UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering, Mathematics & Computer Science

Viability of anomaly detection aid for vessel traffic service operators

S. J. F. Hendriksen

M.Sc. Thesis
October 2022

Supervisors:
dr. A. Asadi
dr. F.A. Bukhsh

J. Daniëls
E. Flapper
from
Saab Technologies B.V.

Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

With the importance of maritime traffic on the global supply chains and the role of vessel traffic service operators in its resilience, improving the digital tools available appears an essential step for the future of maritime traffic management. This work examines the path toward viability for maritime anomaly detection as an aid for vessel traffic service operators, building from the literature towards a proof of concept and evaluating hurdles in the process. Leveraging data of moving vessels from radar and Automatic identification system (AIS) sources, the work discusses maritime anomaly grouping, detection strategy, transparency, and visualization as part of the proof of concept. The framework for maritime anomaly grouping supplies guidance for grouping anomalies and a non-comprehensive set of anomaly types. Anomaly types defined in the framework supply the foundation for the detection strategy used in the proof of concept. The selected maritime anomaly detection strategy, leveraging self-organizing maps and Gaussian mixture model normalized probability density, performs well by consistently highlighting low-density areas. Separate transparency models trained on subsets of the features score individual anomaly types. These transparency models enable the corresponding visualization to provide what anomaly types deviate a vessel from normalcy. The proof of concept and process are evaluated in one-on-one expert sessions, resulting in three aspects to mature the maritime anomaly detection field. These aspects are: defining the behavior constituting a maritime anomaly, focusing on transparency as part of the strategy, and selecting the relevant anomaly types for vessel traffic service operators. With the potential for the viability of maritime anomaly detection on the horizon, the steps suggested by this work will help mature the field toward real-world application.

Contents

List of Figures

List of Tables

1 Introduction

This section provides the motivation, scope, and structure of this work. The primary motivation is the importance of the resilience of the maritime sector and aiming to improve the role of vessel traffic service (VTS) operators within this interconnected system by employing an anomaly detection aid. The motivation is elaborated in Section 1.1, and the scope is discussed in Section 1.2. The impact of this work is briefly discussed in Section 1.3. Finally, Section 1.4 provides a brief overview of the structure for the rest of this work.

1.1 Motivation

The maritime transport sector accounts for 80% of global trade by volume and 70% of global trade by value (on Trade & Development, 2018). In 2020, 10 billion tons of cargo were loaded, with the prediction of further annual growth of 2.4% between 2022 and 2026 (on Trade & Development, 2021). The role of maritime transport in global trade highlights the impact of product prices on consumers and businesses. A combination of various downstream effects from the Coronavirus outbreak caused the container freight rate index to increase by 600% between May 2020 and September 2021 (Placek & Freightos, 2022). The container shipping congestion has highlighted the need for digitization of shipping and an increased focus on the supply chain's resilience. Furthermore, an increase in freight rate will jeopardize the profitability of sectors reliant on low-value commodities (Analytica, 2021).

Critical nodes in the trade routes consist of busy waterways that are singular points of failure due to the lack of available fall-back options that can support the quantity, size, and volume of traffic. The critical nodes in common trade routes such as straits, canals, and ports frequently cope with traffic congestion impacting the risk conditions and the effects of incidents in these waterways (van Meersman et al., 2012). The resilience of maritime trade routes depends on the functioning of all critical nodes within the system. A failure of one of these crucial nodes can cause the collapse of a trade route and take weeks to resolve and clear the backlog. For instance, in March of 2021, the Suez Canal became blocked for six days, and the trade loss was estimated at $400 million worth of trade per hour, resulting in a trade loss of $54 billion (Baker et al., 2021).

The Vessel Traffic Services (VTS) monitor busy waterways to protect the environment, improve safety, and improve the efficiency of maritime traffic (Council, 1996). VTS uses systems such as radar, Automatic Identification System (AIS), and direct communication to monitor the movement of vessels and provide extra navigational safety in monitored waters (Brodje et al., 2010). VTS operators manage traffic and provide information to vessels in the monitored area. These operators are responsible for maintaining situational awareness and providing guidance on movement within their assigned area. The operators rely on their intuition for these tasks. They determine the priority of specific vessels by the preparedness and impression of direct communication and behavioral patterns (Brodje et al., 2010). Many technological innovations developed benefit VTS operators. Innovations such as AIS, advancement in their display technology, radar technology (Lin & Huang, 2006), and path prediction (Macdonald, 2002). One of these has been recent attempts to detect anomalous behavior of vessels to provide warnings, alerts, or priority for attention using artificial intelligence strategies. The behavior seen as anomalous would be any behavior that deviates from how most vessels behave, highlighting how vessels deviate from normal behavior. The maritime anomaly detection field has primarily focused on the technical viability of anomaly detection systems to determine events that would require an operator's attention. While numerous systems have been proposed and validated to detect anomalies in some capacity, validation mostly consists of simple tests and cases.

With the importance of maritime traffic and the role of VTS in the resilience of the global supply chain, improving digital tools seems like an essential step for the future of VTS. This work aims to create a comprehensive proof of concept anomaly detection system focused on providing anomaly detection for VTS operators and determine its usefulness.

1.2 Problem definition

This work develops an anomaly detection proof of concept consisting of detecting various anomalies. The proof of concept helps evaluate the viability and effectiveness of anomaly detection as a VTS operator aid.

The proof of concept uses data sets from the port of X, consisting of combined AIS and Radar data for identified types of anomalies detected to train the models for normalcy. These models provide vessels with a normalcy score for each type of anomaly that is the inverse of the anomaly score. The output of these models is technically validated using examining results to evaluate

the effectiveness in detecting types of anomalies.
To enable validation for the aid of VTS operators, the proof of concept allows user interaction. Supporting validation utilizes sessions with experts in the field. From this problem definition, the following research questions arise:

RQ: What is the viability of anomaly detection aid for VTS operators?

Six sub-questions have been identified to answer the main research question as follows:

SQ1: What types of anomalous behavior exist for maritime traffic?

SQ2: How can we evaluate the effectiveness of maritime anomaly detection?

SQ3: What effective methods exist to detect maritime anomalies?

SQ4: What is the effectiveness of using self-organizing map and Gaussian mixture models to detect maritime anomalies?

SQ5: What level of transparency can be achieved using self-organizing maps and Gaussian mixture models to detect maritime anomalies?

SQ6: How can a maritime anomaly scoring tool aid VTS operators?

1.3 Work contribution

This work aims to connect the practical and the scientific by evaluating the state of the scientific field and relating it to what is lacking for practical benefit. It does this by working through the entire process from a practical approach, attempting to leverage available literature. Since the work looks at the whole process, it contributes to many aspects in areas of maritime anomaly detection. This work contributes to maritime anomaly grouping, the role of transparency, transparency detection strategies, and maritime anomaly visualization. These contributions, the literature, and expert validation further enable comprehensive recommendations for the requirements to achieve practical viability in the field.

1.4 Report structure

Other sections in this report consist of findings from the literature and the implication for study design in Chapter 2. The next section provides an overview of the data relevant to the proof of concept development in Chapter 4.1. Chapter 3 discusses the theoretical background and framework. Moving forward with practical implementation in Chapter 4 and the results in Chapter 5. The conclusion section returns to the research questions, discusses the results, and provides future research recommendations in Chapter 6.

2 Literature

This section discusses the literature on user interaction and maritime anomaly detection. The implications from the literature discussed in this section form the foundation for the proof of concept design. The user interaction section is split up into trust and interface, while the anomaly detection examines the broader fields and then specifically at maritime anomaly detection strategies.

2.1 User interaction

User interaction is the field that studies the interaction of humans with computers and subdivides it into technology acceptance, user satisfaction, trust, computer self-efficacy, personalization, and privacy (Rzepka & Berger, 2018). The viability of the proof of concept depends mainly on the user interaction with the system to enable benefit. The AI-enabled systems field can significantly benefit from incorporating user interaction specific to their purpose (Rzepka & Berger, 2018). Specific recommendations for the interaction of users and systems will depend on the specific field. The work from (Praetorius & Lützhöft, 2012) highlights the need to look beyond the technological ability to create systems for aiding VTS operators. Hence, the proof of concept incorporates various principles. These principles are trust in the system and the interfacing of VTS systems.

2.1.1 Trust

Trust can be present in many forms, generally between the person and their trustee. Trustees can, for example, be other humans, organizations, institutions, or even IT artifacts (Söllner & Leimeister, 2013). Benbasat and Wang (2005) further explains that people also form social and trusting relationships with technological artifacts and move beyond tool utilization.
In algorithmic decision-making and artificial intelligence tools to aid the user's experience and benefits, many considerations are to determine the perceived usefulness. Within this, papers such as Shin (2020) and Shin and Park (2019) show that trust correlates with the FATE (Fairness, Accountability, Transparency, and Explainability) or FAT (Fairness, Accountability, and Transparency) principles. Fairness in this context is having a careful approach to every process of an AI design so that the system will not produce discriminatory results for groups or individuals Shin (2020). Accountability

encompasses who and how is accountable for the actions of AI, as argued by Diakopoulos (2016). The field of artificial intelligence still debates the definition of transparency, but this work adopts the definition by Shin and Park (2019). This work defines transparency as the ability of users to understand how an AI system makes a decision. Lastly, explainability is similar and defined as how the input features of the model are associated with its decision in a manner understandable by humans (Rai, 2019).
This trust will correlate with more perceived usefulness, convenience, and satisfaction. To achieve the benefits desired of a data enhancement tool for users, these can be taken into account early in the process. The work by Shin and Park (2019) provides guidelines. Explaining the system's intention, the sources of data, and the relation between input and outputs form the guidelines for realizing the benefits. The anomaly detection strategy will use these guidelines and pick methods that promote transparency and the interaction of various features to the anomaly score.

2.1.2 Interface

The VTS operator's role requires active monitoring of many information sources to gauge the potential risks in their assigned area. Interface design plays the role of managing the information overflow and information availability to support user interaction. As early as 1988, Baldauf and Wiersma (1998) highlights the importance of the implementation method for alerts and warnings. An option to achieve this can be seen in a priority queue system for messages and alerts, as seen in Gonin et al. (2009). Occupied VTS operators also need the ability to return to earlier notifications, highlighting further importance for a queue system (Riveiro et al., 2008).

2.2 Anomaly detection

Anomaly detection is the field that specializes in the finding of patterns within data that do not correspond with expected behavior. An anomaly consists of finding a nonconforming pattern and providing value in many fields such as fraud detection, cyber-security intrusion detection, fault detection for systems, and surveillance (Chandola et al., 2009). Anomaly scoring and detection can employ many strategies, from histograms to complex machine-learning approaches. The employed strategies might differ based on assumptions regarding the normal and anomalous data (Chandola et al., 2009). In the following section, maritime anomaly detection and its groupings from literature are discussed and continued with the grouping of maritime anomalies and exploring the strategies employed in the literature.

2.2.1 Maritime anomaly detection

Anomaly detection within the maritime sector for detecting anomalous behavior of ships has been a field of study for almost 20 years. Highlighting its importance in improving vessel traffic safety, maritime security, and protecting the environment (Riveiro et al., 2018). Two distinct research areas in this field are visible in maritime anomaly detection. Specifically, the defense and civilian approach to maritime anomaly detection. The defense approach predominately focuses on threats, such as terrorism, smuggling, piracy, and territorial violations (Martineau & Roy, 2011). The civilian approach focuses more on vessel traffic safety, often monitored by VTS operators in busy waters. Anomalies within the maritime field are many different aspects of the behavior seen in vessel movements. An attempt at grouping types of anomalies can be found in Riveiro et al. (2018) as follows:

- **Positional anomalies:** Vessels in unusual areas, outside shipping lanes.
- **Contextual anomalies:** Factors such as weather, ship type, or time period do not correspond to expected behavior.
- **Kinematic anomalies:** The speed and direction or change of speed and direction do not align.
- **Complex anomalies:** Anomalies related to the intention of ships, such as loitering, drug smuggling, and AIS spoofing.
- **Data related anomalies:** Anomalies related to data incompleteness or behavior.

