Spatial ontologies for detecting abnormal maritime behaviour
Arnaud Vandecasteele, Aldo Napoli

To cite this version:

HAL Id: hal-00741035
https://hal-mines-paristech.archives-ouvertes.fr/hal-00741035
Submitted on 5 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract: The upsurge in piracy and the impact of recent environmental disasters have highlighted the need to improve maritime surveillance. Governmental and private initiatives have developed monitoring systems with improved acquisition and analysis capabilities. These systems rely on one major component, namely the detection of abnormal ship behaviour. This implies a detailed formalisation of expert knowledge. However, the quantity of data, the complexity of situations, the failure to take into account their spatial characteristics and the potential for the same scenario to be interpreted in different ways have proved to be significant problems. We therefore propose a new prototype for the analysis of abnormal ship behaviour. The system is based on a spatial ontology associated with a geographical inference engine. It automatically identifies suspicious vessels and associates them with probable behaviours defined by operational staff.

Keywords: Spatial ontology, Maritime surveillance systems, Anomaly detection, Automated reasoning

I. INTRODUCTION

The sea is a principal vector of the global economy and the scene of many criminal and illegal activities. It is estimated that 70% of cocaine is transported by sea and 82% of fish stocks are overfished in the Mediterranean [1]. Neither national authorities nor maritime experts can hope to monitor such a wide area, and such a large volume of maritime traffic, without the help of information systems. However, current systems are principally navigational aids that do not provide an adequate solution to new threats and the ingenuity of criminal organisations. Recent initiatives have defined the characteristics of new systems with improved data acquisition, processing and detection capabilities [2]. One of the key components of this architecture is the automatic identification of abnormal ship behaviour. Optimal identification requires the simultaneous assessment of many parameters, a large amount of data and a detailed modelling of the context [3]. Furthermore, it requires an understanding not only of the phenomenon itself but also of the many environmental interactions (natural or anthropogenic) inherent in the situation as it evolves. These interactions can be difficult to measure and require the integration of expert knowledge.

This article addresses these issues and describes the capabilities of a prototype system for the analysis of abnormal ship behaviour driven by a geographical ontology [4]. Specifically, we focus on the spatial heuristics used by experts in the identification of abnormal behaviour. We describe the development of a behavioural model and an associated knowledge base, which are integrated into a reasoning engine that can analyse ship trajectories. An integrated mapping module provides a visualisation of the results. The aim of this ontology-based system is to improve the analysis of abnormal ship behaviour and thus provide maritime surveillance experts with the necessary information for decision-making.

This article is organised as follows. Section 2 discusses ongoing work into new-generation monitoring systems. It highlights the strengths and limitations of these systems, particularly the need to formalise expert knowledge. In Section 3 we describe the ontology that was built to help in modelling the abnormal behaviour of vessels. Section 4 presents the system architecture of the prototype created from this ontology and discusses its strengths and limitations. Finally Section 5 draws some overall conclusions.

II. CONTEXT AND CHALLENGES

A. New-generation maritime systems

Various initiatives, both governmental and in the private sector, have tried to define the characteristics and scope of new maritime surveillance systems. However, because of the
challenges and economic sensitivity of the technologies involved very little documentation has been made available to the public.

One of the pioneering projects in the field of maritime surveillance is the Predictive Analysis for Naval Deployment Activities program (PANDA) carried out by the United States’ Defense Advanced Research Projects Agency (DARPA). The goal of PANDA is the automatic detection of abnormal ship behaviour. However, no information about the methodologies used or the technologies implemented has been disclosed. In Europe, there have been two initiatives. The ‘SECURITY system to protect people, goods and facilities located in critical MARitime area’ (SECMAR) project, carried out by Thales Underwater Systems aims to secure ports. This project uses statistical methods such as hidden Markov models to study the behaviour of ships. The second initiative is a series of four projects: Surveillance et Contrôle des Activités des Navires en mer (ScanMaris); Traitement et Authentification des Menaces et des Risques en mer (TAMARIS); Système d’Information et de Surveillance MARéitime pour l’Identification des comportements Suspects (SISMARIS), and ‘Integrated System for Interoperable sensors & Information sources for Common abnormal vessel behaviour detection & Collaborative identification of threat’ (I2C), all carried out by a public and private sector consortium.

