LOCAL ANOMALY DETECTION IN MARITIME TRAFFIC USING VISUAL ANALYTICS

by

Fernando Henrique Oliveira Abreu

Submitted in partial fulfillment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
Aug 2020

Table of Contents

List of Tables

List of Figures

Abstract

With the recent increase in sea transportation usage, the importance of maritime surveillance to detect unusual vessel behavior related to several illegal activities has also risen. Unfortunately, the data collected by the surveillance systems are often incomplete, creating a need for the data gaps to be filled using techniques such as interpolation methods. However, such approaches do not decrease the uncertainty of ship activities. Depending on the frequency of the data generated, they may even make the operators more confused, inducing them to errors when evaluating ship activities to tag them as unusual. Using domain knowledge to classify activities as anomalous is essential in the maritime navigation environment since there is a well-known lack of labeled data in this domain. In an area where finding which trips are anomalous is a challenging task using solely automatic approaches, we use visual analytics to bridge this gap by utilizing users' reasoning and perception abilities. In the current work, we investigate existing work that focuses on finding anomalies in vessel trips and how they improve the user understanding of the interpolated data. We then propose and develop a visual analytics tool that uses spatial segmentation to divide trips into *subtrajectories* and give a score for each *subtrajectory*. We then display these scores in tabular visualization where users can rank by segment to find local anomalies. We also display the amount of interpolation in *subtrajectories* with the score so users can use their insight and the trip display on the map to make sense if the score is reliable. We did a user study to assess our tool's usability and the preliminary results showed that users were able to identify anomalous trips.

List of Abbreviations and Symbols Used

AIS Automatic Identification System

COG Course Over Ground

D3 Data-Driven Documents

DBSCAN Density-Based Spatial Clustering of Applications with Noise

DD Diverse Density

DRDC Department of Defense of Canada

ETA Estimated Time of Arrival

GPS Global Positioning Systems

HDC Heterogeneous Curvature Distribution

HDP-HMM Hierarchical Dirichlet Process Hidden Markov Model

HMM Hidden Markov Model

IMO International Maritime Organization

KDE Kernel Density Estimatio

LRF Minimum Description Length

MA Maximum Acceleration

MDL Minimum Description Length

MMSI Maritime Mobile Service Identity

MSOCs Coastal Marine Security Operation Centres

ROT Rate of Turn

S-AIS Satellite-based AIS

SOG Speed Over Ground

SWS Sliding Window Segmentation

TOST Trip Outlier Scoring Tool

VHF Very High Frequency

VTS Vessel Traffic Service

Acknowledgements

I wish to express my sincere appreciation to my supervisor, Dr. Stan Matwin, for offering me the opportunity to pursue my masters, for all the support he gave me through my program, and for the wisdom he shared with me during this period.

I would also like to pay my special regards to Dr. Fernando Paulovich, whose assistance helped me give a direction and shape for my thesis.

I wouldn't be able to thank Dr. Amílcar Soares enough for all the support he gave me inside and outside the academic environment. I wouldn't have started a masters degree if it wasn't for him.

My thanks to all my colleagues and staff from the institute, they contributed in several ways towards my accomplishment.

I also would like extend my gratitude to my friends at the lab and for all the moments we shared in the past two and half years.

My sincere thanks to my parents Marco and Silmara, who encouraged me to pursue this journey, and for pushing me to never give up at the hard times.

Last but not least, I would like to thank my wife Cyndi for being the pillar that supports me, I wouldn't be able to have finished my masters if it wasn't for her.

Chapter 1

Introduction

Maritime transportation is essential nowadays; about 90 percent of everything traded in the world is done by sea [36, 44, 61, 62], and it grows approximately 8.5% per year [12]. Since 2004, vessels of 300 gross tonnages or more which travel internationally, and cargo ships of 500 gross tonnages or more are obligated by the International Maritime Organization (IMO) to have Automatic Identification System (AIS) onboard[1] which produces a constant high volume of data [7, 46]. This technology transmits the vessel destination, speed, position, and many other items of static information [62], such as ship name and Maritime Mobile Service Identity (MMSI), which is used to identify a ship uniquely [36].

The Department of Defense of Canada (DRDC) and surveillance authorities, such as Coastal Marine Security Operation Centres (MSOCs) which are responsible to guarantee coastal safety, have an interest in using this data to uncover several potential issues [31, 42, 22], such as illegal transport of drugs, human trafficking, fishing in illegal areas, illegal immigration, sea pollution, piracy, and even terrorism [8]. These activities have a significant impact on society, environment, and economy, and for such, it is essential to identify these types of events as soon as possible [53, 52].

Vessels involved in these types of illegal activities usually follow specific patterns like unexpected stops, speeding, and deviations from standard routes [8, 23, 36]. Ships that are operating legally commonly travel through the same route, due to regulations [55] and because it is usually the shortest path between ports, which would decrease the vessel fuel consumption. For this reason, ships that navigate non-standard routes or show signals of route deviations can be potentially labeled as presenting anomalous behavior [8].

However, identifying which trips are anomalous is not an easy task for maritime operators due to the large volume of data AIS produces [62], which creates an overload

[1] `http://www.imo.org/en/OurWork/Safety/Navigation/Pages/AIS.aspx`

of instances to be analyzed manually. Currently, operators usually use systems that display vessels on a world map that they can use to track their movements [30]. Although this can help operators reach some awareness of what is going on in the sea, it can prove a difficult task trying to identify anomalous vessels among a large number of normal vessels [22].

There have been many works that focus on finding anomalies in an automated manner by creating alerts or events when a possible anomaly is discovered. However, the problem of automatically identifying anomalies is very complex and not well-defined [41]; additionally, it requires dynamic adaptation since humans will always try to change their *modus operandi* to not get caught, which in turn, makes automatic systems less reliable [39]. Thus, systems that automatically detect anomalies are rarely used in the real world [41, 39]. On the other hand, visualizations make use of humans' inherent ability to perceive patterns and filter information in combination with their creativity and background knowledge [40, 41, 32], which allows them to be able to analyze and understand complex, massive and dynamic data [6].

Secondly, the vast majority of algorithms proposed to identify anomalies automatically may not work for *local anomalies* [59] or they require labeled data to train a model [16, 47]. This means that deviations from normality that happen just in a small portion of a vessel trajectory may be left out when considering the trajectory as a whole, especially when analyzing works in the maritime domain. According to the literature review done in this thesis, most work involving *visual analytics* also doesn't focus on segmenting trajectories to find *local anomalies*, and those who tried to address this issue are very limited.

