Ontology Definition and Cognitive Analysis in Occupational Health and Security (OHS) Environments

Conference Paper · April 2015

DOI: 10.1145/2695664.2695891

4 authors:

Nayat Sánchez-Pi
National Institute for Research in Computer Science and Control

Luis Martí
National Institute for Research in Computer Science and Control

A. C. B. Garcia
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

José M. Molina
University Carlos III de Madrid

Ontology Definition and Cognitive Analysis in Ocupational Health and Security (OHS) Environments

Nayat Sanchez-Pi*
Instituto de Lógica, Filosofia e Teoria da Ciéncia (ILTC)
Niterói (RJ) Brazil
nayat@iltc.br

Luis Martí
Pontifícia Universidade Católica do Rio de Janeiro
Rio de Janeiro (RJ), Brazil
lmarti@ele.pu-rio.br

Jose Manuel Molina
Universidad Carlos III de Madrid
Madrid, Spain
molina@ia.uc3m.es

Ana Cristina Bicharra Garcia
ADDLabs, Universidade Federal Fluminense
Niterói (RJ) Brazil
bicharra@ic.uff.br

ABSTRACT

Events recognition is central to occupational health and safety OHS, since the system can selectively start proper prediction services according to the user current situation and past knowledge taken from huge databases. In this sense, a fusion framework that combines data from multiples sources to achieve more specific inferences is needed. Our proposed model provides the big picture about risk analysis for that employee at that place in that moment in a real world environment. Our main contribution lies in building a causality model for accidents investigation by means of a well-defined spatiotemporal constraints on offshore oil industry domain.

Categories and Subject Descriptors

I.2.m [**ARTIFICIAL INTELLIGENCE**]: Miscellaneous—*Data Fusion*; J.2 [**Computer Applications**]: PHYSICAL SCIENCES AND ENGINEERING

General Terms

Algorithms

Keywords

data fusion, ontology, oil industry

1. INTRODUCTION

OHS is a priority issue for the offshore oil and gas industry and a determining factor in its overall success. Years passed since community takes into account the implications of oil industry to Health, Safety and the Environment but nowadays industries invest a lot of efforts in accidents prevention. With the advances of communication technologies and the novelty researches in Ubiquitous Computing (UC) and Ambient Intelligence (AmI), is almost a fact to think of a Pervasive Offshore Oil Industry Environment.

In this scenario employees are surrounded of intelligent technology capable of not only interacting in an natural way, but also to intelligently reason on the accidents risk picture in order to alert the employees when an risky event is probable to occur in the place where is located. The process of constructing a dynamic risk picture for accident or incident detection and recognition involves contextual reasoning about past events, dynamic context (location, user, profile, etc), as well as relations between them with respect to particular goals,capabilities, and policies of the decision makers.

The system should reason about the situation. Situation assessment or situation awareness (SA) is a key component of any decision-making process [15]. A good definition of SA is found at [6, 32]: *"Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and a projection of their status in the near future."*

In an information fusion process, situation assessment represents a high-level inference level to identify the likely situations given the observed events and obtained data. High-level information fusion studies theories and methods to effectively combine data from multiple sensors and related information to achieve more specific inferences that could be achieved by using a single, independent sensor.

Traditional approaches using observational data and *a priori* models are insufficient to deal with real-world complex problems. In oil industry exists standards to identify and record workplace accidents and incidents in order to provide guiding means on prevention efforts, indicating specific failures or reference, means of correction of conditions or circumstances that culminated in accident.

An information fusion model can help to intelligently predict undesirable events like accidents by taking into account

*correspondant author

SAC'15 April 13-17, 2015, Salamanca, Spain.

http://dx.doi.org/10.1145/2695664.2695891

the event location and time, historical data of past events and user's profile. A solution should involve a machine learning approach to learn from past anomaly events and to predict accidental events in time and space while also exploiting the contextual knowledge available.

Past knowledge (historical data) can be analyzed to find interesting patterns. For this, data mining is the most essential part of the knowledge discovery process which combines databases, artificial intelligence, machine learning and statistics techniques. The basic techniques for data mining include: decision tree induction, rule induction, artificial neural network, clustering and association rules. Data mining can could be applied to any domain where large databases are saved. Some applications are: failure prediction [9], biomedical applications [14], process and quality control [11].

