Detecting Search and Rescue missions from AIS data

Iraklis Varlamis
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: varlamis@hua.gr

Konstantinos Tserpes
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: tserpes@hua.gr

Christos Sardianos
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
Email: sardianos@hua.gr

***Abstract*—The crossing of the Mediterranean by refugees has turned to be an extremely perilous activity. Human operators that handle Search and Rescue (SAR) missions need all the help they can muster in order to timely discover and assist in the coordination of the operations. In this work we present a tool that automatically detects SAR missions in the sea, by employing Automatic Identification System (AIS) data streams. The approach defines three steps to be taken: a) trajectory compression for affordable real time analysis in the presence of big data; b) detection of sub-operations to which a SAR mission is actually decomposed, and; c) synthesis of multiple vessels' inferred behavior to determine an ongoing SAR mission and its details. The evaluation results are promising showing that AIS data carry highly valuable information even in the absence of any other type of data that could make the problem easier (e.g. coast guard signals).**

***Keywords*-search and rescue missions, AIS data. density based clustering, navigation pattern detection, trajectory mining, complex event processing.**

I. Introduction

Maritime surveillance has been a topic that is concentrating a lot of research and commercial interest during the last few years. The main reasons behind this interest are that i) vessels are nowadays generating unprecedented amounts of situational information, and ii) applications such as autonomous navigation that can leverage on this information are on high demand from relevant stakeholders. Traditional surveillance systems have significantly contributed to increasing the efficiency and safety of operations at sea. Their primary functions are to monitor vessels positional characteristics, typically using Automatic Identification System (AIS) [1] data and to analyze it so as to provide situational awareness services.

The research advances in data analytics and mining have contributed many solutions to maritime surveillance problems, such as ship route prediction and anomaly detection [2], [3]. However, environmental (e.g. the opening of the northwest passage) and socioeconomic (e.g. refugee and migrant influx passing through the Mediterranean) factors are bringing risks at sea once again on the rise, thus demanding an evolution in maritime surveillance systems, such that they can support a human operators better understanding of the complex reality and enhance their decision-making. A critical requirement for such systems is that they exhibit the ability to foresee unfolding cautious and potentially hazardous situations, so as to propose measures of danger avoidance.

A situation that demonstrates how a situational awareness system must meet such a requirement is the identification of *Search and Rescue* (SAR) missions taking place in the context of the refugee crisis in the Central and Eastern Mediterranean since 2015 [4]. The unnecessary loss of so many human lives indicates that SAR operations can benefit from situational awareness information either with the intention to identify them on time and to assist in the coordination of rescue ships, but also in order to clarify the role of involved vessels. At the macroscopic level, this knowledge can be even used to assess the impact of related policies.

Since the refugees rubber dinghies are highly unlikely to transmit any signal about their position, a SAR mission must be automatically recognized by studying the behavior of nearby vessels or vessels that approach to assist. The availability and richness of AIS data makes it a comprehensive source for this sort of analysis. It contains periodic, timestamped information about vessel's location and motion characteristics (heading, course, speed, etc.) as well as moderately reliable details about its operation, cargo, destination, etc. Therefore, the research question is shifted to whether AIS data can be used to automatically identify and explain SAR operations and generate insights about rescue behavior and migration patterns.

What makes the problem particularly challenging is the fact that a SAR mission is decomposed into multiple smaller operations, such as patrolling, rushing to the event, directly collecting people at sea, collecting people by other means such as floating devices or smaller boats, safely returning the collected people to the port etc. Recognizing a vessel performing any of those operations is critical into timely understanding incidents, their details and how the complete SAR operation evolves.

