Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Procedia 3 (2014) 720 – 729
Transportation Research Procedia
www.elsevier.com/locate/procedia

17th Meeting of the EURO Working Group on Transportation, EWGT2014, 2-4 July 2014, Sevilla, Spain

Grid size optimization for potential field based maritime anomaly detection

Ewa Osekowska*, Henric Johnson, Bengt Carlsson

Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden

Abstract

This study focuses on improving the potential field based maritime data modeling method, developed to extract traffic patterns and detect anomalies, in a clear, understandable and informative way. The method's novelty lies in employing the concept of a potential field for AIS vessel tracking data abstraction and maritime traffic representation. Unlike the traditional maritime surveillance equipment, such as radar or GPS, the AIS system comprehensively represents the identity and properties of a vessel, as well as its behavior, thus preserving the effects of navigational decisions, based on the skills of experienced seamen.
In the developed data modeling process, every vessel generates potential charges, which value represent the vessel's behavior, and drops the charges at locations it passes. Each AIS report is used to assign a potential charge at the reported vessel positions. The method derives three construction elements, which define, firstly, how charges are accumulated, secondly, how a charge decays over time, and thirdly, in what way the potential is distributed around the source charge.
The collection of potential fields represents a model of normal behavior, and vessels not conforming to it are marked as anomalous. In the anomaly detection prototype system STRAND, the sensitivity of anomaly detection can be modified by setting a geographical coordinate grid precision to more dense or coarse. The objective of this study is to identify the optimal grid size for two different conditions – an open sea and a port area case.
A noticeable shift can be observed between the results for the open sea and the port area. The plotted detection rates converge towards an optimal ratio for smaller grid sizes in the port area (60-200 meters), than in the open sea case (300-1000 meters). The effective outcome of the potential filed based anomaly detection is filtering out all vessels behaving normally and presenting a set of anomalies, for a subsequent incident analysis using STRAND as an information visualization tool.

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014

* Corresponding author. Tel.: +46-734-22-35-10; fax: +46-455-38-50-57.
E-mail address: ewa.osekowska@bth.se

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014
doi:10.1016/j.trpro.2014.10.051

Keywords: anomaly detection; maritime traffic; potential field; AIS;

1. Introduction

Maritime traffic management decisions were in the past taken on the basis of human experience, and often – a hunch. The operator keeping track of the vessel movements faced a highly sophisticated task to monitor the open sea and harbors in order to detect suspicious behavior and critical situations. Today, expert analysis is informed and augmented by advanced automated systems. However, methods for automated anomaly detection with the aim of supporting the operators in their decision making need to be improved (Department of Homeland Security, 2005). In addition, maritime traffic is continuously increasing, and the importance of using the water as a transportation highway has created a demand for better systems to support agencies responsible for surveillance. The demand for increased security can be satisfied by increasing the surveillance capability and by proactively working to minimize the impact from threats to traffic safety. The maritime domain, being a part of a shared logistic chain, functions as an avenue for a multitude of opportunities and threats (Acciaro & Serra, 2013; Bichou et al., 2013). Much of what occurs in the maritime domain is also difficult to observe and assess, with respect to vessel movements, activities, and intentions. Hence, it is imperative to make the best possible use of all available data in order to detect and visualize out-of-the-ordinary behavior. Just like road traffic, the maritime traffic tends to concentrate locally in certain areas. The distances between ships tend to be smallest in harbors and anchorage zones, and largest on open sea. As a consequence, in some areas the traffic is precisely regulated, while in others it may be less restrained or completely unorganized. In case of an accident, the consequences can be serious both in personnel, environmental and economical terms.

A monitoring system receives information from various sensors based on technologies that are used to quantify location, speed, course, et cetera. In addition, there is a multitude of civil and military maritime traffic surveillance operations that extract data for their own purposes using proprietary management systems. One common system is the AIS (Automatic Identification System) – an automated tracking system used on vessels for identifying and locating ships by electronically communicating with nearby vessels, AIS on-shore base stations and satellites. The information provided by AIS includes a unique vessel identification, position, course and speed. Following the letter of the International Maritime Organization's (2002) International Convention for the Safety of Life at Sea, AIS transponders have to be installed on vessels with a gross tonnage of 300 or more, and on all passenger ships regardless of size. This development of global ship tracking systems in the recent years opens possibilities of advancing maritime safety and security beyond simple collision prevention, e.g., to inform about detected anomalies. Therefore, this study focuses on utilizing the increased surveillance capability, and, more directly, on optimizing a novel method for anomaly detection in maritime traffic by the use of AIS information. The goal is to scale the system settings to accommodate the specificity of traffic in dense port areas, as well as the less regulated open sea. A more specific objective of this study is to identify the optimal grid size for areas with intensive and sparse traffic, by examining the influence of the grid precision on the anomaly detection outcomes. The grid size terminology is further described in section 3. This study uses the concept of potential fields (Osekowska et al., 2013) for data modeling and anomaly detection in maritime traffic. The potential fields are traffic patterns built based on the recorded vessel positions and behaviors history. Anomalies are detected as conflicts between the vessel's current potential, and the local potential fields. Visualizing the potential fields using modern rendering techniques can provide the maritime operators with an automated help in identifying traffic situations that merit further investigation.

