Vilniaus universitetas
Duomenų mokslo ir skaitmeninių technologijų institutas
L I E T U V A

INFORMATIKOS INŽINERIJA (07 T)

DOKTORANTO MOKSLINĖ ATASKAITA UŽ 2017/2018 METUS

<Julius Venskus>

2018 m. spalis

Mokslinė ataskaita DMSTI-DS-07T-xx-<ataskaitos nr.>

VU Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4,
Vilnius LT-08663
www.mii.lt

Unsupervised Neural Network Retraining: Real-time Maritime Traffic Anomaly Detection Based on History Data Embedding

Julius Venskus · Povilas Treigys · Jolita Bernatavičiene · Gintautas Tamulevičius · Viktor Medvedev

Abstract The automated identification system of vessel movement receives enormous amounts of data that need to be analyzed to make the proper decision on vessel movement. A vast number of vessels make the process of abnormal movement detection time-consuming and complicated for human analysts and rapid-response algorithms for the decision support system have to be developed to detect the abnormal vessel movement in intense maritime traffic areas. This paper extends the previous study on a self-organizing map application for stream data, received by maritime automated identification system, processing. With the view to maintain the quality of the algorithm results, data batching strategies for the neural network retraining are investigated to detect anomalies in streaming maritime traffic data. Authors estimate the strategies performance on model precision as well as learning rate parameter change; present an experimental investigation that has been performed using the real data set. The results obtained show that proposed model retraining strategies allow decreasing model training time by half while keeping the model sensitivity and precision at minor change.

Keywords Marine traffic anomaly detection · Neural network retrain strategy · Retrain speed-up · Learning rate parameter estimation

1 Introduction

The maritime industry is one of the main sectors in Europe. The total gross weight of goods transported as part of EU short sea shipping is estimated at almost 1.9 billion tonnes of goods in 2016, an increase of 2.6% from the previous year. Despite that the sector is one of the most important areas of human activity, it is one of the most dangerous. The growth of maritime traffic in ports and their surroundings raises the traffic and security control problems and increases the workload for traffic control operators. As the traffic becomes more intense to prevent maritime incidents also becomes very important and more difficult task. To deal with the problem, technological solutions can be applied [28]. Nowadays, the navigation technology is highly developed: the vessels are considerably bigger, faster and are safer in typical situations. However, in non-typical situations, huge numbers of vessels make the process of abnormal movement detection time-consuming and error-prone for human analysts [27]. Data mining and visual analytics are becoming very important to extract useful knowledge from the increasingly available information on vessels and their movements. This enables the automatic detection of anomalies, the prediction of vessel routes, the understanding and mapping of activities at sea.

The scientific community is actively working on the development of algorithms for modelling the regular maritime traffic in ports and surroundings, to improve safety at sea and security in navigation. Anomaly or abnormal movement detection is one of the technique available that improves the domain safety and security. The obtained knowledge can help spotting suspicious vessel movements and provide additional information to traffic control operators. In this paper, anomalies are defined as deviations from the normal state. A maritime trajectory includes motion data, auxiliary data (e.g., speed, rotation, meteorological data) for a ship, and such trajectories can be used for anomalous event detection [7]. Anomalies are detected as deviations between the vessel's registered live data and the history based potential sea region data.

Therefore, most of the existing methods for abnormal movement detection in maritime traffic are not suitable for massive stream data processing. They can hardly be applied in the decision support system because of the high computational cost. To this end, rapid self-learning algorithms for the decision support system have to be developed to detect the abnormal movement in stream data of intense maritime traffic areas.

The paper is structured as follows. In Section 2, the related works on the abnormal movement detection in maritime traffic and the state-of-the-art solutions are reviewed. Section 3 introduces a motivation background for this research. The experimental results of the proposed retraining strategies are demonstrated in Section 4. The last section concludes the paper.

2 Review

In this section, we will present maritime anomaly detection task and review some recent research results in this area.