The objective of this series of projects is to provide an end-to-end system where capabilities range from data acquisition to the processing of a threat at sea. To achieve this, new methods have been deployed, such as surface wave radar (which extends the area that can be monitored), algorithms to detect abnormal behaviour and collaborative approaches [5].

B. The components of new-generation maritime systems

These systems generally consist of four common components centred on three principal functions (monitoring, detection and analysis) [6]. The four components are:

- Extended monitoring: these new systems must be capable of ensuring maritime surveillance regardless of the weather, time of day or night, or whether the ship is willing to cooperate or not. Consequently, they involve the deployment of fixed and mobile sensors (in addition to traditional Automatic Identification Systems and conventional radar). These sensors aim to improve acquisition capabilities by extending the surveillance zone and improving detection of small boats.

- Enrichment and merging of sensor data: the heterogeneity of sensors means that a spatio-temporal correlation and aggregation of the data obtained is necessary. This step refines the data and reduces uncertainties. Once the data is merged, it is possible to enrich it using various information sources (institutional and/or private) related to vessel activity, maritime conditions, etc.

- Analysis of abnormal ship behaviour: this is a key feature of each of the projects. It concerns deducing and categorising a threat on the basis of a set of rules established in conjunction with operational staff. The rules can be based on static (age and type of vessel, etc.) but more particularly mobile (ship kinematics, sea state, etc.) criteria. The spatial characteristics of a situation are an essential element in the detection of abnormal behaviour they make it possible to establish, whether a vessel is in a prohibited area or near a coast (for example).

- Collaborative capacity: once suspicious behaviour has been detected and validated, the operator prepares a report for the competent authorities who can then decide whether to investigate further or initiate any necessary legal proceedings. Such interfaces improve collaborative capacity and enable better dissemination of information in the decision-making chain, both horizontally (expert to expert) and vertically (expert to decision-maker).

Each of these components has been the subject of extensive research. We, however, are particularly interested in the formalisation of expert knowledge for the analysis of abnormal ship behaviour.

C. Approaches to the formalisation of expert knowledge

Abnormal ship behaviour manifests itself as unusual or unexpected manoeuvres, or actions which do not comply with laws and regulations. Typically, detection is carried out by operators who draw upon their domain knowledge and the information available to analyse the situation. However, various limitations have been observed [7]. First, the operator is constrained by their cognitive abilities and attention span. Second, their analysis of the situation depends on their past experience. Given that past experience is particular to the individual, the heuristics used and therefore the results may be different from one person to another. Finally, illegal activities have evolved and their detection has become more difficult. Decision support systems that are specific to the maritime surveillance domain attempt to provide a response to these constraints. Such systems must enable experts to detect, analyse and interpret abnormal situations. This necessitates the formalisation and categorisation of ship behaviour and there are currently two approaches to the issue.

The first is data oriented. Using a large data-set and various statistical methods, profiles (patterns) of behaviour are created. The drawback of this approach is that it implies a potentially long learning curve and the development of algorithms specific to the maritime domain. These difficulties lead to a second approach, based on knowledge engineering.

The knowledge engineering approach consists of encoding the knowledge and heuristics used by experts in an information system. However, encoding runs into a well-known problem in the field of artificial intelligence — the knowledge acquisition bottleneck [8] (an obstacle that was encountered in the projects mentioned earlier). In addition, the methodology is based on a series of interviews with domain experts and this fact makes it difficult to formalise, manage, and effectively exploit knowledge. Several limitations have been observed, notably:
Formalisation: the analysis of abnormal ship behaviour requires the simultaneous assessment of many phenomena which can interact in many different ways. In the current knowledge engineering approach, each of the abnormal behaviours is described in a document that specifies the conditions for triggering an alert. However, this format does not provide the rigour or semantic richness needed to fully describe a scenario. For example, the same concept can be used in different scenarios but with a different meaning, or the interpretation of abnormal behaviour can vary according to the organisation (Customs, Coast Guard, etc.).