In this paper, we attempt to tackle the problem of SAR mission detection by proposing an unsupervised method, which begins with the detection of the navigational patterns that vessels in a SAR mission follow and continues with the

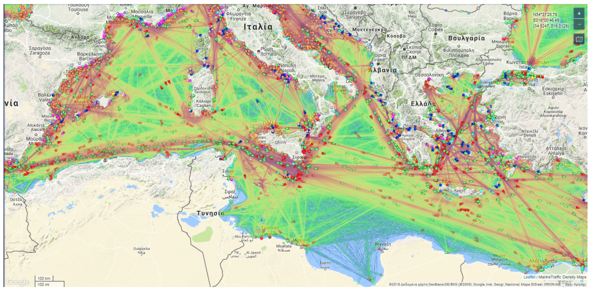


Figure 1. A snapshot of more than 3,000 vessels that sail the Mediterranean sea on a typical day.

identification of a SAR mission event by the composition of pattern information for many vessels in the same area. In the following section, we highlight some key research works in the field. In section III we formally define the problem we attempt to solve and describe the information that we have available. Section IV focuses on the methodology we propose and the tools that we employ and section V provides some initial results. Finally, Section VI, summarizes the findings so far and discusses our next steps.

II. Related Work

The problem of maritime situational awareness using AIS data is not new [3], [5]–[11]. Even though most of these works focus on understanding the navigational patterns from AIS data (e.g. for trajectory prediction), there are also cases in which the authors have focused on a particular operation such as fishing patterns [12]. Those approaches in vessel-at-sea behavior analysis are based on the monitoring of a single vessel at a time. The aim of our work on the field of situational awareness revolves around the identification of SAR missions in particular by means of analyzing the behavior of multiple vessels in the same period.

A search on the on-line maritime monitoring service of MarineTraffic[1] for the west and central Mediterranean region (see Figure 1), reveals more than three thousand vessels that sail in an area of 950,000 square miles and send AIS data every few minutes. The requirement of managing such a large amount of data in such a short period of time sets an unprecedented challenge to the processing task at hand. Data need to be aggregated, synchronized, cleansed, and then placed to an appropriate data structure to be processed at high speeds with the purpose to classify each of the data objects as part of a normal/off-normal behavior. Employing Complex Event Processing (CEP) to summarize the vast amounts of data seems to be a promising approach.

Researchers in the CEP field are compressing the raw data by extracting simple (or instantaneous) events from the timed series of positions [13]. These events are basic and, in the domain we are dealing with, are usually labeled using verbs such as turn, stop, accelerate or decelerate. These events can be formalized using a computer logic language, like Prolog, or an ontology language, like OWL and processed in order to reveal more complex events. To be more compliant with this specific field, a group of semantically rich languages can be used such as SPAseq [14] and RTEC calculus [15].

By summarizing a trajectory, we achieve great reductions in the data volume, because we only need to store critical points along the course of the object, so if the object is following predictable routes, like a naval vessel in open sea, we could replace hours of AIS data with a few selected positions (timestamps etc) and associated labels, that correspond to the position of an acceleration, deceleration or turn event, and omit all the information collected between consecutive events. Patroumpas et. al. [16] mention a 98% compression ratio of the original data, using an event-based data summarization methodology, which takes simple but semantically enriched events, and deduces more complex ones by applying logic rules to a subset of the detected events. Usually, the subset is a based on a large time window, comprising the most recent events for a certain object [16], [17]. For example, when a series of turn events are detected, we can assume that the vessel is either taking a long turn or trying to reverse course with a U-turn, or doing a routine zig-zag scan of an area.

In conclusion, Complex Event Processing can be a useful tool in the detection or generalization of a single vessel's navigation pattern, and the identification of individual events that may be part of an SAR mission. However, when the complex event engages multiple vessels at the same period that act in a wider area, a more complex analysis is necessary, comprising event and pattern correlation and operational intelligence. For this reason, in this work we employ machine learning approaches which are more appropriate when it comes to capturing complex notions.

III. Problem description and dataset

According to the proposed approach, the problem of detecting Search and Rescue missions directly from AIS data is broken down into two discrete subproblems: i) detecting that a vessel performs specific maneuvers that relate to Search and Rescue missions [18], ii) detecting that more than one vessels perform such maneuvers in the same area, within the same time period.