The anomalous vessel movement can be defined as an unreasoned movement deviation from the sea lanes, trajectory, speed or other traffic parameters [12]. As most of the vessels have the automated identification system (AIS) installed, giving the static and dynamic information about the vessel movement, the detection of traffic anomaly comes as the task of data analysis and outlier detection. Traffic data is analysed in point-based or trajectory-based manner [21].

In the first case, every single data point (message from the vessel to the AIS) or a group of them is treated as an independent point. For this purpose, the analysed geographical area is subdivided into independent cells with related AIS messages. These data points in the grid are analyzed using so-called signature based or rule-based techniques. The idea of these techniques is the employment of various association rules to detect specific movement changes [18]. Zhu applied database management, data warehouse, and data mining technologies to analyse AIS data [30]. Deng in [5] extended the features and inserted time stamps. These extensions enabled to employ Markov model for supplementation of rules. While declaring the point-based analysis, Pallotta et al. in [16] have proposed to use a sliding time window to estimate the relationship between successive AIS data points. The obtained waypoints are clustered using Density-Based Spatial Clustering of Applications with Noise methodology and employed for anomaly detection and movement prediction. There, despite the claims about point-based analysis, the authors have implemented the idea of updating the traffic knowledge from the input of AIS messages and the use of historical knowledge. The same clustering methodology was explored in [1]. Here, the historical spatiotemporal data is analysed to detect waypoints of routes.

The main weakness of point-based techniques is the analysis of movement short-term history or disregard of history even. The planned and purposing vessel movement should generate highly-correlated AIS data, and this can be used for movement anomaly detection. On the other hand, a limited number of analysed data points means real-time calculation and decision making. This quality makes point-based anomaly detection techniques attractive for real-time tasks. Nevertheless, at the moment the prevalence of these techniques is quite limited.

Trajectory-based techniques treat the entire traffic data sequence as a whole. Several research directions are analysed in the literature related to the analysis of vessel trajectories: maritime traffic pattern mining, ship collision risk assessment [22], and maritime anomaly detection [14], [29], [19], identification the types of ships [20], combating abalone poaching [4].

In the case of trajectory-based detection, models of normal movement are created (using the entire trajectory data, not part of it) and the anomalies are detected as movement data inadequacy to the model. Thus, these techniques are characterised by a huge amount of AIS data to analyse. This property requires some data pre-processing like compression or clustering.

In [26] a piece-wise linear segmentation is applied to compress the data of vessel trajectories, then the similarity of trajectories (for detection of anomalies) is performed using alignment kernels (dynamic time warping and edit distances, namely). The model by Lei [14] defines spatial, sequential, and behavioural features of the vessel movement. The movement anomaly is detected as the outlying features of the trajectory model, and the degree of suspiciousness is evaluated. The geometrical properties of the trajectory are employed in [24]. Here, the vessel trajectory is compared with the graph search-based path and the difference is estimated by a final score. The threshold value of the score is employed as the decision and labelling value. Another trajectory-based analysis techniques can be found in [17], [13], [23].

Analysis of the entire trajectory gives the advantage of the historical movement data, which can be important for anomaly detection. However, full data analysis requires much more complicated algorithms like neural networks, for example. This complicates the application of trajectory-based analysis for real-time tasks. Also, such algorithms are sensitive to missing data (lost AIS messages, for example).

A comprehensive and categorizing review on maritime anomaly detection can be found in [12], [3], [19].

Analysis of full trajectory data and anomaly detection will require data-driven approaches like artificial neural network-based or statistical approaches. These approaches can perform in an unsupervised or semi-supervised manner (i.e. do not need labelled data) and can cope with large amounts of data. The issue of real-time calculations should be solved using the idea of incremental modelling (retraining, re-estimating, etc.): the model of vessel movement should be updated with respect to recent data to avoid of complete remodelling or model retraining.