Lastly, when trying to analyze vessel trajectories from raw AIS data is that it can be faulty and incomplete. This can happen for multiple reasons. First, one of the frequencies used by AIS transceivers is Very High Frequency (VHF), which makes AIS data unreliable [60]. Second, Vessel Traffic Service (VTS) stations may miss several AIS messages from vessels traveling close to the coast due to information overloading [35]. Third, even though Satellite AIS has become more common, since it can capture longer ranges than shore-based AIS, it is common for the data received by it to have gaps since the Satellite is limited by its field of view and footprint, and the number of messages it can lose increases in regions with a high number of vessels [27].Finally, there are also cases where vessel crew interfere with AIS signal or turn the transponder off so they can cover illegal activities [34]. For this reason, vessel trajectories often need to be interpolated, which can increase algorithm accuracy [14]. However, anomalies found in the interpolated data may be incorrect if the interpolation was not done properly, or when many consecutive data points are missing. Therefore, it would be important to present information related to interpolation if an anomaly is detected in the interpolated region of a trajectory, such as what was the quality of that interpolation, or show the interpolation itself, so one can assess if the interpolation was done properly and if it is indeed an anomaly. It could also the user to further investigate what could possibly happen when there was not signal. However, to my knowledge, there is no work in this field that allows users to explore the potential impact of interpolation on anomalies.

In this work, we propose a tool which aims to tackle the problems mentioned above. We make very few assumptions who the users of this tool could be since we want it to be open source and as accessible as possible. Therefore, it is desired that such a system should be easy to use and learn.

1.1 Research Questions

Based on the problems previously mentioned the current work will try to answer the following research questions:

1. Is it possible to identify local anomalies using one or a combination of features given a port of origin and a port of destination?

2. Is it possible to make sense of the interpolation and the uncertainty it may cause when determining anomalies?

1.2 Proposal

To address both research questions, we propose a visual analytics framework called Trip Outlier Scoring Tool. An overview of this framework can be seen in Figure 1.1. The top portion shows the preprocessing step required every time a new dataset is an input to the system. This step is divided into four phases: (1) Integration, (2) Cleaning, (3) Segmentation, and (4) Feature Extraction. In the integration, trips are extracted from raw positional data and are combined with the voyage information. After that, in the cleaning phase, invalid trips and data are removed from the dataset, such as noisy data points. Then, we fill the trip gaps using kinematic interpolation, and finally, we compute attributes: speed, heading, accumulated travel distance for each data point. In the next phase, we automatically create spatial segments based on the minimum and maximum latitudes and longitudes from all trips data points. And in the last phase, every trip is divided into subtrajectories, one for each spatial segment, and then has features extracted for each of these subtrajectories i.e., maximum speed, distance traveled. The Web Server's main job is serving the visualization requests, but it also computes for each subtrajectory a score for each of the features used.

The visualization aggregates these scores and ranks the trips based on blue the scores; this is then displayed in a table in which the users can explore and select which features and segments they want to use to see the final score. This visualization also displays the percentage of data points that have been created for each segment and for each trip. The original trajectory and the segmentation can also be displayed on a map.

1.3 Contributions

The contributions of this work are the following:

- Proposal and development of a visual analytics tool for finding local anomalies in trip trajectories while also taking into account the trip's interpolation.
- We validate the proposed tool with an evaluation of its effectiveness into finding the most anomalous trips through a study conducted with 10 users.

1.4 Thesis outline

The remainder of this work is structured as follow. Chapter 2 provides the background, Chapter 3 gives an overview and survey on works that look to detect anomalies either in an automated or in a visual way. Chapter 4 describes the proposed tool

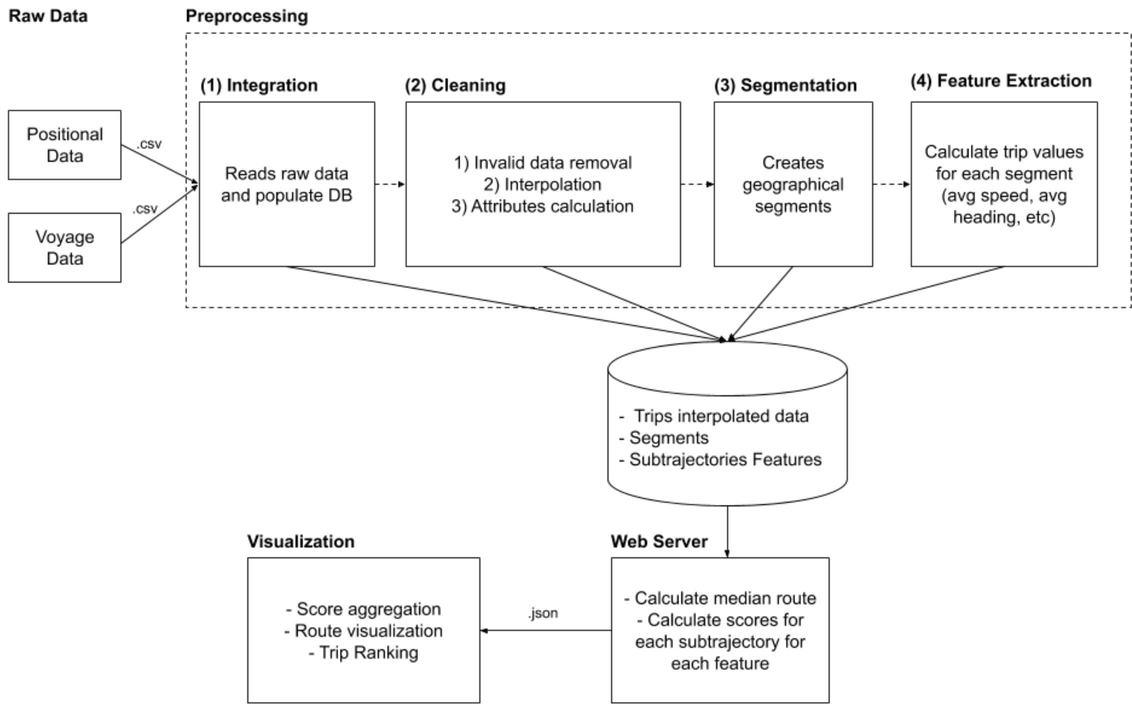

Figure 1.1. Overview of the framework of the Trip Outlier Scoring Tool

and discusses some of the decisions that were made. In Chapter 5 we present the study we have conducted to evaluate the user experience and the effectiveness of our proposed tool. Finally, in Chapter 6, we present a summary of this work and discuss some of our tool's limitations; and we propose some ideas for future work.