In order to evaluate the feasibility of our approach, we employed a dataset, which was provided to us by the community-based AIS vessel tracking system (VTS) of MarineTraffic. Due to provider limitations, the dataset contains information for 25 vessels, with 5 of them being cargo ships that have been hired by NGOs for assisting in SAR missions and the other 20 being randomly selected

[1] MarineTraffic is an open, community-based maritime information collection project, which provides information services and allows tracking the movements of any ship in the world. It is available at: https://www.marinetraffic.com

Figure 2. A snapshot of the area that we monitor in our case study.

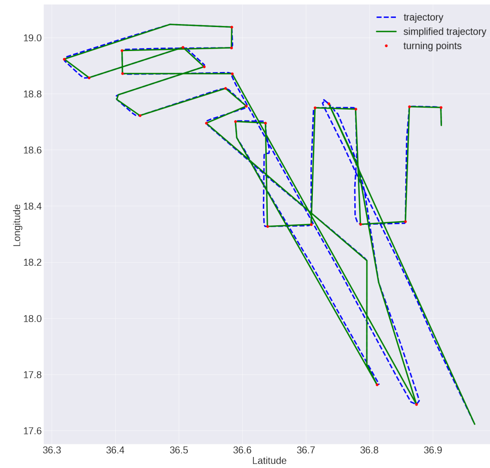


Figure 3. The information kept for a route contains only the turning points, after compression using the RDP algorithm and turn identification.

Table I
DATASET ATTRIBUTES

Feature	Description
ship_id	Unique identifier for each vessel
latitude, longitude	Geographic location in digital degrees
sog	Speed over ground in knots
cog	Course over ground in degrees with 0 corresponding to north
heading	Ship's heading in degrees with 0 corresponding to north
ship_type	Ship's type like: Yacht, Supply Vessel etc
timestamp	Full UTC timestamp
Departure_timestamp	Ship's departure datetime
Departure_port_id	Ship's departure port id
Departure_port_name	Ship's departure port name
Departure_port_type	
Departure_country_code	The country code for ship's starting point
Arrival_timestamp	Ship's departure datetime
Arrival_port_id	Ship's arrival port id
Arrival_port_name	Ship's arrival port name
Arrival_port_type	
Arrival_country_code	The country code for ship's arrival point

ships navigating in the same area. It is among our next steps to use a larger dataset and for this reason the whole approach has been designed to support scalability: redundant AIS information has been removed and an incremental version of the clustering algorithm has been employed. The area that our dataset covers is bounded by a rectangle with the top left point at (45.43612, 6.990125) and the bottom right at (33.06117,23.59917) with most of the traffic being at the part of Mediterranean sea between Italy and Tunisia and Libya, as depicted in Figure 2. The vessels have been monitored for a 3 months period starting at July 1st, 2015 and ending at September 29th, 2015. During that time a great number of incidents happened involving refugee boats.

The dataset contains 211929 AIS records in total each comprising 17 attributes as described in Table I.

