curso. O gerenciamento, controle e tomada de decisão sob situações dessa natureza são favorecidos quando empregado o uso de sistemas computacionais que apoiem estes processos (PATRICK e MORGAN, 2010).

As operações nestes ambientes necessitam de resposta rápida e eficaz além de ajustes nas ações, baseadas na evolução da situação e em seus processos que

se definem por informações comumente originadas de fontes heterogêneas. Algumas destas informações são adquiridas diretamente por meio de sensores humanos e físicos e outras são criadas a partir da inferência e da análise da situação (SEPPÄNEN et al., 2013).

Estas informações, adquiridas em tempo real durante o curso da situação, podem apresentar alto índice de imperfeição, visto a característica incerta destes ambientes. Informações imperfeitas, que não descrevem claramente a situação real do ambiente, reduzem a eficácia dos sistemas e podem contribuir negativamente para a formação do modelo mental de operadores humanos e conseqüentemente para o entendimento adequado da situação (BATINI e SCANNAPIECO, 2006).

A sobrecarga cognitiva dos envolvidos nestas situações, submetidos rotineiramente ao processamento de grande quantidade de informações em tempo real, de múltiplas fontes, além da realização de múltiplas tarefas simultaneamente, também contribui para a perda ou falta de consciência situacional. Este evento é um dos principais fatores de erros humanos em ambientes complexos e dinâmicos (NADERPOUR et al., 2014) (ENDSLEY e JONES, 2012) (BAUMGARTNER et al. 2014).

A habilidade de associar padrões entre os sinais críticos no ambiente e elementos no modelo mental é um fator dependente para se adquirir a *consciência situacional*, precursora do processo de tomada de decisão, definida como uma representação interna, que ocorre na mente do indivíduo sobre o estado do ambiente. É baseado nesta representação que as decisões são tomadas (ENDSLEY, 1988).

As principais classes de modelos de tomada de decisão são o modelo tradicional e o naturalista. O modelo tradicional analisa uma coleção de modelos axiomáticos de incerteza e risco. É uma abordagem simplista e limitada para aplicação em ambientes complexos. Assume que o tomador de decisão tem todo o tempo necessário para análise completa de todas as alternativas (ROY et al., 2007).

Ambientes complexos, dinâmicos e incertos compõem um cenário típico às características decisórias *naturalistas*, que são processos decisórios em que indivíduos utilizam a experiência em seu campo de atuação. Nesse cenário, os problemas (situações) se apresentam mal estruturados, incertos, dinâmicos, com forte possibilidade de ocorrência de ações e reações em cadeia e pelo aspecto de restrição de tempo, usualmente não há possibilidade da geração de múltiplas

alternativas para que o decisor realize comparações (KLEIN, 1993) (WONG e BLANDFORD, 2001).

O processo decisório nesses ambientes apresenta distinção quando subdividido, segundo Leedom (2001), em domínio físico, de informação e cognitivo. No domínio físico centram-se as ações, objetos e eventos da situação, no domínio de informação os modelos e a representação dos dados e informações e no cognitivo o modelo mental para os processos de compreensão.

Sistemas para apoio à tomada de decisão necessitam prover elementos que tratem estas características do domínio, mantendo e enriquecendo a consciência situacional do decisor. Para tanto, devem possibilitar caracterização da situação pela estimativa de aspectos marcantes com base em dados reais recebidos, reconhecimento da situação pela sua classificação quanto ao tipo real ou hipotético, análise da situação pela estimativa e previsão das relações entre suas entidades (ou objetos) e suas implicações e projeção da situação, pela inferência de situações futuras com base na projeção dos dados atuais (LIGGINS et al., 2009).

Pesquisas na área de tomada de decisão objetivam apoiar o decisor humano por meio de técnicas computacionais que abranjam todos os aspectos referentes ao domínio (LEEDOM, 2001) e à situação (LIGGINS et al., 2009). Os trabalhos concentram-se, por exemplo, (a) na representação de situações, (b) aquisição, retomada e manutenção da consciência situacional do operador de tais sistemas, (c) simulação de ambientes e processos, (d) visualização e tratamento da qualidade das informações, (e) fusão de dados e informações, (f) modelagem e automação de processos decisórios.

## 1.2 Motivação e Questões de Pesquisa

Escolher a alternativa mais adequada de resgate, atendimento médico, atendimento policial, dentre outras é o resultado de processos de avaliação de situações emergenciais que conduziram à determinação de tais escolhas. A tomada de decisão para resposta à emergência pode ser definida como o resultado destes processos. Implica também em escolhas alternativas que necessitam ser

consideradas, ou seja, espera-se selecionar a opção que melhor se adapta à situação em curso (SHAN et al., 2012).

Em sistemas de atendimento a Chamadas 190<sup>1</sup> da Polícia Militar, por exemplo, cujos ambientes das ocorrências são dinâmicos, complexos e críticos, o operador recebe grande quantidade e variedade de informações, muitas vezes simultâneas, outras de maneira assíncrona e em tempo real do evento em curso. As informações relatadas são frequentemente transmitidas com alto índice de imprecisão e incerteza, podendo sobrecarregar cognitivamente o operador e comprometer a compreensão da situação que muitas vezes envolve risco à vida. As decisões podem ainda ser tomadas sob forte estresse e impacto do tempo.

Esta sobrecarga cognitiva pode intensificar-se pelos impactos emocionais que o atendente está exposto diante das diversas ocorrências atendidas e à característica multitarefa da atividade – o profissional tem contato com a tela do sistema, com o telefone, rádio, câmeras de vídeo dentre outros dispositivos (REZENDE, 2010).

Informações adquiridas em tempo real da situação em curso, em ambientes complexos e dinâmicos, podem apresentar alto índice de imperfeição. Informações imperfeitas que não descrevem claramente a situação real do ambiente reduzem a eficácia dos sistemas e podem contribuir negativamente para a formação do modelo mental do operador e conseqüentemente para o entendimento da situação (BATINI e SCANNAPIECO, 2006).

Em um sistema automatizado para apoio à tomada de decisão, o produto da avaliação da situação é um estado informacional, ou seja, é uma representação de uma situação real (LIGGINS et al., 2009).

Há um notável esforço em pesquisas nas áreas de consciência situacional e tomada de decisões emergenciais e médicas, áreas que se equiparam pela característica do ambiente dinâmico, pelos processos estocásticos, pelas decisões em sequência e por vezes com resultados incertos, pelo comportamento humano com relevante grau de impacto ao ambiente e pelo tempo como fator de alto nível de criticidade.

---

<sup>1</sup> Chamada 190 - serviço em que operadores humanos alternam-se no atendimento de emergências 24/7/365 (a qualquer momento, o ano todo) à população, nos âmbitos policial, bombeiros e ambulâncias (unificados na cidade de São Paulo em 2017).

Trabalhos como de Xiang e Poh (2002) que trataram aspectos importantes como a dinamicidade do tempo e propuseram um modelo para tratamento de problemas cardíacos e Matheus et. al (2005) apresentaram uma ontologia para a formalização de situações. Day et al. (2010) desenvolveram um sistema de apoio à decisão para agendamento de leitos em tempo real que afeta o uso dos seus recursos, Huq et al. (2011) apresentaram um sistema automatizado robótico que orienta pacientes com AVC (acidente vascular cerebral) a se exercitarem. Gocgun et al. (2011) desenvolveram um modelo matemático para decisões de triagem de pacientes para exame de tomografia computadorizada e Shan et al. (2012) trataram o processo de tomada de decisão de resposta à emergência por uma perspectiva de gestão de serviços. Kokar e Endsley (2012) desenvolveram um modelo de mapeamento cognitivo, Goulionis e Koutsiumaris (2010) propuseram um modelo para gerenciamento de terapias contra o câncer precoce da próstata e Laskowski (2013) e Alagoz et al. (2010) propuseram modelos para resolver problemas de tratamento clínico sequencial sob incerteza. Bennett e Hauser (2013) pesquisaram um simulador para explorar políticas de saúde com base em IA para “pensar como um médico”.

Todos os trabalhos mencionados são exemplos que contemplam apoio à tomada de decisão auxiliando nos aspectos da consciência situacional.

O gerenciamento, controle e tomada de decisão sob situações dessa natureza são favorecidos quando empregado o uso de sistemas computacionais que apoiem estes processos de forma colaborativa, complementando as capacidades humanas e auxiliando em suas limitações (ROY et al., 2007). No entanto a área ainda apresenta desafios e aspectos que até o momento não foram completamente examinados.

Questões como até que ponto a automação de sistemas de apoio à decisão deve atingir e onde se encontra o limite entre a decisão automática e a decisão humana ainda não foram respondidas.

Há um consenso de que os sistemas de apoio à tomada de decisão colaboram para que a sobrecarga cognitiva dos operadores seja mitigada e assim decisões mais assertivas possam ser tomadas. Sob este prisma, mesmo com a abundância de pesquisas e técnicas ainda há uma lacuna entre a automação e a decisão humana. Resta um vasto campo de desafios a serem investigados.

Uma prática bastante controversa e que também compõe a esfera de discussão das questões supracitadas é a terceirização do atendimento às Chamadas 190 no Brasil realizadas por civis - e que já é uma realidade em vários estados e cidades, com forte indicativo de prosseguir a outros locais do país. Esta tendência pode ser um indicativo ainda mais contundente da necessidade de sistemas de apoio à tomada de decisão e apoio à consciência situacional. Sistemas que apoiem o operador humano de sistemas de respostas à emergência, mais especificamente Chamadas 190, automatizando parte do processo decisório, colaborarão efetivamente para diminuir a sobrecarga cognitiva do operador humano e conseqüentemente poderão favorecer a consciência situacional.

Essa realidade corrobora com Frey e Osborne (2017) que analisaram a tendência da automação de vários serviços. Os pesquisadores avaliaram 702 profissões e cargos utilizando um classificador Gaussiano com o objetivo principal de analisar quais ocupações dos EUA estão em risco de extinguir diante a probabilidade de serem informatizadas.

A pesquisa mencionada aborda profundamente questões trabalhistas e salariais, mas no que tange ao contexto do estudo de caso dessa proposta destacam-se alguns aspectos, como a declaração de Brynjolfsson and McAfee (2012) que afirmam que a informatização está atingindo os domínios cognitivos e o valor elevado da probabilidade de informatização de cargos operadores da polícia, bombeiros e ambulâncias está estimado em 49%. Esse fato ratifica os questionamentos em relação ao limite da automação. No entanto, em especial a este serviço efetuado por civis e não policiais militares, eventualmente o questionamento do limite entre a decisão automática e a decisão humana, em se tratando de um tomador de decisão sem a experiência de um policial, o veredito poderia ser diferente.

Resposta à emergência, escopo desse trabalho, é uma área extremamente desafiadora e difícil pela sua própria natureza e finalidade, por envolver vidas e prover a preservação da natureza e do patrimônio público e privado.

A pesquisa desenvolvida empenha-se em contribuir nos aspectos da consciência situacional utilizando técnicas de inteligência artificial para o tratamento da imprevisibilidade das situações com base nos contextos que o operador das chamadas 190 dispõe em 60 árvores de decisão. O objetivo é responder a pergunta “O que está acontecendo?” que é a frase norteadora do atendimento emergencial,

de maneira automática e com o menor número de palavras relatadas. Ou seja, no menor tempo, antecipando a inferência da situação e limitando possíveis problemas nos níveis de consciência situacional e conseqüentemente, na tomada de decisão.

A identificação da situação ocorre na fase inicial do atendimento, de maneira precoce. Desta forma, limitações de qualidade das informações recebidas pelo atendente são impedidas de se propagarem ao longo do processo de aquisição da consciência situacional e conseqüentemente, do processo decisório – já que o primeiro é precursor da decisão. Outro aspecto benéfico pode ser verificado em relação ao tempo de atendimento. O atendente, no direcionamento correto de uma situação em curso, não terá perda no tempo alterando o tipo de árvore de decisão para outra apropriada à situação correta decorrente de relatos incertos.

### **1.3 Objetivos**

Sistemas de apoio à tomada de decisão fazem jus à sua denominação de “apoio” e limitam-se a indicar ao operador aspectos referentes à situação em curso e ou às alternativas em relação à decisão a ser estabelecida. Até o presente momento, o poder decisório encontra-se em mãos do operador humano. Assim sendo, ferramentas e métodos desenvolvidos nesta área buscam aprimorar o processo decisório como um todo a fim de apresentar as melhores alternativas ao decisor humano.

O processo decisório se inicia na percepção dos elementos relevantes do ambiente de acordo com o que está acontecendo, na compreensão da relação destes elementos no ambiente de acordo com seu modelo mental, construído com base em sua experiência, treinamento, personalidade, dentre outros aspectos, gerando uma projeção futura desta situação para só então uma decisão de fato ser tomada e uma ação ser executada.

O propósito basilar deste trabalho é o desenvolvimento de mecanismos de apoio à consciência situacional do operador de sistemas de resposta à emergência, que é um estado que antecede o processo de tomada de decisão, podendo desta forma torná-lo mais assertivo. A contribuição desta tese é apresentar um método eficiente para a identificação precoce da situação em tempo real, em ambiente

dinâmico com objetos sem padrão comportamental, com a criação de ferramental de orientação à implementação de módulos de consciência situacional para os sistemas de apoio à tomada de decisão.

Para que esta tese alcançasse seu objetivo desenvolveu-se um modelo conceitual e um modelo preditivo para automação e identificação precoce da situação para apoio à tomada de decisão em sistemas de respostas emergenciais.

Secundariamente, para que os modelos fossem validados, foram implementados dois métodos de classificação de texto, sendo eles *Bag of Words* e *Naïve Bayes* e seus resultados avaliados e comparados.

## 1.4 Proposições de Pesquisa

Baseado na revisão da literatura sobre fundamentos da consciência situacional, processos de tomada de decisão e métodos para formalização da situação, o presente estudo parte das seguintes proposições de pesquisa:

**P1:** A proposta do Modelo Conceitual facilita a compreensão das etapas fundamentais, bem como os elementos e ferramentas necessários ao desenvolvimento de aplicações que contemplem aspectos de consciência situacional.

**P2:** O Modelo Preditivo de Situações orienta a implementação de módulos de consciência situacional para os sistemas de apoio à tomada de decisão.

**P3:** O uso de técnicas de IA (Inteligência Artificial), como os classificadores de texto, podem auxiliar no processo de formalização de consciência situacional em sistemas de apoio à tomada de decisão de situações emergenciais, especialmente os que utilizam o diálogo como principal fonte de informações, tornando o processo mais eficiente e eficaz.

**P4:** A identificação preliminar da situação, no tipo de escopo da pesquisa, mitiga ou até mesmo impede que limitações de qualidade das informações se propaguem ao longo do processo decisório.

## 1.5 Metodologia de Desenvolvimento

Simplificadamente, a Tabela 1.1 apresenta os métodos utilizados para o desenvolvimento deste trabalho.

**Tabela 1.1 - Principais procedimentos metodológicos utilizados nesta pesquisa.**  
Fonte: elaborado pela autora.

Item	Classificação
Abordagem	Quantitativa, Exploratória e Experimental.
Método de Pesquisa	Estudo de Caso.
Instrumento para coleta de dados	Simulação de dados.
Método de análise	Métricas de qualidade de classificação: precisão, acurácia e cobertura. Teste de Wilcoxon Signed Rank (TAHERI e HESAMIAN, 2013) e estimador de Hodges–Lehmann (HODGES e LEHMANN, 1963).

### 1.5.1 Abordagem e Classificação da Pesquisa

A abordagem da pesquisa realizada neste trabalho é quantitativa, exploratória e experimental, visto que o objeto de estudo foi submetido à medição numérica para testar hipóteses no processo de interpretação (MARTINS, 2010).

A classificação exploratória é identificada na primeira etapa desta pesquisa, que objetivou conhecer o ambiente e o contexto do problema das situações críticas e das tomadas de decisões nestes domínios, além de todas as nuances dos aspectos da consciência situacional que envolviam as variadas esferas e todos os seus elementos.

Compreendida a problemática envolvendo operadores humanos, tomada de decisão, características das situações e de ambientes dinâmicos, e modelos de consciência situacional, a segunda etapa consistiu na concepção dos modelos. A concepção e criação dos Modelos Conceitual e Preditivo orientam o desenvolvimento de sistemas ou módulos para contemplar aspectos da consciência situacional em sistemas de apoio à tomada de decisão.

O caráter experimental desta pesquisa é identificado na terceira fase que submete os dados (parte reais, parte simulados) aos modelos e algoritmos implementados e testados.

### 1.5.2 Método de Pesquisa para Coleta de Dados

O método de pesquisa para coleta de dados utilizado neste trabalho foi um estudo de caso.

Estudo de caso é um tratamento empírico que investiga um acontecimento dentro de um contexto atual e possibilita profundo conhecimento sobre o mesmo, com a lógica do planejamento, da coleta e da análise de dados. Pode incluir tanto estudos de caso único quanto de múltiplos (MIGUEL, 2010) (YIN, 2001).

### 1.5.3 Análise dos Dados

Os algoritmos desenvolvidos foram submetidos a um conjunto de testes, analisando 54 classes. Cada classe representa uma árvore de decisão utilizada pela PMESP (Polícia Militar do estado de São Paulo) para atendimento do serviço 190.

Os dados resultantes dos métodos desenvolvidos foram analisados segundo as métricas Precisão, Acurácia e Cobertura para avaliar as possíveis diferenças entre o desempenho desses métodos. Realizou-se um planejamento experimental que possibilitou detectar diferenças significativas entre os métodos e estimar sua grandeza. Para evitar suposições de que os dados apresentassem uma distribuição normal, foi utilizado o teste de *Wilcoxon Signed Rank* (TAHERI e HESAMIAN, 2013).

As análises estatísticas foram realizadas com auxílio do software *RStudio version 3.3.2 (2016-10-31) Copyright (C) 2016 The R Foundation for Statistical Computing*.

### 1.5.4 Descrição das Atividades Metodológicas

Com o intuito de atingir os objetivos propostos nesta tese, a abordagem metodológica é composta das seguintes atividades:

## **Fase 1: Pesquisas sobre fundamentos em consciência situacional e o estado da arte**

Nesta primeira fase, foram realizados estudos e pesquisas sobre os fundamentos da consciência situacional: conceitos, aplicações, modelos. Esta fase foi intensa e longa, pois demandou tempo de aprendizagem e longas discussões acerca de detalhes das peculiaridades que envolvem os modelos propostos na literatura.

Simultaneamente pesquisou-se o estado da arte em métodos para formalização da situação que representa o elo para o tratamento computacional e automação dos aspectos da consciência situacional propriamente ditos nos sistemas de apoio à tomada de decisão. Analisou-se o tipo de contexto e domínio da situação, área de atuação e desenvolvimento, tipos de dados e informações tratadas.

Diante da variedade de métodos utilizados, cada um com seus atributos e limitações, um aspecto no último quinquênio se destacou: a importância do denominado *soft sensor* (sensor humano) nestes sistemas. Dados humanos, especialmente em redes sociais tornou-se fonte das pesquisas mais diversas.

Neste sentido, o resultado desses estudos convergiram à necessidade de um modelo que abordasse esse tipo de dado textual (gerado pela fonte humana) aliado ao modelo de consciência situacional fortemente aceito pela literatura (ENDSLEY, 1995) que trata da importância do modelo mental do humano neste processo.

A junção desses conceitos (dados textuais de sensores humanos, modelos de consciência situacional e modelo mental) originou a proposta fundamental deste trabalho.

## **Fase 2: Definição do Modelo Conceitual para o desenvolvimento de aplicações que contemplem aspectos da consciência situacional**

Diante dos estudos realizados na Fase 1, iniciou-se as atividades para a definição do Modelo Conceitual proposto para o desenvolvimento de aplicações que contemplem aspectos da consciência situacional.

Estudos a partir da literatura sobre métodos de classificação de texto possibilitaram a escolha dos dois métodos *Bag of Words* e *Naïve Bayes*. A escolha se deu em razão de serem dois métodos amplamente utilizados e com abordagens

diferentes. *Naïve Bayes* é um método probabilístico e *Bag of Word* não, seu resultado advém do somatório da frequência de termos em uma classe.

Considerou-se importante duas abordagens diferentes para uma análise não tendenciosa. Foram realizadas mais dois testes com o método *Redes Bayesianas* e *Redes Bayesianas com Suavização de Laplace*, porém como os resultados foram inferiores aos demais métodos, considerou-se irrelevante ao resultado final e optou-se por não apresentá-los.

### **Fase 3: Desenvolvimento do Modelo Preditivo de Situações**

O Modelo Preditivo de Situações foi desenvolvido a partir do Modelo Conceitual. O modelo é baseado em toda a pesquisa da Fase 1 e é a junção dos conceitos de dados textuais originados do que neste trabalho é denominado *sensor situacional*, o modelo de consciência situacional (ENDSLEY, 1995) e o conceito do modelo mental de Endsley (ENDSLEY e GARLAND, 2000).

### **Fase 4: Avaliação dos Resultados**

A implementação do modelo e a realização de um estudo de caso gerou uma base de dados com os quais foram realizados testes de desempenho dos algoritmos assim também como testes estatísticos.

As avaliações e testes foram realizados por meio de dados simulados, no entanto a base de referência é real e composta por árvores de decisão utilizadas e fornecidas pela PMESP. Detalhes de todo o experimento, testes e demonstração dos resultados são apresentados e detalhados no capítulo seis deste trabalho.

## **1.6 Organização do Texto**

Este trabalho está organizado e dividido em sete capítulos, sendo o primeiro composto pela introdução do tema, justificativas do trabalho, problemas de pesquisa desta tese, a solução proposta e contribuições, os objetivos, proposições de pesquisa e procedimentos metodológicos adotados.

O segundo capítulo apresenta os fundamentos conceituais utilizados nesta tese incluindo consciência situacional, modelos, modelos mentais, processamento de linguagem natural e classificadores de texto.

O capítulo três versa sobre o estado da arte em métodos para formalização da situação em sistemas de apoio à tomada de decisão e aborda as principais pesquisas sobre ontologias, métodos probabilísticos, incluindo classificadores de texto e discute sobre essas abordagens.

No capítulo quatro é apresentado o Modelo Conceitual proposto neste trabalho, assim como suas etapas que incluem: sensores situacionais, pré-processamento de dados e informações, métodos e técnicas, modelos decisórios e consciência situacional. E as fases: entrada, submódulo indexador, sentenças, modelos decisórios, módulo filtro, vocabulário, submódulo classificador e ranking de situações.

O capítulo cinco descreve o modelo preditivo de situações, as aplicações e métodos, com a contextualização do problema que será o estudo de caso e as características dos dados. Demonstra uma tela do sistema e árvore de decisão utilizada atualmente pela PMESP em atendimentos 190. Descreve também o método de trabalho, como foi realizada a filtragem dos dados para a criação do vocabulário e por fim os algoritmos de classificação.

O sexto capítulo apresenta as avaliações conduzidas para validação do Modelo Conceitual proposto e os Resultados da pesquisa. No sétimo e último capítulo são apresentadas as considerações finais desta tese, incluindo o atendimento à questão de pesquisa, aos objetivos e às proposições declaradas; as limitações da pesquisa; e os trabalhos futuros.

Com a conclusão da contextualização do problema, da solução proposta e da apresentação dos elementos que balizaram esta pesquisa, o capítulo subsequente consiste na descrição dos fundamentos bibliográficos utilizados neste trabalho.

# Capítulo 2

## FUNDAMENTOS CONCEITUAIS: CONSCIÊNCIA DA SITUAÇÃO, PROCESSAMENTO DE LINGUAGEM NATURAL E CLASSIFICADORES

---

*Consciência Situacional (Situation Awareness – SAW) é um componente natural da cognição humana. Sua importância nos processos decisórios e estratégicos passou a ser reconhecida e identificada por um piloto da Primeira Guerra Mundial. Este capítulo trata a formalização dos conceitos de SAW, diferenças entre propostas de modelos e a razão de ser considerada elemento fundamental na dinâmica da tomada de decisão. Na sequência o capítulo apresenta os conceitos de Processamento de Linguagem Natural e Classificadores, considerando a categoria de dados utilizados para o estudo de caso deste trabalho.*

### 2.1 Considerações Iniciais

Sistemas de apoio à tomada de decisão são comumente referenciados na literatura e seus propósitos atingem diversificados fins. Independente do grau de interação destes sistemas com o operador e o resultado final da análise das informações, a decisão final fica a cargo do ser humano. Desta forma, o escopo destes sistemas expande-se para além do apoio ao ato final

da decisão em si, mas sim, a todo o processo que se inicia desde a aquisição das informações para compreensão do que está acontecendo. Entender e manter este entendimento durante o processo decisório poderá ser o fator determinante para que decisões sejam tomadas de maneira assertiva.

A compreensão do que está ocorrendo é denominada consciência da situação e será tratada neste capítulo, incluindo os conceitos e elementos que balizaram a proposta de trabalho desta tese. Abrange os fundamentos que formalizam e estruturam o modelo adotado de SAW, assim como a proposta de outros autores, os fundamentos de processamento de linguagem natural e classificadores de texto. Tais conceitos e métodos reunidos criam um arcabouço computacional para que sistemas de apoio à tomada de decisão contemplem aspectos de SAW e assim contribuam para gerar fatores de impactos positivos nos processos decisórios.

## **2.2 Consciência Situacional**

O termo consciência situacional (*Situation Awareness - SAW*) tornou-se importante conceito na dinâmica da tomada de decisão humana. Afirmam Roy et al. (2007) que a existência do fenômeno é consensual, no entanto a definição em palavras concretas é controversa.

Consciência situacional tornou-se objeto de estudo científico formal na década de 1980, não obstante seja parte natural da cognição humana. Este elemento cognitivo origina de tempos pré-históricos, quando indubitavelmente o homem necessitava estar ciente dos sinais do ambiente para ter êxito em suas caçadas e evitar que se tornasse a presa (ENDSLEY e JONES, 2012).

Um dos precursores a reconhecer e identificar a importância de SAW foi o piloto alemão da Primeira Guerra Mundial Oswald Boekle que declarou a importância de “ganhar a consciência do inimigo antes que o inimigo ganhe uma consciência similar” e desenvolveu métodos para realizá-la (JEANNOT, 2000).

Ainda que com poucas variações, não há concordância unânime em relação à definição de consciência situacional (JEANNOT, 2000) (ROY et al., 2007). No entanto um conceito bastante aceito e frequente na literatura é proposto por Endsley (1995) como sendo a “percepção de elementos no ambiente em um volume de tempo e espaço, a compreensão do seu significado e a projeção de seu estado em um futuro próximo”. É estar ciente do que está acontecendo ao redor e entender o que as informações significam no momento presente e no futuro.

Dominguez (1994) relacionou diversos conceitos de SAW disponíveis na literatura, conforme mostra a Tabela 2.1 - Parte 1 e Tabela 2.1 - Parte 2, e os analisando, propôs a definição transcrita a seguir:

“SAW é a contínua extração das informações do ambiente, a integração destas informações com o conhecimento prévio para formar uma imagem mental coerente e o uso desta imagem para direcionar maiores explorações em um contínuo ciclo perceptivo, assim como antecipar eventos futuros”. (DOMINGUEZ, 1994)

A consciência situacional faz parte da cognição humana, e este “recurso” é acessado tão natural e automaticamente que sequer se percebe seu mecanismo nas ações rotineiras do dia a dia como atravessar a rua ou mudar a marcha do carro quando se está dirigindo. No entanto, quando aplicada a situações operacionais onde se deve estar cientes da situação por razões específicas é que se faz melhor definida. Um atleta necessita avaliar seu adversário e as condições do jogo para decidir as ações da partida; para uma ultrapassagem segura o motorista verifica todos os sinais no ambiente dinâmico da estrada (velocidade do seu carro, do carro à frente, tipo de faixa do asfalto, condições da estrada, dentre outros); um policial necessita analisar objetivamente o cenário, condições da vítima, do bandido, avaliar o resultado e todas as implicações da ação e só então a executa.

Sob este prisma, consciência situacional é geralmente conceituada no que se refere às metas e objetivos de uma função ou trabalho específicos. Enquanto os elementos da SAW nesta definição podem variar grandemente de um domínio para outro, a importância da SAW como base para tomada de decisão e execução se aplica a praticamente todos os campos de atuação.

Como área de pesquisa SAW é estudada nos mais diversos campos como, por exemplo, educação, condução de veículos, manutenção de equipamentos, operações militares, autodefesa, recreação, medicina, navegação, equipes esportivas, controladores de tráfego aéreo/aviação, navegação fluvial, usinas de energia, comando e controle policia/militares, gerenciamento de crises, resposta às emergências, dentre outros.

Áreas nas quais as decisões envolvem alto risco (de vida, segurança das pessoas) e alto custo (de equipamentos, garantia de fornecimento de serviços) e com fluxo elevado e variado de dados e informações, vindas de fontes heterogêneas, podem ser beneficiadas se os responsáveis puderem decidir mais rápida e assertivamente.

**Tabela 2.1 – Parte 1 - Definições de SAW. Fonte: traduzido e adaptado de (DOMINGUEZ, 1994).**

Definição	Autor(es)
A percepção contínua de um piloto e da aeronave em relação ao ambiente dinâmico de voos, ameaças e missões e a habilidade de previsão enquanto executa as tarefas baseadas na percepção.	Carroll, 1992
A habilidade de extrair, integrar, avaliar e agir de acordo com a informação relevante das tarefas.	Companion, Corso, Kass e Herschler, 1990
A percepção precisa de fatos e condições que afetam uma aeronave e sua tripulação.	Edens, 1991
A percepção precisa de fatos e condições que afetam uma aeronave e sua tripulação durante um período de tempo.	Schwartz, 1993
A percepção dos elementos no ambiente com um volume de tempo e espaço, a compreensão de seus significados e a projeção de seus estados em um futuro próximo.	Endsley, 1990
A apreciação global do piloto de seu "mundo" atual.	Gibson e Garret, 1990
A habilidade de permanecer ciente de tudo o que está acontecendo ao mesmo tempo e integrar este senso de consciência no que está fazendo no momento.	Haines e Flateau, 1992

**Tabela 2.1 – Parte 2 - Definições de SAW. Fonte: traduzido e adaptado de (DOMINGUEZ, 1994).**

Definição	Autor(es)
Onde se refere a consciência espacial... O que caracteriza a identidade da consciência ou o conhecimento do piloto da presença de ameaças e objetivos, (tão bem como) os parâmetros de desempenho do voo e do estado do motor. Quem está associado a qual responsabilidade, ou automação da consciência; que é o conhecimento de quem é o comando. Finalmente, quando significa a consciência temporal e o conhecimento de eventos como a missão evolui.	Hardwood, Barnett e Wickens, 1988
A habilidade de vislumbrar a posição corrente e próxima de forças amigas e inimigas.	Masters, Mc Taggart e Green, 1986
Consciência das condições e ameaças imediatas do entorno.	Morishige e Retelle, 1985
A habilidade de manter a percepção precisa do ambiente ao redor, interno e externo à aeronave assim como identificar problemas e/ou problemas em potencial, reconhecer a necessidade de ação, notar desvios da missão e manter consciência do desempenho das tarefas.	Prince e Salas, 1993
SAW significa que o piloto tem um entendimento integrado dos fatos que contribuirá para um voo seguro da aeronave sob condições normais e anormais.	Regal, Rogers e Boucek, 1988
SAW se refere à habilidade de ter consciência rápida daquelas características que se desenvolvem durante o voo.	Wickens, 1992.
O conhecimento do piloto sobre o ambiente à luz dos objetivos de sua missão.	Whitaker e Klein, 1988

SAW é o modelo interno decisório sobre o estado do meio ambiente (Figura 2.1) e com base nesta representação, o tomador de decisão pode definir o que fazer sobre a situação e realizar ações necessárias. Acontecimentos reais são vivenciadas por humanos ou capturadas por sensores, e estes (humanos e sensores) percebem os sinais mais relevantes que serão processados para que a compreensão da situação ocorra. Subsequentemente, compreendendo a situação em curso, a projeção do seu estado futuro poderá ser gerada.

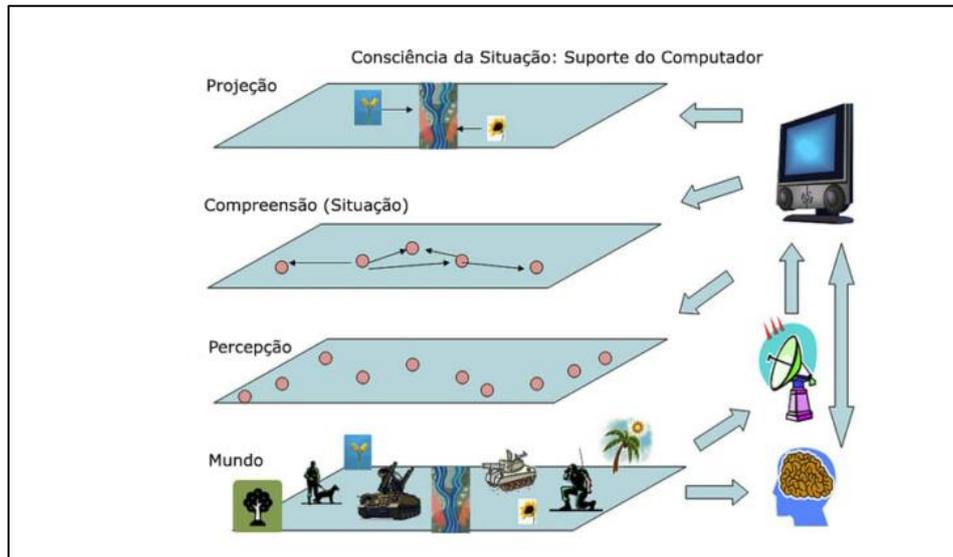


Figura 2.1 – Situações e percepções. Fonte: traduzido e adaptado de (KOKAR et al., 2009).

