Introducing the Pattern of Life (PoL) Concept for Maritime Traffic

Alya Itani*, Cyril Ray*

*Naval Academy Research Institute, 29240 Brest Cedex 9, France
{alya.itani, cyril.ray}@ecole-navale.fr,
https://www.ecole-navale.fr

Résumé. Pattern of Life (PoL) is a common concept based on understanding the behaviour of a certain moving object, through analysing its past and present navigational patterns, for the purpose of potentially predicting its future actions. In this paper, we introduce the Pattern of Life concept to the domain of maritime navigation traffic. The novelty of this concept lies in representing vessels navigational patterns at an aggregated level through spatio-temporal directed networks in order to simplify the post-related data mining. The idea behind this method is to achieve the existing domain research goals but with a certain level of abstraction, we mention of them: analysis and prediction of trajectories, anomaly detection, exposing unlawful activity at sea, etc.

1 Introduction

There is a world-wide increasing interest in understanding maritime navigation traffic, especially with it being the most used mean of global trade. Moreover, there is a considerable rise in navigation exploitable data with the increasing installation of positioning report systems on board, e.g. Automatic Identification Systems (AIS), of various vessel types (Vespe et al., 2012). In addition to AIS there exists other available information sources such as Synthetic Aperture Radar (SAR) images and coastal radars. Moreover, environmental, contextual and geographical data became easily attainable nowadays (Ray et al., 2018). The aggregation and analysis of rather complementary and correlative data empowers greatly the research on maritime situational awareness. Indeed, there is an expeditiously progressing research on predicting vessel trajectories, detecting navigational anomalies, detecting illegal human activities and much more related topics (Hexeberg et al., 2017). All prior research endeavours addressed these challenges through analyzing instant positioning data of vessels, which is showing promising progress. However, in our proposed Pattern of Life (PoL) we aim at re-implementing prior analysis in a simplified form. Hence, the simplification and novelty of our proposed concept PoL lies in representing the navigational life patterns of multiple vessels as connected spatio-temporal networks.

This paper describes a preliminary introduction to the PoL concept setting the bedrock to the current work in progress. The idea of resuming the movement of objects into a sequential pattern of positions has been addressed and applied on various domains, even though the naming changes (Giannotti et al., 2007; Do et Gatica-Perez, 2014; Parent et al., 2013). However, the idea behind this proposed concept is to construct and analyse a semantic graph that represents the spatio-temporal connection between maritime route sequences in the intention relatively optimising the incorporated analysis. The paper is organised as follows, section 2, provides insights on the project of which this work endeavour belongs. In section 3, we offer a briefing on existing related research Whereas, section 4 introduce the PoL concept in general and the specifics of its different phases. Finally, we conclude the proposed study concept and mention the main future perspectives in section 5.

2 Towards a European Information Sharing Environment

European Commission and EU/EEA member states have accomplished efforts in designing and implementing maritime information systems. Vessel movements have therefore become available to surveillance authorities through the use of several automated technologies. Merged and analysed, together with a large variety of contextual information of different natures and types, positioning data can be used to monitor maritime mobility, uncover maritime activities, illegal trafficking or risks for the environment, living resources and the navigation. While lots of maritime data are nowadays available, methodologies and processing architectures for data exchanges, standards to harmonise data and meta-data description are still missing at the European scale. Moreover, ensuring interoperability of national legacy big data systems is still a problem. To address these issues at European level, a Common Information Sharing Environment (CISE) has been initiated and is currently being developed jointly by the European Commission and EU/EEA member states (EU, 2010).

CISE is a voluntary collaborative process in the European Union to further promote and enhance relevant data and information sharing between authorities involved in maritime surveillance at the horizon of the year 2020. It will integrate existing surveillance systems and networks and will provide to all concerned authorities an access to the information they need for their missions [1]. CISE will make different national legacy systems interoperable so that data and other information can be exchanged easily. This common environment first relies on a shared architectural principle which connects legacy systems to CISE core architecture though adaptors and gateways. Secondly it relies on a data model providing a standardised description of maritime data thus supporting interoperability of European information systems. Each gateway can be developed together with operational functionalities and services supporting operational missions of EU/EEA member states.