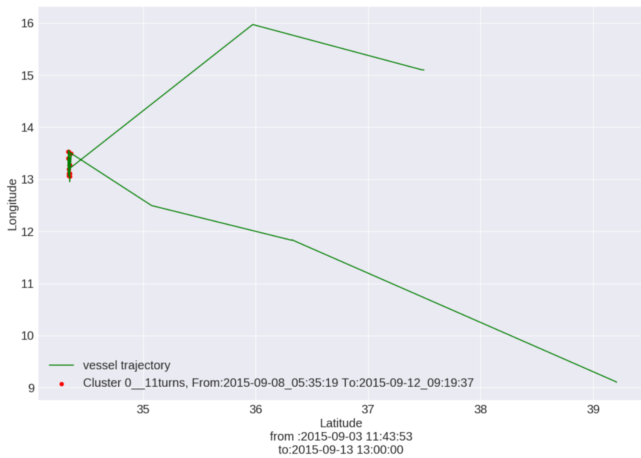
IV. PROPOSED APPROACH

The first step of the proposed approach is to reduce the amount of collected information by removing records that do not provide much information concerning the vessel trajectory. For this purpose, we applied the Ramer–Douglas–Peucker (RDP) algorithm [19], which simplifies a ship trajectory by removing intermediate points that slightly decline for the vessel course. This step significantly improved the processing time of our approach, without loosing important information. The algorithm is applied dynamically to the AIS signals collected for a vessel as soon as a new record arrives and decides whether it should be kept or not.

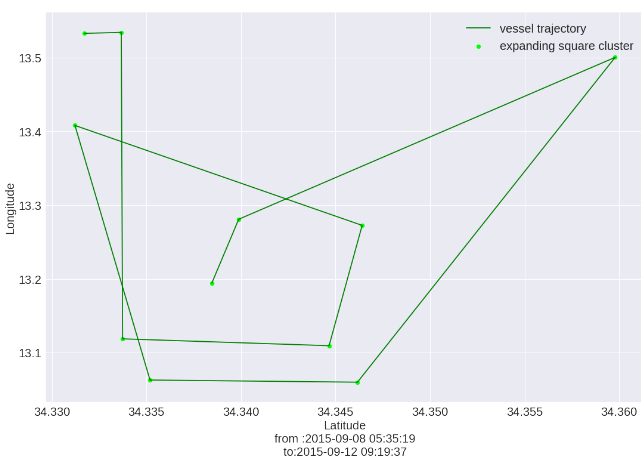
The second step is to extract the simple events of interest for a vessel. In our case, we consider only the major changes in the vessel course, which we call "turns". A turn is defined as a change in the course that is greater than 40°. Although it would probably be useful to keep also the points of acceleration or deceleration of the vessel, we omit them at this step and focus on turning points only. The data that we finally keep for a route is significantly smaller than the original data and the information they contain is much richer. Figure 3 presents the actual route of a vessel, the simplified route after applying the RDP algorithm and the turning points, which are used as input in the third step.

The third step aggregates multiple turn events into more complex events in order to detect interesting navigation patterns. For this purpose we applied a density based clustering algorithm to the turn points of a vessel's route. We implemented an incremental version of the Spatiotemporal DB-Scan (ST-DBScan) clustering algorithm [20] using a minimum range of 40 Km and a minimum number of 8 turns within this range (i.e. eps=40 and MinPts=8 as input parameters to DBScan). The incremental version allows us to update the clusters as long as we collect records for a vessel. Since, the departure and destination ports of a vessel, is among the features of our dataset, it is easy to detect when the vessel begins a trip and when it reach the destination port. Based on this information, the incremental clustering process is reset for each vessel, when it arrives to the arrival port.

The result of this process when it is applied to the route of a vessel is depicted in Figure 4a. A zoom on the detected cluster is presented in Figure 4b, where an expanding square movement pattern is depicted. The expanding square is a typical search pattern in SAR missions.



(a) A vessel's route, after detecting and clustering turns.



(b) A zoom on the cluster reveals the tactical movement of the vessel.

Figure 4. An example of a vessel that participates in a SAR mission.

The last step in a data clustering process is the labeling of clusters, which is usually done by abstracting the features of the individual instances that cluster together. For example, when data points (represented as feature vectors) are clustered together, the cluster centroid that averages the values in each feature can be used as the cluster label. In the case of document clusters, a set of representative (i.e. frequent, or with high tf-idf score) terms are selected to label the cluster. In our case, the clusters are group of points that cannot simply labeled by a representative point (e.g. the centroid), since they define a navigation pattern. So, the cluster label must describe the navigation pattern and in the case of SAR mission it can either be a simple annotation that distinguishes between SAR maneuvers and normal maneuvers, or a fine grained classification of the pattern to predefined types of SAR maneuvers [18]. Cluster labeling is part of our ongoing work and will be handled as a pattern classification problem. In this work, we show some indicative navigation patterns that are related with SAR missions, which have been manually identified in the dataset.