These categories outline the categorization of examples for anomalies, enabling the grouping of various papers into types of anomalies found in the following section regarding the analytical methods used for anomaly detection in literature. The grouping for anomalies allows a more straightforward comparison of the anomalies detected in the literature. Many anomaly detection strategies also combine portions of these grouping, and this work establishes a more specific framework. This process rescopes some anomaly

types to fit more specific anomalies relevant to civilian anomaly detection, focusing on contextual and complex anomalies. These changes are because contextual anomalies are seen as a sub-group of all anomaly types, providing an extra layer of detail to other anomalies. Factors such as the ship type, time of day, and weather influence whether or not a certain speed or location may be normal. The grouping found in the literature identifies the complex anomalies as the vessel's intention in this grouping. These are difficult to validate and more relevant for maritime surveillance from a defense standpoint rather than the civilian angle. These aspects mean this work will adopt altered groupings for anomalies and not use those outlined in Riveiro et al. (2018).
The definitions adopted for this work are as follows:

- **Positional anomalies:** Vessels that are in a location deviating from normal.
- **Kinematic anomalies:** Vessels with an unusual course, speed, or combination.
- **Data related anomalies:** Anomalies relating to data incompleteness or transmission behavior.
- **Routing anomalies:** Deviations from normal routing paths or sudden changes in speed or direction.

2.2.2 Anomaly detection strategies

This section explores literature that discusses various strategies for detecting maritime anomalies. A review summarises various aspects of the discussed papers. Summarising aspects found in Table 1 are the types of anomalies detected, the data used, the techniques used, and an evaluation of the results. The review groups the types of anomalies detected in the literature according to those in Section 2.2.1. The techniques consist of the primary methods used to create the model and score the anomalies. Finally, each paper's outcome and validation portions are summarized to indicate the approach's effectiveness in scoring anomalies. The review consists of papers from the keywords *maritime anomaly detection* and references found in respective papers.
The result of the literature summary is in Table 1. With the anomaly types indicated with the first three letters of the grouping provided in the previous

Table 1: Maritime anomalies detection strategies review

Paper	Anomaly types				Analytical tools	Effectiveness
	Pos	Kin	Dat	Rou		
Zhao and Shi (2019)	P	P	N	Y	Recurrent neural network	Successful in individual cases
Zhen et al. (2017)	Y	Y	N	Y	Naïve Bayes classifier	Successful in individual cases
Osekowska (2014)	Y	Y	N	N	two-dimensional Histogram	Useful frequency map
Handayani et al. (2013)	N	P	N	P	SVM, interpolation	Interpolation helps
Kowalska and Peel (2012)	N	N	P	Y	Gaussian Process	Success on larger displacement
Vespe et al. (2012)	Y	N	N	Y	Custom code for building track	Shipping lanes and turning points
Will et al. (2011)	Y	Y	P	N	Gaussian processes, KD-trees	Resource intensive
Lane et al. (2010)	N	N	P	P	Bayesian network	None determined
Roy (2010)	Y	Y	N	N	Rule-based reasoner	Dependent on rules
Riveiro et al. (2008)	Y	Y	Y	N	Self-organizing map and GMM	Effective representation

subsection. With a label of either Y, N, or P referencing yes, no, or partial respectively. In the groups of anomaly types that each paper has attempted to detect, it becomes evident that most of the strategies employed focus on a small subset of all detectable anomaly types. The literature frequently aims to detect a specific type or example of an anomaly. Notably, every paper further uses different analytical tools. Another concern is the lack of validation on the effectiveness of the papers that are part of the review since none of the papers have validation beyond individual cases. This lack of validation makes comparison challenging. Another aspect is that all papers use AIS source data, except the rule-based system seen in Roy (2010) that does not outline any specific data.

Other papers that can serve as a basis for comparison between anomaly detection methods are a set of review papers outlining the different strategies and their respective benefits and applicability. Riveiro et al. (2018) highlight the prevalence of statistical anomaly detection due to its complex and ill-defined nature. This anomaly detection strategy utilizes a normalcy model using various tools and tackling anomaly detection. Many different methods can create this statistical normalcy model. In an earlier paper looking at specific machine learning techniques, Obradovic et al. (2014) highlights a direct comparison of the benefits and detriments. This review looks specifically at neural networks, Bayesian networks, Gaussian processes, Gaussian mixture models, and support vector machines (SVM). Application of neural networks in maritime anomaly detection can be found in papers such as Zhao and Shi (2019) to help predict the future state of a vessel, allowing comparison to the actual state. SVM is a supervised approach, shown in Handayani et al. (2013) to impact the applicable scope significantly. The SVM approach primarily allows the detection of changes in vessels' behavior veering from normal shipping lanes. As seen in Lane et al. (2010), Bayesian networks promise the ability to incorporate expert knowledge and transparency but lack a maritime domain-specific attempt. Both Gaussian-based methods have the benefit that the implementations do not require codified expert knowledge, as these strategies are methods of determining the multivariate normal distribution of the data, using the probability density to score data points. When analyzing the attempts for maritime anomaly detection in the literature, issues regarding various strategies arise. The supervised strategies lack specificity or ability to generate training data stemming from the lack of anomalous labeled data points (Riveiro et al., 2018). Within other anomaly detection fields, Cansado and Soto (2008) does offer the ability to generate random data to label anomalous. Codifying these anomalies will enable the detection of anomalies as seen in Roy (2010). The need for expert knowledge of each anomaly limits the generalization. Most research focuses on unsupervised anomaly detection using various strategies with these limitations. With the transparency aspect in mind, this further reduces the promising approaches. The prevalence of AIS as a data source with anomaly detection is also evident. The reliance on self-reported systems does highlight another potential issue (Chandola et al., 2009). In comparison, the density-based clustering and recurrent neural network combination in Zhao and Shi (2019) does enable the segmentation of some distinct anomaly types. A combination of various methods would be required to detect all anomaly types, and this combination would need to include separate models for individual point and trajectory scoring. Individual point scoring detects many anomaly types in various groups but lacks track-based routing anomalies. Track-based anomaly detection strategies enable routing inconsistencies but cannot detect a wide range of positional and kinematic anomalies.

3 Methodology

This section covers the theoretical background of the frameworks, maritime anomaly detection strategies, visualization principles, and expert sessions employed in this work. This background serves as the foundation for the proof of concept. The implementation in Chapter 4 builds upon the theoretical foundation with practical choices, providing the results found in Chapter 5. The chapter is divided into five sections, first discussing the research framework applied in this work in Section 3.1. It then Continues by discussing the anomaly types detected in the existing literature and how those will be adapted in Section 3.2. Then, Section 3.3 motivates the maritime anomaly detection strategy choice and provides a thorough explanation of its components and formulas. Visualization, in Section 3.4, covers the principles incorporated in the visualization that is part of the proof of concept. The section that discusses the expert sessions covering the resulting complete proof of concept is covered last in Section 3.5.

3.1 Research framework

The research framework adopted for this work is the Cross Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al., 2000). CRISP-DM guides data science processes using six distinct steps iteratively, as seen below. A suggestion for the interaction of these steps is also made in the reference architecture, as seen in Figure 1. CRISP-DM was selected first for this work as it is the most frequently used and can be considered the standard (Mariscal et al., 2010). Furthermore, it also provides extensive documentation and practical overviews of the various intuitive steps to help limit missteps in the process.

1 - Business understanding: Analyse the business needs and objectives, determining important factors early in an early stage and ensuring the project answers the right questions.

2 - Data understanding: Examining the data to connect the available data and the project's needs.

3 - Data preparation: Process the data to produce to ensure all the correct data is in one place in the right format.

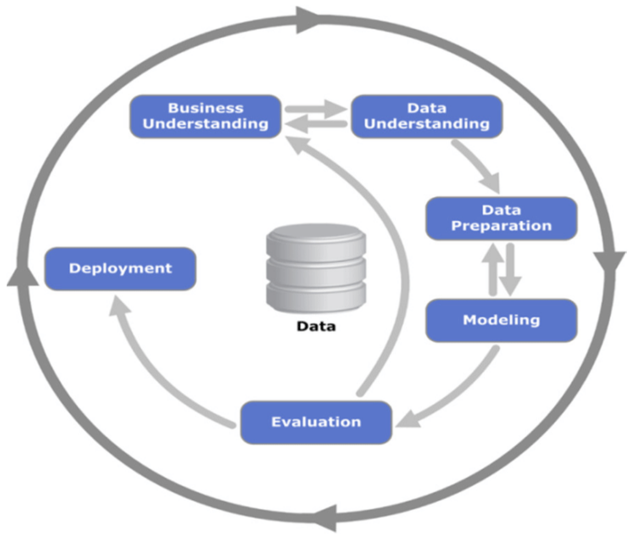


Figure 1: Steps of the CRISP-DM Methodology (Chapman et al., 2000), figure from: Kizilkan et al. (2015)

4 - Modeling: Determine, test, and create the model for the project.

5 - Evaluation: Evaluate the model's results and review them against the intended goals.

6 - Deployment: Determine the format of deployment of the results, develop a post-deployment strategy and deploy the results.

The six steps from the CRISP-DM framework clarify important steps toward a successful project. For this work, the CRISP-DM framework serves as a set of separate CRISP-DM cycles, focusing on the frequent evaluation of the alignment of the prototype with reality. These separate CRISP-DM cycles align with the various aspects of the overall prototypes: Anomaly types, anomaly detection strategy, and visualization. Lastly, the viability study provides a final evaluation of the three cycles combined. Step one, business understanding, is discussed predominately in Chapter 2 with some broader scope in Chapter 1. The data understanding and data preparation steps are discussed together in Chapter 4.1. Modeling selection is split into planning in Chapter 3 and specific tools selected in Chapter 4. Step five, evaluation can be found in Chapter 5. Step six, deployment is not reached as part of this work. These steps are discussed separately for determining anomaly types, anomaly detection strategy, visualization, and expert sessions. The expert sessions serve as the final evaluation step after the smaller cycles.

3.2 Role of maritime anomaly framework

A clear grouping and definition of each type of anomaly supports the selection for the proof of concept and simplifies feature selection. For this, the groupings discussed in Section 2.2.1 serve as the foundation for the framework for grouping and anomaly types used in this work.

3.3 Maritime anomaly detection tools

As seen in the background section, many different maritime anomaly detection methods exist in the literature. These methods aim to tackle the problem with various tools and strategies with differing degrees of success. The strategy selection combines various aspects in the literature and data availability to select the strategy for the proof of concept. The primary aspects affecting the choice in anomaly detection strategy are the lack of labeled or concrete manners of labeling data, ease of adding various context features, the possibility to achieve some level of transparency for the output, and promising existing literature. Due to the lack of labeled data and difficulty labeling data, the strategy needs to leverage an unsupervised approach. The specific groupings and types identified for maritime anomalies require modularity, feature selection for each identified anomaly type, and a method of scoring the combination of all features. The features picked should coincide with user-understandable concepts, in line with the transparency principles (Shin & Park, 2019).

Two primary contenders arise with these aspects: the combination of Self-organising maps (SOM) and a Gaussian Mixture Model (GMM) (Riveiro et al., 2008) and the recurrent neural network approach (Zhao & Shi, 2019) that showed distinction in the type of detected anomalies. With this, it became clear that the combination of SOM and a GMM allow more flexibility in feature selection as part of maritime anomaly detection. The SOM and GMM approach allows a ground-up approach by defining various types of anomalies and connecting these with associated features, thus enabling the selection of a reasonable subset of the anomaly types identified.

The SOM and GMM strategy trains a model for future data points by creating a probability density function from the training data set. In this work, the training data set consists of data points, where each point denotes the state of a vessel at a specific time. The strategy first clusters similar data points, training a SOM to fit all data points. This trained SOM consists of neurons representing a group of similar data points in the training set, considered a cluster. The distribution and means of data within each cluster are then extracted and used in the GMM, weighed according to their prevalence, to create a single distribution for the entire problem space. This single distribution can then score new data points by determining the probability density for its features. The remainder of the maritime anomaly strategy section elaborates on the individual components and formulas that achieve the steps to create the probability density function from the train data set and further elaborates on the validation of this strategy.

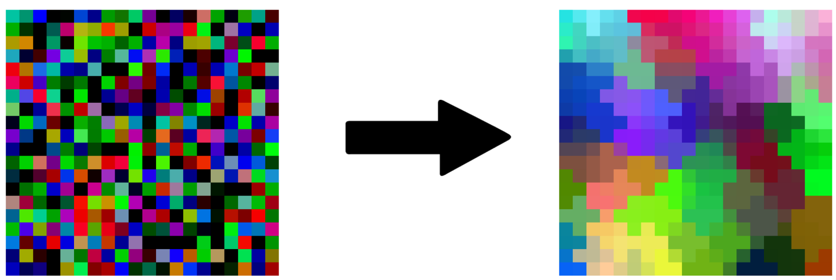


Figure 2: Training a SOM to cluster random RGB colors

SOM is an unsupervised, competitive learning neural network to create artificial groupings within data. The resulting grid of neurons represents clusters in the data after training, with the number of neurons denoting the number of clusters. The SOM initializes random values to group similar data points (See left in Figure 2). Through an iterative process, the SOM measures the distance to a training data point and maps it to the closest neuron on the map. This closest neuron is the best matching unit (BMU) and is determined using the Euclidean distance. After determining which neuron is the BMU, the weight of the BMU neuron adjusts towards the data point. The weight of neighbors close to the BMU neuron further changes toward the new data point. Changing the weights of neurons reduces further in the SOM training iterations. This simple step repeats on every data point for many cycles to have the SOM represent groupings in the data (See right in Figure 2). With this, the SOM can cluster similar data points and reduce the dimensionality of similar values without context (Yin, 2007).