2. Background and related work

The detection of anomalies in the maritime domain is a complex process. It can be broken down into several steps including data acquisition, information fusion, situational awareness and anomaly detection. The problem of anomaly detection is also generic, i.e., not tied to the maritime domain, and is present in many other domains. One can view the data acquisition as the first step, providing the raw input material for anomaly detection. Normal behavior models and, consequently, anomalies are then derived from that data. Therefore the success of the

following steps largely depends on the data collection and the quality of that data. If the data is absent or insufficient, the rest of the process may be compromised.

The AIS monitoring system developed in recent years was not to deal with anomaly detection, but rather to register as much traffic data as possible, with the focus on remote traffic oversight and situational awareness. The problem of being flooded with data is referred to as "data overload" (Kessler, 2009), and it is a major motive for the development of anomaly detection, which role is to find and highlight elements worthy of interest. As previously mentioned, anomaly detection is a complex process, which design should ensure providing the operators with the most suitable situational awareness aids to support them in decision-making. The operator's work normally begins where the system's works ends. Within the problem of anomaly detection in maritime traffic two main types of approaches emerge, one focusing on defining anomalous behavior, the other on exceptions from the modeled normal behavior. In the former case, the definition of anomalous behavior is often defined based on expert knowledge, and then used as a base for defining anomaly recognition and detection rules. In the latter case – the normal behavior is captured in form of a model, and the deviations from it are considered anomalous. It is preferable to leverage the advantages of human expertise in combination with computing techniques, since machines are suitable of processing large quantities of data; humans are not. However, the human reasoning capacity is normally better to that of the machine. Therefore, the purpose of the research is to improve the performance of the operators, not to replace them.

In this paper, the previously presented STRAND (Seafaring TRansport ANomaly Detection) system (Osekowska et al., 2013) is further developed and improved in terms of parameter settings, i.e., finding the optimal grid size for areas of different traffic intensity. The system builds a model of normal behavior, based on the past AIS vessel tracks. This results in a visual presentation that displays distinct patterns of normal behavior inherent in the recorded maritime traffic data. Based on the created model of normality, the system can then perform the anomaly detection on real-life maritime traffic data. The objective is further to provide early detection and prediction of possibly malicious events. For the system to be useful, the predictions must be made at such an early stage to give time for the authorities to handle a critical situation. For example, a vessel travelling in the wrong direction could be a potential and severe collision risk.

Although the research community in this domain uses different approaches to handle the problem of anomaly detection, they all share the same objective: to exploit information from sources within the maritime domain with the end purpose of improving the security of people and environment. A number of studies has been focused on this problem. For instance, in Van Laere and Nilsson's publication (2009) anomalies are identified through assembling expert knowledge by practitioners. Ristic et al. (2008) proposed a statistic method to extract normal behavior from raw data, in which they demonstrated the presence of anomalous trajectories, i.e., vessel passing through locations not belonging to the normal model. In another publication a vectorial traffic characterization is proposed indicating that Gaussian mixture modeling is difficult to use (Baldacci & Carthel, 2009). Therefore, the authors propose a simplified version of the problem: traffic characterization of main sea-lanes only. Most of the published works use AIS data, acquired through different sources. This has inspired research into developing a unified AIS anomaly simulator. With the tool presented by Baldacci (2008) it is possible to study and simulate AIS status, kinematic and positional anomalies.

In an article by Kazemi et al. (2013) a visualization tool is presented that makes use of open maritime data to detect anomalies. A case study was also performed and evaluated by the Swedish coast guard with real data from the Baltic Sea. In another publication using expert systems, Riveiro and Falkman (2009) proposed applying interactive visualization techniques to enrich their rule-based anomaly detection. Another joint work by Riveiro, Falkman and Ziemke (2008) combines a visual approach (self-organizing maps) with non-parametric statistics (density estimation with Gaussian mixture model) and probabilistic theory (Bayes theorem). In a publication by Perera and Oliveira (2012) machine-learning techniques, e.g., neural networks are presented. Presentations of proprietary solutions for normalcy learning can also be found in the maritime security literature, e.g., that by Rhodes et al. (2008).

Detection of anomalous trajectories is an alternative formulation of the traffic safety problem that Laxhammar and Falkman (2013) addressed by presenting their Inductive Conformal Anomaly Detector (ICAD), which is a parameter-light anomaly detection algorithm. Finally, the problem of anomaly detection in the maritime domain has also been a subject for multi-agent systems (Brax et al., 2013), in which the system uses a rule engine and evaluates vessels based on successive number of alerts. Invalidated alerts (determined by an operator) are then returned to the multi agent system for a learning step, retraining the detection engine to improve its accuracy.

