12th International Conference on Information Fusion
Seattle, WA, USA, July 6-9, 2009

Belief Modeling for Maritime Surveillance

Aaron Hunter
School of Computing Science
Simon Fraser University
Burnaby, BC, Canada
hunter@cs.sfu.ca

Abstract – *In maritime surveillance, the volume of information to be processed is very large and there is a great deal of uncertainty about the data. There are many vessels at sea at every point in time, and the vast majority of them pose no threat to security. Sifting through all of the benign activity to find unusual activities is a difficult problem. The problem is made even more difficult by the fact that the available data about vessel activities is both incomplete and inconsistent. In order to manage this uncertainty, automated anomaly detection software can be very useful in the early detection of threats to security. This paper introduces a high-level architecture for an anomaly detection system based on a formal model of beliefs with respect to each entity in some domain of interest. In this framework, the system has beliefs about the intentions of each vessel in the maritime domain. If the vessel behaves in an unexpected manner, these intentions are revised and a human operations centre worker is notified. This approach is flexible, scalable, and easily manages inconsistent information. Moreover, the approach has the pragmatic advantage that it uses expert information to inform decision making, but the required information is easily obtained through simple ranking exercises.*

Keywords: Anomaly detection, maritime surveillance, belief revision.

1 Introduction

Global maritime surveillance involves the monitoring of several hundred vessels in many cases. Using existing sensors, it is possible to monitor an individual ship very effectively if we are aware that it may pose a threat to security. However, due to the volume of information with which we are faced, it is often difficult for a human observer to determine which ships should be subjected to detailed monitoring. As such, the study of automated anomaly detection systems has emerged as an important topic in maritime surveillance. In this paper, we present an approach to anomaly detection based on a formal model of belief change that has been developed in the Artificial Intelligence community. We present our approach as a high-level architecture, built to complement an existing rule-based expert system for anomaly detection [13].

This paper makes two main contributions to existing literature. First, we illustrate how an existing approach to designing anomaly detection software can be improved to manage inconsistency. It is well known that anomaly detection is a significant problem in maritime surveillance. The second contribution of this work is the introduction of a practical application of formal belief change operators. This contribution is significant in as much as the applications of belief change operators are currently lagging behind the theoretical foundations.

2 Background

2.1 Anomaly Detection in the Maritime Domain

Monitoring maritime activities is clearly a critical concern for security. The kinds of activities that we would like to detect are numerous, ranging from relatively minor threats such as illegal fishing to more serious terrorist attacks. In the past, it was difficult to perform automated anomaly detection because maritime data was too heterogenous. A military operations centre typically has access to many sources of information, including self reports, satellite imagery, direct observation, and radar. Before this information can be analyzed effectively, there is a difficult *information fusion* problem that must be addressed. In particular, the information must be used to identify *tracks* representing individual vessels, along with all available data about the vessels. This information fusion problem has been addressed to varying degrees by different military surveillance groups around the world. In this paper, we will make the assumption that we are dealing with well-defined tracks that have been output by a suitable fusion.

Given a set of tracks, several different approaches to anomaly detection have been studied, including machine learning approaches [6] and rule-based expert systems [13]. We suggest that both of these approaches are incomplete, for different reasons. Approaches developed based on machine learning do an excellent job of detecting unusual behavior, but most military commanders are still not comfortable depending entirely on software for security-critical problems. For this reason, there is still a preference for military applications to be developed as *mixed-initiative* tools, where a human expert is in the decision making loop. As such, al-

978-0-9824438-0-4 ©2009 ISIF

though machine learning can help detect unusual behavior in the maritime domain, it is still useful to supplement such approaches with expert input.

The development of a rule-based expert system for anomaly detection can be valuable, as it incorporates expert knowledge in the detection of anomalies. However, the problem with such a system is that it only incorporates the rules an expert uses to draw *new conclusions*. However, in many cases these new conclusions will conflict with existing facts. Therefore, a proper implementation of an expert system for anomaly detection should incorporate some mechanism for resolving these conflicts. One of the main goals of this paper is to illustrate one natural methodology for resolving this problem.