The BMU equation, found in Equation (1), determines the BMU neuron given a distance function and SOM for any data point. Provided a data point x, given the SOM s and distance function d. I and J are the width and height of SOM s, using i and j to indicate specific neurons in the SOM. The BMU neuron (s_{ij}) matches with a data point if it is the closest data point according to distance function d.

$$BMU(x|s,d) = (d(s_{ij},x) = \min(\sum_{i=1}^{I}\sum_{j=1}^{J} d(s_{ij},x))) \rightarrow s_{ij} \tag{1}$$

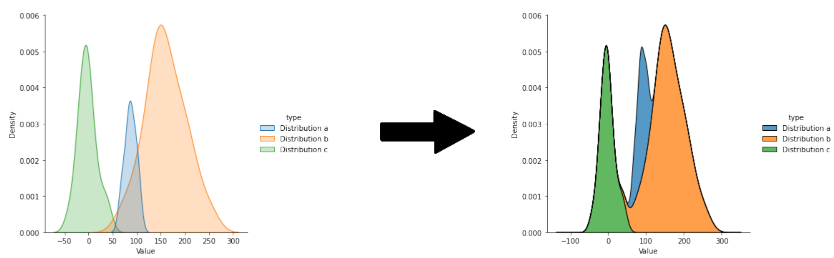


Figure 3: A Gaussian Mixture Model combining three distributions into a single distribution

A mixture model is a type of model that represents the subgroups within the overall distribution. The GMM is a mixture model that assumes that the distribution is a combination of normal distributions. Lastly, the added dimensionality from multiple dimensions represents the need for the multivariate version of the GMM. Effectively, the multivariate GMM is a tool to create a probability density function from a set of grouped data points. Each neuron in the SOM has its multivariate Gaussian distribution, which combines distributions according to their importance in the training set to create a single probabilistic model for the entire model space. The importance of a neuron is the frequency of the cluster, calculated by the percentage of training data points that have the neuron as BMU.
The multivariate Gaussian probability density function, seen in Equation (2), determines the probability density for each distribution in the GMM. The distribution's mean μ and covariance matrix Σ determine the probability density for the data points x. n represents the number of dimensions

of the data points and multivariate distribution. The mixing proportion π represents the contribution of the individual distribution to the mixture.

$$\mathcal{N}(x|\mu, \Sigma, \pi) = \frac{1}{\sqrt{(2\pi)^n}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \quad (2)$$

Equation (3) provides the formula determining p_{ij}, which calculates the probability that a specific neuron and distribution represent a data point. The probability, also called the mixing proportion in GMM, represents whether a specific neuron is the BMU for a data point in the training set x.

$$p_{ij} = P(BMU(x|s) = s_{ij}|x) \quad (3)$$

The combination of SOM and GMM uses the adjusted multivariate Gaussian probability density function, seen in Equation (4), to represent each specific neuron's contribution toward the combined distribution. It replaces the mixing proportion, previously π, with p_{ij} as seen in Riveiro et al. (2008) and shown in Equation (3). This definition of the mixing proportion ensures the height of the probability density is adjusted by the corresponding contribution of f_{ij} towards the combined distribution as indicated by p_{ij}. Further, μ indicates the average of points with s_{ij} as BMU, representing the resulting weights of s_{ij} in the trained SOM. Σ is the covariance matrix determined by the points matched to s_{ij}. Lastly, n is the number of features in each data point. These aspects use a probability density made for each neuron s_{ij} that given data points $(x_1, ..., x_n)$ will determine each data point's probability density.

$$f_{ij}((x_1, ..., x_n)|\mu, \Sigma, p_{ij}) = \frac{1}{\sqrt{(2p_{ij})^n}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \quad (4)$$

To combine the probability density functions f_{ij} that represent each respective s_{ij}. The probability density functions combine according to the probability p_{ij} as seen in Equation (5), here I and J reference the height and width of the corresponding SOM.

$$f(x) = \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} f_{ij}(x) \tag{5}$$

The model built with the combination of SOM and GMM describes the normalcy of behavior. This model creates a probability density function that combines d multivariate Gaussian distributions, where d is the number of neurons in the SOM. Thus, a low probability density score for a data point indicates it is far from assessed clusters within the training dataset. Subsequently, a low normalcy score is determined anomalous.

Leveraging SOM for clustering and the multivariate GMM for combining distributions, the resulting probability density function represents normal vessel behavior. Separate SOM and GMM models are used for each subset to determine the contribution of that subset of features. These models achieve transparency by indicating what anomaly types, connected to the subset of features, contribute to the anomaly score.

Finally, the result is normalized using a binning system to move from a probability density function to a more easily understandable value. With this normalization, the resulting value will be between zero and one, where one represents the highest seen probability density and zero is the lowest. Following this normalization, an overall score of one indicates a very normal point, and a score near zero indicates an anomalous point.

The model's behavior will be examined to determine if low-scoring data points are outliers, validating the SOM and GMM approach. Furthermore, examining the correlation between the transparency scores and overall scores determines the interaction between various anomaly types and the overall score. Lastly, expert sessions improve validation by verifying interpretations of these results.

3.4 Visualization principles

The visualization compliments the proof of concept by providing an interface to the output of the scores from the SOM and GMM anomaly detection approach. This interface allows a more accessible examination of the output from the maritime anomaly detection strategy and aids expert validation. The visualization takes the form of a simulated real-time minimal viable product combining the core interaction requirements and maritime anomaly detection specific features. The literature review highlights the need for an interface to explain anomalous scores and a queue to indicate vessel priority. These principles will be the basis for creating an interface design based on the created requirements that will be part of the design. An iterative process with experts highlights the need for various features to provide a minimum viable product to interact with the system.

3.5 Expert validation purpose

The viability serves as an expert opinion validation series of the maritime anomaly detection prototype and visualization results. It aims to validate the results beyond the ability of the researcher. For this, there are four components: discussing the process, the overall scores, the transparency strategy, and the future of the field. For all of these, the visualization serves as the interface. The process is a brief presentation regarding the anomaly types and proof of concept choices. A set of low-scoring example vessels is the example of the overall model, with a specific point for each vessel examined for transparency using a reference transparency strategy. The final component discusses the still missing aspects and future of maritime anomaly detection.

4 Implementation

This chapter elaborates on the practical choices and implementation building built on the choices in Chapter 3, creating the models for the results in Chapter 5. First, this chapter discusses the data used for the proof of concept in Section 4.1. Then, the anomaly framework created for this work is elaborated on in Section 4.2. Thirdly, in Section 4.3 it discusses the choices for the maritime anomaly detection proof of concept. Section 4.4 discusses the design of the visualization proof of concept. Finally, this chapter discusses the format for the expert sessions in Section 4.5.

4.1 Data

The data used for the proof of concept is from the port of X. For this area, there are two primary sources of data: AIS transmission and a combined source of radar and AIS data that VTS uses. This combined radar and AIS data are the most accurate for VTS usage (Lin & Huang, 2006). The data set thus uses the merged data, supplementing with AIS data to fill in the fields that are not present. The methods of vessel detection and processing are left out due to confidentiality. The data used for this work is of the highest quality available at the time. The data consists of points that update the state of a vessel at a specific time. These data points can be connected to a specific vessel using its identifier, which is a random string that represents that vessel. The time delta between updates for vessels is determined by connecting all data points and checking for the difference in time between the last data point. This time delta further serves as the indicator to determine the change in speed and course per second since the last data point.
This data set is further filtered to include only vessels present in the area of interest, moving vessels (>0.1 m/s), valid identifier, and an expected update interval for the system. With this, the data set consists of all moving vessels that transmit AIS within the area. The information inside a data point is in Table 2, including the name, type, unit, usage, and an elaboration on the contribution of that column.

4.2 Defining a maritime anomaly framework

A distinction for various specific types within the groupings discussed in Chapter 2 supports the modular approach for maritime anomaly detection.

Table 2: Data sample

Column	Type	Units	Use	Elaboration
Identifier	String	-	Identifier	Vessel identifier
Time	Datetime	-	Various	The timestamp for the vessel update
Breadth	Float	meters	Visualization	-
Length	Float	meters	Visualization	-
Longitude	Float	knots	Detection	Current position
Latitude	Float	knots	Detection	Current position
Time delta	Float	seconds	Intermediate value	Time since last update
Speed	Float	knots	Detection	Current speed of vessel
Speed delta	Float	knots/s	Detection	Difference in speed since last update
Course	Float	degrees	Detection	Movement direction of vessel
Course delta	Float	degrees/s	Detection	Difference in course since last update
Orientation	Float	degrees	Visualization	Facing direction of vessel
AIS delta	Float	seconds	-	Time difference between 2 most recent updates
Draught	Float	meters	-	Depth of vessel below waterline

This section discusses the starting point for the grouping of maritime anomaly types. Respectively these groups are the data-related, kinematic, positional, and routing anomalies. The earlier chapter provides descriptions for these groupings, but the anomaly types that fit in those groups further elaborate on the grouping. This context could consist of information on the type of vessel, weather patterns, specifics of the area, and other potential avenues to enrich the information on whether or not specific behavior are to be considered anomalous. This work discusses some avenues to include context, but the final implementation does not contain any context.

Here are the descriptions and types that correspond to each anomaly group elaborating on the earlier descriptions and starting with Data-related anomalies. The vessel's behavior does not cause data-related anomalies. Instead, data-related anomalies relate to the transmission or detection of vessels. The consistency of data information refers to anomalies detected using misaligned information regarding a vessel. Meanwhile, the consistency of transmission refers to the consistency of AIS transmission of specific vessels, which may be connected to the tampering or hindering of AIS transmission devices (d'Afflisio et al., 2021; Katsilieris et al., 2013). Kinematic anomalies relate to a vessel's current speed or direction (Note that kinematic anomalies do not include changes in speed or direction). The positional anomalies are vessels present in an unexpected area, given the context. The positional anomalies consist of the following types: unexpected arrival, unauthorized access, low-depth waters, and unexpected area. Mentions of unexpected arrival can be seen in Lane et al. (2010) and consist of vessels that present themselves at an unexpected time or place. Further mentions can be found in d'Afflisio et al. (2021), consisting of a vessel entering or heading towards an area that the vessel does not have permission to access. Low-depth water anomalies consist of vessels close to or heading towards areas that lack sufficient water depth compared to their draught. Lastly, the unexpected area is the anomaly that encompasses the vessel's location. Finally, routing anomalies consist of two different groupings of anomalies. The first grouping is anomalies corresponding with changes and the deviation from tracks seen in the literature (Lane et al., 2010). The second grouping consists of change anomalies outside the vessel's ordinary behavior changes. These recognized types are the changes in direction and speed. Deviations from tracks consist of vessel behavior that does not align with the expected navigation from one point to another. Proximity anomalies and leaving expected track anomalies are part of the deviation from track anomalies. The proximity anomalies refer to close passes with either land, shallow water, or other vessels. In comparison, expected track anomalies are vessels that do not use a logically used path from one point to another. To define logically used paths, Vespe et al. (2012) offers the option to build a database of typical tracks in an area.

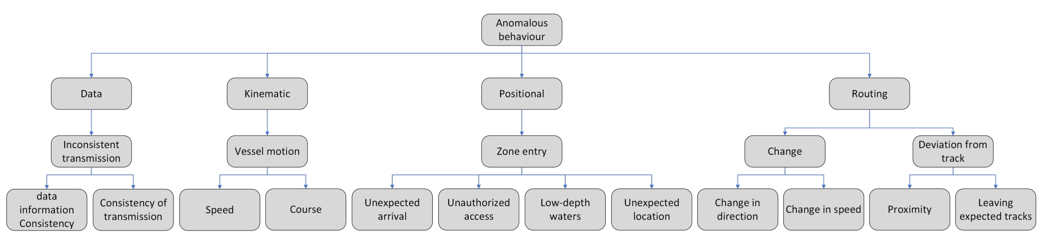


Figure 4: The types of maritime anomalies used for the proof of concept

Figure 4 shows all the groupings and types of maritime anomalies. Data-related, kinematic, and positional anomalies only have a single sub-group of anomalies. Nevertheless, other types of anomalies can expand the scope of these groupings. Note that this grouping is non-comprehensive but does elaborate on various specific types of anomalies that will be detectable from data commonly used to detect anomalies. The proof of concept uses five different anomaly types from the types defined here. These are the speed, course, location, speed change, and course anomalies. The types included in the proof of concept correspond to at least a portion of three groupings.

