A semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behavior analysis

Arnaud Vandecasteele, Rodolphe Devillers, Aldo Napoli

A semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behavior analysis

Arnaud Vandecasteele[1], Rodolphe Devillers[1] & Aldo Napoli[2]

[1] Department of Geography, Memorial University of Newfoundland, St. John's, NL, Canada A1B 3X9
a.vandecasteele@mun.ca, rdeville@mun.ca
[2] MINES ParisTech, Centre for Research on Risk and Crises (CRC), Rue Claude Daunesse,
CS 10207, 06904 Sophia-Antipolis Cedex, France
aldo.napoli@mines-paristech.com

Abstract

Detection of abnormal movements of mobile objects has recently received a lot of attention due to the increasing availability of movement data and their potential for ensuring security in many different contexts. As timely detection of these events is often important, most current approaches use automated data-driven approaches. While these approaches have proved to be effective in specific contexts, they are not easily accepted by operators in charge of surveillance due, among other reasons, to the lack of user involvement during the detection process.

To improve the detection and analysis of maritime anomalies this paper explores the potential of spatial ontologies for modeling maritime operator knowledge. The goal of this research is to facilitate the integration of human knowledge by modeling it in the form of semantic rules to improve confidence and trust in the anomaly detection system.

Introduction

The increasing ubiquity of positioning devices (e.g., GPS, AIS) on mobile objects, often combined with data collected from other sensors (e.g., camera, radar), allows for the collection of large amounts of movement data. Such data can be used for various applications such as video surveillance, traffic analysis or animal migration monitoring (Spaccapietra *et al.*, 2008). In the maritime domain, research projects and government initiatives are increasingly using such data to produce dynamic, accurate and comprehensive pictures of maritime environments (Maritime Domain Awareness) (Bürkle and Essendorfer, 2010).

However, exploring and analyzing such a large amount of data in a short timeframe is often an unrealistic task for maritime operators. Constraints such as the time pressure, the uncertainty and the heterogeneity of the data and the complexity of the situations also have to be considered. All of these factors have a critical impact on decision quality and can lead to a cognitive overload for the operators (Riveiro *et al.*, 2008). To support operators in their tasks, different methods and algorithms mainly based on data-driven approaches that automatically detect anomalies have been investigated (Laxhammar, 2008; Martineau and Roy, 2011). Despite significant research efforts and the benefits provided by these approaches, automated anomaly detection methods have not yet been widely adopted by operators (Riveiro and Falkman, 2009). This lack of user adoption has different causes, including the frequent absence of users' involvement during the detection process and the use of complex detection techniques that are often hard to understand by the operator and hard to explain and justify at a management level (Van Laere and Nilsson, 2009). In this paper we propose a semi-supervised learning framework for detecting and analyzing abnormal maritime behaviors. The proposed framework uses a semantic approach for modeling the maritime surveillance domain associated to a rule language for describing operator knowledge. Once identified, potential events are stored and semantically enriched in a dedicated ontology. Then, events related to a vessel are analyzed using a Case Based Reasoning (CBR) (Schank, 1983) approach to identify similar previous behaviors. Finally the potential events and behaviors are displayed onto a geovisual analytics platform to allow for a human operator's input. The aim of this platform is both to offer a cartographic representation of the suspicious situation that facilitates its understanding and to allow the enrichment of the knowledge rule database depending on the validation or rejection of a detected behavior by the operator. This validation step updates the weight of the rule using expert feedback to reduce false alarms.

The next section presents the proposed ontological model developed for maritime anomaly detection. First a brief description of ontologies is given. Then the proposed model based on spatial ontologies is presented. The last section summarizes the approach and presents future research directions.

Maritime anomaly detection with ontologies

Converting raw movement data into semantic objects requires trajectories to be described using specific formalisms that can model concepts. Currently, most methods and tools that allow this to be done come from the field of ontologies. Although originally associated with the domain of philosophy, we approach the concept of ontology from its use in artificial intelligence (AI). In this context, an ontology can be described as a computational artifact used to formally model a domain and provide a shared and common understanding of this domain between and among peoples and systems (Studer *et al.*, 1998). In the next paragraph we discuss how spatial ontologies and spatial reasoning can be combined to improve maritime surveillance.