Representada como elemento precursor na tomada de decisão, uma boa SAW não garante necessariamente uma boa decisão. Contudo, melhorar a SAW aumenta a probabilidade de selecionar o curso de ação mais apropriado na maioria das situações, declaram Endsley e Jones (2012). Sistemas de apoio à tomada de decisão que contemplam aspectos de SAW podem contribuir para isto (Figura 2.2).

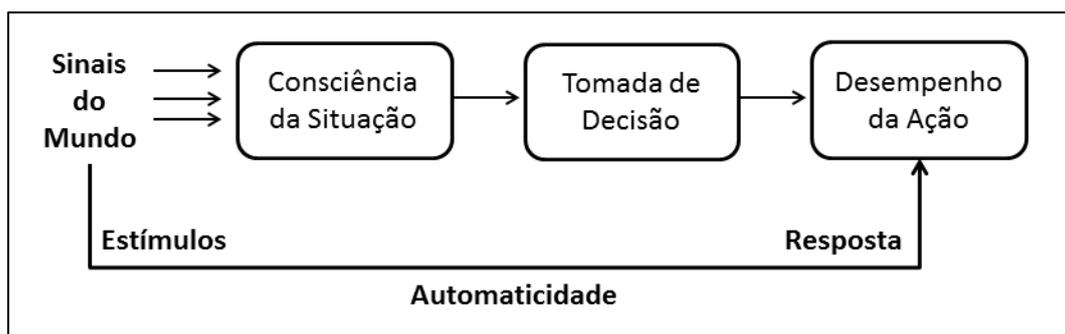


Figura 2.2 – Automaticidade no Processo Cognitivo. Fonte: traduzido e adaptado de (ENDSLEY e GARLAND, 2000).

No domínio da aviação, os controladores de tráfego aéreo geralmente se referem à SAW como “a imagem” – que é uma representação mental da situação na qual se baseiam suas decisões (ENDSLEY e GARLAND 2000).

Para os controladores, "ter a imagem" é o primeiro pré-requisito para realizar a tarefa. "Perder a imagem" é relatado como um dos maiores perigos para estes profissionais, uma vez que é a fonte de vários riscos, tais como não ser capaz de prever a evolução da situação, não conseguir detectar precocemente um problema ou um conflito, não escolher a melhor resolução e em casos extremos, não evitar que se estabeleçam incidentes ou acidentes (JEANNOT, 2000).

Inicialmente relacionado ao domínio da aviação, o potencial de generalização e expansão de estudos da SAW pode ser confirmado desde os trabalhos de Gaba et al. (1995) que descreveram uma forte analogia entre os requisitos de SAW na aviação e na anestesiologia. Ambos os campos implicam dinamismo, complexidade, uma alta carga de informação e risco de vida. No trabalho, os autores fornecem exemplos de situações reais ou simuladas do campo da anestesiologia em que a consciência da situação é fundamental para a prestação de cuidados de saúde ideais e mapeiam os elementos de consciência de situação em um modelo do processo cognitivo do anestesista. O estudo identificou três aspectos das situações nos quais o decisor deve permanecer consciente: sinais sutis, situações em evolução e elementos de conhecimento especial.

Esses aspectos, considerando a natureza das situações e dos domínios em que foram pontuadas, apresentam potencial para serem empregados em outros ambientes igualmente dinâmicos e complexos, como por exemplo, atendimento à emergência.

Estudos de Shelton et al. (2013) mostraram que na medicina, a consciência situacional é um fator humano de importância crítica para a segurança do paciente. No entanto, em simulação de treinamento clínico com objetivo de examinar e investigar estes fatores, a SAW não pode ser analisada minuciosamente, pela dinamicidade do ambiente. Esta constatação fez com que os pesquisadores projetassem uma ferramenta de medição de SAW, em tempo real, assim a cena em curso não seria interrompida. Denominada SPAM (*Situation Present Assessment Method – Método de Avaliação da Situação Presente*) o procedimento se define pela entrega de um questionário para o

participante que deve respondê-lo com precisão e rapidez. A diferença de tempo entre o recebimento do questionário e as respostas recebidas representam os valores de SAW. Estes dados são disponibilizados para o professor do curso, estendendo o escopo da ferramenta para o campo da educação médica.

No domínio da navegação, a importância da SAW se destaca, pois de acordo com Okazaki e Ohya (2012) existem poucos sistemas que controlam o tráfego marítimo como o sinal de intersecção no tráfego rodoviário. É função do marinheiro compreender o movimento de outros navios por meio do radar marítimo, bem como um piloto para o sistema de controle de voo. Detectar risco de colisão em fase inicial é fundamental, pois navios dispõem muito tempo para realizarem a manobra de virada, por exemplo.

Para que a consciência situacional do humano seja enriquecida e mesmo beneficiada em atividades como nas das áreas exemplificadas, sistemas de apoio à tomada de decisão necessitam contemplar seus elementos - e um dos grandes desafios é a “impossibilidade de quantificar alguns fatores de SAW, diante de sua natureza tão complexa” (VOGEL, 1994).

Com o propósito de transpor este e outros desafios, modelos de SAW foram desenvolvidos tratando diferentes especificidades. A seção 2.2.1 descreve o modelo de Endsley (1995), utilizado como fundamento deste trabalho, apontando aspectos que justificam sua escolha.

### **2.2.1 Modelos**

Os modelos conceituais de SAW são apresentados sob diferentes perspectivas. As discordâncias em torno desta questão transpõem um debate que se inicia desde a natureza de SAW como processo ou produto e se estende aos termos de sua análise psicológica subjacente, comentam Salmon et al. (2007) ao analisar e comparar os modelos de Endsley (1995), Smith e Hancock (1994) e Bedny e Meister (1999).

Em síntese, o modelo de Smith e Hancock (1994) descreve que a SAW não está nem no mundo nem na pessoa, mas sim na interação da pessoa com

o mundo, em modelos mentais realizados internamente, contendo informações sobre determinadas situações. Esse processo de interação direcionada e modificada continua em um caráter cíclico infinito (Figura 2.3). No modelo deste ciclo perceptual a interação com o mundo (denominado “explorações”) é direcionado por esquemas realizados internamente. O resultado da interação modifica os esquemas originais e assim direciona a exploração adicional.

Esses modelos mentais facilitam a antecipação de eventos situacionais, direcionando a atenção do indivíduo aos estímulos do ambiente e direcionando o curso de ação. O indivíduo então realiza verificações para confirmar se a evolução da situação está de acordo com as suas expectativas. Eventos inesperados servem para solicitar mais pesquisa e explicação, modificando o modelo existente do indivíduo.

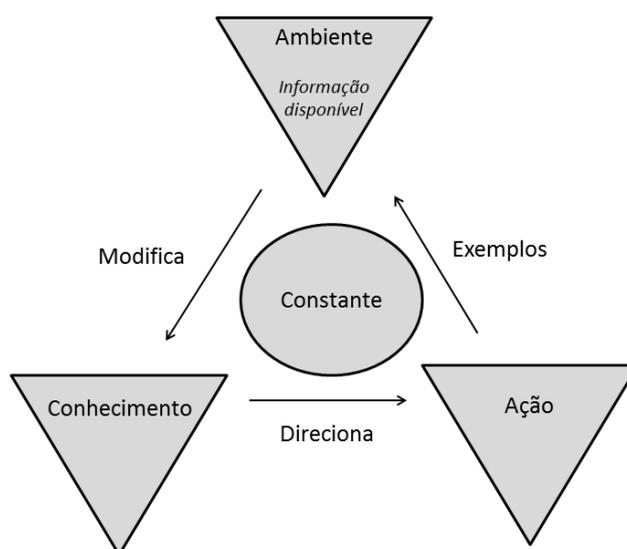


Figura 2.3 - Modelo Ciclo Perceptual de Niesser. Fonte: traduzido e adaptado de (SMITH e HANCOCK, 1994).

O modelo de Bedny e Meister (1999) representa a SAW por meio da teoria da atividade que descreve os diversos processos cognitivos associados ao comportamento humano. A teoria da atividade propõe que os indivíduos possuam objetivos que representam uma imagem ideal ou o estado final desejado da atividade, os motivos que direcionaram para o estado final e os

métodos ou ações das atividades que permitam a realização destes objetivos (Figura 2.4).

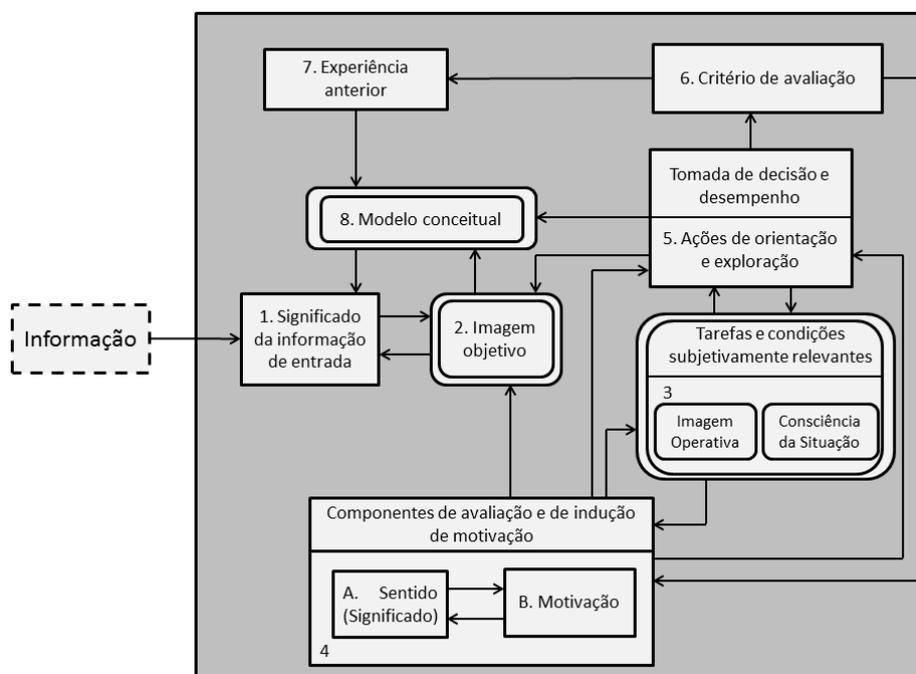


Figura 2.4 - Modelo de SAW de Bedny e Meister. Fonte: traduzido e adaptado de (BEDNY e MEISTER, 1999).

Cada um dos blocos do modelo desempenham uma função para o desenvolvimento e manutenção de SAW. A interpretação dos dados de entrada (bloco 1) é influenciado por objetivos de um indivíduo (bloco 2), pelo modelo conceitual da situação atual (bloco 8) e pela experiência anterior (bloco 7). Esta interpretação modifica os objetivos, experiências e o modelo conceitual da situação atual de um indivíduo. Recursos ambientais críticos são identificados (bloco 3) com base na importância da tarefa e motivação do indivíduo para o objetivo (bloco 4) que direciona sua interação com o mundo (bloco 5). A medida na qual os proventos individuais envolvem os objetivos da tarefa é determinada por seus objetivos (bloco 2) e por sua avaliação da situação atual (bloco 6). A experiência resultante da interação do indivíduo com o mundo é armazenada como experiência (bloco 7) e este informa o seu modelo conceitual (bloco 8) (BEDNY e MEISTER, 1999).

Os principais processos envolvidos na aquisição de SAW de acordo com o modelo são a função do bloco 8 (modelo conceitual), do bloco 2 (imagem do objetivo) e a função do bloco 3 (condições e tarefa relevantes).

O modelo de Endsley (1995), por ter se popularizado, tem recebido diversas críticas de uma parte da comunidade científica desde a sua criação (BEDNY e MEISTER, 1999) (DEKKER e HOLLNAGEL, 2004) (SORENSEN et al., 2011) e reconhecimento de outra (PARASURAMAN et al., 2008) (WICKENS, 2008). Endsley (2015) argumenta que na tentativa de apresentar novos modelos, surgem críticas por equívocos em relação ao entendimento do modelo proposto pela autora em 1995 e neste recente trabalho aponta, discute e defende cada ponto criticado de seu modelo.

Para Endsley (1995) a SAW é composta por um núcleo de elementos subsequentes que se relacionam em três níveis influenciados por fatores externos (individuais, dos sistemas e das tarefas) como demonstra a Figura 2.5.

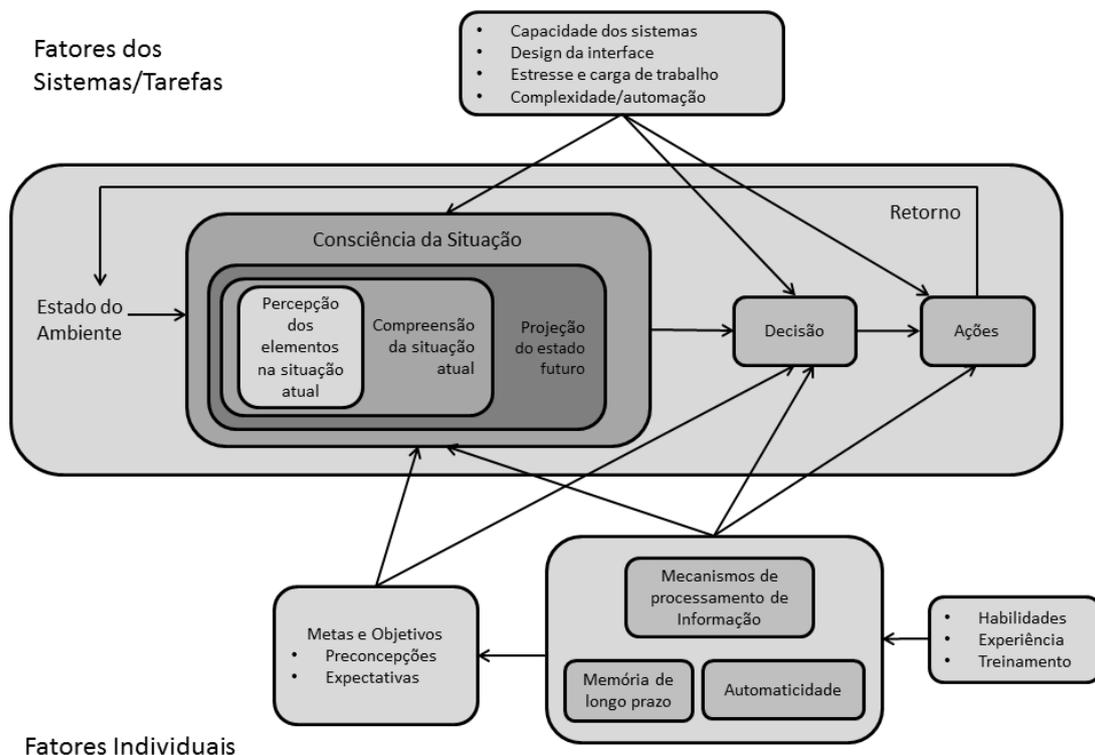


Figura 2.5 - Modelo de Consciência de Situação em tomada de decisão dinâmica. Fonte: traduzido e adaptado de (ENDSLEY, 1995).

- **Nível um: percepção dos elementos do ambiente.** A primeira etapa para obter SAW é perceber o estado, os atributos e a dinâmica dos elementos relevantes do ambiente. A percepção das informações se realiza de forma visual, auditiva, tátil, olfativa, pelo paladar, ou pela combinação dos sentidos. Seleciona os objetos relevantes do ambiente para a situação em curso (Figura 2.6).

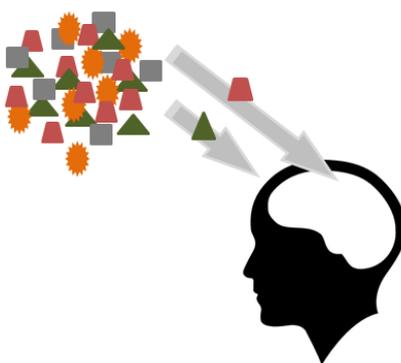


Figura 2.6 – Percepção dos elementos da situação. Fonte: adaptado de (ENDSLEY e JONES 2012), pág. 16.

- **Nível dois: compreensão da situação atual.** Refere-se a compreender o que os dados e as pistas/sinais percebidos no ambiente significam em relação às metas e objetivos da situação em curso.

- **Nível três: projeção do estado futuro.** Conhecendo o que os elementos são e o que eles significam em relação ao corrente objetivo, se é capaz de prever seu estado futuro.

Endsley e Connors (2008) esclarecem que o nível três só é alcançado tendo um bom entendimento da situação (Nível dois) e do funcionamento e da dinâmica do sistema computacional (caso exista) com o qual se está trabalhando. Entender a situação corrente para fazer projeções requer um entendimento muito bom do domínio – e isso implica em alto desenvolvimento

do modelo mental do humano. Sem a experiência suficiente ou um sistema de informação bem projetado, humanos podem falhar nas fases iniciais da SAW e nunca prosseguirem para o nível três.

O nível três requer a extrapolação de informações para frente no tempo para determinar como ele afetará os estados futuros do ambiente operacional. Funde o que o indivíduo sabe sobre a situação atual com os seus modelos mentais do sistema para prever o que é provável de acontecer em seguida (ENDSLEY e CONNORS, 2008).

No modelo de Endsley (1995) (Figura 2.5), os sinais ou estados do ambiente são entradas para o cerne formado pelos níveis de percepção, compreensão e projeção. Propositamente o módulo de decisão é separado do núcleo, pois SAW é considerada a principal precursora do processo de tomada de decisão. SAW é um modelo interno do operador sobre o estado do ambiente, e baseado nesta representação decide o que fazer sobre a situação em curso. A execução das ações sucede a tomada de decisão, porém, um pode afetar o outro de maneira cíclica.

Por serem módulos independentes, mesmo tendo sido tomada uma boa decisão, agentes externos podem influenciar uma ação, ocasionando uma reação ruim no ambiente. Assim, mesmo o processo tendo sido iniciado positivamente, poderá resultar algo não satisfatório.

O núcleo de SAW recebe influências diretas dos fatores externos dos indivíduos, das tarefas e dos sistemas. Os sistemas computacionais influenciam o núcleo SAW por meio de sua interface e capacidades, aliado às características de cada tarefa no que tange à sua complexidade, automação, estresse e carga de trabalho que provoca no indivíduo.

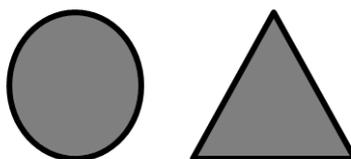
Fatores externos podem prover de maneira a facilitar ou dificultar que o indivíduo alcance a SAW. Este por sua vez influenciará o núcleo de SAW por meio de suas habilidades em executar as tarefas, seu nível de experiência e treinamento. Objetivos e metas direcionam a atenção para os sinais do ambiente assim como as expectativas de cada um, uma vez que as pessoas tendem a ver o que elas querem ver (ENDSLEY e JONES, 2012).

Certas particularidades do indivíduo que tanto podem influenciar a SAW positiva ou negativamente igualmente fazem parte de seu modelo mental e este aspecto será tratado na seção 2.2.2.

## **2.2.2 Modelos Mentais e Consciência da Situação**

Desenvolver mentalmente uma representação ou um modelo em pequena escala da realidade externa com as ações possíveis a serem realizadas, fará com que o indivíduo experimente várias alternativas antes de concluir qual é a melhor. Fará com que reaja a situações futuras antecipadamente, utilize o conhecimento de eventos passados para lidar com os eventos presentes e futuros de forma mais completa, segura e competente (CRAIK, 1943 apud JOHNSON-LAIRD 2013). Para Johnson-Laird (2004), um componente central do raciocínio é a geração de possibilidades, assim, as pessoas raciocinam por meio de modelos mentais.

O pressuposto fundamental da Teoria do Modelo de Johnson-Laird (JOHNSON-LAIRD, 2001) é que cada modelo mental representa o que é comum a um conjunto de possibilidades. Tomando como exemplo a afirmação “o triângulo está do lado direito de um círculo”, esta sentença possui um modelo mental comum representado pela Figura 2.7.



**Figura 2.7 - Círculo e triângulo. Mera ilustração representativa da Teoria do Modelo de (JOHNSON-LAIRD, 2001).**

O eixo da esquerda para a direita do modelo corresponde ao eixo da esquerda para a direita de uma cena, e à disposição do triângulo e do círculo no modelo corresponde à sua disposição para uma cena em que a afirmação é verdadeira. O modelo pode representar um número indefinido de possibilidades

(como por exemplo, tamanho dos objetos, cor, perspectiva, distância entre eles, ângulo de inclinação, dimensionalidade das formas, dentre outras) que não desempenham papel fundamental no modelo de raciocínio, pois apenas tem em comum o fato do triângulo estar à direita de um círculo (JOHNSON-LAIRD 2001).

Sob a ótica da ciência cognitiva, os modelos mentais têm propósitos de prever e explicar o comportamento de um sistema e servir como dispositivos mnemônicos para lembrar suas relações e eventos (WILLIAMS et al., 2014).

Na intersecção dos aspectos e conceitos teóricos dos modelos mentais, e tomando uma perspectiva funcional, Rouse e Morris (1986) os relacionaram a três finalidades: descrever, explicar e prever. Também os definiram como mecanismos pelos quais os seres humanos geram descrições da finalidade e forma de um sistema, explicações de seu funcionamento, observações e previsões de seus estados futuros. Rasmussen (1979) corrobora estes conceitos.

Endsley e Jones (2012) esclareceram que a mente humana trabalha com dois tipos de memórias: de curto prazo ou de trabalho e de longo prazo. Na memória de trabalho, apenas uma quantidade limitada de informações independentes pode ser mantida e manipulada e o indivíduo deve trabalhar ativamente para mantê-la ou irá perdê-la. A informação recém-recebida é combinada com conhecimentos existentes na memória de trabalho para criar uma nova imagem mental ou atualizá-la da mudança de situação. Estas partes de informação são então processadas e usadas para criar projeções do que possa ocorrer no futuro. Estas projeções ajudam o indivíduo a decidir que ações tomar para obter um resultado de acordo com seus objetivos e metas.

As memórias de longo prazo estruturam o conhecimento como esquemas e os modelos mentais desempenham um papel significativo na melhoria da SAW do indivíduo, conforme ilustra a Figura 2.8. Esquemas são “estados prototípicos do modelo mental que proporcionam maior eficiência de processamento” (ENDSLEY e JONES 2012). Segundo as autoras, as pessoas utilizam um processo definido como correspondência de padrões para vincular

as sugestões recebidas da situação atual aos esquemas e desta forma escolher pela melhor combinação das que estão disponíveis.

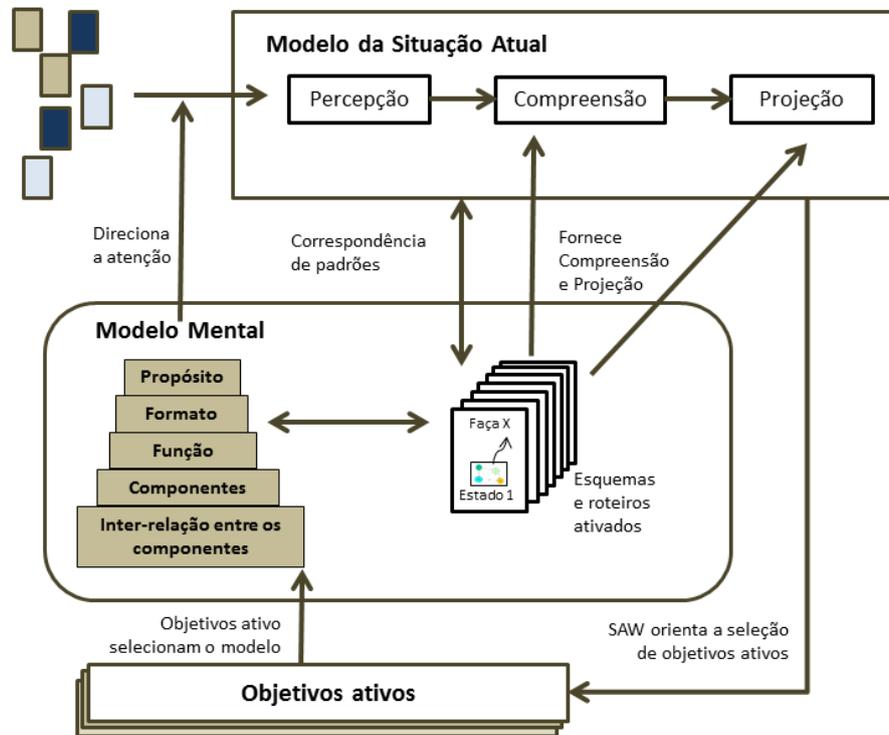


Figura 2.8 - Esquema, modelos mentais e SAW. Fonte: traduzido e adaptado de (ENDSLEY e JONES, 2012).

Uma das principais premissas do modelo de Endsley (1995) é o papel crucial dos modelos mentais no desenvolvimento e manutenção de SAW. As características do ambiente são mapeadas para os modelos mentais do humano que atuam como facilitadores do desenvolvimento de SAW, dirigindo a atenção para os elementos críticos no ambiente, integrando os elementos para a compreensão do seu significado e gerando possíveis estados futuros dos eventos (SALMON et al. 2007).

Jeannot (2000) estabeleceu uma relação entre modelos mentais e SAW quando declarou que SAW é uma construção interna e também faz parte da atividade cognitiva geral do indivíduo e desta forma SAW é influenciada e, em troca influencia outros constructos psicológicos.

A construção do modelo mental é um processo dinâmico em que o conhecimento e as habilidades adquiridas por um humano serão utilizados na construção e atualização dos futuros modelos mentais. Estes modelos mentais podem direcionar a aquisição do conhecimento com a interação dinâmica entre memória de longo prazo e memória de trabalho, favorecendo a junção dos diferentes níveis de informação (WICKENS e HOLLANDS, 2000). Estes modelos ajudam os indivíduos a determinarem quais informações são importantes para atender seus objetivos e ajudá-los na formação das suas expectativas – o que é esperado em resposta. Possibilitam atingir os altos níveis de SAW (compreensão e projeção do modelo de Endsley (1995) sem esforço demais das capacidades da memória de trabalho, por meio dos esquemas e *scripts* (conjunto de instruções a serem seguidas)).

Os esquemas permitem que o indivíduo classifique rapidamente e entenda uma informação percebida por meio de sensações anteriores. Uma vantagem significativa dos modelos mentais e esquemas é que as situações presentes não necessitam ser exatamente como as anteriores para que sejam reconhecidas. As pessoas podem usar mapeamento de caracterização que realiza uma espécie de ajuste entre as características percebidas na situação atual e a do esquema construído. A habilidade de associar padrões entre os sinais críticos no ambiente e elementos no modelo mental também é um fator dependente para se alcançar SAW. (ENDSLEY e JONES, 2012)

As pesquisadoras ainda elucidaram que os objetivos da situação auxiliam a determinar quais elementos do ambiente receberão maior atenção para que os indivíduos desempenhem suas tarefas. Diferentes objetivos conduzirão a interpretações completamente diferentes dos mesmos dados e informações. Modelos mentais, esquemas e *scripts* são desenvolvidos com o tempo e experiência em um domínio específico. Desta forma, indivíduos com menos experiência necessitarão dispende mais tempo em processamento mental classificando os acontecimentos para desenvolver SAW e provavelmente sobrecarregando a memória de trabalho, levando a falhas significativas de SAW.

São diversos os contextos em que a SAW desempenha um papel crítico, como por exemplo, para pilotos de aeronaves, controladores de tráfego aéreo, operadores de sistemas complexos, emergenciais, táticos e estratégicos. Em grandes e complexos sistemas como, por exemplo, o controle de uma usina

nuclear, os operadores devem observar diversos parâmetros, estados e quaisquer alterações de padrões que possam revelar anomalias sobre seu funcionamento e indícios que possam indicar sinais de previsão de seu estado futuro, diante de aspectos temporais e de restrição, como temperatura e pressão.

Para operadores de sistemas táticos e estratégicos citando bombeiros, policiais de unidades militares e de comando e controle, a SAW representa uma base importante para a tomada de decisão, favorecendo a identificação de pontos críticos e melhores ações em diferentes situações (WIRSTAD, 1981).

Sistemas de apoio à tomada de decisão em domínio que a SAW desempenha um papel fundamental possuem algumas características em comuns, como por exemplo, dados e informações dinâmicas como entrada e necessidade de reação ao ambiente cujo estado pode ser alterado independente da reação humana (KOKAR e ENDSLEY, 2012). Nestes domínios o operador deve ser capaz de ir além de simplesmente perceber os estados destes dados no ambiente, pois tipicamente as tarefas que exigem SAW abrangem demandas múltiplas e simultâneas, metas concorrentes e tarefas que deverão ser executadas sob alto nível de estresse e carga cognitiva (ENDSLEY, 2000).

Niessen et al. (1999) evidenciaram o proeminente papel dos modelos mentais para a manutenção da SAW, na função de antecipar estados futuros do ambiente e coordenar eventos simultâneos em curso e elencam cinco características peculiares:

1. propiciar a representação do mundo externo, refletindo as relações entre seus elementos de modo análogo;

2. ser o resultado de uma construção ativa das relações espaciais, temporais e finais significativas entre os elementos do ambiente da tarefa (esta construção é baseada na interação entre a situação externa do momento e os esquemas de domínio específico);

3. ser atualizável - a representação é modificada pela integração de novas informações que possam reorganizar a estrutura relacional do modelo mental;

4. ser incompleto no que se refere ao conteúdo de informação e transitório (a representação é construída por esquemas, a fim de servir funções correntes e não está armazenado na memória de longo prazo) e

5. ser manipulável por inferências para se fazer previsões, para decidir o processamento ou a ação a ser tomada.

Estas características foram baseadas em tarefas de controle de tráfego aéreo, objeto de estudo dos autores, mas podem ser estendidas aos modelos mentais na execução de outras tarefas e em outros domínios devido às características comuns destes ambientes.

Fundamentado no conceito de Endsley (2000) de que SAW é o estado atual de um modelo mental, à medida que a realização de uma tarefa evolui, diferentes partes do modelo mental na memória de longo prazo podem ser ativadas e anexadas à situação atual para ajudar a formar e manter a SAW. Desta forma, SAW pode refletir a estrutura e o conteúdo de um modelo mental especialmente em um ambiente dinâmico em que sua manutenção também é um processo em constante mudança.

Quando o indivíduo tem um desenvolvimento completo do modelo mental para sistemas ou domínio específicos, este modelo provê (i) orientação dinâmica da atenção para sinais críticos do ambiente, (ii) expectativas em relação aos estados futuros do ambiente baseado no mecanismo de projeção do modelo e (iii) uma direção, uma ligação entre classificações de situações reconhecidas e ações típicas, possibilitando que a tomada de decisão seja feita de maneira rápida (ENDSLEY e JONES, 2012).

Diante do exposto, o modelo de Endsley (ENDSLEY, 1995) se apresentou adequado para fundamentar este trabalho na proposta de um modelo para apoio e enriquecimento da SAW do operador de sistemas de

resposta à emergência, diante das características subsequentemente relacionadas:

- **modelo dinâmico e iterativo**: o ambiente real das ocorrências das situações analisadas é igualmente dinâmico. Baseado nos objetivos, compreensão e projeção da situação, o tomador de decisão tem a possibilidade de reavaliar os dados, ou seja, retornar ao nível um, como ilustra a Figura 2.9. Observado na imagem do modelo, a SAW modifica dinamicamente a interação com o mundo e, em seguida, a interação com o mundo modifica dinamicamente a SAW (ENDSLEY, 2015). A busca de novos dados relevantes para ampliar a compreensão da situação é tarefa comum em processo de tomada de decisão crítica, como por exemplo, em atendimentos emergenciais do serviço 190 da PMESP.