Another labeling problem that arises from this work, is the labeling of the whole vessel route that is probably engaged in a SAR mission. It is evident from this primary analysis, that vessels that participate actively in such missions, especially in rescue missions after they have received a notification, follow a very specific navigation route: they head to the rescue point, they perform several maneuvers and they head back to a port. This can also be treated as a classification problem that takes as input the whole route information.

V. Results

A. *Detection of SAR navigation patterns*

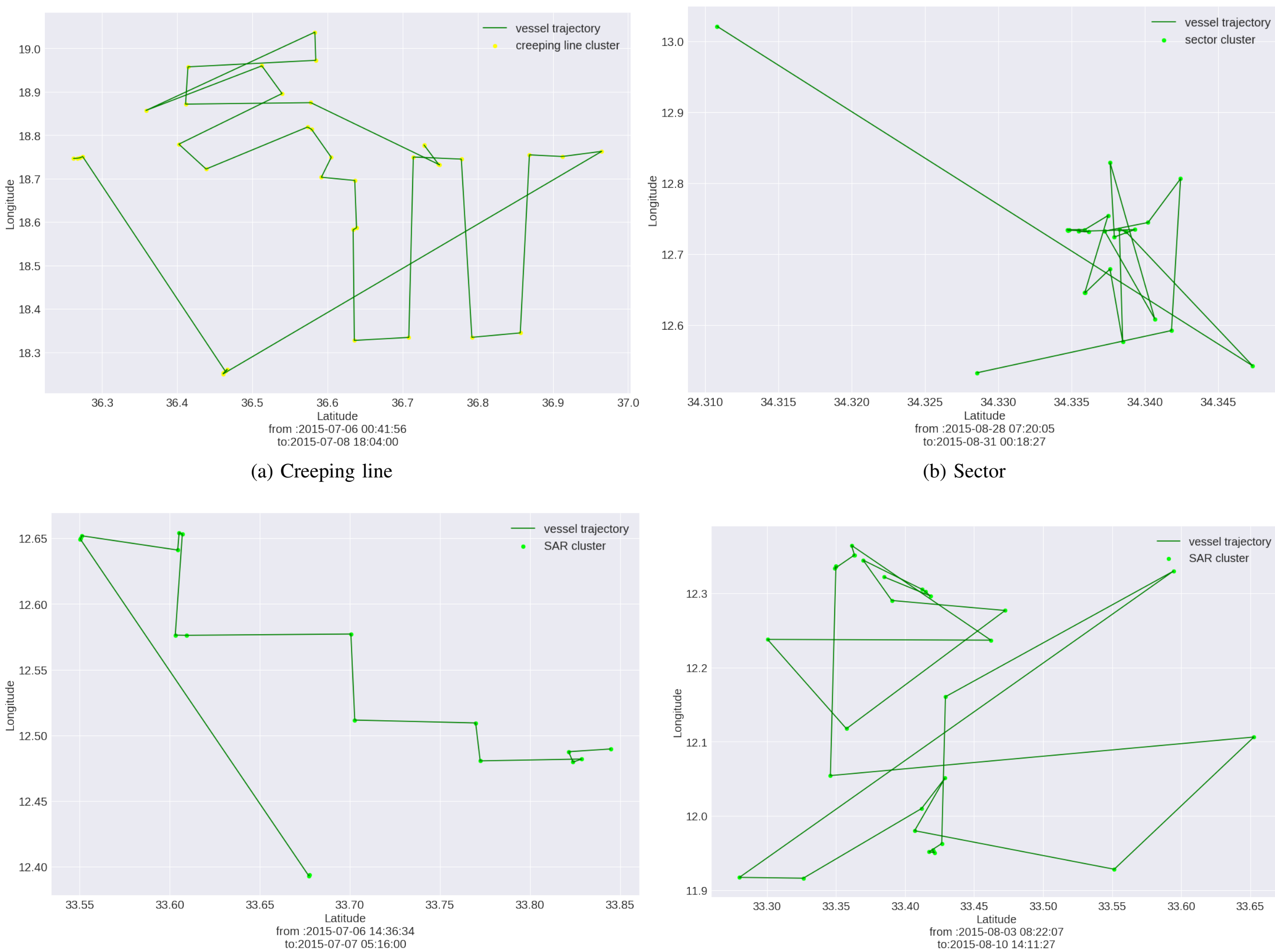
By applying the proposed method to the AIS data of all the 25 ships for the 3 months period we detected several clusters of interest, and in them we have noticed all the main search (i.e. Expanding square at Figure 4b, Creeping line at Figure 5a, Sector at Figure 5b) and rescue (see Figures 5c and 5d) navigation patterns for SAR missions as depicted in Figure 5.