The main objective of the French CISE project (CISE DMS, started in January 2018 for two years) and has therefore proposed several Data Mining Services (DMS) on top of CISE architecture and data model, thus offering to the relevant maritime administrations (of member states) services to analyse mobility patterns, anomalies or contradictory information. The data mining service discussed in this paper will be integrated within the national gateway, connected to the CISE architecture and will rely on data coming from national legacy information systems.

Maritime positioning data mining covers a large spectrum of interesting use cases. In order to match concrete needs, a limited number of possible services have been defined by the CISE

1. EU CISE 2020. www.eucise2020.eu, accessed 29.11.2018

project managers and proposed to maritime experts. That initial set of 13 mining services was grouped in 5 categories : understanding and prediction of trajectories, management and control of ships identities, monitoring of areas of interest, environmental protection, detection and reporting of risk events including hacking events (Ray et al., 2015). Finally, after various exchanges with experts and feedbacks from them (coast guards, maritime affairs, ...), this list converged to 5 main services to be studied and implemented. These mining services are :

— Ship motion prediction ;
— AIS and Earth observation coupling ;
— Nominative databases correlation ;
— Pattern of life ;
— Extraction and monitoring of areas of interest.

In the scope of this paper, we address the Pattern of Life use case as part of the main 5 CISE : DMS services.

3 Related Literature

There has been a considerable rise in maritime navigation exploitable data, especially with the introduced legal obligation of installing positioning report systems on board of vessels (Pallotta et al., 2013). This information flood is orienting all domain experts, from authorities to industrial organisations and research centres, towards the necessity of maritime traffic analysis and surveillance. Nowadays, in addition to the available nominative and positioning information broadcast by collaborative systems (e.g., AIS, VMS, LRIT), other information related to environmental, contextual and geographical data are also at hand and can be aggregated as complementary and correlative sources to the exploitable data (Ray et al., 2018). There exists plenty of research topics in the literature related to the domain of maritime navigation traffic analysis, we categories them as follows :

— Vessels trajectory analysis and prediction : This mainly includes research on analysing AIS and other complementary sources to predict trajectories of vessels. Various predictive analysis techniques where deployed for that matter (Borkowski, 2017; Gan et al., 2016; Xu et al., 2011).
— Anomaly detection : This work endeavour encompasses mostly clustering of trajectories to distinguish any irregularity in behaviour. This is often used to detect threats or entry to restricted areas (Roberts, 2014; Liu et al., 2015; Lane et al., 2010; Soleimani et al., 2015).
— Analysing human activity at the sea : Such as fishing activity, illegal traffic (drugs, refugees, goods), piracy, etc. (Silveira et al., 2013; Petrossian, 2015).
— Port traffic forecasting : This includes port volume handling forecast, cargo throughput forecast (Jugović et al., 2011).
— Collision prevention : analysing risk of ships collisions through analyzing ships navigational behaviour (Johansen et al., 2016; Bukhari et al., 2013).
— Multi-source information fusion : it involves fusion of multiple data sources (AIS, radar, satellite, sensors) for solving contradiction, redundancy, and uncertainty in available data. This can be helpful in achieving better accuracy in analysis (for example surveillance) Guerriero et al. (2008); Mazzarella et al. (2013); Roy et Bosse (1998); Battistello et Koch (2010).

maritime navigational patterns of life

The related research topics are diverse and expeditiously progressing. Many analysis techniques are applied to attain the above major goals using maritime navigational data. Nevertheless, in the Pattern of life case study we will be addressing the first four main topics with the possibility of extending the work in the future to cover broader areas. The common aspect of the above addressed research motivations is analysing the data on a micro level to achieve the expected needs. Pattern of Life (PoL) comes from the concept of understanding and monitoring the behavioural patterns of a certain subject (in this case ships) to analyse and use for potentially predicting its future actions. In the maritime domain PoL addresses the understanding of navigation data at an aggregated (meso, macro) level.

