THE JOURNAL OF NAVIGATION (2017), **70**, 847–858. © The Royal Institute of Navigation 2017
doi:10.1017/S0373463317000066

Study of Automatic Anomalous Behaviour Detection Techniques for Maritime Vessels

Abdoulaye Sidibé[1] and Gao Shu[1,2]

[1](*School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430063, China*)
[2](*Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan 430063, China*)
(E-mail: abdoul@bk.ru)

The maritime domain is the most utilised environment for bulk transportation, making maritime safety and security an important concern. A major aspect of maritime safety and security is maritime situational awareness. To achieve effective maritime situational awareness, recently many efforts have been made in automatic anomalous maritime vessel movement behaviour detection based on movement data provided by the Automatic Identification System (AIS). In this paper we present a review of state-of-the-art automatic anomalous maritime vessel behaviour detection techniques based on AIS movement data. First, we categorise some approaches proposed in the period 2011 to 2016 to automatically detect anomalous maritime vessel behaviour into distinct categories including statistical, machine learning and data mining, and provide an overview of them. Then we discuss some issues related to the proposed approaches and identify the trend in automatic detection of anomalous maritime vessel behaviour.

KEYWORDS

1. Anomalous vessel behaviour. 2. Automatic Identification System (AIS). 3. Vessel movement behaviour.

Submitted: 14 October 2016. Accepted: 12 January 2017. First published online: 8 March 2017.

1. INTRODUCTION. The maritime domain is the most utilised environment for bulk transportation, making maritime safety and security an important concern. Maritime safety and security is achieved via Maritime Situational Awareness (MSA), which is supported by surveillance and tracking systems such as the Automatic Identification System (AIS). AIS was conceived primarily as a navigational safety tool for collision avoidance and for supporting vessel traffic services in ports and harbours. Vessels equipped with AIS systems periodically exchange messages with nearby vessels and AIS base stations to enhance tracking and monitoring vessel movements. Collected sequences of AIS messages form vessel movement data or vessel trajectory data. AIS vessel movement data can be used to

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

perform tasks including vessel movement prediction (Last et al., 2014), characterisation of marine traffic patterns (Silveira et al., 2013), Collision Avoidance (Tijardovic, 2009) and automatic anomalous vessel movement behaviour detection. Automatic anomalous vessel movement behaviour detection could support the operator of maritime surveillance systems in detecting suspicious situations or threats in the maritime domain and in taking appropriate action.

Recently many efforts have been made in automatic anomalous maritime vessel movement behaviour detection based on AIS data to achieve an effective maritime situational awareness. In this paper we present a review of the state-of-the-art automatic detection of anomalous maritime vessel behaviour based on AIS data. First, we categorise some approaches proposed between 2011 and 2016 for automatic anomalous detection of maritime vessel behaviour into distinct categories including statistical, machine learning and data mining, and others) and provide an overview. Then we will discuss some issues related to the proposed approaches and identify trends in automatic anomalous maritime vessel behaviour detection.

The remainder of the paper is organised as follows. Section 2 presents a summary of recent approaches proposed in scientific literature on automatic detection of anomalous maritime vessel behaviour. Analysis and discussion of the proposed approaches is presented in Section 3. Conclusions and future directions for automatic anomalous maritime vessel behaviour detection are reported in Section 4.

2. REVIEW OF RECENT RESEARCH WORKS ON AUTOMATIC ANOMALOUS MARITIME VESSEL BEHAVIOUR DETECTION.

In this section we review some recent research works on automatic anomalous maritime vessel behaviour detection.

Automatic anomalous maritime vessel behaviour detection consists of finding anomalous movement behaviour of vessels moving in maritime areas such as in harbours or in open sea using computational methods. Anomalous vessel behaviour refers to vessel movement behaviour that differs from the typical or normal vessel movement behaviour. For instance, anomalous vessel behaviour can refer to a sudden change in vessel kinematic behaviour (such as unusual speed or location), deviation from standard sea lanes, unexpected port arrivals, close approach, and zone entry (Lane et al., 2010). Anomalous vessel behaviour detection is one of the main tasks for MSA, allowing the detection of suspicious situations or threats in the maritime domain and taking appropriate action.

According to the literature, the approaches for anomalous vessel behaviour detection are divided into three categories: data-driven, knowledge-based, and hybrid.

2.1. *Data-driven approaches for anomalous vessel behaviour detection.*

Typically, data-driven approaches are composed of two phases: 1) vessel behaviour normalcy model learning, 2) matching new vessel motion data to the learned normalcy model for anomalous behaviour detection. In the first phase, a model representing normal vessel movement behaviour is learned from vessel historical movement data. In the second phase, new vessel movement data are matched to the learned normalcy model of vessel behaviour, considering any mismatch as anomalous behaviour.