Chapter 2

Background and Terminology

In this chapter we will present the background and definitions of important concepts used in this thesis.

2.1 Automatic Identification System (AIS)

AIS is a self-reporting device which is capable of transmitting information about its vessel to other vessels and to coastal authorities. It was initially created with the intent to help avoid collisions between vessels at sea, but nowadays it is heavily used by maritime authorities to find potential *threats* at sea.

AIS works by integrating Very High Frequency (VHF) transceivers with Global Positioning Systems (GPS) and ship sensors, such as gyrocompass and rate of turn indicator, to broadcast information every 2 to 10 seconds depending on the vessel speed and every 3 minutes if it is anchored. The messages consist of *dynamic kinematic* data, such as vessel speed, position, rate of turn, and a Maritime Mobile Service Identity (MMSI) number which uniquely identifies each device. It also sends *dynamic non kinematic information*, which is voyage related information, such as destination, time of arrival, together with *static* information about the vessel, such as the type of ship, the vessel name, and International Maritime Organization (IMO) number. An overview of the information sent by AIS messages is shown in Figure 2.1.

The broadcast information can usually be received by other vessels equipped with a receiver, which is used to avoid collisions especially when they are navigating in conditions of restricted visibility. It is also collected by coastal receivers which can receive signals from vessels up to 40 nm away [10]. Due to this coverage limitation, Satellite-based AIS (S-AIS) has been also used to receive messages that are out of range of coastal stations. However, S-AIS is less consistent and has a lower update rate when compared to terrestrial AIS [27]. An overview of how AIS works can be seen on Figure 2.2.

Type of Information	Information	Broadcasting Rate
Dynamic Information	Maritime Mobile Service Identity number (MMSI) Ship position Speed over ground Course over ground Navigational status Time Etc.	At anchor or moored (<3 kts): 3 min At anchor or moored (>3 kts): 10 s Ship 0–14 kts: 10 s Ship 0–14 kts and changing course: 3.3 s Ship 14–23 kts: 6 s Ship 14–23 kts and changing course: 2 s Ship >23 kts: 2 s
Static and Voyage-related Information	MMSI number Ship type Length and beam Estimated time of arrival Destination Etc.	Every 6 min and on request

Figure 2.1. AIS sensor message information and update rates [21].

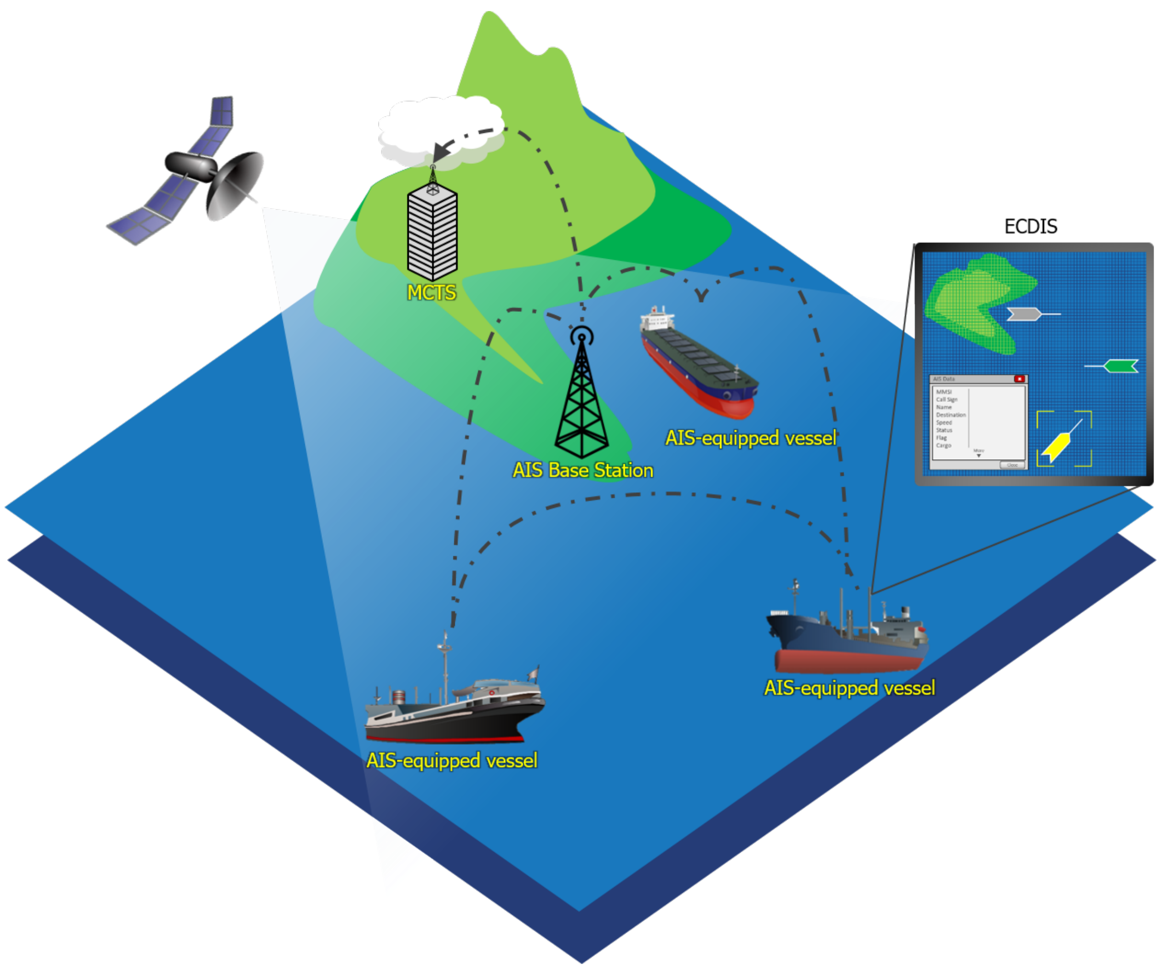

Figure 2.2. Overview of the Automatic Identification System (AIS)[1].

2.2 Anomaly detection

2.2.1 Types of anomalies

The term *anomaly* can have different interpretations depending on the context used. In this work, we will use a similar definition given by [42] in which something is considered anomalous if it is deviating from what is usual, normal, or expected. To decide what is normal we will aggregate all vessel data from the same type of vessel, given that different classes of vessel can have different behaviour [57]. And we will consider that values that deviate from this aggregation will be considered an *anomaly*. Example of anomalies are vessels of high tonnage travelling at high speed near the coast or vessels that don't travel on sea lanes.