Association rule learning is a popular and well-researched set of methods for discovering interesting relations between entities in a large databases. It is intended to identify strong rules discovered in databases using different measures of interestingness. Many algorithms for generating association rules were presented over time.

In this work we describe an approach to deal with this problem involving high-level information fusion, ontologies and rule mining. Our proposed architecture provides the big picture about how to integrate past knowledge in risk analysis for that employee at that place in that moment in a real world environment, that we named cognitive analysis.

The paper is organized as follows. After providing an introduction toinformation fusion processes, section 3 focuses on architecture and domain mode. Section 5 depicts the reasoning process and mining information used.Finally, Section 6 presents some conclusive remarks and outlines the current and future work been carried out in this area.

2. FOUNDATIONS

Data fusion has been defined in [33] as "*a multi-level process dealing with the association, correlation, combination of data and information from single and multiple sources to achieve refined position, identify estimates and complete and timely assessments of situations, threats and their significance*". Data fusion (DF) and information fusion (IF) has been treated similarly in literature but when talking about data fusion is because it represents raw data, and when referring to information fusion, is because it implies a higher semantic level of fusion. The problem of information fusion has attracted significant attention in the artificial intelligence community, trying to innovate in the techniques used for combining the data and to refine state estimates and predictions.

Information Fusion can be classified depending on the level of abstraction [27]: low-level fusion, medium level fusion, high level fusion and multi-level fusion. In the process of low level fusion the raw data are directly provided as an input to the data fusion process. The medium level fusion is a feature level where features are fused to obtain other features that could be employed for other tasks. In the high level fusion a combination of symbolic representation is the entry of the fusion process. And in the multi-level fusion the entry comes from different levels of abstractions.

Others classifications are proposed: Dasarathy's Functional Model [13] or JDL conceptual model proposed by the American Department of Defense [33]. The JDL classification model consists into five processing levels in the transformation of input signals to decision-ready knowledge. These levels are: level 0 or source pre-processing; level 1 or object refinement; level 2 or situation assessment; level 3 or impact assessment and level 4 or process refinement.

High-level fusion starts at level 2. Situation assessment (SA) aims to identify the likely situations given the observed events and obtained data. It establishes relationships between the objects. Relations (i.e., proximity, communication) are valued to determine the significance of the entities or objects in a specific environment. The aim of this level includes performing high-level inferences and identifying significant activities and events (patterns in general). The output is a set of high-level inferences. Situation assessment is an important part of the information fusion process because it is the purpose for the use of IF to synthesize the multitude of information, it provides an interface between the user and the automation, and (3) focuses data collection and management.

Intensive research has been done in past years focused on low-level information fusion, nowadays the focus is currently shifting towards high-level information fusion [8]. Compared to the increasingly mature field of low-level IF, theoretical and practical challenges posed by high-level IF are more difficult to handle. Some of the applications that involve high-level fusion are: Defense [1, 2, 10, 16, 23], Computer and Information Security [12,17], Disaster Management [24–26,28], Fault Detection [3–5], Environment [20–22]. Also techniques for using contextual information in high-level information fusion architectures has been studied at [18].

In the context of oil and gas industry is increasingly concerned with achieving and demonstrating good performance of occupational health and safety (OHS), through the control of its OHS risks, which is consistent with its policy and objectives. In oil industry exist standards to identify and record workplace accidents and incidents to provide guiding means on prevention efforts, indicating specific failures or reference, means of correction of conditions or circumstances that culminated in accident. So, events recognition is central to OHS, since the system can selectively start proper prediction services according to the user current situation and past knowledge taken from huge databases. In this sense, a fusion framework that combines data from multiples sources to achieve more specific inferences is needed [30, 31].

3. A DATA FUSION ARCHITECTURE FOR OHS ENVIRONMENT

The architecture of our context-based fusion framework is depicted in Fig. 1. The context-aware system developed has a hierarchical architecture with the following layers: Services layer, Context Acquisition layer, Context Representation layer, Context Information Fusion layer and Infrastructure layer. The hierarchical architecture reflects the complex functionality of the system as shown in the following brief description of the functionality of particular layers:

- *Infrastructure Layer*: The lowest level (L0 in JDL Model) of the location management architecture is the Sensor Layer which represents the variety of physical and logical location sensor agents producing sensor-specific location information.