2.2 Belief Revision

Belief revision refers to the process that an agent uses to incorporate new information about the state of the world. In this section, we briefly outline the AGM approach to belief revision [1]. This approach has been highly influential in the Artificial Intelligence literature, and it has lead to a large body of work on related approaches.

Let **F** be a set of boolean variables that describe various features of the world. Informally, the set **F** is understood to represent all aspects of the world that are relevant for a given domain. Hence, an assignment of values to the variables in **F** represents a complete description of a particular configuration of the world. A *proposition* is either an element of **F**, or a combination of propositions using standard connectives: $\neg$ (not), $\wedge$ (and), $\vee$ (or). A *state* is a function s that assigns the value $true$ or $false$ to each variable in **F**. The notion of *truth* in a state is defined for propositions using standard propositional truth tables; hence $\phi_1 \wedge \phi_2$ is true in s just in case both ϕ_1 and ϕ_2 are true in s.

Typically, an agent will not have complete knowledge about the state of the world, so the values of some variables will be unknown. A *belief set* is a consistent set of propositions, representing all of the facts that an agent believes to be true. In this paper, we will only be concerned with finite belief sets, which can actually be represented as a single proposition by taking a conjunction. The fundamental problem in belief revision, is the following. How can a new fact ϕ be incorporate into a pre-existing belief set Φ? Clearly, if ϕ is consistent with Φ, then this is straightforward: we simply add the proposition to the pre-existing beliefs. This is known as *belief expansion*. However, if ϕ is not consistent with Φ, then the problem is more difficult. Our goal is to produce a new belief set Φ' that incorporates ϕ, while keeping as much of Φ as consistently possible.

An AGM belief revision operator is a function $*$ that maps a belief set and a proposition to a new belief set, while satisfying a set of logical postulates. For example, one of the AGM postulates states that ϕ must be contained in the revised belief set $\Phi * \phi$. The complete list of postulates is given in [1]; we will not produce the list here, since the details are not relevant for the present work. Instead, we give a semantic characterization of the postulates due to Grove [8]. It can be demonstrated that every revision operator depends on a total pre-order of states associated with each belief set. Hence, in order to compute $\Phi * \phi$, we first need to define an ordering $<$ over states where the minimal states are precisely those where the propositions in Φ are true. The revision operator then computes $*$ by taking the $<$-minimal states that also satisfy ϕ.

Example Suppose that an intelligence agency is tracking the location of a target individual, by monitoring the country where he or she is located. Let **F** be the set of fluents *inCOUNTRY* where *COUNTRY* ranges over all countries in the world. Suppose that it is believed that the individual is currently located in North America. This can be represented by the belief set:

$$\Phi = \{inCANADA \vee inUSA \vee inMEXICO\}.$$

Now suppose we are told that the individual is not in any of these countries. Hence, we want to compute $\Phi * \phi$, where ϕ is the following proposition:

$$\phi = \neg(inCANADA \vee inUSA \vee inMEXICO).$$

In order to define the $*$ operator, we first need to determine a plausible ordering on states. In this case, geographic proximity is a reasonable candidate. Hence, the most plausible alternative locations are those that are geographically closest to the countries in North America. Hence, we may plausibly define $\Phi * \phi$ to be the following belief set:

$$\{inGREENLAND \vee inRUSSIA \vee inGUATAMALA\}.$$

While the preceding example is very simplistic, it illustrates the kind of reasoning that can reasonably be carried out by an automated belief revision system. The only human-level involvement required is the specification of an appropriate ordering over alternative states. However, this can be done offline.