A second finding from the analysis of the dataset is that a vessel's route during a Search and Rescue mission, is similar to the one depicted in Figure 4a, where the vessel moves to a specific area, performs a specific search or rescue movement and then returns to the destination port. This results in a single cluster for the whole route, which corresponds to a specific maneuver type. In addition to this, the navigation patterns before, during and after this maneuver differ: the average speed before and after the maneuver are much higher than the speed during the maneuver. The distance and area covered also differ.

B. *Multiple vessel analysis and Validation of events*

In order to validate our primary findings, we performed a spatio-temporal query in the annotated dataset, which comprises only the cluster points detected for all the vessels. The query retrieved clusters that partially overlap (in the area they cover) within the same day, which means that more than one vessels have performed SAR movements in the same day in the same area.

Unfortunately, we retrieved only one incident for the three months period, in which two of the five cargo vessels have taken part. It was on the night from the 5th to the 6th of August, where two of the ships have participated in a SAR



(a) Creeping line

(b) Sector

(c) Track crawl

(d) Rescue

Figure 5. Different types of ship movement (maneuvers) during a SAR mission

mission. This information was verified by a search in the news of that period, giving us evidence that our approach is working.

Although we had detected only one event that engaged more than one vessel, we verified many of the clusters that we detected by searching on the news for the specific dates and ships. We also detect additional clusters, which may correspond to SAR movements but have not been reported in the news.

Finally, it is important to notice that in some of the routes of the 20 randomly selected ships, our method detected dense clusters of turns. However, in all these clusters the vessels were not following a particular navigation pattern, but rather moved in a random way. In some cases, the vessels were anchored in the port or in a place near the port and in other cases they were moving slowly to their destination (e.g. the sailing boats).

VI. Conclusions and next steps

From an application point of view, we investigate the problem of applying semantic meaning to located events. Extensive work has been already contacted in identifying concrete rules that can categorize an observed moving pattern into simple events. CEP provides specific rules for identifying a turn or a stop event just by looking at the latest two positions of a vessel and calculating its movement vector, which is the speed and direction of the movement.

Our aim is to apply semantic meaning not only on simple events such as these but on more complex events that relate to the ”SAR operation” or even vessel in distress. This is a challenge that can be tackled employing two options, either training a classification or clustering algorithm in the available data and leaving the algorithm to judge if a recent behavior is similar to a known behavior or by composing complex events from a series of simple events.

This work emphasizes on the first option as it will allow us to identify baseline patterns that could feed the second. Our immediate future work involves experimentation with CEP techniques for the same objectives.

As a result, the analysis that we performed so far, allowed us to detect basic navigation patterns that possible relate to SAR missions. It also allowed us to detect the beginning and end of a route and understand whether it contains a SAR movement pattern or not. Finally, we are in the process of generating a description for the navigation of a vessel, before, during and after a cluster, which comprises speed, area and distance covered and vessel type information in an attempt to classify a vessel route as a SAR or normal route.

