12th International Conference on Information Fusion
Seattle, WA, USA, July 6-9, 2009

Anomaly detection in sea traffic - a comparison of the Gaussian Mixture Model and the Kernel Density Estimator

Rikard Laxhammar
Informatics Research Centre
University of Skövde
Skövde, Sweden
rikard.laxhammar@his.se
&
Saab Systems
Saab AB
Järfälla, Sweden

Göran Falkman
Informatics Research Centre
University of Skövde
Skövde, Sweden
goran.falkman@his.se

Egils Sviestins
Saab Systems
Saab AB
Järfälla, Sweden
egils.sviestins@saabgroup.com

This paper presents a first attempt to evaluate two previously proposed methods for statistical anomaly detection in sea traffic, namely the Gaussian Mixture Model (GMM) and the adaptive Kernel Density Estimator (KDE). A novel performance measure related to anomaly detection, together with an intermediate performance measure related to normalcy modeling, are proposed and evaluated using recorded AIS data of vessel traffic and simulated anomalous trajectories. The normalcy modeling evaluation indicates that KDE more accurately captures finer details of normal data. Yet, results from anomaly detection show no significant difference between the two techniques and the performance of both is considered suboptimal. Part of the explanation is that the methods are based on a rather artificial division of data into geographical cells. The paper therefore discusses other clustering approaches based on more informed features of data and more background knowledge regarding the structure and natural classes of the data.

Keywords: Anomaly detection, sea surveillance, Density estimation, Gaussian Mixture Model, adaptive Kernel Density Estimation

1 Introduction

Anomaly detection has been identified as a critical component in order to achieve Situation Awareness in the context of information fusion and maritime surveillance ([1], [2], [3], [4], [5], [6] and [7]). Anomaly detection can be regarded as a method that supports the Situation assessment process at JDL level 2 by indicating objects and situations that, in some sense, deviate from the expected, known or "normal" behavior and thus may be of interest for further investigation.

Conceptually, the methods that have been proposed and implemented for anomaly detection in the maritime domain are based on statistical modeling of kinematical properties and behavior of individual vessels. In particular, the majority of the proposed feature models comprise a combination of momentary kinematical state features such as position, course/heading, speed, velocity vector, acceleration and angular velocity ([1], [3], [4], [6] and [7]). Furthermore, behavior over time is "captured" by clusters of vessel trajectories [5] or discrete state transition models where discrete vessel position state is a function of the previous position and velocity state [2].

Technically, the models and algorithms proposed for anomaly detection in the maritime domain are more or less data driven in the sense that normalcy is determined by machine learning algorithms analyzing a relative large set of historical data assumed to reflect normalcy. Generally, the methods are based on either neural networks that learn what is normal by unsupervised/semi-supervised learning ([1] and [2]), or more refined and transparent statistical/probabilistic models ([3], [4], [5] and [6]), or a hybrid where neural networks are used for determining parameters of a statistical model [7]. Considering the statistical methods, these can be categorized as parametric ([3] and [6]) or non-parametric [4], where the parametric methods assume that the model for (normal) data has a particular structure or belongs to a family of parameterized models. Structure and parameter setting can be purely data driven, e.g. unsupervised learning of structure and parameter estimation based on available data using machine learning techniques ([3] and [5]), or it can be a hybrid approach supporting the incorporation of human expert knowledge together with unsupervised/supervised learning [6].

So far, there has been no attempt to further evaluate and compare different methods for anomaly detection in the maritime domain. In this paper, we are concerned with a first attempt to investigate and compare the performance of two previously proposed statistical models for anomaly detection in sea traffic, namely the Gaussian Mixture Model (GMM) [3] and the adaptive Kernel Density Estimator (KDE), also known as the Parzen Window method [4].

The rest of the paper is structured as follows. First we describe a general approach to statistical modeling and anomaly detection based on vessel position and velocity vector, followed by a presentation of the GMM and KDE approaches as well as description of cell based normalcy modeling and anomaly detection. We then proceed to the experimental part of the paper where a novel performance measure is proposed and evaluated for the two models. The results are followed by a discussion and conclusion including directions for future work.