Management: to be useful, the knowledge base must contain the largest possible number of scenarios. In the current approach it quickly becomes difficult to manage such a large quantity of data. It may also be the case that certain scenarios can only be differentiated by some conditions. The problem then arises of how to distinguish between two, a priori similar situations?

Operations: in current decision support systems the document describing the scenario does not directly exploit the expert’s knowledge. Instead is the transcription of an interview with an expert. This transcription can lead to the loss of data.

These points underline the fact that current decision support systems for the detection and analysis of abnormal ship behaviour do not provide all the necessary mechanisms for knowledge formalisation.

Such limitations led us to another solution — specifically, spatial ontologies, which are both semantically richer and have a greater inference capability than current technologies. This approach makes it possible to formalise expert knowledge and then integrate it into a system for the analysis of abnormal ship behaviour in order to provide decision support. The following sections describe the approach in more detail.

III. AN ONTOLOGY-BASED FRAMEWORK FOR MODELLING BEHAVIOR

The representation of knowledge in a readable form is essential to understanding a phenomenon. Ontologies provide a way to achieve this. This section briefly describes the origins of ontology and the shortcomings of current ontology languages, which do not provide a means to capture spatial characteristics. This section also describes a method to overcome these limitations.

A. Ontology

Although originally associated with philosophy, this paper approaches the concept of ontology from the perspective of artificial intelligence. Designed as a response to the problems of integrating knowledge in information systems, they became a key element in the automatic processing of information at the semantic level. Defined as “an explicit and formal specification of a shared conceptualization” [9] they originate in the challenges posed by modelling the cognitive processes of individuals in the implementation of expert systems. It became necessary to develop new formalisms able to master both terminology and the associated semantics. This spawned a new form of modelling based on abstractions that could subsume or be subsumed by other abstractions — and the new discipline of knowledge engineering.

We have already highlighted the need to incorporate spatial characteristics (ships’ positions, mooring zones, etc.) into maritime surveillance ontologies, and although there is ongoing work by the World Wide Web Consortium (W3C) into the formalisation of a Web Ontology Language (OWL) for the publishing and sharing of ontologies, it does not as yet provide types, concepts or spatial relationships. Consequently, OWL cannot be used to associate a geographical location with a particular abstraction, nor can the topological relationships between entities be inferred [10]. Many previous studies have proposed ways to extend OWL to include this spatial dimension. They include the work of Hiramatsu and Reitsma on the integration of a spatial ontology into GIS software [11], Stocker and Sirin on spatial inference [12], Karmacharya et al. on the definition of a spatial SWRL query and rule engine [13], and Dia Miron et al. on the exploitation of semantic information using a spatial-temporal reasoner [14]. In addition to their technical merits, these studies demonstrate the feasibility and potential of implementing spatial semantics in ontologies.

The integration of a spatial dimension into our ontology was necessary both to formalise expert knowledge and to provide a way to automatically analyse abnormal ship behaviour. The following sections describe how this was achieved.

B. The knowledge base

The first step was to formalise expert knowledge in such a way as that it could be exploited by an information system. For this, we drew upon transcripts of interviews with experts who specified the steps necessary to identify an alert, the various relevant factors and their weight in the analysis. Protégé software² was used to process the information gathered. The result was a maritime surveillance ontology that consisted of around forty small classes, grouped into four principal classes (Figure 1):

Figure 1. The principal classes of the maritime surveillance ontology

1 http://www.w3.org/2005/Incubator/geo/
2 http://protege.stanford.edu/
• **Vessel**: this class consists of all identified vessels and all possible vessel characteristics. These include the type of vessel, (fishing boat, oil tanker, etc.), its departure point, and its position (stored in Well-known text format). Certain other characteristics are inferred (criticality level, potential alerts, etc.) from this data.