Technically, data-driven approaches for automatic anomalous vessel behaviour detection can be divided into distinct categories of approaches: statistical, machine learning, data mining and other approaches.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

2.1.1. *Statistical approaches for anomalous vessel behaviour detection.* The main concept of any statistical anomaly detection method is: "*An anomaly is an observation which is suspected of being partially or wholly irrelevant because it is not generated by the stochastic model assumed*" (Anscombe and Guttman, 1960).

Statistical anomaly detection techniques are based on the assumption that normal data instances occur in high probability regions of a stochastic model, while anomalies occur in the low probability regions of the stochastic model (Chandola et al., 2009).

Statistical approaches to anomalous vessel behaviour detection fit a statistical model representing normal vessel behaviour to the given historical vessel movement data and then apply a statistical inference test to determine whether a new vessel observation belongs to this model or not. Observations that have a low probability of being generated from the learned model, based on the applied test statistic, are declared as anomalous behaviours. The given historical vessel movement data are assumed to be Independent and Identically Distributed (IID) observations, and reflect vessel normal behaviour.

Some recent research works have adopted statistical approaches to address automatic maritime vessel anomalous behaviour detection. These research works are summarised below.

Anneken et al. (2015) have used a non-simulative anomalies dataset based on real data, to provide an evaluation and comparison of maritime anomaly detection algorithms, specifically the Gaussian Mixture Model (GMM) and the Kernel Density Estimator (KDE). According to the result of their experiment, the rate of false alarm as well as the rate of anomalies which are not found is high for both algorithms.

Kowalska and Peel (2012) combined Gaussian Processes and an Active Learning approach to deal with anomalous vessel behaviour detection. In their approach, a model of normal vessel behaviour is created from historical AIS data using Gaussian Processes (GPs), and to deal with the high computational complexity in training, the authors used the Active Learning approach.

Smith et al. (2012; 2014a) have introduced an abnormality detection technique using a combination of Gaussian Processes, extreme value theory and divergence measurement, to identify anomalous vessel behaviour in both streaming and batch data.

Mascaro et al. (2014) have addressed the anomalous vessel behaviour detection issue using Bayesian Networks. By training Bayesian Networks with real world AIS data combined with additional real world contextual information such as weather and time, as well as vessel interactions, they have taught two Bayesian Network-based models at two different time scales, namely in the form of the time series and track summary models. As a training tool, the authors used a machine learner named CaMML (Causal discovery via Minimum Message Length). The authors have found that adding further real world attributes helped to create a better model of normalcy.

To preserve port safety by classifying the movements of different ships in the scene, Castaldo et al. (2015) have applied Bayesian Networks to the analysis of vessel behaviour and ship-to-ship interactions in port areas.

2.1.2. *Machine learning and data mining approaches to anomalous vessel behaviour detection.* The main concept of the use of machine learning or data mining approaches for automatic anomalous maritime vessel behaviour detection is to create a model of normal vessel behaviour by applying machine learning or data mining techniques (such as unsupervised or supervised clustering algorithms) to vessel movement data. This normal

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

model is used to detect anomalous vessels, which are vessels that deviate from the learned normal model.

Depending on the format of input data, machine learning and data mining approaches for anomalous vessel behaviour detection can be divided into two main classes, namely point-based anomalous behaviour detection approaches and trajectory-based anomalous behaviour detection approaches. Point-based anomalous behaviour detection approaches highlight individual anomalous points while trajectory-based anomalous behaviour detection approaches, receive as their input entire vessel trajectories and automatically classify them as anomalous or not (Pallotta and Jousselme, 2015).

In the following we present some machine learning and data mining approaches that have been recently proposed for automatic anomalous maritime vessel behaviour detection. These approaches are summarised in Table 1.

Wang et al. (2014) have performed vessel route anomaly detection using Density-Based Spatial Clustering of Applications with Noise considering Speed and Direction (DBSCANSD) (Liu et al., 2014) and parallel meta-learning algorithms on the Hadoop MapReduce environment cluster. According to their experiment, the authors have mentioned that the vessel route anomaly detector accuracy and execution time improves linearly with the number of nodes in the cluster.

Pallotta et al. (2013a; 2013b) proposed an approach based on an incremental variant of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al., 1996) to learn maritime traffic routes from vessel AIS data to detect anomalous vessel behaviours and predict vessels' future positions. The proposed approach is called Traffic Route Extraction for Anomaly Detection (TREAD). The same TREAD approach was used in Pallotta and Jousselme (2015) to perform track association to vessel traffic routes for hierarchical maritime anomaly detection. In their approach, a tool implementing the TREAD approach is first used to extract vessel traffic routes from historical AIS data of areas of interest.