Over the years, several anomaly detection systems have been developed, used and evaluated. The following paragraph remarks some of them. The SeeCoast system (Seibert, 2006) is built on the Hawkeye system and installed in the Joint Harbour Operations Center in Portsmouth, Virginia. The SCANMARIS system (Morel et al., 2009) creates a traffic picture that is accessed by both a rule engine and a learning engine. The prototype is tested at the Centre Regional Operationnel de Surveillance de Sauvetage Corsen. The LEPER system (Griffin, 2009) was supported by the Office of Naval Research, USA, and was tested at the Joint Interagency Task Force South (JIATF South). The SECMAR system (Gehant et al., 2009) is Thales underwater systems for port security and a prototype has been tested in the strategic harbor of Fos-sur-Mer in France. Finally, the MALEF system (Tozicka et al., 2008) is a generic framework used for distributed machine learning and data mining, in which the architecture allows agents to improve using information from other learners in the system.

3. Potential fields

The general idea of the proposed method, applying potential fields to maritime traffic, is for the geographical traces of vessel movements to assign charges to all passed locations. A collection of charges distributed over an area generates a potential field, which is locally weaker or stronger depending on the density and strength of surrounding charges. The three main concepts are the total strength of a local charge, the decay of potential fields, and the distribution of a potential field around its charged source (Osekowska et al., 2013).

Table 1. Listing of variables used in equations 1 to 4.

Variable	Description
c	Elementary charge
C	Total charge
P	Potential value
k, i	A point in the traffic model
lat_k, lon_k	Geographical coordinates (latitude and longitude) of k
τ	The time period of charge accumulation
σ	Standard deviation of the Gaussian distribution
a	Ratio constant compensating the difference in the geographical length per a unit of longitude vs. latitude
$d(t)$	The non-increasing charge decay function

Table 1 lists the parameters used in equations 1 to 4. The strength of a charge c is a metric nearest to the AIS surveillance data. Each vessel tracked by AIS is characterized by a collection of n numerical and verbal properties including vessel's static parameters (e.g., identification number, call sign, name), as well as the current state of its dynamic behavior (e.g., speed or course).

$$c_{lat_k lon_k} = \left\langle c^1_{lat_k lon_k}, c^2_{lat_k lon_k}, \ldots, c^n_{lat_k lon_k} \right\rangle \tag{1}$$

The total charge C at a location k is counted as the sum of all local charges c accumulated over a time period τ (2). The more vessels visits are reported in a location, the higher potential builds up in it, and around it.

$$C_{lat_k lon_k} = \sum_{t=0}^{\tau} c_{lat_k lon_k} \tag{2}$$

The potential field formed by a single charge is strongest at the location of the charge, and dissipates within a radius around it. Areas, where a potential is very strong, represent a traffic pattern and belong to the model of normal

behavior. On the other hand, areas where a potential is very weak or none, signalize absence of normal behavior, and therefore – anomaly. In this study the anomaly levels are determined using minimal potential thresholds. The total potential at location k is the superposed potential generated by all surrounding charges in locations i, decreased by the distance between these locations. Here the potential distribution P is described by two-dimensional Gaussian smoothing, using Euclidean distance for measuring the radius between two points:

$$P_{lat_k lon_k} = \sum_i \frac{1}{2\pi\sigma^2} e^{-\frac{(lat_k - lat_i)^2 + (a(lon_k - lon_i))^2}{2\sigma^2}}, \tag{3}$$

where a is a constant multiplier compensating the proportional difference between the Euclidean distance per a unit of longitude versus a unit of latitude, and the standard deviation σ is assumed same for both dimensions.

It is desirable for the potential fields modeling maritime traffic, to evolve over time and reflect real-world traffic patterns changes. Instead of addressing the continuity of real time by applying constructs such as sliding time frame or data window (Ristic et al., 2008), (Brax et al., 2010), in this investigation a field decay factor is used. It enables to continuously update and retrain the model by representing charge at a location as a function of time:

$$C_{lat_k lon_k}(t) = \sum_{t=0}^{\tau} d(t) c_{lat_k lon_k}, \tag{4}$$

where $d(t)$ is a non-increasing decay function with limit at zero, describing the decrease of a local charge over time.

4. Experiment setup

The central aim of the study and purpose of the developed anomaly detection prototype is to provide warnings about unusual behaviors (potential incidents) in maritime traffic that may impact its safety. In terms of detection performance, the goal is to avoid false negative results and minimize false detections. Deficient information on objectively labeled incidents on the sea forces researchers in this area to treat the traffic surveillance data as an unlabeled dataset. The evaluation of the detection accuracy is in many cases based on subjective judgment or demonstrative cases, less often on expert opinion. The investigations in this study target a wider perspective and therefore do not focus on specific individual cases, but overall quantifiable relation of anomalous behavior to the normal, benign traffic. This study concentrates primarily on investigating the influence of the different geographic distance granulation sizes on the anomaly detection output. The aim is to find an optimal grid size for discretizing geographical locations in maritime anomaly detection.

