18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Ship Movement Anomaly Detection Using Specialized Distance Measures

Bo Liu*, Erico N. de Souza*, Cassey Hilliard‡ and Stan Matwin*†
‡ *Dalhousie University, Canada*
* *Faculty of Computer Science, Dalhousie University, Canada*
† *Institute for Computer Science, Polish Academy of Sciences, Poland*
Email: boliu@dal.ca, erico.souza@dal.ca, casey.hilliard@dal.ca, stan@cs.dal.ca

***Abstract*—This paper provides a solution for anomaly detection in maritime traffic domain based on the clustering results presented in a previous work. That work created clusters for vessels moving close to shores by associating vessel movements with International Maritime Organization Rules (especially Traffic Separation Scheme Boundaries). In this paper, we show how three division distances with the clusters can detect anomalous navigational behaviors. The proposed method decides for each trajectory point if the vessel is anomalous, considering longitude, latitude, speed and direction. Although the approach is point-based, which is applicable for real-time AIS surveillance, it is also flexible enough for analysts to set their own threshold for labeling whole trajectories.**

***Keywords*-anomaly detection; clustering; trajectory mining; maritime surveillance**

I. Introduction

In [1], a method was presented to cluster Automatic Identification System (AIS) data, with the objective to use the clusters to identify possible anomalous behavior. This work continues the work presented in [1], by creating a method to generate an anomaly score, which can be used by a human analyst to decide whether a certain track is anomalous or not. This work considers that anomalies are all the GPS (Geographic Positioning System) points in a track that deviate from normal speed or direction, or are relatively distant from a cluster. After applying the model to the real AIS data, the experiment results not only prove the effectiveness of the model, but also validate our previous work done in [1].

Maritime Anomaly Detection techniques primarily fall into two categories: statistical modelling [2][3][4] and predictive modelling [5][6][7]. The general idea of using statistical techniques for anomaly detection is to fit a statistical model for normal behaviors with the given data set and then apply a statistical inference test to determine if an unseen instance belongs to the model [8]. Approaches based on predictive models usually predict future status information (e.g. position, speed and course) of a particular vessel and then compare the real data with the prediction to decide the abnormality.

In the maritime domain, the majority of statistical models are built upon the momentary kinematical features (position, course, speed and acceleration rate) of individual vessels. Laxhammar [2] used a Gaussian Mixture Model (GMM) and a greedy version of Expectation-Maximization (EM) for clustering. A GMM can be regarded as an ensemble model of K multivariate Gaussian distributions (mixture components). The greedy EM algorithm is employed to determine the parameter set for all the K distributions. In [3], the authors propose to use adaptive Kernel Density Estimator (KDE) for estimating unknown probability densities and modelling arbitrary sea lanes. In the anomaly detection phase, the anomaly detector is sequentially applied to the incoming data. The value of new incoming point's density is calculated under the null hypothesis (no anomaly) and this value is then compared with a detector parameter related to false alarms for deciding the new point's abnormality. A comparison between the two approaches above is given in [9], demonstrating that the anomaly detection results from both models are not satisfactory. As the two models detect the anomalous segments at a significant distance from the point where anomaly behavior happens (three kilometer and four kilometer respectively) while an expected effective anomaly detector should detect such behaviors at a shorter distance [9]. In [4], Gerben $et\ al.$ propose a technique based on Machine Learning models. In their work, different trajectory alignment kernels (Dynamic Time Warping and Edit Distance) are applied with one-class SVMs (Support Vector Machine) for detecting the outlying trajectories. This trajectory-based method, however, is not applicable for real-time AIS surveillance, unlike our proposed point-based method.

Pallotta $et\ al.$ [5] suggest the use of rule-based and low-likelihood models for anomaly detection. The rules defined in this approach are similar to those in other knowledge-based work, which require maritime domain experts' knowledge. As an example, the maximum speed pre-defined in a port area can only be accurately estimated by a specialist knowledgeable about the area. For low-likelihood detection, a Weibull model was employed (a parametric exponential-like model), along with a sliding time window technique to avoid problems with incomplete and intermittent tracks.

Similar to the work done in [5], Nevell [6] proposes to use a Bayesian approach to predict the future route of a particular vessel for comparison. This methodology is based on a node-sparse network, built from different kinds of coastal nodes. In [7], the authors insist that an overall threat

is indicated by a sequence of the individual behaviours. Therefore, five specific anomalies are introduced to extend Nevell's work [6] to assess the probability of a higher-level threat based on a constructed Bayesian Network.