2 Statistical anomaly detection based on vessel position & velocity vector

The feature model we propose in this paper is identical to that previously proposed in [3] and [4], i.e. sea traffic is characterized by the momentary position in latitudinal and longitudinal coordinates and the velocity in latitudinal and longitudinal directions of individual vessels, i.e. a four-dimensional feature space. Considering "normal" or routine sea traffic, we assume that values of the four features constituting a single data vector, i.e. a single vessel observation, can be statistically modeled by a joint probability density function (PDF) that captures correlations between the features. Given such a PDF, we calculate the data likelihood for a particular vessel observation and the assumption that it is normal and that it is independent of previous observations, i.e. it constitutes an iid (independent and identically distributed) random sample from the underlying normal PDF. But, assuming that we have no corresponding PDF for abnormal/anomalous sea traffic (and prior probabilities for normal and anomalous vessels), we cannot determine the Bayesian posterior probability that the vessel observation is normal or anomalous. However, we can interpret the likelihood as an indication of the degree to which the corresponding observation is normal; if the likelihood is below a particular threshold, we assume that it is very unlikely that it was generated from the normal PDF and thus it is considered to be an anomaly.

As is the case in most real world applications, we do not have the true PDF for the feature values of normal data. In fact, we do not even know the parametric-form of the PDF, if such exists. The only thing we have is a (large) data sample that we assume has been generated from the PDF we seek. Thus, we need to estimate the PDF from our sample data without assuming any particular parametric form of the PDF.

2.1 Gaussian Mixture Model & Expectation-Maximization

The GMM is perhaps the most commonly used parametric density model for approximating arbitrary continuous multivariate PDFs when there is no particular knowledge or assumption regarding the parametric form of the density. It consists of K multivariate Gaussian distributions known as mixture components, where each component k has its own parameter set $\theta_k = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ where $\boldsymbol{\mu}_k$ is the mean value vector and $\boldsymbol{\Sigma}_k$ is the covariance matrix of the multivariate Gaussian. Each component of the mixture also has an associated mixing weight π_k and all weights are non-negative and sum to one. The probability density function for the Gaussian mixture model is given in equation (1):

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k N(\mathbf{x}; \theta_k) \qquad (1)$$

The total parameter set $\theta = \{\theta_1, ..., \theta_K\}$ for a *fixed* number of components K can be estimated from a training data set using the Expectation-Maximization (EM) algorithm originally proposed by Dempster *et.al.* [8]. Given a randomly initialized parameter set, the algorithm iteratively estimates the optimal parameter set $\hat{\theta}$ that maximizes the average log likelihood of the training set. However, depending on the initialization, the EM algorithm may converge to a local maximum that is different from the global maximum. Therefore, in order to suppress the sensitivity to initialization, we execute multiple runs of the EM algorithm with random initializations for a fixed number of components K and store the best estimate $\hat{\theta}_K$, i.e. the total parameter set that yields the highest average likelihood, as the optimal K-component model.

In order to determine a suitable number of components K, we estimate multiple mixture models with different number of components, i.e. with different values of K, and use the holdout method to determine when we are starting to over-fit the data. This is done in an incremental manner by, starting with K=1, estimate the optimal K-component model and the optimal (K+1)-component model and compare the average likelihood of a separate and non-correlated validation set; if the likelihood has decreased since adding another component, we assume that the (K+1)-component model is over-fitting the data and thus the K-component model is considered the optimal solution.

2.2 Adaptive Kernel Density Estimation

Generally, vessel traffic more or less follows sea lanes that can be described as sequences of straight line segments. Adopting the multivariate Gaussian for the spatial distribution of vessels implies that the location and extension of these sea lane segments are characterized by a Gaussian mean vector and covariance matrix, respectively, in the two-dimensional plane. However, one may argue that the two-dimensional Gaussian is not an optimal density model if we assume that the distribution of vessel positions along the major axis of the sea lane segments is approximately uniform, even though the variance along the axis perpendicular to the sea lane would nicely capture the vessel position offset relative the sea lane.