Secondly for associating and classifying tracks (i.e., partial trajectories), the same authors have adopted a hierarchical reasoning, where new tracks are first associated to existing routes based on their positional information only, and "off-route" vessels are detected as vessels not following an existing route. Then, for on-route vessels (vessels following an existing route) further anomalies are detected such as speed anomaly and heading anomaly.

To provide a more complete and interactive Maritime Situational Picture (MSP) to operational authorities and policy-makers to support the decision-making process, Arguedas et al. (2015) have proposed a spatio-temporal data mining framework based on the TREAD approach for route-based anomalous vessel detection. In the proposed framework, the authors used the TREAD approach to build a synthetic representation of all maritime traffic routes; afterward they perform maritime traffic anomaly detection.

Lei (2015) proposed a framework for maritime trajectory modelling and anomaly detection. First, the author defines three outlying features of abnormal movement behaviours: spatial, sequential, and behavioural features extracted from the sequential occurrence in trajectory data collection, and then builds a maritime trajectory model for anomaly detection. Finally, the author develops an anomaly detection algorithm based on the learned trajectory model, in which an indicator is used to evaluate suspicious behaviour and scores trajectory behaviour according to the defined outlying features.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

Table 1. Summary of machine learning and data mining approaches to anomalous vessel behaviour detection.

Research work	Methods	Approach type	Used attributes of vessel motion data	Used data set
Pallotta et al. (2013a; 2013b)	Incremental DBSCAN algorithm	Point-based	Vessel location, speed, course and location timestamp.	Recorded AIS data
Arguedas et al. (2015)	Incremental DBSCAN and DBSCAN-based algorithms	Point-based	Vessel location, speed, course and location timestamp.	Recorded AIS data
Pallotta and Jousselme (2015)	Incremental DBSCAN algorithm	Point-based	Vessel location, speed, course and location timestamp.	Recorded AIS data
De Vries and Van Someren (2012)	one-class SVMs algorithms	Trajectory-based	Vessel MMSI, location and location timestamp.	Recorded AIS data
Guillarme and Lerouvreur (2013)	The OPTICS clustering algorithm	Trajectory-based	Vessel MMSI, location, speed, course, and location timestamp.	Recorded AIS data
Wang et al. (2014)	DBSCAN-based algorithm and parallel Meta learning algorithm with MapReduce	Point-based	Vessels position, speed and course.	Recorded AIS data
Lei (2015)	Grid-based clustering algorithm	Trajectory-based	Vessel MMSI, location and location timestamp.	Recorded AIS data
Radon et al. (2015)	Density based clustering algorithm similar to DBSCAN and Line segment clustering algorithms.	Trajectory-based	Vessel MMSI, location, speed, course, and location timestamp.	Recorded AIS data
Smith et al. (2014b)	Conformal predictor-based, algorithms.	Trajectory-based	Vessel MMSI, location and Speed.	Recorded AIS data
Smith et al. (2015)	Multi-class Conformal predictor-based algorithm	Trajectory-based	Vessel MMSI, type, location, Speed and location timestamp.	Recorded AIS data
Laxhammar and Falkman (2015)	Inductive conformal predictor-based algorithm with Local Outlier Factor-based non conformity measure.	Trajectory-based	Vessel MMSI, location and location timestamp	Recorded AIS data
Laxhammar and Falkman (2012)	Sliding Window based on Conformal Anomaly Detection and Local Outlier Factor.	Trajectory-based	Vessel MMSI, location and location timestamp	Recorded AIS data
Laxhammar and Falkman (2011)	Anomaly detection method based on conformal predictors and Hausdorff distance	Trajectory-based	Vessel MMSI, location and location timestamp.	Simulated trajectory data and real video trajectories

Vessel location is composed of longitude and latitude; Used attributes of vessel motion data refers to the vessel attributes that are used as the parameters for determining vessels' moving pattern and anomalous behaviour detection; Approach type refers to the input format of the data used in the approach; Point-based approach considers individual vessel observations as input data; Trajectory-based approach considers the entire vessel trajectory as input data; Recorded AIS data refers to historical vessel motion data provided by the Automatic Identification System.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

One of the issues in anomalous vessel behaviour detection is the high rate of false positive anomalies. Some research works have been focused on addressing this issue in anomalous vessel behaviour detection.