One issue with the work described in [6][7] is that it cannot incorporate speed into deciding if a trajectory is anomalous; instead the judgement is based only on position. Another problem is that a pre-defined network may not be applicable in many near-port regions due to the nature of port traffic. Traffic in near-port areas is usually highly variable and the vessels are not always following straight lanes (optimal routes in [6][7]). Both of these problems are handled with our approach.

This work uses the results of the clustering framework presented in [1], which generates arbitrary shapes of moving patterns and stopping areas. The proposed method is point based but is capable of handling trajectory tracks. For each track the algorithm will return an anomaly ratio. The ratio is based on three types of distances to take position, direction and speed into consideration.

The rest of this paper is organized as follows. Section 2 gives an overview of the framework and a brief introduction to the normal traffic extraction model. Section 3 proposes the anomaly detection model. Section 4 presents the results of experimental evaluation. In the final section, we conclude with a summary and discuss our method's limitations and the potential future work.

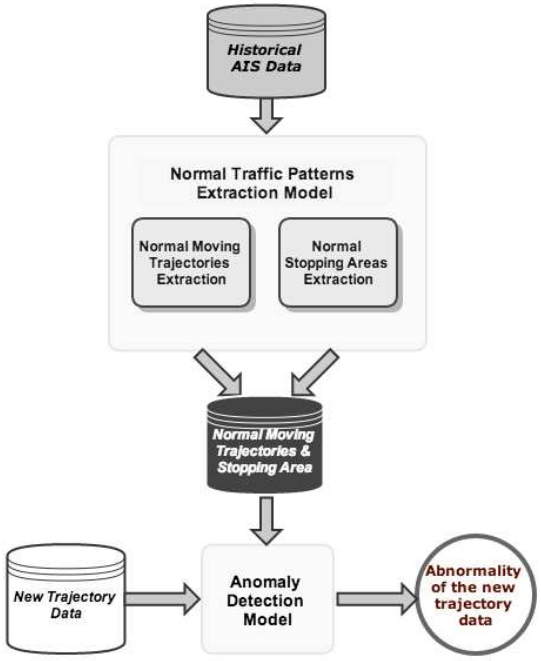

Figure 1. The framework of Maritime Anomaly Detection Model. This can be roughly regarded as a two-component procedure. First, the system extracts the normal traffic patterns from the historical AIS data repository. Then the extracted GVs and SSPs can be employed by the anomaly detection model to decide the abnormality of the new given trajectory.

II. Model Overview

In this paper, we propose a clustering-based maritime traffic anomaly detection model. The approach includes two components: the normal traffic patterns extraction model and the anomaly detection model. As shown in Figure 1, historical AIS data is first sent to the normal patterns extraction model which generates, as output, a set of gravity vectors (GV) and stopping sampling points (SSP) via its two sub-components. Afterwards, when new ship trajectory data is to be judged, the second component (anomaly detection model) will be applied based on the normal traffic patterns results to decide the new data's abnormality.

The work for normal traffic patterns extraction is presented in [1]. Before proposing our second part of the framework, a brief overview of the first phase is necessary.

The proposed clustering-based normal traffic patterns extraction model first divides the AIS data into moving and stopping parts respectively based on a stopping SOG (Speed Over Ground) threshold of 0.5 knots.

For the case where SOG is not less than 0.5 knots, we propose DBSCANSD as the basic algorithm to detect the main traffic lanes within the data. DBSCANSD is based on DBSCAN [10] algorithm, modified to consider that for each point of a cluster, the neighborhood of a given radius has to contain at least a minimum number of points with similar SOGs (Speed Over Ground) and COGs (Course Over Ground). The algorithm's output is a set of Gravity Vectors (GV), which are vectors formed by 5 features: average COG, average SOG, average Latitude, average Longitude and Median Distance. Before calculating the GVs of one cluster, the cluster area is first partitioned into a grid of multiple cells. The partitioning is based on an experimentally decided width to ensure that each cell contains a sufficient number of trajectory points. So each cell has one particular GV. The Median Distance of a GV is the median of all the distances between the points in the cell and the cell's average geographical point. More details about the process can be found in [1].

For the case where the SOG is less than 0.5 knots (stop areas), the original DBSCAN algorithm is executed because speed and direction are not important factors. Another type of vector is created as output, labelled Sampled Stopping Point (SSP), dependent only on the geographic shape of the region.

III. Abnormal Trajectory Detection

In this section, we present our anomaly detection model. Similar to the normal traffic pattern extraction model, we also treat stopping and moving separately. A ship trajectory in a particular area, especially in the near-port areas, can consist of both stopping points and moving points. Hence, given one incoming trajectory dataset (one sequence of

trajectory points with same identification), every point in this trajectory should first be labeled as stopping or moving according to the SOG threshold (0.5 knots) and then be checked for abnormality. In the following part of the section, three division distances are first presented and then our abnormal detection model is introduced.