Anomalies were divided by Roy [42] into two categories *static* and *dynamic* anomalies. *Static* anomalies are related to vessel information that should not change, such as its name, its id given by IMO. *Dynamic* anomalies were divided into two subcategories: *kinematic* and *non-kinematic*. Some anomalies that are categorized as *non-kinematic* are associated with missing or wrong information about the vessel crew, cargo or about its passengers. Whereas *kinematic* anomalies are related to vessel location, speed, course and maneuvers.

2.2.2 Anomaly detection by vessel type

There are several types of vessel, such as cargo, passenger, tanker and many others[2]. When looking for the normal behaviour, we need to compare vessels that belong to the same type since vessels that belong to the same class travel at similar speed [42] and have similar maneuvering behaviour [57]. Large vessels are also obligated to travel in in specific *routes*[3] [55] created by IMO.

However, anomalies are not always a threat, such as *piracy*, illegal fishing and many others [30]. It is of interest to the operators to receive recommendations of vessels that are having some type of anomalous behaviour, which will trigger further investigation on the part of the operator to decide if it is a threat or not [30].

In this thesis we will work only with *kinematic* anomalies, more specifically, we will

[1] `https://www.marinfo.gc.ca/e-nav/docs/ais-index-eng.php`
[2] `https://www.marineinsight.com/guidelines/a-guide-to-types-of-ships/`
[3] `http://www.imo.org/en/OurWork/Safety/Navigation/Pages/ShipsRouteing.aspx/`

look into anomalies related to speed, course, zone and navigability between vessels of same type.

2.3 Trajectory segment and subtrajectory

Differently from most works in the field, we define *segment* as a spatial *region* because we want the user to be able to identify anomalies that may happen more in one are than other, and in potential areas of interest. Then, trips that travel through these *segments* have their AIS data checked against the normal behaviour. An example of segments can be seen in Figure 2.3

Definition 1 (Segment) A *segment* is a 2-dimensional polygon with straight sides $S = (p_1, p_2, \ldots, p_n)$ where each p is a point with a *latitude* and *longitude* in Cartesian plane.

Definition 2 (Subtrajectory) A trajectory is a finite sequence $T = ((x_1, t_1), (x_2, t_2), \ldots, (x_m, t_m))$, where x is a set of <*TripId, Longitude, Latitude, Bearing, Speed, Travel_Distance, Interpolated*> and t_i is the timestamp such that $t_i < t_{i+1}$ *for* $i = 1, \ldots, m-1$. A subtrajectory is a subset of the trajectory T such that it only contains *points* which are inside the boundaries of a *segment*.

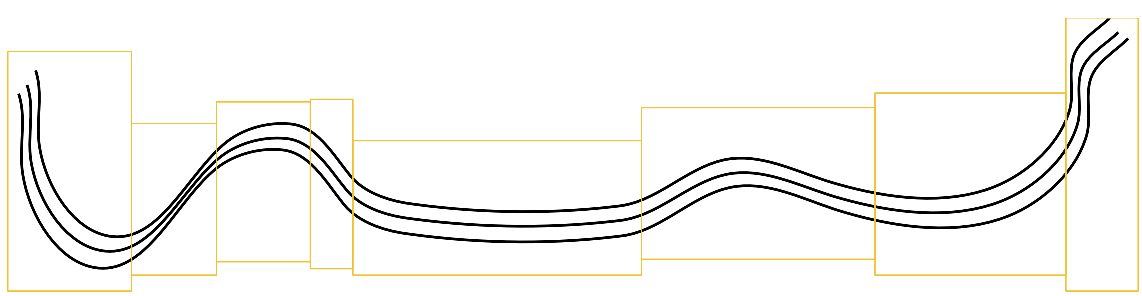

Figure 2.3. Segments represented by yellow rectangles.

2.4 Global and Local anomaly detection

Different works have different definitions of what they consider to be a local anomaly [3, 59]. In this work we consider as *global detection* algorithms that use the whole trajectory to find anomalies, while *local detection* divides trajectories into sub trajectories and find anomalies in those *subtrajectories*. An example can be seen in Figure 2.4, we can see that most trajectories are very similar, but one of the trajectories has a small deviation which could be detected as normal if a model used the whole trajectory.

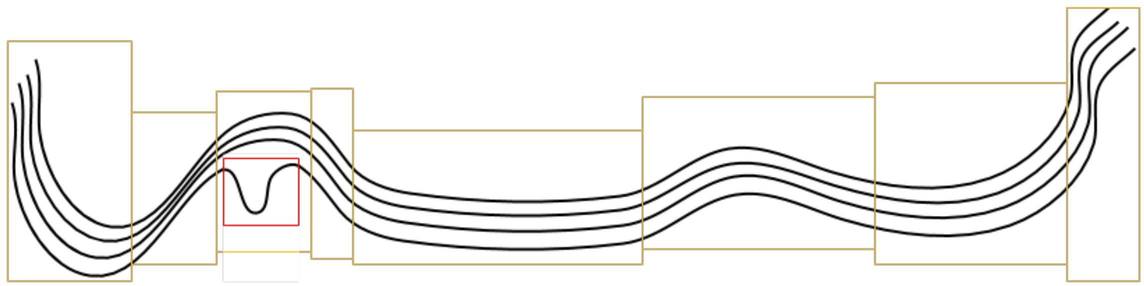

Figure 2.4. A short local anomaly in a long trajectory.

2.5 Visual Analytics

Visual Analytics uses interactive visual interfaces to help the user make decisions in a more efficient and effective way [49] by combining interactivity with automated visual analysis [20]. It is a particularly good solution for problems which cannot be solved by a totally automated tool, nor is it solvable by humans without the cost of a huge cognitive overload. These types of problems are not well-defined; therefore users are not sure they can trust the system output. However, visual analytics uses input from users and allows some degree of exploration which increases the user's reliability on the system [20], the potential of using visual analytics is shown in Figure 2.5. Since finding anomalies is not a well defined problem, and the maritime operators lack trust in fully automated systems [41, 39], using *visual analytics* seems a suitable decision in this domain.