In simpler words, all prior research concentrates on semi instant positioning data of ships which gets the goal attained but, in our opinion, stays limited to short term predictions and can be vulnerable to noise in data. Whereas the vision we have in mind includes a more abstract representation of ship routes.

4 The Pattern of Life Concept

In this section we introduce the PoL concept's flowchart with its composing phases. Figure 1 provides a simplified illustration of the PoL concept. In the following subsections we describe each phase briefly.

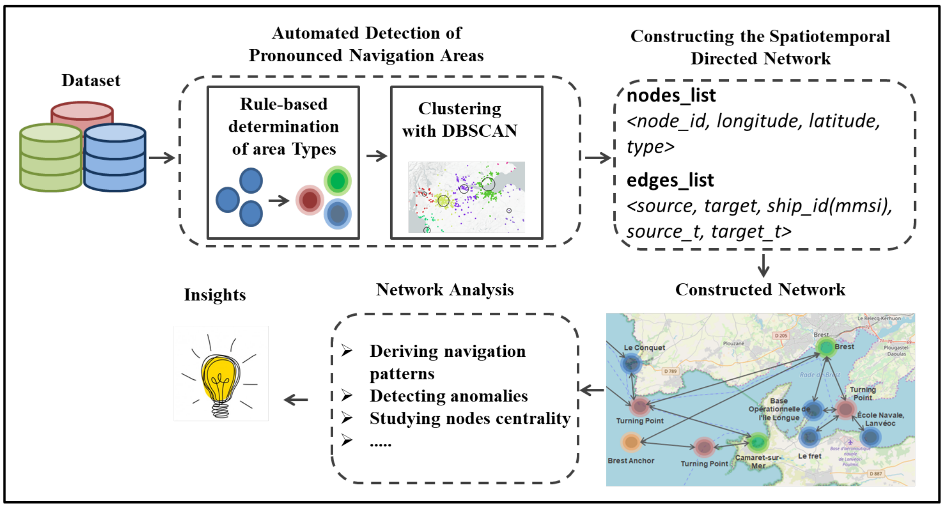

FIG. 1 – *The PoL Concept Flowchart*

4.1 Introducing the Dataset

The concerned dataset of this research study includes historical traces of maritime vessels collected through the AIS, created together with correlated data spatially and temporally aligned. The incorporated data characterises the vessels, the areas were they navigate, and the situation at sea (Ray et al., 2018). The information contained in this dataset can be grouped into four categories : navigation data (AIS receiver collected data of vessel positions, geographic data (cartographic, topographic or regulatory context of vessel navigation), and environmental data (climatic and sea state related information) It covers a time period of six months, from October 1^{st}, 2015 to March 31^{st}, 2016 and provides ship positions over the Celtic sea, the North Atlantic ocean, the English channel, and the Bay of Biscay (France). Moreover, its is possible to enrich the current dataset with supplementary data sources such as nautical charts. Nautical charts are simply a rich form of maps that are very useful in navigation aid. They includes supplementary information such as depth of water, coastline details, heights of land, tides, harbours, etc. These charts are not typically open source (they are licensed), however, to work around this issue a technical descriptive document of these charts is made available online (C.Ray, 2018). Table 1 describes the specifications of the heterogeneous data set. Also, other statistical information on the intervention of the CROSS (Regional Operational Centers of Surveillance and Rescue) rescue can a be used to enrich the current dataset [2].

Specifications Table

Subject area	Maritime Mobility and Transportation
Specification of subject area	Intelligence, Surveillance, and Reconnaissance of vessels' navigation
Type of data	Positions of ships, geographical, contextual and environmental data related to maritime navigation.
Acquiring the dataset	The dataset was created by combining publicly available datasets and data from AIS (Class A SAAB R4)
Data format	Comma-Separated Values (CSV), ESRI Shapefile
Experimental factors	Spatial, temporal and message_based filtering of received data
Data source location	Celtic Sea, North Atlantic Ocean, English Channel, Bey of Biscay (France)
Data accessibility	Heterogeneous Integrated Dataset for Maritime Intelligence, Surveillance, and Reconnaissance (HIDMISR) Identifier : doi ; 10.5281/zenodo.1167594 Usage rights : Creative Commons Attribution-Non Commercial-Share Alike 4.0 International (CC BY-NC-SA 4.0)

TAB. 1 – *Dataset Specifications Table*

4.2 Automated Detection of Pronounced Navigation Areas

In this phase, we intend to automate the detection of all main components of a vessel's route. These components include any pronounced area in a vessel's PoL (example : ports,

2. https ://carte.snosan.fr, accessed 30/11/2018

anchorage areas, turning points, etc.). Table 2 illustrates the initial list of possible pronounced areas, which will represent the nodes of our PoL network.