A. Three Division Distances

In order to label every data point, three division distances are proposed in our approach, Absolute Division Distance (ADD), Relative Division Distance (RDD) and Cosine Division Distance (CDD). ADD is employed in the stopping points abnormality detection phase while RDD and CDD are used for the moving part. And the definitions of the three division distances are given in the following. Here we use **target point**s to represent the points of a new coming trajectory to be labeled.

The Absolute Division Distance between a target point P_t and a Sampled Stopping Point P_s is defined as:

$$D_{absolute} = Distance((P_t.Lat, P_t.Lon), (P_s.Lat, P_s.Lon)) \quad (1)$$

As can be seen in Definition 1, ADD is actually the Geographical Distance [11] between Latitude and Longitude values of the target point(P_t) and the sampled stopping point(P_s).

The Relative Division Distance between a target point P_t and a Gravity Vector GV is defined as:

$$D_{relative} = \frac{Distance((P_t.Lat,P_t.Lon),(GV.Lat,GV.Lon))}{GV.MedianDistance} \quad (2)$$

For moving trajectory points, we adopt RDD ($D_{relative}$) rather than ADD ($D_{absolute}$) because the normal moving lanes could have different widths according to their surrounding geographical environment. The routes in a narrow strait can be more cramped than those in the open sea. Relative distance, the ratio of a point's distance from the centroid to the median distance of all the points in one cluster from the centroid, is one efficient metric in clustering-based approaches for detecting outliers [12]. We replace the centroid of each cluster with our Gravity Vectors, which can be regarded as surrogates for a centroid when we only consider Longitude and Latitude. As stated in [1], every cluster can have more than one GV, in which case the median distance of one GV can be used to measure the width variance of a moving cluster. Another work that confirms this approach is presented by Ethiene *et al* [13] in which it is shown that the choice of the statistical decile used to compute the spatio-temporal channel can give a tolerable estimate of this channel's width.

The next division distance is CDD, which is employed to involve direction and speed of moving objects. The Cosine Division Distance between a target point P_t and a Gravity Vector GV is defined as:

$$D_{cosine} = \cos\alpha \times \frac{min(P_t.SOG,GV.SOG)}{max(P_t.SOG,GV.SOG)} \quad (3)$$

where:

α is the angle between the two directions, that is, the difference between P_t's COG and GV's COG.

The intuition behind CDD is to combine the angle ($\leq 180°$) between the two directions (angle α is defined by COG differences between P_t and GV) and the difference between the two speeds. Figure 2 shows two abnormal cases that consider COG and SOG. In Figure 2(a), we can see that the speeds of the two vectors are in the same length L while the angle a is too large. This could be explained as a ship crossing a normal lane in a nearly opposite direction, which may cause a collision. In Figure 2(b), the speed of the target point P_t is much lower than GV's although they have similar COG. In this case, the slowly sailing vessel may also be in a high risk of collision in relation to other ships moving at a normal rate of speed. As a consequence, Cosine Division Distance (CDD) is proposed. It can be easily seen that $cos\ \alpha$ accounts for the angles' difference and the ratio between two SOGs can reflect the speeds' difference.

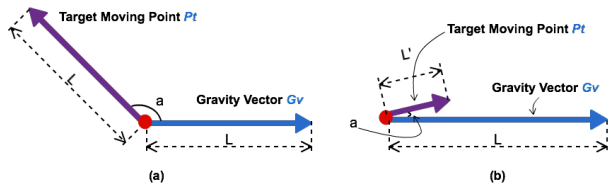

Figure 2. Two Abnormal Cases after considering COG and SOG

The Cosine Division Distance proposed in this work combines both COG and SOG in its calculation. A possible alternative is to calculate the distance of each component separately and then check the normality, but this alternative may increase the misclassification error. One example is presented in Figure 3. In this scenario, it is assumed that a certain point P_t is in the middle of two different grids belonging to different clusters. If this approach of separately calculating each component is employed, the target point, P_t, will first be considered normal with respect to G_v' on the basis of direction, and also with respect to G_v on the basis of length (value of speed). However, P_t is an abnormal point because it is in the direction of G_v', but with a much faster speed.

B. Anomaly Detection Model

After having the three division distances defined, we can present our anomaly detection model.