The adaptive KDE, also known as the Parzen Window method, is another technique for estimating unknown probability densities which, in contrast to the GMM, is purely non-parametric in the sense that it makes no assumption at all regarding the parametric form of the true density; the form of the estimated PDF is explicitly determined by the training data. This property gives the KDE an advantage over the GMM regarding the ability to accurately model arbitrary sea lanes as discussed above.

The probability density function for the KDE, given by equation (2), is determined simply by placing a kernel function on each and every observation $\mathbf{x}_k$ of the training set having size *K*, where each kernel is parameterized by its adaptive window width h_k estimated from the training set as described in [4].

$$p(\mathbf{x}) = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{(h_k)^d}\phi\left(\frac{\mathbf{x}-\mathbf{x}_k}{h_k}\right) \quad (2)$$

Similarly to [4], we adopt the multivariate Gaussian kernel with zero-mean and fixed covariance matrix Σ given in equation (3), where the kernel covariance is estimated as the sample covariance of the whole training set.

$$\phi(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{|\Sigma|}}\exp\left\{-\frac{1}{2}\mathbf{x}^T\,\Sigma^{-1}\,\mathbf{x}\right\} \quad (3)$$

2.3 Cell based normalcy modeling and anomaly detection

When the size and the complexity of the training set grows, parameter estimation and anomaly detection may become infeasible due to the complexity of the statistical models and algorithms. In the case of the KDE, the computational complexity of the algorithms for calculating the adaptive window widths and anomaly detection are quadratic and linear in the size of training set, respectively.

In order to suppress the computational complexity, we previously proposed that the surveillance area should be discretized into a uniformly sized grid where each cell has a local PDF estimate based on the local training data [3]. If the number of samples of the local training set is below specific threshold, the amount of training data is considered to be insufficient for normalcy learning and thus no PDF is estimated for that particular cell. Thus, observations within such cells would be considered anomalous, regardless of their features. Because grid discretization is done without considering the actual geographical distribution of data (e.g. without considering what is land and sea), there will most likely be one or more cells that lack a sufficient amount of data for normalcy modeling.

3 Experimental evaluation

In order to evaluate the GMM and KDE for statistical anomaly detection, we have trained and evaluated the implemented models based on a set of recorded AIS[1] data assumed to reflect normal/typical sea traffic. This data has been preprocessed, resulting in a set of trajectories that have been further divided into a training set for estimating the PDFs and an evaluation set for evaluating the performance of the estimated PDFs. The experiment consist of two parts corresponding to two novel performance measures we have defined for this application; *normalcy modeling performance* and *anomaly detection performance*.

3.1 Normalcy modeling performance

By normalcy modeling performance we mean the ability to estimate the *true* PDF for normal data; recall that our goal is to find a PDF that approximates the true PDF for normal feature values. Assuming that the observations from our evaluation set are normal, we would expect that their likelihood under the normal hypothesis in general is rather high. Therefore, we assess normalcy modeling performance based on the normal evaluation data likelihood for the GMM and KDE model; the model that assigns the largest likelihood for previously unseen normal data is assumed to better estimate the true normal PDF and is therefore regarded as superior.

3.2 Anomaly detection performance

The normalcy modeling performance proposed above is not enough if we are to evaluate two systems ability to detect anomalies. Yet, assessing anomaly detection performance is not straight forward; as discussed in previous work [3], there exists no established benchmark comprising a set of well defined maritime scenarios that are considered anomalous by domain experts. In fact, one may argue that assessing anomaly detection performance in this way is not appropriate as we are biased towards evaluating the systems ability to detect a particular class of anomalous situations, i.e. from a goal driven perspective, which stands in conflict with our definition of an anomaly as something prior unknown/ill-defined [3].

In this paper, we have addressed anomaly detection performance by evaluating the models ability to distinguish simulated trajectories, which have been generated by an arbitrary stochastic process, from real recorded trajectories that are assumed to have been generated by the true normal process, i.e. according to the normal PDF we are trying to estimate. The evaluation metric we use in this context is the number of consecutive observations required from an arbitrary anomalous trajectory segment in order to classify it as anomalous; the model that requires the least number of observations is regarded as superior.