3 Belief-Based Anomaly Detection

There are two kinds of maritime anomalies that concern operations centre workers. The first kind of anomaly is an anomaly that occurs regularly, but has minor consequences. Examples of this kind of anomaly include illegal fishing anomalies and illegal immigration anomalies. While both of these activities are important to monitor, the consequences of missing them are relatively insignificant. For this reason, anomalies of this sort can be detected in real time as they occur. The second kind of anomaly is an anomaly that occurs very rarely, but has major consequences if it goes undetected. Examples of this kind of anomaly include terrorist attacks and vessel hijacking.

The problem with anomaly detection systems based purely on machine learning algorithms is that we may not have sufficient training data to detect the second kind of anomaly. Certain kinds of terrorist attack have simply never occurred previously, but we would like to detect them the first time that they occur. There is no guarantee that training a machine learning algorithm based on historical data will converge to detect such attacks, and the risk is too great to simply trust such an approach. For this reason, we propose that automated anomaly detection systems need to incorporate expert knowledge in the detection of anomalies. In fact, we propose that anomaly detection systems should, to the greatest extend possible, simply mimic the reasoning of an operations centre worker. The problem with manual anomaly detection is that workers can not look at every track. An automated anomaly detection system should act as a reasoning prosthetic for military experts, by applying expert knowledge in the analysis of each track.

Once we take this perspective on anomaly detection, it becomes clear that a simple rule-based approach is not sufficient. Expert reason using more sophisticated forms of knowledge than simple if-then rules. In this section, we set out a basic model of expert reasoning about maritime anomalies that is more sophisticated than a basic rule-based approach. However, we suggest that the reasoning is still simple enough to be tractable.

3.1 High-Level Architecture

A high-level architecture for a the rule-based anomaly detection system is provided in Figure 1. This figure is essentially based on the system described in [13], although the architecture is generic enough to describe any standard production system for anomaly detection.

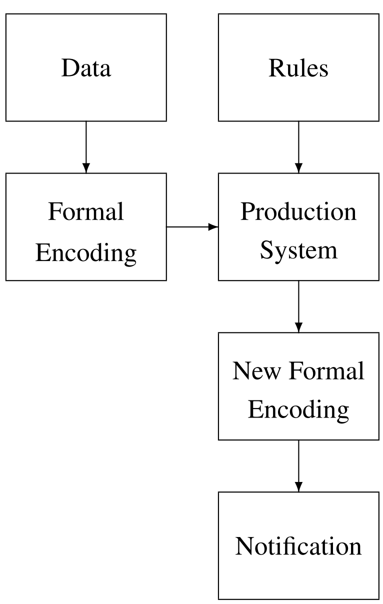

Figure 1: Rule-Based Anomaly Detection

The system has two main inputs: data and rules. The data input consists of a variety of reports, including 24-hour ship reports, 96-hour ship reports, and AIS reports. The rules have been defined by military experts to describe situations that a military operations centre worker would consider unusual. The system proceeds by first translating the data into a suitable format, then using a production system to produce a new formal description of the world which is visualized for an operations centre worker.

The problem with this architecture is that it does not involve any representation of inconsistency or conflict. This problem is obscured in [13] because the system appears to mainly involve rules of the form:

$$\textit{If } CONDITION \textit{ then } ANOMALY.$$

This kind of rule is only able to conclude that a certain kind of anomaly has occurred, it does not postulate new facts. This kind of system detects the anomaly *when it occurs*. Instead, we need to detect the anomaly *before it occurs*.

This is a very conservative model of the reasoning performed by an operations centre worker. In realistic situations, an operations centre worker would follow a procedure more like the following:

1. Examine data from sensors and reports.
2. Project future activities.
3. Apply expert knowledge to see if anything strange is happening.
4. If something strange is happening, modify projected future activities.

In terms of an automated system, we can abstract these steps into the following basic steps:

1. Get data.
2. Input data to projection engine.
3. Apply rules.
4. Revise projected beliefs.

In order to model this process, we propose the extended architecture for anomaly detection presented in Figure 2. In this architecture, we make it explicit that the incoming data from reports and sensors gives an incomplete picture of the world. Therefore, the encoding of the world is actually a set of possible configurations of the world, which we call a *belief set*. In this framework, the production system outputs a collection of new facts that may or may not be consistent with our current beliefs. Incorporating these new facts requires the system to revise the previous beliefs, using a new belief revision subsystem. The outcome of belief revision is used to produce a new belief set, which is then presented to the operator using a suitable vizualization.