VII. Acknowledgements

This work has been developed in the frame of the MASTER project, which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777695

This work was supported in part by MarineTraffic which provided data access for research purposes.

References

[1] Automatic identification system. Accessed: 2010-09-30. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Automatic_identification_system&action=history

[2] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.

[3] B. Ristic, B. La Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in ais data: Anomaly detection and motion prediction," in *information fusion, 2008 11th international conference on*. IEEE, 2008, pp. 1–7.

[4] S. Trevisanut, "Search and rescue operations in the mediterranean: factor of cooperation or conflict?" *The International Journal of Marine and Coastal Law*, vol. 25, no. 4, pp. 523–542, 2010.

[5] V. Avram, U. Glässer, and H. Y. Shahir, "Anomaly detection in spatiotemporal data in the maritime domain," in *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference on*. IEEE, 2012, pp. 147–149.

[6] N. Zorbas, D. Zissis, K. Tserpes, and D. Anagnostopoulos, "Predicting object trajectories from high-speed streaming data," in *Trustcom/BigDataSE/ISPA, 2015 IEEE*, vol. 2. IEEE, 2015, pp. 229–234.

[7] A. Valsamis, K. Tserpes, D. Zissis, D. Anagnostopoulos, and T. Varvarigou, "Employing traditional machine learning algorithms for big data streams analysis: the case of object trajectory prediction," vol. 127, pp. 249–257.

[8] N. Le Guillarme and X. Lerouvreur, "Unsupervised extraction of knowledge from s-ais data for maritime situational awareness," in *Information Fusion (FUSION), 2013 16th International Conference on*. IEEE, 2013, pp. 2025–2032.

[9] C. D. Rogers, "System and apparatus for automatic and continuous monitoring, proactive warning and control of one or more independently operated vessels," May 16 2006, uS Patent 7,047,114.

[10] D. Gentle, S. Cameron, and T. Kennedy, "Comparative ice drift and tow model analysis for target marine structure," Jan. 7 2016, uS Patent App. 14/741,455.

[11] F. Mazzarella, M. Vespe, D. Damalas, and G. Osio, "Discovering vessel activities at sea using ais data: Mapping of fishing footprints," in *Information Fusion (FUSION), 2014 17th International Conference on*. IEEE, 2014, pp. 1–7.

[12] E. N. de Souza, K. Boerder, S. Matwin, and B. Worm, "Improving fishing pattern detection from satellite ais using data mining and machine learning," *PloS one*, vol. 11, no. 7, p. e0158248, 2016.

[13] A. Buchmann and B. Koldehofe, "Complex event processing," *IT-Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik*, vol. 51, no. 5, pp. 241–242, 2009.

[14] S. Gillani, A. Zimmermann, G. Picard, and F. Laforest, "A query language for semantic complex event processing: Syntax, semantics and implementation."

[15] A. Artikis, M. Sergot, and G. Paliouras, "An event calculus for event recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 4, pp. 895–908, 2015.

[16] K. Patroumpas, E. Alevizos, A. Artikis, M. Vodas, N. Pelekis, and Y. Theodoridis, "Online event recognition from moving vessel trajectories," *GeoInformatica*, vol. 21, no. 2, pp. 389–427, 2017.

[17] N. Giatrakos, A. Artikis, A. Deligiannakis, and M. Garofalakis, "Complex event recognition in the big data era," *Proceedings of the VLDB Endowment*, vol. 10, no. 12, pp. 1996–1999, 2017.

[18] International Maritime Organization, *IAMSAR manual : International Aeronautical and Maritime Search and Rescue manual*. IMO, London, 2008.

[19] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.

[20] D. Birant and A. Kut, "St-dbscan: An algorithm for clustering spatial–temporal data," *Data & Knowledge Engineering*, vol. 60, no. 1, pp. 208–221, 2007.