A semi-supervised ontological framework for maritime anomaly detection

Most existing maritime anomaly detection processes are based on automatic data-driven methods where an anomaly is considered as being a deviation from normality (Laxhammar, 2008). While these methods support the discovery of unknown patterns and can handle large volumes of data, a number of reasons make their use challenging with real-world problems (Hunter, 2009). One of those reasons is a lack of user involvement during the decision process associated to a lack of support for analytical reasoning process (Riveiro and Falkman, 2010). As ontologies can be used to represent different types of knowledge, recent research efforts explored the potential of ontologies for maritime surveillance (Vandecasteele and Napoli, 2012; van Hage *et al.*, 2011). Using ontologies, maritime operators' knowledge can be modeled and integrated using specific concepts. A crucial point in the development of such types of systems is to provide a clear conceptual framework that minimizes complexity while maintaining flexibility. While most current research focuses on semantic trajectories, we propose to expand this framework to integrate the concepts of semantic events and semantic behaviors. Such a three-level framework (Figure 1) provides a convenient way to describe both the maritime domain, but also semantic events and behaviors. This framework extends the ontological framework developed by Yan and his colleagues (2012) that proposed using three main ontologies to capture semantics for trajectory data.

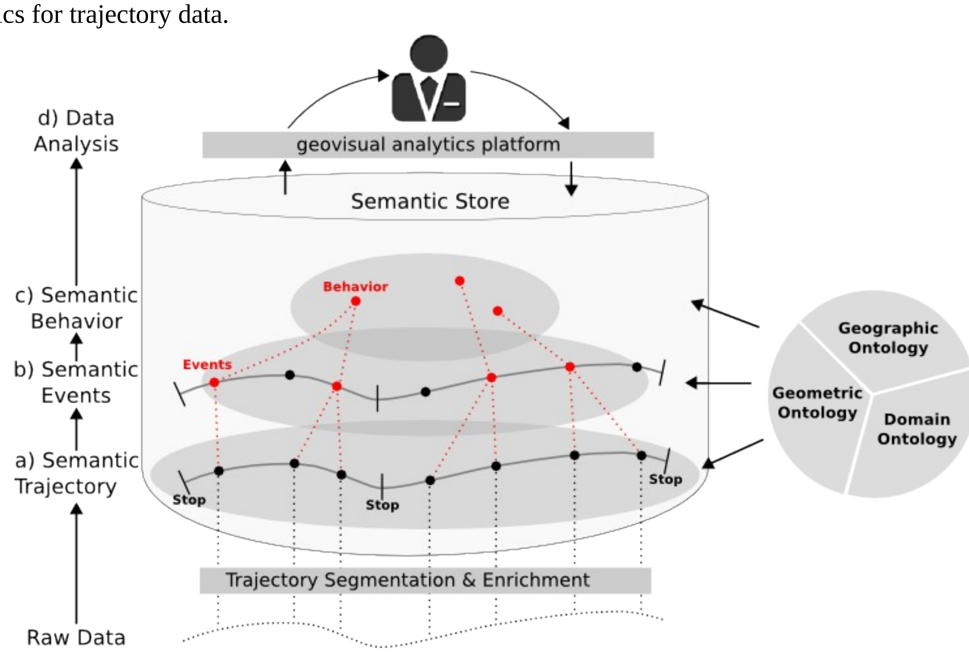

Figure 1. Conceptual architecture of the system.
Raw data are extracted, enriched and then analyzed to identify potential abnormal behavior

The components of the proposed model are illustrated using the example of a foreign vessel conducting illegal fishing activities.