From a practical perspective, one important feature of this architecture is that it makes it clear that we need a more sophisticated knowledge elicitation process. In a standard rule-based systems, we need to interview military experts and abstract a set of rules that capture the knowledge obtained from the interviews. This is a relatively straightforward process, because if-then rules are relatively easy to understand from a non-techincal perspective. In our proposed architecture, we need three different forms of expert knowledge. First we need sufficient information to define a projection engine. Second, we need if-then rules specifying anomalies. Third, we need enough information to define a revision engine. In the following sections, we discuss each of these inputs as well as the difficulty of encoding the relevant expert knowledge.

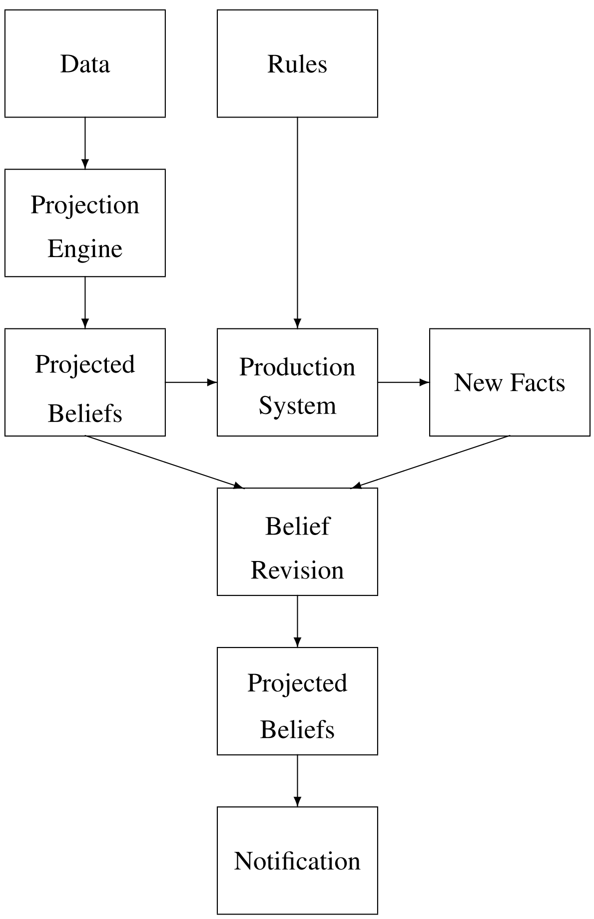

Figure 2: Belief-Based Anomaly Detection

3.2 Deriving Projections

Following information fusion, the data available at a maritime operations centre includes a vessel type, along with a path that the vessel has been following. Moreover, there are existing ontologies and databases of all possible types of vessels that are relevant for developing maritime domain awareness [7]. As such, we need a system that takes a vessel path and a vessel type as input, then outputs a projected future path.

Projections can be derived by humans, or also by machine learning algorithms. In one recent study, it was estimated that approximately 80% of observable vessels behave in a manner than can be predicted by a human operations centre worker [6]. There are several reasons for this fact. For example, for commercial vessels there are pragmatic reasons for following a predictable path. Specifically, commerical vessels tend to follow a “great circle route” between two points, as this is known to be the shortest path. Although some deviation may occur due to bad weather, the general pattern is relatively stable. Similarly, fishing vessels exhibit consistently predictable behavior.