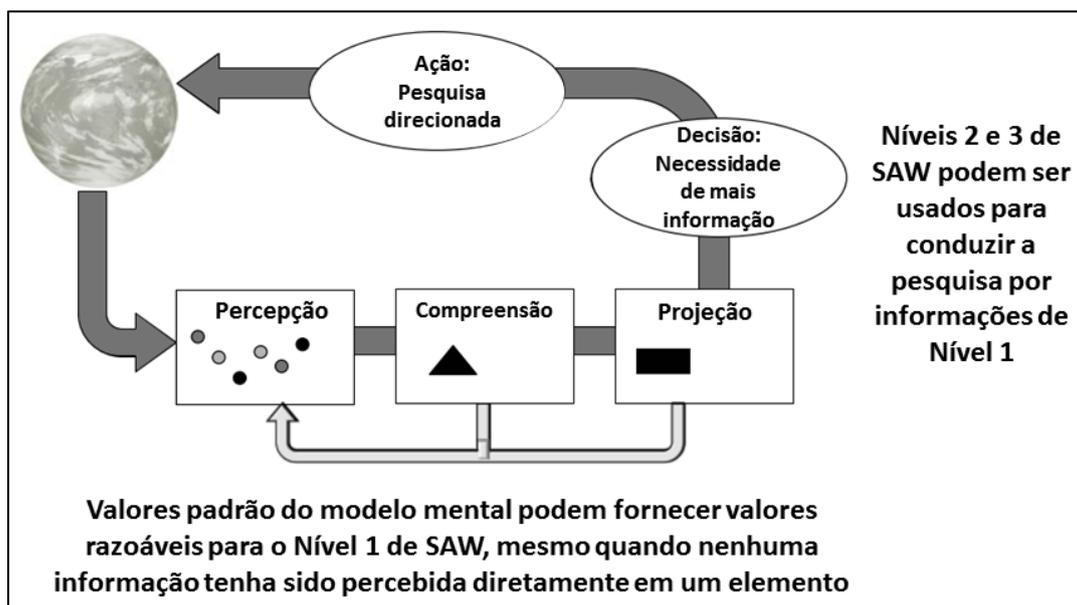


Figura 2.9 - O nível superior de SAW pode ser usado para direcionar a busca de dados e fornecer valores padrão quando a informação não está disponível.  
Fonte: traduzido e adaptado de (ENDSLEY, 2015).

- **modelo mental integrado:** diferentemente dos demais modelos, estrutura de forma clara, simples e direta a relação cíclica entre a obtenção de sinais e informações do ambiente e seu intercâmbio com a imagem da situação formada pelo modelo mental (resultado de experiências pessoais, treinamento, habilidades, dentre outros atributos). Para o contexto das situações, o ambiente em que elas ocorrerão e o perfil do humano, esta característica é bastante relevante.

- **módulos independentes:** para que sistemas de apoio à tomada de decisão contemplem aspectos da SAW humana, é necessária formalização por meio de tratamento computacional das atividades cognitivas humanas (tarefa não trivial). A estrutura modularizada contribui de maneira a facilitar que os elementos internos que compõe a SAW, assim como todos os outros que integram o processo decisório, possam ser tratados formal e computacionalmente.

## **2.3 Elementos do Processamento de Linguagem Natural**

Linguagem natural é a linguagem utilizada pelos seres humanos a fim de se comunicarem. O Processamento de Linguagem Natural (PLN) é uma área de pesquisa que reúne técnicas computacionais motivada a explorar a forma como os computadores podem ser usados para entender e manipular textos ou linguagem natural para uma variedade de tarefas e aplicações úteis (LIDDY, 2001) (CHOWDHURY, 2003). Essas tarefas frequentemente convergem para a aquisição do conhecimento em termos de classificação de texto, recuperação de informação e extração de informação (RUSSELL e NORVING 2013).

Compreender o que se quer expressar é o cerne de qualquer tarefa que envolve PLN, assegurou Chowdhury (2003) e o processo de construção de sistemas que compreendem linguagem natural abrange três principais questões: (i) os processos do pensamento, (ii) a representação e o significado do material linguístico e (iii) o conhecimento mundial que envolve a expressão.

Liddy (1998) recomenda que para entender a linguagem natural se faz necessário identificar os sete níveis que as pessoas utilizam para extrair o significado do texto ou da linguagem falada:

- Nível fonético ou fonológico: trata da pronúncia;
- Nível morfológico: trata as partes menores das palavras que trazem significado, os sufixos e prefixos;
- Nível lexical: trata com o significado lexical de palavras e partes da análise da fala;
- Nível sintático: trata a gramática e a estrutura das frases;
- Nível semântico: trata o significado de palavras e frases;
- Nível de discurso: trata da estrutura de diferentes tipos de texto usando estruturas de documentos;
- Nível pragmático: trata o conhecimento que origina do mundo exterior, ou seja, de fora do conteúdo do documento.

Um sistema de PLN ou que utiliza técnicas de PLN pode envolver todos ou alguns desses sete níveis de análise.

Um aspecto comum nas tarefas que envolvem PLN é o uso de modelos de linguagem - que são modelos que preveem a distribuição de probabilidade das expressões de linguagem. Quando expressa de forma escrita, o texto é composto por caracteres que podem ser letras, dígitos, pontuação e espaços. Assim, uma sequência de caracteres escritos de comprimento  $n$  é denominado de  $n$ -grama (da raiz grega para escrita ou letra). Um modelo de distribuição de probabilidade de  $n$  sequências de letras é então denominado de modelo de  $n$ -grama (RUSSELL e NORVING, 2013).

### **2.3.1 Classificadores**

Os classificadores executam a tarefa de classificar ou categorizar textos, ou seja, dado um tipo de texto ou um conjunto de caracteres, identificam a qual conjunto predefinido de classes este pertence.

A classificação de texto tem sido aplicada nas mais diversas áreas desde descoberta de *spam*, análise do sentimento do usuário de uma rede social para a classificação de um filme ou utilização para a revisão de um produto como pesquisa de opinião de empresas comerciais, dentre outros. Estes aspectos de pesquisas serão detalhados na seção 3.4.

Russell e Norving (2013) descreveram duas formas de tratar a classificação de texto: (i) como modelagem de linguagem, definindo um modelo de linguagem  $n$ -grama, utilizando um conjunto de treinamento disponível (aprendizagem supervisionada) e utilizando abordagem probabilística para classificação do texto; (ii) como abordagem de aprendizagem de máquina em que representa-se a mensagem como um conjunto de pares (documento e conjunto de categorias) de características ou valores e aplica-se um algoritmo de classificação  $h$  à característica de vetor  $X$ . Esta segunda forma foi a adotada neste trabalho.

O modelo *Bag of Words* é um dos métodos de representação mais populares para a categorização de objetos segundo Zhang e Jin (2010). Destaca-se por sua simplicidade e eficiência. É um método estático que ignora a ordem das palavras e a categorização é resultante do somatório da frequência das palavras em uma determinada classe (WALLACH, 2006).

Formalmente, no modelo *Bag of Words* um documento  $d_i$  é representado por um vetor de pesos para seus termos (ou palavras), ou seja,  $d_i = \{w_{1i}, \dots, w_{xi}\}$ , onde  $x$  é a cardinalidade do vocabulário da coleção  $D$  usada pelo classificador. Os pesos poderão ser definidos na forma *booleana*, onde  $w_{ki} = 1$  corresponde a ocorrência e  $w_{ki} = 0$  corresponde a ausência do termo  $w_k$  em um documento  $d_i$ . Comumente é utilizado um número real para esta representação.

Neste trabalho, quando o termo pesquisado ocorre no documento, sua frequência é acumulada. Detalhes do procedimento são descritos na seção 5.3.1.

O classificador *Naïve Bayes* parte do princípio bayesiano  $p(c_j|d) =$  probabilidade da classe  $c_j$ , dado que observamos  $d$ . Utiliza o teorema de Bayes, conforme Fórmula 2.1:

$$p(c_j|d) = \frac{p(d|c_j) p(c_j)}{p(d)} \quad (2.1)$$

- $p(c_j|d)$  = probabilidade de ocorrência de  $d$ , dado  $c_j$
- $p(d|c_j)$  = probabilidade de ocorrência de  $c_j$ , dado  $d$
- $p(c_j)$  = probabilidade de ocorrência de  $c_j$
- $p(d)$  = probabilidade de ocorrência de  $d$

Em uma simplificação do teorema, os classificadores *Naïve Bayes* assumem que os atributos possuem distribuições independentes e a equação fica conforme Fórmula 2.2 (DUDA et al., 2001):

$$p(d|c_j) = p(d_1|c_j) * p(d_2|c_j) * ... * p(d_n|c_j) \quad (2.2)$$

- A probabilidade de ocorrência de  $c_j$ , dado  $d$
- A probabilidade de ocorrência de  $c_j$ , dado recurso 1, multiplicado por ...
- A probabilidade de ocorrência de  $c_j$ , dado recurso 2, multiplicado por ...

O classificador *Naïve Bayes* possui benefícios que diante do objeto de estudo deste trabalho, duas delas convergem para sua escolha: rapidez no treinamento e rapidez na classificação (MCCALLUM e NIGAM, 1998).

O algoritmo *Naïve Bayes* está descrito na seção 5.3.1 e um exemplo está detalhado na seção 5.3.2.

## **2.4 Considerações Finais**

Não há um consenso quanto à definição de consciência situacional, embora o conceito de Endsley (ENDSLEY, 1995) seja bastante aceito, especialmente pela comunidade de Fatores Humanos.

Da mesma forma, seu modelo de três níveis, amplamente difundido em estudos e pesquisas, recebeu críticas especialmente de Salmon et al. (2007) em relação a sua natureza estática, em oposição ao modelo de Smith e Hancock (SMITH e HANCOCK, 1994) e de Bedny e Meister (BEDNY e MEISTER, 1999) descritos como dinâmicos. O primeiro, na medida em que se refere tanto ao processo (amostragem contínua do ambiente) como ao produto (esquema continuamente atualizado) de SAW e o segundo em que SAW modifica dinamicamente a interação com o mundo e em seguida, a interação com o mundo modifica dinamicamente a SAW. Este modelo é classificado como lógico por ir além da perspectiva estática feita pelo modelo de Endsley, porém como há falta de evidência empírica para apoiá-lo, recebe menos atenção.

Outra crítica em relação ao modelo de Endsley se refere ao fato de ser denominado produto de um processo de avaliação da situação. Esta distinção entre produto e processo, sugerindo ambos separadamente, constrói-se contraditória, uma vez que os níveis (percepção, compreensão e projeção) podem ser tomados como processos envolvidos para se atingir SAW (SALMON et al., 2007).

Paralelo às críticas, Stanton et al. (2010) declararam haver um elemento “da verdade” em todos eles e Endsley (ENDSLEY, 2015) publicou um trabalho argumentando e pontuando cada uma das críticas em relação ao seu modelo.

O modelo de Endsley (1995) por se tratar de um modelo de processos, o detalha como atividades internas individuais envolvidas no desenvolvimento de SAW. Por oferecer descrição pura e intuitiva, permite aos pesquisadores medir de forma simples os construtos e os requisitos de SAW em cada um dos três níveis.

Os sinais do ambiente que são “insumos” do modelo estão cada vez mais voltados ao fornecedor humano. Particularmente sistemas que utilizam postagens em redes sociais como principal fonte de informação dispensam especial atenção. Diante deste cenário, os classificadores de texto têm se tornado importantes métodos para estruturarem o arcabouço de ferramentas para o desenvolvimento de sistemas que consideram aspectos de SAW.

Neste contexto, outros sistemas em que o humano é a principal fonte de informação podem se beneficiar com o avanço das pesquisas nestas áreas.

SAW faz parte da atividade cognitiva humana e a todo o momento ela é desenvolvida em atividades de tomada de decisão rotineiras e cotidianas. Utilizar esta habilidade humana é um desafio quando transposta aos domínios computacionais.

Desenvolver sistemas munidos com esta capacidade, fornecendo meios ao tomador de decisão para que suas ações sejam realizadas de maneira mais segura e assertiva tem sido um dos grandes desafios desta área de pesquisa. Trabalhos buscam contribuir para a geração do que Liggins et al. (2009) denominaram “estado informacional” – o produto de um sistema automatizado para avaliação da situação.

O capítulo três apresenta o estado da arte em métodos para formalização da situação, aspecto fundamental no desenvolvimento de sistemas que se propõem a contemplar aspectos de SAW.

# Capítulo 3

## ESTADO DA ARTE EM MÉTODOS PARA FORMALIZAÇÃO DA SITUAÇÃO EM SISTEMAS DE APOIO À TOMADA DE DECISÃO

---

*Para que os sistemas de apoio à tomada de decisão contemplem aspectos da consciência situacional é preciso que a situação seja formalizada por meio de tratamento computacional. Este capítulo apresenta o estado da arte em métodos para formalização da situação na profusão de áreas em sistemas de apoio à tomada de decisão. O processo para localizar e selecionar os trabalhos presentes neste capítulo deu-se em cinco etapas, descritas concisamente a seguir. (1) Os trabalhos foram encontrados em repositórios digitais disponibilizadas no Portal da CAPES<sup>2</sup>. Utilizou-se o acesso remoto CAFE (Comunidade Acadêmica Federada), serviço provido pela Rede Nacional de Ensino e Pesquisa (RNP) às instituições participantes. (2) O termo de busca foi composto por palavras-chave referentes ao tema e escopo da pesquisa e como requisito deveria estar presente no título ou no resumo do artigo. (3) Foram inseridos filtros de pesquisa como de ano de publicação, tipo de material e área de conhecimento. (4). Os artigos, exibidos em ordem decrescente de ano de publicação, após escolha por título e resumo foram pré-selecionados. (5) Realizou-se então a leitura completa do artigo, para que sua classificação ou descarte de acordo com a relevância ao projeto.*

---

<sup>2</sup> <https://www.periodicos.capes.gov.br/>

### **3.1 Considerações Iniciais**

A Teoria da Situação de Barwise e Perry (1981) expressa que situações são onipresentes, pois ininterruptamente sempre se está em alguma situação. A atividade cognitiva humana categoriza estas situações em condições de objetos, tendo esses, atributos e estando permanentemente em relações uns com os outros em determinadas localizações em regiões conectadas no espaço e no tempo.

Matheus et al. (2005) referem-se a situações como partes delimitadas de uma realidade que as pessoas percebem e ou vivenciam, denominada de região de interesse. O conhecimento sobre os objetos da região de interesse da situação são tipicamente proporcionados por sensores (mecânico e ou humano) que executam a caracterização, monitoramento e identificação dos relacionamentos entre estes objetos. Tal conhecimento constitui a base para a consciência situacional (FURNO et al., 2010).

Uma vez que se estabeleça a situação composta por objetos, propriedades podem ser a eles associadas e técnicas e métodos incorporados para que se obtenham abordagens computacionais e se instaure conceitos de consciência situacional nos sistemas propostos de apoio à tomada de decisão (MATHEUS et al., 2003). Da mesma forma, o modelo de SAW de Endsley (ENDSLEY, 1995), adotado neste trabalho, ao segmentar em três níveis principais o processo de consciência situacional precedente ao ato da decisão e ação, sistematiza um processo cognitivo e assim como a Teoria da Situação, contribui para que a SAW possa ser formalizada computacionalmente. Contudo, “deve-se considerar um número potencialmente infinito de possíveis relações entre os objetos que compõem as situações”, e a aplicação de restrições, diretrizes e objetivos precisa e previamente definidos pelo usuário sinaliza possíveis soluções na modelagem das situações (MATHEUS et al., 2005). Isto porque informações obtidas destas situações podem fornecer respostas a apenas algumas, porém não todas, das perguntas sobre o mundo e o ambiente em que estão inseridas.

Associado ao amplo escopo de relações entre os objetos Gross et al. (2014) relatam que uma situação de interesse é geralmente obscurecida pelo grande volume de dados sem real importância para o observador. Todos esses dados podem ser segmentados em duas categorias, tendo sua origem de sensores físicos e de observação *soft* ou humana. Os dados de observação *soft* são coletados por meio da interação humana, ou inteligência humana (HUMINT - *Human Intelligence*).

Nesse âmbito, Kotevska et al. (2016) denominam como *sensores sociais* dados originados de redes sociais e *microblogging*. Na contemporaneidade trata-se de valiosos dados que, se capturados e analisados de forma adequada, podem contribuir significativamente com a consciência situacional em uma diversidade de contextos, como por exemplo, resposta a emergência, cidades inteligentes e sistemas para colaborações coletivas. Nesse contexto o emprego de técnicas de Aprendizado de Máquina, como os classificadores, favorecem a análise e manipulação desses dados.

Este capítulo apresenta o estado da arte para a representação da situação, expressa sob perspectivas semântica e probabilística bem como a maneira com que os dados de observações humanas têm sido analisados e tratados para a construção e identificação de situações. O capítulo também descreve os principais trabalhos que contribuíram e contribuem cientificamente para a área.

Considera-se relevante registrar que não há uma consonância entre os autores em relação à denominação conceitual utilizada para a expressão “representação formal da situação”, sendo comum nos trabalhos referências a “modelar a situação” (FENZA E FURNO, 2010), “representar” (DAPOIGNY e BARLATIER, 2013) ou mesmo “descrever” (TAMEA et al., 2014).

## **3.2 Ontologias**

O termo ontologia, quando empregado na representação do conhecimento, refere-se à descrição formal e explícita de conceitos. Seus componentes básicos são: classes (organizadas em uma taxonomia), relações (representam o tipo de interação entre os conceitos de um domínio), axiomas (usados para modelar sentenças sempre verdadeiras) e instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados) (NOY e KLEIN, 2004).

Ontologias são métodos semânticos utilizados na modelagem de situações, por meio de conhecimento de domínio estruturado em classes, objetos, seus atributos e relações entre eles (CHEN et al., 2004). Oferecem estrutura que proporciona decomposição compreensiva de complexas relações metafísicas como relações espaço-temporais, relações lógicas, relações psicológicas (intencionais), relações causais e vários tipos de relações de dependência. Essas relações são importantes para caracterizar, prever e inferir sobre situações, detectando inconsistências e derivando novos conhecimentos, compartilháveis e reutilizáveis por seres humanos e máquinas (YE et al., 2011).

No domínio marítimo Roy e Davenport (2010) desenvolveram um protótipo utilizando ontologias para representar o conhecimento pericial marítimo e assim apoiar o operador de sistemas de segurança na detecção de anomalias, classificação de navios de interesse e categorização de ameaças. O modelo ilustrado pela Figura 3.1 utiliza experiência de especialistas para construir a base do conhecimento de uma ontologia de vigilância marítima.

Os tópicos de interesse do domínio tratados pela ontologia incluem detecção de anomalias, embarcações de interesse, análise de situação e ameaças, narcotráfico marítimo e situações marítimas específicas. O papel do humano especialista neste ofício é de extrema importância e este trabalho não tem a finalidade substituí-los, afirmam os autores, pois são reiteradamente melhores do que os computadores na compreensão do significado dos dados em atividades de vigilância.

O trabalho objetivou detectar automaticamente, por meio de sensores, situações consideradas “normais” e “anômalas”, auxiliando os operadores em seu trabalho, que é sobrecarregado pela enorme quantidade de dados.

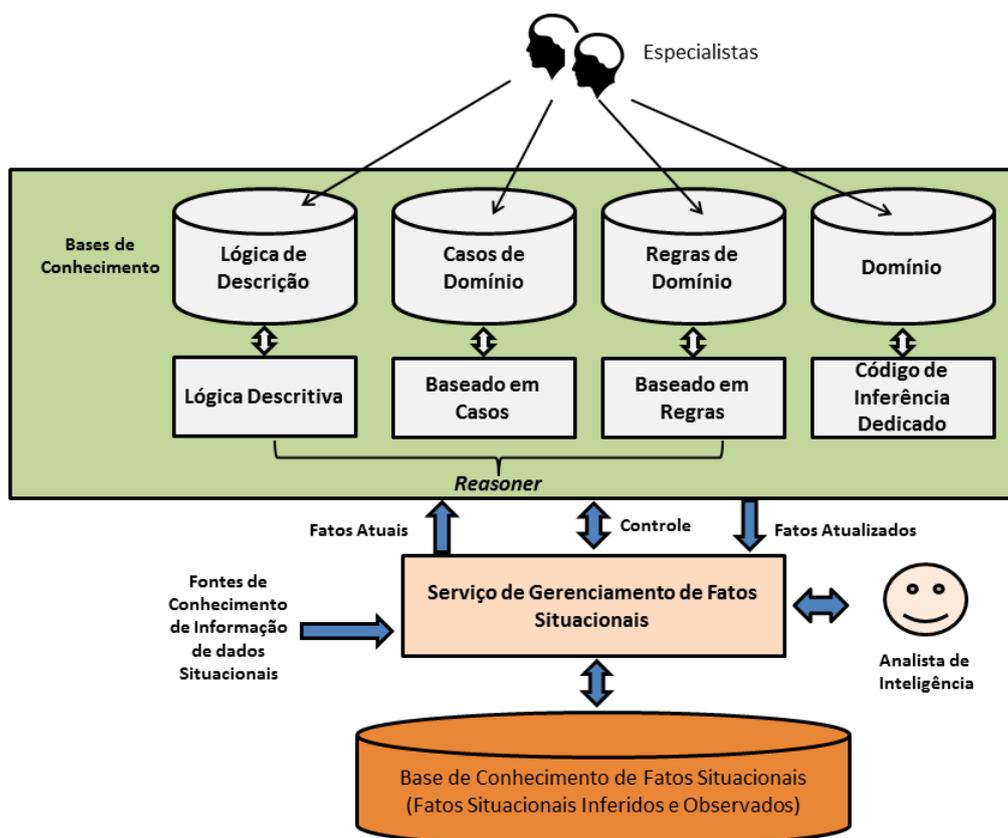


Figura 3.1 – Inferência de fatos situacionais por meio de raciocínio automático. Fonte: traduzido e adaptado de (ROY e DAVENPORT, 2010).

Como limitações a interface conceitual não suporta dados em tempo real e enfoca apenas “o que é verdadeiro” (grifos dos autores). Não utiliza dados de chamadas telefônicas de comunicações marítimas e serviços de tráfego 24 horas, assim como não trata fontes conflitantes – fatos contrários.

Baumgartner et al. (2014), discutiram a importância do apoio aos operadores dos grandes centros de controle, como por exemplo, de tráfego rodoviário, responsáveis por gerenciar ambientes altamente dinâmicos. Processam grandes quantidades de informações de fontes heterogêneas,

sobre um grande número de objetos do mundo real. Para gerenciar estes sistemas, os operadores humanos ficam vulneráveis à sobrecarga de informação. Desta forma, podem não estar cientes do significado geral da informação disponível e de suas implicações causada por uma decisão equivocada (como, por exemplo, interromper o tráfego de umas das pistas de uma rodovia sem considerar o alto fluxo de veículos que causará um engarrafamento), destacam os pesquisadores.

*BeAware* é a proposta de um *framework* que apoia o desenvolvimento de aplicações de conscientização de situação para centros de controle, integrando ontologias criadas em trabalho iniciado anteriormente (BAUMGARTNER et al., 2010). O sistema *BeAware* recebe dados de sensores de tráfego, de sistemas de manutenção e de sistemas de informações de acidentes. Baseado nos tipos de ocorrências reais, o sistema resulta sete tipos de situações, como por exemplo, (i) o acidente causa um engarrafamento e (ii) acidente em área de obras rodoviárias. Neste trabalho, embora o escopo seja o controle de tráfego para as decisões de operadores humanos em centros de controle, ainda não há suporte para controle em tempo real.

STO que é um acrônimo para *Situation Theory Ontology* (Ontologia da Teoria da Situação). Desenvolvida por Kokar et al. (2009), o trabalho parte do princípio colaborativo entre agentes computacionais e humanos para que compartilhem as responsabilidades nas tarefas - humanos desenvolvem modelos mentais e computadores executam modelos computacionais - modelos cognitivos da mesma situação. Esta ontologia foi utilizada para representar uma situação de emergência e para isso, a STO foi estendida por algumas classes, relações e regras adicionais.

Fenza et al. (2010) aplicaram a STO para modelar situações em domínio de segurança aeroportuária para apoiar a consciência situacional do operador de segurança para gerir esse ambiente em cenários de risco. Este trabalho baseia-se em uma abordagem cognitiva que modela a ontologia e introduz uma

arquitetura baseada em agentes para abordar o problema. As informações, como por exemplo, localização da aeronave, velocidade e tempo procedem de sensores, radares e controladores humanos. Neste trabalho, o sistema modela a situação envolvendo duas aeronaves em operações distintas em uma pista compartilhada e a ontologia extrai os elementos e as situações decorrentes deste cenário. O objetivo é obter todas as situações que possam ocorrer e as relações entre os objetos da cena. Com base nessa representação, os operadores de segurança aeroportuária podem decidir o que fazer em relação às situações de risco e realizar as ações necessárias.

SAW *Core Ontology* (Ontologia do Núcleo de SAW) é uma ontologia desenvolvida por Matheus et al. (2003), cuja estrutura principal compõe-se de uma *Situação* que contém um *Objetivo* ou uma coleção deles, *Objetos Situação* e *Relações*. *Objetos Situação* são todas as entidades em uma situação (físicas ou abstratas) com atributos e que podem participar de relacionamentos. As *Relações* definem os relacionamentos entre os conjuntos de *Objetos Situação*, como por exemplo, *naDireçãode* ( $X, Y$ ) – um objeto físico  $X$  está na direção do objeto físico  $Y$ . As *Relações* e *Atributos* precisam ser associados a valores que podem mudar com o tempo. Um novo *Valor Propriedade* é criado para cada atributo sempre que uma *Notificação de Evento* surgir e “afetar” o *Atributo/Relação*. *Notificações de Evento* são informações sobre eventos na situação do mundo real observado por uma fonte sensorial em um momento específico e que afetará uma relação ou atributo (de um *Objeto Situação* específico), definindo ou restringindo seu *Valor Propriedade*. Essas são as entidades que indicam mudança na situação, portanto veículos pelos quais as mudanças nos atributos e nas relações que representam a situação são realizadas (Figura 3.2).

Quando algum evento ocorre no tempo  $t1$ , resulta na geração de uma *Notificação de Evento*  $t1$  por algum sensor. Esta *Notificação de Evento* afeta o *Atributo1* ou *Objeto1*, atribuindo-lhes um valor e uma certeza instanciados como *Valor Propriedade* e assim sucessivamente até o último tempo. Mesmo

que nenhuma *Notificação de Evento* que afete a posição  $t+1$  seja recebida, é razoável supor que a posição do objeto mudou.

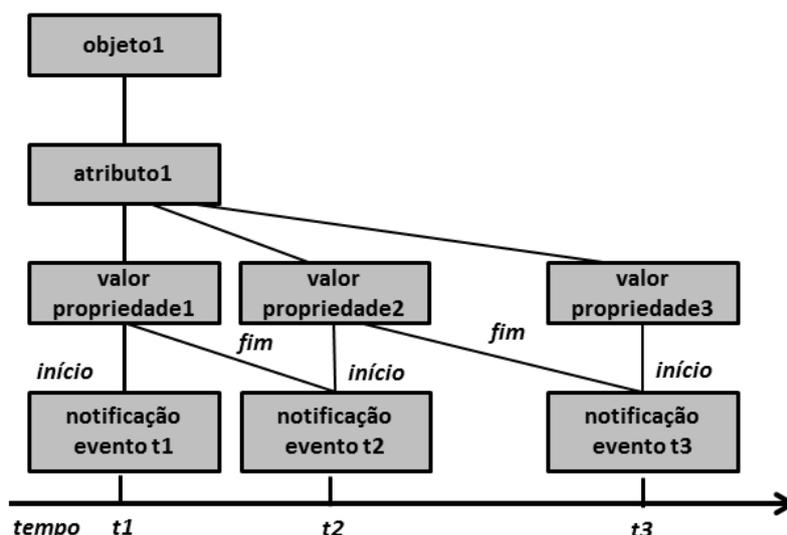


Figura 3.2 - Valores propriedade delineados por notificações de evento. Fonte: traduzido e adaptado de (MATHEUS et al., 2003)

Na ausência de informação adicional (por exemplo, trajetória, velocidade) estima-se que o objeto continue evoluir em seu último valor observado, até ser informado de outra forma, aumentando a incerteza sobre os dados com o passar do tempo (MATHEUS et al., 2003).

Em uma extensão do trabalho de Fenza et al. (2010), Furno et al. (2010) instanciam e integram dois meta-modelos ontológicos, STO e SAW Core *Ontology*. A possibilidade de modelar a incerteza foi considerada pela introdução da modelação *fuzzy* na STO, originando o que denominaram FSTO.

Os autores relatam a dificuldade em descrever o conhecimento incerto na construção de ontologias, visto a complexidade e dinamismo dos ambientes em que a consciência situacional se aplica e, assim sendo, a incerteza torna-se primária. Dentre os principais resultados desse trabalho destaca-se a melhora dos processos pré-táticos e táticos de segurança aeroportuária, a possibilidade que o operador de segurança fazer escolhas conscientes ao decidir executar o plano de ação iniciado ou decidir por planejar outro. Ressaltam que ter um alto

nível de consciência situacional é talvez o fator mais crítico para alcançar sucesso na aviação, referindo-se a Endsley e Garland (2000).

No domínio de C4ISR (*Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance* - Comando, Controle, Comunicações, Computadores, Inteligência, Vigilância e Reconhecimento), o modelo de dados JC3IEDM (*Joint Consultation, Command and Control Information Exchange Data Model* - Modelo de Dados de Intercâmbio de Informações de Consulta Conjunta, Comando e Controle) é amplamente utilizado para troca de informações que especifica o conjunto mínimo de dados necessário para a interoperabilidade entre os Sistemas de Informação de Comando e Controle e o mecanismo de troca de informações (VALIENTE et al., 2011).

Visto que a *SAW Core Ontology* suporta a modelagem da situação formalizando o conhecimento e JC3IEDM é um modelo de dados visando fornecer interoperabilidade entre sistemas de informação heterogêneos, a proposta da pesquisa de Valiente et al. (2011) é a integração de ambas para fornecer uma fusão de informação de alto nível para a percepção da situação nos sistemas C4ISR.

A fusão dos elementos do modelo de dados JC3IEDM (*Contexto, Dados Relatório, Item Objeto, Tipo Objeto, Localização*) com a ontologia *SAW Core Ontology* gerou uma nova ontologia *JC3IEDM OWL Ontology* que servirá de estrutura específica para aplicações de apoio à SAW em sistemas C4ISR. O desenvolvimento de sistemas que suportem o compartilhamento de consciência situacional é considerado uma tarefa fundamental para a melhoria da tomada de decisão em C4ISR.

Baumgartner et al. (2010) utilizaram a *SAW Core Ontology* para desenvolverem um protótipo de sistema de apoio ao operador de controle de tráfego rodoviário. A fonte de informações utilizada foi uma base real processada no sistema de Gestão de Tráfego Rodoviário (*RTM - Road Traffic Management*) da Agência Austríaca de Rodovias (*ASFINAG - Austrian*

*Highways Agency*). Essa base de dados fornece informações de trânsito a partir dos seguintes sistemas: gestão de obras rodoviárias, detecção de engarrafamentos, gestão de incidentes e uma estação de radiodifusão nacional. O sistema proposto classifica dez situações críticas e propõe uma análise global das informações por um conceito mais amplo denominado “vizinhança” e não apenas em um escopo de eventos observados de modo focalizado como por meio de uma “lupa”, discutem os autores.

### **3.3 Métodos Probabilísticos**

Além da quantidade, a qualidade dos dados e informações procedentes das fontes heterogêneas (sensores físicos e humanos) é um requisito fundamental no processo de aquisição e manutenção de consciência situacional dos operadores de sistemas de apoio à decisão.

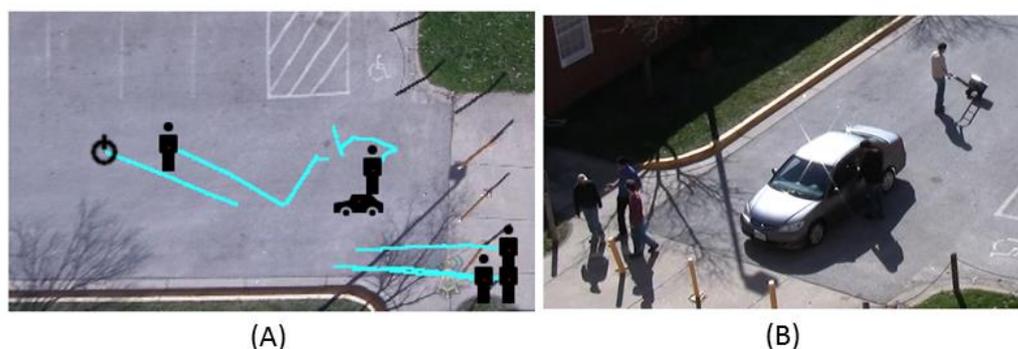
Realizar o tratamento desses dados e informações, a fim de monitorar situações em ambientes dinâmicos, requer ferramentas para a interpretação automática (em tempo real) dos eventos em curso a fim de prover suporte ao operador nos processos de percepção e avaliação da situação e tomada de decisão. Esses aspectos podem ser alcançados utilizando métodos lógicos e probabilísticos (FISCHER e BEYERER, 2012). No entanto, os autores sinalizam que limitar-se aos métodos lógicos dificulta o tratamento de problemas de qualidade como incertezas e ambiguidades. Neste sentido pode-se citar modelos probabilísticos como Redes Bayesianas e Redes Bayesianas Dinâmicas como exemplos de soluções para modelar dependência entre entidades observáveis (como a posição de um objeto) e entidades não observáveis (percepção de um comportamento anômalo).