3 Motivation

Noting that the vessel movement can be treated differently regarding the context. For example, if the ship is quite distant from the seaport then even high decline from its course cannot be indicated as an anomaly: weather condition, stormy sea, etc. may have great influence on vessel trajectory. On the other hand, is vessel movement is observed at the seaport surroundings even small deviation from the course may be thought as abnormal vessel activity. To this purpose method used for traffic anomaly detection has to have a feature that allows different region scaling at different maritime traffic observation areas. Such scaling property has a self-organising map (SOM) method. SOM is a neural network-based method that is trained in an unsupervised way using a competitive learning [10], [9], [11], [6], [8]. This type of neural networks can be used for both visualisation and clustering of multidimensional data [15].

In the previous research [25] the modified SOM algorithm for maritime vessel movement data classification into normal and abnormal classes is presented. The modification is achieved by incorporating virtual pheromone intensity calculations at the last epoch of model training. During the model validation stage, the pheromone intensity threshold is established by applying a gradient descent method. The dependence of the network neighbouring function on the classification results was investigated, the best classification accuracy is achieved using the Mexican hat neighbouring function. The influence of different SOM grid dimensions on the classification results of the proposed algorithm has been investigated. It was proved experimentally that the algorithm achieved the best precision using grid dimension 25x25. This knowledge will be further used as a starting point for network data batching and training strategies investigation presented in this paper.

With the growth of maritime traffic especially near the seaports, the complete retraining of the SOM algorithm becomes costly in terms of training time. The need for algorithm retrain is quite straightforward: the more vessel movement data is observed and fed to the algorithm the better precision of the algorithm is. Besides, retraining ensures the inclusion of the most recent movement data that reflects actual conditions and context. To maintain high algorithm precision and sensitivity, approaches to data streaming, batching and model retrain strategies has to be explored [2]. In this article, authors explore three model train strategies. Strategy I presents data batching and algorithm training whenever the new batch becomes available as if no model history data would be available. Strategy II presents algorithm performance while using pre-trained model parameters on previously trained data with the newly arriving data batches. Strategy III presents different data batch shuffling techniques and the use of previously pre-trained model parameters. All three strategies investigate the learning rate parameter influence on the model performance and train time as well. The above presented strategies should be investigate due to the nature of the vessel data passed to AIS system. Data passed from a vessel can be viewed as a stream that contains facts regarding vessel movement trajectories. Those may depend on seasonal data, the shipping routes, schedules and so on. Thus, the abnormality detection model has to be developed by analysing vessel movement trajectories (as well as historical data) in an incremental manned based on the up-to-date data it receives.

4 Experiments

In this section, we present a detailed description of the SOM network retraining strategies and results of the experiments using real data sets.

4.1 Data Preparation

The detailed description of the previous study of SOM size and modification by introducing the SOM evaporation functions are presented in [25]. Data from the region of medium maritime traffic at the Klaipeda seaport were selected for the analysis of the proposed retraining strategies of the SOM network. During the experiments, two data sets were used: Cargo vessels and Passenger vessels. Each item (point) of vessel's streamed data is described by longitude, latitude, heading, vessel speed, wind direction, wind speed, wave direction, and wave height values. The Cargo data set is represented by 180300 and the Passenger data set is described by 43879 vessel movement observation items that were registered in a streamed manner. All experiments in this section were carried out with the Cargo data set; later the data batching strategies were tested on the Passenger data set (see Table 8).

At first, 20% of the Cargo vessel data set is randomly selected for the general model error evaluation. Then, the resulting 80% of the data set items are used for the data batching Strategy investigation. Those 80% of data items are split in such manner: 20% of this part of data is used for Strategy testing and the rest 80% of data were used for T1, T2, and T3 data batch splitting. Batches are used in the experiments to imitate the continuous data arrival with the view to investigate different SOM network retraining strategies and learning rate parameter selection. The scheme of data split is shown in Figure 1.

All the data items are sorted with respect to data send time ascending. The SOM network of size 25x25 is taken according to the SOM size investigation published in [25].