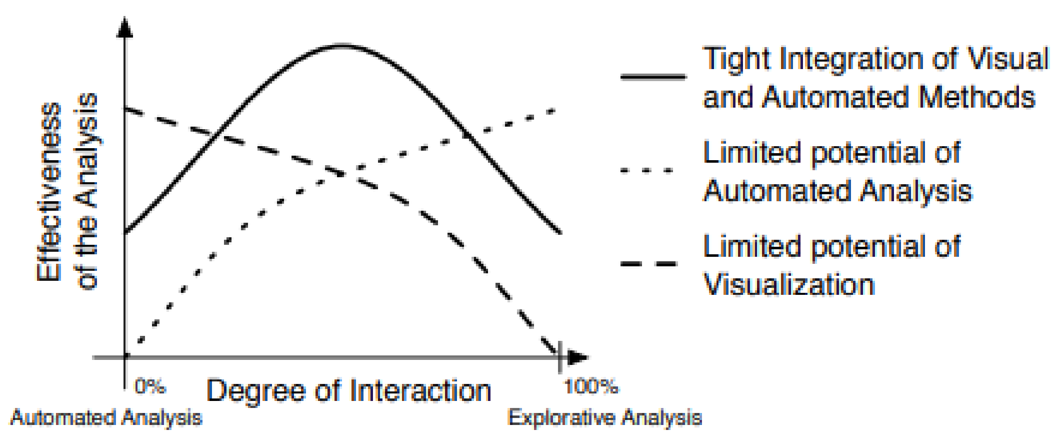

Figure 2.5. Potential of Visual Analytics [20]

Chapter 3

Related Works

3.1 Automated Anomaly Detection of Vessel Trajectories

Since AIS data has been made publicly available many researchers have started working on tools to analyze and detect anomalous vessel behaviours. The vast majority of work done in this field is related to automated detection. The papers discussed in this section can be analyzed into two major aspects as shown in Table 3.1 and are fully described in Section 3.1.1. And, in Section 3.1.2, we evaluate several works from the literature under these aspects.

Aspect	**Values**
Method	data-driven/signature-based/hybrid
Normalcy Extraction	parametric/non-parametric/clustering
Local Anomaly Detection	yes/no
Interpolation Factor	yes/no

Table 3.1. Automated anomaly detection aspects

3.1.1 Analyzed aspects

The first aspect is about the anomaly detection method which can be *signature-based*, *data-driven* or *hybrid*. The *data-driven* approaches use historical data to learn what is the normal behaviour of a trajectory and based on that they classify if a new trajectory is abnormal. *Signature-based* systems use of operators' knowledge of what they consider to be an abnormal behaviour to create rules, e.g. IF speed >25mph THEN *high-speed alert*, and use them to automatically identify anomalies while also handling large quantities of data [51]. Lastly, hybrid approaches combine both types into the same system, usually each focusing on a different type of anomaly.

A second important aspect is the normalcy extraction which can be parametric, non-parametric or clustering. Parametric and non-parametric are statistical methods that can be used to find a probability density function. The parametric method assumes a finite set of parameters for a normal distribution, whereas non-parametric methods don't make such assumptions, they have no bound on a fixed number of parameters, and the distribution can be of any shape. Clustering methods divide the data points into groups based on the similarity between them; one common measure of similarity is distance.

Local Anomaly Detection refers to whether the anomaly detection algorithm used focuses on finding anomalies in subsegments of a trajectory. And Interpolation Factor considers if a proposed method takes interpolation into account when finding anomalies and if it is displayed in any way to the user.

3.1.2 Papers on Automated Anomaly Detection of Vessel Trajectories

In this section we will briefly describe a few of the papers analysed and in One of the works in automated anomaly detection of vessel trajectories was conducted by Pallotta et al. [36]. They proposed a methodology called TREAD which reads AIS data from AIS data streams, then, it uses Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to extract routes. Traffic anomalies are detected by comparing a new route with a group of routes that have the same start and end location. In order to remove outliers from the group of trajectories that will be used to compare against other routes Kernel Density Estimatio (KDE) is used.

Data from AIS was also used to detect anomalous behaviours in vessel trajectories by Mascaro et al.[31].Their work is different from [36] as their solution works with historical data which is cleaned and merged with other sources of data, such as weather data. It clusters trajectories which is a similar approach done in [36], but it uses a different tool called Snob. Then, they use causal discovery via MML(CaMML) to learn Bayesian Networks (BN) from this data.

Trajectory clustering and Bayesian methods are used to classify anomalous behaviour by Zhen et al. [61] which is similar to what [31] does. However, different from [31, 36] it uses k-medoids to cluster vessel trajectories. It then uses a Naive Bayes Classifier to label the routes.

There is a focus on decreasing the error rate when identifying anomalies in vessel trajectories by using Non Conformal Prediction on streaming AIS data in Laxhammar, Rikard and Falkman work [25]. They use kinematic features, such as position and velocity, to classify vessels into a vessel type, such as cargo ship, tanker or passenger ship. In case no known class seems plausible, a vessel is considered anomalous.

A framework was proposed by Yang et al. [59] based on trajectory segmentation and multi-instance learning to identify local outliers. It tests a combination of different segmentation algorithms, representation models, and multi-instance learning. There are four possible segmentation methods Minimum Description Length (MDL), Maximum Acceleration (MA), Minimum Description Length (LRF), and Heterogeneous Curvature Distribution (HDC); the segmentation produced by each of these methods is evaluated based on measuring how different the subtrajectories are from each other and the quantity of segments created in order to avoid over segmentation. The subtrajectories can be represented as either Hidden Markov Model (HMM) or Hierarchical Dirichlet Process Hidden Markov Model (HDP-HMM). And to detect the anomalies either Diverse Density (DD) or Citation kNN can be used; if a subtrajectory is classified as anomalous the whole trajectory is classified as such.

Different to previous approaches, Kazemi et al. [18] propose a system that uses expert knowledge through rules to detect dynamic non kinematic anomalies which are displayed for the user in a map with the vessel trajectory. Similarly, Idiri et al. [15] also use a rule-based approach to identify anomalies, however it differs from the previous work by trying to automatically extract the expert knowledge from history data using a rule learning technique on a database with Maritime accidents, the Marine Accident Investigation Branch (MAIB) database.

3.1.3 Comparative Analysis and Discussion

Most papers analyzed use data-driven methods as can be seen in Table 3.2, only the works done by Kazemi et. al [18] and Idiri et al. [15] use a signature-based approach, although both of them differ on how the expert knowledge is obtained.

When we look at the works which use data-driven methods, clustering the trajectories to extract the normalcy, I believe this is due to the popularity of techniques such as k-means, and more recently DBSCAN which has the advantage of not needing a predefined number of clusters compared to the former. Each of these papers use a different clustering technique: Pallotta et al. [36] use DBSCAN, while Zhen et al. [61] use k-medoids and Mascaro et al. [31] use Snob.