4.3 Transparent maritime anomaly detection

This section discusses the maritime anomaly detection implementation for the proof of concept. The implementation builds upon the strategy selection and theoretical foundation, discussed in Section 3.3, towards a functioning anomaly scoring proof of concept. The implemented maritime anomaly detection tool creates a probability density-based model for anomaly scoring and corresponding transparency scores. Here the overall model provides the general score for each data point, seen as the anomaly or normalcy score. The transparency models provide additional scores for each implemented anomaly type, explaining which ones contribute to the overall data point score. The proof of concept trains a SOM based on an extensive training data set, creating clusters used in the GMM to build the distribution for the entire problem space. As a final step, a binning system normalizes the probability density scores to values between 0.00 and 1.00. In developing the maritime anomaly detection proof of concept, the primary steps consist of i) system design ii) feature selection iii) tool selection iv) conditions to achieve stability. The implementation portion of maritime anomaly detection will follow these steps to clarify the exact steps taken to create a proof of concept for a SOM and GMM maritime anomaly detection model using the data explained in Chapter 4.1.

The system can handle a classic train and test split, which simulates training on historical data and testing on real-time data. Enabling the training of the SOM, building the GMM model, and the bins for the model created ahead of scoring. Each of these steps is resource intensive but can support state checkpoints in a small file format, no longer requiring the training data set. The system design can be found in Figure 5, starting at the top left with the data set outlined in Chapter 4.1. This data set splits between the training data of all data points fitting the filter criteria from the month of training data and the simulated real-time input consisting of the next two days. The train and test data represent a sizeable recent time frame, where the training is most of the data set to provide sufficient data for the model. Meanwhile, the test set is large enough to have sufficient variety in the time of day to examine the model behavior. The training uses the training data, with the SOM for all features (SOM_{all}) and the models for subsets of overall features representing the transparency models (e.g., SOM_{speed}, SOM_{course}). The trained SOMs, in combination with the training data set, build the GMM. Each GMM model (GMM_{all}, GMM_{speed}, and further transparency models) requires the storage

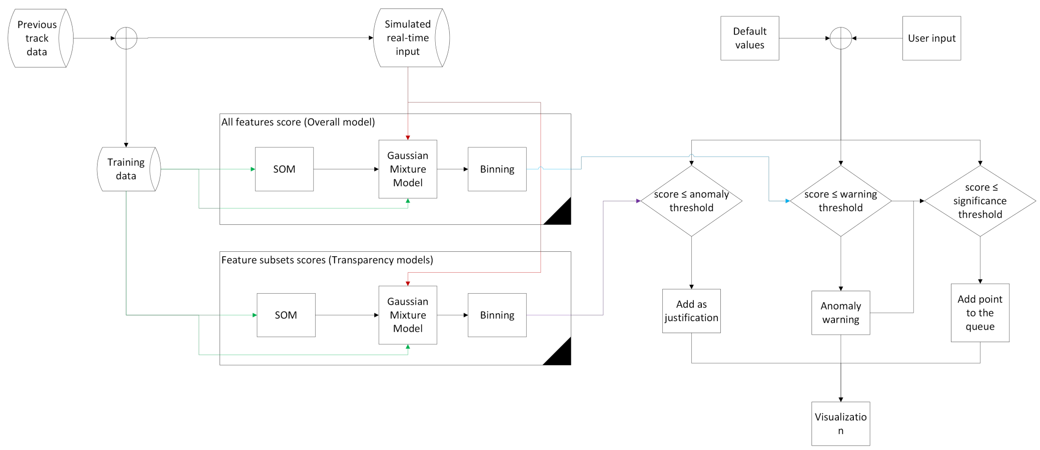


Figure 5: Architecture for the system

of the characteristics of all distributions that are a part of the GMM. These characteristics consist of the frequency of the distribution, the mean, and the covariance matrix. These are the given parameters to provide the probability density function described in Equation 4. The GMM_{all} model calculates the overall probability density of the test data points. In combination with the transparency scores from the transparency models. Lastly, the binning normalization normalizes the probability density to a value between 0.00 and 1.00. The bins correspond to 0.1% intervals of the probability density. Ensuring the lowest 0.1% of data points will correspond to the value 0.000, with the next 0.1% corresponding to 0.001. Deciding these bins continues for each of the 0.1% intervals up to 1. The normalized probability density values represent the scores for the data point, having the overall anomaly score and the transparency score for each. The visualization shows these normalized values, and the set thresh holds.

The SOM and GMM combination fits as a base layer, allowing expandability with more features and types. The final implementation only used the green types, as seen in Figure 6. The selected types consist of speed anomalies, course anomalies, location anomalies, change in direction anomalies, and change in speed anomalies. The final implementation left out the orange types, transmission consistency, and low-depth waters. Similarly, the implementation left out the size and type of ships. However, all of these lack the continuous data requirement creating over-fitting characteristics skewing the overall model drastically.

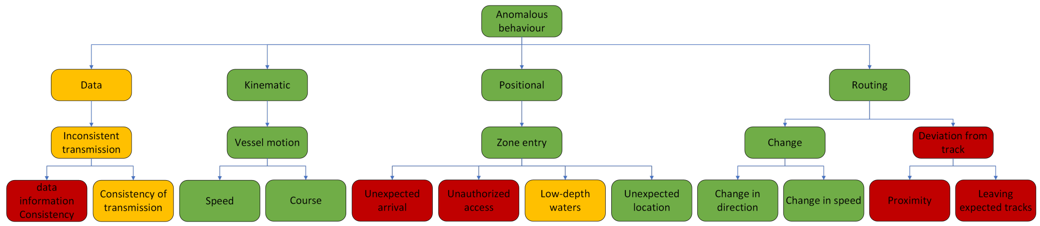


Figure 6: Selected types of anomalies for the prototype

The maritime anomaly detection prototype was developed in Python 3.10 (Rossum & Drake, 2009) to enable easy access to libraries and data science tools. The miniSOM (Vettigli, 2018) package is a robust and straightforward SOM implementation. The miniSOM package needs a small change to allow it to provide time estimates for larger training data sets without crashing on completion. This change avoids creating a $q \times q$ matrix, where q is the train data set size when training using a verbose setting. All SOMs trained using the same parameters, using a size of 50 by 50 neurons, an initial distance between neurons of 1.00, a learning rate of 0.5, and one million iterations. The training of six different SOMS resulted in one for each type selected and one for the overall model. The 50 by 50 size was chosen as a size limiting stability filtering of distributions without making the size overly large. The learning rate was chosen at 0.5 to represent high plasticity in the model (Yeo et al., 2005). This learning rate ensures that the SOM finds sufficient new clusters in the data set. The initial distance between neurons represents the initial distance considered for updating near a winning neuron. This distance is set to 1.00 to ensure clusters initially avoid interacting broadly, with the distance between clusters decreasing to ensure cluster interaction later in training (Mersiovsky et al., 2018). A large number of iterations improved the improvement beyond 100000 iterations while maintaining less than ten hours for each overall or transparency model in a single-threaded manner on a powerful single machine server. Optimization of learning rate and sigma is not part of the scope due to the limited ability to compare similar results with the validation strategies available. Training a single SOM and then extracting feature subsets was also attempted. This extraction proved unsuccessful due to disconnection between clustering and extraction of subsets.

The GMM implementation was self-implemented in two steps after examining available packages that did not offer all features required. First, the trained SOM is used to determine the win map for the training data. This win map determines the means, covariance, and frequency of each neuron in the SOM. Determining the win map consists of each neuron collecting all data points with that neuron as its Best-matching unit (BMU). For which the mean will correspond to the respective neuron weight in the SOM. The covariance is determined using all data points connected to each neuron. Finally, the percentage of data points connected to that neuron compared to the total amount of data points determines the frequency. The next step selects distributions that are a part of the distribution, excluding distributions with too much sensitivity. Firstly, this step removes all neurons with a single corresponding data point since these do not have a respective covariance. Then the distributions that lack significance due to too much sensitivity resulting in infinity probability densities are checked to that their covariance matrix has a sufficiently low condition number. Calculating the condition number checks that the distribution has enough significance to not result in issues by providing infinity values. In Equation 6 the formula for a condition value is provided. The condition number of a matrix is a measure of the sensitivity of a matrix to change. The formula calculates the condition number by multiplying the matrix's absolute determinant with the inverse matrix's absolute determinant. Literature by Strang (1988) provides a rule of thumb for the condition number, it creates a value k for a matrix A, which is in Equation 7, it creates the rule of thumb value k given using the condition value for matrix A. For this, the matrix loses k decimal places in round-off in Gaussian elimination, where k was set to 10 for all distributions, maximizing included distributions while minimizing issues. A low k filters too many distributions, whereas a very high (> 20) for k would let in distributions that would result in infinity probability density values for some test data points. Generally, less than 0.01% of data points were excluded from the model's large training samples using the selected k.

$$cond(A) = ||A|| \cdot ||A^{-}1|| \tag{6}$$

$$k(A) = log(cond(A)) \tag{7}$$

The characteristics from all distributions are combined with the self-implemented function as seen in Equation 4 and Equation 5. This function is provided a list of data points and will return a corresponding list of probability density values accordingly. The function achieves by treating the entire scoring as a vector operation. After the GMM_{all} model determines the probability density and the transparency models determine the transparency scores, those scores determine the bins. This model thus provides a score between 0.000 and 1.000 for the overall score and each anomaly type. The visualization then shows the scores given to all the data points corresponding to the vessels.

A different transparency strategy, using depth-first re-scoring of values, enables validation of the transparency models. This separate transparency strategy is a point of reference for comparing the transparency models to an alternative strategy. This separate transparency strategy employs a much more resource-intensive but more straightforward approach. It starts with the initial point and examines the highest possible scores by changing individual features to their optimal positions in an iterative cycle. The different strategy aims to find what can be considered a normal data point at the location of the inspected data point. The strategy validates specific data points for their transparency by providing a reference for the transparency models. The different strategy starts with all the original features of the data. Then the different strategy changes one feature at a time for a wide range of values. The strategy stores the highest normalcy score and feature value for each feature. The feature with the highest normalcy score is then selected, assuming the score is improved and has not reached the normalcy threshold (0.1). This normalcy was chosen based on the assumption that a minimum of 90% of data points are normal, aligning with a visual inspection of data points scored between 0.1 and 0.15. Suppose none of the features reach the normalcy threshold. In that case, the data point examined is changed for the selected feature to the value corresponding to the highest normalcy score. The strategy examines all features again, with the best feature changed, to look for the highest normalcy score in the next iteration. A graph also complements this strategy, enabling quicker comparison. In this graph, nodes correspond to the changed feature and the depth. At the same time, the vertices in the graphs between the nodes provide information about the change.