The technical limitation for tracking data update period is 90 seconds, which means that every time a snapshot of the state of maritime traffic is acquired, it takes another 90 seconds to acquire the next one. In that time a vessel moving with speed of 2-5 knots, which is typical for harbor and anchorage areas, will change its position by approximately 90 to 230 meters. On open sea, a typical slow steaming cargo ship with speed of 10-20 knots will move 460 to 930 meters in the same time (Cariou & Notteboom, 2011). Therefore, it appears rational to investigate the traffic in restrained and regulated circumstances of harbor-like areas separate from freely sailing open waters. The working hypothesis is that the grid size between 50m and 200m is suitable for areas with dense traffic (e.g., harbors), and grid size between 500m and 1000m is more suitable for open sea areas.

The difference in distances crossed by slower and faster moving vessels is additionally a factor that could cause imbalance in the strength of potential. Namely, potential created by slow and static ships would be focused in a very small area, which would make it locally very intensive, while faster moving ships would create a track of weak potential stretched over a larger area. That imbalance is compensated by assigning weights to the charges dropped by vessels. In the current implementation of the STRAND system the weight of a charge is proportional to the square root of the ships speed. That way the charges generated by slow and static vessels are made moderately weaker, and by faster vessels – stronger.

Computational description of the proposed method entails discretization of all continuous variables, i.e., time, space, speed, course and other related. Course and speed over ground are stored without alterations with precision to 0.1° and 0.1 knot respectively. For maritime traffic modeling and detection, however, they are grouped by ranges. Course is divided into 8 equal 90 degrees intervals: N, NE, E, SE, S, SW, W, and NW. Speed ranges are not equal in size, and correspond to a speed division levels common for maritime traffic (in knots): Static (0-1), Very slow (1-7), Slow (7-14), Medium (14-22), Fast (22-30), Very fast (30-45), Ultra fast (45-60), and Probably flying (beyond 60). Speeds exceeding 60 knots are achievable only by relatively small vessels with very powerful engines. At these velocities most of the vessel's body is lifted above the water surface, and the draught of decreasingly submerged vessel is minimal. The AIS system is also present among helicopters, seaplanes and other aircrafts connected to the sea, therefore all AIS reports with speed 60 knots or more are categorized as probably flying.

The precision of course and speed is reduced to a nominal scale for two reasons: to build an understandable parameter selection in user interface, and to define the speed and course value granularity for data modeling and detection sensitivity. Besides these metrics the system uses nominal attributes: vessel type (Passenger, Cargo, Tankers, etc.) and time of day. The time of day divides 24 hours into four equal daytime groups: morning (6-12), afternoon (12-18), evening (18-0), and night (0-6).

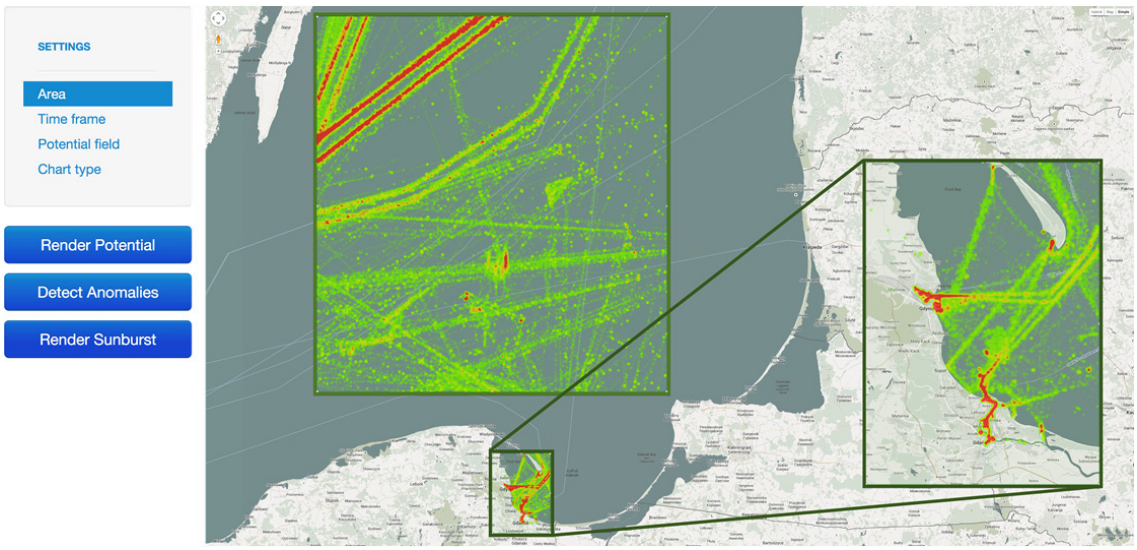

Fig. 1. STRAND interface with marked focus areas: open Baltic Sea to the north and harbor area of Gdansk bay at the Polish coast.