3.3 AIS data description and preprocessing

The normal data used for both experiments in this paper has been generated from a large set of historical AIS data provided by Saab Transponder Tech. The data set corresponds to about three weeks of continuously recorded AIS traffic data that has been collected from vessels along the west coast of Sweden (Figure 1). All observations are unlabelled in the sense that there is no explicit label telling if a particular observation is normal or anomalous; however, we assume that the data set more or less reflect normal/routine sea traffic.

[1] Automatic Identification System (AIS). For more information on AIS: http://www.imo.org

From the raw AIS data we have extracted the MMSI[2], latitudinal and longitudinal position, course, speed and absolute timestamp for each recorded AIS report, where speed and course are transformed to longitudinal and latitudinal velocities. Thus, we have vessel reports containing six attributes corresponding to the vessel ID, timestamp, position and velocity in latitudinal and longitudinal space.

Even though we are not explicitly characterizing trajectories in our statistical model, we still want to group vessel reports corresponding to particular vessel trajectories for practical and experimental reasons. This "tracking" of the vessels is based on their MMSI which is assumed to uniquely identify each vessel present in the data set. In order to suppress the size of the data set without losing significant information, vessel reports from the trajectories are sampled at a fixed rate; whenever a tracked vessel has travelled a distance equal to or larger than the sampling distance, in these experiments set to 200 m, the current vessel report is added to the trajectory. In principle, this sampling for a particular vessel is continued until either no new vessel reports are received within a particular time interval (e.g. due to the fact the vessels leaves AIS-coverage), or it has remained stationary for more than 5 min in which case it is assumed to be moored[3]. In any of these cases, the trajectory is terminated and a new trajectory initiated when the vessel appears again or leaves its stationary state and starts to move again (i.e. it initiates a new route). In total, approximately 4,500,000 observations were sampled from the data, generating a total of 36,370 trajectories.

For each cell, we extract local trajectory segments from the part of the trajectories that are within the cell. The local trajectory segments are then divided into two sets corresponding to local training data, used for PDF estimation, and local evaluation data used for evaluating the normalcy modeling and anomaly detection performance. In our experiments, 80 % of the local trajectory segments were randomly selected for training and the rest used for evaluation. By partitioning the set of trajectory segments rather than the set of all observations, we suppress correlation between the training set and evaluation set (which would be the case if both contained observations from the same trajectory).

3.4 Simulation of anomalous trajectories

The trajectories used for evaluating the anomaly detection performance each consist of two parts; the first part corresponds to a normal segment and the second part an anomalous segment. The normal segment of each evaluation trajectory is constructed by first selecting a random trajectory from the set of normal evaluation trajectories, and then selecting a random break point along this trajectory; the subpart of the selected trajectory that extend to the selected breakpoint constitutes the normal segment. The anomalous segment, extending from the selected breakpoint, is generated according to a discrete stochastic process. Starting at t_0, the process samples new values for speed and course (independently of the previous values of the normal segment) from the two corresponding uniform PDFs within the intervals 0-30 knots and 0-360°, respectively. In the next time step t_1, the latitudinal and longitudinal coordinates of the trajectory are updated based on the current course, speed and time difference $\Delta t = t_1 - t_0$. Analogously to the sampling of the AIS trajectories described in section 4.1, we chose the next sampling time point t_1 such that the corresponding geographical distance between the sampling points equals a predefined sampling rate, in this case 200 m. Thus, Δt is a dynamic variable that depends on the current speed. At each new time step, there is a 10 % probability that new values for the speed and course, independently of each other, are sampled from the same uniform PDFs mentioned previously. Moreover, if the trajectory is about to enter a cell lacking a model or leaves the grid, a new course is sampled that ensures that the trajectory does not violate the boundaries. This process is repeated until the anomalous trajectory segment has a reached a predefined length.