To reduce the rate of false alarm in anomalous maritime vessel behaviour detection, Radon et al. (2015) have proposed an approach in two phases. In the first phase, the normal vessel movement patterns are extracted from historical vessel tracks (trajectories within a particular origin and destination) using trajectory partition, the DBSCAN and line segment clustering algorithms. In the second phase potential anomalies are detected and then real anomalies are filtered out from potential anomalies based on contextual verification. In the same line of false alarm rate reduction, Laxhammar and Falkman (2011; 2012; 2015) and Smith et al. (2014b; 2015) have used an anomalous vessel behaviour detection approach based on conformal predictors (Vovk et al., 2005) known as conformal anomaly detection. The conformal anomaly detection approach assumes that all the training data are generated by the same probability distribution and makes no further assumptions about the distribution. For each new test data, the likelihood that it is coming from the same probability distribution as the training data is estimated. If this likelihood is below the specified confidence level, then the new object is declared as an anomaly. The advantage of the conformal anomaly detection approach is that the rate of false alarm is well-calibrated and it is at most equal to the specified confidence level.

Based on the result of their preview work presented in Liu et al. (2014; 2015) have proposed a clustering-based anomaly detection model for vessel movement data. In their approach, the authors have proposed three distance measures which are the core of the anomaly detection algorithm. The three distance measures are, Absolute Division Distance (ADD), Relative Division Distance (RDD) and Cosine Division Distance (CDD). ADD is used in the stopping points abnormality detection phase while RDD and CDD are used for the moving part. To evaluate the effectiveness of the proposed model, the authors have used a highly imbalanced data set from the Juan de Fuca Strait area.

2.1.3. *Other data-driven approaches to anomalous vessel behaviour detection.* Apart from statistical, machine learning, data mining, and visual analytics approaches, recently other approaches related to anomalous vessel behaviour detection have been proposed, and some of them are presented in what follows and are summarised in Table 2.

Osekowska et al. (2014; 2015) used an approach based on potential fields to perform vessel anomalous behaviour detection. In their approach, every vessel generates potential charges, the value of which represents the vessel's behaviour, and drops the charges at locations it passes. Each AIS report is used to assign a potential charge at the reported vessel positions. A model of normal vessel behaviour (traffic patterns) is represented by the collection of potential fields created based on the recorded vessel motion data and behaviour history. Vessels whose current potential deviates from the local potential fields are flagged as anomalous vessels.

Soleimani et al. (2015) have proposed a new method for finding anomalies in ships' movements. The proposed method analyses the trajectory of ships from a geometrical perspective. The trajectory of the ship is compared with a near-optimal path that is generated by a graph search algorithm. To make this comparison the proposed method extracts some scale-invariant features from the real trajectory and from the optimal movement pattern, and then it compares the two sets of features to generate an abnormality score.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

Table 2. Summary of other data-driven approaches to anomalous vessel behaviour detection.

Research work	Methods	Used attributes of vessel motion data	Used data set
McAbee et al. (2014)	Hough transformation	Vessel location	Recorded AIS data
Osekowska et al. (2014; 2015)	Potential field-based method	Vessel location, speed, course, vessel type, and location timestamp.	Recorded AIS data
Soleimani et al. (2015)	Geometrical-based method	Vessel MMSI, location and location timestamp.	Recorded AIS data
Wu et al. (2014)	Topology preserving mapping method	Vessel MMSI, location, speed, course, and location timestamp.	Recorded AIS data
Terroso-Saenz et al. (2015)	Complex Event Processing-based method	Vessel MMSI, location, speed and location timestamp.	Recorded AIS data
Venskus et al. (2015)	Bio-inspired virtual pheromone-based method	Vessel name, class, location, speed and heading.	Recorded AIS data

Vessel location is composed of longitude and latitude; Used attributes of vessel motion data refers to the vessel attributes that are used as the parameters for determining vessels moving pattern and anomalous behaviour detection; Recorded AIS data refers to historical vessel motion data provided by Automatic Identification System.

Venskus et al. (2015) proposed an adaptive self-learning algorithm based on a bio-inspired virtual pheromone method for the detection of anomalous maritime traffic. The proposed algorithm first learns normal vessel traffic pattern grids from vessel movement data provided by the surveillance data fusion subsystem, then detects a potential incident where the moving vessel path does not match with learned pattern grids. The authors have verified that the algorithm provides rapid self-learning and fast adaptation characteristics.

2.1.4. *Knowledge-based approaches to anomalous vessel behaviour detection.* Knowledge-based anomalous vessel behaviour detection approaches refer to the generation of alerts based on a set of rules (Roy, 2008), such as maximum speed allowed in a port, presence in areas restricted to navigation or inconsistencies between the ship's claimed and actual activity. These approaches are composed of two phases. The first phase consists of creating and specifying predefined patterns that one wants to detect using rules. These rules can be predicates or any type of logic. In the second phase the new data are matched to the predefined rules to detect predefined anomalies of interest, which are those data that deviate from the predefined rules.