Figure 7 provides an example of the graph corresponding to the different transparency strategy. The results of the depth-first re-scoring are shown by having the nodes correspond to the shortened anomaly type names with a number for the iteration. The vertices provide the results for the highest normalcy score for that feature. In the graph seen in Figure 7, the red node

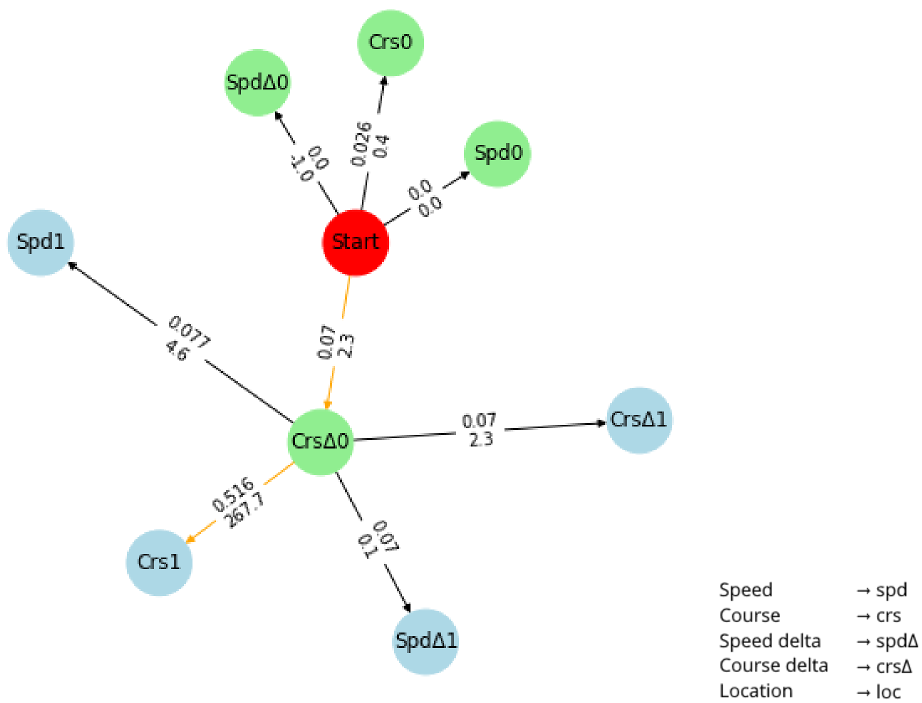


Figure 7: A depth-first re-scoring approach towards normalcy

labeled *Start* corresponds to the original values of the data point examined. Where the vertex to $course\Delta 0$ shows that the highest normalcy score obtainable with a single feature changed is 0.07 if the $course\delta$ would have a value of 2.3 m/s^2. The orange arrows also indicate the direction of the path towards normalcy for this data point. Comparison between this graph and the anomaly scores of a data point allows validation for the transparency level achieved using the transparency models of the proof of concept.

4.4 Visualization design

The visualization described in this section compliments the maritime anomaly detection strategy by providing an interface with the anomaly and transparency scoring models to the proof of concept. It visualizes the environment of moving vessels, where each data point has an anomaly score and corresponding transparency scores. A set of requirements support a minimum viable product. The requirements divide into two main categories:

core interaction and features specific to anomaly detection. The relevant requirements can be found in Table 3 and Table 4 and their importance for a minimum viable product. The tool selection aims to satisfy all requirements with minimal effort, enabling interaction with the model and completing the proof of concept.

Table 3: Core requirements visualization

Core interaction	
Requirement	**Value**
Map panning	Move around the map to inspect specific areas
Map zoom	Zoom in to inspect smaller and larger areas
Automatic updates	Automatic timeline to simulate real-time
Manual time change	Ability to alter time manually for inspection
Track vessels	Ability to track vessels between different timestamps
Select vessels	Live information of selected vessel

Table 4: Anomaly requirements visualization

Anomaly detection	
Requirement	**Value**
Queue	Ability to track vessels between different timestamps
Transparency	visualization for contributing factors to score
History	Graphs over time to see the previous state of vessel

The requirements allow for a good enough experience to allow the inspection of all relevant aspects in the visualization itself. Primarily, smooth movement in combination with live updating timing required a high degree of flexibility in the tools. The visualization uses a pure JS approach, combined with Mapbox JS (Mapbox, 13). These tools allow for maximum flexibility and a web page as an interface. The web page hosts anywhere, requiring the serving of the JSON data separately or locally as a single HTML page that would load the JSON data from a file.

An initial design to support all the requirements, leveraging the abilities of the selected tools, is found in Appendix A.1. Further iterations improved upon this design by taking feedback on the preferred information and avoiding specific design choices. During this, the primary changes were the ability to have many graphs open simultaneously, the choice from many colors to simple colors, and moving from a pie chart to a bar graph, clearly indicating the respective score of each anomaly type rather than the relativistic scores.

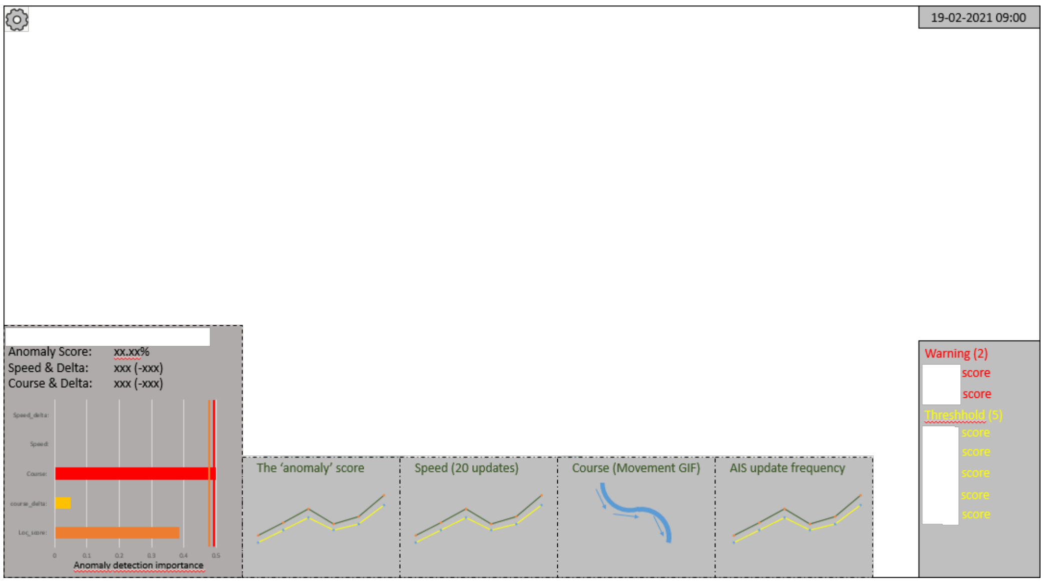


Figure 8: The final visualization design

The final design, seen in Figure 8, supports all requirements. Most core interactions do not have a representation in the final design, focusing on features specific to anomaly detection. Selected vessels are colored blue, showing the information in the bottom left, and the queue is on the bottom right as part of the anomaly detection features. The coloring of specific values in the vessel information screen and the bar graph is a part of the transparency strategy. The additional graphs represent the history of the selected vessel, showing the past vessel behavior and anomaly score. These graphs represent the overall score, speed, course, and location over the previous minute.

4.5 Expert validation outline

The expert sessions serve as a validation step on the results from the framework and proof of concept components. This validation consisted of one-on-one sessions with an expert in VTS systems, maritime artificial intelligence, and VTS standards. The sessions discuss the process and tools used to achieve the results. The sessions then moved forward to presenting and discussing results in a guided manner to limit the amount of system-specific knowledge needed. The primary discussion points are the anomaly type framework, maritime anomaly detection strategy, transparency strategy, and visualization. These discussion points aim to provide better validation and recommendations using the feedback provided by the experts.

5 Results

This section discusses the results that arose from the theoretical foundation and practical implementation in Chapter 3 and Chapter 4. These results divide into four aspects: The types of anomalies identified, the technical prototype, the visualization prototype, and the expert validation results. The chapter discusses the framework used to separate types of anomalies first. Second, the results from the test set of the trained model built upon the framework follow. Then the visualization complimenting the model is discussed. Finally, the last section discusses feedback from the expert sessions discussing the other results.

5.1 Novel framework for maritime anomalies

The framework for anomaly types created for this work consists of four different groups of anomalies. The framework groups have various types detectable from point-based anomaly detection strategies. The framework is the foundation for understanding maritime anomalies and determines the subset of the anomaly types included in the proof of concept. The selected types create a combined result for the maritime anomalies. Each anomaly type gets a score to help explain what type contributed to an anomaly as part of the transparency models.

The framework improves the work of Riveiro et al., 2018, adapting it to better include context as a different aspect for all anomaly types and adding routing anomalies. Further, the scoping expands with explicit anomaly types that provide examples. Elaboration on the process and extra grouping information can be found in Section 4.2.

These groups are data anomalies, kinematic anomalies, positional anomalies, and routing anomalies. The data anomalies divide into detection and transmission anomalies. Speed and course anomalies split kinematic anomalies. Location-based anomalies take the form of positional anomalies. Finally, the changes and deviations from expected tracks fall under routing anomalies. This framework for anomaly types can be seen in Figure 9 and also in Section 4.2.

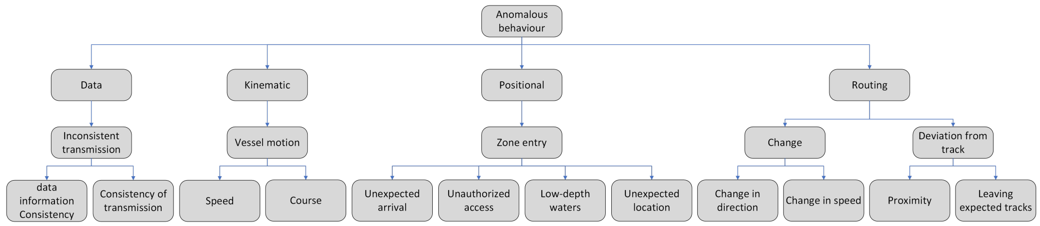


Figure 9: The types of maritime anomalies used

5.2 Transparent anomaly model

A maritime anomaly detection proof of concept combines self-organizing maps and Gaussian mixture models. A framework for anomaly groupings and types is the foundation for selecting types for the proof of concept. The primary goal is to use transparency models to explain what anomaly types contributed to the overall anomaly score. The visualization is also part of the proof of concept to visualize the anomaly and transparency scores. This visualization is elaborated in Section 5.3. Section 3.3 elaborates on the theory of self-organizing maps and Gaussian mixture models that serve as the tools used for the proof of concept. Finally, Section 4.3 discusses the implementation and parameter selection to achieve the used model.
This section will provide the results from the proof of concept, examining the overall model scores and their relation to the scores from the transparency models. These are also referred to as the *all* model for the overall model and the *speed*, *course*, *speed delta*, *course delta*, *location* transparency models.
The model's training was done with a month of data as outlined in Chapter 4.1. The first two days of the next month serve as test data. The anomaly types from the anomaly grouping, framework type, and used features are part of the final model that is summarized as follows:

Overall: Longitude, Latitude, Speed, Course, Speed delta, Course delta
Speed: Longitude, Latitude, Speed
Course: Longitude, Latitude, Course
Speed delta: Longitude, Latitude, Speed delta
Course delta: Longitude, Latitude, Course delta
Location: Longitude, Latitude

The results examine the test data in subsets based on their respective scores and their correlation. A set of 39 vessels with more than 200 scores below 0.01 were further selected, with the individual examined data points consisting of the lowest score for each of these vessels in the test data. This set represents vessels with many low normalcy scores, where the individual examined data points represent the least normal. These 39 vessels are examined and classified as anomaly, normal, or data quality issue. The lowest data points for each vessel serve as validation points for the success of the transparency strategy. The first part of validation starts by examining the 39 vessels for their anomaly scores.

The 39 vessels represent 49.56% of all scores below 0.01 in this test set. Allowing manual examination of a manageable sample size representing a large portion of the resulting detected anomalies. These 39 vessels represent a large portion of anomalous behavior. The manual examination of the 39 vessels aims to validate that the behavior detected constitutes outliers. The 39 vessels have their identifier replaced with a number. For each vessel, the number of data points within the test data set is in the total column. The percentage of data points for each vessel that scored under 0.01 and 0.03 is in the columns warning and threshold. A score of 0.01 and 0.03 indicates the probability density of the data point is lower than 99% and 97% of scored points, respectively. The 0.01 and 0.03 as default values for warning and threshold represent less than ten vessels with scores below the threshold at one time while presenting highlighted vessels most of the time. The threshold and warning subgroup of data points are further explored in later results, comparing them to other subsets of the scored data points. Note that the threshold is thus consistently higher since it includes all the warning data points. The vessels are organized based on the number of data points below 0.01. All the vessels classify as anomaly, normal, or data according to whether they can be considered an outlier following a manual inspection. This classification considers both data and anomaly as anomalous, where data constitutes anomalous behavior unlikely to be the vessel movement.

Table 5 contains the first 12 of 39 vessels as a sample of the examined vessels. The table of all 39 vessel is in Appendix A.2. From the manual examination resulting in the table, all 39 vessels behave like outliers in the detection moments. However, 11 out of 39 vessels selected (28.2%) can be considered likely to be detection data quality issues. These 11 out of 39 vessels are likely not behavior from the vessel but in detecting vessel behavior before becoming a part of the data set. Some examples of these data quality issues include non-moving vessels with speed, large course changes, and non-connected trajectories.