3.5 Parameters and grid setup

For the parameter estimation of the GMM, 20 % of the available trajectory segments of the corresponding local training set were randomly and exclusively selected for the hold-out model validation, i.e. determining the appropriate number of mixture components. This value was chosen arbitrarily but considered to be enough for model validation.

For the normalcy modeling experiment, we have used the total data set, covering the west coast of Sweden, and applied a global grid of size 12 and 24 cells in longitudinal and latitudinal direction, respectively, resulting in cells having approximately quadratic form of size 22 km. However, only 112 out of 288 cells contained enough data for normalcy modeling; cells with less than 100 observations in the evaluation data were simply excluded from normalcy learning. The size of the global grid was chosen with consideration to the computational complexity of the parameter estimation in general and the KDE adaptive window widths estimation in particular; given the current grid size, estimation of adaptive window widths took approximately 24 h on a PC laptop.

The initial anomaly detection experiments presented in this paper have, for practical reasons, been limited to a local area outside of the harbor of Gothenburg. The size of the local grid was set to 6 (longitude) and 4 (latitude), where each cell has a length of approximately 2 km. The reason for choosing a smaller cell size was because of the relatively high traffic density around the harbor.

---

[2] Maritime Mobile Service Identity

[3] The attribute 'Navigational status' available in the AIS reports was empirically found to be unreliable.

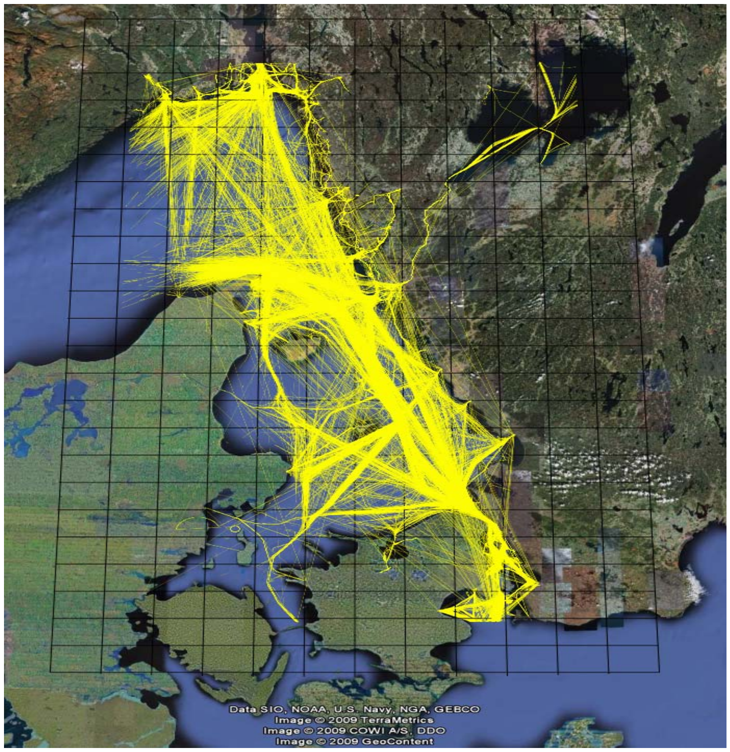

Figure 1: Plot of vessel trajectories extracted from the AIS-dataset, covering the western coast of Sweden[2]. The grid of size 12x24 used during the experiments is overlaid, where each cell is approximately quadratic with length 22 km

Figure 2: Plot of vessel trajectories extracted from the AIS-dataset covering an area outside the harbor of Gothenburg[4] located in the upper right corner. The grid of size 6x4 used during the experiment is overlaid, where each cell is approximately quadratic with length 2 km

[4] Screenshot taken from the Google Earth application.

Two thresholds, one for the GMM and one for the KDE, are used during the anomaly detection evaluation, were each has been tuned to a level where 1 % of the normal trajectory segments of the evaluation set contain one or more anomalous observations. This threshold is interpreted as an estimated false alarm rate of 1 per every 100 arbitrary normal trajectories observed by the models which we considered a reasonable operational alarm rate.

3.6 Results

Results for the normalcy modeling experiment is summarized in Table 1 below, where for each model (columns) the median and the first percentile (rows) of the log likelihood for all observations from the set of all available normal evaluation trajectories is presented.