Although most of the recent works on anomalous vessel detection are data-driven, some knowledge-based approaches have been proposed.

Maggi et al. (2013) have demonstrated the use of process mining for analysing the behaviour of vessels, and for detecting anomalies, i.e., deviations from the normal behaviour. To model the behaviour of vessels, the authors have used the Declare language which is a constraint-based language that can specify complex behaviour in terms of a couple of restrictions. Declare combines a formal semantics for analysis purposes with a graphical representation for users.

Brax et al. (2012) have presented a Multi-Agent System (MAS) dedicated to anomalous behaviour detection and alerts triggering in the maritime surveillance area. Anomalies are detected based on the implementation of maritime regulation.

Vandecasteele et al. (2013) have proposed a semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behaviour analysis.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

For maritime anomaly detection, the authors have used spatial ontology for modelling a maritime operator's knowledge.

2.1.5. *Hybrid Approaches to anomalous vessel behaviour detection.* The hybrid approach combines the data-driven and knowledge-based approaches to perform anomalous maritime vessel behaviour detection. Recently, only limited effort has been devoted to the use of hybrid approaches for anomalous maritime vessel behaviour detection. Shahir et al. (2014; 2015) combined data-driven machine learning methods with maritime background domain knowledge to detect anomalous vessel interaction patterns. In Kazemi et al. (2013) a hybrid framework for anomaly detection based on the use of open data (contextual information and background knowledge) is proposed. Based on the proposed anomaly detection framework and the algorithms implementing the expert rules, the Open Data Anomaly Detection System (ODADS) was developed.

3. ANALYSIS AND DISCUSSION.

In this section we analyse and discuss the recent proposed approaches to anomalous vessel behaviour detection in order to highlight the main issues related to them. In addition, we point out some scarcities of literature on automatic detection of anomalous vessel behaviour.

The issues related to most of the recent approaches to automatic anomalous vessel behaviour detection are mainly scalability issues and the issue of offline anomalous behaviour detection.

3.1. *Scalability issue of machine learning and data mining approaches to anomalous vessel behaviour detection.* A review of recent studies on automatic anomalous vessel behaviour detection reveals that the majority (Arguedas et al., 2015; Pallotta et al., 2013a; 2013b; Radon et al., 2015; Wang et al., 2014) of proposed data mining approaches are based on the DBSCAN clustering algorithm or its adapted version. However, the DBSCAN clustering algorithm is sensitive to the input parameter values (Eps, MinPts) the determination of which requires complex and time consuming computation or domain expert knowledge. In addition, the time complexity of the DBSCAN algorithm is $O(n^2)$ where n is the number of the data point in the training data set, which is an issue when dealing with a large amount of data. It has been pointed out in Ester et al. (1996) that the average time complexity of the DBSCAN algorithm can be reduced to O(n log n) by using spatial indices such as the R*-tree structure. However, to the best of our knowledge no recent research work on anomalous vessel behaviour detection has used spatial indices in the DBSCAN algorithm.

The DBSCAN algorithm, and other machine learning and data mining algorithms (Guillarme and Lerouvreur, 2013; De Vries and Van Someren, 2012; Smith et al., 2014a; 2015; Laxhammar and Falkman, 2011), used for automatic detection of anomalous vessel behaviour also suffer from the scalability issue when they are applied to a large volume dataset because the training time is too long. Using big data technologies such as the Hadoop MapReduce paradigm could be one of the potential solutions to scalability issues.

3.2. *The scarcity of literature on online anomalous vessel behaviour detection.* Almost all the recent proposed approaches for anomalous vessel behaviour detection have focused on offline anomalous detection, whereas only a little effort has been devoted to sequential anomalous vessel behaviour detection (Laxhammar and Falkman, 2011; 2012; 2015; Smith et al., 2014b; 2015; Pallotta et al., 2013a) or online anomalous vessel behaviour detection (Smith et al., 2012; 2014a). In offline anomalous detection both the vessel normalcy model learning and anomalous behaviour detection are performed offline,

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

which is a serious limitation in surveillance applications because it delays anomaly alarms and thus the ability to take action on impending events (Laxhammar and Falkman, 2011). Therefore, the development of algorithms that perform real time (online) or near real time vessel normalcy model learning and anomalous vessel behaviour detection could enhance the level of maritime safety and security. Moreover as the AIS vessel motion data can be considered as time series or asynchronous streaming data, some algorithms for anomaly detection in a time series or stream data can be applied to vessel AIS data in order to perform real time or near real time anomalous vessel behaviour detection.