Redes Bayesianas dinâmicas foram utilizadas para desenvolver um módulo de interpretação automática para apoiar o processo de avaliação da situação a fim de reconhecer situações de interesse no conjunto de dados de um sistema de vigilância com câmeras - VIRAT (*Video and Image Retrieval*

and Analysis Tool - Ferramenta de Recuperação e Análise de Vídeo e Imagem) (FISCHER e BEYERER 2012).

As situações são caracterizadas por informações coletadas ao longo de um período de tempo, porém elas só existem em um ponto especial desse tempo. Sua existência no próximo ponto tem de ser verificada novamente. Devido à semântica das situações, existe uma série de dependências entre elas, sendo assim, situações podem ser inferidas a partir de outras situações, por exemplo, se uma pessoa está perto de um carro e acaba de aparecer, a situação que poderia ser inferida é “a pessoa está saindo do carro”. Várias situações podem existir em paralelo ou a existência de uma situação pode excluir a existência de outra situação, ou ainda uma situação pode ser dependente da outra, elucidam Fischer e Beyerer (2012).

No sistema VIRAT, para calcular a probabilidade de existência de todos os tipos de situações, as dependências citadas entre as mesmas têm de ser modeladas. Para tanto, dois tipos de cálculos de probabilidade de situações são possíveis: de nível um (inferência direta dos valores dos atributos do espaço de configuração) e de níveis superiores (probabilidade das situações dependentes da probabilidade de outras situações). Atributos dos objetos das imagens de vídeo da base VIRAT como, por exemplo, identificação, posição, tipo (pessoa, veículo ou outro) são extraídos de imagens, como exemplifica a Figuras 3.3, então um espaço de configuração ( $\omega$ ) é definido.



**Figura 3.3 - Imagem da base de dados VIRAT com (A) representação trajetórias dos objetos de interesse e (B) instantâneo de vídeo do conjunto de dados, exibindo o mesmo cenário ao mesmo tempo. Fonte: adaptado de (FISCHER e BEYERER, 2012).**

A partir desses atributos é possível inferir a existência de uma situação de nível um (se elas são verdadeiras ou falsas), como por exemplo, “uma pessoa está perto de um veículo ou se a sua distância ao veículo é inferior a certo limiar”.

Pilato et al. (2012) propuseram um sistema de vigilância marinha que recebe dados de sensores (Figura 3.4) e infere sobre esses dados por meio de uma rede bayesiana e emite um alarme quando uma ameaça é detectada. Os dados fornecem informações tais como: tipo de embarcação, distância, tipo da manobra, velocidade do navio.

O sistema é composto também por uma ontologia que fornece a descrição do domínio (dos navios e suas características) e regras enriquecidas com fatores probabilísticos – usada para construir redes bayesianas. Redes bayesianas são utilizadas para representação da experiência e conhecimento dos especialistas, visando identificar situações críticas que a experiência os faz detectar facilmente, como por exemplo, o comportamento de navios piratas.

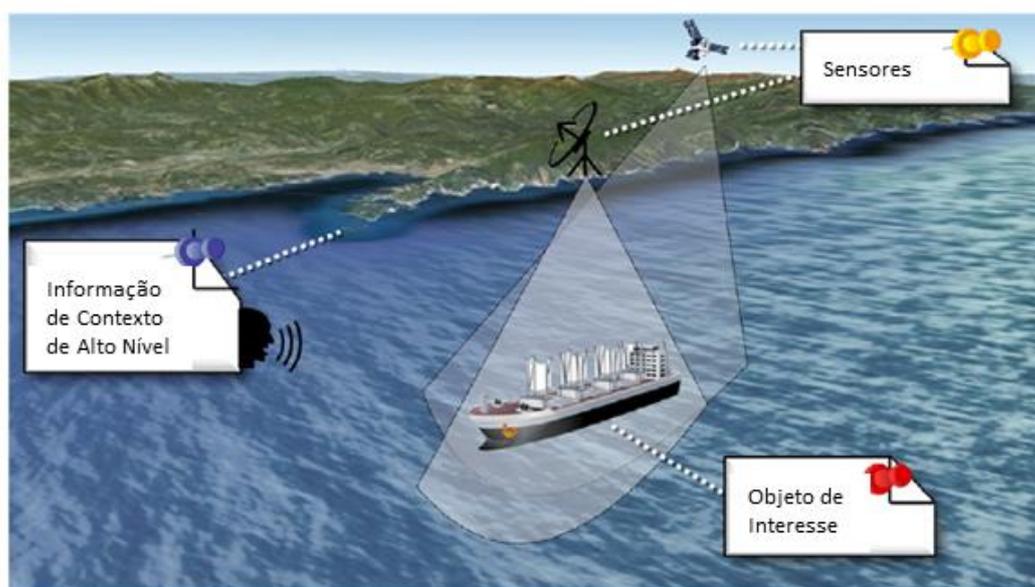


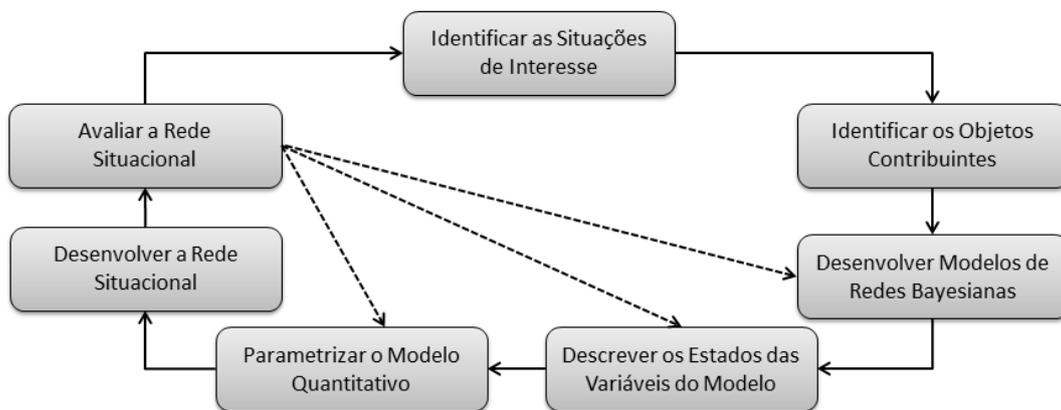
Figura 3.4 – Ilustração de cenário de consciência situacional marítima ideal.  
Fonte: traduzido e adaptado de (Snidaro et al., 2013).

Redes Bayesianas podem ser utilizadas para modelar as dependências entre entidades observáveis (por exemplo, a posição de um objeto) e entidades não observáveis (por exemplo, detecção de comportamento anormal) (FISCHER e BEYERER 2012).

Anagnostopoulos e Hadjiefthymiades (2008) comentam, fundamentado no conceito de Billings (1995), que SAW é uma abstração da mente humana e a observação de dados imperfeitos, como por exemplo, dados que induzem a avaliação errônea, conduz a raciocínios e estimativas inexatas, configurando um problema a aquisição de SAW. Nesta perspectiva propõem o uso da teoria *fuzzy* para inferir sobre essas situações. O benefício da integração métodos *fuzzy* e redes bayesianas podem ser observados em estudos nas áreas de segurança e monitoramento.

Em 2008, um acidente em uma usina química dos EUA motivou a pesquisa de Naderpour et al. (2014). O estudo relaciona situações anômalas na unidade da usina e as modelam em uma rede situacional utilizando redes bayesianas. As informações são fornecidas por sensores que reportam a temperatura dos tanques a cada minuto e sensores de temperatura ambiente exibem a temperatura da unidade de produção. Lógica *fuzzy* é utilizada para simular o pensamento do operador humano quando confrontado com estas situações anormais. A rede situacional bayesiana combinada com o sistema de lógica *fuzzy* torna possível ao operador avaliar dinamicamente situações anômalas para obter consciência da situação mais precisa, afirmam os pesquisadores.

A Figura 3.5 representa o método utilizado no desenvolvimento da rede bayesiana para situações anormais. Para identificar essas situações perigosas, realizou-se uma análise combinando procedimentos de engenharia cognitiva e métodos de identificação de perigos. Estes métodos incluem: a observação do desempenho do operador, a análise dos materiais escritos e a documentação, a elucidação de especialistas e os questionários formais.



**Figura 3.5 - Ciclo para construir a rede situacional usando redes bayesianas.**  
Fonte: traduzido e adaptado de (NADERPOUR et al., 2014)

O trabalho usa redes bayesianas e modela situação com base em modelos de um operador de sala de controle, e depende também da projeção de nível de risco para mostrar se a situação é anormal ou não, fornecendo prioridades, ou seja, uma situação pode prevalecer sobre outra requerendo mais atenção. Desta forma pode ajudar os operadores a entender a hierarquia das investigações - uma situação com maior risco tem prioridade sobre outras situações a serem investigadas.

O trabalho de Castaldo et al. (2016) propõe um sistema de apoio ao operador humano responsável pela segurança de grandes locais públicos de transporte como portos e aeroportos. O estudo de caso deste projeto serviu-se de um cenário marítimo real e os dados da trajetória dos navios, fornecidos por uma *web-cam* instalada na área, determinaram as situações normais e anômalas.

O sistema que se baseia em redes bayesianas e na análise de comportamentos dos objetos de interesse possibilita extrair uma imagem do cenário observado. Neste ambiente, comumente ocorrem diferentes tipos de interações entre diferentes embarcações (Figura 3.6).



Figura 3.6 – Movimentos normais no canal. Fonte: traduzido e adaptado de (CASTALDO et al., 2016).

Por exemplo, embarcações que transportam pessoas dividem o canal com navios de carga e pequenos barcos em alta velocidade e diferentes tipos de manobras e trajetórias. Assim sendo é necessário identificar situações que comprometam a segurança da área.

Para o treinamento da rede bayesiana, a trajetória e os dados normais de navegação das embarcações são usados para treinar um mapa topológico e diferentes modelos probabilísticos baseados em eventos, respectivamente. Uma vez que os modelos estão disponíveis, à medida que os dados dos novos navios são coletados, a inferência sobre os modelos pode ser realizada para a avaliação da situação (Figura 3.7).

Os pesquisadores afirmam que o sistema poderá representar um apoio útil para os operadores humanos encarregados de grandes áreas como portos, canais ou aeroportos, onde muitas vezes uma única pessoa é responsável pela segurança, controlando ao mesmo tempo muitos monitores e indicadores relacionados com área.

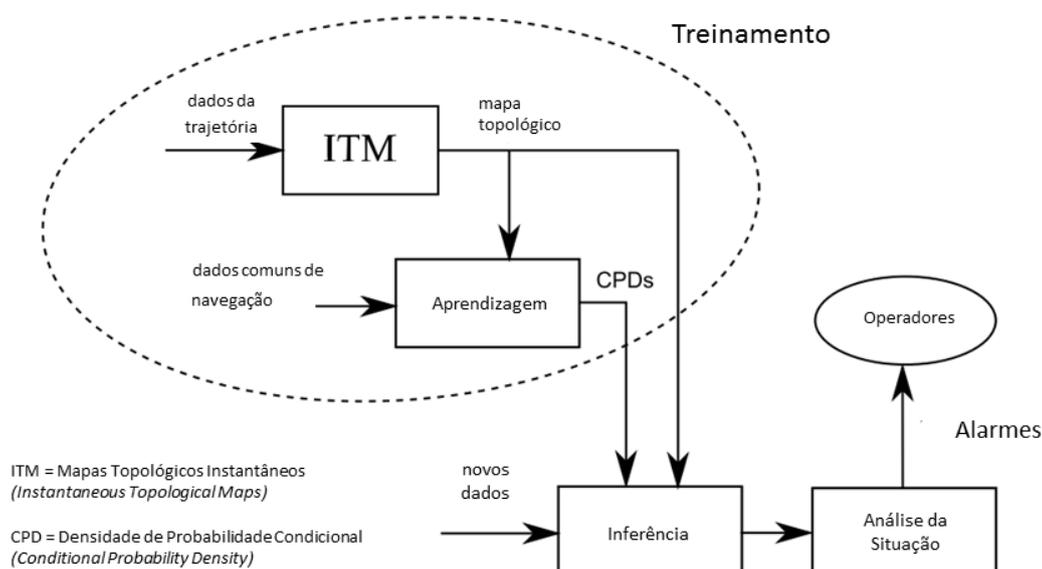


Figura 3.7 – Fluxograma dos dados dos navios e representação da rede bayesiana. Fonte: traduzido e adaptado de (Castaldo et al., 2016).

### 3.4 - Classificadores

Na última década a popularização das redes sociais e *microblogging* estabeleceu um novo significado e valor aos dados denominados HUMINT, ou seja, que são fornecidos pela inteligência humana (*Human Intelligence*). Assim como no domínio das mídias sociais, dados HUMINT também são fontes fundamentais nos ambientes de respostas à emergências que utilizam os relatos dos solicitantes como base para o atendimento. Este último ambiente compõe o objeto de estudo desta tese.

Os textos das postagens (*posts*) compartilhados nas mídias sociais, convergem-se em importantes fontes de informações, pois o conteúdo pode representar muito além de comentários para fins de entretenimento. Múltiplas áreas têm utilizado *posts* como elemento de pesquisa e empresas como

subsídio em tomadas de decisão estratégicas – mapeando a opinião do consumidor, por exemplo.

Fontes de dados *soft* ou dados de observadores humanos, ricos em conteúdo e grau de detalhamento, por estarem sob a forma textual ou em linguagem natural exigem processamento apropriado, como por exemplo, algoritmos de Aprendizagem de Máquina, para convertê-los em uma forma que sistemas computacionais possam utilizá-los (REIN e BIERMANN, 2013).

Khan et al. (2016) utilizaram o classificador *Naïve Bayes* e analisaram os comentários do repositório de vídeos *YouTube*, a fim de identificar a opinião dos autores sobre dois sistemas de celulares. Os pesquisadores partiram da suposição de que palavras em torno de palavras-chave relacionadas a uma opção em particular são suficientes para entender os sentimentos dos usuários, ou seja, para criar uma classe de comentários. Esta classe poderá servir de base de expressivo indicativo para auxílio na tomada de decisão em diversas áreas de negócios.

No contexto de Cidades Inteligentes, Kotevska et al. (2016) asseguram que integrar informações de diferentes fontes em tempo real é uma necessidade para tomar decisões mais inteligentes para a cidade como um todo. Sob esta perspectiva apresentam um trabalho de criação de um modelo para a identificação automática dos tópicos do *Twitter* como um sensor social objetivando contribuir para a consciência da situação sobre a cidade.

A classificação das publicações auxilia a localizar informações relevantes para tipos de categorias predefinidas, como por exemplo, *comida*, *esporte*, *música*, *saúde*, *viagem*, dentre outras. Utilizam técnicas de aprendizagem supervisionada com os classificadores *Naïve Bayes*, *Support Vector Machines* (SVM) e *Random Forest* (RF) para avaliar o desempenho em medidas de precisão. Para ilustração dos dados, a Figura 3.8 exibe as palavras em gráfico do tipo “*word cloud*” (nuvem de palavras) para as categorias de *tweets* Alimento e Esporte. Observa-se em ambas as nuvens de palavras, as



Pesquisadores indianos desenvolveram um sistema para detectar terremotos no Japão e notificar a população com até um minuto e meio de antecedência do evento ocorrer, classificando publicações do *Twitter* (DEIVA RAGAVI e USHARANI, 2014). Em se tratando de um desastre natural tão forte quanto terremotos, este tempo pode ser determinante para a vida das pessoas. Para o desenvolvimento do algoritmo de classificação foram definidos três grupos de características para cada *tweet*: o número de palavras da mensagem e a posição da palavra-chave, o número total de palavras e palavras que estão antes e após a palavra-chave. Usando o modelo obtido, o novo *tweet* é então classificado como correspondente a uma classe positiva – que atende àquele conjunto de requisitos - ou a uma classe negativa. A Figura 3.19 demonstra o modelo para detecção de eventos por meio de sensores humanos, pelo *Twitter*.

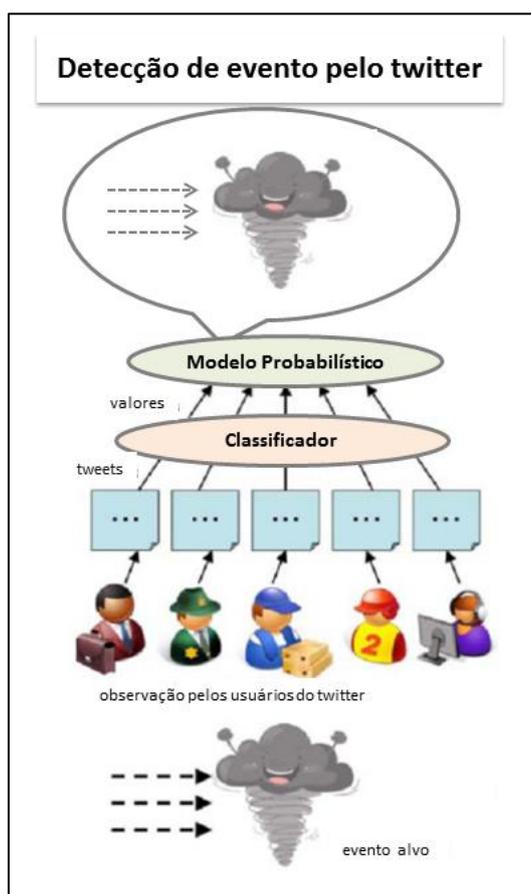


Figura 3.9 – Detecção de eventos pelo Twitter. Fonte: traduzido e adaptado de (DEIVA RAGAVI e USHARANI, 2014)

Os numerosos usuários do *microblogging* são considerados sensores que fornecem dados para que o modelo probabilístico espaço-temporal encontre a localização central do evento mais rápido do que a Agência Meteorológica do Japão, afirmam Deiva Ragavi e Usharani (2014).

Yin et al. (2012) propuseram uma arquitetura de sistema que utiliza mídias sociais para melhorar a consciência situacional no campo do atendimento a emergência. Mensagens do *Twitter* são enviadas durante desastres naturais e crises, fornecendo informações importantes sobre a conscientização da situação, como respostas da comunidade a avisos de emergência, notificação em tempo praticamente real dos incidentes e relatos de primeira mão do impacto de um incidente. Utilizaram os classificadores *Naïve Bayes* e SVM. O sistema fornece ferramenta de visualização com informações sobre as postagens como progressão das postagens minuto a minuto. Também estão disponíveis indicadores de volume de tweets capturados em um local especificado com cores, e os usuários podem clicar em cada marcador para exibir *tweets* recentes desse local. Essas informações, se extraídas e analisadas adequada e rapidamente, podem contribuir efetivamente para aumentar o nível de percepção da situação dos usuários deste sistema.

Aramaki et al. (2011) na mesma direção dos trabalhos supracitados, propuseram um sistema para extrair informações do *Twitter* e detectar epidemias de gripe utilizando o classificador SVM. Os autores destacam duas vantagens quanto ao uso de sensores e redes sociais: o grandioso volume de dados e a possibilidade da vigilância direta e em tempo real. Diante dos resultados, os autores concluem que os textos do *Twitter* refletem com precisão o mundo real, além disso, possibilitar a detecção de epidemias de gripe em estágio inicial é de suma importância para os alertas à população e aos órgãos responsáveis.

Para além do contexto de sensores sociais, o trabalho de Agarwal et al. (2015) apresenta novos conceitos para os testes escolares tradicionais. O algoritmo utiliza o classificador *Naïve Bayes*, se adapta ao conhecimento do usuário identificando o grau de dificuldade da questão posterior e auxilia-o gradualmente a melhorar seu conhecimento. O assunto da avaliação foi dividido em categorias e estas são parte de um teste, com cada parte composta por séries de perguntas objetivas de múltipla escolha. O usuário responde a um conjunto de perguntas e as próximas se adaptam dinamicamente de acordo com seus conhecimentos. O nível de dificuldade do próximo conjunto de questões das categorias do referido assunto é determinado usando o classificador. Ao final de três conjuntos de perguntas, o usuário recebe uma análise de desempenho nas categorias predefinidas. Baseada nessa revisão poderá trabalhar em suas áreas deficitárias e refazer o teste. Os autores acreditam que o uso prolongado desse tipo de sistema de teste pode colaborar com estudantes na melhora gradual de seu conhecimento.

Atualmente, dados originados de fontes heterogêneas compõem as informações que são a base para a formalização e compreensão da situação, no entanto com a propagação e popularização das redes sociais e *microblogging*, os sensores humanos têm se destacado como importante fonte de informações. Os trabalhos apresentados nesta seção estão sintetizados na Tabela 3.1 (Parte 1 e Parte 2) bem como os domínios de aplicação, os métodos utilizados para formalização da situação, dados de entrada e se fazem tratamento de situações em tempo real. Retratam e destacam o estado da arte e serviram de inspiração para o desenvolvimento desta tese.

**Tabela 3.1 – Parte 1- Resumo dos trabalhos estudados: domínio e métodos utilizados para formalização da situação. Fonte: elaborado pela autora.**

Trabalho	Domínio	Método	Objetivo	Fonte de Dados	Tratamento em tempo real
Roy e Davenport (2010)	Segurança Marítima	Ontologia	Vigilância para detecção de situações anômalas	Sensores marítimos e observação humana	Não suporta dados em tempo real: chamadas telefônicas de comunicação marítima, serviço de tráfego 24 horas
Baumgartner et al. (2014), (2010)	Centros de Controle de Tráfego Rodoviário		Suporte a grandes centros de controle rodoviário	Sensores de tráfego, sistemas de manutenção, sistemas de acidentes	Não há
Fenza et al. (2010)	Segurança Aeroportuária		Gestão de cenários de risco.	Sensores, radares e controladores humanos	Não há referência
Furno et al. (2010)		Ontologia / Fuzzy	Apoio a decisões em segurança aeroportuária	Sensores, radares e controladores humanos	Há suporte
Valiente et al. (2011)	Militar – C4ISR		Integração de Ontologias para percepção da situação em C4ISR	-	Não há referência
Pilato et al. (2012)	Segurança Marítima	Ontologia / Rede Bayesiana	Vigilância para detecção de ameaças e emissão de alarme	Sensores, radares e controladores humanos	Não há referência

**Tabela 3.1 – Parte 2- Resumo dos trabalhos estudados: domínio e métodos utilizados para formalização da situação. Fonte: elaborado pela autora.**

Trabalho	Domínio	Método	Objetivo	Fonte de Dados	Tratamento em tempo real
Fischer e Beyerer (2012)	Segurança por Imagens	Rede Bayesiana	Monitoramento por Imagens de Câmeras de Segurança	Dados extraídos de imagens de câmeras de segurança	Há suporte
Castaldo et al. (2016)	Vigilância de Local de Transporte Público		Apoio ao operador em segurança de grandes locais públicos	Dados extraídos de imagens de <i>web-cam</i>	Não há referência
Naderpour et al. (2014)	Segurança de Usina Química	Rede Bayesiana / Fuzzy	Modelagem de situações anômalas	Sensores	Há suporte
Aramaki et al. (2011)	Saúde pública	Classificadores de Texto	Detecção precoce de epidemias: gripe	Posts do Twitter	Não há referência
Hanifah et al. (2014)	Informações do tráfego rodoviário		Método de filtragem e classificação sobre congestionamento	Posts do Twitter	Não há
Yin et al. (2012)	Resposta a emergência		Suporte à consciência da situação no atendimento a emergências	Posts do Twitter	Há suporte
Deiva Ragavi e Usharani (2014)			Detecção de terremoto no Japão	Posts do Twitter	Há suporte
Agarwal et al. (2015)	Ambiente Educacional		Simulador de avaliação	Conjunto de perguntas e respostas	Não se aplica
Khan et al. (2016)	Análise e identificação dos sentimentos dos usuários		Pesquisa de opinião sobre dois celulares	Posts do Youtube	Não se aplica
Kotevska et al. (2016)	Cidades Inteligentes		Twitter como sensor social sobre a cidade	Posts do Twitter	Não se aplica

O Modelo Preditivo de Situações compartilha com os demais trabalhos o objetivo de prover suporte à consciência situacional do tomador de decisão que utiliza um sistema computacional para gerir processos decisórios. Para os

trabalhos em resposta à emergência e segurança, tanto os que possuem ou não suporte em tempo real, ambos os grupos utilizam objetos da situação com padrão de comportamento, seja de um navio, de um carro, avião, de epidemias ou mesmo de elementos químicos ou naturais. O trabalho apresentado nesta tese propõe utilizar como objeto principal da situação o elemento humano, cujo comportamento invariavelmente não segue um determinado padrão de comportamento ou propende à imprevisibilidade.

### **3.5 Discussões sobre Métodos de Formalização da Situação**

A riqueza de detalhes que caracteriza uma situação real simultaneamente à sua natural complexidade, propicia a utilização de uma variedade de métodos para a sua formalização. No entanto é necessário atentar às características da situação, do domínio de aplicação e do escopo do apoio que o sistema fornecerá ao operador na tomada de decisão para que diante de tais atributos, os métodos utilizados sejam os mais adequados.

Apesar da característica dinâmica dos ambientes dos trabalhos de formalização da situação apresentados, em sua maioria, os objetos analisados apresentam um padrão de comportamento, ou seja, as ações que estes executam se repetem diante de situações semelhantes. Esses padrões podem ser percebidos e caracterizados nos trabalhos de identificação de situações de risco em indústria química pela análise da temperatura dos recipientes, e manobras de navios e de aviões e dados de tráfego rodoviário em rodovias (Tabela 3.1 - Parte 1 e Tabela 3.1 - Parte 2). Para estes ambientes, os atributos dos objetos são obtidos por meio de sensores físicos e humanos e compreendem valores de tempo, medida, localização ou distância, por exemplo.

Interpretar informações de sensores humanos requer dedução de significado semântico e também participação humana na interpretação. Este

desafio é presente nas interpretações de solicitações de serviços de emergências (VINCEN et al., 2009).

Na mesma direção converge o desafio da formalização das situações cujos objetos não possuem padrão de comportamento, como por exemplo, comportamento humano ou atitudes relacionadas a ele como é o caso de ocorrências policiais.

Métodos semânticos como as ontologias se mostram importantes soluções para a modelagem de situações. No entanto, apesar de suas principais características serem o compartilhamento e reuso ainda se faz necessário o desenvolvimento de ontologias mais abrangentes para que mais situações, ao menos do mesmo domínio, possam ser representadas sem a necessidade da criação de extensões (KOKAR e ENDSLEY, 2012).

Apesar dos avanços expressivos ainda falta similaridade no repertório de conceitos utilizados em ontologias para formalizar situações, como declaram Matheus et al. (2003) e Kokar et al. (2009), especialmente quando se referem às complexas e dinâmicas, pois não oferecem um padrão de comportamento dos objetos que as compõe. Outro aspecto ainda desafiador dessa perspectiva reside no “potencial infinito de possíveis relações entre os objetos das situações” (MATHEUS et al., 2005).

Ainda sobre ontologias, Roy e Davenport (2010) consideram que não se mostram totalmente adequadas para formalizar situações em sistemas de apoio à tomada de decisão em ambiente dinâmico, visto a natureza instável dos eventos ou fatos que nele se apresentam. O trabalho de Furno et al. (2010), extensão do trabalho de Fenza e Furno (2010), desenvolveu o aspecto da dinamicidade e dos ambientes e situações em tempo real associando lógica *fuzzy* à arquitetura proposta.

Métodos probabilísticos, comumente empregados para avaliação automática da situação, similarmente viabilizam o tratamento de problemas de qualidade como incertezas e ambiguidades que apenas o emprego de métodos lógicos não atenderia (FISCHER e BEYERER, 2012). Inicialmente questões de qualidade não serão consideradas nesta tese, mas é um atributo significativo para a determinação do método utilizado neste trabalho.

Para a formalização de situações cujas informações originam fundamentalmente de sensores humanos, foco deste trabalho, métodos probabilísticos de aprendizagem de máquina, como os classificadores, especialmente o *Naïve Bayes*, destacam-se pelo bom desempenho na classificação de partes de textos curtos e ruidosos, pela simplicidade em sua estrutura e facilidade de implementação (MARUCCI-WELLMAN et al., 2015). Por estas características, está presente em grande parte dos trabalhos desta categoria.

Árvores de decisão são estruturas na qual cada nó interno representa uma escolha entre um número de alternativas (ALAGOZ et al., 2010). Embora também sejam consideradas por alguns autores método para modelar situações (ALAGOZ et al., 2010), a escassez de trabalhos alinhados às características dos demais métodos desse capítulo determinou sua exclusão da relação. O método possui limitações em sua capacidade de modelar situações complexas, especialmente quando os eventos ocorrem ou reaparecem ao longo do tempo - características de ambientes dinâmicos

Apesar da diversidade de métodos semântico e probabilístico, a formalização de uma situação real por meio de tratamento computacional ainda encontra desafios no sentido da integração de sistemas de apoio à tomada de decisão, especialmente os de respostas à emergência, à perspectiva da consciência situacional.

Os métodos classificadores são utilizados no tratamento de dados provenientes de redes sociais, fundamentalmente pela natureza das fontes textuais, sendo elas em tempo real ou não. O diferencial fomentador desta tese é a utilização destas mesmas técnicas adaptadas em um contexto de chamadas emergenciais diante do reconhecimento de seu potencial em formalizar situações nos aspectos da consciência situacional propostos neste trabalho.

### **3.6 Considerações Finais**

Tarefa crucial para os sistemas de apoio à tomada de decisão é a inferência sobre comportamentos que se desenvolvem em tempo real, visto a necessidade dos operadores humanos ao acesso imediato às informações sobre situações que exigem rápida intervenção (CASTALDO et al., 2016). Esse processo se inicia com a formalização da situação por meio de métodos computacionais relacionados neste capítulo. Nessa perspectiva, este capítulo apresentou os métodos utilizados para formalização da situação por meio de tratamento computacional semântico utilizando ontologias e probabilístico empregando métodos de aprendizagem de máquina, redes bayesianas e classificadores.

Analisando os trabalhos apresentados na literatura com variações de métodos empregados, observa-se que essas ocorrências se dão, sobretudo pelos requisitos do domínio da aplicação e pela natureza das informações que servirão na composição da situação.

Por serem métodos semânticos, ontologias estabelecem relação bem próxima à estrutura teórica das situações e são aplicadas em um número considerável de trabalhos. No entanto, há ainda desafios em aplicá-las a ambientes complexos e dinâmicos (KOKAR e ENDSLEY, 2012), (ROY e DAVENPORT, 2010), (KOKAR e NG 2009).

Métodos probabilísticos se apresentam em extenso número de trabalhos para a formalização da situação, da mesma forma a relação de pesquisas corrobora com Sakai et al. (2015) quando observam que o aumento da consciência de situação no mundo real usando dados de sensores sociais é um dos temas de pesquisa mais atraentes. É indubitável a relevância deste tipo de informação na contemporaneidade.

Nas reflexões de Blasch et al. (2013) em trabalho sobre os desafios de sistemas de Fusão de Informações em Alto Nível, os pesquisadores observam que encontram nos ambientes complexos e dinâmicos novas modalidades de entrada como texto e linguagem natural e que estas apresentam desafios

particulares. Os autores avaliam que se deve compreender quais aspectos deste problema podem ser resolvidos com métodos automatizados de processamento. Os autores se referem aos sistemas de Fusão de Informações em Alto Nível, no entanto, as considerações se alinham perfeitamente aos sistemas de apoio à tomada de decisão atuando em ambientes como os citados, ou seja, complexos e dinâmicos. A diferença se apresenta quando os autores se referem às novas modalidades de entradas. Para esta tese, texto e linguagem natural não se tratam de uma nova modalidade, mas sim um tipo específico de informação que será a base fundamental do sistema.

Ao analisar os métodos empregados, bem como todo o escopo que envolve a formalização da situação, constatou-se a oportunidade de explorar esses métodos tendo como fonte primária de informações não os denominados sensores sociais (KOTEVSKA et al., 2016), mas sim os humanos que recorrem aos atendimentos emergenciais.

Esses métodos classificadores serão utilizados para formalizar e antecipar situações em ambiente dinâmico, e em tempo real, em sistemas que possuem como fundamental fonte de informação o humano, denominado neste trabalho como “sensor situacional”.

Diante do exposto, o capítulo quatro apresenta a proposta de um Modelo para Automação e Previsão da Situação como Apoio à Consciência Situacional em Sistemas de Respostas a Emergências.