As shown in Algorithm 1, the anomaly detection process is completed in two steps. Lines 3-6 employ ADD to label

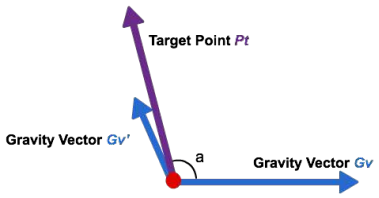

Figure 3. One case that there are two moving clusters in one specific area or point.

the stopping points of the trajectory and Lines 7-14 use RDD and CDD to decide the moving points' labels. Lastly, the ratio of abnormal points to the number of all points is returned as the abnormality (Lines 15-18). The abnormality can be interpreted as a confidence ratio and it is beneficial for the end users to choose a level of confidence while retrieving the abnormal results of the whole system.

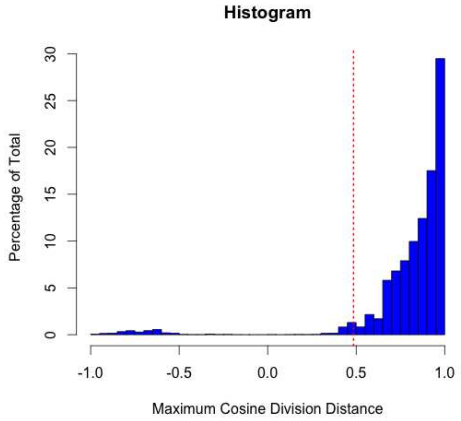

Figure 4. The maximum CDD distribution in the area of Juan de Fuca Strait.

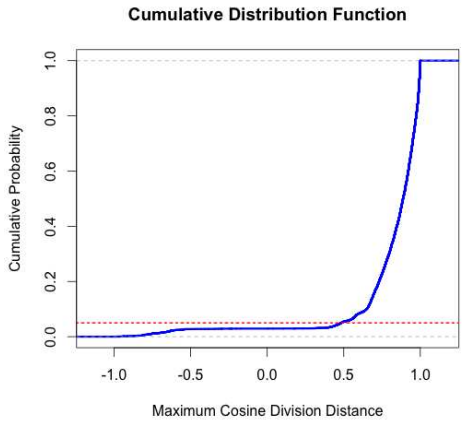

Figure 5. The Cumulative Distribution Function of the maximum CDD in the area of Juan de Fuca Strait.

This algorithm requires 3 thresholds as inputs, which can be estimated through experiments. In this work, we first take out one month of data in a specific area and calculate all the division distances, then select specific values based on distribution and quartile analysis. Figure 4 shows the distribution of the maximum CDD of moving points in the area of Juan de Fuca Strait. And Figure 5 illustrates the CDF (Cumulative Distribution Function) of the maximum CDD in the area. It is obvious that over 90% of the data points' CDD are greater than 0.5. Thus, in this case, we can choose 0.5 or a smaller value as the CDD threshold for this area and then when a new trajectory dataset is given, we can use this value as the basis for anomaly detection.

Algorithm 1 Detect abnormality of the target trajectory

Input: (1) The target trajectory dataset, D; (2) The lists of Sampled Stopping Points and Gravity Vectors from the previous model, SSP and GV; (3) Three thresholds, $add\_threshold$, $rdd\_threshold$ and $cdd\_threshold$

Output: The abnormality rate, $abnormality$

1: Separate D into two sub-datasets based on the speed threshold, moving dataset D_m and stopping dataset D_s
2: Initialize all labels of points in D_m and D_s as Normal
▷ label stopping points of the target trajectory
3: **for** each data point S in D_s **do**
4: $ADD\_s \leftarrow$ minimum(ADD(S,SSP))
5: **if** $ADD\_s > add\_threshold$ **then**
6: $S.label \leftarrow$ Abnormal
▷ label moving points of the target trajectory
7: **for** each data point M in D_m **do**
8: $RDD\_m \leftarrow$ minimum(RDD(M,GV))
9: **if** $RDD\_m > rdd\_threshold$ **then**
10: $M.label \leftarrow$ Abnormal
11: **else**
12: $CDD\_m \leftarrow$ maximum(CDD(M,GV))
13: **if** $CDD\_m < cdd\_threshold$ **then**
14: $M.label \leftarrow$ Abnormal
15: $count\_ab \leftarrow$ the number of the abnormal points in D
16: $count\_all \leftarrow$ the total number of all points in D
17: $abnormality \leftarrow count\_ab/count\_all$
18: **return** $abnormality$

IV. Experiments

In this section, we evaluate the effectiveness of our maritime anomaly detection model in the region of Juan de Fuca Strait. The evaluation work contains two parts, the first one is conducted with the non-labeled data while the second one is done after labelling the data. The results of the first experiment are shown visually and the second experiment compares our model's results with the labels by the expert.