3.3. *The scarcity of literature on big data technologies in anomalous vessel behaviour detection.* Another noteworthy fact is the scarcity of published literature on the use of big data technologies in anomalous maritime vessel behaviour detection. According to our study, only a few works (Cazzanti et al., 2015; Wang et al., 2014; Wijaya and Nakamura, 2013) have used big data technologies to perform anomalous vessel behaviour detection. Since machine learning and data mining approaches for anomalous vessel behaviour detection are divided into point-based and trajectory-based approaches according to the format of input data, and point-based approaches treat each AIS position message as a point on a geographical grid, big data technology such as the Hadoop MapReduce paradigm can be applied to the point-based approaches in order to speed up and scale up the process of anomalous vessel behaviour detection.

4. CONCLUSION AND FUTURE DIRECTION.

Good maritime situational awareness can be achieved through the analysis of vessel movement data provided by tracking systems such as the Automatic Identification System (AIS). Recently, many approaches have been proposed in the literature on automatic anomalous maritime vessel movement behaviour detection based on vessel AIS data. In this work we have presented a review of some of these proposed approaches, discussed some of their issues, and highlighted some scarcities of literature on some aspects related to automatic anomalous vessel behaviour detection.

Through our study we have found that the recent proposed approaches to automatic anomalous vessel behaviour detection can be divided into the following categories: statistical, machine learning, data mining and other. Most of these approaches take into account only vessel kinematic features such as vessel location (longitude, latitude), Speed Over the Ground (SOG) and Course Over the Ground (COG). The datasets used are mainly vessel movement data provided by AIS transponders.

The main issues related to some of these proposed approaches are: 1) scalability issues and 2) the issue of offline anomalous vessel behaviour detection. This study has also revealed some scarcity in the literature on anomalous vessel behaviour detection such as the scarcity of the use of big data technologies, the scarcity of real time anomalous vessel behaviour detection, and the scarcity of the use of both vessel kinematic attributes and static characteristics in anomalous vessel behaviour detection.

Although extensive research work has been done on anomalous vessel behaviour detection to achieve an efficient maritime situational awareness, the use of big data technologies in anomalous vessel behaviour detection, and real-time anomalous vessel behaviour detection still remain a challenge. Therefore the use of big data technologies (such as the Hadoop MapReduce framework) and real-time anomalous vessel behaviour detection constitute the direction of current and future research related to automatic anomalous vessel movement behaviour detection.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

ACKNOWLEDGMENTS

This research is supported by the National Natural Science Foundation of China (project No: 51479155).

REFERENCES

Anneken, M., Fischer, Y. and Beyerer, J. (2015). Evaluation and comparison of anomaly detection algorithms in annotated datasets from the maritime domain. *Proceedings of IEEE SAI Intelligent Systems Conference (IntelliSys)*, 169–178.

Anscombe, F.J. and Guttman, I. (1960). Rejection of outliers. *Technometrics*, **2**(2), 123–147.

Arguedas, V.F., Mazzarella, F. and Vespe, M. (2015). Spatio-temporal data mining for maritime situational awareness. *Proceedings of the IEEE Conference OCEANS 2015-Genova*, 1–8.

Brax, N., Andonoff, E. and Gleizes, M.-P. (2012). A Self-adaptive Multi-Agent System for Abnormal Behavior Detection in Maritime Surveillance. *In: G. Jezic et al. (Eds.), KES-AMSTA 2012, LNAI. Springer-Verlag Inc*., 174–185.

Castaldo, F., Palmieri, F.A.N. and Regazzoni, C. (2015). Application of Bayesian Techniques to Behavior Analysis in Maritime Environments. *In: Bassis, S., Esposito, A., Morabito, F.C. (eds.), Advances in Neural Networks: Computational and Theoretical Issues. Springer International Publishing Inc., Switzerland,* 175–183, DOI: 10.1007/978-3-319-18164-6_17.

Cazzanti, L., Millefiori, L.M. and Arcieri, G. (2015). A Document-based Data Model for Large Scale Computational Maritime Situational Awareness. *Proceedings of the IEEE International Conference on Big Data (Big Data)*, 1350–1356.

Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, **41**(3), 1–58.

De Vries, G.K.D. and Van Someren, M. (2012). Machine learning for vessel trajectories using compression, alignments and domain knowledge. *Expert Systems with Applications*, **39**(18), 13426–13439.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the 2nd International Conference Knowledge Discovery and Data Mining (KDD'96)*, 226–231.