# Capítulo 4

## MODELO PARA AUTOMAÇÃO E PREVISÃO DA SITUAÇÃO COMO APOIO À CONSCIÊNCIA SITUACIONAL E AO PROCESSO DECISÓRIO EM SISTEMAS DE RESPOSTA À EMERGÊNCIA

---

*Este capítulo apresenta o Modelo para Automação e Previsão da Situação proposto como apoio ao processo decisório. O modelo objetiva sistematizar a implementação de sistemas ou módulos que cooperem com software ou protocolos de instituições com atendimentos emergenciais, como por exemplo, corpo de bombeiros e serviço 190. A Seção 4.1 introduz o modelo com as considerações iniciais do capítulo; a Seção 4.2 descreve o Modelo Conceitual desenvolvido; a Seção 4.3 versa sobre o Modelo Preditivo, suas etapas e especificidades e a Seção 4.4 encerra o capítulo com as considerações finais.*

## **4.1 Considerações Iniciais**

Consciência situacional é o modelo mental do tomador de decisão sobre o estado do ambiente. Com base nessa representação, o operador pode decidir o que fazer e realizar as ações necessárias sobre a situação em curso. Mesmo possuindo as etapas de decisão e ação separadas do núcleo SAW, o modelo de Endsley (1995), descrito no capítulo dois, seção 2.2.1 (Figura 2.5), integra os processos e evidencia a consciência situacional como principal fator precursor para a tomada de decisão. A melhoria do processo decisório humano pode ser atribuída como fator altamente relacionado com a melhoria da consciência situacional (ROY et al., 2007).

O processo decisório humano fundamenta-se na aplicação de diferentes modelos de tomada de decisão como, por exemplo, naturalista, racional ou processual - cada um deles apropriado a um determinado tipo de situação.

O modelo de tomada de decisão naturalista indica que indivíduos optam por uma alternativa que lhes parece satisfatória por meio de um processo de reconhecimento de padrões ao contrário de comparar diversas alternativas possíveis para encontrar a melhor alternativa de decisão e ação. Este modelo é apropriado para representar o processo de tomada de decisão em situações caracterizadas por fatores como estresse, pressão de tempo, objetivos conflitantes e problemas mal definidos (WONG e BLANDFORD 2001).

Este conceito alinha-se ao trabalho de Gibson (1979), sobre os princípios da percepção direta, sendo dois deles destacados por Roy et al. (2007): (i) toda informação necessária para a percepção está contida no ambiente e (ii) a percepção é imediata e espontânea.

Apesar de princípios inquestionáveis, a teia de relações que uma situação é composta pode oferecer dificuldades, tornando a percepção de elementos relevantes à sua identificação algo nem sempre trivial, dependendo do tipo de ambiente em que se desenvolve, como por exemplo, ambientes complexos e dinâmicos.

Além da peculiaridade do modelo das decisões naturalistas e da complexidade para, muitas vezes, identificar informações relevantes na caracterização de situações, Endsley e Garland (2000) qualificam a possibilidade de classificação da situação como um arranjo sistemático em grupos ou categorias de acordo com critérios estabelecidos, cujos modelos podem estar disponíveis a priori.

Associados a estes modelos podem estar esquemas de situações prototípicas. As pistas críticas no ambiente podem ser combinadas com tais esquemas a priori para indicar situações prototípicas que fornecem classificação e compreensão instantânea da situação. (ENDSLEY e GARLAND, 2000)

Estes três parâmetros (modelo decisório naturalista, ambientes complexos e dinâmicos e situações classificadas a priori) caracterizam o tipo de decisão, ambiente e situações que juntos constituem o objeto de estudo desta tese e conduziram a construção do Modelo Conceitual e Preditivo de Situações. Especificidades de ambos os modelos serão descritos e discutidos nas seções seguintes deste capítulo.

## **4.2 Modelo Conceitual**

O modelo conceitual proposto neste trabalho constitui-se de cinco etapas, sendo quatro conjuntos de atos e uma resultante do processo (Figura 4.1). Os quatro conjuntos de atos ilustrados pelo modelo objetivam orientar o desenvolvimento de ferramentas que contemplem aspectos da consciência da situação em sistemas de apoio à decisão que operem em situações críticas, como por exemplo, resposta a emergência.

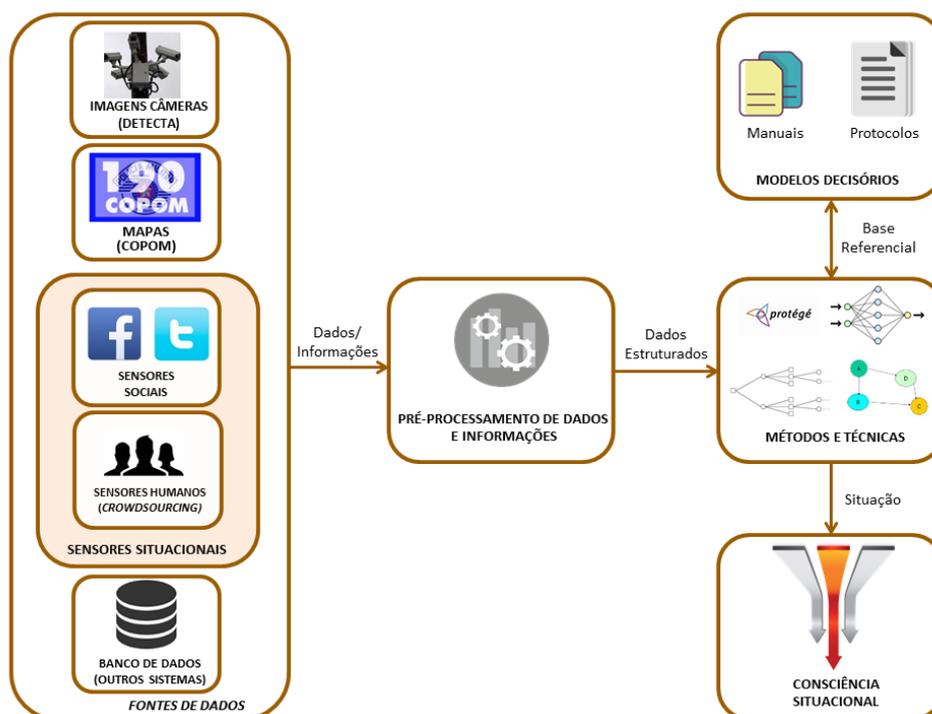


Figura 4.1 – Modelo conceitual para o desenvolvimento de ferramentas orientadas à consciência da situação. Fonte: elaborado pela autora.

#### 4.2.1 Fontes de dados

Diferentes fontes de dados podem ser consideradas no modelo conceitual, tais como câmeras de monitoramento distribuídas na cidade, mapas com dados de localização de viaturas em tempo real, incidência criminal das cidades (denominada mapas de calor), bancos de dados de veículos e pessoas, dados de outras chamadas que podem tratar da ocorrência de um mesmo evento, além de dados gerados pela população, que podem ser obtidos de redes sociais, *microbloggings* e de aplicativos de *crowdsourcing*<sup>3</sup> (denominados neste trabalho de sensores situacionais).

Sensores situacionais foi o termo estabelecido para tipificar fontes de dados e informações sobre situações nos ambientes complexos e dinâmicos.

<sup>3</sup> *crowdsourcing* - tipo de aplicação em que informações, inclusive geográficas, sobre um evento são geradas e disseminadas por indivíduos voluntários da população, por meio de aplicativos para diferentes propósitos: de condições de tráfego em grandes cidades, padrões de fluxo de pedestres, a compartilhamento de carro.

Essas fontes compreendem dois grupos principais: os Sensores Sociais (KOTEVSKA et al., 2016) e os Sensores Humanos. Sensores Sociais são dados originados de redes sociais e *microblogging* (textos das postagens dos usuários). Sensores Humanos são fontes de dados originados de pessoas, independente de seu formato e meio de comunicação (por exemplo, *crowdsourcing*).

#### **4.2.2 Pré-Processamento de Dados e Informações**

As fontes de informações representadas no Modelo Conceitual incluem: (i) dados do ambiente físico em narrativa em formato de texto e ou de áudio durante a situação em curso; (ii) dados do ambiente físico em narrativa em formato de texto das postagens nas redes sociais durante e após a ocorrência da situação.

Considerando a riqueza de detalhes concomitantemente aos inúmeros problemas diante da natureza estrutural da narrativa de uma postagem ou de um relato falado, esta etapa de pré-processamento se faz necessário. Nesta fase serão estruturados os dados e informações para que se adequem às bases de dados que a ferramenta utilizará ou até mesmo à linguagem de programação adotada. Assim sendo, a fase de pré-processamento deve ser prevista com o objetivo de adequar e estruturar os dados para que possam ser utilizados como entrada da etapa Métodos e Técnicas. Da mesma forma, os dados das diferentes fontes representadas no modelo conceitual são também submetidos à fase de pré-processamento podendo requerer a aplicação de diferentes técnicas.

#### **4.2.3 Métodos e Técnicas**

A terceira etapa do Modelo Conceitual consiste do conjunto de métodos e técnicas utilizados para a formalização e ou identificação da situação. Como

tratado no capítulo três desta tese, diferentes abordagens podem ser empregadas para essa tarefa.

O complexo número de relações entre os objetos que compõem uma situação real associada à sua natural complexidade amplia as possibilidades de exploração de métodos para a sua formalização e análise. Não obstante, é necessário estar ciente das reais características da situação, do domínio de aplicação e do limite de escopo do apoio que o sistema em desenvolvimento fornecerá ao operador na tomada de decisão. Desta forma, diante de tais atributos, os métodos utilizados corresponderão adequadamente aos quesitos que contemplem aspectos de SAW.

Independente do método utilizado, o modelo conceitual proposto contempla situações prototípicas (ENDSLEY e GARLAND, 2000) que são representados pelos Modelos Decisórios utilizados pela instituição do agente humano decisor. Este conceito será mais detalhado na seção 4.3 que trata do Modelo Preditivo de Situações.

#### **4.2.4 Modelos Decisórios**

A proposta deste módulo prevê instituições que já adotem determinadas normas, condutas decisórias e ou determinados protocolos de atendimento que ao final de um conjunto de procedimentos culminem em uma tomada de decisão.

Conforme tratado na subseção 4.2.3, o Modelo Conceitual proposto atribui aos Modelos Decisórios o elemento “situações prototípicas”. Esta característica atribui ao modelo mental do tomador de decisão o estado atual da situação, ou o modelo de situação. Nesta abordagem, sua função é servir como no modelo mental do operador humano, ou seja, um esquema arquivado para consulta e imediata identificação da situação.

Quando da automação da situação em sistemas de apoio à tomada de decisão, no módulo Métodos e Técnicas, os Modelos Decisórios terão o papel dos esquemas e roteiros do modelo mental (descrito na seção 2.2.2, Figura

2.8) provendo a compreensão e projeção da situação e conseqüentemente a Consciência Situacional.

Os esquemas agem como atalhos, fornecendo a compreensão e projeção da situação em um simples passo, especialmente quando as situações envolvem grandes volumes de dados. Baseado neste conceito, os Modelos Decisórios atuarão como esquemas.

#### **4.2.5 Consciência Situacional**

Resultado final do processo, esta última etapa do Modelo Conceitual apresenta a proposta objetivada pela etapa Métodos e Técnicas que pode apresentar a projeção da situação futura ou análise da situação em curso, dependendo do método empregado e dos objetivos da ferramenta em desenvolvimento.

### **4.3 Modelo Preditivo de Situações**

A concepção do Modelo Preditivo de Situações foi inspirada nos três parâmetros citados na seção 4.1: (i) modelo decisório naturalista, (ii) ambientes complexos e dinâmicos e (iii) situações classificadas a priori. A descrição inicial do problema pode ser representada de forma simples pela Figura 4.2.

A mensagem representa os sinais e as informações do ambiente que serão recebidas pelo agente humano. A compreensão dessas informações será integrada ao seu modelo mental sobre a situação em curso (que poderá existir ou não, dependendo de sua experiência ou treinamento, por exemplo) e às situações classificadas a priori (ou modelos decisórios, dependendo do domínio) subsidiando a consciência situacional do agente humano – tomador de decisão.

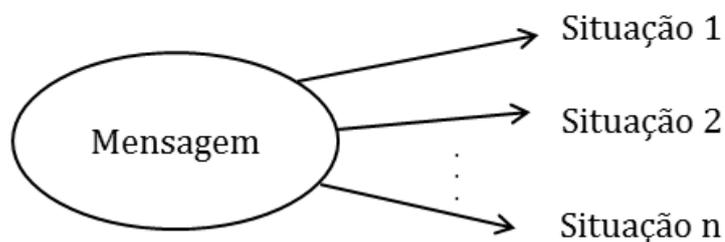


Figura 4.2 – Descrição do problema. Fonte: elaborado pela autora.

A instabilidade que a situação pode apresentar é uma característica predominante em ambientes dinâmicos. Conseqüentemente, uma situação pode sofrer mudanças durante seu curso, no decorrer do tempo. Como a proposta do Modelo Preditivo de Situações contempla esses ambientes, a saída será a classificação ou ranqueamento das situações (Figura 4.3) e não apenas uma única possibilidade.

O modelo proposto objetiva sistematizar a implementação de ferramentas ou módulos que, integrados a sistemas já existentes, contribuam para adquirir e ou prover a manutenção da consciência situacional do operador em sistemas de apoio à tomada de decisão em atendimentos emergenciais, como por exemplo, o serviço 190 (integrado ao do corpo de bombeiros e Samu em 2017 em São Paulo).

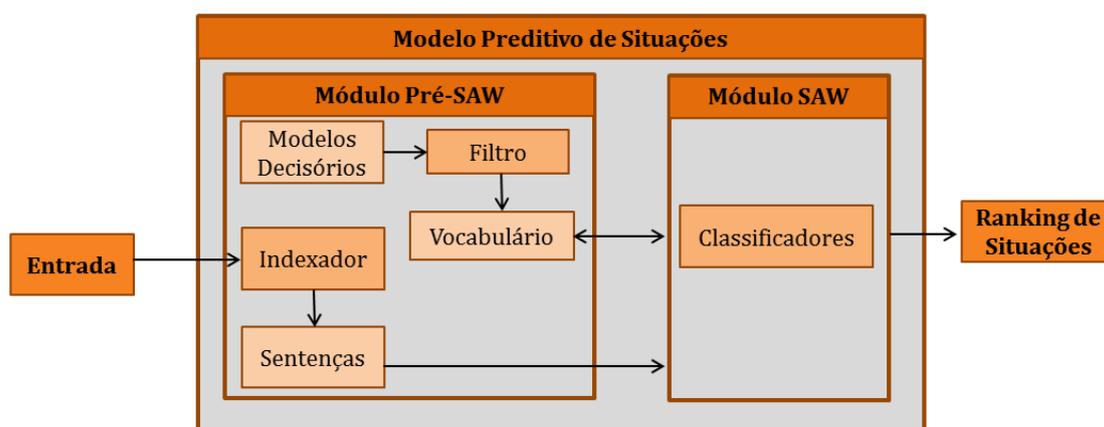


Figura 4.3 – Modelo Preditivo de Situações, considerando os sensores situacionais. Fonte: elaborado pela autora.

O Modelo Preditivo de Situações, ilustrado pela Figura 4.3, constitui-se de dois blocos, sendo eles o Módulo Pré-SAW que prepara os dados de entrada recebidos e os organiza e o Módulo SAW que os classifica, identificando as situações e as ranqueando pela ordem de relevância. O Modelo Preditivo considera os dados de sensores situacionais – solicitantes do atendimento 190 da PMESP.

O Módulo Pré-SAW é composto por cinco submódulos, sendo eles: Filtro, Indexador, Modelos Decisórios, Sentenças e Vocabulário. O Módulo SAW compõe-se do submódulo Classificadores. A descrição de cada componente será detalhada nas próximas subseções.

### **4.3.1 Entrada**

A Entrada representa os dados do ambiente que neste modelo são predominantemente no formato de áudio. Os dados são originados de sensores situacionais humanos. Esta denominação justifica-se pela singularidade da inspeção dos sinais do ambiente no único propósito da caracterização da situação.

O humano no papel de sensor percebe os sinais do ambiente por meio de seus sentidos, visão, audição, tato, paladar e olfato ou a combinação deles. Sua experiência e conhecimento prévio de situações semelhantes anteriores poderá determinar a seleção dos elementos relevantes que comporão os objetos da situação.

É fato que em determinados ambientes, alguns desses sentidos podem ser prejudicados por fatores naturais como ruído, fumaça ou mesmo a dinamicidade da situação que está sempre em mudança. No entanto, em contextos de respostas a emergências, dados de sensores humanos compõe a principal fonte de informação dos sistemas de apoio a estes serviços.

Ferramentas que analisam dados em linguagem natural tem se mostrado eficazes nas mais diversas áreas de atuação, como descrito na

seção 3.4. Para tal, no modelo proposto, é necessário que se faça a conversão dos dados da forma de áudio para a textual na execução do Módulo Indexador descrito a seguir.

### 4.3.2 Submódulo Indexador

Como neste modelo os dados são predominantemente no formato de áudio, o Módulo Indexador é responsável pela execução da transcrição dos dados de entrada em forma de áudio para a forma textual e posteriormente sua estruturação em sentenças.

O Módulo Indexador utilizará parte do módulo *HUMINT Data Acquisition*<sup>4</sup> (BOTEGA et al., 2017) que utiliza técnicas de Processamento de Linguagem Natural para a transcrição do áudio em estrutura de *string* semelhante à fornecida pelo *Google* (CHELBA et al., 2012). O módulo é responsável pela limpeza e pré-classificação, identificando objetos e atributos, disponibilizando os dados para o uso dos outros componentes.

Após receber as sentenças, o estágio de análise gramatical as analisa em tempo real, adicionando rótulos como substantivos, números, objetos e outras classificações. O módulo *HUMINT Data Acquisition* segue com outras funcionalidades, mas que não são essenciais à proposta deste trabalho. A transcrição para texto, limpeza, organização, classificação e extração de elementos chave em formato de sentenças é o que o Modelo Preditivo requer.

### 4.3.3 Sentenças

Sentenças, como descritas na subseção 4.3.2 é o resultado da execução do módulo Indexador. A estrutura de *strings* será a composição das sentenças e a nova entrada do Módulo SAW.

---

<sup>4</sup> Os pesquisadores referenciados e a autora desta tese trabalham em parceria em pesquisas colaborativas.

As sentenças serão compostas por conjuntos de palavras da mensagem original e a partir dessas sentenças todo o processamento será realizado.

#### **4.3.4 Modelos Decisórios**

Modelos Decisórios representam o acervo, ou seja, conjunto de documentos e ou protocolos com os quais os operadores dos sistemas de apoio à tomada de decisão em atendimentos emergenciais balizam suas decisões e ações, utilizando-os para identificar as situações em curso.

Configuram as *situações prototípicas* denominadas por Endsley e Garland (2000). Atuarão como os esquemas do modelo mental do agente humano, direcionando para a compreensão e projeção da situação de maneira rápida, analogicamente aos esquemas e roteiros do modelo mental.

Compreensão e projeção da situação ocorrendo de forma rápida podem gerar, conseqüentemente, tomadas de decisão igualmente rápidas. Esse aspecto representa uma particularidade importante devido às características críticas dos ambientes que ocorrem as situações emergenciais. O tempo de resposta é um fator, muitas vezes, determinante para direcionar o encerramento satisfatório ou não de um atendimento de situação crítica, envolvendo, por exemplo, uma vida.

A padronização de correspondência entre sinais críticos do ambiente e elementos no modelo mental é uma capacidade individual e esta relação influencia na obtenção de SAW. Por esta perspectiva, o modelo que o indivíduo tem da situação é o estado atual do modelo mental. Esse modelo situacional representa os vários parâmetros da situação, além de como eles se relacionam em termos de forma, função, a inter-relação de seus componentes, criando uma compreensão geral do “estado” da situação (ENDSLEY e GARLAND 2000). Isso é que proporciona, além da compreensão, a projeção dos estados futuros da situação. A relação entre modelos mentais e consciência da situação foi descrita da subseção 2.2.2.

O Modelo Preditivo de Situações foi inspirado no modelo *Função dos Objetivos e Modelo Mental em SAW* de Endsley e Garland (2000) e é esperado

que exista acervo de modelos decisórios, sejam eles documentos de quaisquer tipo. No entanto, caso não existam, isso não se tornará um impeditivo para que o modelo seja utilizado (Figura 4.4).

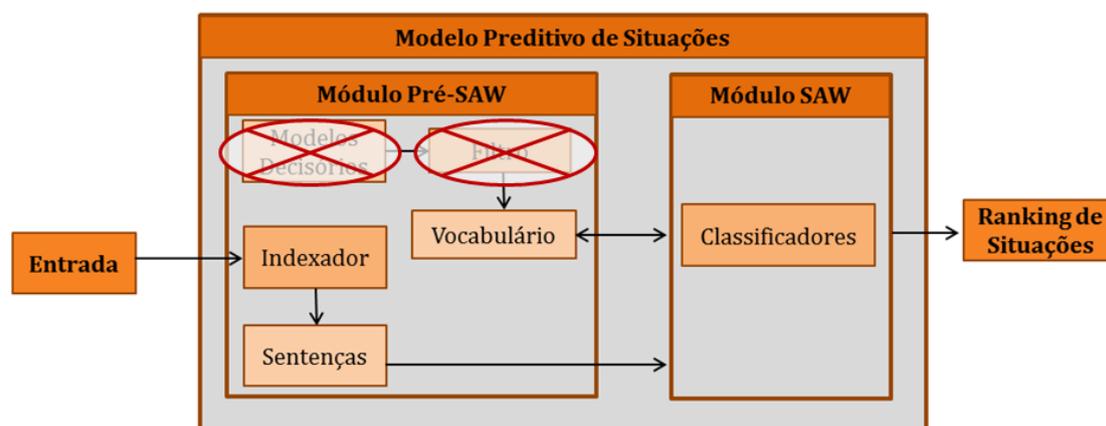


Figura 4.4 – Modelo Preditivo de Situações prevendo a não existência de Modelos Decisórios. Fonte: elaborado pela autora.

De outra forma um Vocabulário poderá ser criado para que o Módulo Classificador se adapte às novas condições.

Modelos Decisórios podem ser compostos por diversos tipos de documentos e por esta razão é necessário que sejam estruturados para que os métodos classificadores os utilizem de forma adequada. Essa estruturação é executada pelo Módulo Filtro.

#### 4.3.5 Módulo Filtro

O Módulo Filtro estrutura os Modelos Decisórios padronizando-os em um mesmo formato de texto e seleciona o que de fato é relevante para a classificação das situações. Endsley (2000) denomina de “lacuna de informação” este espaço entre todo o conteúdo produzido e o que de fato são informações necessárias.

A função do módulo Filtro é executar a limpeza do conteúdo, com a eliminação de caracteres especiais, como por exemplo, acentos e espaços em

branco e padronizar a forma da letra do texto para que não haja problemas na execução dos algoritmos.

Esse módulo também é responsável por eliminar palavras repetidas com frequência acima de 100 ocorrências devido a não relevância em relação à identificação da situação. Palavras e termos com valores irrelevantes para o contexto, como por exemplo, artigos definidos e indefinidos, certos verbetes, dentre outros, são eliminadas. Isto feito, as palavras são ordenadas alfabeticamente e se ainda verificada a existência de duplicatas, estas são excluídas.

O módulo Filtro prepara o Vocabulário para que se torne a base de referência para o módulo Classificadores.

#### **4.3.6 Vocabulário**

O Vocabulário é o resultado da execução do módulo Filtro, reunindo todos os Modelos Decisórios em uma base estruturada. Após a limpeza e estruturação dos dados o resultado foi uma relação de 467 palavras. A imagem do resultado completo do vocabulário pode ser verificada na Figura A.7 do Apêndice A.

#### **4.3.7 Submódulo Classificadores**

O submódulo Classificadores executa métodos de classificação de textos e apresenta a identificação da situação ou das situações em curso. O Módulo prevê a execução de mais de um método de classificação. Inicialmente foram implementados os métodos *Naïve Bayes* e *Bag of Words*. Isto se dá em razão dos sistemas se caracterizarem como apoio à tomada de decisão, ou seja, a automação e mesmo a predição da situação objetiva diminuir a carga cognitiva do operador, auxiliar na obtenção e manutenção da SAW e indicar possíveis caminhos de solução. No entanto, a decisão é tomada pelo especialista humano e os sistemas somente o apoiam. Desta forma, é razoável que o modelo ofereça ao tomador de decisão, quando implementado,

alternativas que contemplem a dinamicidade tanto do ambiente quanto das situações em que ele e o solicitante estão envolvidos.

Como os Modelos Decisórios farão o papel dos esquemas e roteiros do Modelo Mental do operador, os operadores humanos menos experientes poderão usufruir deste recurso – esses modelos colaborarão para formar as *situações prototípicas* quando ausentes no modelo mental de operadores inexperientes.

O resultado esperado deve conter a probabilidade da situação ou das situações em um panorama dinâmico e a experiência do agente humano identificará a situação indicada àquele cenário com maior grau de assertividade.

#### **4.3.8 Ranking de Situações**

O resultado do Método Classificadores como descrito na seção 4.3.7 será uma listagem das situações em ordem decrescente dos valores de suas probabilidades de pertencerem à classe dos modelos decisórios. Como elucidado, isso se torna necessário devido à dinamicidade do ambiente, da própria situação e da experiência do operador humano decisor.

### **4.4 Considerações Finais**

Este capítulo apresentou o Modelo Conceitual e o Modelo Preditivo de Situações desenvolvidos com o objetivo de sistematizar a implementação de sistemas ou módulos que cooperem com sistemas de apoio à tomada de decisão que contemplem aspectos de SAW.

Os trabalhos propostos foram inspirados no modelo das funções dos objetivos e modelo mental em SAW (ENDSLEY e GARLAND, 2000) em que experiências orientam os operadores humanos como roteiros e esquemas em seu modelo mental. Instituições que trabalham com atendimentos emergenciais

possuem protocolos e manuais de atendimento e todo o procedimento passa por estes “roteiros”.

A experiência do operador, personalidade e características pessoais de forma geral pode influenciar em todo o processo decisório, visto que SAW é uma condição individual.

Fornecer ferramentas que possam auxiliar este processo pode contribuir para que o resultado do processo de tomada de decisão como um todo gere ações mais assertivas.

O capítulo cinco apresenta a aplicação do modelo desenvolvido e o método para preparação dos dados.

# Capítulo 5

## MODELO PREDITIVO DE SITUAÇÕES: APLICAÇÃO E MÉTODO

---

*Este capítulo apresenta a aplicação e o método do Modelo Preditivo de Situações. Inicialmente na seção 5.1 será contextualizado o problema do atendimento às chamadas 190, como a PMESP o realiza e o método de árvore de decisão empregado no processo. Na seção 5.2 a metodologia de trabalho para a aplicação do modelo proposto será descrita, como os dados foram filtrados e o vocabulário criado, considerando dados de áudio, fonte primária para o atendimento 190. A seção 5.3 especifica os algoritmos dos métodos classificadores que foram implementados e as considerações finais são apresentadas na seção 5.4.*

### **5.1 Contextualização do Problema e Características dos Dados**

O serviço às chamadas 190 da PMESP é realizado da seguinte forma: o atendente recebe as chamadas telefônicas e procede ao atendimento por meio de um sistema de informação que utiliza o método de árvores de decisão. Exemplos de árvores de decisão utilizadas como modelo de atendimentos são demonstrados na Figura A.1, Figura A.2 e Figura A.3 do Anexo A.

Após preenchimento das informações sobre o endereço do solicitante, todo atendimento inicia com uma pergunta do tipo “O que está acontecendo?”,

“Está acontecendo agora?” ou “Onde está acontecendo?”, dependendo do tipo de ocorrência – que no sistema é referido como Assunto.

Diante das informações iniciais do solicitante, o atendente deve selecionar a árvore de decisão correspondente ao tipo de ocorrência, que pode variar em mais de 206 tipos, agrupados em 60 árvores<sup>5</sup>, e continuar o seu preenchimento e atendimento. A Figura 5.1 ilustra uma das telas de atendimento do sistema da PMESP exemplificando situação de roubo e a Figura 5.2 representa parte da árvore de decisão para roubo. As elipses vermelhas na Figura 5.2 são grifos da autora deste trabalho para indicarem a ordem de atendimento seguida e representada pela tela do sistema na Figura 5.2.

The screenshot displays the 'WAtende - Atende' interface for the São Paulo Military Police. The top section shows the system title 'Sistema de Informações Operacionais da Polícia Militar' and the date/time '16/07/13 14:50'. Below this, there are fields for 'CAD' (SAO PAULO), 'Serviço' (RÁDIO PATRULHA), 'Atendente' (CB PM RABELLO), and 'PA' (28). The main form contains fields for 'Telefone' (29413093), 'Categoria' (NÃO CLASSIFICADO), 'Solicitante' (NILZA LEMOS CATRANZARO), and 'Anônimo' (checkbox). The 'Município' is set to 'SAO PAULO'. The 'Logradouro' is 'RUA IVAI', 'Número' is '318', and 'Compl.' is 'AP 42 B'. The 'Bairro' is 'TATUAPE'. The 'Referência' field is empty. The 'Delegacia' is '52º D.P. PARQUE S. JORGE'. The 'Assunto' is 'Roubo'. The 'Histórico inicial' section contains a series of questions: 'Está acontecendo agora?' (Sim), 'Qual é a Vítima / Estabelecimento?' (Comercio), and 'Há Reféns?' (Não tem certeza). The 'Prioridade' is 'NORMAL' and the 'Data/hora do fato' is '16/07/2013 14:50'. The 'Classificação de chamadas' section lists six options: 1 - Assunto Adm/Operacional, 2 - Chamada não concluída, 3 - Informação, 4 - Trote Adulto, 5 - Trote Criança, and 6 - Transferência. The 'Classificação de ocorrências' section has two checkboxes: 'Alerta Geral?' and 'Alerta Parcial?'. Below these are buttons for 'Policial em perigo', 'Gerar Ocorrência', and 'SOP'. At the bottom, there are buttons for 'Transferir Ligação', 'Atalho...', 'Mensagens' (0), and 'Ficar Indisponível'.

Figura 5.1 – Sistema e Informações Operacionais de Polícia Militar do estado de São Paulo para atendimento às chamadas 190. Fonte: PMESP (ALMEIDA, 2013).

<sup>5</sup> Este trabalho teve acesso a 54 árvores de decisão da PMESP, totalizando 206 tipos de ocorrências.

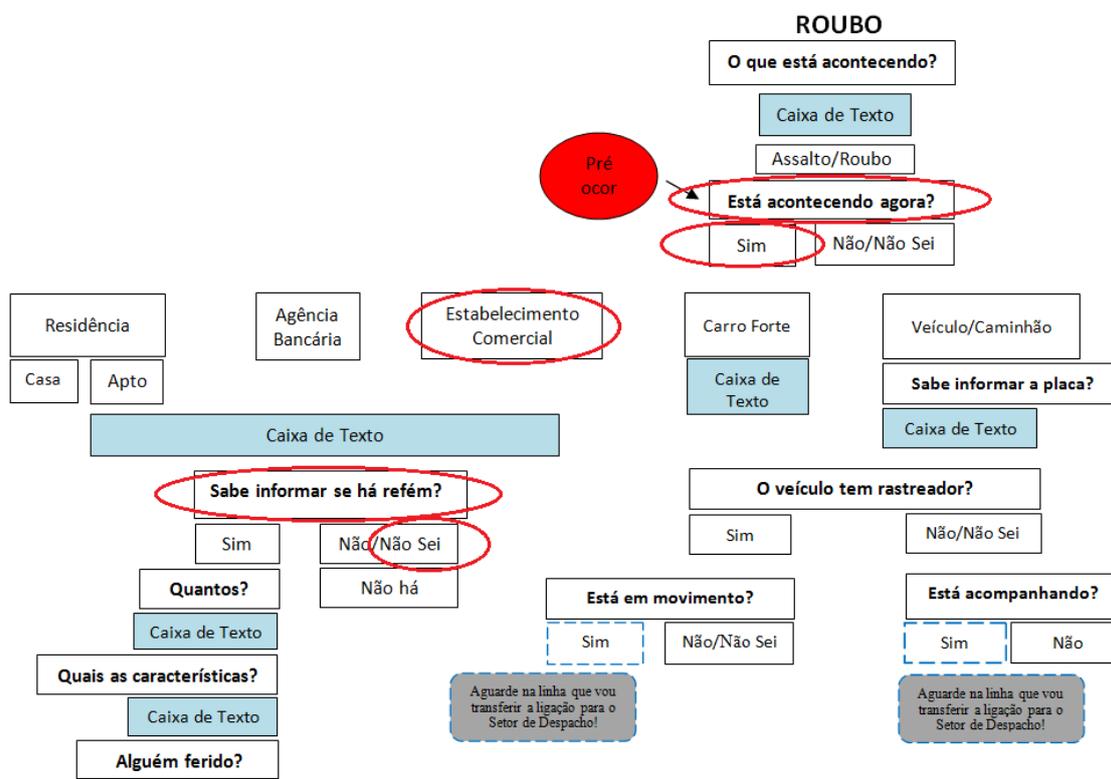


Figura 5.2 – Parte da árvore de decisão para situação de Roubo. Fonte: PMESP.

De acordo com o COPOM/Campinas-SP é comum durante o preenchimento de determinada estrutura de árvore, o solicitante ou o próprio atendente verificar que a situação relatada não corresponde especificamente àquela árvore de decisão selecionada. Por exemplo, pode-se citar o caso de estar sendo preenchida a árvore de decisão correspondente a um Roubo, quando o atendente pelo relato percebe de fato que a ocorrência se refere a um Furto. Neste caso o atendente necessita abandonar o preenchimento daquela árvore de decisão e iniciar a correta desde o princípio. Esta operação demanda tempo, fator crucial neste tipo de atendimento.