• **Alert**: this class defines the various alerts to be analysed. The alert typology was defined according to characteristics such as weather (rough seas, strong winds, etc.), ship behaviour (e.g. coming to stop when at sea) or spatial elements (entry into a prohibited area or a restricted fishing area, etc.). Each alert is given a criticality index indicating the danger level. In the case of spatial alerts, two further properties were added. The first (**hasSpatialAnalysis**) described the spatial representation to be used. Region connection calculus (**RCC8**) was used to define the possible spatial representations [15]. RCC8 implements eight basic relations between objects, including their intersection (partially overlapping) or proximity (externally connected). The second property (**hasData**) defined the data to be used in the context of the analysis.

• **Context**: the context class consists of elements that enrich alerts. It provides information related to the environment (sea state, wind speed, etc.) and makes it possible to more accurately categorise a situation. For example, the fact that a vessel stops in calm waters does not have the same meaning as the same manoeuvre in rough seas. In the first case, the vessel may be involved in drug trafficking while in the second it may be damaged.

• **Scenario**: this class represents the possible interpretation of a situation given the number of alerts and contextual elements. For example, a bulk carrier that stops near the coastline in calm seas and good weather may be identified as a goods or drug trafficking scenario. The **Scenario** class draws upon members of the **Context** and **Alert** classes. Eventually we hope to be able to establish the most likely scenario for a particular type of vessel.

The following example is a concrete illustration of the approach and demonstrates the potential of the ontology to describe and define a model of abnormal behaviour.

A fishing boat leaves port in a country known to be a flag of convenience and rapidly makes it way to a protected fishing zone. It stays there some time before leaving for another port. This situation is modelled as follows. First, three new instances of the **Alert** class are created: a ‘high-speed’ instance, a ‘stop at sea’ instance and a ‘no fishing’ instance. One or more of these alerts are then assigned to the boat in question depending on its characteristics (speed, heading, etc.). Then an ‘illegal fishing’ scenario is added which is linked to the previously defined alert instances through the **hasAlert** attribute.

However, to be useful in an analysis system, the ontology requires the addition of a rule base in order to identify and characterise situations.

### C. Detection rules

We used the Semantic Web Rule Language (SWRL) to represent and integrate rules into the ontology. SWRL, which is a combination of OWL-DL and RuleML [16], enriches the semantics of an ontology. Designed to support reasoning based on Descriptive Logic Programs (DLP) and Horn logic, SWRL rules take the form of an implication between an antecedent (body) and consequent (head). In other words, whenever the conditions specified in the antecedent hold, then the conditions specified in the consequent must also hold. Unlike OWL, SWRL only allows the addition of relationships or existing properties according to whether or not they meet the rule. In addition to the OWL predicates, SWRL has additional built-in functions (mathematical, Boolean, etc.) that extend the initial OWL functionality and make comparisons and calculations possible. We adopted the idea of the ‘built-in’ function to integrate spatial capabilities into our SWRL rules.

The rule base was defined in collaboration with domain experts and from a review of the literature [17]. Two types of rules were distinguished, characterised by whether they were non-spatial or spatial. The first (non-spatial) rules directly exploited the reasoning capabilities of SWRL. For example, suppose the aim is to detect a ship moving at a speed that is excessive for its type. The request, translated into SWRL is shown below and reads as follows, “If a vessel (?vessel) of an identified type (?vesselType) has a speed (?speedVessel) greater than (greaterThan) the maximum speed allowed for that type of vessel (?speedTypeVessel), then trigger an alert (Alert_Speed_HighSpeed).”

```
Vessel(?vessel),
OP_Vessel_hasVesselType(?vessel, ?vesselType),
DP_Vessel_hasSpeed(?vessel, ?speedVessel),
greaterThan(?speedTypeVessel, ?speedVessel)
→ OP_Vessel_hasAlert(?vessel, Alert_Speed_HighSpeed)
```