Although human observers can project the future behaviour of a vessel with a high degree accuracy, it is also possible to automate this process using machine learning algorithms. This is the approach taken by DARPA in the development of the Predictive Analysis for Naval Deployment Activities program (PANDA) [6]. We suggest that this approach is preferable over human projections for the simple reason that it is difficult for humans to precisely describe routes in ordinary language. This is a common problem in the development of decision support systems for military use; acquiring expert knowledge in a suitable formal language can be difficult [10]. For this reason, the projection engine in Figure 2 can be defined using learning algorithms that have learned the most likely path for any given vessel, given sufficient historical data.

3.3 If-Then Rules

Anomaly detection to support the analysis of potential threats requires the explicit specification of anomalies that are important to military users. In a sense, one could simply use a machine learning algorithm to determine “normal” behaviour, and then flag all “not-normal” behaviour as an anomaly. There is, however, a problem with this approach. It is possible for a learning algorithm to be trained over time to accept undesirable behaviour as normal behaviour.

Example Suppose that every fishing vessel from a given country follows the same pattern of behaviour. First the vessels go to a designated, legal fishing area and fish for three days. Upon completion of three days of fishing, the vessels return home and continue fishing the entire way. Since the path home includes controlled fishing zones, this activity is illegal.

The preceding example is a plausible case where an ”anomalous” fishing activity may not be flagged by a learning algorithm. The reason it will not be flagged is because the behaviour happens every time the vessels return home. As such, this behaviour may come to be accepted as normal behaviour that is not worth flagging. Clearly this is a problem, as we would like to detect this activity.

In order to flag this kind of activity, it is preferable to incorporate some expert defined rules in the system that explicitly specify anomalies that should be detected. In this case, one might check the speed at which ships return home to see if it is consistent with fishing. Eliciting expert knowledge in the form of rules is a topic that has been studied extensively using established techniques, such as questionnaires [2], [11].

3.4 Revising Expert Beliefs

The final reasoning component of our high-level system for anomaly detection is related to the revision of projected beliefs. In our framework, this process will occur frequently because projected behaviours of vessels will not be perfect. As such, it will often be the case that new observed data will conflict with the projections. This is exactly the context in which formal belief revision operators are appropriate: the initial projections are understood to be error-prone, whereas the observed data is understood to be very reliable.

As stated in the introduction, belief revision requires a total pre-order over all possible states of the world. Eliciting expert knowledge that allows us to define such an ordering is not a straightforward process. We need to present military experts with two initial pieces of information:

1. An initial believed state of the world.
2. A new piece of data that is not consistent with 1.

Given these pieces of information, we need to ask the military expert to *rank* all possible alternative states of the world that explain the new piece of data. There is an established method to eliciting this kind of ordering, based on card sorting [5]. Hence, it is possible to elicit expert knowledge given a particular initial believed state. Carrying out this exercise for every possible configuration of the world would be possible given enough time with a military expert.

4 Prototype Formalism

4.1 Entities and Relations

In order to demonstrate our approach, we introduce a simplified prototype formalism. We assume a multi-sorted location vocabulary involving the following kinds of entities:

- L: A set of locations, represented by integer coordinate pairs.
- T: A set of discrete time points, including a distinguished *current time* t_0.
- V: A set of vessels.
- C: A set of vessel classes.

We assume further the following predicates:

- $isAt(v, l, t)$: Indicates that the vessel v is at the location l at time t.
- $isType(v, c)$: Indicates that the vessel v is in class c.

Assume that the set of locations and time points are both finite, so that we can represent all formulas as ground propositional formulas. This is a somewhat artificial restriction, but it allows us to use propositional belief revision operators.

Assume $V = \{ship_n \mid n < 1000\}$, and assume that $C = \{merchant, fishing, ferry, terrorist\}$. For each vessel $v \in V$, assume that there is exactly one true formula $isType(v, c)$ and one true formula $isAt(v, l, t)$. For ferries and fishing vessels, it is possible in practice to specify the set of coordinates where they can be expected to travel. Hence, we can define two new classes of formulas:

- $fishing(l)$
- $ferry(l)$

The restriction of vessels to these zones can be represented by axioms of the form:

$$isType(v, ferry) \Rightarrow (isAt(v, l, t) \Rightarrow ferry(l).$$

We have an parallel axiom for fishing vessels.