Area Categories	**Description**
Ports	Areas where vessels are usually anchored or moored for either professional or leisure purposes
Anchorage Areas	Areas where ships are at anchor usually at proximity of a port
Waiting Areas	Areas where cargo vessels wait until an available spot is open for them
Fishing Areas	Areas where fishing boats navigate with low speed
Offshore Infrastructures	Such as wind, gas, electricity, or oil platforms
Rendez-vousing Areas	An area where vessels stop to meet for a period of time
Turning Points	Areas where vessels usually change their direction (Course Over Ground)

TAB. 2 – *Description of Pronounced Area Types*

First, with the help of domain experts, we declare the rules that define each area category. As example a port can be defined as a stationary area with at least X stationary vessels a day, also a vessels's turning point is defined by a repeated change in vessels CoG (Course over Ground) with Y degrees, etc. The choice of X and Y and other rule parameters are either data-driven or selected through experts recommendation. Then, we train a density based clustering algorithm to detect these areas in an unsupervised manner, we consider using DBscan for being one of the most convenient clustering algorithm for geospatial datasets (Li et al., 2016).

There exists in the literature some applied research on unsupervised port and stationary areas detection. In their paper (Millefiori et al., 2016), Millefiori *et al.* applied a data driven approach to defining seaport's extended area of operation based on AIS collected data. They used kernel density estimate (KDE) algorithm. Vespe *et al.* (Vespe et al., 2016) also used density estimates on AIS data to map fishing activities on the scale of Europe. Furthermore, Pallotta *et al.* introduced the TREAD (Traffic Route Extraction and Anomaly Detection) methodology used for anomaly detection and route prediction (Pallotta et al., 2013). In the route extraction fold of their work they apply a data driven approach for defining route components including ports and turning points. In this context we re-implement this concept of data driven route component area detection (stationary areas and turning points) as a bedrock to our pattern of life network representation of routes.

Figure 2 illustrates a comparison between different derived zones within the Bay of Brest in France. The dotted bright yellow colour is of zones defined by the official nautical chart for maritime navigation. The green zones are those derived by the TREAD method for comparison purposes, however it is important to mention that the TREAD zones where computed at a different time-span than that of all other zones. Moreover, the violet marked zones illustrate the stationary areas of cargo ships, while the blue coloured zones are stationary areas of passengers transportation boats.

After defining all pronounced navigation areas in the previous phase, the initial dataset is then used to construct a routes network that connects these areas. The network should re-