As we have already underlined, the analysis of the abnormal behaviour of vessels requires an understanding of the spatial characteristics of the situation. Therefore, the second type of rule extended the traditional functions of SWRL, and defined new built-in spatial functions (intersects, touches, etc.). For example, some areas may be closed to shipping. The entry of a ship into such an area should automatically trigger an alarm. This rule is expressed in SWRL by:

```
Vessel(?vessel),
DP_Vessel_hasPosition(?vessel, ?vesselPosition),
DP_Analysis_DataPath(Alert_Area_Restricted, ?geoData),
intersects(?geoData, ?vesselPosition)
→ OP_Vessel_hasAlert(?vessel, Alert_Area_Restricted)
```

Well-known text is a text mark-up language developed by the Open Geospatial Consortium for representing vector geometry.
In this example, we record an intersection (intersects) between the position of the ship (?position) and geographical data specified in the Alert class (?geoData). If the result is positive, we add an object property (OP) to link the detected vessel and the specified alert (Alert_Area_Restricted). Clearly, depending on the alert to be detected, other data or spatial functions (touches, overlaps, etc.) may be selected. Changing the spatial function simply entails changing the keyword in the SWRL rule. Changing data involves specifying which data to use from the data property (DP) DP_Analysis_DataPath.

IV. A SYSTEM FOR THE ANALYSIS OF ABNORMAL SHIP BEHAVIOUR DRIVEN BY AN ONTOLOGY

Once the ontology has been built, it is integrated into a system for the analysis of ship behaviour, in order to automatically detect abnormal activities.

A. System architecture

The analysis consists of five stages of information processing (Figure 2). First, as new sensor data arrives it is merged and the sensor database is updated (Step 1). The merged data is then fed into the ontology processing engine (Step 2). Next, the inference engine processes this new information using the predefined SWRL rules and characterises the behaviour of the vessel (Step 3). The result is stored in the detected threat ontology (Step 4) and displayed by the mapping module (Step 5). This section describes in detail Steps 3, 4 and 5 which form the heart of our system.

The ontology is updated with new information arriving from sensors. Two types of information can be distinguished: static data that is not expected to change (for example a new model of ship in the Vessel class); and dynamic data (typically vessel properties such as position or velocity). Updates are carried out via the OWL API (an open-source Java library) which enables the creation and manipulation of ontologies.

The ontology cannot be exploited until processed with a reasoning engine. The purpose of the engine is to infer new facts from existing data. In our case, it is used to infer a situation according to the position and characteristics of a vessel. Amongst the reasoners currently available, we decided to use Pellet. Although it is not the most efficient option available [18], it has very good SWRL support, it is compatible with OWL 2, it integrates seamlessly with the Jena framework and the OWL API, and more importantly, to the best of our knowledge it is the only reasoning engine that makes it possible to add new built-in functions [19]. The reasoner applies the specified SWRL rules to the ontology to determine whether or not a ship is behaving abnormally. Depending on the result of the inference, new attributes and new relationships are automatically added, and the ship in question is assigned a criticality level.

The display module connects to the detected threat ontology. Once the ontology is updated, the results are displayed on a map interface. This enables the analysis to be visualised more intuitively. The information displayed can relate to the ship’s properties, its spatial characteristics or to a specified alert. This is illustrated by the two examples shown in Figure 3.

4 http://owlapi.sourceforge.net/
5 http://clarkparsia.com/pellet/
6 http://incubator.apache.org/jena/
Figure 3. Examples of abnormal ship behavior. A: detection with a spatial rule. B: detection by attributed rules.

The first case corresponds to the SWRL rule discussed in Section III C which describes the entry of a ship into a restricted zone (the area represented by the light-coloured polygon). The inference engine has automatically detected an offense based on the ship’s position, and on the DataPath property of the restrictedArea instance of the Alert class. On the basis of this information the vessel is reported as acting suspiciously. At the same time, the ship is automatically allocated a danger level based on the criticality index of the alert.