4.2 Training Strategies of the SOM Network

Strategy I. For the SOM network training and validation, we use T1, T2 and T3 data batches. The learning rate parameter is set to 0.5. Then, after the network is

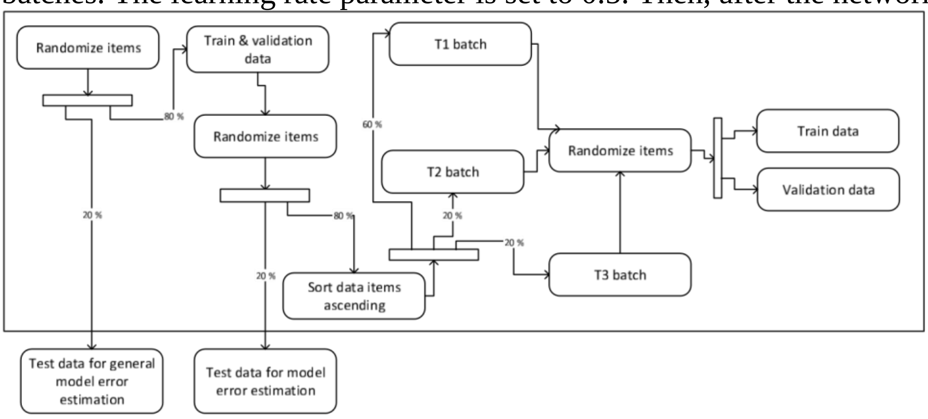

Fig. 1 Data split scheme

trained and validated with the T1 data batch, the new data is fed to the network in such way: the T1 and T2 batch data are merged together and the algorithm is trained from the initial random state using all items from T1 and T2. The same scheme is applied with the T3 data batch.

In order to get the best network performance, the learning rate parameter can be adjusted. Initial investigation let us to divide the learning rate parameter search into these intervals and step sizes: in the interval [0.005;0.04] step is set to 0.005; in the interval [0.04;0.1] step size is increased to 0.01; and in the interval [0.1;0.5] step size is set to 0.1 (see Table 1). In such way, the training experiment of the Strategy I is repeated while every learning parameter value is tested in order to achieve the best algorithm performance. After the model is trained, it is being tested with the model test data set and the data set that allows to evaluate the general model error. The best-obtained model characteristics with model test data set are presented in Table 1 (bold line).

The statistics of the best Strategy I model using test data for general model error estimation and test data for model error estimation is presented in Table 2. The time needed for the algorithm retraining is 40769 seconds.

Strategy II. The initial algorithm trained 10 times with the T1 batch data. During each training, the weights of the SOM network were generated randomly, and the best performing network has been selected while keeping fixed learning rate parameter at the value of 0.5. The performance of the investigated network on repetitive Strategy II (using only T1 data set) model evaluation and testing is presented in Table 3. The line marked in bold shows the best network obtained. Quite a small deviation of the precision and the sensitivity rates shows the network stability. Then, the best-obtained network parameters are used as an initial weights for the network to be trained with T2 batch data. Finally, imitating the new data portion arrival, the best model obtained with T2 batch data is retrained with the T3 batch data. The results of the additional experiment have shown that best performance network is obtained with learning rate 0.025.

The statistics of the best model with model test and general model error evaluation data is presented in Table 4. The time needed for the model training is 18229 seconds.

Table 1 Selection of learning rate

Learning rate	TP	FP	TN	FN	Precision	Sensitivity
0.005	924	519	26648	757	0.6403	0.5497
0.010	943	505	26662	738	0.6512	0.5610
0.015	957	498	26669	724	0.6577	0.5693
0.020	963	487	26680	718	0.6641	0.5729
0.025	968	478	26689	713	0.6694	0.5758
0.030	976	471	26696	705	0.6745	0.5806
0.035	986	468	26699	695	0.6781	0.5866
0.040	998	461	26706	683	0.6840	0.5937
0.050	1025	445	26722	656	0.6973	0.6098
0.060	1066	413	26754	615	0.7208	0.6341
0.070	1109	394	26773	572	0.7379	0.6597
0.100	1197	303	26864	484	0.7980	0.7121
0.200	1431	135	27032	250	0.9138	0.8513
0.300	1486	81	27086	195	0.9483	0.8840
0.400	1500	55	27112	181	0.9646	0.8923
0.500	**1510**	**52**	**27115**	**171**	**0.9667**	**0.8983**
0.600	1507	54	27113	174	0.9654	0.8965
0.700	1502	59	27108	179	0.9622	0.8935