Merchant vessels are not typically restricted to specific zones, but they do travel predictable routes. Hence, we can introduce merchant restriction axioms of the form:

$$isType(v, merchant) \Rightarrow (isAt(v, l, t) \Rightarrow isAt(v, l', t').$$

Axioms of this form describe the restricted paths taken by merchant ships.

4.2 Information Update

A *state* is a set of formulas that assigns every vessel a class and a location at the current point in time. Vessels can also be assigned locations at future points in time: these are understood to be projections from the current point. A *belief state* is a set of states, which represents the set of all states that are currently possible. In this manner, uncertainty about the current state can be represented.

An *information update* is a set of formulas assigning a location or vessel type to some subset of the vessels in V. The main computational task that is faced in this simple formalism is how to incorporate the new information. As suggested previously, we propose that an AGM revision operator $*$ should be used. Hence, if α is he initial belief state and I is the information update, then the new belief state should be $\alpha * I$. The question is then: how should the operator $*$ be defined?

Rather than formally defining $*$, we simply describe a heuristic that could be used. In general, vessel locations are given through automated systems that track ships. These systems are often voluntary, but the fact remains that spoofing the ship location can be difficult. By contrast, information about the ship class is typically obtained from ship reports that are offered by the ship operator themselves. This is very easy to spoof. As a result, $*$ is defined so that formulas about the ship class are viewed as less reliable than formulas about ship location. Hence, if an information update indicates that a fishing ship is going well outside a fishing zone, then the ship type is revised. Depending on the content, it might be revised to a merchant, ferry, or terrorist vessel. This determination depends on the prevalence of each kind of ship in a particular context. The important point is that it is easy to define an appropriate AGM operator $*$ in any case.

5 Discussion

5.1 Advantages of the Belief Model

We have proposed an approach to anomaly detection that combines an expert system with a formal model of belief. In order to justify this approach, we must justify both components. There are basically two main justifications for using an expert system. First, as compared with a machine learning approach, an expert system does not require a large volume of training data. This is not an issue for military developers with access to large volumes of classified surveillance data. However, for academic explorations, the lack of data can cause practical difficulties. The second advantage of an expert system is the fact that the operation of the system is transparent to military users. This is an important pragmatic consideration if we hope to convince military users to use anomaly detection software in safety critical situations. Additional advantages of expert systems for anomaly detection are presented in [13].

For the purposes of the present paper, we are primarily interested in demonstrating the advantages of the proposed belief model. As stated previously, introducing belief change operators allows an automated anomaly detection system to use projections more intelligently. Specifically, it allows us to use projections to reclassify vessels and vessel behaviour *before* problems occur. This is a valuable feature in maritime surveillance. The main reason that a belief model can be useful is because ship behaviour is so predictable based on the declared class or activity being performed. Given that experts can identify specific patterns for specific ship types, it seems worthwhile to use this information for projections. However, projections are inherently fallible; therefore some mechanism for resolving erroneous beliefs is required.

Using the AGM model of belief change has several advantages. First of all, an expert system is inherently defined in terms of if-then logic. As such, it is natural to define belief change in the same kind of logical framework. However, the fact that is seems “natural” does formally not justify the use of a logical approach over, say, a quantitative probabilistic approach; but there is a pragmatic justification. In particular, defining a belief revision operator just requires an ordering over the plausibility of different states or an ordering over the reliability of different information sources. For an academic researcher, it is difficult to define such an ordering in a reliable manner. However, if an expert system is to be created, then experts must be consulted and interviewed in any case. With this in mind, it is worthwhile to try and determine if the expert opinions can be used to define an appropriate revision operator to resolve inconsistencies. Asking a military expert to assign quantitative probabilities to all possible events in not realistic, and it is likely that such probabilities would not be reliable. On the other hand, it is more reasonable to a military expert to put different sequences of events in ranked order of likelihood. As such, it seems that our approach is based on the most reliable information we can fairly expect a human expert to provide.