Após o preenchimento destas informações, este boletim é transformado em pré-ocorrência e caso exista algum erro ou dado faltante, o policial designado como inlucor os altera ou corrige. Somente então será encaminhado ao policial despachador que transmitirá à viatura policial ou a outro órgão responsável ao atendimento como Serviço de Atendimento Móvel de Urgência (SAMU), Corpo de Bombeiros, Defesa Civil, dentre outros.

A PMESP utiliza atualmente métodos de árvore de decisão para representar os procedimentos padrão de todos os tipos de eventos em chamadas 190, apresentando o que deve ser perguntado ao solicitante durante a ligação.

O método de árvore de decisão propõe um roteiro ou um fluxo de perguntas indicando as informações que devem ser obtidas em uma ordem determinada, assim como também orientando as decisões que o atendente deve proceder.

A compreensão da situação pode ser construída de acordo com este fluxo, porém, pela natureza crítica e dinâmica da situação, não há garantia de que todas as informações sejam obtidas de acordo com a ordem estrutural deste método. Isto se dá pelo fato de que os relatos das ocorrências podem ocorrer sob forte estresse tanto dos solicitantes que podem estar envolvidos na situação como vítima ou testemunha, quanto dos atendentes que trabalham sob fatores de tensão como, por exemplo, turnos longos e exaustivos, necessidade de extrema atenção aos detalhes, rapidez de raciocínio e compreensão, estado de alerta constante, forte carga cognitiva. Outro fator limitante do método de árvore de decisão é que se uma informação crucial estiver ausente, o atendente será impedido de prosseguir com o questionamento e preenchimento. Desta forma, atualmente, a identificação da situação depende exclusivamente da compreensão do atendente em relação ao relato do solicitante.

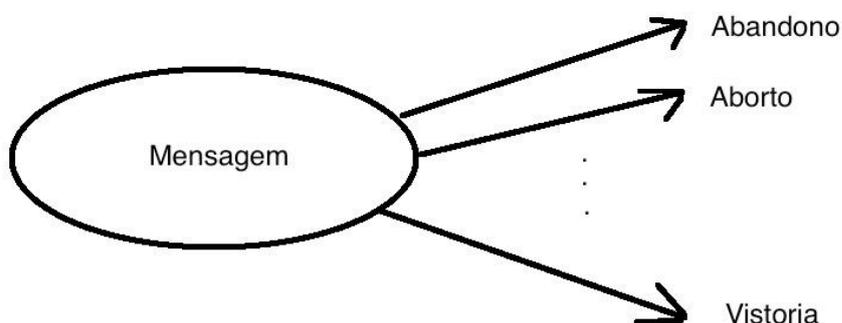
Exposto o contexto, este trabalho propõe, além da automação de parte do processo de identificação da situação, a antecipação da identificação dessa situação em curso antes que problemas de qualidade de informações se propaguem nos níveis de consciência situacional e, conseqüentemente, nos processos de tomada de decisão.

Apesar de dividir opiniões, há mais de uma década no Brasil a terceirização dos atendimentos emergenciais 190 em alguns Estados com a colocação de civis neste serviço já é realidade (PACÍFICO e CARVALHO, 2013) (DUARTE, 2010). Com a proposta deste trabalho, a integração de modelos de predição da situação, particularmente em atividades de operadores

civis, poderá contribuir no apoio à consciência situacional dos operadores destes sistemas, atenuando alguns problemas que podem ocorrer neste tipo de atividade, como por exemplo, agilidade e qualidade no atendimento.

## 5.2 Método de Trabalho

Para o conjunto de árvores de decisão (nesta seção serão denominadas fluxogramas) utilizadas pela PMESP nos atendimentos de chamadas 190, a tarefa será classificar uma mensagem hipotética, ou seja, encontrar em um dos 54 fluxogramas o que possuir a maior probabilidade de corresponder à chamada em curso. A Figura 5.3 representa a descrição do problema.



**Figura 5.3 - Descrição do problema. Fonte: elaborada pela autora.**

Neste processo, cada fluxograma foi transformado em um conjunto de palavras relevantes e para isso, uma filtragem dessas palavras foi realizada e um vocabulário de dados foi criado. Esse método será descrito nas próximas seções.

### 5.2.1 Filtragem dos Dados e Criação do Vocabulário

A primeira etapa desta metodologia consistiu em calcular a frequência de palavras para cada fluxograma. Todos os dados de todos os fluxogramas foram filtrados para que todas as palavras repetidas fossem eliminadas. Uma das ferramentas utilizadas nesta etapa foram *scripts* em *bash* no sistema operacional Linux.

Todos os fluxogramas encontravam-se no formato `.doc` (exemplos das árvores de decisão são apresentadas no Anexo A), sendo todas as suas palavras extraídas, convertidas em arquivo `.txt` e armazenadas em um mesmo diretório. Utilizou-se o comando `cat` cuja função é listar o conteúdo de arquivos (Figura A.1 do Apêndice A).

Na sequência, foi necessário ordenar e apresentar apenas as palavras únicas e para tal os comandos `sort` e `uniq` foram executados.

O resultado da execução dos três comandos foi armazenado em um arquivo denominado `dicionário.txt` utilizando o redirecionador `>`. Este redirecionador é utilizado para criar um novo arquivo recebendo o fluxo de saída de outros comandos. A Figura A.2 do Apêndice A ilustra este procedimento.

Frequentemente algumas palavras estavam em letras maiúsculas e isso poderia dificultar a execução do algoritmo. Desta forma, utilizou-se o comando conversor `tr '[:upper:]' '[:lower:]'` para que as letras fossem padronizadas. A Figura A.3 do Apêndice A ilustra o comando utilizado nos fluxogramas para esta etapa.

A ocorrência de palavras com acentos e ou caracteres especiais fez necessário nova filtragem para eliminá-los e o comando `sed` foi executado no conjunto de palavras. As Figuras A.4 e A.5 do Apêndice A representam o *script* completo para fazer todas as conversões descritas anteriormente e o comando para a remoção de espaços em branco, respectivamente. Por fim foi realizada a conversão ISO8591-1 para UTF-8 para padronizar o sistema de caracteres utilizado e a Figura A.6 do Apêndice A apresenta o *script* para este procedimento.

Após a filtragem das palavras que ocorrem nos fluxogramas, foi necessário retirar as menos relevantes. Isso se deu pelo fato da listagem conter inúmeras palavras e termos de valor irrisório ao processo, como por exemplo, artigos definidos e indefinidos, certos verbetes, dentre outros. Utilizando as ferramentas *R* e *RStudio*<sup>6</sup>, foi desenvolvido um código para identificar as palavras mais frequentes presentes nos fluxogramas. Quanto maior a frequência da palavra em cada fluxograma, menor o auxílio para a identificação da situação nos algoritmos de classificação de texto.

A Figura 5.4 exhibe o resultado do cálculo da frequência de palavras na forma de visualização do tipo *world cloud*. O histograma (Figura 5.5) mostra as palavras com as maiores frequências de ocorrências nos fluxogramas.



Figura 5.4 - Frequência de palavras visualizadas no *Word Cloud*. Fonte: elaborada pela autora.

<sup>6</sup> R-Tools Technology Inc. Disponível em: <http://www.r-studio.com/pt/>

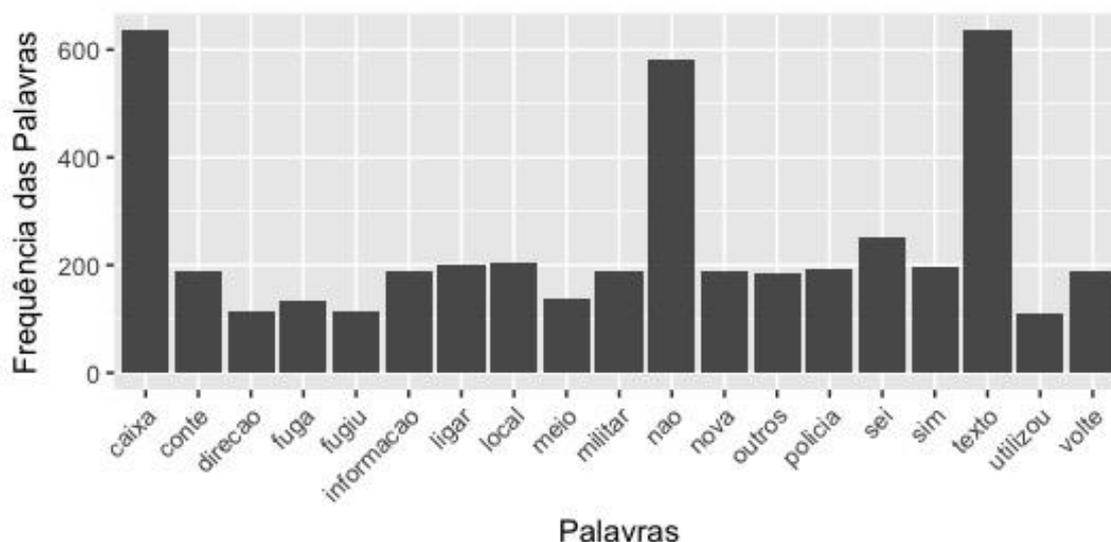


Figura 5.5 - Histograma com palavras com frequência maior que 100 ocorrências. Fonte: elaborada pela autora.

Identificadas as palavras menos relevantes a partir do *word cloud* e histograma, estas foram removidas do vocabulário. Assim sendo, esta primeira etapa finaliza-se com a criação do vocabulário gerado a partir dos 54 fluxogramas utilizados pela PMESP para atendimento às chamadas emergenciais 190. O vocabulário completo com 467 palavras, assim como exemplos de fluxogramas antes do processo de filtragem encontram-se no Apêndice A nas Figuras A.7 e A.8, respectivamente.

### 5.3 Métodos de Classificação

Esta seção apresenta os algoritmos dos métodos classificadores *Bag of Words* e *Naïve Bayes* que foram implementados e testados neste trabalho. Também foram implementados e testados os métodos Redes Bayesianas e Rede Bayesiana com Suavização de Laplace. No entanto como os resultados de ambos foram inferiores em relação aos primeiros, optou-se por suprimir sua descrição.

Dentre os diversos métodos classificadores de texto, optou-se por testar dois que diferem em sua abordagem. O método *Naïve Bayes* é um método probabilístico e sua escolha foi pautada na característica dinâmica do ambiente em que as situações ocorrem. Pela natureza dinâmica, essas situações podem iniciar de uma forma e no transcorrer do tempo se transformar em outra. Julga-se importante que os sistemas de apoio à decisão considere este fator em alinhamento ao poder decisório humano. Assim, oferecer alternativas de probabilidade para a ocorrência de situações em curso parece adequado no contexto do ambiente e do domínio em que a decisão será tomada.

Diferentemente do método *Naïve Bayes*, o método *Bag of Word* não é probabilístico e seu resultado provém do cálculo da frequência do termo no texto analisado. A inclusão deste método objetivou testar uma abordagem diferente da probabilística e assim confrontar o resultado em relação ao tratamento bayesiano.

### 5.3.1 Bag of Words

O método *Bag of Words* (descrito na seção 2.3.1) calcula a frequência de cada palavra em uma determinada frase. Neste trabalho, para calcular a frequência das palavras nas frases relatadas em chamadas 190 (como exemplo foram utilizados casos reais localizados na Internet), o algoritmo utilizou o vocabulário gerado e descrito na primeira fase do trabalho (Seção 5.2.1 e Figura A.8 do Apêndice A).

Realizou-se a contagem da frequência de cada palavra do vocabulário, em cada fluxograma, gerando uma matriz cujas linhas são os 54 fluxogramas e as colunas as 467 palavras do vocabulário. A Figura 5.6 ilustra parte da matriz *fluxogramas x vocabulário*.

Por fim, dado o texto que se pretende classificar, identificando-o como informações de uma possível situação a identificar dentre as possíveis representadas pelos fluxogramas, o método *Bag of Word* analisa a frequência das palavras do texto em cada fluxograma utilizando a matriz *fluxograma x vocabulário* (Figura 5.6).

	abandonado	aberto	aborto	abuso	acabou	acidente	dolescent	adulto	afogamento	afronta	agente
ABORTO.txt	0	0	2	0	0	0	0	0	0	0	0
AFOGAMENTO.txt	0	0	0	0	0	0	0	0	4	0	0
ALARMEDISPARADO.txt	0	0	0	0	0	0	0	0	0	0	0
ANIMAISAVES.txt	1	0	0	0	0	0	0	0	0	0	0
APOIO.txt	0	0	0	0	0	0	0	0	0	0	0
ARMADEFOGO.txt	0	0	0	0	0	0	0	0	0	0	0
ATOOBSCENO.txt	0	0	0	0	0	0	0	0	0	0	0
Arvore.txt	0	0	0	0	0	0	0	0	0	0	0
AVERIGUACAO.txt	0	0	0	0	0	0	0	0	0	0	0
BEBE.txt	1	0	0	0	0	0	0	0	0	0	0
CALAMIDADES.txt	0	0	0	0	0	0	0	0	0	0	0
CONFLITO.txt	0	0	0	0	0	0	0	0	0	0	0
CONTRABANDODESCAMINHO.txt	0	0	0	0	0	0	0	0	0	0	0
CORRUPCAODEMENORES.txt	0	0	0	0	0	0	0	1	0	0	0
CRIMECONTRAEECONOMIAPOPULAR.txt	0	0	0	0	0	0	0	0	0	0	0
CADAVER.txt	0	0	0	0	0	0	0	0	0	0	0
DESVIDECONDUTADOAGENTEPUBLICO.txt	0	0	0	1	0	0	0	0	0	0	3
DESVIDECONDUTADOCIDADA.txt	0	0	0	0	0	0	0	0	0	0	0
ENTORPECENTE.txt	0	0	0	0	0	0	0	0	0	0	0
ESTELIONATOFRAUDE.txt	0	0	0	0	0	0	0	0	0	0	0
ESTUPRO.txt	0	0	0	1	0	0	0	0	0	0	0
EXPLORACAOSEXUAL.txt	0	0	0	0	0	0	0	0	0	0	0
EXPLOSIVOS.txt	0	0	0	0	0	0	0	0	0	0	0
FALSIFICACAO.txt	0	0	0	0	0	0	0	0	0	0	0

**Figura 5.6 - Parte da matriz fluxograma X vocabulário com a frequência de cada palavra no respectivo fluxograma (árvore de decisão). Fonte: elaborado pela autora.**

O fluxograma cujo somatório das frequências de palavras analisadas possuir o maior valor será considerado com maior probabilidade de representar a situação que contextualiza aquele texto.

É importante salientar que se considera com maior probabilidade de representar a situação e não se afirma representar a situação fundamentando-se em dois princípios: (i) a situação é dinâmica e não estática, está exposta ao ambiente idênticamente dinâmico e, portanto poderá alterar-se no decorrer do tempo e (ii) é atribuído ao operador humano a decisão e aos sistemas computacionais compete apenas o apoio. Desta forma, a experiência, o treinamento, o modelo mental e até mesmo a intuição pessoal do operador poderá orientá-lo a outras inferências que não sejam as indicadas pelo resultado do método.

Ressalta-se que em se tratando de sistemas de automação para tomada de decisão, os sistemas apenas apoiam o processo decisório. Desta forma o especialista humano é o detentor do conhecimento necessário para a tomada de decisão em si. Esta é a razão pela qual os resultados dos métodos sempre irão indicar a maior probabilidade de um fluxograma representar a situação do evento em curso e, diante desta indicação o operador optará pelas melhores decisões frente os diversos elementos que se delineiam ao cenário.

Tomando como exemplo a frase contendo as palavras “**Arma Pessoa Abandonada**” iniciando a análise pelo fluxograma *Abandono.txt* e seguindo aos demais, o método Bag of Word somará a frequência de ocorrências da palavra *Arma* no fluxograma *Abandono.txt*. Em seguida calculará a ocorrência da palavra *Pessoa* no fluxograma *Abandono.txt* e somará a ocorrência da palavra *Abandonada* no fluxograma *Abandono.txt* e assim sucessivamente para todas as palavras da frase em todos os outros 53 fluxogramas.

O resultado será o somatório das ocorrências de todas as palavras da frase em cada fluxograma e aquele com o maior valor terá maior probabilidade de representar a situação daquela frase.

O algoritmo *Bag of Word* é ilustrado na Figura 5.7 e será descrito na sequência.

```
Algoritmo Bag of Word

Entrada: Matriz Fluxograma_Vocabulário
         Frase a ser analisada
Saída: Fluxograma com maior probabilidade de representar a
        situação

/* BG armazenará o somatório da frequência das palavras da
frase de cada fluxograma */

Inicializa BG = [0, 0, ..., 0]

Para cada Fluxograma f
  Para cada palavra da frase p
    BG[f] += Fluxograma_Vocabulário[f][p]

situação = Maior(BG)
retorna situação
```

Figura 5.7 - Algoritmo Bag of Words. Fonte: elaborado pela autora.

O algoritmo recebe como entrada uma frase a ser analisada e uma matriz denominada Matriz Fluxograma x Vocabulário que é composta pela Frequência das Palavras do Vocabulário em relação aos Fluxogramas.

Cada linha desta matriz representa um fluxograma e as colunas representam as palavras do vocabulário gerado.

O resultado de saída será o Fluxograma com maior probabilidade de identificar e representar a Situação correspondente à frase analisada, atribuída à variável Situação.

O vetor denominado BG é inicializado com 54 índices.

Para cada fluxograma e para cada palavra da frase o vetor BG acumula o somatório da Frequência da Palavra do fluxograma

Ao final a variável Situação recebe o maior valor do vetor BG, cujo índice corresponderá ao Fluxograma representante da Situação.

A variável Situação é apresentada e indicará o fluxograma com maior probabilidade de identificar a situação representada pela frase analisada.

### 5.3.2 Naïve Bayes

O método *Naïve Bayes* (descrito na seção 2.3.1) calcula a probabilidade de uma frase relatada em chamadas do serviço 190 da PMESP, pertencer a uma determinada situação representada por um dos fluxogramas utilizados nesses atendimentos emergenciais.

Na primeira etapa do algoritmo, é calculada a probabilidade de cada palavra do vocabulário pertencer a cada classe (fluxograma).

Para exemplificar este método serão utilizados os dados apresentados na Figura 5.8 que representa uma tabela com a frequência de algumas palavras do vocabulário em cada fluxograma. Os fluxogramas utilizados são Roubo, Furto e Homicídio, sendo 1, 2 e 3, respectivamente.

Árvores	Vocabulário				Classificação
	moto	branca	veiculo	vitima	
1	1	1	2	2	Roubo
2	1	1	1	1	Furto
3	1	1			Homicídio

Figura 5.8 - Exemplo de tabela de frequência das palavras do vocabulário nas árvores de decisão (fluxogramas). Fonte: elaborado pela autora.

O algoritmo calculará primeiramente a probabilidade da palavra *moto* pertencer à *Árvore 1* (fluxograma Roubo), a probabilidade da palavra *moto* pertencer à *Árvore 2* (fluxograma Furto) e assim sucessivamente para todas as palavras do vocabulário e para todas as árvores. A seguir a demonstração do cálculo utilizando o exemplo apresentado na Figura 5.8.

$$\begin{array}{lll}
 P(\text{moto}|1) = \frac{1+1}{6+4} = \mathbf{0,2} & P(\text{moto}|2) = \frac{1+1}{4+4} = \mathbf{0,25} & P(\text{moto}|3) = \frac{1+1}{2+4} = \mathbf{0,33} \\
 P(\text{branca}|1) = \frac{1+1}{6+4} = \mathbf{0,2} & P(\text{branca}|2) = \frac{1+1}{4+4} = \mathbf{0,25} & P(\text{branca}|3) = \frac{1+1}{2+4} = \mathbf{0,33} \\
 P(\text{veiculo}|1) = \frac{2+1}{6+4} = \mathbf{0,3} & P(\text{veiculo}|2) = \frac{1+1}{4+4} = \mathbf{0,25} & P(\text{veiculo}|3) = \frac{0+1}{2+4} = \mathbf{0,16} \\
 P(\text{vitima}|1) = \frac{2+1}{6+4} = \mathbf{0,3} & P(\text{vitima}|2) = \frac{1+1}{4+4} = \mathbf{0,25} & P(\text{vitima}|3) = \frac{0+1}{2+4} = \mathbf{0,16}
 \end{array}$$

Baseado nas probabilidades das palavras do vocabulário, dado uma frase a ser analisada, por exemplo, a frase contendo as palavras “**moto vitima**”, o método calcula a probabilidade da frase pertencer a cada classe (fluxograma). A seguir são apresentados os cálculos do produtório das probabilidades da frase para cada classe.

$$P(1): p(\text{moto}|1) * p(\text{vitima}|1) = \mathbf{0,2 * 0,3 = 0,06}$$

$$P(2): p(\text{moto}|2) * p(\text{vitima}|2) = \mathbf{0,25 * 0,25 = 0,0625}$$

$$P(3): p(\text{moto}|3) * p(\text{vitima}|3) = \mathbf{0,33 * 0,16 = 0,0528}$$

Por fim, o algoritmo realiza a ordenação dos valores obtidos na etapa anterior e o fluxograma com o maior valor terá a maior probabilidade de representar a situação que a frase indica.

Como o modelo é dinâmico, à medida que as palavras são mencionadas durante o atendimento, as probabilidades se ajustam às possíveis ocorrências e as probabilidades são recalculadas à medida que mais informações (palavras) são repassadas ao algoritmo.

A Figura 5.9 representa a classificação para a frase com as palavras “moto vitima”. No exemplo, a maior probabilidade deu-se para a situação identificada como classe 2, Furto. O algoritmo *Naïve Bayes* está ilustrado na Figura 5.10 e será descrito em seguida.

Ordenação		
P(2)	0,0625	Furto
P(1)	0,06	Roubo
P(3)	0,0528	Homicídio

Figura 5.9 - Resultado da classificação *Naïve Bayes*. Fonte: elaborado pela autora.

```

Algoritmo Naive Bayes
Entrada: Matriz Fluxograma_Vocabulário
         Frase a ser analisada
Saída: Fluxograma com maior probabilidade de representar a situação

/*MatrizProbabilidade armazenará a probabilidade de cada palavra do vocabulário
pertencer a cada fluxograma.*/

Inicializa MatrizProbabilidade

Para cada Fluxograma f
  Para cada palavra do vocabulário p
    /*frequência da palavra p no fluxograma f*/
    frequenciaPalavra = (Fluxograma_Vocabulário(f,p)

    tamanhoVocabulario = |Vocabulário|

    /*somatório das frequências de todas as palavras
do vocabulário no fluxograma f */
    totalFrequencia = Soma(Fluxograma_Vocabulário(f,))

    MatrizProbabilidade(f,p) = (frequenciaPalavra+1) /
                               (tamanhoVocabulário+totalFrequencia)

/*NB contém a probabilidade da frase pertencer a cada fluxograma*/
Inicia NB = (1,1,...,1)
Para cada Fluxograma f
  Para cada palavra da Frase p
    NB(f) *= MatrizProbabilidade(f)(p)

situação = Maior(NB)
retorna situação

```

Figura 5.10 - Algoritmo *Naïve Bayes*. Fonte elaborado pela autora.

O Algoritmo *Naïve Bayes* recebe como entrada uma frase a ser analisada e uma matriz composta pela frequência das palavras do vocabulário em relação aos fluxogramas. Cada linha desta matriz representa um fluxograma e as colunas representam as palavras do vocabulário.

O resultado de saída será o fluxograma com maior probabilidade de identificar a Situação correspondente à frase analisada.

O método *Naïve Bayes* como primeira etapa deste algoritmo, armazenará na MatrizProbabilidade a probabilidade de cada palavra do vocabulário pertencer a cada fluxograma (dos 54 existentes) e para tanto, será inicializada.

A seguir será calculada a probabilidade de cada palavra pertencer a cada fluxograma.

Para simplificar, algumas variáveis auxiliares serão utilizadas, são elas: FrequênciaPalavra (armazenará a frequência da palavra no fluxograma), TamanhoVocabulário (armazenará a quantidade de palavras do vocabulário) e TotalFrequência (armazenará o somatório da frequência de todas as palavras do vocabulário no fluxograma).

Para cada fluxograma e para cada palavra do vocabulário a MatrizProbabilidade é resultante da  $(\text{FrequênciaPalavra} + 1)$  dividido pelo  $(\text{TamanhoVocabulário} + \text{TotalFrequência})$ .

O final do ciclo, a próxima etapa é o cálculo do produtório para calcular a probabilidade da frase (com apenas as palavras a serem analisadas) pertencer a cada classe (fluxograma).

Para isso, o vetor NB é inicializado e para cada fluxograma e para cada palavra da frase a ser analisada NB recebe o produtório das probabilidades de cada palavra da frase no fluxograma.

Ao final do ciclo a variável Situação receberá o maior valor de NB.

A variável Situação é apresentada e indicará o fluxograma com maior probabilidade de identificar a situação representada pela frase analisada.

## 5.4 Considerações Finais

Neste capítulo foram apresentados o método e o processo de implementação do Modelo Preditivo de Situações.

Conforme apresentado no capítulo quatro, o modelo prevê a utilização de modelos decisórios (neste estudo as árvores de decisão) que compõem o módulo de pré consciência da situação, onde os módulos e blocos atuam no papel do modelo mental e preparam as informações para que sejam classificadas pelo módulo principal. A subseção 5.2.1 descreveu este processo quando da filtragem e criação do vocabulário. Dois métodos classificatórios com abordagens distintas foram desenvolvidos e os resultados testados e avaliados.

A descrição completa dos testes e avaliação dos resultados será apresentada no capítulo seis.

# Capítulo 6

## RESULTADOS E DISCUSSÕES

---

*Neste capítulo são apresentados e discutidos os resultados obtidos com a implementação dos métodos para desenvolvimento do Modelo para Automação e Previsão da Situação propostos neste trabalho. Este capítulo está organizado em seis seções e a primeira apresenta de forma geral como os resultados foram obtidos. A seção 6.2 descreve como é composta a matriz de confusão e quais as métricas utilizadas para avaliação de desempenho dos algoritmos desenvolvidos. A seção 6.3 especifica as instâncias de testes e a seção 6.4 detalha o planejamento experimental. Resultados são relatados e discutidos na seção 6.5 e por fim, as conclusões do capítulo são apresentadas na seção 6.6.*

### 6.1 Considerações Iniciais

Os experimentos realizados neste trabalho consistem na implementação e execução dos algoritmos baseados nos métodos de classificação de texto *Bag of Words* e *Naïve Bayes*. Os algoritmos foram submetidos a um conjunto de testes, analisando as 54 classes. Cada classe representa uma árvore de decisão utilizada pela PMESP para atendimento do serviço 190.

Os resultados foram avaliados segundo as métricas precisão, acurácia e cobertura. Para avaliar as possíveis diferenças entre o desempenho dos métodos propostos, um planejamento experimental foi utilizado possibilitando detectar diferenças significativas entre os métodos e estimar a magnitude destes.

Em uma primeira análise, foi considerada apenas a primeira classe retornada pelos algoritmos. No entanto, para corroborar a qualidade das soluções, um segundo teste foi aplicado, no qual os algoritmos retornam cinco possíveis classes para as informações de entrada. Este teste se justifica uma vez que, a consciência da situação é estabelecida com base em uma análise pessoal da realidade, considerando o modelo mental do atendente do serviço 190 – objeto de estudo. Desta forma, ao apresentar cinco possíveis classes (representando cinco situações possíveis para o evento em curso), os algoritmos colaborarão para mitigar a sobrecarga cognitiva do operador humano, contribuindo para a consciência da situação e para o processo de tomada de decisão.

## 6.2 Matriz de Confusão e Métricas para Avaliação de Desempenho

A maioria das métricas de qualidade de classificação de texto pode ser obtida a partir da matriz de confusão (FAWCETT, 2006), dentre elas: precisão, acurácia e cobertura. Essas medidas foram aplicadas neste trabalho para avaliar os algoritmos baseados nos métodos *Bag of Words* e *Naïve Bayes*, propostos neste trabalho.

Para cada algoritmo, foram executadas 30 replicações, no qual foram computados os seguintes resultados: verdadeiros positivos (Vp), verdadeiros negativos (Vn), falsos positivos (Fp), falsos negativos (Fn). Com eles é possível compor a matriz de confusão. Nesta abordagem, optou-se por construir uma matriz de confusão binária para cada uma das classes e analisá-las separadamente. A classificação binária é composta pela análise dos dados de classificação definidos corretamente e incorretamente para cada classe avaliada.

Considerando as classes estudadas neste trabalho, utilizando a classe “Furto” como exemplo, pode-se dizer que:

- Vp: quantidade de vezes que o algoritmo classificou amostras da classe “Furto” corretamente.
- Vn: quantidade de vezes que o algoritmo classificou corretamente amostras de classes diferentes de “Furto” como sendo diferentes de “Furto”.
- Fp: quantidade de vezes que o algoritmo classificou incorretamente amostras de outras classes como sendo “Furto”.
- Fn: quantidade de vezes que o algoritmo classificou incorretamente amostras da classe “Furto” como sendo de outra classe.

A Tabela 6.1 apresenta um exemplo de matriz de confusão construída com os dados obtidos da classe “Furto” pelo método de classificação *Bag of Words*. Das 1.620 amostras utilizadas, 30 pertenciam à classe “Furto”.

Nos resultados demonstrados na Tabela 6.1, 29 testes retornaram corretamente a classe “Furto”, pois as amostras de fato se referiam a esta classe. A apenas um teste retornou erroneamente outra classe, pois a amostra era referente à classe “Furto”.

Outros 16 testes retornaram erroneamente a classe “Furto”, pois as amostras pertenciam de fato a outras classes e 1.574 retornaram assertivamente outra classe, sendo as amostras referentes de fato a outras classes. Assim sendo, quanto maior os valores contidos na diagonal principal da tabela (valores em destaque que representam “acertos”), melhor o desempenho do algoritmo.

**Tabela 6.1 - Exemplo de Matriz de Confusão binária da classe “Furto”. Fonte: elaborada pela autora.**

Classe Real	Classe Prevista	
	Positivo	Negativo
Positivo	<b>29</b>	1
Negativo	16	<b>1574</b>

Uma vez definidas as matrizes de confusão binárias de cada classe, as métricas precisão, acurácia e cobertura podem ser facilmente aplicadas (MANNING, RAGHAVAN e SCHÜTZE, 2008) (RUSSELL e NORVING, 2013).

- **Precisão:** calcula o número de predições corretas dividido pelo número total de predições positivas (Fórmula 6.1):

$$precisão = \frac{Vp}{Vp + Fp} \quad (6.1)$$

- **Acurácia:** calcula a taxa de acertos do algoritmo. Quanto mais próximo de um, melhor o valor da acurácia (Fórmula 6.2):

$$acurácia = \frac{Vp + Vn}{Vp + Fp + Vn + Fn} \quad (6.2)$$

- **Cobertura:** indica a quantidade de classificações positivas que o modelo retorna (Fórmula 6.3):

$$cobertura = \frac{Vp}{Vp + Fn} \quad (6.3)$$

### 6.3 Instâncias de Teste

As instâncias são problemas-teste criados para um problema específico, tentando recriar situações reais. Assim, as instâncias servem de base para avaliar e comparar os algoritmos propostos para solucioná-las.

Para os testes deste trabalho foram utilizadas quatro tipos de instâncias diferentes, denominadas frases<sup>7</sup>. Uma frase é composta por dez palavras selecionadas aleatoriamente a partir de dois conjuntos:

1. Vocabulário geral: composto por palavras contidas nas 54 árvores de decisão.
2. Vocabulário específico: palavras oriundas da árvore de decisão de uma determinada classe.

As instâncias de testes foram divididas em quatro grupos diferentes de acordo com a composição da frase de entrada, demonstrada pela Tabela 6.2.

**Tabela 6.2 – Grupos de instâncias de acordo com a composição das frases de entrada dos métodos de classificação. Fonte: elaborada pela autora.**

Instâncias	Descrição
Teste_100	Frase contendo 100% de palavras do vocabulário específico
Teste_50	Frase contendo 50% de palavras do vocabulário específico e 50% do vocabulário geral
Teste_30	Frase contendo 30% de palavras do vocabulário específico e 70% do vocabulário geral
Teste_10	Frase contendo 10% de palavras do vocabulário específico e 90% do vocabulário geral

A instância Teste\_10 possui o grau de maior dificuldade de classificação correta, visto a quantidade reduzida de palavras próprias do vocabulário.

Para cada uma das 54 classes analisadas, foram gerados quatro tipos de testes diferentes, específicos para cada classe. Assim, nos experimentos, os algoritmos foram então avaliados a partir de 216 instâncias (54 classes multiplicados por 4 tipos de frases) de testes diferentes.

<sup>7</sup> Neste trabalho, frase consiste em um conjunto de palavras criadas pela autora que remete a alguma situação. Não necessariamente apresenta um sentido completo, conforme exigem normas gramaticais.