The hypothesis is tested using the anomaly detection prototype system STRAND, with the parameters and detection set to satisfy the assumptions and fit the open sea and harbor area scenarios. The traffic detection output is measured as the count of anomalies of specific types and the total number of examined vessels. The detection is performed for areas of specific type, starting with a preliminary selection of a rectangular area in the Baltic region containing both – harbor and open sea areas. Subsequently, two subsections of that area are run under detection separately, one covering open sea, the other marking a harbor area.

The measurements resulting from the experiment are then compared in form of superposed plots and analyzed in a quantitative manner. The analysis is additionally supported by results of further experiments specifically targeting selected detection types: course and speed.

The experiment was performed for an open sea area (large rectangle in figure 1) delimited by approximate coordinates: 55.0S, 16.9W, 56.8N, and 19.9E, using a data set capturing 20 days of traffic. The harbor area includes two major ports of Gdynia and Gdansk, within the Gdansk bay and the bay of Puck, along with a stretch of coast and a fragment of the Hel peninsula (small and magnified rectangle in figure 1). The approximate coordinates are 54.3S, 18.4W, 54.7N, and 18.9E.

5. Results and analysis

The figures presenting results display the anomaly detection statistics for all maritime traffic in the selected areas as a function of grid size.

5.1. Open sea and harbor case comparison

The plots in figures 2 and 3 represent the numbers of detections of types: waypoint, course, speed and daytime. The total is a sum of all positive detections regardless of type. The course, speed and daytime are detections made respectively based on an observation of a time of day, course or speed, with which the ship travels, and which is unusual for ship's present location. It is important to note that these detections may overlap, e.g., a ship may travel with anomalous course and time of day, but at a speed that is normal for its location. The waypoint detection signalizes the most severely anomalous behavior, and is triggered when a vessel is observed in an area in which no prior visit of any other vessel was ever observed. As a consequence this type of anomaly also indicates anomalous speed, course and daytime, increasing their detection count.

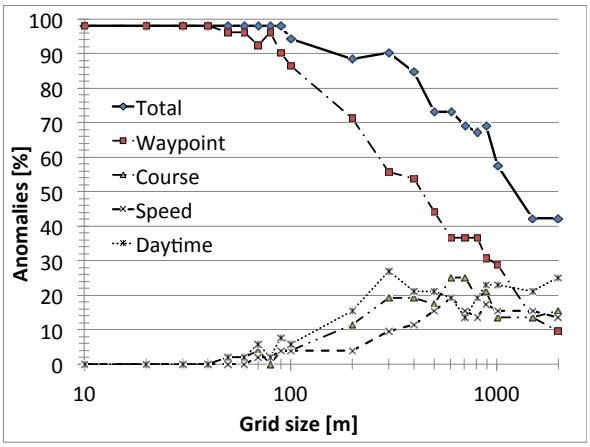

Fig. 2. Open sea detection percentage as a function of grid size.

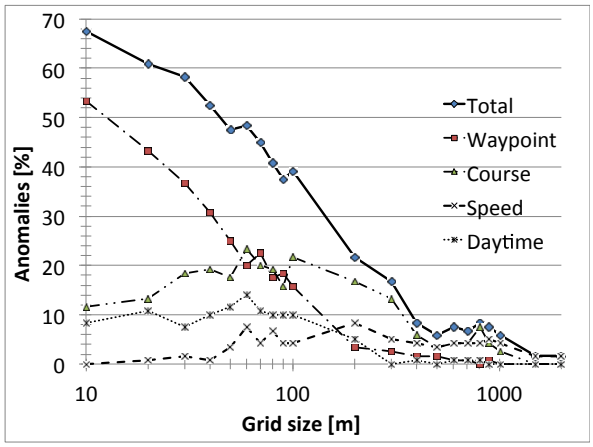

Fig. 3. Harbor area detection percentage as a function of grid size.

In practice the STRAND system recognizes the waypoint detections as a deficit of any manner of recorded vessel presence at an examined location. The detections specific to speed, course and daytime and carried out if a waypoint detection result was not positive. The observed tendency is that for oversensitive detection settings (where the grid is too dense), the proportion between the general (waypoint) and specific detection (course, speed, daytime) is strongly skewed towards the former type. The reason being that in a high resolution grid the distances between waypoints recorded in fact relatively close to one another, are so many grid nodes apart, that they cannot create a common potential field (i.e., a traffic pattern).