The only paper analyzed which uses non-parametric methods was Laxhammar et al. [25], while Laxhammar et al. [26] and Yang [59] use parametric methods.

One of the papers analyzed focused on identifying local anomalies given our definition (see Section 2.4), Yang et al. [59], which uses trajectory segmentation and considers a local anomaly an anomaly which happens in a sub trajectory.

It is important to point out that there was no paper that takes interpolation into consideration when detecting anomalies. This may be less relevant to tools that work online to certain degree, in case they focus on *point anomalies* [5] and do not use any type of interpolation.

Work	Methods	Normalcy Extraction	Local Anomaly Detection	Interpolation Factor
Pallotta et al. (2013) [36]	data-driven	clustering	no	no
Mascaro et al. (2014) [31]	data-driven	clustering	no	no
Zhen et al. (2017) [61]	data-driven	clustering	no	no
Laxhammar et al. (2010) [25]	data-driven	non-parametric	no	no
Yang et al. (2013) [59]	data-driven	parametric	yes	no
Kazemi et al. (2013) [15]	signature-based	-	no	no
Idiri et al. (2012) [15]	signature-based	-	no	no

Table 3.2. Aspects values for each of the papers that use automated anomaly detection.

Although there are a lot of works that focus on using *data-driven* approaches to find trajectory anomalies in maritime, they may work well for abnormal patterns that have been seen on data before [22]; however, this is based on the assumption that all normal behavior is contained in the dataset used to train the algorithm [40], which is not what happens in reality. Thus these systems may generate a large number of false positives. Another issue with this type of work is that it is hard to codify the knowledge the operators have [41]. Furthermore, a problematic aspect when using this type of approach is that the results are not transparent to the user [1, 37]. In other words, it is difficult for the users to understand the reason why a trajectory was flagged as anomalous, thus decreasing their trust in this type of system [20].

With respect to *signature-based* systems, for them to work correctly, they need all possible scenarios to be thought of beforehand; however, this is not what happens in the real world [37] due to the lack of knowledge from experts and the difficulty in representing all possible scenarios [24].

3.2 Visual Anomaly Detection of Vessel Trajectories

The works here presented can be analyzed and compared in diverse aspects, which can be seen in Table 3. Section 2.2.1. describes the aspects analyzed in the topic of visual anomaly detection of vessel trajectories, while Section 2.2.2 discusses the papers in this field.

Aspect	Values
Domain	maritime/non-maritime/generic
Anomaly scope	global/local
#Attributes used	1/2/3+
Prioritization	yes/no
Interpolation Factor	yes/no

Table 3.3. Visual Anomaly Detection paper aspects.

3.2.1 Analyzed Aspects

The first aspect analyzed in this section is the domain for which the tool was created. In this context, we divide into three possible domains: maritime scenario, non-maritime (e.g., trajectory anomalies on roads), or generic. This category value is given by how the authors classify their solution.

The second aspect is the anomaly scope, which can be global or local. Here we define as a local scope a solution that segments and analyzes bits of a trajectory; this is different from solutions that take into account trajectories as a whole, which may miss local anomalies, while most of the trajectory may be normal.

The third aspect is the number of attributes used; some solutions only use trajectory coordinates to define if a trip is anomalous or normal. Others use the positions and another attribute, like speed. And some solutions use several attributes that can be derived from AIS like bearing, average speed, etc.

The fourth aspect is used to describe if a solution utilizes any form of prioritization of the anomalies found. This is important to give the operator an idea of priority and also certainty, since some trips may be more anomalous than others, and by doing so, it gives the operator the ability to decide which trajectories need a more in-depth investigation.

The last aspect is the same as described in 3.1.1 with the addition that the interpolation can be displayed using some sort of visualization.

3.2.2 Papers on Visual Anomaly Detection of Vessel Trajectories

One of the most cited works in this field is the Visualization of Vessel Movements proposed by Willems et al. [55]. This work uses kernel density estimation (KDE) to show ships area usage, such as sea highways and anchoring zones, and it can be used to identify the most common paths used by vessels. It uses a smaller kernel to display changes in the speed of vessels, and this is used to highlight possible anchoring zones, as can be seen in Figure 3.1. However, this visualization is not interactive and is more focused on area usage rather than finding outliers. The work by Scheepens et al. [43] was based on [55] and extended it to allow users to take multiple attributes when creating the density maps. A density field is created by filtering a subset of the data by selecting a combination of attribute range values; the user also defines weight, radius, and color through a color map for the density field. The user can also select the type of aggregation for the density fields or the image composition. One of the possible aggregates is named D (anomaly), which can be used to find outliers between a density field, which contains normal behavior, and another which the user wants to compare to. In this case, an anomaly would be represented where the density field values are low. An example can be seen in Figure 3.2, which displays the result of applying D anomaly aggregation between a density field with data from 6 days and a density field from only two hours.

Another visualization (see Figure 3.3) was created by Willems et al. [56]. This tool also focuses on understanding the movement of vessels like [55]; more specifically, it aims to detect spatiotemporal patterns by visually testing hypotheses. This is done through the combination of visual analytics with web semantics. This system transforms trajectories into the Simple Event Model (SEM), which can be queried by the visualization, which uses a trajectory contingency table to display how the trajectory changes based on different attributes combinations.