- *Context Acquisition*: The link between sensors (lowest layer) and the representation layer (L1 in JDL Model)

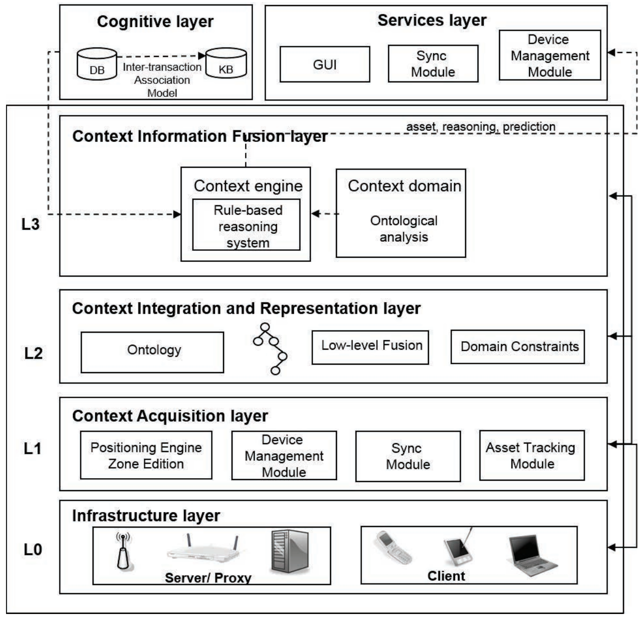

Figure 1: Information fusion framework architecture.

- *Context Representation*: This is where the situation is represented by means of an ontology (L2 in JDL Model)
- *Context Information Fusion Layer*: This is where the high-level information fusion occurs. It is here where reasoning about context and events of the past takes place (L3 in JDL Model)
- *Cognitive Layer*: This is where the intra e inter transaction association rules are extracted from the database.
- *Services Layer*: This layer interacts with the variety of users of the system (employees) and therefore needs to address several issues including access rights to location information (who can access the information and to what degree of accuracy), privacy of location information (how the location information can be used) and security of interactions between users and the system.

4. ONTOLOGY DEFINITION FOR OHS ENVIRONMENT

Normally, ontology represents a conceptualization of particular domains. In our case, we will use the ontology for representing the contextual information of the offshore oil industry environment.Ontologies are particularly suitable to project parts of the information describing and being used in our daily life onto a data structure usable by computers. Using ontologies provides an uniform way for specifying the model's core concepts as well as an arbitrary amount of subconcepts and facts, altogether enabling contextual knowledge

An ontology is defined as "an explicit specification of a conceptualization" [19]. An ontology created for a given domain includes a set of concepts as well as relationships connecting them within the domain. Collectively, the concepts and the relationships form a foundation for reasoning about the domain. A comprehensive, well-populated ontology with classes and relationships closely modeling a specific domain represents a vast compendium of knowledge in the domain. Furthermore, if the concepts in the ontology are organized into hierarchies of higher-level categories, it should be possible to identify the category (or a few categories) that best classify the context of the user. Within the area of computing, the ontological concepts are frequently regarded as classes that are organized into hierarchies. The classes define the types of attributes, or properties common to individual objects within the class. Moreover, classes are interconnected by relationships, indicating their semantic interdependence (relationships are also regarded as attributes). We built a domain ontology for the Health, Safety and Environment (HSE) (see Fig. 2 of oil and gas domain [29]). We also obtain the inferences that describe the dynamic side and finally we group the inferences sequentially to form tasks. Principal concepts of the ontology are the following:

- Anomaly: Undesirable event or situation which results or may result in damage or faults that affect people, the environment, equity (own or third party), the image of the Petrobras System, products or production processes. This concept includes accidents, illnesses, incidents, deviations and non-conformances.
  - Neglect: Any action or condition that has the potential to lead to, directly or indirectly, damage to people, to property (own or third party) or environmental impact, which is inconsistent with labor standards, procedures, legal or regulatory requirements, requirements management system or practice.
    * Behavioral neglect: Act or omission which, contrary provision of security, may cause or contribute to the occurrence of accidents.
    * Non-behavioral neglect: Environmental condition that can cause an accident or contribute to its occurrence. The environment includes adjective here, everything that relates to the environment, from the atmosphere of the workplace to the facilities, equipment, materials used and methods of working employees who is inconsistent with labor standards, procedures, legal requirements or normative requirements of the management system or practice.
  - Incident: Any evidence, personal occurrence or condition that relates to the environment and / or working conditions, can lead to damage to physical and / or mental.
  - Accident: Occurrence of unexpected and unwelcome, instant or otherwise, related to the exercise of the job, which results or may result in personal injury. The accident includes both events that may be identified in relation to a particular time or occurrences as continuous or intermittent exposure, which can only be identified in terms of time period probable. A personal injury includes both traumatic injuries and illnesses, as damaging effects mental, neurological or systemic, resulting from exposures or circumstances prevailing at the year's work force. In the period for meal or rest, or upon satisfaction of other physiological needs