The transparency of the graph was measured by whether or not the alterna-

Table 5: Overall statistics (12 of 39 vessels)

Nr	Warning	Threshold	Overall
1	72%	89%	Anomaly
2	10%	19%	Data
3	7%	16%	Anomaly
4	17%	39%	Data
5	6%	10%	Anomaly
6	78%	92%	Anomaly
7	9%	17%	Anomaly
8	11%	20%	Anomaly
9	17%	18%	Anomaly
10	26%	37%	Anomaly
11	14%	19%	Anomaly
12	12%	25%	Data

Table 6: Transparency alignment (12 of 39 vessels)

Nr	Bar graph	Alternative method
1	Location	Location
2	Course delta	Course delta
3	Course delta	Course delta, Course
4	Course delta, Course	**Course delta, Course**
5	Location	Location
6	Location	Course (Large), Speed (Large)
7	Speed delta, Course delta	Course, Course delta, Speed
8	Speed delta	Speed delta
9	Location	Location
10	Course delta, Course	Course delta, Speed or Speed delta
11	Course delta	Course delta, Course
12	Course delta	Course delta

tive transparency tool outlined in Section 4.3 aligned with the transparency models from the proof of concept. The transparency alignment validation

point is the lowest value for each of the 39 selected vessels. With this, the different transparency strategy graphs serve as a reference tool to determine the direction of the data point. Furthermore, we have the correlation and histograms between various individual feature scores for subsets of the entire test set. The direction of the reference tool toward normalcy and the bar graphs is in Table 6 and the entire table in Appendix A.3. This table compares the interpretation of the bar graph to the interpretation of the depth-first graph-based search. Both results for the alternative method required manual intervention to find the optimal explanation. For these, a close second option on the first iteration produced better results on the second iteration. The interpretation nuances can be found in Section 4.3.
Table 6 compares the transparency to the reference transparency tool for the lowest data point of each of the 39 vessels. This reference transparency tool compares how the data point deviates from normal by re-scoring data points with different feature values to determine the local optimum. The transparency models are trained separately and use a subset of features to indicate the contribution of the corresponding anomaly type. Commas separate multiple features in order of priority for the bar graph column in the table. Commas for the alternative method show the next iteration step, with 'or' indicating multiple options towards normalcy for that depth. Bold results for the alternative method indicate that it resulted from scoring another first step with less change in the data point feature.

Table 7: Correlation values between overall score and transparency scores

Anomaly type	0.00 - 1.00	0.00 - 0.01	0.00 - 0.03	0.00 - 0.10	0.50 - 1.00	0.90 - 1.00
location	0.120	0.147	0.127	0.086	0.034	-0.019
Speed	0.099	0.128	0.108	0.128	-0.014	-0.082
Course	0.524	0.161	0.034	0.100	0.310	0.208
Speed delta	0.209	0.216	0.212	0.153	0.082	0.019
Course delta	0.226	0.275	0.273	0.149	0.102	0.020

Table 7 includes the correlation between various transparency scores and the overall score. The correlation values are calculated by how the transparency model scores relate to the overall score for various subgroups. The subgroups consist of below 1%, below 3%, below 10%, above 50%, and above 90%, all comparing the correlation to the overall score. From this correlation table, it is clear that the course significantly impacts the overall score. Further, it shows the lack of impact of other anomaly types in the more normal data points.

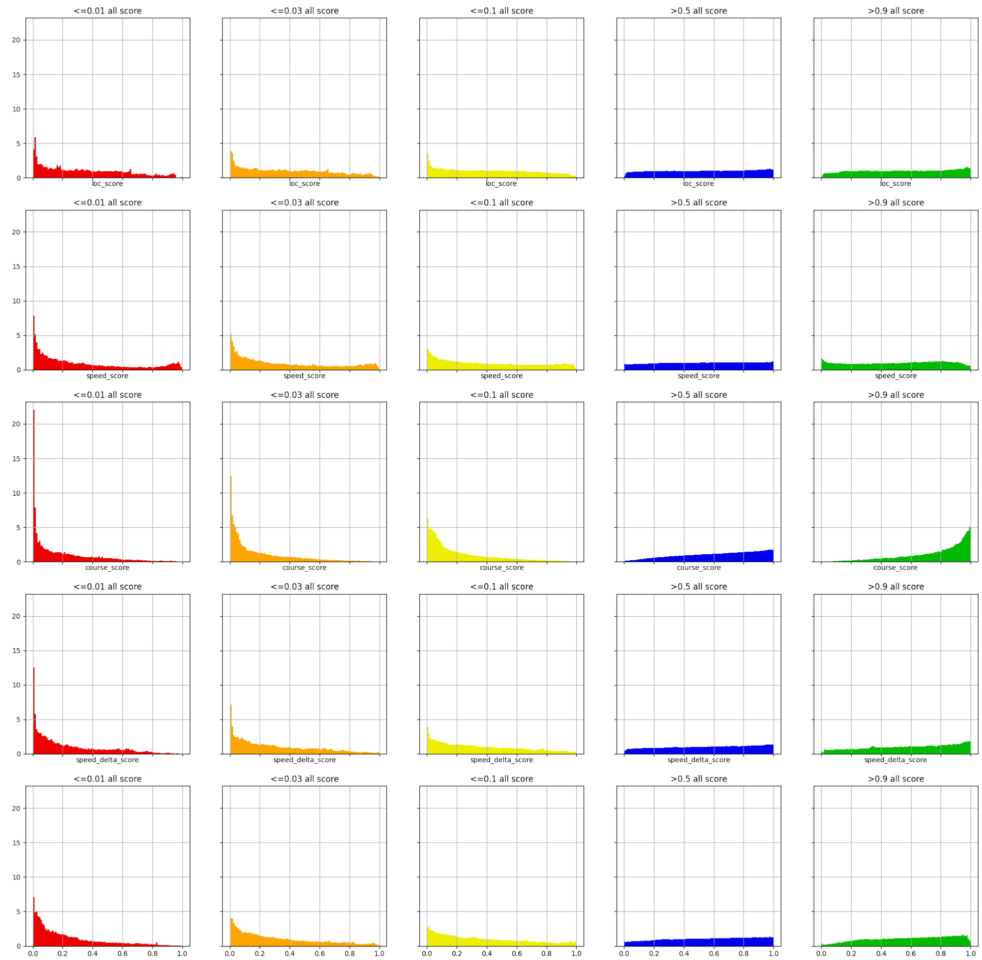


Figure 10: Histograms for feature scores, subsets of the results

The histograms of various features for different subsets of the tests serve as a validation result. Examining these histograms shows the effects of various anomaly types on the overall score. By definition, the normalization step ensures that the histograms are equal in each distribution for the entire test set. Thus, the histograms do not include those for the entire test set. The histograms below 1%, below 3%, below 10%, above 50%, and above 90% show trends in the data. Each histogram shows the frequency of different transparency model scores in the subset of data, sharing the same y-axis and x-axis between all 25 histograms. Each row of histograms corresponds to a different transparency model, where the columns have different colors corresponding to the same subset. The histograms show that a single low score does not have to impact the overall model drastically. The histograms also show the impact course has by both the frequency of low course scoring data points seen as very anomalous and the inverse interaction on high scoring data points.

5.3 Maritime anomaly visualization

To complement the anomaly detection proof of concept, according to the design found in Section 4.4, a visualization was created using JavaScript and MapBox as the map provider. This visualization served as an interface to interact with the proof of concept and to examine the required components to support a maritime anomaly detection implementation. This visualization has three primary anomaly detection-related features: the queue, transparency bar graph, and vessel history.

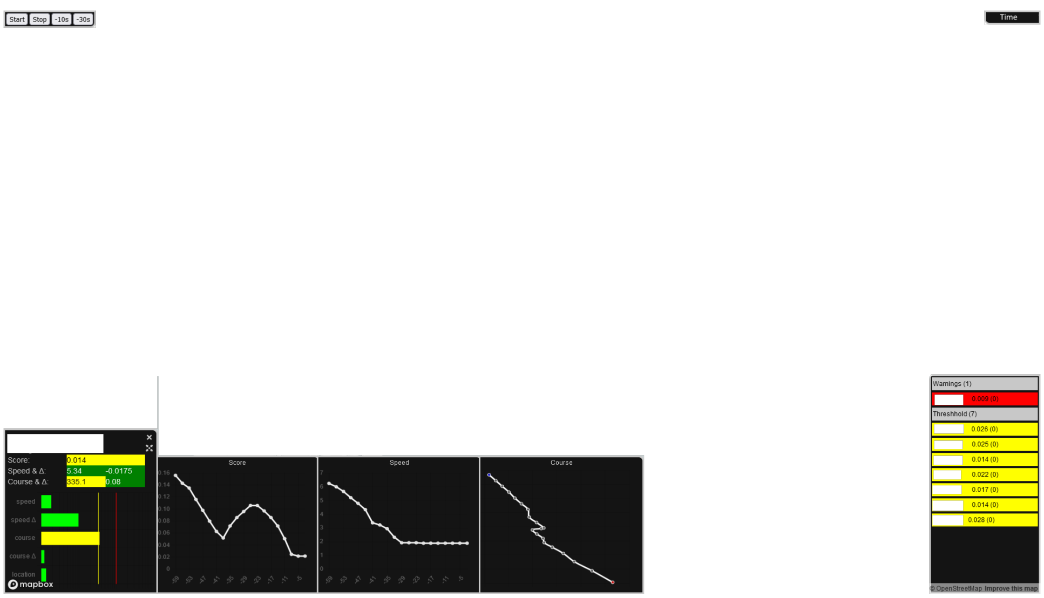


Figure 11: Visualization, with a vessel selected

Figure 11 is the resulting visualization. The bottom right has the queue system, while the bottom left has the vessel information and the transparency bar graphs. Lastly, the bottom has a set of graphs explaining the last minute of history.

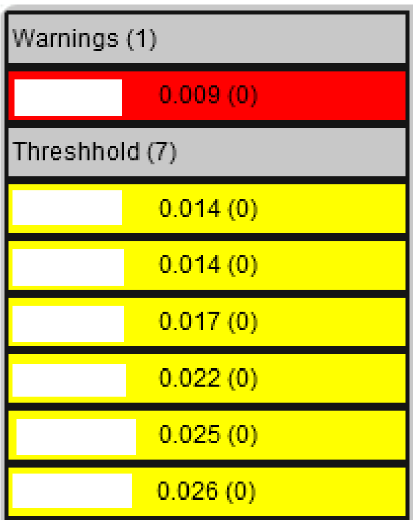


Figure 12: The queue for detected anomalies

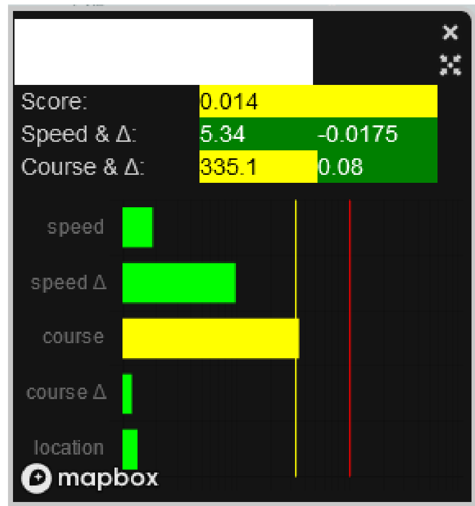


Figure 13: Vessel info and transparency

The queue, seen in 12, shows data points that got low scores from the anomaly model, in this case, grouped in two categories *warning* and *threshold*. *Warning* and *threshold* correspond to $< 1\%$ and $< 3\%$ respectively. Clicking on specific vessels in the queue will select them and shift the map's focus towards the data point. The history, vessel information, and transparency bar graph provide further information about that specific vessel.

The transparency bar graph in Figure 13 shows at a glance what features contribute towards an anomaly, using the separately scored subsets of features. The size of the bars is the inverse of the result to make the size appear more noticeable when it is more anomalous. The coloring of the bars is according to the same thresholds as set for the queue. This bar graph shows that course is a factor in the low score of the vessel.