For the analysis to be performed, the raw data collected by different sensors must be segmented into appropriate representations of the vessel's trajectory (step (a) in Figure 1). Applying these models to real mobile objects is an intensive task that can only be realized with an appropriate trajectory model. While different trajectory models have been proposed, most of them focus on spatio-temporal components (x,y,t) rather than semantic components (e.g., movement in a restricted area) (Spaccapietra *et al.*, 2008). Spatio-temporal positions alone do not provide sufficient information about the context in which mobile objects evolve, making it difficult to have a global interpretation of movement behaviors (Yan *et al.*, 2012). Consequently, we decided to consider not only the spatio-temporal positions of the trajectory but also semantic trajectory units (e.g., *begin, stop, moves, end*). These semantic units can be enriched with different types of knowledge (e.g., spatio-temporal, geographic, domain) to provide end-users with high-level semantic descriptions of trajectories and a better understanding of the situation. Using the illegal fishing scenario described above, the semantic trajectory of the vessel could be represented as: *(begin, port) → (move, sea, high-speed, 1 hour) → (move, restricted fishing area, fishing, 2 hours) → (end, port)*, where semantics units were enriched with geographic (e.g., sea, restricted fishing area), spatio-temporal (e.g., move, begin) and domain (e.g., fishing) properties to provide a better understanding of the vessel's trajectory.

This first step allows performing further analyses of trajectories and identifying potential alerts that could relate to abnormal movements (step (b) in Figure 1). Depending on the human knowledge integrated in the surveillance system, different alerts can be automatically detected. Human knowledge can be encoded in the surveillance system in the form of semantic rules. In the maritime domain, these rules can be simple, such as "if the speed of the vessel is above X, then generate an event of type Y", or can be more complex and include different spatio-temporal elements such as "if two vessels are moving in parallel within a certain distance and during a certain time, then generate an event of type Z". Using the previous example a typical rule could be "if a foreign fishing vessel enters a restricted fishing area, then generate an event of type 'illegal fishing' ". Once all of the rules are defined and the semantic trajectories are stored in the knowledge database, links with potential alerts have to be specified. This step is done using a spatio-temporal inference engine that will automatically analyze semantic trajectories with the rules to identify potential suspicious events. While many inference engines exist (e.g., Jena, FaCT++, Pellet), only few of them are able to handle spatio-temporal data. For this research, further tests will be needed to select the most appropriate inference engine. Finally, the alerts detected will be described using the Simple Event Model proposed by Van Hage *et al.* (2011) that provides the minimal set of classes, properties and constraint necessary to describe events.

The ultimate goal of the proposed framework is the interpretation of vessels' activities and behaviors (step (c) in Figure 1). This step uses a CBR (Schank, 1983) approach to compare previous behaviors defined by the operators with the current facts of the knowledge base. CBR comes from the field of AI and has been selected for its capacity to implement standard reasoning procedures similar to human reasoning and capture domain knowledge, even if the domain is imperfectly understood or hard to codify. CBR relies on the assumption that successful solutions used to solve past problems can be reused to solve new similar problems. Similarity between new and past problems is computed between sets of characteristics. For example, if a previous case similar to the illegal fishing scenario described above exists, then the identification of this situation as a potential illegal fishing behavior will be automatically proposed to the maritime operator.

At this point, semantic events and semantic behaviors are just facts stored in the semantic store. To be analyzed, they need to be integrated through a specific user interface that provides a better visualization and a better understanding of the maritime situation (step (d) in Figure 1). The Hybrid Spatio-Temporal Filtering (HSF) interface developed by Enguehard *et al.* (2012) has been chosen due to its two major components that can be used to understand why potential events have been detected. These components are the geovisualization components and the interactive hybrid filtering components. The geovisualization component uses the Java World Wind framework and offers a three-dimensional view of the data. The interactive hybrid filtering component can be used to filter out uninteresting movement data and to isolate specific movement patterns. This interface allows the operator's attention to focus on potentially interesting data and then improves the understanding of the situation. To provide a better understanding of the situation, HSF will be extended by two other components. First, a query engine that allows maritime operators to query facts stored in the knowledge database and then identify specific data that can be used to improve the understanding of the maritime situation. Secondly, a validation engine that reduces false alarms by a weighted score-based rule adaptation mechanism through expert feedback.

Future works and conclusion

This paper has described a novel semi-supervised framework based on spatial ontologies that can be used for maritime surveillance. An important contribution of our approach is to offer a consistent framework based on ontologies that allows modeling not only semantic trajectory but also semantic events and behaviors. Semantic trajectories are obtained using different ontologies (e.g., geographic, domain) and semantic events and behaviors are automatically detected by using maritime operator's knowledge stored in a semantic store. Therefore the user interface allows the maritime operator to understand the reasons why thr abnormal events detected are considered suspect. Finally, to reduce the number of potential false alarms, the validation or rejection of detected facts is taken into account by the framework through expert feedback.