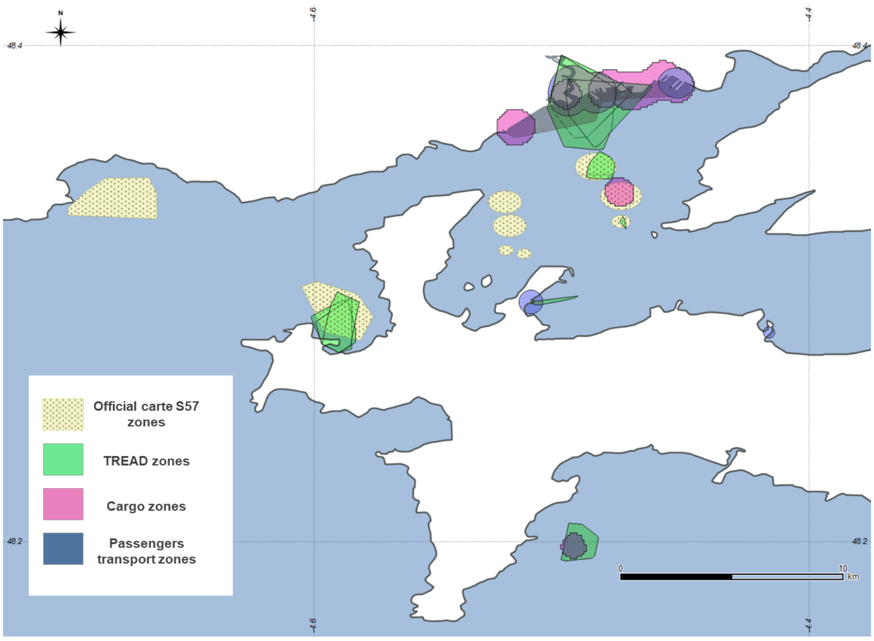

FIG. 2 – *Stationary Zones in the Bay of Brest (TREAD areas computed on different timespan than other areas)*

present the spatio-temporal information accompanied with usual routes. In order to do that we represent this network with two (data frames) CSV type files, one holds information on nodes and the other on edges. The structure of each data frame and its columns is as follows :

nodes_list : $< node\_id, longitude, latitude, node\_type >$

edges_list : $< source, target, ship\_id(mmsi), source\_time, target\_time >$

Hence the nodes hold information about the DBscan estimated areas (there position, their type, possibly name, etc.). Whereas the edges represent a directed route passed by various vessels. The edge will hold weight information annotating the number of times the route has been passed. Moreover inspecting a specific edge the user should be able to have a statistical chart showing the distributed percentages of this taken route relative to different ship types. Moreover the user should be able to display routes specific to a certain period of time ot to a certain ship ID (MMSI). The proposed representation of the network components allows including the temporal and spatial aspects to routes which in return permits the long-term prediction of routes. It also permits analysing centrality of ports, identifying repeated route patterns by similar typed ships and correspondingly detecting anomalies to these patterns, etc. Moreover a dynamic web-based application is currently under development, allowing the user to manipulation and inspect this network, through dynamically visualising routes at specified times, filtering routes by ship type, visualising statistics on specified routes, etc. It is important to note that the proposed route network vision is still preliminary and will be optimised upon implementation following the needs of domain experts.

4.2.1 Local and Global Implementation Scales

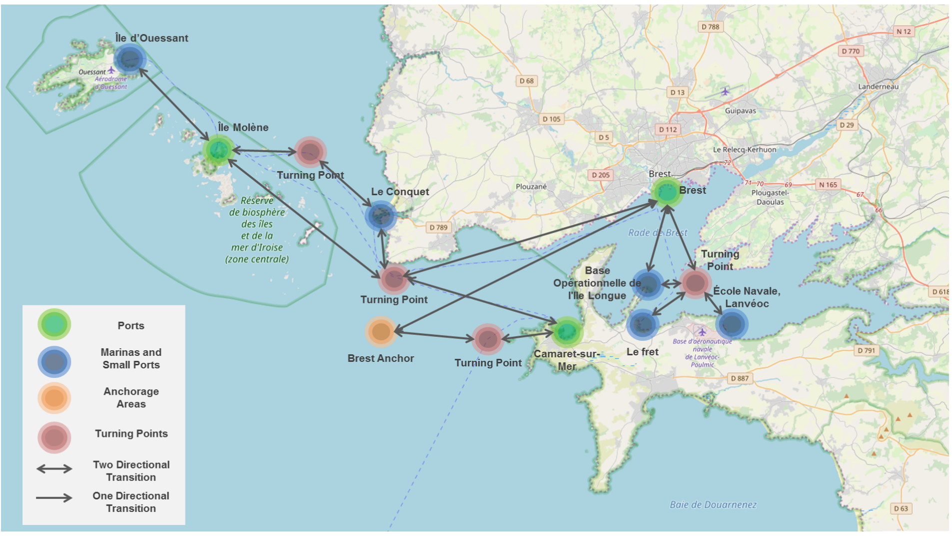

FIG. 3 – *Pattern of Life (PoL) Network of the Brest Bay Region*

There are two main steps in the PoL implementation phase : the local and the global implementation. The reason of this separation comes from the idea of training and validation. It is notable that the stationary and turning points of vessels and their navigational behaviour in general vary significantly with the varied types of ships and areas of navigation understudy. Therefore we find that training the pronounced detection algorithms and constructing the PoL on a local well known maritime scale, at first, would allow us to realistically and closely asses and justify the outcome of our proposed approach.