Guillarme, N.L. and Lerouvreur, X. (2013). Unsupervised extraction of knowledge from S-AIS data for maritime situational awareness. *Proceedings of the 16th IEEE International Conference on Information Fusion (FUSION),* 2025–2032.

Kazemi, S., Abghari, S., Lavesson, N., Johnson, H. and Ryman, P. (2013). Open data for anomaly detection in maritime surveillance. *Expert Systems with Applications*, **40**(14), 5719–5729.

Kowalska, K. and Peel, L. (2012). Maritime Anomaly Detection using Gaussian Process Active Learning. *Proceedings of the 15th IEEE International Conference on Information Fusion (FUSION),* 1164–1171.

Lane, R.O., Nevell, D.A., Hayward, S.D. and Beaney, T. W. (2010). Maritime Anomaly Detection and Threat Assessment. *Proceedings of the 13th IEEE Conference on Information Fusion (FUSION),* 1–8.

Last, P., Bahlke, C., Hering-Bertram, M. and Linsen, L. (2014). Comprehensive Analysis of Automatic Identification System (AIS) Data in Regard to Vessel Movement Prediction. *The Journal of Navigation*, **67**, 791–809. DOI: 10.1017/S0373463314000253.

Laxhammar, R. and Falkman, G. (2011). Sequential Conformal Anomaly Detection in Trajectories based on Hausdorff Distance. *Proceedings of the 14th IEEE International Conference on Information Fusion (FUSION)*, 1–8.

Laxhammar, R. and Falkman, G. (2012). Online Detection of Anomalous Sub-trajectories: A Sliding Window Approach Based on Conformal Anomaly Detection and Local Outlier Factor. *In: Iliadis, L., Maglogiannis, I., Papadopoulos, H., Karatzas, K., and Sioutas, S. (eds.), AIAI 2012, Part II. IFIP AICT, vol. 382,* 192–202, Springer, Heidelberg.

Laxhammar, R. and Falkman, G. (2015). Inductive conformal anomaly detection for sequential detection of anomalous sub-trajectories. Annals of Mathematics and Artificial Intelligence, **74**, 67–94, DOI: 10.1007/s10472-013-9381-7.

Lei, P.-R. (2015). A framework for anomaly detection in maritime trajectory behavior. *Knowledge and Information Systems*, **47**(1), 189–214.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

Liu, B., De Souza, E.N., Matwin, S. and Sydow, M. (2014). Knowledge-based clustering of ship trajectories using density-based approach. *Proceedings of the IEEE International Conference on Big Data (Big Data)*, 603–608.

Liu, B., De Souza, E.N., Hilliard, C. and Matwin, S. (2015). Ship movement anomaly detection using specialized distance measures. *Proceedings of the 18th IEEE International Conference on Information Fusion (Fusion)*, 1113–1120.

Maggi, F.M., Mooij, A.J. and van der Aalst, W.M.P. (2013). Analyzing Vessel Behavior using Process Mining. *In: Van de Laar, P., Tretmans, J. and Borth M. (eds.), Situation Awareness with Systems of Systems. Springer International Publishing Inc.,* 133–148, DOI: 10.1007/978-1-4614-6230-9_9.

Mascaro, S., Nicholso, A.E. and Korb, K.B. (2014). Anomaly detection in vessel tracks using Bayesian networks. *International Journal of Approximate Reasoning*, **55**(1), 84–98.

McAbee, A., Scrofani, J., Tummala, M., Garren, D. and McEachen, J. (2014). Traffic pattern detection using the Hough transformation for anomaly detection to improve maritime domain awareness. *Proceedings of the 17th IEEE International Conference on Information Fusion (FUSION),* 1–6.

Osekowska, E., Johnson, H. and Carlsson, B. (2014). Grid size optimization for potential field based maritime anomaly detection. *Transportation Research Procedia*, **3**, 720–729.

Osekowska, E., Johnson, H. and Carlsson, B. (2015). Learning Maritime Traffic Rules Using Potential Fields. *In: Corman, F., Voß, S. and Negenborn, R.R. (eds.), ICCL 2015, LNCS. Springer International Publishing Inc.,* 298–312, DOI: 10.1007/978-3-319-24264-4_21.

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

Pallotta, G., Vespe, M. and Bryan, K. (2013b). Traffic Knowledge Discovery from AIS Data. Proceedings of the 16th IEEE International Conference on Information Fusion, pp. 1996–2003.

Pallotta, G. and Jousselme, A.-L. (2015). Data-driven detection and context-based classification of maritime anomalies. *Proceedings of the 18th IEEE International Conference on Information Fusion,* 1152–1159.