Table 2 Training Strategy I performance at learning rate 0.5

Stage	TP	FP	TN	FN	Precision	Sensitivity
Testing (model error)	1510	52	27115	171	0.9667	0.8983
Testing (general error)	1868	69	33890	233	0.9644	0.8891

Table 3 Strategy II performance on model test data

No.	TP	FP	TN	FN	Precision	Sensitivity
1	1364	241	26926	317	0.8498	0.8114
2	1329	280	26887	352	0.8260	0.7906
3	1359	252	26915	322	0.8436	0.8084
4	1364	274	26893	317	0.8327	0.8114
5	1356	253	26914	325	0.8428	0.8067
6	1335	253	26914	346	0.8407	0.7942
7	1314	251	26916	367	0.8396	0.7817
8	1332	258	26909	349	0.8377	0.7924
9	**1367**	**237**	**26930**	**314**	**0.8522**	**0.8132**
10	1338	240	26927	343	0.8497	0.7960
				max	0.8522	0.8132
				min	0.8260	0.7817
				average	0.8413	0.8011
				stdev	0.0079	0.0115

Strategy III. The scheme of the model training validation and testing is the same as that in Strategy II description except of two things. Firstly, from T2 and T3 batches, there are produced 4 data batches (Tm2-Tm5) every containing quarter of both of T2 and T3 data (see Table 5). Secondly, as previously described, after every model training and validation, the parameters of the best obtained model are used for every next Tm2-Tm5 batch training, except the model training data aggregation. For every retraining test data for model error estimation of data is used as described in previous Strategy I and Strategy II. The half of items

Table 4 Retraining Strategy II performance at learning rate 0.025

Stage	TP	FP	TN	FN	Precision	Sensitivity
Testing (model error)	1500	98	27069	181	0.9387	0.8923
Testing (general error)	1836	122	33837	265	0.9377	0.8739

from Tm2-Tm5 data batches is compound of items from T2 and T3 as shown in Table 5 (Tm2-Tm5) while another part of data is selected proportionally, with respect to that data points attached to the previous best model SOM winning neurons. Approach guarantee that the knowledge of frequently passed sea regions is incorporated into the next model training because it is not frequent for the ships to change their sea routes. Experiments have depicted that the best model is obtained with the learning rate being 0.03.

Table 5 Partitioning of data set (Strategy III)

Data batches	% of train and validation data	New data items	All data itms
T1	60%	69235	69235
Tm2	10%	11539	23078
Tm3	10%	11539	23078
Tm4	10%	11539	23078
Tm5	10%	11539	23078

The statistics of the Strategy III best model obtained using test data for general model error estimation and test data for model error estimation is presented in Table 6.

Table 6 Retraining Strategy III performance at learning rate 0.003

Stage	TP	FP	TN	FN	Precision	Sensitivity
Testing (model error)	1527	73	27094	154	0.9544	0.9084
Testing (general error)	1866	91	33868	235	0.9535	0.8881

The time needed for the algorithm retraining was 27854 seconds. The summary of relative time needed for the training Strategies I-III is presented in Table 7.

Table 7 Retraining Strategies I-III performance on Cargo data set

Strategy	Precision	Sensitivity	Relative time
Strategy I	0.9644	0.8891	1
Strategy II	0.9377	0.8739	0.4471
Strategy III	0.9535	0.8881	0.6832

The same data batching Strategies I-III described above has been tested on the Passenger data set as well. The results are presented in Table 8.

Table 8 Retraining Strategies I-III performance on Passenger data set

Strategy	Precision	Sensitivity	Relative time
Strategy I	0.9795	0.8897	1
Strategy II	0.9802	0.8870	0.4478
Strategy III	0.9817	0.8888	0.6817

From the results shown in Table 7 and Table 8 it can be seen that by applying different SOM model retraining Strategies, while keeping the same data batch sizes, it is possible to substantially decrease the time for maritime traffic abnormal movement detection while retraining the model precision and sensitivity at very high values. The results obtained show that the SOM network could be retrained in halftime while keeping precision and sensitivity at almost the same high values. The results presented in Table 8 prove the correctness of the training strategies investigation.