The data set which was prepared for normal traffic patterns extraction phase comprises two months of trajectory

data from November 1 to December 31 in 2012 and contains 67,850 trajectory points. The whole data set is not used for extracting normal patterns, instead 46,000 records (40,000 moving points and 6,000 stopping points distinguished by the SOG threshold 0.5 knots) are selected and then the rest of the data set are used for estimating the 3 thresholds for the anomaly detection phase.

Afterwards, to evaluate our anomaly detection model in both experiments, we chose the first half of January (January 1st to January 15th) in 2013, as our target trajectory data set. This second dataset consists of 284 different trajectories with 17,431 points.

A. *Experiment On Unlabeled Data Set*

After applying the normal traffic extraction model, 16 different clusters, including 15 moving clusters and 1 stopping cluster are identified. Figure 6 shows the gravity vectors of the moving clusters and the sampled stopping point of the stopping cluster. After this extraction phase, only 388 points are generated which include 388 GVs and only one SSP.

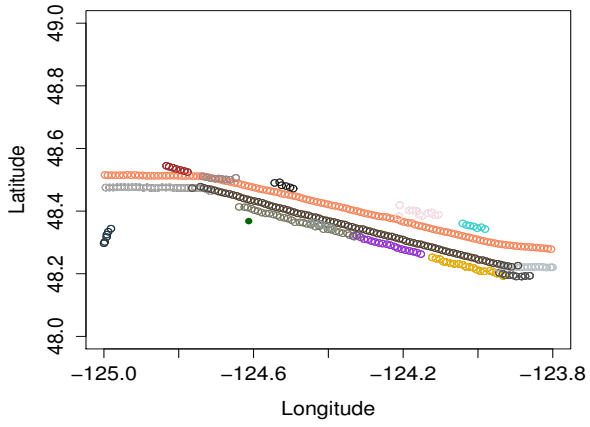

Figure 6. The Gravity Vectors (open circles) and Sampled Stopping Points (filled circles) extracted from the clusters in JUAN DE FUCA STRAIT area.

The next step before detecting anomalous trajectories is to estimate the 3 thresholds used in Algorithm 1. As stated before, the remaining 23,850 trajectory records are chosen for this phase. There are 10,825 stopping points and 11,025 moving points in this subset.

The quartile values of ADD and RDD of the subset are shown in Table I. Authors tested different thresholds to be used as anomaly detector. After various tests, for the area of Strait of Juan de Fuca, the best threshold value was to consider 95% of the data as normal in relation to distance, and from this sub-set another 95% of the data to be considered normal in relation to speed and direction. The main objective is to reduce the number of false alarms (vessels considered abnormal, while they are normal), and this filtered data will later be evaluated by a human expert that will give the final decision. The model is flexible to allow changing this threshold value depending on the geographical area under evaluation.

Table I
QUARTILE STATISTICS OF THE 3 DIVISION DISTANCES

Statistic	ADD	RDD	CDD
Min	0.13	0.00537	-0.9937
1st Quartile	3.00	0.70300	0.7642
Median	4.46	1.04000	0.8876
Mean	36.97	1.81500	0.8104
3rd Quartile	6.89	1.58400	0.9612
Max	44250.00	52.86000	0.9999

So we choose the sample quantiles of 0.95 for both ADD and RDD. The corresponding thresholds of ADD and RDD in this case are 97.290 and 5.938. After calculating the RDD threshold, the statistic for CDD is obtained (shown in the 3rd column in Table I). Then we select 0.05 as the possibility to decide the third threshold (0.485) which can be employed as our CDD threshold.

With the extracted normal patterns and the thresholds estimated, we start to evaluate the capacity of detecting abnormal trajectories.

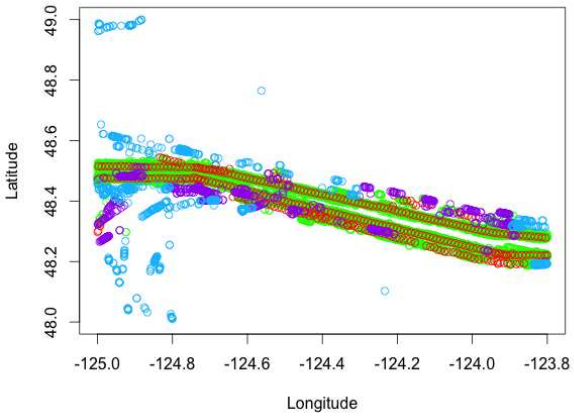

Figure 7. The anomaly labeling results of the trajectory data points in JUAN DE FUCA STRAIT area. GVs and SSPs are in red and the normal points are in green. The two types of abnormal points are in blue (abnormal in relation to ADD or RDD) and purple (abnormal in relation to CDD).