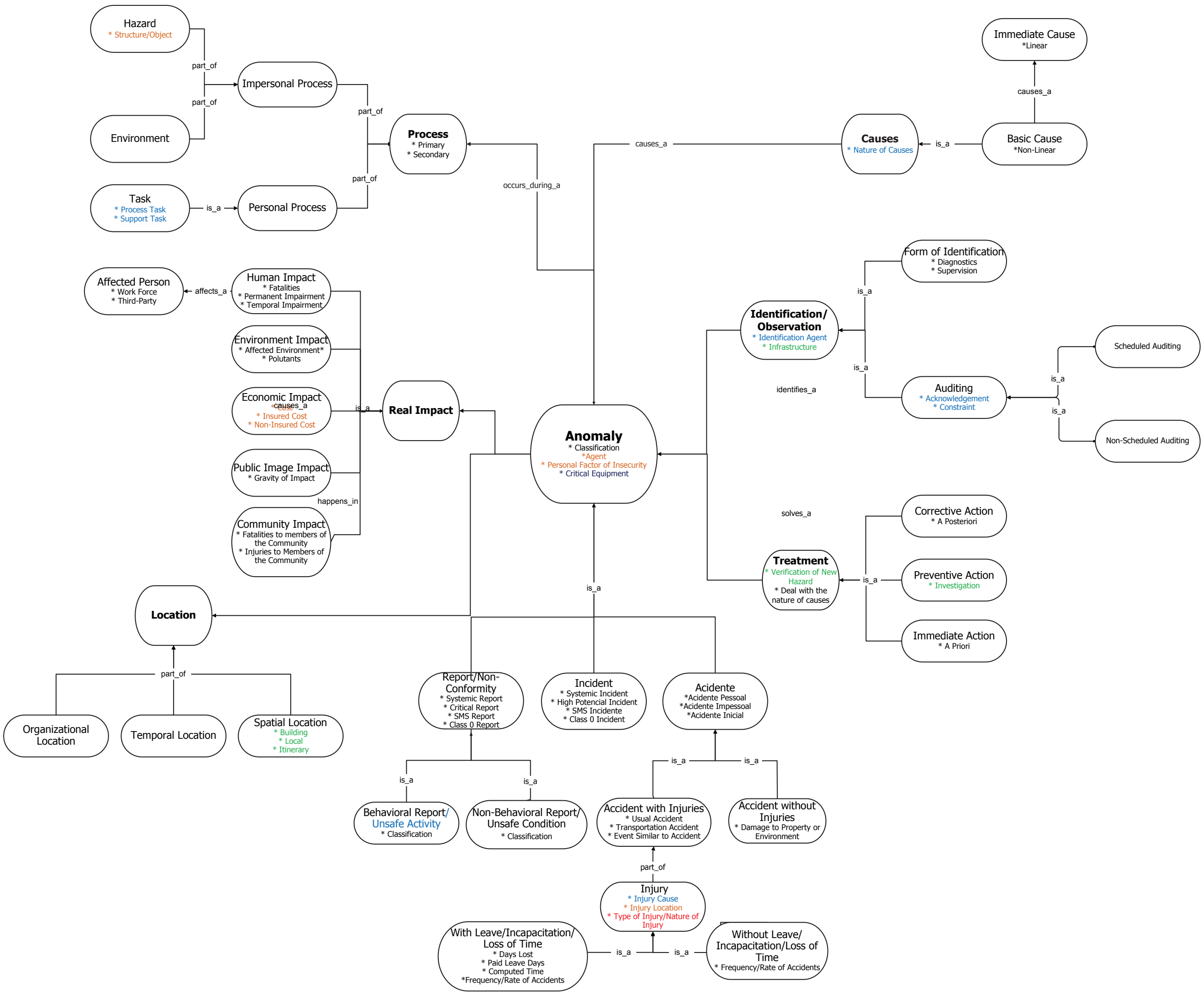

Figure 2: Occupational Health and Security (OHS) ontology.

at the workplace or during this, the employee is considered in carrying out the work.

* Accident with injury: It's all an accident in which the employee suffers some kind of injury. Injury: Any damage suffered by a part of the human organism as a consequence of an accident at work.
  * With leave: Personal injury that prevents the injured from returning to work the day after the accident or resulting in permanent disability. This injury can cause total permanent disability, permanent partial disability, total temporary disability or death.
  * Without leave: Personal injury that does not prevent the injured to return to work the day after the accident, since there is no permanent disability. This injury, not resulting in death, permanent total or partial disability or total temporary disability, requires, however, first aid or emergency medical aid. Expressions should be avoided "lost-time accident" and "accident without leave", used improperly to mean, respectively, "with leave injury" and "injury without leave."
* Accident Without Injury: Accident causes no personal injury.

5. COGNITIVE ANALYSIS FOR OHS ENVIROMENT

Standard ontology reasoning procedures can be performed within the ontologies to infer additional knowledge from the explicitly asserted facts. By using an inference engine, tasks such as classification or instance checking can be performed. Figure 3.

Risk prevention is a paradigmatic case of inductive reasoning. Inductive reasoning begins with observations that

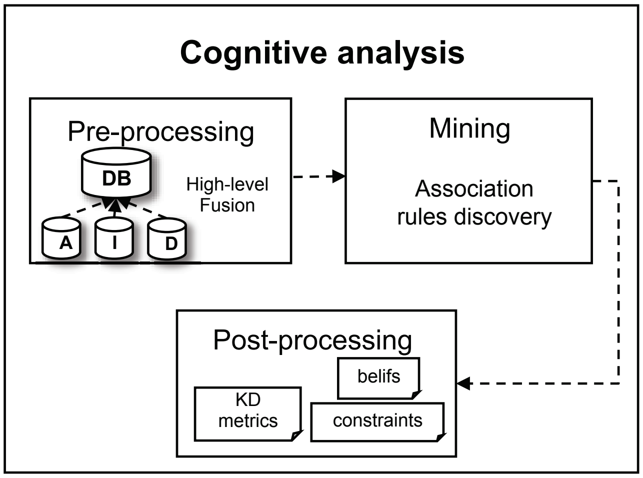

Figure 3: Cognitive analysis.

are specific and limited in scope, and proceeds to a generalized conclusion that is likely, but not certain, in light of accumulated evidence. You could say that inductive reasoning moves from the specific to the general. Much scientific research is carried out by the inductive method: gathering evidence, seeking patterns, and forming a hypothesis or theory to explain what is seen.

In our framework, inductive rules formally represent contextual, heuristic and common sense knowledge to accomplish high-level scene interpretation and low-level location refinement.

Once an employee enters the network, it immediately connects with a local proxy, which evaluates the position of the client device and assign a role to the employee. A preprocessing step begins then filtering the relevant features that are selected to participate in the process of knowledge discovery by type of employee (role). The association rules mining process starts with the selected configuration and the set of resulting rules can be analyzed. Later a postprocessing step starts. It is an important component of KDD consisting of many various procedures and methods for pruning and filtering the resulting rules.

The fusion engine implements an association rules model that combines dynamically feature selection based on the role of the user in order to find spatiotemporal patterns between different types of anomalies (or event sequence, ex. neglects, incidents, accidents) that match with the current location of the user.