5 Conclusions

In this article, three different unsupervised SOM network retraining strategies have been presented and investigated. It was shown that the SOM network could be retrained in halftime while keeping model precision and sensitivity varying not more than 3% in unusual maritime traffic detection.

The results of the performed experiments show that:

- if the model is trained from initial random weights of the SOM network the best performance is observed, however the training time is the longest. Model precision reaches 0.979 and sensitivity 0.889 at learning rate 0.5.
- if the model is trained on top of the pre-trained model weights the precision and sensitivity slightly drops but the training time decreases by half at learning rate 0.025.
- if the model is trained on top of the pre-trained model weights and newly arrived data batch is proportionally mixed with that winning neurons, training time can be decreased by one third while keeping the same model performance at learning rate 0.03.

The independent experiment on different data set confirms results correctness and allows to conclude that by applying batched data the SOM network training can be shortened to halftime by selecting learning rate parameter from the interval [0.025;0.03] while keeping the same model performance.

References

1. Arguedas, V.F., Mazzarella, F., Vespe, M.: Spatio-temporal data mining for maritime situational awareness. In: OCEANS 2015 - Genova, pp. 1–8 (2015). DOI 10.1109/OCEANSGenova.2015.7271544
2. Bernataviciene, J., Dzemyda, G., Bazilevicius, G., Medvedev, V., Marcinkevicius, V.,Treigys, P.: Method for visual detection of similarities in medical streaming data. International Journal of Computers Communications and Control **10**(1), 8–21 (2015)
3. Cazzanti, L., Pallotta, G.: Mining maritime vessel traffic: Promises, challenges, techniques. In: OCEANS 2015 - Genova, pp. 1–6 (2015). DOI 10.1109/OCEANS-Genova.2015.7271555
4. Dabrowski, J.J., de Villiers, J.P., Beyers, C.: Context-based behaviour modelling and classification of marine vessels in an abalone poaching situation. Engineering Applications of Artificial Intelligence **64**, 95–111 (2017)
5. Deng, F., Guo, S., Deng, Y., Chu, H., Zhu, Q., Sun, F.: Vessel track information miningusing AIS data. In: 2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems (MFI), pp. 1–6 (2014). DOI 10.1109/MFI.2014.6997641
6. Dzemyda, G., Kurasova, O., Zilinskas, J.: Multidimensional Data Visualization: Methodsˇ and Applications, vol. 75. Springer Science & Business Media (2012)
7. Fu, P., Wang, H., Liu, K., Hu, X., Zhang, H.: Finding abnormal vessel trajectories usingfeature learning. IEEE Access **5**, 7898–7909 (2017)
8. Kamimura, R.: Input information maximization for improving self-organizing maps. Applied intelligence **41**(2), 421–438 (2014)
9. Kayacik, H.G., Zincir-Heywood, A.N., Heywood, M.I.: A hierarchical SOM-based intrusiondetection system. Engineering applications of artificial intelligence **20**(4), 439–451 (2007) 10. Kohonen, T.: The self-organizing map. Neurocomputing **21**(1-3), 1–6 (1998)
11. Kurasova, O., Molyte˙, A.: Quality of quantization and visualization of vectors obtainedby neural gas and self-organizing map. Informatica **22**(1), 115–134 (2011)
12. Lane, R.O., Nevell, D.A., Hayward, S.D., Beaney, T.W.: Maritime anomaly detection andthreat assessment. In: Information Fusion (FUSION), 2010 13th Conference on, pp. 1–8. IEEE (2010)
13. Laxhammar, R., Falkman, G.: Inductive conformal anomaly detection for sequential detection of anomalous sub-trajectories. Annals of Mathematics and Artificial Intelligence **74**(1), 67–94 (2015). DOI 10.1007/s10472-013-9381-7. URL https://doi.org/10.1007/s10472-013-9381-7