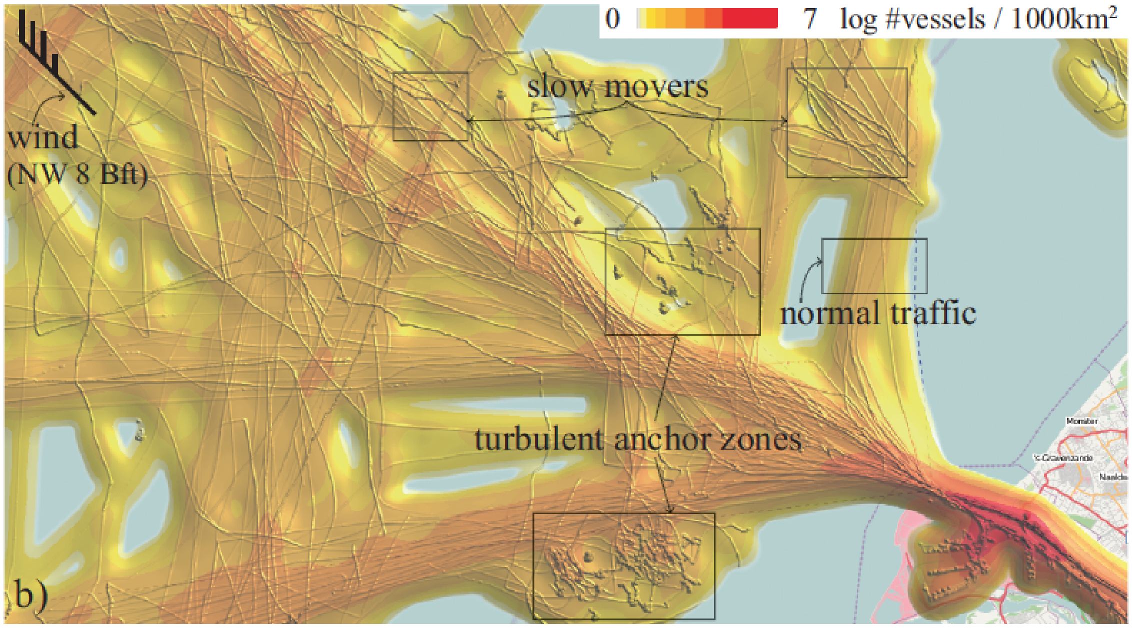

Figure 3.1. Sea lanes and anchoring zones highlighted in [55]

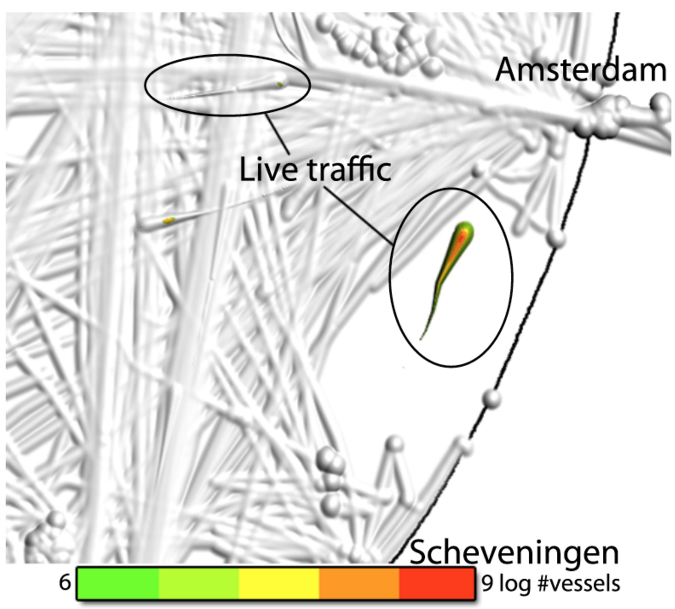

Figure 3.2. Anomaly detected by D anomaly between two density fields in [43].

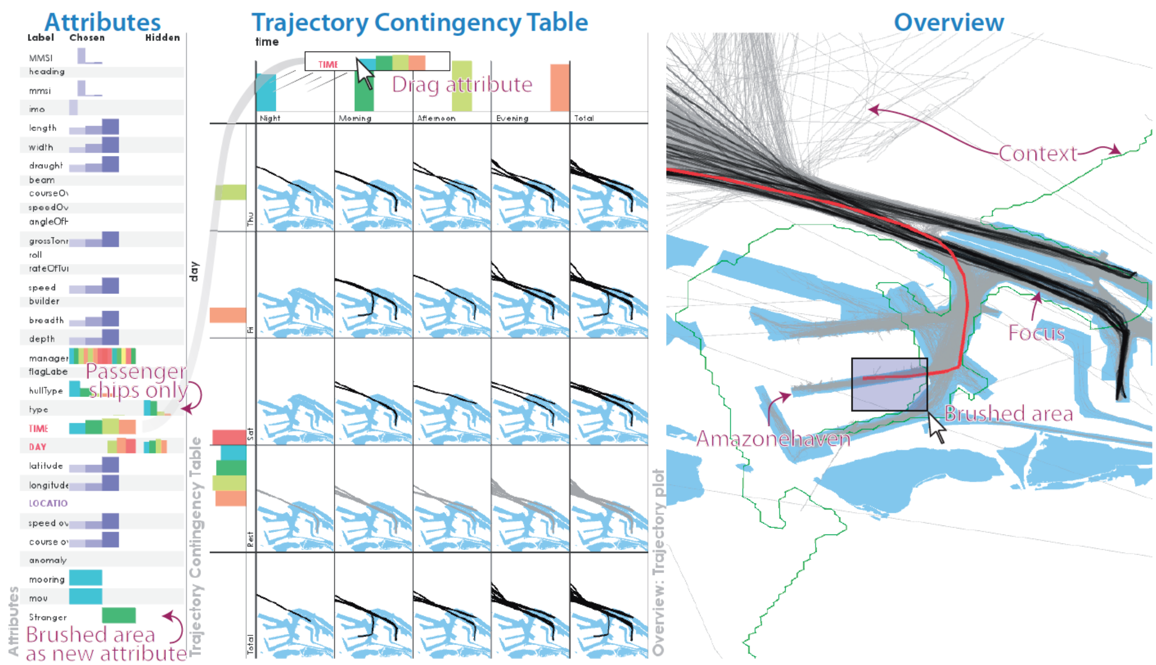

Figure 3.3. Willems et al. [56] visualization tool.

Maritime Visual Analytics Prototype (MVAP) [22] is a prototype created by Defence R&D Canada to allow maritime operators to find anomalies and to analyze vessels of interest (VOI). This prototype contains different widgets and each one has a different purpose. It enables the user to create and analyze a group of vessels that works as the starting point of this tool. In a widget with map and vessel positions, it shows vessels encountered which were automatically found, and by clicking on the vessel, it shows the path the ship traveled against the expected path (see Figure 3.6). This idea of comparing a vessel trajectory with another path is similar to what we want to propose, however here, the "optimal" path in this work is simply a straight line from the origin to destination, while in ours we compute based on all other trajectories of the same group. In another widget, there is a magnet grid that the user can add attributes as magnets, and depending on how high the value for the vessel is, the more it will be attracted to the grid, it can be seen in Figure 3.7. However, during validation, they found that this visualization wouldn't be useful since it lacked the data to make it effective. In contrast, we will precompute all information that will be displayed in our visualization from AIS data, so it won't depend on the user having access to external data sources.