## 6.4 Planejamento Experimental

Esta seção apresenta o planejamento experimental adotado para avaliação dos resultados obtidos pelos algoritmos baseados nos métodos *Bag of Words* e *Naïve Bayes*. Com o planejamento do experimento e a análise estatística dos resultados é possível averiguar se há diferenças significativas entre o desempenho dos algoritmos e mensurar dessas diferenças.

Para cada algoritmo foram realizadas 30 replicações para cada classe. Os testes foram executados de maneira independente e foram avaliadas as seguintes medidas de qualidade: Precisão (P), Acurácia (A) e Cobertura (C), descritas na seção 6.2.

Para cada métrica foi aplicado o mesmo teste estatístico, buscando encontrar as diferenças significativas entre os algoritmos - considerando cada uma delas. Como foram utilizadas diferentes instâncias de teste, o modelo estatístico que descreve os dados deste experimento pode ser escrito de acordo com a Fórmula 6.4.

$$y_{ijk} = \mu_i + \beta_j + \epsilon_{ijk} \begin{cases} i = 1, \dots, a \\ j = 1, \dots, b \\ k = 1, \dots, r \end{cases} \quad (6.4)$$

O valor de  $y_{ijk}$  corresponde ao valor da métrica obtida pelo  $i$ -ésimo algoritmo na  $j$ -ésima instância da  $k$ -ésima replicação. O valor de  $\mu_i$  representa a média geral dos valores obtidos pelo  $i$ -ésimo algoritmo dada a métrica em questão e  $\beta_j$  corresponde ao efeito que a  $j$ -ésima classe provoca na variável de resposta. A variável aleatória  $\epsilon_{ijk}$  consiste no erro aleatório experimental, isto é, as variações provocadas por outros fatores que podem influenciar no valor de resposta e que não foram considerados no experimento.

Os parâmetros  $a$ ,  $b$  e  $r$  representam o número de algoritmos avaliados, o número de classes analisadas e o número de replicações do experimento, respectivamente, sendo que  $a = 2$ ,  $b = 54$  e  $r = 30$ .

Para melhorar a estimativa da diferença no desempenho entre os algoritmos avaliados, o experimento foi realizado em blocos completos, no qual cada classe corresponde a um bloco. O termo "bloco" refere-se a uma unidade experimental relativamente homogênea.

A "blocagem" reduz ou elimina alguns fatores inconvenientes que podem influenciar a resposta, mas não são interessantes para a análise. Neste trabalho, as diferenças entre as classes (tamanho do vocabulário específico, frequência das palavras do vocabulário específico, por exemplo) podem interferir nos resultados obtidos pelos algoritmos e, quando se aplica a "blocagem" dos experimentos por instância, é possível remover o fator "instância" da análise (MONTGOMERY, 1991).

Para cada métrica considerada, a hipótese nula indica a ausência de diferença entre os algoritmos avaliados *Bag of Words* (1) e *Naïve Bayes* (2) contra sua hipótese alternativa de que exista diferença entre os algoritmos.

Caso os algoritmos apresentem o mesmo desempenho e não haja diferença significativa, a hipótese nula é aceita. Por outro lado, caso haja diferenças significativas no desempenho dos algoritmos avaliados a hipótese nula é rejeitada. Para a análise dos experimentos foi considerado um nível de significância de 95% o que corresponde a um valor de  $\alpha = 0.05$ .

Para evitar suposições de que os dados apresentam uma distribuição normal, foi utilizado o teste de *Wilcoxon Signed Rank* (TAHERI e HESAMIAN, 2013). Este teste é uma alternativa onde não é possível afirmar ou justificar sobre a normalidade dos dados. Uma vez finalizado o teste de hipótese, caso a hipótese nula seja rejeitada, isso implica que há uma diferença significativa entre os algoritmos implementados. Assim, precisa-se estimar a magnitude dessa diferença. Neste trabalho o tamanho de efeito é obtido utilizando o estimador de Hodges–Lehmann (HODGES e LEHMANN, 1963).

## 6.5 Avaliações

Os testes foram executados em um computador Dell-PC, processador Intel® CORE™ i7-3612QM CPU 2.10GHz, com 8,00 GB de memória RAM e sistema operacional de 64 Bits Windows 7 Home Premium – Service Pack 1. Utilizou-se para desenvolvimento dos algoritmos de classificação o *software* MATLAB *Student Licence* R2017a.

Os testes de hipótese foram realizados para as três métricas Precisão Acurácia e Cobertura isoladamente e da mesma forma os resultados foram avaliados separadamente para cada grupo de instância: Teste\_100, Teste\_50, Teste\_30 e Teste\_10, descritos na seção 6.3. O resultado não demonstrou diferença estatisticamente significativa entre as amostras.

Conforme descrito na seção 6.3, as instâncias do grupo Teste\_100 são frases compostas por 10 palavras selecionadas aleatoriamente do vocabulário específico de cada classe. Sendo assim, são instâncias nas quais os algoritmos apresentam maior probabilidade de acerto.

As instâncias do grupo Teste\_10 são compostas por apenas uma palavra selecionada aleatoriamente do vocabulário específico e nove palavras do vocabulário geral, sendo, portanto, as instâncias nas quais os algoritmos têm maior probabilidade de erro.

Para cada grupo de instâncias, foram obtidas 3.240 amostras para cada métrica, uma vez que foram avaliados dois algoritmos em 54 classes diferentes com 30 replicações ( $2 \times 54 \times 30 = 3.240$ ). Conforme mencionado no planejamento experimental (seção 6.4), em todos os testes foi considerado um nível de significância  $\alpha = 0.05$ .

Os resultados dos testes de hipótese, para os grupos de instâncias Teste\_100, Teste\_50 e Teste\_30 indicam que a hipótese nula não deve ser rejeitada, pois não existem evidências suficientes para tal, em todas as métricas avaliadas. Porém, para as instâncias Teste\_10, todas as métricas apontam para rejeição da hipótese nula, indicando diferenças estatisticamente significativas entre os algoritmos *Bag of Words* e *Naïve Bayes*.

Todos os p-valores obtidos nas análises estatísticas são apresentados na Tabela 6.3. Os p-valores abaixo de 0.05 indicam rejeição da hipótese nula e estão destacados.

**Tabela 6.3 - P-valores obtidos após teste de hipótese. Fonte: elaborado pela autora.**

Instâncias	Métricas		
	Precisão	Acurácia	Cobertura
Teste_100	0.759	0.3725	0.4123
Teste_50	0.9019	0.8657	0.7889
Teste_30	0.625	0.3893	0.5242
Teste_10	<b>0.01293</b>	<b>6.414e-14</b>	<b>&lt; 2.2e-16</b>

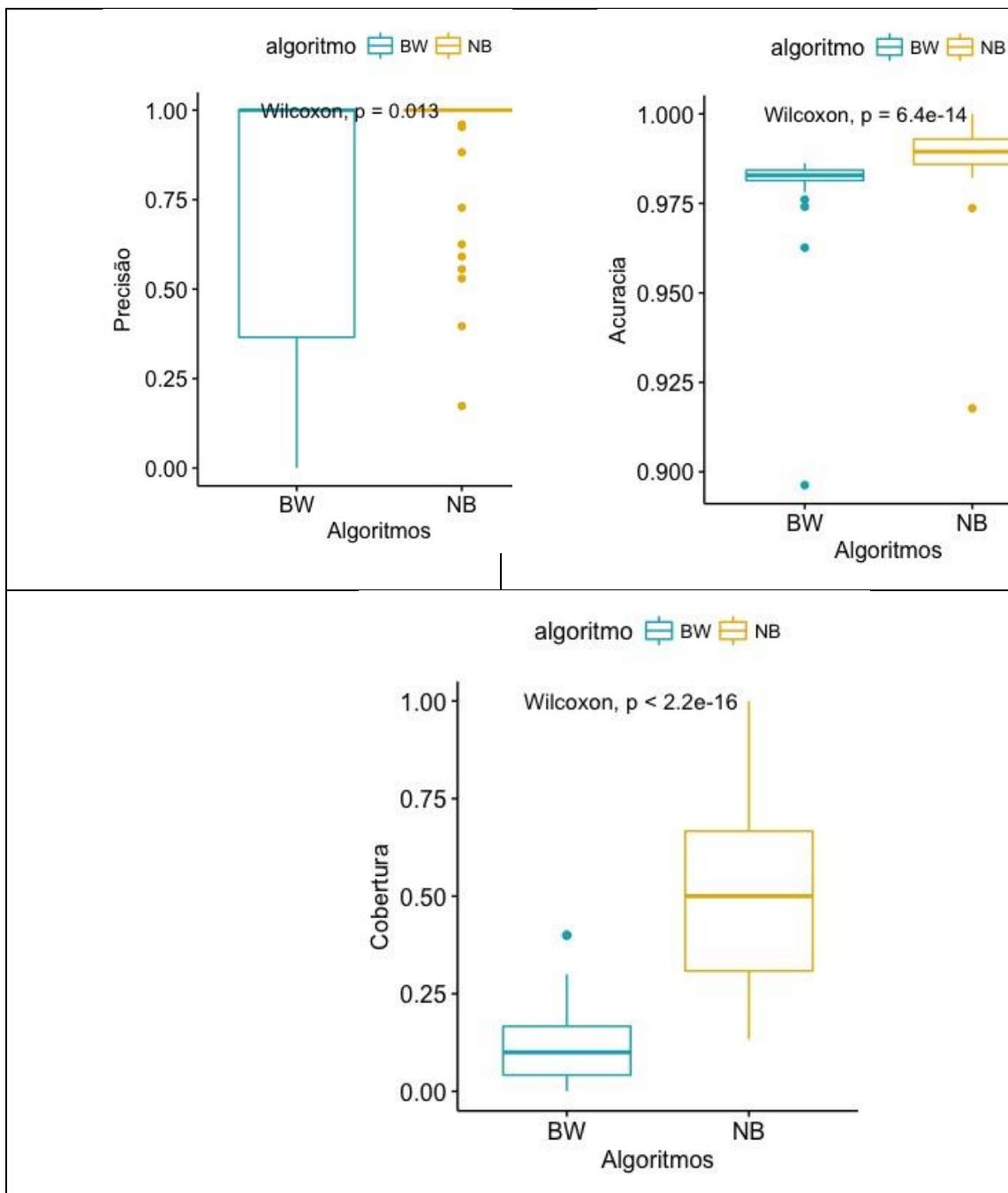
A Figura 6.1 apresenta o *boxplot* e os p-valores das métricas analisadas para o grupo de instâncias Teste\_10. Nos gráficos é possível observar a superioridade do algoritmo baseado no método de classificação *Naïve Bayes* uma vez que a mediana obtida nas métricas Acurácia e Cobertura é visivelmente maior.

Quanto à métrica Precisão a mediana entre os algoritmos não é tão aparente, no entanto é possível perceber maior variabilidade da Precisão entre as classes analisadas.

A fim de estimar a magnitude das diferenças entre os métodos, foi utilizado o estimador de Hodges–Lehmann (HODGES e LEHMANN, 1963). Para as métricas Precisão e Acurácia a diferença entre os algoritmos foi menor que 0.01. Para a métrica Cobertura essa diferença foi de 0.36, ou seja, o algoritmo *Naïve Bayes* (NB) é 36% superior ao *Bag of Words* (BW).

A superioridade do algoritmo baseado no método *Naïve Bayes* pode observada na Tabela 6.4 que apresenta as médias da taxa de acerto dos algoritmos nas 30 replicações, para cada classe analisada. Na Tabela 6.4, as métricas não estão sendo apresentadas e sim a quantidade de classes

classificadas corretamente, bem como o percentual de acerto em relação ao total de amostras.



**Figura 6.1 - Boxplot e p-valores para as métricas Precisão, Acurácia e Cobertura para o grupo de instâncias Teste\_10. Fonte: elaborado pela autora.**

Os dados apresentados corroboram com a análise experimental, principalmente ao observar a classe de instâncias com maior probabilidade de

erro, as instâncias Teste\_10, com percentual de classificação correta em 49,57% para o algoritmo *Naïve Bayes*.

**Tabela 6.4 - Média de taxa de acerto dos algoritmos BW e NB para todos os grupos de instâncias analisados. Fonte: elaborado pela autora.**

Algoritmos	Instâncias			
	Teste_100	Teste_50	Teste_30	Teste_10
Bag of Word	25,38	20,46	14,46	3,74
	84,6%	68,2%	48,2%	<b>12,47%</b>
Naïve Bayes	26,25	20,75	15,27	14,87
	87,53%	69,2%	50,93%	<b>49,57%</b>

Pesquisas na área de tomada de decisão envolvendo situações críticas em ambientes dinâmicos buscam, dentre tantos desafios, mitigar o excesso de carga cognitiva e problemas com qualidade de informação originados da situação em curso. No entanto, como a própria denominação dos sistemas propõe, o objetivo é de suporte aos operadores em suas decisões, uma vez que até o momento, nenhum algoritmo substituiu certas capacidades e habilidades humanas como a intuição e improvisação, por exemplo. Diante disso, a decisão final ainda é sempre do operador e especialista humano e pode contrariar todo indicativo do algoritmo do sistema de suporte.

Dito isto e conforme mencionado na seção de considerações iniciais, com o objetivo de contribuir na qualidade deste estudo aplicou-se um segundo teste em que o retorno aceitável como correto foram as cinco primeiras classes retornadas para as instâncias de entrada.

Pelo objeto do estudo se tratar de situações dinâmicas que podem se alterar conforme a narração do solicitante, ou mesmo conforme o desenrolar dos fatos durante o tempo, uma análise sob esta perspectiva mostrou-se válida e pertinente.

Também se considerou o processo de aquisição e manutenção da consciência da situação do atendente, de acordo com o modelo de Endsley (ENDSLEY, 1995) sendo precursor do processo decisório.

Desta forma, ao apresentar cinco possíveis classes (representando cinco situações possíveis para o evento em curso), os algoritmos colaborarão para mitigar a sobrecarga cognitiva do humano, contribuindo para a consciência da situação e ao processo de tomada de decisão.

A Figura 6.2 demonstra parte da tabela com os valores retornados para as instâncias Teste\_100, Teste\_50, Teste\_30 e Teste\_10, descritos na seção 6.3.

Classes (Árvores de Decisão)	Métodos							
	Bag of Words				Naïve Bayes			
	Instâncias				Instâncias			
	100%	50%	30%	10%	100%	50%	30%	10%
ABANDONO.txt	30	29	21	8	30	30	23	27
ABORTO.txt	30	29	24	13	30	29	25	22
AFOGAMENTO.txt	30	30	28	16	30	30	27	30
ALARMEDISPARADO.txt	30	30	28	15	30	30	25	26
ANIMAISAVES.txt	30	30	29	23	30	30	29	29
APOIO.txt	30	30	28	11	30	30	25	27
ARMADEFOGO.txt	30	30	29	18	30	30	29	26
⋮								
PRODUTOSQUIMICOS.txt	30	30	25	12	30	30	30	27
ROUBO.txt	30	27	17	8	30	29	22	25
SEQUESTROECARCEREPRIVADO.txt	30	29	26	11	30	29	28	25
SUICIDIO.txt	30	30	24	11	30	30	28	27
VEICULO.txt	30	29	25	16	30	29	25	29
VISTORIA TECNICA DE BOMBEIRO.txt	30	30	30	26	30	30	30	30
<b>Média de Acertos (Entre 5 Classes)</b>	<b>30</b>	<b>29,62</b>	<b>26,7</b>	<b>13,21</b>	<b>30</b>	<b>29,62</b>	<b>26,94</b>	<b>26,7</b>

Figura 6.2 – Recorte de parte da tabela de resultados para todos os grupos de instâncias analisados, com retorno dos cinco melhores em total de acertos.

Fonte: elaborado pela autora.

Neste segundo teste, as médias de acerto do método *Naïve Bayes* em todos os grupos de instâncias foram superiores a 25 vezes o número de

acertos das 30 iterações dos testes realizados.

Destaca-se o resultado para a instância Teste\_10 do método *Naïve Bayes* cujo desempenho ultrapassou o dobro da média do número de acertos em relação ao método *Bag of Words*. Além de se destacar a relevância das opções de situações ordenadas em classificações, como se afirmou anteriormente, visto a dinamicidade dos ambientes das situações críticas, o algoritmo do método *Naïve Bayes* ainda reafirmou-se como superior para estas classificações das situações.

## 6.6 Considerações Finais

Quanto maior a quantidade de palavras a ser analisada maior a probabilidade de acerto de se identificar uma situação (Tabela 6.4). Porém, em um contexto de situação crítica, onde o tempo é fator crucial, é necessário que esta identificação seja feita com a maior rapidez possível e no estudo de caso deste trabalho, com a menor quantidade de palavras.

Assim sendo, os testes foram realizados com 10 palavras apenas. Dessas 10 palavras (que podem formar uma frase), os métodos identificaram cinco possíveis fluxogramas – situações. Cada fluxograma é uma árvore de decisão que representa a situação principal da ocorrência que o policial atendente das chamadas emergenciais que o serviço 190 deve tratar.

Os métodos classificadores de texto avaliados se mostraram com potencial de implementação de um módulo para comporem sistemas de apoio à tomada de decisão de atendimentos emergenciais.

O capítulo sete apresenta as conclusões desta tese.

# Capítulo 7

## CONSIDERAÇÕES FINAIS

---

*Este capítulo apresenta as considerações finais deste trabalho. A Seção 7.1 apresenta o atendimento à questão de pesquisa que delineou esta tese, aos objetivos e às proposições. A Seção 7.2 trata as limitações da pesquisa. A Seção 7.3 versa sobre trabalhos futuros.*

### **7.1 Atendimento à questão de pesquisa, aos objetivos e às proposições**

Fundamentalmente, o objetivo deste trabalho pautou-se no desenvolvimento de mecanismos de apoio à consciência da situação do operador de sistemas de respostas à emergência. Neste contexto, dois modelos foram desenvolvidos: um conceitual e um preditivo no intuito de apoiar a implementação de aplicações que contemplem aspectos de SAW.

Para testar e avaliar os modelos, foram desenvolvidos dois métodos de classificação de textos, *Naïve Bayes* e *Bag of Words* e um estudo de caso com documentos e manuais fornecidos pela PMESP. Os resultados indicaram comprovações para as proposições desta tese.

*P1: A proposta do Modelo Conceitual facilita a compreensão das etapas fundamentais, bem como os elementos e ferramentas, necessários ao desenvolvimento de aplicações que contemplem aspectos de consciência da situação.*

O arcabouço que envolve o desenvolvimento de aplicações para SAW é bastante vasto. Neste sentido, o modelo conceitual que determine quais os elementos básicos presentes nesta estrutura direcionou para uma construção bem sucedida de uma aplicação para SAW. Os módulos exemplificaram objetivamente as etapas e indicaram os dados resultantes de cada uma delas. Assim foi possível compreender o fluxo e seus materiais e métodos.

*P2: O Modelo Preditivo de Situações orienta a implementação de módulos de SAW para os sistemas de apoio à tomada de decisão.*

O modelo possibilitou o desenvolvimento do módulo preditivo de forma eficiente. A simplicidade do modelo permitiu facilidade de compreensão das relações entre os submódulos. O bloco Modelos Decisórios (que atua como um modelo mental de situações prototípicas) alinha-se perfeitamente com instituições que prestam serviços emergenciais e que já possuem protocolos de atendimento, mas o modelo não é dependente desta característica (Figura 4.7). Este aspecto compôs com o estudo de caso utilizado neste trabalho que utilizou manual e árvores de decisão para atendimento da PMESP nas chamadas 190. Estas árvores de decisão geraram um vocabulário (Figura A.7 do Apêndice A) que se tornaram a fonte de consulta do modelo preditivo. Desta forma o modelo atendeu ao seu objetivo de orientar a implementação de módulos de SAW.

*P3: O uso de técnicas de IA como os classificadores de texto, pode auxiliar no processo de SAW em sistemas de apoio a tomada de decisão de situações emergenciais, especialmente os que utilizam o diálogo como principal fonte de informações, tornando o processo mais eficiente e eficaz.*

Classificadores de texto têm se tornado importantes métodos para identificação de situações quando os dados primários são textuais. O relato de

solicitantes em sistemas de respostas a emergência, como é o caso de serviços 190, atende a este quesito, pois mensagens de voz podem ser convertidas em texto e desta forma são analisadas e classificadas por estes métodos. Os testes foram realizados com os dois classificados, *Naïve Bayes* e *Bag of Words* que realizaram a classificação de frases hipotéticas baseadas no vocabulário gerado. O método *Naïve Bayes* como resultado apresentou as cinco situações com maior probabilidade de acerto. Desta forma o atendente tem a possibilidade de confrontar o resultado do método, com o relato do solicitante e ainda convergir com a sua experiência de atendimento. A automação de parte do processo de atendimento poderá contribuir para diminuir a sobrecarga cognitiva do atendente e contribuir para melhorar o processo de tomada de decisão como um todo.

*P4: A identificação preliminar da situação, no tipo de escopo da pesquisa, mitiga ou até mesmo impede que limitações de qualidade das informações se propaguem ao longo do processo decisório.*

No domínio de atendimento às emergências, o tempo é fator crucial e determinante. Além disso, o solicitante do atendimento, independente de ser vítima ou testemunha, pode encontrar-se em estado de extremo estresse, medo, pânico e tais fatores psicológicos podem influenciar e alterar as informações que transmitirá ao atendente. Vários problemas decorrentes desse fato podem ser percebidos: atraso no atendimento causando sérias consequências (dependendo da ocorrência pode até colocar em risco uma vida), incompreensão de informações, registro de dados errados, falta de informações, deslocamento desnecessário de viaturas, dentre outras. Se a informação sobre a situação da ocorrência for compreendida erroneamente e o registro for concluído, a decisão tomada em relação àquela situação conseqüentemente será diferente do que seria se estivesse no registro da situação correta. Por exemplo, a situação indicava o envio de uma viatura do SAMU, mas pela compreensão foi direcionada uma viatura da PM. Este fato ocasiona perda de tempo e pode gerar mais estresse. Diante disso, identificar precocemente a situação em que o solicitante está envolvido ou está relatando

pode mitigar e até mesmo impedir problemas como os relatados. Os testes realizados foram feitos com apenas 10 palavras e foram bastante satisfatórios. Desta forma, a proposição foi atendida.

## 7.2 Limitações da Pesquisa

O Modelo Preditivo de Situações apresenta as seguintes limitações:

- Visualização de Informação: o Modelo Preditivo de Situações trabalha com simulações, no entanto um protótipo com interface de visualização da informação tanto nos módulos de entrada como de saída agregarão grandemente ao projeto.
- Medição de SAW: tão importante quanto apoiar à consciência da situação, possibilitar que o sistema produza medidas de SAW é fundamental. Durante o desenvolvimento desta tese foram estudadas ferramentas de medição que deverão ser incorporadas ao projeto. Não estão presentes por se tratar de uma simulação.

## 7.3 Trabalhos Futuros

Como trabalhos futuros desta tese pretende-se inicialmente:

- Criação de metodologia para desenvolver aplicações a partir do modelo conceitual.
- Teste em ambiente real: apesar do módulo decisório e o vocabulário estarem balizados nas árvores de decisão utilizadas pela PMESP e este trabalho ter tido todo apoio da instituição, não foi possível testar em situações reais e em tempo real.

- Desenvolvimento da interface para a aplicação.
- Integração com módulo HUMINT Data Acquisition (Botega et al., 2017)
- Utilização de outras fontes de dados, como postagens de redes sociais.
- Validação dos testes em ambiente real, utilizando a estrutura do COPOM-Campinas. Em ambiente real, em diálogo com os operadores do serviço 190, realizar uma análise rigorosa da contribuição desta pesquisa em relação à atual árvore de decisão utilizada pela PMSP.
- Aplicação do Modelo Preditivo de Situações em domínios diferentes como, por exemplo, Corpo de Bombeiros ou SAMU.
- Avaliação de consciência situacional, com técnicas do tipo SART (*Situation Awareness Rating Technique* – Técnica de Pontuação de Consciência da Situação) (Taylor, 1990) que se baseia em um questionário em que os operadores indicam o quão cientes estiveram durante um experimento. A medição de SAW é de importante para aferir a eficiência e a contribuição da aplicação.
- A partir dos resultados obtidos com a medição e avaliação da SAW, analisar o grau de informatização possível com o ferramental.

# REFERÊNCIAS

---

AGARWAL, S.; JAIN, N.; DHOLAY, S. Adaptive testing and performance analysis using naive bayes classifier. **Procedia Computer Science**, v. 45, n. C, p. 70–75, 2015.

ALAGOZ, O. et al. Markov decision processes: a tool for sequential decision making under uncertainty. **Medical decision making : an international journal of the Society for Medical Decision Making**, v. 30, n. 4, p. 474–83, 2010.

ALMEIDA, A. S. **Relatório de Manutenção do Sistema RMS - SIOPM - CORPE** Estado de São Paulo Polícia Militar do Estado de São Paulo, , 2013.

ANAGNOSTOPOULOS, C.; HADJIEFTHYMIADES, S. Enhancing Situation-Aware Systems through Imprecise Reasoning. **IEEE TRANSACTIONS ON MOBILE COMPUTING**, v. 7, n. 10, p. 1153–1168, 2008.

ARAMAKI, E.; MASKAWA, S.; MORITA, M. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. **Computational Linguistics**, v. 2011, p. 1568–1576, 2011.

BARWISE, J.; PERRY, J. **Situations and Attitudes** *Journal of Philosophy*, 1981. Disponível em: <<http://www.jstor.org/stable/2026578>>

BATINI, C.; SCANNAPIECO, M. **Data Quality**. [s.l.: s.n.].

BAUMGARTNER, N. et al. BeAware!—Situation awareness, the ontology-driven way. **Data & Knowledge Engineering**, v. 69, n. 11, p. 1181–1193, nov. 2010.

BAUMGARTNER, N. et al. A tour of BeAware – A situation awareness framework for control centers. **Information Fusion**, v. 20, p. 155–173, nov. 2014.

BEDNY, G.; MEISTER, D. Theory of Activity and Situation Awareness. **International Journal of Cognitive Ergonomics**, v. 3(1), n. 1, p. 63–72, 1999.

BENNETT, C. C.; HAUSER, K. Artificial Intelligence in Medicine Artificial

intelligence framework for simulating clinical decision-making : A Markov decision process approach. **Artificial Intelligence In Medicine**, v. 57, n. 1, p. 9–19, 2013.

BILLINGS, C. **Situation Awareness Measurements and Analysis- A Commentary**. Conf. Experimental Analysis and Measurement of Situation Awareness. **Anais...**1995

BLASCH, E. et al. Revisiting the JDL model for information exploitation. **Proceedings of the 16th International Conference on Information Fusion, FUSION 2013**, p. 129–136, 2013.

BOTEGA, L. C. et al. **Quality-Aware Human-Driven Information Fusion Model**. 20th International Conference on Information Fusion. **Anais...**Xi'an, China: IEEE Computer Society, 2017

BRYNJOLFSSON, E.; MCAFEE, A. **Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy**MassachusettsDigital Frontier Press Lexington, , 2012.

CASTALDO, F.; PALMIERI, F. A. N.; REGAZZONI, C. S. Bayesian Analysis of Behaviors and Interactions for Situation Awareness in Transportation Systems. **IEEE Transactions on Intelligent Transportation Systems**, v. 17, n. 2, p. 313–322, 2016.

CHELBA, C. et al. Large Scale Language Modeling in Automatic Speech Recognition. **CoRR**, v. abs/1210.8, 2012.

CHEN, H. et al. **SOUPA : Standard Ontology for Ubiquitous and Pervasive Applications \***. Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services. **Anais...**IEEE Computer Society, 2004

CHOWDHURY, G. G. Natural Language Processing. In: **Annual Review of Information Science and Technology**. [s.l: s.n.]. v. 37 (1)p. 51–89.

DAPOIGNY, R.; BARLATIER, P. **Formal foundations for situation awareness based on dependent type theory**Information Fusion, 2013.

Disponível em: <<http://dx.doi.org/10.1016/j.inffus.2012.02.006>>

DAY, R. W. . et al. Improving patient flow in a hospital through dynamic allocation of cardiac diagnostic testing time slots. **Decision Support Systems**, v. 49, n. 4, p. 463–473, 2010.

DEIVA RAGAVI, M.; USHARANI, S. Social Data Analysis for Predicting Next Event. **International Conference on Information Communication and Embedded Systems - IEEE**, n. 978, 2014.

DEKKER, S.; HOLLNAGEL, E. Human factors and folk models. **Cognition, Technology & Work**, v. 6, n. 2, p. 79–86, 2004.

DOMINGUEZ, C. **Can SA be Defined?**Ohio: Armstrong LaboratoryEm: Vidulich, M.; Dominguez, C.; Vogel, E.; McMillan, G. (Eds.). Situation awareness: papers and annotated bibliography, , 1994.

DUARTE, N. **Terceirização do Atendimento do 190 Gera Polêmica entre Militares**São Paulo-SP, 2010. Disponível em: <<http://g1.globo.com/Noticias/Brasil/0,,MUL1473902-5598,00-TERCEIRIZACAO+DO+ATENDIMENTO+DO+GERA+POLEMICA+ENTRE+MILITARES.html>>

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. 2nd. ed. Nova Jersey: Wiley and Sons, 2001.

ENDSLEY, M.; GARLAND, D. **Situation Awareness Analysis and Measurement**. 10. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2000a.

ENDSLEY, M. R. Situation awareness global assessment technique (SAGAT). **Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National**, p. 789–795, 1988.

ENDSLEY, M. R. Toward a theory of situation awareness in dynamic systems. **Human Factors**, v. 37(1), p. 32–64, 1995.

ENDSLEY, M. R. Situation Models: An Avenue to the Modeling of Mental

Models. **Congress of the International Ergonomics Association and h Annual Meeting of the Human Factors and Ergonomics Society**, v. 44, n. 1, p. 61–64, 2000a.

ENDSLEY, M. R. Theoretical Underpinnings of Situation Awareness: A Critical Review Process. In: ENDSLEY, M. R.; GARLAND, D. J. (Eds.). . **Situation Awareness Analysis and Measurement**. [s.l.] Mahwah, NJ: Lawrence Erlbaum Associates, 2000b. p. 1–24.

ENDSLEY, M. R. Situation Awareness Misconceptions and Misunderstandings. **Journal of Cognitive Engineering and Decision Making**, v. 9, n. 1, p. 4–32, 2015.

ENDSLEY, M. R.; CONNORS, E. S. Situation Awareness: State of the Art. **Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE**, p. 1–4, 2008.

ENDSLEY, M. R.; GARLAND, D. J. Pilot Situation Awareness Training in General Aviation. **Proceedings of the Human Factors and Ergonomics Society Annual Meeting**, v. 44, n. 11, p. 357–360, 2000b.

ENDSLEY, M. R.; JONES, D. G. **Designing for Situation Awareness: An Approach to User-Centered Design**. [s.l.] Boca Raton-FL: CRC Press, 2012.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, p. 861–874, 2006.

FENZA, G. et al. Agent-based cognitive approach to airport security situation awareness. **CISIS 2010 - The 4th International Conference on Complex, Intelligent and Software Intensive Systems**, p. 1057–1062, 2010.

FENZA, G.; FURNO, D. Agent-based Cognitive approach to Airport Security Situation Awareness Awareness. **Computer Society IEEE**, n. February, 2010.

FISCHER, Y.; BEYERER, J. Defining dynamic Bayesian networks for probabilistic situation assessment. **Information Fusion (FUSION), 2012 15th International Conference on**, p. 888–895, 2012.

FREY, C. B.; OSBORNE, M. A. THE FUTURE OF EMPLOYMENT : HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION? **Technological Forecasting and Social Change**, v. 114, p. 254–280, 2017.

FURNO, D.; LOIA, V.; VENIERO, M. A fuzzy cognitive situation awareness for airport security. **Control and Cybernetics**, v. 39, n. 4, p. 959–982, 2010.

GABA, D. M.; HOWARD, S. K.; SMALL, S. D. **Situation awareness in anesthesiology** *Human factors*, 1995.

GIBSON, J. J. **The Ecological Approach to Visual Perception**. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

GOCGUN, Y. et al. Artificial Intelligence in Medicine A Markov decision process approach to multi-category patient scheduling in a diagnostic facility. **Artificial Intelligence In Medicine**, v. 53, n. 2, p. 73–81, 2011.

GOULIONIS, J. E.; KOUTSIUMARIS, B. K. Partially observable Markov decision model for the treatment of early Prostate Cancer. **OPSEARCH**, v. 47, n. 2, p. 105–117, 2010.

GROSS, G.; NAGI, R.; SAMBHOOS, K. A fuzzy graph matching approach in intelligence analysis and maintenance of continuous situational awareness. **Information Fusion**, v. 18, p. 43–61, 2014.

HANIFAH, R.; SUPANGKAT, S. H.; PURWARIANTI, A. Twitter information extraction for smart city. **Proceedings - 2014 International Conference on ICT for Smart Society: “Smart System Platform Development for City and Society, GoeSmart 2014”, ICISS 2014**, p. 295–299, 2014.

HODGES, J. L.; LEHMANN, E. L. Estimates of Location Based on Rank Tests. **Mathematical Statistics**, v. 34–2, p. 598–611, 1963.