In this step, we first apply our model to the trajectory data points; the labeling results of the data points are shown in Figure 7. The red points stand for the GVs and the SSP in this area. The green points are normal, while the blue and purple ones are abnormal. More specifically, a blue point means it is too far away from the corresponding GV or SSP, while a purple point represents that its speed or direction is too aberrant in the specific location. In this case, 1,534

Table II
LABELS AND DESCRIPTIONS BY THE EXPERT

Label	Description
Bad_Pos	Track contains questionable point, far outside track, looks like bad GPS return
In_Excl_Zn	Track has significant portion within the exclusionary zone between traffic lanes
$XING\_TSS$	Track appears to be crossing lanes of TSS [14]
$XING\_NShor$	Track appears to be crossing lanes of near shore two way traffic area
Odd_Mvmt	Track shows unusual movement without other explanation
$Leave\_Lane$	Vessel was in traffic lane, then veered outside
$Harbour$	Track seems to describe in-harbour navigation or moored vessel
$Normal$	Normal Movement

points (872 in blue and 662 in purple) of the 17,431 points are finally considered as abnormal.

After finishing the labeling process, we calculate the abnormality ratio of each trajectory. We use a threshold of 0.5 as the minimum confidence rate to extract the abnormal trajectories. Noteworthy, the number 0.5 is adjustable and was chosen based only on the experiments to reduce the presence of false alarms (normal trajectories considered abnormal by our model). The result is that 22 trajectories are labeled abnormal among the total 284 trajectories, in other words, the abnormality rates of the 22 trajectories are over 0.5. In Figure 8, six examples of abnormal cases are illustrated. As can be seen in the figures, both purple points and blue points contribute to the final abnormality of one trajectory.

B. Experiment On Labeled Data Set

In this experiment, the same data set is labeled by an expert who has multiple years of experience in maritime data analysis. The labelling process is not biased by our model's results since only the raw AIS data has been provided to the expert.

A track division method is employed by the expert before labelling. More specifically, the AIS data points are divided into distinct tracks on the basis of vessel ID (MMSI), and where temporally sequential points are separated by no more than 4.5 minutes of time. Thus one track, as defined above, may be divided into multiple-sub tracks on the basis of time. Additionally, this process can result in tracks comprised of single points. For the one-point track case, the length of the track is 0 nautical mile and the track is not assigned with any labels. The final labels and their descriptions provided by the expert are shown in Table II.

The labels assigned by the expert are not based on the points; instead, they are based on the whole sub-tracks. In other words, as long as one sub-track shows an anomalous pattern, the whole set of points inside the sub-track will be assigned as one kind of abnormal label. Another noteworthy point is that the expert has not taken SOG (speed) into account except for Harbour behavior during his labeling process. That is, whether the speed of the vessel is too fast or too slow near the lane is not considered, but our algorithm can take this into account.

Table III
CONFUSION MATRIX

	Abnormal (Our Model)	Normal (Our Model)
Abnormal (Expert's Label)	4	10
Normal (Expert's Label)	127	1301

From Table II, we can firstly assume the label of $Normal$ as normal patterns based on the description. Then the label $Harbour$ can also be considered as a normal case because it is reasonable for a vessel to moor in harbour. Lastly, we observe that all the sub-tracks with the label of $Leave\_Lane$ only have tiny changes from their route and they still navigate strictly within the normal lanes. So we also classify the label of $Leave\_Lane$ as normal.

After dividing the 284 tracks (284 different MMSIs), 2,122 sub-tracks are generated. Among them, 680 sub-tracks contain only one point (Length=0 nautical mile). Then 14 sub-tracks are classified as abnormal labels (other than $Leave\_Lane$, see Table II) by the expert and the remaining 1,428 tracks are all normal patterns. Thus we can see that the data set is a highly imbalanced data set which can make our work extremely challenging.

At this point we can use our algorithm to label the data set and compare the results with the expert's labels. To compare the results, we first apply the same division method to separate the tracks. We can then use a threshold to decide the whole sub-track's label. In this experiment, we employ 60% as the threshold value. For example, if the portion of abnormal points in one track is greater than 60%, we will label this whole track as abnormal. Using this approach, we find that 131 sub-tracks are classified as abnormal and the remaining 1,311 sub-tracks are all normal. Table III is the confusion matrix for the experiment.

From Table III we can see that 4 sub-tracks are classified as abnormal by both the expert and our model and 1301 sub-tracks are classified as normal by both too. On the other hand, another 10 sub-tracks are labeled as abnormal by the expert while normal by our model. The remaining 127 sub-tracks are classified as abnormal by our model while normal by the expert. And the overall accuracy of this detection result is 90.49%.