For the harbor case there is a range between 60 meters and 200 meters where an increased number of course and speed anomalies are found. At the same time the number of waypoint anomalies decreases and flattens out after 200 meters grid size. For the open sea case the range of the optima starts at 300 meters and decreases after 1000 meters for course and speed. The waypoint is steadily decreasing until 2000 meters. Interestingly, the tendency of the detection specific to daytime yields very similar results in that range.

5.2. Analysis of the speed and course range binning

In order to analyze the discretization of course and speed, the acquired outcomes of detection are compared to the additional two result sets for course and speed. The added sets are produced by performing detection on data with

altered speed and course. One of the sets is computed for the real course (say NE) altered as if the ship was travelling with course slightly more to the left (the altered course is N). In the second set, the course is altered to the next right (e.g., NE altered to E). In case of speed, the next lower and the next higher speeds are tested, with the exception for the lowest and highest speed, which remain unaltered (as no lower or higher speed is available).

The diagrams in figures 4 and 5 plot the results of detections performed on the traffic tracking data with the real speed and course, and with their altered (increased and decreased) values. It is instantly apparent that the amount of detections for the true observed speed and course is lower than for their altered values (in diagrams labeled as *Low* and *High* for the altered speeds, and *Left* and *Right* for the courses). This is true for all grid sizes.

An interesting observation can be made about the course. The plotted results for the altered courses: *Left* and *Right* are nearly the same, and at least a couple percent larger than for the real Course in both cases. The trends for all three course plots appear similar for grid sizes smaller that 700m. In the open sea case (figure 4), for coarser grids there is a decreasing trend in Course, while the increase for the neighboring courses continues. The observed distance between the real course plot and the altered courses shows that there are fewer positive detections for the true 45º interval of vessels' courses, especially evident in the open sea case (figure 4), but also clearly visible in the harbor case (figure 5).

Spoken plainly, if the observed vessels were sailing at open sea with course set 45º more to the left or to the right, approximately 3 times more of them would be marked as anomalous (figure 4). This finding suggests that the course scale binning succeeds with differentiating the correct and faulty vessel courses. The goal is to maximize the distance between the true course and the false courses. In the open sea case, grids greater than 200m yield results with at least 10 percent difference in detection rate (figure 4). In the harbor case there is a difference of at least three percent for grids smaller than 90m (figure 5). These observations show similarity to the general open sea and harbor case comparison in figures 2 and 3. Once more there is a lower limit for adequate grid sizes at open sea, and an upper limit in the harbor.

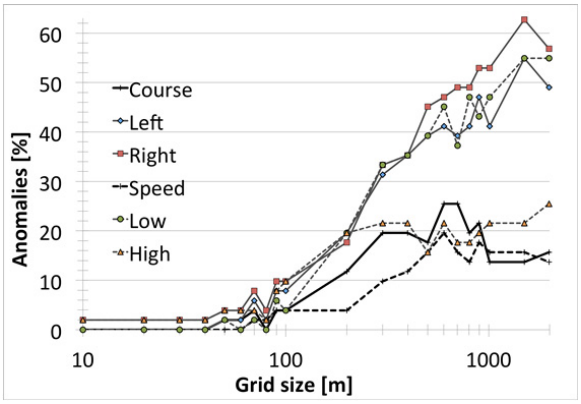

Fig. 4. Open sea: course and speed comparison with altered values.

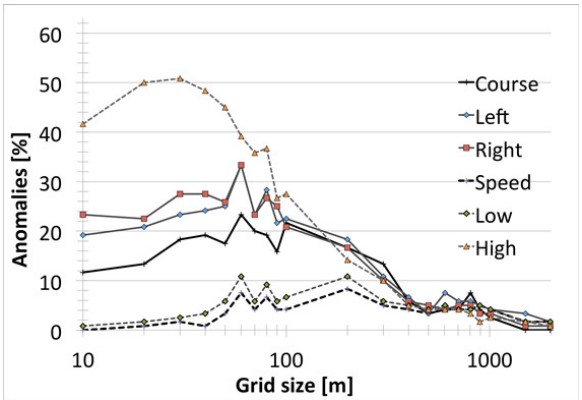

Fig. 5. Harbor: course and speed comparison with altered values.

Unlike the altered courses, the altered speeds (*Low* and *High*) do not yield pairwise similar results. At open sea (figure 4) the higher speed grows faster for lower grid sizes, but for coarser grids it is just slightly higher than the true speed, while the lower speed increases as steeply as the altered courses, completely deviating from the true speed trend. In the harbor case (figure 5) that relation is reversed. Namely, the lower speed is consistently only slightly greater than the true speed, while the higher speed plot shows much higher initial values and an intensive decrease. These findings infer that at open sea there is a higher tolerance for increased speed, and slowing down is unusual. On the other hand, in the harbor it is less anomalous to sail slower, than to speed up.