In the second case, the inference engine has detected two non-spatial alerts; the first is related to the age of the vessel and the second to its speed. Both these alerts are strongly linked to the type of vessel (for example, the maximum speed of a container ship is less than that of a ferry). This case is interesting because it illustrates the situation where a vessel raises more than one alert. We have decided (for the time being) to define the danger level of a vessel according to the alert with the highest criticality index. Visually, this progressive increase in criticality is represented by colour differences. Vessels are displayed surrounded by a halo of yellow, orange or red to represent increasing levels of criticality.

B. Strengths

The prototype was based on a data set containing the routes taken by 2,200 vessels in the Mediterranean and totalled more than 36,000 coordinates. Consequently, one of the principal challenges was to provide a scalable system able to withstand and adapt to the high volume of data inherent in this field. Previous research has only used a limited number of instances and properties. For example, Roy and Davenport’s paper [20] on the development of a spatial analysis module and the analysis of the kinematics of vessels (KIGAM) was based on around a hundred recordings. Like our project, the authors aimed to identify suspect vessels and categorise potential threats. Although the volume of data they analyse is smaller, their ontology is (at the present time) more complete than ours.

Our approach provides domain experts with great flexibility in the development of a model. Previously defined keywords (intersects, touches, etc.) provide the foundations of a topological logic. This intuitive, keyword-based method enables operators to take account of spatial characteristics when modelling scenarios. Although the method can be improved, we have found it to be more suitable than traditional Geographic Information Systems, which require specific training and knowledge that maritime surveillance experts do not necessarily have. Moreover, these systems do not have the same modelling capability that ontologies offer.

C. Limitations

Although the prototype is now functional, various areas for improvement are already apparent.

At the present time, the prototype described here can process a database containing details of several thousand ships, using around a dozen simple rules, in under a minute. Scalability tests, which involve adding a spatial index (similar to that found in databases), are in progress. Initial results have shown significant gains in computing time.

It would also be useful to make it easier for the operator to create detection rules. Currently, these are added directly in SWRL using the Protégé interface. This can be a difficult task for the person responsible for developing the model. It therefore seems useful to create an interface that would make it possible to draw these rules graphically (as described by Snoggle [21]). The expert could then draw a scenario that would be translated automatically into SWRL and integrated into the ontology.

Finally, calculating the probability of a scenario has not yet been addressed. This step would measure the semantic similarity between, on the one hand, the characteristics of a particular vessel and the other hand, the characteristics of the scenario. We are considering two approaches to calculating this similarity: the first is based on the informational content of objects (a node-based approach), and the second is edge-based [22]. We plan to test these two solutions in order to determine which offers the most relevant results.
V. CONCLUSIONS

This paper described the development of a prototype to demonstrate the potential of spatial ontologies for the detection of abnormal ship behaviour. To achieve this it was necessary to extend the initial capabilities of SWRL rules in order to include a spatial dimension. This provides maritime experts with the ability to describe behaviours while at the same time ensuring maximum expressiveness. The ontology and the associated rule-base were processed by an inference engine in order to provide automatic scenarios corresponding to the situation in question. In the case of abnormal behaviour, the system automatically determines the criticality level and adds this information to the ontology. Finally, the results of the inferences are displayed on a map interface so that the expert can better analyse and understand the situation.

The scenarios described here show the potential of our approach in the domain of maritime surveillance. The examples given show the potential of the system to analyse both spatial and non-spatial situations. A meeting with domain experts (Marine Surveillance and Rescue Centres, Coast Guard, etc.) is planned; this will enable us to enrich our approach, add new types of alerts and complete our analysis.

Beyond the prototype described here, this article has highlighted the need for ontologies to take into account spatial characteristics. At present, there is no ontology of spatial types (line, point, etc.). For this reason, our prototype recorded the location of vessels in textual form (Well-known Text). However, this storage method does not exploit all the semantics of spatial information (spatial functions, indexes, etc.). It would therefore be interesting to focus further work on the definition and implementation of new spatial types and predicates, as is currently the case in databases.

REFERENCES