The proposed framework is still a work in progress and a number of questions remain to be answered. The next step will be to link these components together and test the semantic model with real situations. Our future work will not only focus on the analysis of the spatio-temporal events associated to a vessel, but also to its relationship with others elements of the context (e.g., other vessels, proximity of an area of interest). Although the proposed framework has been designed for maritime anomaly detection, its approach is generic enough to make it applicable to a broader range of movement data and its application to other domains would also be interesting to explore.

References

Bürkle, A. and B. Essendorfer (2010), "*Maritime surveillance with integrated systems*". *In*: Waterside Security Conference (WSS 2010), IEEE, Marina di Carrara, Italy: 1–8.

Enguehard, R.A., O. Hoeber and R. Devillers (2012), "Interactive exploration of movement data: A case study of geovisual analytics for fishing vessel analysis". *Information Visualization, 12*:85–101.

Gruber, T.R. (1993), "A translation approach to portable ontology specifications". *Knowledge Acquisition*, 5:199–220.

Hunter, A. (2009), "*Belief modeling for maritime surveillance*". *In*: 12th International Conference on Information Fusion (Fusion 2009), IEEE, Burnaby, BC, Canada: 1926–1932.

Laxhammar, R. (2008), "*Anomaly detection for sea surveillance*". *In*: 11th International Conference on Information Fusion (Fusion 2008), IEEE, Cologne, Germany: 1–8.

Martineau, E. and J. Roy (2011), *Maritime Anomaly Detection: Domain Introduction and Review of Selected Literature*, Defence R&D Canada, Canada.

Riveiro, M. and G. Falkman (2009), "*Interactive Visualization of Normal Behavioral Models and Expert Rules for Maritime Anomaly Detection*". *In*: Sixth International Conference on Computer Graphics, Imaging and Visualization (CGIV 2009), Singapore: 459–466.

Riveiro, M. and G. Falkman (2010), "*Supporting the Analytical Reasoning Process in Maritime Anomaly Detection: Evaluation and Experimental Design*". *In*: 14th International Conference on Information Visualisation (IV 2010), London, UK: 170–178.

Riveiro, M., G. Falkman and T. Ziemke (2008), "*Visual Analytics for the Detection of Anomalous Maritime Behavior*". *In*: 12th International Conference on Information Visualisation (IV 2008), London, UK: 273–279.

Schank, R.C. (1983), *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. Cambridge University Press, New York, NY, USA.

Spaccapietra, S., C. Parent, M.L. Damiani, J.A. de Macedo, F. Porto and C. Vangenot (2008), "A conceptual view on trajectories". *Data & Knowledge Engineering*, 65:126–146.

Studer, R., V.R. Benjamins and D. Fensel (1998), "Knowledge engineering: Principles and methods". *Data & Knowledge Engineering*, 25:161–197.

Vandecasteele, A. and A. Napoli (2012), "*An Enhanced Spatial Reasoning Ontology for Maritime Anomaly Detection*". *In*: 7th International Conference on System Of Systems Engineering (SOSE 2012), IEEE, Genova, Italy: 247–252.

Van Hage, W.R., V. Malaisé, G.K.D. De Vries, G. Schreiber and M. Van Someren (2011), "Abstracting and reasoning over ship trajectories and web data with the Simple Event Model (SEM)". *Multimedia Tools and Applications*, 57:1–23.

Van Laere, J. and M. Nilsson (2009), "*Evaluation of a workshop to capture knowledge from subject matter experts in maritime surveillance*". *In*: 12th International Conference on Information Fusion (Fusion 2009), IEEE, Seattle, USA: 171–178.

Yan, Z., D. Chakraborty, C. Parent, S. Spaccapietra and K. Aberer (2012), "Semantic Trajectories: Mobility Data Computation and Annotation". *ACM Transactions on Intelligent Systems and technology*, 9:1–34.