The comparison of the harbor and open sea course and speed results (figures 4 and 5) implies even stronger distinctions than in figures 2 and 3. While the true course and speed plots are kept unchanged, the altered courses

and speeds present quite contradictory trends: increase for open sea, and decrease for harbor. The results also further support the deduced optimal grid size ranges.

6. Discussion

One way of detecting anomalies is to follow deviations from the route of the vessel. By plotting transmitted AIS waypoints it is possible to represent previous positions and the number of vessels that visited the same locations. In the same way, by plotting courses and speed, it is possible to detect more volatile changes in the positions of the vessels. These parameters are heavily dependent on the actual grid size for the investigated area.

The two investigated cases include the open sea and the harbor. The latter is depicted by proximity to land, narrow passages, a higher density of vessel traffic, and regulated as well as irregular traffic patterns and speed limits. Vessels sailing at a slower speed generate subsequent waypoints at shorter distances between one another, but with possible larger changes in course and speed. The collected positions are therefore closer in a given area for the harbor case, suggesting a denser grid.

The analysis of the case study outcomes indicates that the number of suspected waypoint anomalies decreases when the grid size is enlarged, however, a bigger grid size may result in false negatives. The need to minimize the false detections without excluding any actual deviations emphasizes the importance of setting detection sensitivity. The optimal range for the grid size (as a detection parameter) was shown to vary for different study cases.

The findings for both cases appear to correspond with common sense. Open sea, where vessels sail at higher speeds, more stable courses and longer distances from each other, requires a larger grid size. The more densely traveled harbor waters, where many routes cross and ships sail closer to one another, require a denser grid.

The analysis of results leads to further insights regarding specific parameters. In harbor areas, speed is regularly limited by law, obstacles and (or) sailing practice, therefore exceeding it is more dangerous and less common than slowing down even more. On the contrary, at open sea it is common to sail with particular minimal speeds, and exceeding them is not disadvantageous. Hence, in that case slowing down is considered potentially more harmful and unusual than increasing speed.

Real anomalies leading to accidents or being indicators for terrorist acts or smuggling are rare, but identifying them in timely matter is vital for traffic safety. The potential filed based method helps by drawing user's attention to potentially problematic traffic behaviors, and eliminating the need to examine them all. The great advantage of the method is the white-box structure, allowing to observe the detections, as well as the basis on which they were made. The ability to display the violated traffic pattern in a visual, map-based form, along with the detection, makes the method uniquely suited for prompt incident analysis.

The introduced anomaly detection tool, STRAND, should be regarded as a visualization tool for human experts, providing another level of information compared to traditional equipment, like radar or GPS. It can be said that, through the use of the AIS vessel tracking data, STRAND incorporates seaman's experiences in estimating the changing traffic and operating environment conditions.

7. Conclusion

The method's novelty lies in employing the concept of a potential field for AIS vessel tracking data abstraction and maritime traffic representation, where the collection of potential fields, represents a model of normal behavior, and vessels not conforming to it are marked as anomalous. The geographical grid size is a factor impacting the sensitivity of the anomaly detection prototype system STRAND, implementing the potential fields.

The performed experiment examined how a changing grid size influences anomaly detection for two different Baltic areas, an open sea and a port area. A noticeable shift could be observed between the results. For both total and waypoint anomalies there was a decline with an increasing grid, with the harbor area declined faster than the open sea area. Course, speed and daytime parameters increase for lower grid sizes and decrease towards the end of plots, with the optimal amount of anomalies in range between 60 and 200 meters for the harbor case and 300 to 1000 meters in the open sea case. The STRAND prototype acts as a traffic operator supporting visualization tool, filtering out all normally occurring traffic and focusing the attention of a human expert on the proposed anomalies. In both presented cases, the system succeeds to exclude the majority of "uninteresting" traffic.

There are various potential benefits and practical applications of the method, depending on the user. From a ship navigator point of view, the presentation of patterns of correct or normal behavior, aids the choice of the safest and most optimal path. From a traffic safeguarding perspective, the anomaly detection based on potential fields may help quickly and comprehensively inspecting possible traffic incidents. Finally, from authorities' point of view, the clear overview of traffic may help to recognize traffic regulation and legislation issues.

References

Acciaro, M. & Serra, P. (2013). Maritime Supply Chain Security: A Critical Review. In: IFSPA 2013, Trade Supply Chain Activities and Transport: Contemporary Logistics and Maritime Issues, pp. 636.

Baldacci, A. (2008). Automated Identification System Anomaly Simulator, Memorandum Report, NURC-MR-2008-002.

Baldacci, A. & Carthel, C. (2009). Maritime Ttraffic Characterization with the Automated Identification System, Technical Report, NURC-FR-2009-008.

Bichou, K., Szyliowicz, J.S. & Zamparini, L. (2013). Maritime Transport Security: Issues, Challenges and National Policies. Edward Elgar Publishing Ltd., UK.