Radon, A.N., Wang, K., Glässer, U., Wehn, H. and Westwell-Roper, A. (2015). Contextual verification for false alarm reduction in maritime anomaly detection. *Proceedings of the IEEE International Conference on in Big Data (Big Data)*, 1123–1133.

Roy, J. (2008). Anomaly Detection in the Maritime Domain. *In Proceedings of SPIE 6945, Optics and Photonics in Global Homeland Security IV*, 69450W, DOI: 10.1117/12.776230.

Shahir, H.Y., Glasser, U., Nalbandyan, N. and Wehn, H. (2014). Maritime Situation Analysis A Multi-vessel Interaction and Anomaly Detection Framework. *Proceedings of the IEEE Joint Intelligence and Security Informatics Conference (JISIC)*, 192–199.

Shahir, H.Y., Glasser, U., Shahir, A.Y. and Wehn, H. (2015). Maritime Situation Analysis Framework: Vessel Interaction Classification and Anomaly Detection. *Proceedings of the IEEE International Conference on Big Data (Big Data)*, 1279–1289.

Silveira, P.A.M., Teixeira, A.P. and Guedes Soares, C. (2013). Use of AIS Data to Characterise Marine Traffic Patterns and Ship Collision Risk off the Coast of Portugal. *The Journal of Navigation*, **66**, 879–898. DOI: 10.1017/S0373463313000519.

Smith, M., Reece, S., Roberts, S. and Rezek, I. (2012). Online maritime abnormality detection using Gaussian process and extreme value theory. *Proceedings of the 12th IEEE International Conference on Data Mining (ICDM)*, 645–654.

Smith, M., Reece, S., Roberts S., Psorakis I. and Rezek, I. (2014a). Maritime abnormality detection using Gaussian processes. *Knowledge and Information Systems*, **38**, 717–741.

Smith, J., Nouretdinov, I., Craddock, R., Offer, C., and Gammerman, A. (2014b). Anomaly detection of trajectories with kernel density estimation by conformal prediction. *In: Iliadis, L., Maglogiannis, I., Papadopoulos, H., Sioutas, S., Makris, C. (eds.), AIAI 2014 Workshops, IFIP AICT. IFIP International Federation for Information Processing, Inc.,* 271–280.

Smith, J., Nouretdinov, I., Craddock, R., Offer, C., and Gammerman, A. (2015). Conformal anomaly detection of trajectories with a multi-class hierarchy. *In: Gammer man, A., Vovk, V., Papadopoulos, H. (eds.) SLDS 2015, LNAI. Springer International Publishing Inc., Switzerland,* 281–290, DOI: 10.1007/978-3-319-17091-6_23.

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

Terroso-Saenz, F., Valdes-Vela, M. and Skarmeta-Gomez, A. F. (2015). A complex event processing approach to detect abnormal behaviours in the marine environment. *Information Systems Frontiers*, **18**(79), 1–16.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press

Tijardovic, I. (2009).The Use of AIS for Collision Avoidance. *The Journal of Navigation*, **62**(1), 168–172. DOI: 10.1017/S0373463308005055.

Vandecasteele, A., Devillers, R. and Napoli, A. (2013). A Semi-Supervised Learning Framework Based on Spatio-Temporal Semantic Events for Maritime Anomaly Detection and Behavior Analysis. *Proceedings of the 11th International Symposium for GIS and Computer Cartography for Coastal Zones Management.*

Venskus, J., Kurmis, M., Andziulis, A. and Lukošius, Ž. (2015). Self-Learning Adaptive Algorithm for Maritime Traffic Abnormal Movement Detection based on Virtual Pheromone Method. *Proceedings of the IEEE International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS),* 1–6.

Vovk, V., Gammerman, A. and Shafer, G. (2005). *Algorithmic learning in a random world.* Springer.

Wang, X., Liu, X., Liu, B., De Souza, E.N. and Matwin, S. (2014). Vessel route anomaly detection with Hadoop mapreduce. *Proceedings of the IEEE International Conference on Big Data (Big Data),* 25–30.

Wijaya, W.M. and Nakamura, Y. (2013). Predicting Ship Behavior Navigating Through Heavily Trafficked Fairways by Analyzing AIS Data on Apache HBase. *Proceedings of the IEEE first International Symposium on Computing and Networking,* 220–226.

Wu, Y., Patterson, A., Santos, R.D.C., and Vijaykumar, N. L. (2014). Topology preserving mapping for maritime anomaly detection. *In: Murgante, B. et al. (Eds.): ICCSA 2014, Part VI, LNCS. Springer International Publishing Inc.*, 313–326, DOI: 10.1007/978-3-319-09153-2_24.

https://doi.org/10.1017/S0373463317000066 Published online by Cambridge University Press