To understand the result, we investigate the 14 sub-tracks designated as abnormal by the expert. Among these, we find that the 4 which are further designated as abnormal by the algorithm are so labeled solely because of their direction. After investigating other tracks, we find that even if the ships deviate far from the lane the expert may still label them as normal. The intuition behind this is straight forward,

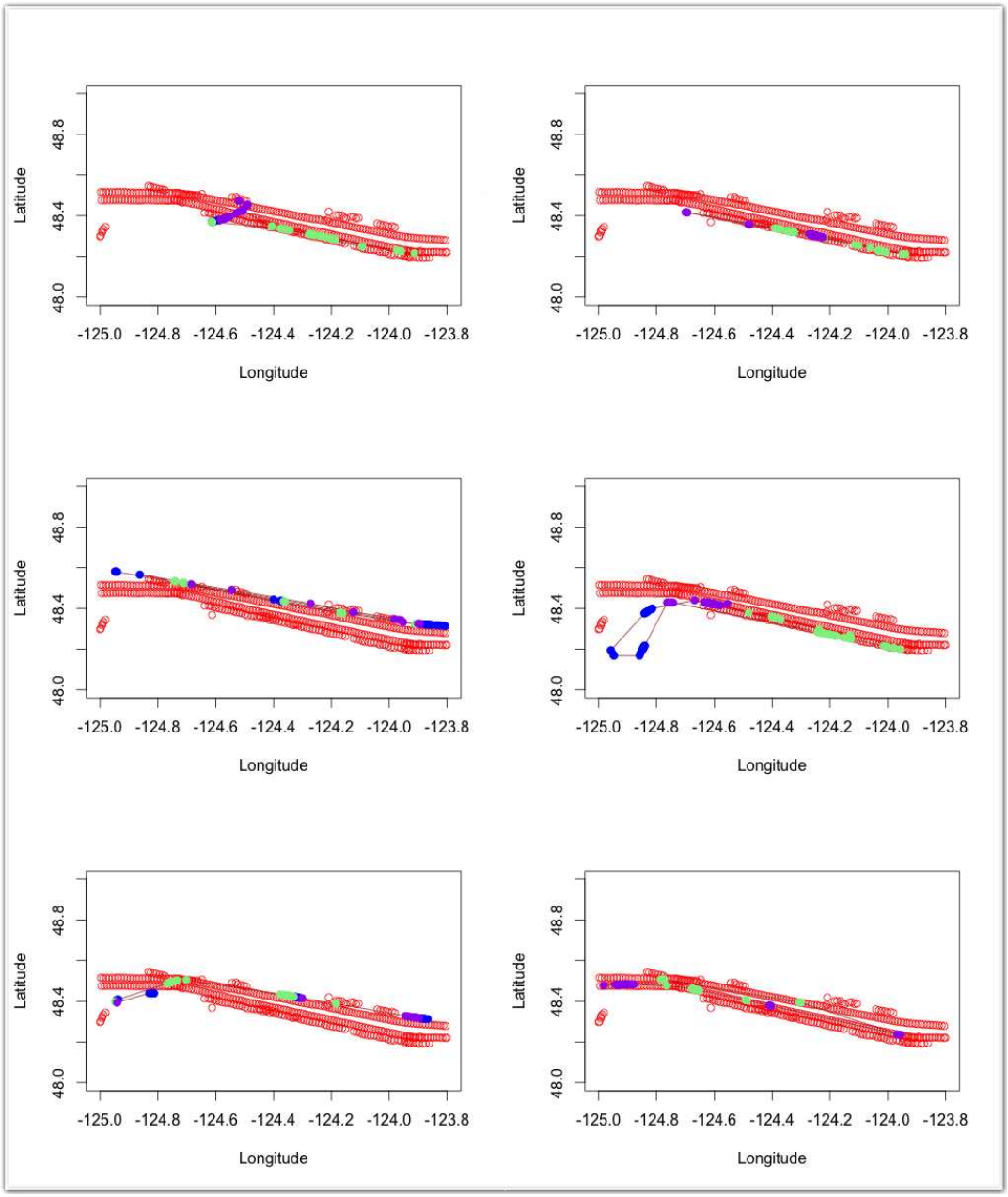

Figure 8. Six Examples of the abnormal trajectories detected by our framework in JUAN DE FUCA STRAIT area. GVs and SSPs are in red and the normal points are in green. The two types of abnormal points are in blue (abnormal in relation to ADD or RDD) and purple (abnormal in relation to CDD).