Taking for example the Bay of Brest, the maritime navigation of this Bay is very well known to implicated experts in the area. It includes for example repetitive well organised passenger transportation routes that can be used at this step to asses and tune the algorithms. Moreover, the described public dataset at hand, introduced in 4.1 incorporates the Brest area within its covered scale making the proposition attainable. In that context, figure 3 illustrates the PoL concept around the Bay of Brest region, it offer a display of the preliminary network with its different node types and directed edges. Refer to table 2 for a better understanding of the different displayed node types.

Once the local implementation is complete, the concept can be tested on a larger scale for validation. Any necessary tuning of parameters would be applied at this stage. The testing phase will be either on the scale of Europe or worldwide depending on the availability of the corresponding data.

4.3 Analysing the Network

After the network is constructed, it becomes possible to apply standard graph analysis techniques to mine the network for all possible insights. This can include studying the centrality of ports, predicting vessels' trajectories, identifying anomalies, rendez-vous, suspicious behaviours (drug or human trafficking). Moreover standing maritime trajectory analysis methods can be re-attempted on a more abstract level of predicting next stop or destination rather than future positions in the next period of time.

5 Conclusion

We describe in this paper the research motivation towards the definition and implementation of a Pattern of Life concept applied to the domain of maritime navigation as part of the CISE European project. The intention behind the proposed concept is to represent maritime mobility routes as spatio-temporal directed networks at an aggregated level, in order to facilitate and insure the post-related data mining services. We proposed a two step implementation phase at a local and then global scale to include a realistic aspect of assessment to the proposed approach. Although, the concept is in its initial phase still, it catches closely the attention of operational experts (coast guards, maritime affairs, ...) and is considered among the most prioritised identified use-cases.

Future work on the topic includes applying graph analysis technique to mine the network for predicting vessels' trajectories, identifying anomalies, vessels rendez-vousing zones, suspicious behaviours (drug or human trafficking), etc.

Acknowledgement

This work belongs to CISE DMS project, which has received funding from the COSME program of the European Union under grant agreement Number EASME/EMFF/2016/1.2.1.9 /03/SI2.763845 - CISE DMS.

Références

Battistello, G. et W. Koch (2010). Knowledge-aided multi-sensor data processing for maritime surveillance. In *GI Jahrestagung (2)*, pp. 796–799.

Borkowski, P. (2017). The ship movement trajectory prediction algorithm using navigational data fusion. *Sensors 17*(6), 1432.

Bukhari, A. C., I. Tusseyeva, Y.-G. Kim, et al. (2013). An intelligent real-time multi-vessel collision risk assessment system from vts view point based on fuzzy inference system. *Expert systems with applications 40*(4), 1220–1230.

C.Ray, A. (2018). Integration of nautical charts in maritime dataset. Technical note, datAcron H2020.

Do, T. M. T. et D. Gatica-Perez (2014). The places of our lives: Visiting patterns and automatic labeling from longitudinal smartphone data. *IEEE Transactions on Mobile Computing 13*(3), 638–648.

EU (2010). Roadmap towards establishing the common information sharing environment for the surveillance of the eu maritime domain. Com/2010/0584 final, European Commission.

Gan, S., S. Liang, K. Li, J. Deng, et T. Cheng (2016). Ship trajectory prediction for intelligent traffic management using clustering and ann. In *Control (CONTROL), 2016 UKACC 11th International Conference on*, pp. 1–6. IEEE.

Giannotti, F., M. Nanni, F. Pinelli, et D. Pedreschi (2007). Trajectory pattern mining. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 330–339. ACM.

Guerriero, M., P. Willett, S. Coraluppi, et C. Carthel (2008). Radar/ais data fusion and sar tasking for maritime surveillance. In *Information Fusion, 2008 11th International Conference on*, pp. 1–5. IEEE.