The history graphs, seen in Figure 14, show the overall score, speed, and path visualize the history of the selected vessel. In a live VTS system, the graphs would likely be incorporated with history systems already present. The overall score over time graph on the left shows that the score decreased from 0.16 to 0.05, then increased before decreasing further towards the current low score in the previous minute. The speed graph shows the vessel first

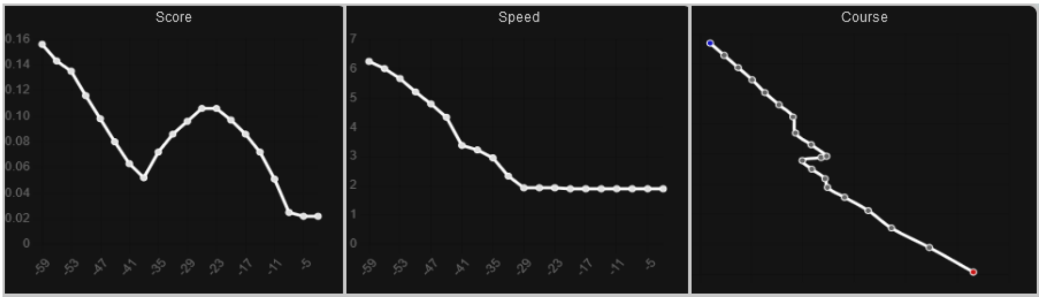


Figure 14: The history graphs for the selected vessel

slowed and then maintained a steady speed just below 2 knots. The third graph shows that the ship's course was heading northwest from the red to the blue dot.

5.4 Expert validation

A set of four one-on-one expert sessions put the other results into a better perspective and allow improved future work recommendations in Chapter 6. These sessions are not meant to validate the model, but discuss the future of the field and provide improved recommendations. This means the sessions merely support future speculation and not a validation of the proof of concept.
Some comments were made by the experts, for example that an essential missing aspect is the importance of larger vessels and segmenting of vessel types. The people more accustomed to VTS workflow also point towards the importance of vessel interaction and having the tool take up limited operator focus. The current feature set selected more towards enforcing rules and possibly too much movement in visualization. Overall the experts seemed optimistic that the proof of concept shows promise towards an aid that could help VTS operators prioritize. Since it showed consistent overall results and the transparency is a valuable addition.

6 Conclusion

This chapter concludes this work by connecting back to the research questions from Section 1.2, answering them individually, and discussing the method's limitations and results. Each research sub-question also touches on future work, which comprehensively elaborates on the direction of the maritime anomaly detection field. The results, found in Chapter 5, serve to answer the research questions discussed in this chapter. These results are obtained through a process starting from existing literature found in Chapter 2, exploring the theoretical basis in Chapter 3. Building on the theoretical basis and using the data described in Chapter 2, the caveats for implementation and elaboration on design are in Chapter 4. The structure of the results, implementation, and methodology chapters correspond to four categories aiming to answer specific research questions: the anomaly framework, maritime anomaly detection, visualization, and expert validation. The anomaly framework answers SQ1, whereas maritime anomaly detection discusses SQ3 and SQ4. The combination of the transparency strategy and the visualization covers SQ5. Visualization further serves as a basis for the expert validation that helps answer SQ6. Finally, the sub-questions combine to answer the main research question.

SQ1: What types of anomalous behavior exist for maritime traffic?

This work provides a framework for grouping maritime anomalies and a non-comprehensive set of individual anomaly types. Section 5.1 presents this framework, with further elaboration in Section 4.2. This framework builds upon groupings found in literature, as explored in Section 2.2.1. The result is a clear and well-segmented grouping of anomaly types. The framework supports four groups: data anomalies, kinematic anomalies, positional, and routing anomalies. These groups have a clear description and have between two and four anomaly types. However, the individual types will and should have more types added as relevant for future works. This framework was previously lacking in the field, although some attempts at grouping existed.

SQ2: How can we evaluate the effectiveness of maritime anomaly detection?

The possible evaluation techniques are first considered in the literature review, Section 2.2.1. Section 3.3 discusses the choice of strategy and criteria for evaluation. The validation of the proof of concept results considers these evaluation criteria, providing validation beyond single-point examples. The effectiveness of maritime anomaly detection is challenging and absent in the area. Due to the lack of concrete definitions for a maritime anomaly, most evaluations are individual examples. This work used this approach by attempting to sample scores and visualize them per vessel. With this, comparisons could be made within these vessels to understand what behavior the model considers anomalous. This validation method allows the evaluation of low probability density-based anomaly detection methods. However, it lacks connection with real-world effectiveness since it can not prove the applicability of the results or if these low probability density areas correspond to actual anomalies. Future work should follow one of two paths: 1) Scope a better definition of an anomaly, starting from what would be useful vessels for VTS operators. 2) A large pilot study in real-time VTS systems to codify expert input into a grouping or labeled data set.

SQ3: What effective methods exist to detect maritime anomalies?

A literature review discusses existing strategies, determining how these relate to each other, found in Section 2.2.1. Unfortunately, the effectiveness of most directions is unclear, and the strategies that come with concrete results tend to have no reference material. This work selected two promising methods from the literature review based on the importance of transparency and expandability. The Self-organizing map and Gaussian mixture model approach shown in (Riveiro et al., 2008) served as the maritime anomaly detection strategy for this work. It is also interesting to look into the recurrent neural network approach from a more recent promising paper (Zhao & Shi, 2019). Another exciting avenue is the ability to build expected tracks (Vespe et al., 2012), possibly in combination with path prediction could combine anomaly detection and deviations.

SQ4: What is the effectiveness of using self-organizing map and Gaussian mixture models to detect maritime anomalies?

The self-organizing map (SOM) and Gaussian mixture model (GMM) are effective tools for creating probability density functions in an unsupervised manner. The accuracy of the resulting model was primarily validated by examining low-scoring vessels for outlier behaviour. In combination with the transparency models, the model determines the impact of specific features. The combination of SOM and GMM allows the scoring of the problem space. However, the tools also require high data quality in large quantities and have some drawbacks. These drawbacks are that the approach requires continuous features to correlate with location and can only detect the entire problem space. As can be seen by the choices to exclude some features for the final model and to filter non-moving vessels, some segmentation may be required to ensure that the problem space does not have junk data points. Whether or not a feature is continuous enough can be debugged during model building using the additional checks mentioned in Section 4.3. If the following projects use the self-organizing map and Gaussian mixture model strategy, they should incorporate vessel types. Training separate models is an option, but this is resource-intensive and will require an even more extensive training data set. The Self-organizing map and Gaussian Mixture model strategy performed well in detecting outliers, however whether the outliers constitute maritime anomalies is up to the maritime anomaly definition used. The strategy also cannot incorporate previous data points into the model unless it includes a feature referencing the previous state.

SQ5: What level of transparency can be achieved using self-organizing maps and Gaussian mixture models to detect maritime anomalies?

The level of transparency achieved using separate models, combining the self-organizing map and Gaussian mixture model, is high, as found by comparing the transparency models to a reference tool leveraging re-scoring using the overall model. However, the nature of specific features in every model makes it difficult to distinguish these. Primarily location is an essential factor for all subsets and the overall model. Since location anomalies indicate a lack of data in that area, other features are unlikely to have distributions that fit these areas. As such, the location is somewhat over-represented. Assuming an expert is looking at the bar graph designed to fit with the transparency strategy, it aligns well with a different more resource-intensive strategy to determine what would make the point more normal. The transparency models and the overall model have difficulty perfectly aligning because they are separately trained models, and matching their clustering density or even clusters is thus impossible. Using the transparency models as a subset of the trained self-organizing map was unsuccessful due to the clustering no longer aligning accurately with the test data and thus lacking accuracy. The field of maritime anomaly detection does not yet recognize the importance of transparency. The promising avenue toward adoption appears to be for VTS operators to trust the systems by providing insight into the reason for the detection.

SQ6: How can a maritime anomaly scoring tool aid VTS operators?

This work uses one-on-one expert sessions to determine how maritime anomaly scoring can benefit VTS operators, with the resulting proof of concept as a reference. Further elaboration on the expert sessions can be found in Section 4.5. The feedback gathered from these sessions and lessons during the process aim to provide information on how maritime anomaly scoring can aid VTS operators in live systems. Maritime anomaly scoring has a few different avenues. With no classification of an anomaly currently existing in current literature, no system can gauge the accuracy of their strategy. With this, the most likely use case would be to have a priority list that does not provide warnings. With the priority list approach, the model highlights and reasons why a vessel requires the attention of a VTS operator. More extensive user testing would be required for an implementation to determine whether such a priority list would have a positive effect on their workflow rather than a distraction.

RQ: What is the viability of anomaly detection aid for VTS operators?

The field of maritime anomaly detection, especially concerning the real-world application, is in its infancy. The most promising short-term avenue would consist of a priority list that allows the VTS operator to have guidance on areas or vessels that require their attention. More intricate systems would require a real-time deployment to monitor what users consider worthwhile anomalies. Alternatively, a concerted effort to define what may be considered anomalous does not rely on detection tools. Transparency should be a consideration for future projects since products in this area are unlikely to achieve adoption without being able to represent what makes detected anomalies anomalous accurately. Data quality is also a significant factor for viability. Since erroneous data are considered outliers by a probability density-based model, causing false positives for anomalies. This work has various contributions in defining a maritime anomaly, working towards transparency, evaluation techniques, and even considerations regarding actual implementation and an expert validation of strategies. There are still many gaps to be filled to achieve a product that can aid VTS operators. Sessions with field experts identified promising directions for a viable priority list. For this, the importance of vessel interaction and early warnings is evident. Further, more open areas could also be interesting since it provides the VTS operators more time to respond to maritime anomalies.

6.1 Contributions and limitations

This work contributes in various aspects by looking at the entire maritime anomaly detection field, resulting in recommendations for the whole field. This section discusses the various contributions by relating them to existing literature and continues by discussing the limitations of this work.

The contributions of this work consist of the framework, validation strategies, transparency importance, transparency strategy, visualization principles, and the steps for the maritime anomaly detection field toward practical viability. The maritime anomaly framework outlines specific groupings, building upon existing work to better differentiate detectable and viable anomaly detection. The validation strategies improve upon lacking validation strategies in the literature to provide better insight into the effectiveness of maritime anomaly detection models. Transparency was identified as an essential aspect of the maritime anomaly detection field in the role of adoption and trust. With this transparency focus, a strategy is proposed building upon existing self-organizing maps, and Gaussian mixture models approach to enable transparency for anomaly types identified in the maritime anomaly framework. To interact with the proof of concept, an accompanying visualization provided the interface for the transparency and incorporated maritime anomaly detection components highlighted by the literature. Finally, the main contribution of this work is the recommendations for the future of the maritime anomaly detection field consisting of shortcomings that will help the field mature toward practical applicability.

The study has limitations, primarily imposed by taking a broad view of the process. This expansive view restricts the depth of different aspects. As such, only a single maritime anomaly detection strategy provides the basis for enabling transparent maritime anomaly detection. Further, while validation is improved, it still lacks comparison between various systems and more interpretation besides the researcher. This work also treats the data as-is without improving the existing data quality. Earlier feedback sessions with vessel traffic management experts may have enhanced focus on the relevant anomaly types for the end-users of maritime anomaly detection systems. Lastly, another important limitation is the approach to expert validation. While the expert sessions provide valuable information, these do not serve as sufficient validation for the viability of the proof of concept of this work.

References

Analytica, O. (2021). Container shipping congestion will disrupt world trade. *Emerald Expert Briefings*. https://doi.org/10.1108/oxan-db260025

Baker, J., Watkins, E., & Osler, D. (2021). Suez canal remains blocked despite efforts to refloat grounded ever given. *Lloyd's List*. Retrieved February 9, 2022, from https://lloydslist.maritimeintelligence.informa.com/LL1136229/Suez-Canal-remains-blocked-despite-efforts-to-refloat-grounded-Ever-Given

Baldauf, M., & Wiersma, E. (1998). Risk recognition and collision avoidance by vts operators. *IFAC Proceedings Volumes*, *31*, 239–244. https://doi.org/10.1016/s1474-6670(17)40099-1

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, *6*. https://doi.org/10.17705/1jais.00065

Brodje, A., Lützhöft, M., & Dahlman, J. (2010). The whats, whens, whys and hows of vts operator use of sensor information. *Human Performance at Sea (HPAS)*.

Cansado, A., & Soto, A. (2008). Unsupervised anomaly detection in large databases using bayesian networks. *Applied Artificial Intelligence*, *22*, 309–330. https://doi.org/10.1080/08839510801972801

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Comput. Surv.*, *41*(3), pp 1–58. https://doi.org/10.1145/1541880.1541882

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T. P., Shearer, C., & Wirth, R. (2000). Crisp-dm 1.0: Step-by-step data mining guide.