	GMM	KDE
Median of log likelihood for total evaluation set	-1.8270	-1.3162
1st percentile of log likelihood for total evaluation set	-11.6569	-7.7414

Table 1: Summary of the log likelihood performance of the GMM and KDE over the total evaluation set

Table 2 below summarizes the anomaly detection experiment for the local area outside of Gothenburg, where the evaluation is based on 10,000 simulated trajectories that have been generated as described in section 3.4. Recall that each simulated trajectory has two parts; an initial normal segment that transits to an anomalous segment at a randomly selected point. Having fixed the rate of false alarms, i.e. rate of normal segments detected as anomalous as described in section 3.5, we calculate, for each model, the number observations it requires before it detects the first anomalous observation of each anomalous segment. The results are summarized as the mean and median for all the 10,000 anomalous segments.

	GMM	KDE
Mean number of observations required for detection of anomalous segment	17.72	17.43
Median number of observations required for detection of anomalous segment	12	12

Table 2: Summary of the number of anomalous observations required before detecting corresponding anomalous segment

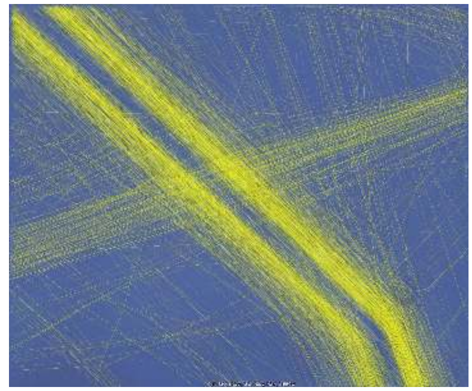

Figure 3: Plot of vessel trajectories in a cell of the global grid, illustrating a two-directional major sea lane and another minor sea lane crossing the major

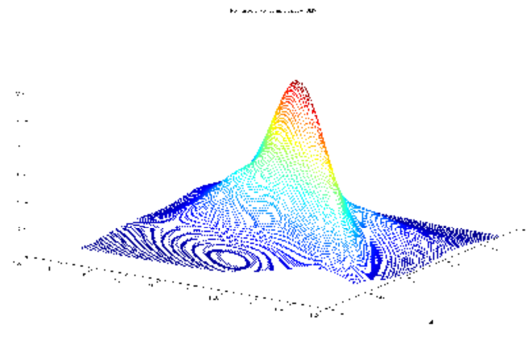

Figure 4: Visualization of normalized PDF in position space for the GMM for the cell in Figure 3, based on the mean and the covariance in position space for each component of the GMM. Note the unimodal peak halfway along the two parallel sea lanes, hiding the separation between them

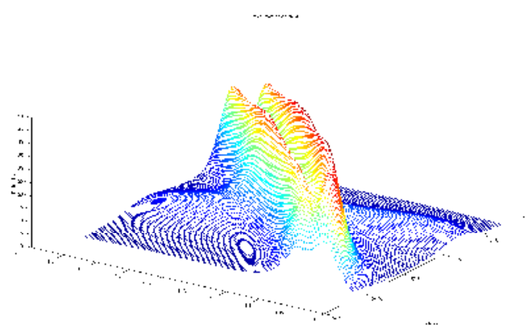

Figure 5: Visualization of normalized PDF in position space for the KDE for the cell in Figure 3, based on each kernel location and total sample covariance in position space. Note how this PDF, in contrast to the one in Figure 4, nicely discriminates the two parallel lanes and approximates a uniform density along them

4 Discussion

4.1 Normalcy modeling performance

Looking at Table 1, we may interpret the results as the KDE being superior in the sense that the median data likelihood of a normal observation is higher than for the GMM model; the median of the likelihood for the normal evaluation set is approximately 5/3 times the corresponding likelihood for the GMM model. This ratio is even larger, approximately 50:1, in favor of the KDE when considering the first percentile, i.e. the least likely observations encountered in the evaluation set. These results, together with the plotted PDFs in Figure 4 and Figure 5, support the hypothesis that the KDE more accurately captures features in normal data related to e.g. the spatial distribution of the vessel location along sea lanes.