HUQ, R. . et al. **A decision-theoretic approach in the design of an adaptive upper-limb stroke rehabilitation robot**. IEEE International Conference on Rehabilitation Robotics. **Anais...Zurich**: 2011Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0->

80055060839&partnerID=40&md5=5f2ff70d8601eb2398dcce1da57f4b07>

JEANNOT, E. **Situation Awareness: Synthesis of Literature SearchEUROCONTROL Experimental Centre**. Brétigny-sur-Orge: [s.n.]. Disponível em:

<<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:REvisiting+the+Swiss+Cheese+model+of+accidents#0>>.

JOHNSON-LAIRD, N. P. **The history of Mental Models Psychology of reasoning: Theoretical and historical perspective**, 2004.

JOHNSON-LAIRD, P. Mental Models and Cognitive Change. **Journal of Cognitive Psychology**, 2013, n. March, p. 37–41, 2013.

JOHNSON-LAIRD, P. N. Mental Models and Deduction. **Trends in Cognitive Sciences**, v. 5, n. 10, p. 434–442, 2001.

KHAN, A. U. R.; KHAN, M.; KHAN, M. B. Naive Multi-label Classification of YouTube Comments Using Comparative Opinion Mining. **Procedia Computer Science**, v. 82, n. March, p. 57–64, 2016.

KLEIN, G. **Naturalistic Decision Making: Implications for Design**. Information Analysis Center Wright-Patterson Air Force Base. **Anais...Ohio**: 1993

KOKAR, M. M.; ENDSLEY, M. R. Situation Awareness and Cognitive Modeling. **IEEE Intelligent Systems**, v. 27, n. 3, p. 91–96, 2012.

KOKAR, M. M.; MATHEUS, C. J.; BACLAWSKI, K. **Ontology-based situation awareness Information Fusion**, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.inffus.2007.01.004>>

KOKAR, M. M.; NG, G. W. High-level information fusion and situation awareness. **Information Fusion**, v. 10, p. 2–5, 2009.

KOTEVSKA, O.; PADI, S.; LBATH, A. Automatic Categorization of Social Sensor Data. **Procedia - Procedia Computer Science**, v. 98, n. WoTBD, p. 596–603, 2016.

LASKOWSKI, M. A Prototype Agent Based Model and Machine Learning Hybrid System for Healthcare Decision Support. In: **Digital Advances in Medicine, E-Health, and Communication Technologies**. [s.l: s.n.]. p. 230–232.

LEEDOM, D. K. **Sensemakin Symposium Final Report**. Sensemakin Symposium - Command and Control Research Program Office of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence. **Anais...2001**

LIDDY, E. D. Enhanced Text Retrieval Using Natural Language Processing. **Bulletin of the American Society for Information Science**, p. 3, 1998.

LIDDY, E. D. Natural Language Processing. In: MARCEL DECKER, I. (Ed.). . **Encyclopedia of Library and Information Science**. 2nd. ed. NY: [s.n.]. v. 37p. 51–89.

LIGGINS, M. E.; HALL, D. L.; LLINAS, J. **Handbook of Multisensor Data Fusion- Theory and Practice**. Second Edi ed. [s.l.] CRC Press, 2009.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to Information Retrieval**. [s.l.] Cambridge University Press, 2008.

MARTINS, R. A. **Abordagens quantitativa e qualitativa**. (P. A. . (Org. ). In: MIGUEL, Ed.)In: MIGUEL, Paulo A.M (Org.). **Anais...Rio de Janeiro: Elsevier, 2010**

MARUCCI-WELLMAN, H. R.; LEHTO, M. R.; CORNS, H. L. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Na??ve Bayes algorithms. **Accident Analysis and Prevention**, v. 84, p. 165–176, 2015.

MATHEUS, C. J. et al. Lessons learned from developing SAWA: A situation awareness assistant. **2005 7th International Conference on Information Fusion, FUSION**, v. 2, p. 969–976, 2005a.

MATHEUS, C. J. et al. SAWA: An Assistant for Higher-Level Fusion and Situation Awareness. **SPIE Conference on Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications**, p. 75–85, 28 mar. 2005b.

MATHEUS, C. J.; KOKAR, M. M.; BACLAWSKI, K. A core ontology for situation awareness. **Information Fusion, 2003. Proceedings of the Sixth International Conference of**, v. 1, p. 545–552, 2003.

MCCALLUM, A.; NIGAM, K. A Comparison of Event Models for Naive Bayes Text Classification. **AAAI/ICML-98 Workshop on Learning for Text Categorization**, p. 41–48, 1998.

MIGUEL, P. A. C. **Adoção do estudo de caso na engenharia de produção**. In: MIGUEL, Paulo A.M. (Org.). **Anais...**Rio de Janeiro: Elsevier, 2010

MONTGOMERY, D. C. **Design and Analysis of Experiments**. [s.l.] John Wiley, 1991.

NADERPOUR, M.; LU, J.; ZHANG, G. The explosion at institute : Modeling and analyzing the situation awareness factor. **Accident Analysis and Prevention**, v. 73, p. 209–224, 2014.

NIESSEN, C.; EYFERTH, K.; BIERWAGEN, T. Modelling cognitive processes of experienced air traffic controllers. **Ergonomics**, v. 42, n. 11, p. 1507–20, 1999.

NOY, N. F.; KLEIN, M. Ontology Evolution : Not the Same as Schema Evolution. n. July 2002, p. 428–440, 2004.

OKAZAKI, T.; OHYA, M. A study on situation awareness of marine pilot trainees in crowded sea route. **Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on**, p. 1525–1530, 2012.

PACÍFICO, F.; CARVALHO, M. Estado de SP Planeja Terceirização de Serviço de Atendimento 190 da PM. p. 3, set. 2013.

PARASURAMAN, R.; SHERIDAN, T.; WICKENS, C. Situational Awareness, Mental Workload and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. **Journal of Cognitive Engineering and Decision Making**, v. 2, n. 2, p. 140–160, 2008.

PATRICK, J.; MORGAN, P. L. Approaches to understanding, analysing and

developing situation awareness. **Theoretical Issues in Ergonomics Science**, v. 11, n. October 2014, p. 41–57, 2010.

PILATO, G. et al. Integration of ontologies and Bayesian networks for maritime situation awareness. **Proceedings - IEEE 6th International Conference on Semantic Computing, ICSC 2012**, p. 170–177, 2012.

RASMUSSEN, J. On the Structure of Knowledge - a Morphology of Mental Models in a Man- Machine System Context. **Psychological Bulletin**, v. 100(3), n. November, p. 349–363, 1979.

REIN, K.; BIERMANN, J. Your high-level information is my low-level data - A new look at terminology for multi-level fusion. **2013 IEEE International Conference on Intelligence and Security Informatics**, p. 412–417, 2013.

REZENDE, F. S. **S.O.S. POLÍCIA! - ESTRESSE NO TRABALHO Estudo no Centro Integrado de Comunicações Operacionais (Atendimento de Urgência 190) da Polícia Militar de Minas Gerais na Região Metropolitana de Belo Horizonte**. Belo Horizonte-MG: [s.n.].

ROUSE, W. B.; MORRIS, N. M. **On looking into the black box: Prospects and Limits in the Search for Mental Models**. Center for Man-Machine Systems Research. Georgia Institute of Technology. **Anais...Atlanta, Georgia: 1986** Disponível em: <<http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.100.3.349>>

ROY, J.; BRETON, R.; ROSSEAU, R. Decision Making Models. In: **Concepts, Models, and Tools for Information Fusion**. 10. ed. Norwood, MA: Artech House, 2007. p. 11–24.

ROY, J.; BRETON, R.; ROUSSEAU, R. Situation Awareness and Analysis Models. In: **In: BOSSÉ, É.; ROY, J.; WARK, S. Concepts, Models, and Tools for Information Fusion**. [s.l.] Norwood-MA: Artech House, 2007. p. 27–64.

ROY, J.; DAVENPORT, M. Exploitation of maritime domain ontologies for anomaly detection and threat analysis. **2010 International Waterside Security Conference, WSS 2010**, 2010.

RUSSELL, S.; NORVING, P. **Artificial Intelligence**. 3rd. ed. [s.l.] Pearson Education, Inc., 2013.

SAKAI, T.; TAMURA, K.; KITAKAMI, H. Density-based Spatiotemporal Analysis System with Photo Image Classifier Using the BoF Model. v. 1, n. 4, p. 85–94, 2015.

SAKAKI, T. et al. Real-time event extraction for driving information from social sensors. **Proceedings - 2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, CYBER 2012**, p. 221–226, 2012.

SALMON, P. M. et al. What really is going on? Review of situation awareness models for individuals and teams. **Theoretical Issues in Ergonomics Science**, v. 9, p. 297–323, 2007.

SEPPÄNEN, H. et al. Developing shared situational awareness for emergency management. **Safety Science**, v. 55, p. 1–9, 2013.

SHAN, S. ; WANG, L. ; LI, L. . Modeling of emergency response decision-making process using stochastic Petri net: An e-service perspective. **Information Technology and Management**, v. 13, n. 4, p. 363–376, 2012.

SHELTON, C. L. et al. Real-time situation awareness assessment in critical illness management: adapting the situation present assessment method to clinical simulation. **BMJ Quality & Safety**, v. 22, p. 163–167, 2013.

SMITH, K.; HANCOCK, P. A. Situation Awareness is Adaptive, Externally-Directed Consciousness. In: D.GILSON, R. D.; J. GARLAND, D. J.; M. KOONCE, J. M. (Eds.). . **Situation Awareness in Complex Systems**. Daytona Beach, FL.: Embry-Riddle Aeronautical University Press, 1994. v. 53p. 160.

SNIDARO, L.; VISENTINI, I.; BRYAN, K. Fusing uncertain knowledge and evidence for maritime situational awareness via Markov Logic Networks. **INFORMATION FUSION**, 2013.

SORENSEN, L. J.; STANTON, N. A.; BANKS, A. P. Back to SA school:

contrasting three approaches to situation awareness in the cockpit. **Theoretical Issues in Ergonomics Science**, v. 12, n. 6, p. 451–471, 2011.

STANTON, N. A. et al. Is situation awareness all in the mind? **Theoretical Issues in Ergonomics Science**, v. 11, n. March 2014, p. 29–40, 2010.

TAHERI, S. M.; HESAMIAN, G. A generalization of the Wilcoxon signed-rank test and its applications. **Statistical Papers**, v. 54, n. 2, p. 457–470, 2013.

TAMEA, G. et al. Situation awareness in airport environment based on Semantic Web technologies. **IEEE International Inter-Disciplinary Conference on**, p. 174–180, 2014.

TAYLOR, R. M. **Situational Awareness Rating Technique(SART)- the development of a tool for aircrew systems design**. In: SITUATIONAL AWARENESS IN AEROSPACE OPERATIONS (AGARD-CP-478). **Anais...NATO Science and Technology Organization**, 1990

VALIENTE, M.; MACHÍN, R.; GARCÍA-BARRIOCANAL, E. An Ontology-Based Integrated Approach to Situation Awareness for High-Level Information Fusion in C4ISR. **Springer Verlag**, v. 83 LNBIP, p. 513–527, 2011.

VINCEN, D.; STAMPOULI, D.; POWELL, G. **Foundations for System Implementation for a Centralised Intelligence Fusion Framework for Emergency Services**. International Conference on Information Fusion. **Anais...Seattle, WA, USA: 2009**

VOGEL, E. **SA- An Operational Point of View** Ohio: Armstrong Laboratory Em: Vidulich, M.; Dominguez, C.; Vogel, E.; McMillan, G. (Eds.). Situation awareness: papers and annotated bibliography, , 1994.

WALLACH, H. M. **Topic Modeling : Beyond Bag-of-Words**. Proceedings of the 23rd International Conference on Machine Learning. **Anais...Pittsburgh, PA: 2006**

WICKENS, C. D. Endsley\_Review. v. 50, n. 3, p. 397–403, 2008.

WICKENS, C. D.; HOLLANDS, J. G. Attention in Perception and Display Space. In: ROBERTS, N. (Ed.). . **Engineering psychology and human performance**. 3rd. ed. New Jersey: Prentice Hall, 2000. p. 69–118.

WILLIAMS, M. D.; HOLLAN, J. D.; STEVENS, A. L. Human Reasoning About a Simple Physical System. In: GENTNER, D.; STEVENS, A. L. (Eds.). . **Mental Models**. New York, NY: Psychology Press, 2014. p. 131–153.

WIRSTAD, J. **On knowledge Structures for Process Operators**. [s.l: s.n.].

WONG, B. L. W.; BLANDFORD, A. **Naturalistic Decision Making in Emergency Ambulance Command and Control** University of Otago Information Science Discussion Papers Series No. 2001/11, , 2001.

XIANG, Y.; POH, K. L. Time-critical dynamic decision modeling in medicine. **Computers in Biology and Medicine**, v. 32, n. 2, p. 85–97, 2002.

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

YIN, J. et al. Using social media to enhance emergency situation awareness. **IEEE Intelligent Systems**, v. 27, p. 52–59, 2012.

YIN, R. K. **Estudo de Caso: planejamento e método**. 2. ed. ed. São Paulo: Bookman, 2001.

ZHANG, Y.; JIN, R. Understanding bag-of-words model: a statistical framework. **Int. J. Mach. Learn. & Cyber**, p. 43–52, 2010.

# Apêndice A - CÓDIGOS

---

```
$ cat *.txt  
adolescente  
apto  
bombeiro  
samu  
casa  
comercial  
crianca  
deficiente  
deficiente  
fisico  
mental  
ensino  
estabelecimento  
estabelecimento  
ferimento  
idoso  
material  
publica  
recem  
nascido  
residencia  
responsavel  
aborto  
aborto  
agressao  
agressor  
....
```

Figura A.1 - Listagem de todas as palavras de todos os fluxogramas. Fonte: elaborado pela autora.



```
cat abandono.txt | sed '/^$/d;s/[[:blank:]]//g' > abandono1.txt
```

Figura A.5 - Remoção de espaços em branco. Fonte: elaborado pela autora.

```
#!/bin/bash

for file in *.txt; do
    iconv -t utf-8 "$file" > "${file%.txt}.txt1"
done

rm *.txt
for file in *.txt1;do
    iconv -t utf-8 "$file" > "${file%.txt1}.txt"
done

rm *.txt1
```

Figura A.6 - Conversão de caracteres ISO8591-1 para UFT-8. Fonte: elaborado pela autora.

abandonado aberto aborto abuso acabou acidente adolescente adulto afogamento afronta agente  
agressao agressivo agressor agua aguarde alarme alcoolica alerta alheia alto ambientais ambiental  
ambiente ameaca ameacada animal apartamento apartamentoto apartameto apoio apreendido  
apreensao apropriacao aproximadamente apto arbitraria arma armada armado armados arrastao  
arrombamento artefato artificio asilo assalto ataque atentado atingido atingiu atitude ato  
atropelamento atual autoridade auxilio ave averiguacao avistado azar baile balao baldio bancaria  
banco bandeira bando bar barulho bebe bebida belico bicho bingo bomba bombeiro branca  
briga caca cachoeira cadaver calamidade calunia caminhao caminho caps carcere cardiaca  
carro casa casal cassino cativo cd celular cemiterio cerimonia cerol cet choque cidadao  
cinco cliente clinica comercial comunidade concussao conduta conflito congestionamento  
constrangimento contrabando corporal corregedoria corrupcao crianca crianas crime criminosa  
culposa culposo dano deficiente depredacao desabamento desacato desaparecida  
desaparecimento descaminho desinteligencia deslizamento desmatamento desmoroamento  
desobediencia despejo desprotegido destruicao desvio detencao detentos difamacao dinamite  
dinheiro direcao discriminacao discussao disparado disparo distrito diverso documento  
documentos doenca doloso domicilio drogas dvd economia edificacao educativa eletrica  
eletrico embarcacao embriaguez emergencia enchente encontrado enforcamento ensino  
entorpecente entulho envenenamento envolvidas envolvido envolvidos equipe escola escrita  
espancamento esquenta estabelecimento estacionado estelionato estrutura estupro exercicio  
exploracao explosao explosivo explosivos exposicao extorsao faculdade faixas falsa falsidade  
falsificacao falso familia ferida ferido ferimento festa fiacao filho fisicas fisico floresta fogo  
formacao forte fraude fuga fugiu funcao funcionario funeraria funk furtados furto galho gaz  
gcm gestante glp golpe granada grande gravidade greve gritando homens homicidio  
homofobia hospital idade identidade ideologica idoso ilegal iminente importunacao inalacao  
incapaz incendio inconveniente indebita individuo individuos industria industrial induzido  
infanticidio infecciosa informar infracao ingestao inidonea injecao injuria inquilino inseto  
instrumentos interdicao interditando interior intolerancia intoxicacao intrusao inundacao  
invadindo invasao ira jogando jogo judicial justica lanchonete latrocínio lenocínio lesao  
liberdade libidinosos lixo localizada localizado mae mal manancial manifestacao mar  
masturbando mata material mau mendicancia menor menores mental mercadoria mergulho  
morta morte morto moto motorista movel movimento mulher mulheres municipal musicais  
nascido natural niquel nome nua numeracao objeto obsceno ocultacao ofendida ofensa  
ofensiva oficial omissao operacional orientar overdose pacifica pancadao panico particular  
passageiro passeata patrimonio pe peculato pequena perda perdido periclitacao perigosa  
perigoso perturbacao pessoa pessoas piquete piscina placa plantacao policial popular porte  
posse potavel preconceito preso pressao prevaricacao prisional privacao privado procurado  
produtos profissao propriedade proprietario prostituicao publica publico pudor quadrilha queda  
queima queimada queimadura queixa quimico quimicos racismo rapto rastreador rebeliao  
recem recipiente reclamante rede refem reiteracao reivindicacao repouso represa resgate  
residencia resistencia responsavel restricao reuniao rg rio risco rixa rodovia roubo ruptura  
rural salto samu saque seguro sepultura sequestro sexual sinalizar situacao socorrida socorro  
soltura som sossego soterramento subito substancias subtracao suicidio superauecimento  
suspeita suspeito tampinha tecnica tempo tentativa terra terreno tortura trafico trajés  
transeunte transito transporte trato trauma tumulto ubs usurpacao vadiagem vapor vazamento  
veiculo velorio venda via viatura vida vilipendio violacao violencia violenta violento visita  
vistoria vitima vitimas vizinho vizinhos

**Figura A.7 – Vocabulário completo filtrado com 467 palavras. Fonte: elaborado pela autora.**

Árvore de Decisão

Animais/Aves

M04 - Inseto Agressivo  
 N12 - Animal Agressivo  
 N26 - Animal em Situação de Risco  
 N28 - Animal

C-04 C-02 C-06 A-10  
 A-15 A-16 A-17

ANIMAIS/AVES

O que está acontecendo?

Inseto  
 Aves  
 Animal

Qual é o animal?

Caixa de Texto  
 Em que situação está?

Morto

Abandonado  
 Preso/cativeiro

Ferido  
 Outros  
 Situação de Risco  
 Agressivo  
 Perguntas do Bombeiro/SAMU

Perguntas Polícia Ambiental  
 Caixa de Texto

Qual local está?

Casa  
 Estabelecimento  
 Comercial  
 Outros  
 Residência

Via Pública

Caixa de Texto  
 Comunidade  
 Apto

Caixa de Texto  
 Quais as características?  
 Visualizou alguém no local?

Não/Não Sei  
 Sim

Caixa de Texto

Se tiver nova informação volte a ligar, conte com a Polícia Militar.

Figura A.8 – Palavras extraídas da Árvore de Decisão “Animais/Aves”. Fonte: PMESP.

# APÊNDICE B - PUBLICAÇÕES

---

SOUZA, J. O.; BOTEGA, LEONARDO CASTRO; SANTAREM SEGUNDO, JOSE EDUARDO; BERTI, C. B. User-Driven Methodology for Data Quality Assessment in the Context of Robbery Events. **ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING**, v. 1, p. 161-170, 2015.

SOUZA, J. O.; Botega, Leonardo; SEGUNDO, JOSE EDUARDO SANTAREM; BERTI, C. B.. A Methodology for the Assessment of the Quality of Information from Robbery Events to Enrich Situational Awareness in Emergency Management Systems. **PROCEDIA MANUFACTURING**, v. 3, p. 4407-4414, 2015.

SOUZA, J. O.; BOTEGA, L. C.; SANTAREM, J. E.; BERTI, C. User-driven Methodology for Data Quality Assessment in the Context of Robbery Events. In: **3rd World Conference on Information Systems and Technologies**, 2015, Azores. *Advances in Intelligent Systems and Computing*, 2015.

SOUZA, J. O.; BOTEGA, L. C.; SANTAREM, J. E.; BERTI, C. A Methodology for the Assessment of the Quality of Information from Robbery Events to Enrich Situational Awareness in Emergency Management Systems. In: **6th International Conference on Applied Human Factors and Ergonomics**, 2015, Las Vegas. *6th International Conference on Applied Human Factors and Ergonomics*. 2015.

BOTEGA, L. C.; BERTI, C; ARAUJO, R. B.; DE NERIS, VÂNIA PAULA. A Model to Promote Interaction Between Humans and Data Fusion Intelligence to Enhance Situational Awareness. In: **16th International Conference on Human-Computer Interaction**, 2014, Creta. *16th International Conference on Human-Computer Interaction*, 2014. Creta, 2014.

# ANEXO A - EXEMPLOS DE ÁRVORES DE DECISÃO DA PMESP

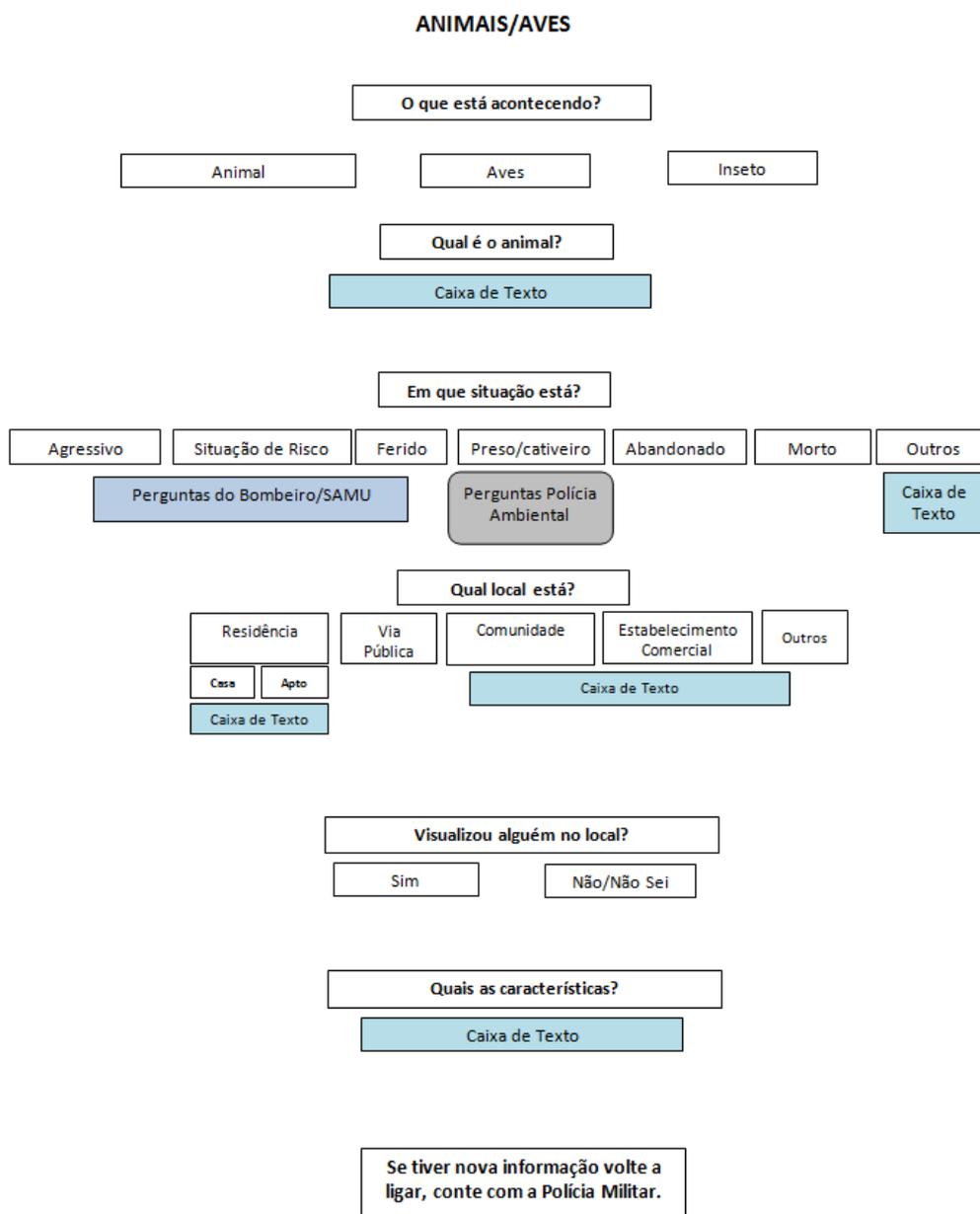


Figura A.1 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação envolvendo animais. Fonte: PMESP.

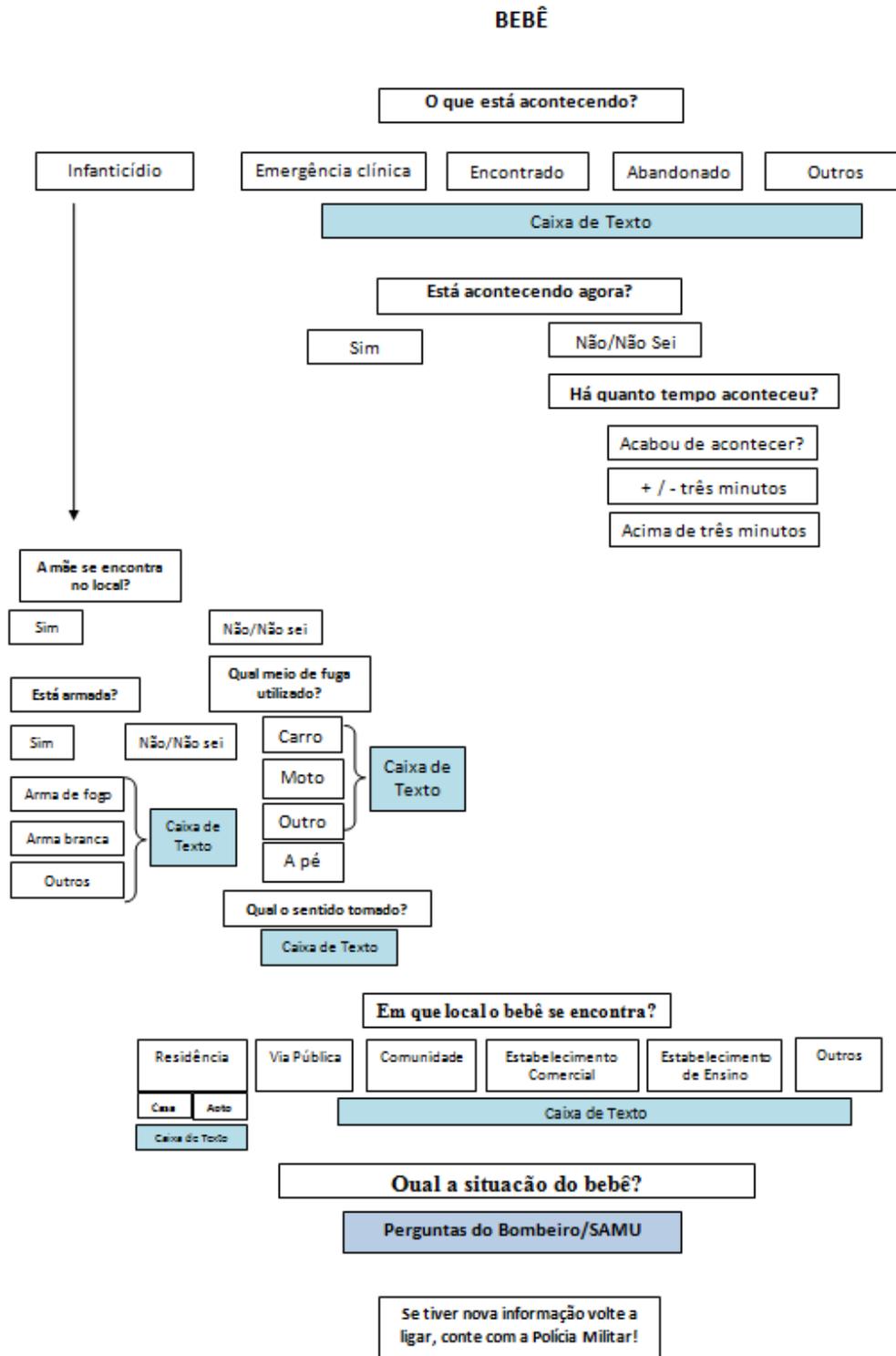


Figura A.2 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação envolvendo bebê. Fonte: PMESP.

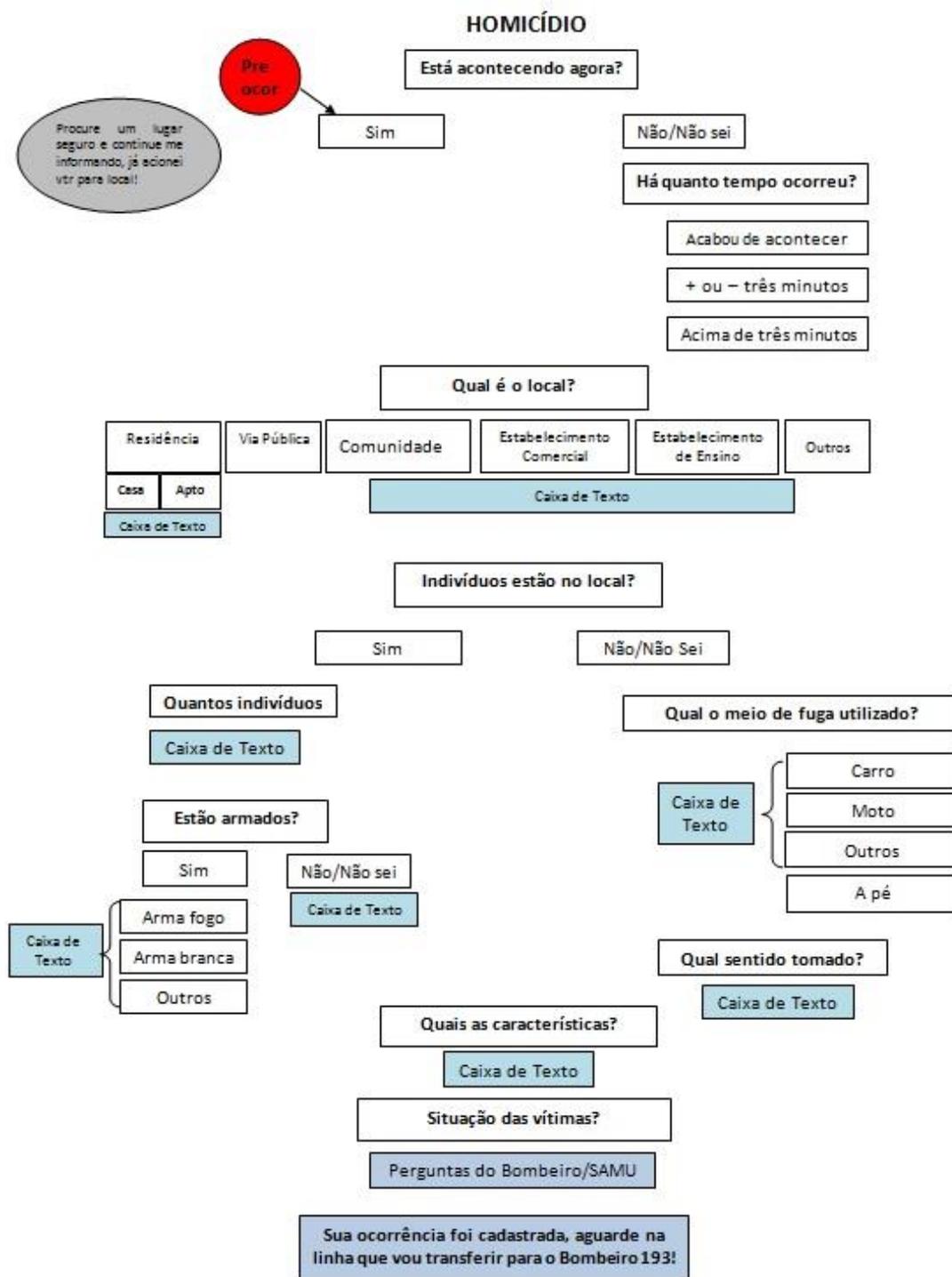


Figura A.3 - Exemplo de estrutura de árvore de decisão utilizada pela PMESP para orientar atendimentos emergenciais pelo serviço 190 para situação de homicídio. Fonte: PMESP.