14. Lei, P.R.: A framework for anomaly detection in maritime trajectory behavior. Knowledgeand Information Systems **47**(1), 189–214 (2016)
15. Medvedev, V., Kurasova, O., Bernataviˇciene˙, J., Treigys, P., Marcinkeviˇcius, V., Dzemyda,G.: A new web-based solution for modelling data mining processes. Simulation Modelling Practice and Theory **76**, 34–46 (2017)
16. Pallotta, G., Vespe, M., Bryan, K.: Vessel pattern knowledge discovery from AIS data: Aframework for anomaly detection and route prediction **15**, 2218–2245 (2013)
17. Radon, A.N., Wang, K., Glsser, U., Wehn, H., Westwell-Roper, A.: Contextual verification for false alarm reduction in maritime anomaly detection. In: 2015 IEEE International Conference on Big Data (Big Data), pp. 1123–1133 (2015). DOI 10.1109/BigData.2015.7363866
18. Ristic, B.: Detecting anomalies from a multitarget tracking output. IEEE Transactions on Aerospace and Electronic Systems **50**(1), 798–803 (2014). DOI 10.1109/TAES.2013.130377
19. Riveiro, M., Pallotta, G., Vespe, M.: Maritime anomaly detection: A review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **2018**, 1–19 (2018)
20. Sheng, K., Liu, Z., Zhou, D., He, A., Feng, C.: Research on ship classification based ontrajectory features. The Journal of Navigation **71**(1), 100–116 (2018)
21. Sidib´e, A., Shu, G.: Study of automatic anomalous behaviour detection techniques formaritime vessels. The Journal of Navigation **70**(4), 847–858 (2017)
22. Silveira, P., Teixeira, A., Soares, C.G.: Use of AIS data to characterise marine trafficpatterns and ship collision risk off the coast of portugal. The Journal of Navigation **66**(6), 879–898 (2013)
23. Smith, J., Nouretdinov, I., Craddock, R., Offer, C., Gammerman, A.: Conformal anomalydetection of trajectories with a multi-class hierarchy. In: A. Gammerman, V. Vovk, H. Papadopoulos (eds.) Statistical Learning and Data Sciences, pp. 281–290. Springer International Publishing, Cham (2015)
24. Soleimani, B.H., Souza, E.N.D., Hilliard, C., Matwin, S.: Anomaly detection in maritimedata based on geometrical analysis of trajectories. In: 2015 18th International Conference on Information Fusion (Fusion), pp. 1100–1105 (2015)
25. Venskus, J., Treigys, P., Bernataviˇciene˙, J., Medvedev, V., Voznak, M., Kurmis, M., Bulbenkiene˙, V.: Integration of a self-organizing map and a virtual pheromone for real-time abnormal movement detection in marine traffic. Informatica **28**(2), 359–374 (2017)
26. de Vries, G.K.D., van Someren, M.: Machine learning for vessel trajectories using compression, alignments and domain knowledge. Expert Systems with Applications **39**(18), 13426 – 13439 (2012). DOI https://doi.org/10.1016/j.eswa.2012.05.060
27. Will, J., Peel, L., Claxton, C.: Fast maritime anomaly detection using kd-tree gaussianprocesses. In: IMA Maths in Defence Conference (2011)
28. Yu, Q., Luo, Y., Chen, C., Wang, X.: Trajectory outlier detection approach based oncommon slices sub-sequence. Applied Intelligence pp. 1–20 (2017)
29. Zhen, R., Jin, Y., Hu, Q., Shao, Z., Nikitakos, N.: Maritime anomaly detection withincoastal waters based on vessel trajectory clustering and Na¨ıve Bayes classifier. The Journal of Navigation **70**(3), 648–670 (2017)

30. Zhu, F.: Mining ship spatial trajectory patterns from ais database for maritime surveillance. In: 2011 2nd IEEE International Conference on Emergency Management and Management Sciences, pp. 772–775 (2011). DOI 10.1109/ICEMMS.2011.6015796