One could argue that the assumption that the recorded AIS-data strictly reflects normal traffic is not realistic and feasible. In fact, there may be occurrences in this data that some people, in some contexts, would consider anomalous and worthy an alert. However, by taking the median and not the mean value of the (log) likelihoods for all observations, our assessment of the normalcy modeling performance is robust with regard to such anomalies. Consider for example the case that we actually have a true anomaly in our evaluation data. A “good” normalcy model would then assign this observation a (very) low likelihood, while a “bad” normalcy model might assign it a considerably larger likelihood. Thus, taking the mean would penalize models that detect these true anomalies in favor for models that do not.

4.2 Anomaly detection performance

Having observed the intermediate results from Table 1, it would seem a reasonable hypothesis that that the KDE have an advantage when it comes to discriminating observations sampled from PDFs other than the normal; if areas in feature space covered by normal data have a relatively high density, it must follow that other (anomalous) areas have a relatively small density as the integral of the PDF is normalized to one. Yet, considering Table 2, it appears that there is no significant difference between the models when it comes to actually discriminating observations from arbitrary trajectories from those of normal trajectories; the mean and the median of the number of observations required before the detection of an anomalous segment is approximately 17 and exactly 12, respectively, for both models. Translating these results to distances implies that the models detect the anomalous segments after 3,4km (mean) and 2,4km (median), respectively. This may be considered rather far given the rather constrained behavior of the normal vessel trajectories in the area and the random sweeping behavior of the anomalous trajectories; we expect that a reasonably effective anomaly detector should detect such anomalous behavior at an earlier stage.

4.3 Limitations of the cell based approach

One reason for the rather disappointing anomaly detection results may be that we have modeled the problem incorrectly in one or more aspects. The cell division model described in this paper serves well for suppressing the size of the data set used for statistical modeling. However, it is rather naïve in the sense that it does not take into account the actual vessel traffic distribution. Moreover, one may argue that simply partitioning the data based on only geographical proximity of individual observations, regardless of other contextual information such as trajectory origin and destination, vessel type etc, is rather artificial and do not reflect a natural partitioning of the data. Consider for example the cell in column 2 and row 2 (starting from the top) in Figure 2. Clearly, there appears to be one major sea lane and one minor sea lane that extends in north-westwards direction from the major. Because of the relatively low traffic intensity in the minor sea lane in this particular cell, traffic along and close to the lane will be regarded as rather unlikely, perhaps even anomalous, even though there clearly is a sea lane that is part of normalcy. This issue highlights that an anomaly is perhaps more of a relative concept that should be considered in the right context. In the example discussed above, it would probably be more appropriate to model the two sea lanes separately by two different statistical models and evaluate new observations by first determining the most likely sea lane (model), e.g. based on a Bayesian posterior analysis using prior probabilities for the two sea lanes and their likelihoods for the observation, and then do anomaly detection relative the most likely model.

Generally, when attempting statistical modeling of large and complex data sets, we believe that it is important that we use all available background knowledge regarding the structure and natural classes in data. In the context of statistical modeling of AIS data, this implies that information regarding origin, destination, vessel type etc that are directly accessible from the AIS data, together with other contextual information, such as season and time of day, should be exploited during modeling. Yet, we still believe that partitioning data into cells in some way could prove a suitable approach as part of a hierarchical clustering when dealing with rather long and complex trajectories.

4.4 Future work

Following the discussions of the previous section, a natural direction for future work is to consider alternative approaches to partitioning of the vessel reports, such as considering vessel class and the origin and the destination of the corresponding trajectories. In [4], the authors propose extracting “motion patterns” from the training set where each motion pattern corresponds to a cluster of trajectories having more or less the same origin. Assuming that vessels originating from the same area, typically harbor or area where coverage/tracking starts, have similar trajectories and behavior to a larger extent than arbitrary vessels, this partitioning is indeed an interesting approach to implement and evaluate. Such a trajectory clustering could potentially be part of a hierarchical clustering that also involves further clustering based on destination and cell division as discussed in section 3.7.