Brax, C., Karlsson, A., Andler, S.F., Johansson, R. & Niklasson, L. (2010). Evaluating precise and imprecise State-Based Anomaly detectors for maritime surveillance, In: Proceedings of the 13th Conference on Information Fusion (FUSION).

Brax, N., Andonoff, E., Gleizes, M.P. & Johnson H. (2013). In: Agent and Multi Agent Systems. Technologies and Applications.

Cariou, P. & Notteboom, T. (2011). Bunker costs in Container Liner Shipping: Are slow Steaming Practices Reflected in Maritime fuel Surcharges? In: Notteboom T. (Ed.), Current Issues in Shipping Ports and Logistics (pp. 69-82). Antwerp: Antwerp University Press.

Department of Homeland Security (2005). National plan to achieve maritime domain awareness. Technical report, USA.

Gehant, M., Roy, V., Marmorat, J.P. & Bordier, M. (2009). A Behavior Analysis Prototype for Application to Maritime Security. NATO Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness (NATO MSA 2009), NATO Undersea Research Centre, Italy.

Griffin, C. (2009). Learning and Prediction for Enhanced Readiness: An ONR Office 31 Program, Presentation to TTCP MAR AG-8.

International Maritime Organization (2002). Guidelines for the onboard operational use of shipborne automatic identification systems.

Kazemi, S., Abghari, S., Lavesson, N., Johnson, H. & Ryman, P. (2013). Open data for anomaly detection in maritime surveillance. In: Expert Systems with Applications, 40.14, (pp. 5719–5729).

Kessler, O. (2009). Anomalous Behaviors in Maritime Environments. In: Proceedings of the workshop on Detection Anomalous Behaviors in Maritime Environments, Carnegie Mellon University.

Laxhammar, R. & Falkman, G. (2013). Inductive Conformal Anomaly Detection for Sequential Detection of Anomalous Sub-Trajectories. In: Annals of Mathematics and Artificial Intelligence, Springer.

Laxhammar, R., Falkman, G. & Sviestins, E. (2009). Anomaly detection in sea traffic - A comparison of the Gaussian Mixture Model and the Kernel Density Estimator. In: Proceedings of the 12th International Conference on Information Fusion.

Morel, M., George, J.P., Jangal, F., Napoli, A., Giraud, M.A., Botalla, M. & Littaye, A. (2009) SCANMARIS Project – Detection of Abnormal Vessel Behaviors. In: Workshop on Data Fusion and Anomaly Detection for Maritime Situational Awareness, NATO Undersea Research Centre, Italy.

Osekowska, E., Axelsson, S. & Carlsson, B. (2013). Potential Fields in Maritime Anomaly Detection. In Proceedings of the 3rd International Conference on Models and Technologies for Intelligent Transportation Systems.

Perera, L.P. & Oliveira, P. (2012). Maritime Traffic Monitoring Based on Vessel Detection, Tracking, State Estimation, and Trajectory Prediction. In: IEEE Transactions on Intelligent Transportation Systems, 13.3, (pp. 1188–1200).

Rhodes, B.J., Bomberger, N.A., Seibert, M. & Waxman, A.M. (2006). SeeCoast: Automated Port Scene Understanding Facilitated by Normalcy Learning. In: Military Communications Conference.

Ristic, B., La Scala, B., Morelande, M. & Gordon, N. (2008). Statistical Analysis of Motion Patterns in AIS Data: Anomaly Detection and Motion Prediction. In: Proceedings of the 11th International Conference on Information Fusion.

Riveiro, M. & Falkman, G. (2009) Interactive Visualization of Normal Behavioral Models and Expert Rules for Maritime Anomaly Detection. In: Proceedings of the Sixth International Conference on Computer Graphics, Imaging and Visualization.

Riveiro, M., Falkman, G. & Ziemke, T. (2008). Visual Analytics for the Detection of Anomalous Maritime Behavior. In: Proceedings of the 12th International Conference Information Visualisation.

Seibert, M., Rhodes, B.J., Bomberger, N.A., Beane, P.O., Sroka, J,J., Kogel, W., Kreamer, W., Stauffert, C., Kirschner, L., Chalom, E., Bosse, M. & Tillson, R. (2006). SeeCoast port surveillance, In: Proceedings of SPIE Vol. 6204: Photonics for Port and Harbor Security II Orlando, FL, USA.

Tozicka, J., Rovatsos, M., Pechoucek, M. & Urban, S. (2008). MALEF: Framework for Distributed Machine Learning and Data Mining. In International Journal of Intelligent Information and Database Systems, vol. 2, pp. 6-24.

Van Laere, J. & Nilsson, M. (2009). Evaluation of a workshop to capture knowledge from subject matter experts in maritime surveillance. In: Proceedings of the 12th International Conference on Information Fusion.