Two categories of association mining are employed: intra-anomaly and inter-anomaly [7]. Intra-anomaly associations are the associating among items within the same type of anomaly, where the notion of the transaction could be events where the same user participate. However, inter-anomaly describes relationships among different transactions. That means between incidents, accidents and neglects.

In this case the data pre-processing before mining is pretty straightforward, as the interest is to discover relationships between the values of different attributes and the possible presence of probabilistic implication rules between them. The results of applying the rule mining algorithms are post-processed to eliminate cyclic rules and to sort them according to an interestingness criterion.

6. FINAL REMARKS

In this work we have presented an information fusion framework for providing context-aware services related to risk prevention in offshore oil industry environment. The proposal put forward aims at providing context-based information related to accidents and their causes to users depending on their profiles and location.

Our approach relies on a domain ontology to capture the relevant concepts of the application and the semantics of the context in order to create a high-level fusion of information. Along with that we have introduced an innovative use of rule mining for provisioning knowledge for situation assessment and decision making regarding risk an accidents prevention. This form of rule mining is capable of an online high-level knowledge extraction that represents relations between different kinds of anomalies that have taken place at the user location and that the system has determined that had lead to an accident.

Acknowledgements

Authors want to acknowledge the support of CNPq BJT Project 407851/2012-7 and CNPq PVE Project 314017/2013-5 and projects MINECO TEC2012-37832-C02-01, CICYT TEC 2011-28626-C02-02.

7. REFERENCES

[1] S. Ahlberg, P. Hörling, K. Johansson, K. Jöred, H. Kjellström, C. Mårtenson, G. Neider, J. Schubert, P. Svenson, P. Svensson, et al. An information fusion demonstrator for tactical intelligence processing in network-based defense. *Information Fusion*, 8(1):84–107, 2007.

[2] T. Aldinger and J. Kao. Data fusion and theater undersea warfare-an oceanographer's perspective. In *OCEANS'04. MTTS/IEEE TECHNO-OCEAN'04*, volume 4, pages 2008–2012. IEEE, 2004.

[3] A. Bashi. *Fault Detection for Systems with Multiple Unknown Modes and Similar Units*. PhD thesis, University of New Orleans, 2010.

[4] A. Bashi, V. P. Jilkov, and X. R. Li. Fault detection for systems with multiple unknown modes and similar units-part i. In *12th International Conference on Information Fusion (FUSION'09)*, pages 732–739. IEEE, 2009.

[5] O. Basir and X. Yuan. Engine fault diagnosis based on multi-sensor information fusion using dempster-shafer evidence theory. *Information Fusion*, 8(4):379–386, 2007.

[6] G. Bedny and D. Meister. Theory of activity and situation awareness. *International Journal of Cognitive Ergonomics*, 3(1):63–72, 1999.

[7] C. Berberidis, L. Angelis, and I. Vlahavas. Inter-transaction association rules mining for rare events prediction. In *Proc. 3rd Hellenic Conference on Artificial Intellligence*, 2004.

[8] E. Blasch, J. Llinas, D. Lambert, P. Valin, S. Das, C. Chong, M. Kokar, and E. Shahbazian. High level information fusion developments, issues, and grand challenges: Fusion 2010 panel discussion. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[9] M. L. Borrajo, B. Baruque, E. Corchado, J. Bajo, and J. M. Corchado. Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises. *International Journal of Neural Systems*, 21(04):277–296, 2011.

[10] C.-Y. Chong, M. Liggins, et al. Fusion technologies for drug interdiction. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI'94)*, pages 435–441. IEEE, 1994.

[11] M. Conti, R. D. Pietro, L. V. Mancini, and A. Mei. Distributed data source verification in wireless sensor networks. *Information Fusion*, 10(4):342–353, 2009.

[12] I. Corona, G. Giacinto, C. Mazzariello, F. Roli, and C. Sansone. Information fusion for computer security: State of the art and open issues. *Information Fusion*, 10(4):274–284, 2009.

[13] B. V. Dasarathy. Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, 85(1):24–38, 1997.

[14] J. F. De Paz, J. Bajo, V. F. López, and J. M. Corchado. Biomedic organizations: An intelligent dynamic architecture for KDD. *Information Sciences*, 224:49–61, 2013.