Another direction for future work is development of the feature model used for the statistical analysis. As pointed out in [3], the momentary position- and velocity vector model is rather limited in the sense that 1) change of individual vessel behavior over time and 2) spatiotemporal relations to other vessels are not statistically modeled. Moreover, data sources and data types related to information other than vessel kinematics, such as various ship data bases and other intelligence, could potentially enhance the prior data partitioning of vessel trajectories, or serve as basis for statistical modeling in other feature spaces.

5 Conclusion

In this paper we have presented and evaluated two methods previously proposed for statistical anomaly detection in sea traffic; the GMM & EM, and the adaptive KDE. These two methods have been used for estimating the joint PDF for the position-velocity vector of normal vessel trajectories within confined geographical cells. A novel performance measure for evaluating anomaly detection performance, together with an intermediate performance measure related to normalcy modeling performance, have been proposed and evaluated based on recorded AIS data, assumed to reflect normal traffic, and simulated stochastic trajectories assumed to be anomalous.

Intermediate results indicate that the KDE is superior to the GMM in the sense that the median likelihood of a previously unseen normal observation is higher for the KDE; this supports our prior hypothesis that the KDE more accurately characterizes features in the data, such as sea lanes. However, results from the anomaly detection experiments were somewhat surprising in the sense that there was no significant difference between the two methods, even though intermediate results regarding normalcy modeling performance suggest that KDE potentially has some advantage. Moreover, anomaly detection results for both models were quite disappointing as the expected trajectory distance observed before detection was considered rather large given the rather constrained normal trajectories and the sweeping and stochastic character of the anomalous trajectories.

One explanation for the suboptimal anomaly detection performance may be that we have modeled the problem incorrectly in one or more aspects. In particular, we question the rather blunt cell division approach to partitioning of the vessel reports, considered as rather artificial and not reflecting a natural partitioning of the data. Rather, other clustering approaches based on more informed features of data and more background knowledge regarding the structure and natural classes of the data should be used. Hierarchical clustering, considering first trajectory origin and second trajectory destination would probably be more appropriate.

Acknowledgements

This research has been supported by the graduate school Intelligent Systems for Robotics, Automation and Process Control (RAP) of the University of Örebro, Sweden, Saab AB, Sweden, and the University of Skövde, Sweden.

References

[1] B.J. Rhodes, N.A. Bomberger, M.C. Seibert, A.M. Waxman, "Maritime situation monitoring and situation awareness using learning mechanisms", Military Communications Conference 2005, Atlantic City, NY, USA, October 17-20, 2005.

[2] B.J. Rhodes, N.A. Bomberger, M. Zandipour, "Probabilistic Associative Learning of Vessel Motion Patterns at Multiple Scales for Maritime Situation Awareness", The 10th International Conference on Information Fusion, Quebec, Canada, 2007.

[3] R. Laxhammar, "Anomaly Detection for Sea Surveillance", The 11th International Conference on Information Fusion, Cologne, Germany, 2008.

[4] B. Ristic, B. La Scala, M. Morelande, N. Gordon, "Statistical Analysis of Motion Patterns in AIS Data: Anomaly detection and Motion Prediction", The 11th International Conference on Information Fusion, Cologne, Germany, 2008.

[5] A. Dahlbom, L. Niklasson, "Trajectory Clustering for Coastal Surveillance", The 10th International Conference on Information Fusion, Quebec city, Canada, 9-12 July, 2007

[6] F. Johansson, G. Falkman, "Detection of vessel anomalies – a Bayesian network approach", 3rd International Conference on Intelligent Sensors, Sensor Networks and Information Processing, 2007.

[7] J.B. Kraiman, S.L. Arouh, M.L. Webb, "Automated anomaly detection processor", Proceedings of SPIE: Enabling Technologies, for simulation science VI, A. Sisti, D. Trevisani, Eds., July 2002, Vol. 4716, p. 128-137.

[8] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", Journal of the Royal Statistical Society, Series B, 39(1):1–38.