5.2 Future Work

At present, the approach outlined in this paper has not been implemented, so one direction for future research would be the development of a prototype system. The easiest way to develop such a system would be to start with an existing expert system for maritime surveillance, and then extend the reasoning component to manage beliefs appropriately. There are practical problems with this approach, however, in the sense that military classification makes it difficult to apply the information obtained for one project to another project. As such, a more feasible approach to prototype development would be to start “from scratch” by conducting interviews and observations of military experts. Since anomaly detection for military surveillance is a problem of great practical interest, it is likely that it would be relatively easy to get access to appropriate operations centre workers.

The second direction for future research would be to improve the logical characterization of the maritime domain. Note that in our prototype formalism we have used some predicates that represent properties of entities and some predicates that represent class relationships. It is well known that classical logic does not make this distinction explicit, but it does seem to be important in this domain . As such, it would be worth reformulating our prototype in a suitable description logic or formal ontology. However, belief revision in the context of an ontology is a difficult task. As such, this theoretical improvement requires significant effort.

6 Conclusion

Maritime surveillance is an important topic from the perspective of security, and anomaly detection is one aspect of maritime surveillance that might benefit from the application of automated reasoning tools. In this paper, we have suggested that existing work based on expert systems could be improved by introducing suitable belief revision operators. We presented a high-level architecture as well as a basic prototype formalism. Although we believe this approach has a great deal of promise, there are many practical hurdles that must be overcome before we can develop a true practical implementation.

References

[1] C. Alchourron, P. Gardenfors, and D. Makinson 1985. On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic* 50(2):510–530.

[2] L. Bainbridge. 1979. Verbal reports as evidence of the process operator's knowledge. *International Journal of-Man-Machine Studies*, 11, 411-436.

[3] R. Bradley. 2007. The kinematics of belief and desire. *Synthese* 156(3):513–535.

[4] C. Carthel, S. Coraluppi, and P. Grignan 2007. Multisensor tracking and fusion for maritime surveillance *Proceedings of the 10th International Conference on Information Fusion*. 1-6.

[5] E. Cordingley. 1989. Knowledge elicitation techniques for knowledge-based systems. Knowledge elicitation: Principles, techniques and applications. Chichester, England: Ellis Horwood Ltd.

[6] DARPA/IXO. 2005. Predictive Analysis for Naval Deployment Activities (PANDA) Proposal Information Package. BAA05-44.

[7] M. Davenport. 2008. MISR-Vis Maritime Intelligence, Surveillance, and Reconnaissance Knowledge Base Final Report. Defence Research and Development Canada Technical Report. (DRDC-Valcartier). Document No. CR 2007-543.

[8] Grove, A. 1988. Two modellings for theory change. *Journal of Philosophical Logic* 17(2):.

[9] X. Li, J. Han, S. Kim, and H. Gonzalez. 2007. ROAM: Rule- and Motif-Based Anomaly Detection in Massive Moving Object Data Sets. *Proceedings 2007 SIAM International Conference on Data Mining.*

[10] N. Noy and W. Grosso. 2000. Knowledge-Acquisition Interfaces for Domain Experts: An Empirical Evaluation of Protege-2000. *Proceedings of the 12thInternational Conference on Software Engineering and Knowledge Engineering*.

[11] Naval Office of Office of Training Technology. 1998. Task Analysis. Chief of Naval Operations' Office of Training Technology Document.

[12] M. Riveiro, G. Falkman, and T. Ziemke 2008 Visual Analytics for the Detection of Anomalous Maritime Behavior *Proceedings of the 12th International Conference Information Visualisation*. 273-279.

[13] J. Roy 2008. Anomaly detection in the maritime domain. *Proceedings of Optics and Photonics in Global Homeland Security IV*. SPIE Vol. 6945.