[15] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64, 1995.

[16] A. Gad and M. Farooq. Data fusion architecture for maritime surveillance. In *Proceedings of the Fifth International Conference on Information Fusion (FUSION'02)*, volume 1, pages 448–455. IEEE, 2002.

[17] G. Giacinto, F. Roli, and C. Sansone. Information fusion in computer security. *Information Fusion*, 10(4):272–273, 2009.

[18] J. Gómez-Romero, J. Garcia, M. Kandefer, J. Llinas, J. Molina, M. Patricio, M. Prentice, and S. Shapiro. Strategies and techniques for use and exploitation of contextual information in high-level fusion architectures. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[19] J. Gómez-Romero, M. A. Patricio, J. García, and J. M. Molina. Ontological representation of context knowledge for visual data fusion. In *12th International Conference on Information Fusion (FUSION'09)*, pages 2136–2143. IEEE, 2009.

[20] U. Heiden, K. Segl, S. Roessner, and H. Kaufmann. Ecological evaluation of urban biotope types using airborne hyperspectral hymap data. In *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pages 18–22. IEEE, 2003.

[21] L. Hubert-Moy, S. Corgne, G. Mercier, and B. Solaiman. Land use and land cover change prediction with the theory of evidence: a case study in an intensive agricultural region of france. In *Proceedings of the Fifth International Conference on Information Fusion (FUSION'02)*, volume 1, pages 114–121. IEEE, 2002.

[22] A. Khalil, M. K. Gill, and M. McKee. New applications for information fusion and soil moisture forecasting. In *8th International Conference on Information Fusion (FUSION'05)*, volume 2, page 7. IEEE, 2005.

[23] M. E. Liggins, A. Bramson, et al. Off-board augmented fusion for improved target detection and track. In *1993 Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 295–299. IEEE, 1993.

[24] E. G. Little and G. L. Rogova. Ontology meta-model for building a situational picture of catastrophic events. In *8th International Conference on Information Fusion (FUSION'05)*, volume 1, pages 1–8. IEEE, 2005.

[25] J. Llinas. Information fusion for natural and man-made disasters. In *Proceedings of the Fifth International Conference on Information Fusion (FUSION'02)*, volume 1, pages 570–576. IEEE, 2002.

[26] J. Llinas, M. Moskal, and T. McMahon. Information fusion for nuclear, chemical, biological & radiological (NCBR) battle management support/disaster response management support. Technical report, Center for MultiSource Information Fusion, School of Engineering and Applied Sciences, University of Buffalo, USA, 2002.

[27] R. C. Luo, Y. C. Chou, and O. Chen. Multisensor fusion and integration: algorithms, applications, and future research directions. In *Mechatronics and Automation, 2007. ICMA 2007. International Conference on*, pages 1986–1991. IEEE, 2007.

[28] J. Mattioli, N. Museux, M. Hemaissia, and C. Laudy. A crisis response situation model. In *10th International Conference on Information Fusion (FUSION'07)*, pages 1–7. IEEE, 2007.

[29] N. Sanchez-Pi, L. Martí, and A. C. Bicharra Garcia. Text classification techniques in oil industry applications. In A. Herrero, B. Baruque, F. Klett, A. Abraham, V. Snášel, A. C. Carvalho, P. García Bringas, I. Zelinka, H. Quintián, and E. Corchado, editors, *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, volume 239 of *Advances in Intelligent Systems and Computing*, pages 211–220. Springer International Publishing, 2014.

[30] N. Sanchez-Pi, L. Martí, J. M. Molina, and A. C. B. Garcia. High-level information fusion for risk and accidents prevention in pervasive oil industry environments. In *Highlights of Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection*, pages 202–213. Springer, 2014.

[31] N. Sanchez-Pi, L. Martí, J. M. Molina, and A. C. B. Garcia. An information fusion framework for context-based accidents prevention. In *Information Fusion (FUSION), 2014 Proceedings of the 17th International Conference on*, pages 1–8. IEEE, 2014.

[32] K. Smith and P. Hancock. Situation awareness is adaptive, externally directed consciousness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):137–148, 1995.

[33] F. E. White. Data fusion lexicon. Technical report, DTIC Document, 1991.