Hexeberg, S., A. L. Flåten, E. F. Brekke, et al. (2017). Ais-based vessel trajectory prediction. In *Information Fusion (Fusion), 2017 20th International Conference on*, pp. 1–8. IEEE.

Johansen, T. A., A. Cristofaro, et T. Perez (2016). Ship collision avoidance using scenario-based model predictive control. *IFAC-PapersOnLine 49*(23), 14–21.

Jugović, A., S. Hess, et T. Poletan Jugović (2011). Traffic demand forecasting for port services. *Promet-Traffic&Transportation 23*(1), 59–69.

Lane, R. O., D. A. Nevell, S. D. Hayward, et T. W. Beaney (2010). Maritime anomaly detection and threat assessment. In *Information Fusion (FUSION), 2010 13th Conference on*, pp. 1–8. IEEE.

Li, S., S. Dragicevic, F. A. Castro, M. Sester, S. Winter, A. Coltekin, C. Pettit, B. Jiang, J. Haworth, A. Stein, et al. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS journal of Photogrammetry and Remote Sensing 115*, 119–133.

Liu, B., E. N. de Souza, C. Hilliard, et S. Matwin (2015). Ship movement anomaly detection using specialized distance measures. In *Information Fusion (Fusion), 2015 18th International Conference on*, pp. 1113–1120. IEEE.

Mazzarella, F., A. Alessandrini, H. Greidanus, M. Alvarez, P. Argentieri, D. Nappo, et L. Ziemba (2013). Data fusion for wide-area maritime surveillance. In *Proceedings of the COST MOVE Workshop on Moving Objects at Sea, Brest, France*, pp. 27–28.

Millefiori, L. M., D. Zissis, L. Cazzanti, et G. Arcieri (2016). Scalable and distributed sea port operational areas estimation from ais data. In *Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on*, pp. 374–381. IEEE.

Pallotta, G., M. Vespe, et K. Bryan (2013). Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction. *Entropy 15*(6), 2218–2245.

Parent, C., S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, et al. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR) 45*(4), 42.

Petrossian, G. A. (2015). Preventing illegal, unreported and unregulated (iuu) fishing: A situational approach. *Biological Conservation 189*, 39–48.

Ray, C., R. Dréo, E. Camossi, A.-L. Jousselme, et C. Iphar (2018). Heterogeneous integrated dataset for maritime intelligence, surveillance, and reconnaissance. *Data In Brief , Accepted for publication*.

Ray, C., R. Gallen, C. Iphar, A. Napoli, et A. Bouju (2015). Deais project: Detection of ais spoofing and resulting risks. In *OCEANS 2015 - Genova*, pp. 1–6.

Roberts, S. (2014). *Anomaly detection in vessel track data*. Ph. D. thesis, Oxford University, UK.

Roy, J. et E. Bosse (1998). Sensor integration, management and data fusion concepts in a naval command and control perspective. Technical report, Defence Research Establishment Valcartier (Québec).

Silveira, P., A. Teixeira, et C. G. Soares (2013). Use of ais data to characterise marine traffic patterns and ship collision risk off the coast of portugal. *The Journal of Navigation 66*(6), 879–898.

Soleimani, B. H., E. N. De Souza, C. Hilliard, et S. Matwin (2015). Anomaly detection in maritime data based on geometrical analysis of trajectories. In *Information Fusion (Fusion), 2015 18th International Conference on*, pp. 1100–1105. IEEE.

Vespe, M., M. Gibin, A. Alessandrini, F. Natale, F. Mazzarella, et G. C. Osio (2016). Mapping eu fishing activities using ship tracking data. *Journal of Maps 12*(sup1), 520–525.

Vespe, M., I. Visentini, K. Bryan, et P. Braca (2012). Unsupervised learning of maritime traffic patterns for anomaly detection.

Xu, T., X. Liu, et X. Yang (2011). Ship trajectory online prediction based on bp neural network algorithm. In *Information Technology, Computer Engineering and Management Sciences (ICM), 2011 International Conference on*, Volume 1, pp. 103–106. IEEE.