Table IV
CONFUSION MATRIX

	Abnormal (Our Model)	Normal (Our Model)
Abnormal (Expert's Label)	4	10
Normal (Expert's Label)	52	1376

the labelling process is based on Traffic Separation Scheme (TSS) [14] boundaries and the expert cannot affirm that a trajectory point far from the TSS Boundaries is abnormal. Considering this fact, in the following experiment, we only use the abnormal labels cased by CDD during the evaluation and we choose a lower threshold for deciding whole sub-tracks' labels. In the previous experiment we use 60% while we choose 10% here. It should be noted that the threshold can be adjusted based on the input from domain experts. In real-time application, it is not necessary to have the threshold

while labelling the new incoming points instead of tracks.

Table IV presents the improved results and we can see that the overall accuracy has been increased from 90.49% to 95.70% while keeping the same recall for abnormal cases.

V. Discussion And Conclusions

In this paper, we extend the work in [1] to propose a clustering-based anomaly detection model for maritime traffic data. An abnormality detection algorithm is presented based on three division distances (ADD, RDD and CDD). Our model is a fairly straightforward point-based approach, and is capable of handling complicated maritime traffic situations. One advantage is that the clustering process is associated with TSS Boundaries [14] which can assure a reliable clustering result for the following anomaly detection work. Another critical advantage is that besides position information (Longitude and Latitude), the model can also take speed and direction into account while deciding the abnormality of a single trajectory point. The model is also flexible enough for analysts to set their own thresholds for labeling whole trajectories. To evaluate the effectiveness of the model, a highly imbalanced data set from Juan de Fuca Strait area is used. There are 2,122 trajectories while only 14 of them are abnormal (imbalance rate $\approx 0.66\%$). Fortunately, as shown in Table IV, our model can detect 28.57% of the abnormal tracks while maintaining a relatively high overall accuracy (95.70%).

One limitation of this work is that the labelling process applied by the expert does not consider speed while our work takes this into account. This leads to another possible future direction, that is, more work should be done while labelling the data set to consider speed, which should reduce the false alarm rate. Another limitation is that the experiments are only conducted with data from Juan de Fuca Strait area and as a result, more experiments in other regions should be done to better illustrate the effectiveness of our approach.

Another possible future work is to try some classification algorithms designed for handling imbalanced data sets and the proposed specialized division distances can be used as the features of the classification models. Then comparisons between the results and our model's could be done to illustrate the effectiveness of the proposed division distances.

As stated in Section I, approaches based on predictive models usually predict future status information of a particular vessel and then compare the real data with the prediction to decide the abnormality. Thus, to improve the performance of the model, an ensemble model which incorporates a predictive model can be developed in the future. Specifically, once a trajectory point needs to be labeled, we can consider both its anomalous score and its deviation from the predicted position to get a more confident result.

Acknowledgment

The authors acknowledge the generous support of GSTS, Inc., ExactEarth, Inc. Marine Environmental, Observation, Prediction And Response Network (MEOPAR) and the Natural Science and Engineering Research Council of Canada for this research.

References

[1] B. Liu, E. N. de Souza, S. Matwin, and M. Sydow, "Knowledge-based clustering of ship trajectories using density-based approach," in *Big Data (Big Data), 2014 IEEE International Conference on*. IEEE, 2014, pp. 603–608.

[2] R. Laxhammar, "Anomaly detection for sea surveillance," in *Information Fusion, 2008 11th International Conference on*. IEEE, 2008, pp. 1–8.

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

[4] G. K. D. De Vries and M. Van Someren, "Machine learning for vessel trajectories using compression, alignments and domain knowledge," *Expert Systems with Applications*, vol. 39, no. 18, pp. 13 426–13 439, 2012.

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

[6] D. Nevell, "Anomaly detection in white shipping," *Mathematics in Defence*, 2009.

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

[8] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[9] R. Laxhammar, G. Falkman, and E. Sviestins, "Anomaly detection in sea traffic-a comparison of the gaussian mixture model and the kernel density estimator," in *Information Fusion, 2009. FUSION'09. 12th International Conference on*. IEEE, 2009, pp. 756–763.

[10] M. Ester, H. peter Kriegel, J. S, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise." AAAI Press, 1996, pp. 226–231.

[11] C. Veness. Calculate distance, bearing and more between latitude/longitude points.

[12] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining, (First Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2005.

[13] L. Etienne and T. Devogele, "Spatio-temporal trajectory analysis of mobile objects following the same itinerary," *Advances in Geo-Spatial Information Science*, vol. 17, no. 1, pp. 11–34, 2012.

[14] "Colreg.2/circ.57 new and amended existing traffic separation schemes," http://www.imo.org/blast/blastDataHelper.asp?data_id=14761&filename=57.pdf, accessed: 2006-05-26.