Council, N. R. (1996). *Vessel navigation and traffic services for safe and efficient ports and waterways : Interim report*. National Academy Press. Retrieved February 8, 2022, from https://doi.org/10.17226/926

d'Afflisio, E., Braca, P., & Willett, P. (2021). Malicious ais spoofing and abnormal stealth deviations: A comprehensive statistical framework for maritime anomaly detection. *IEEE Transactions on Aerospace and Electronic Systems*, *57*(4), 2093–2108. https://doi.org/10.1109/TAES.2021.3083466

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, *59*, 56–62. https://doi.org/10.1145/2844110

Gonin, I., Johnson, G., Shalaev, R., Tetreault, B., & Alexander, L. (2009). *Uscg development, test and evaluation of ais binary messages for en-*

hanced vts operations (2009 International Technical Meeting of The Institute of Navigation).

Handayani, D. O. D., Sediono, W., & Shah, A. (2013). Anomaly detection in vessel tracking using support vector machines (svms). *International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, *2*, 213–217. https://doi.org/10.1109/ACSAT.2013.49

Katsilieris, F., Braca, P., & Coraluppi, S. (2013). *Detection of malicious ais position spoofing by exploiting radar information*.

Kizilkan, Ö., Kucuksille, E. U., & Kabul, A. (2015). Data mining approcah for performance analysis of variable speed refrigeration system. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *19*. https://doi.org/10.19113/sdufbed.53755

Kowalska, K., & Peel, L. (2012). *Maritime anomaly detection using gaussian process active learning* (Vol. 15). IEEE. Retrieved May 5, 2022, from https://ieeexplore.ieee.org/document/6289940

Lane, R. O., Nevell, D. A., Hayward, S. D., & Beaney, T. W. (2010). *Maritime anomaly detection and threat assessment* (Vol. 13). IEEE. https://doi.org/10.1109/ICIF.2010.5711998

Lin, B., & Huang, C.-H. (2006). Comparison between arpa radar and ais characteristics for vessel traffic services. *Journal of Marine Science and Technology*, *14*. https://doi.org/10.51400/2709-6998.2072

Macdonald, J. (2002). A note on the shore applications of ais technology. *Maritime Studies*, *2002*, 25–27. https://doi.org/10.1080/07266472.2002.10878672

Mapbox. (13). Mapbox gl js. *Mapbox*. Retrieved May 16, 2022, from https://docs.mapbox.com/mapbox-gl-js/guides/

Mariscal, G., Marbán, Ó., & Fernández, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, *25*, 137–166. https://doi.org/10.1017/s0269888910000032

Martineau, E., & Roy, J. (2011). Maritime anomaly detection: Domain introduction and review of selected literature. *Defense Technical Information Center*. Retrieved May 2, 2022, from https://apps.dtic.mil/sti/citations/ADA554310

Mersiovsky, T., Thekkottil, A., Hanne, T., & Dornberger, R. (2018). Optimal learning rate and neighborhood radius of kohonen's self-organizing map for solving the travelling salesman problem. *Proceedings of the*

2nd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence - ISMSI '18. https://doi.org/10.1145/3206185.3206194

Obradovic, I., Milicevic, M., & Zubrinic, K. (2014). Machine learning approaches to maritime anomaly detection. *Nase More*, *61*, 96–101. Retrieved May 5, 2022, from https://www.researchgate.net/publication/288458284_Machine_learning_approaches_to_maritime_anomaly_detection

on Trade, U. N. C., & Development. (2018). Review of maritime transport 2018 (J. Hoffmann, Ed.). *unctad.org*. Retrieved February 9, 2022, from https://unctad.org/webflyer/review-maritime-transport-2018

on Trade, U. N. C., & Development. (2021). Review of maritime transport 2021 (S. N. Sirimanne, Ed.). *unctad.org*. Retrieved February 9, 2022, from https://unctad.org/webflyer/review-maritime-transport-2021

Osekowska, E. (2014). *Design and implementation of a maritime traffic modeling and anomaly detection method* (Doctoral dissertation). Retrieved May 5, 2022, from https://www.dissertations.se/dissertation/99b773c0c5/

Placek, M., & Freightos. (2022). Global container freight index. *Statista*. Retrieved February 9, 2022, from https://www.statista.com/statistics/1250636/global-container-freight-index/

Praetorius, G., & Lützhöft, M. (2012). Decision support for vessel traffic service (vts): User needs for dynamic risk management in the vts. *Work*, *41*, 4866–4872. https://doi.org/10.3233/wor-2012-0779-4866

Rai, A. (2019). Explainable ai: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*, 137–141. https://doi.org/10.1007/s11747-019-00710-5

Riveiro, M., Falkman, G., & Ziemke, T. (2008). Improving maritime anomaly detection and situation awareness through interactive visualization. *International Conference on Information Fusion*, *11*, 1–8. Retrieved March 25, 2022, from https://ieeexplore.ieee.org/abstract/document/4632191

Riveiro, M., Pallotta, G., & Vespe, M. (2018). Maritime anomaly detection: A review. *WIREs Data Mining and Knowledge Discovery*, *8*. https://doi.org/10.1002/widm.1266

Rossum, V., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

Roy, J. (2010). Rule-based expert system for maritime anomaly detection (E. M. Carapezza, Ed.). *Sensors, and Command, Control, Commu-*

nications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IX, *7666*. https://doi.org/10.1117/12.849131

Rzepka, C., & Berger, B. (2018). User interaction with ai-enabled systems: A systematic review of is research. *39*.

Shin, D. (2020). User perceptions of algorithmic decisions in the personalized ai system:perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media*, *64*, 1–25. https://doi.org/10.1080/08838151.2020.1843357

Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. https://doi.org/10.1016/j.chb.2019.04.019

Söllner, M., & Leimeister, J. M. (2013). What we really know about antecedents of trust: A critical review of the empirical information systems literature on trust. Nova Science Publishers. Retrieved April 2, 2022, from https://ssrn.com/abstract=2475385

Strang, G. (1988). *Linear algebra and its applications*. Harcourt, Brace, Jovanovich, Publishers.

van Meersman, H., van de Voorde, E., & van Elslander, T. (2012). Port congestion and implications to maritime logistics. In D. Song & P. Photis M (Eds.). Emerald Group Publishing Limited. https://doi.org/10.1108/9781780523415004

Vespe, M., Visentini, I., Bryan, K., & Braca, P. (2012). *Unsupervised learning of maritime traffic patterns for anomaly detection* (Vol. 2012). https://doi.org/10.1049/cp.2012.0414

Vettigli, G. (2018). Minisom: Minimalistic and numpy-based implementation of the self organizing map. Retrieved July 6, 2022, from https://github.com/JustGlowing/minisom/

Will, J., Peel, L., & Claxton, C. (2011). *Fast maritime anomaly detection using kd tree gaussian processes* (Vol. 2). IMA. Retrieved May 5, 2022, from https://piratepeel.github.io/pubs/FastMaritimeAnomalyDetect.pdf

Yeo, N., Lee, K., Venkatesh, Y., & Ong, S. (2005). Colour image segmentation using the self-organizing map and adaptive resonance theory. *Image and Vision Computing*, *23*, 1060–1079. https://doi.org/10.1016/j.imavis.2005.07.008

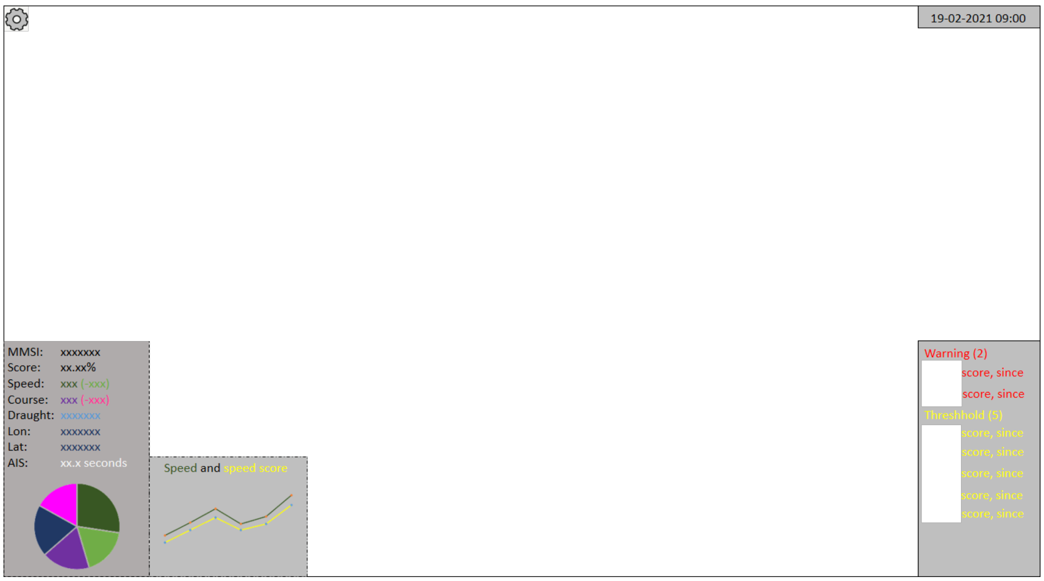
Yin, H. (2007). Nonlinear dimensionality reduction and data visualization: A review. *International Journal of Automation and Computing*, *4*, 294–303. https://doi.org/10.1007/s11633-007-0294-y

Zhao, L., & Shi, G. (2019). Maritime anomaly detection using density-based clustering and recurrent neural network. *Journal of Navigation*, *72*, 894–916. https://doi.org/10.1017/s0373463319000031

Zhen, R., Jin, Y., Hu, Q., Shao, Z., & Nikitakos, N. (2017). Maritime anomaly detection within coastal waters based on vessel trajectory clustering and naïve bayes classifier. *Journal of Navigation*, *70*, 648–670. https://doi.org/10.1017/s0373463316000850

A Appendix

A.1 Initial visualization design



A.2 Complete overall statistics

Nr	Warning	Threshold	Overall
1	72%	89%	Anomaly
2	10%	19%	Data
3	7%	16%	Anomaly
4	17%	39%	Data
5	6%	10%	Anomaly
6	78%	92%	Anomaly
7	9%	17%	Anomaly
8	11%	20%	Anomaly
9	17%	18%	Anomaly
10	26%	37%	Anomaly
11	14%	19%	Anomaly
12	12%	25%	Data
13	8%	13%	Data
14	13%	44%	Anomaly
15	24%	58%	Anomaly
16	2%	5%	Anomaly
17	1%	5%	Anomaly
18	86%	95%	Anomaly
19	3%	6%	Anomaly
20	28%	31%	Anomaly
21	3%	5%	Anomaly
22	5%	6%	Data
23	9%	24%	Anomaly
24	4%	9%	Anomaly
25	11%	28%	Data
26	7%	11%	Anomaly
27	5%	9%	Anomaly
28	9%	16%	Data
29	11%	19%	Anomaly
30	5%	12%	Anomaly
31	4%	12%	Anomaly
32	2%	4%	Data
33	10%	15%	Anomaly
34	4%	9%	Data
35	1%	1%	Data
36	3%	8%	Anomaly
37	14%	19%	Anomaly
38	4%	11%	Anomaly
39	5%	9%	Data

A.3 Complete transparency alignment

Nr	Bar graph	Alternative method
1	Location	Location
2	Course delta	Course delta
3	Course delta	course delta, course
4	Course delta, course	**Course delta, course**
5	Location	Location
6	Location	Course (Large), speed (Large)
7	Speed delta, course delta	Course, Course delta, Speed
8	Speed delta	Speed delta
9	Location	Location
10	Course delta, course	Course delta, speed or speed delta
11	Course delta	Course delta, course
12	Course delta	Course delta
13	Location	Location
14	Course delta	Course(large change), Course delta
15	Course delta, course	Course
16	Speed, speed delta, location	**Speed delta, course (large change)**
17	Course delta, speed delta	Course delta, speed delta or course
18	Location	Course
19	Course delta, course, speed	Course delta, course or speed
20	Speed delta	Course (large change), Speed (large change)
21	Speed delta, Location	Location
22	Course	Course
23	Speed delta, course	Course, speed
24	Course delta	Course delta
25	Location	Location
26	Course	Course
27	-	Course
28	Course	Course
29	Course delta	Course delta (large change)
30	Course delta	Course delta (large change)
31	Speed delta, course delta	**Course delta, course**
32	Speed delta, course delta	Speed delta(large change), course or course delta
33	Location, course delta, speed	Course delta (large change)
34	Course delta	Course delta
35	Course delta, speed, speed delta	Course, speed delta, speed (large change)
36	Speed, course	Course (large change)
37	Course	Course (large change)
38	Course, speed	Course(large change)
39	Course delta	Course delta