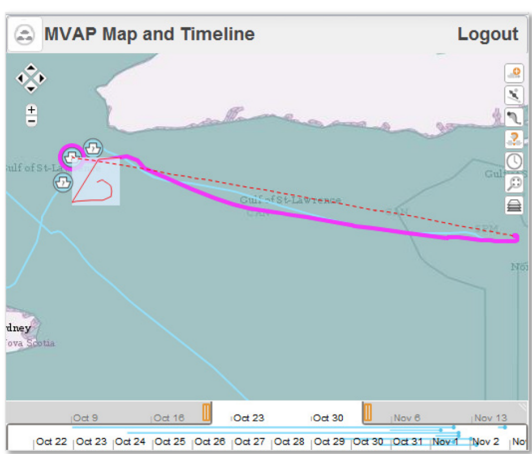

Figure 3.4. Map with Route Ribbons [22].

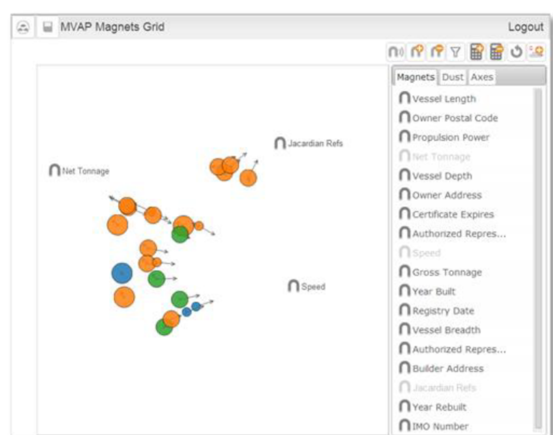

Figure 3.5. Magnet Grid [22].

A tool was created to find anomalous trajectories by Wang et al. [54]. It works by grouping trajectories based on their pairwise distance, and then for each of these clusters, it chooses N equally spatially distributed sample points (see Figure 3.6) and then it classifies as anomalous routes that have points with low probabilistic density, and displays them in a map as seen in Figure 3.7. This approach is somewhat similar to what we propose, which is, instead of comparing and analyzing the whole route, break it into segments that may allow finding local anomalies. However, this work may miss some local anomalies depending on the number of samples chosen, whereas we use all relevant points of a trajectory to calculate the deviation from other trajectories. Furthermore, we also take into account trajectories from different types of ship and also other AIS derived attributes like speed and bearing, while [54] only uses the AIS position to find anomalies.

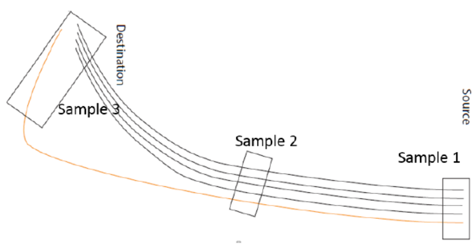

Figure 3.6. Anomaly detection process [54].

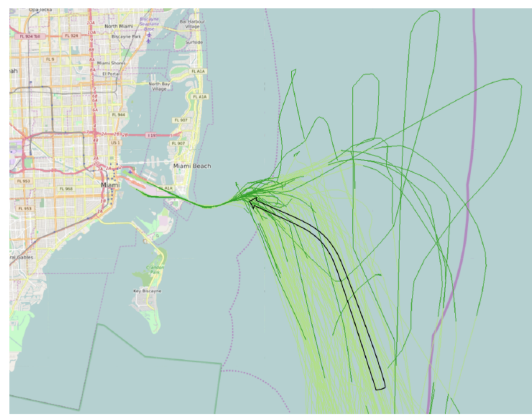

Figure 3.7. Anomalous trajectories highlight [54].

A framework was proposed by Riveiro et al. [41], that uses a hybrid approach between data-driven, signature-based, and visual analytics called VISAD, its graphical interface can be seen in Figure 3.8. It uses Self Organizing Maps with Gaussian Mixture Models to find anomalies in kinematic data and use rules for non-kinematic anomalies. It then highlights anomalous vessels on a map and allows the user to interact and adjust the model by interacting with the mixing proportions of the Self Organizing Maps visually in the case of an incorrect anomaly being detected. However, this work had two problems. First, according to Martineau and Roy [30] the number of false positives created by this framework is too large. Second, according to this paper, the operators from maritime traffic control centers would not be allowed to update the normal models since some changes could decrease the model efficiency. In our solution, we don't use any traditional AI model to classify anomalies because we want the operator to be able interact with the system and change the way the anomaly trajectories are detected.

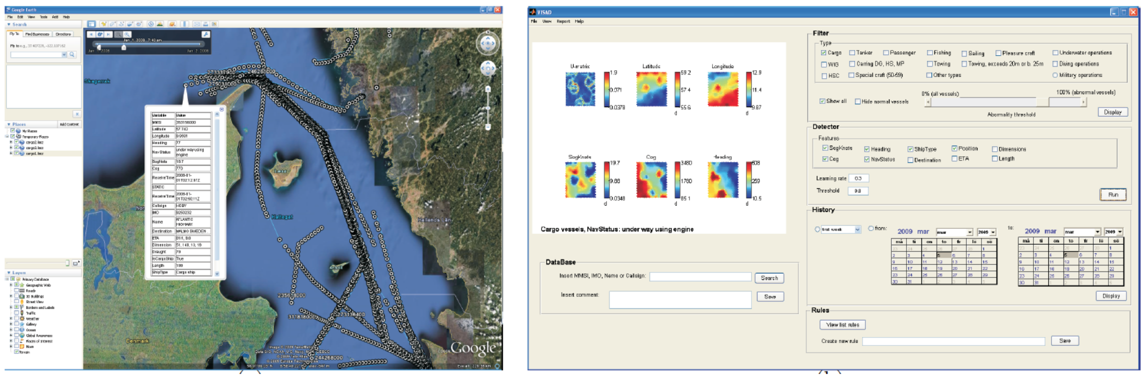

Figure 3.8. VISAD visual interface [41].

There were other papers that, although they do not focus on anomaly detection on the maritime domain, are very important in the visualization field. For example, Tripvista [13] is a visual tool to analyze traffic patterns (see Figure 3.9). One of its visualizations is a parallel coordinate to visualize multiple attributes of multi-dimensional data, which can be very useful to filter certain trajectories based on specific attributes and to quickly identify outliers. TripVista also allows the users to draw a shape that they may want to filter and investigate trajectories with similar shapes. However, it only works because the number of possible shapes is very limited, which is not necessarily the case in the maritime domain.

The goal of the work proposed by Lu et al. [29] aims to understand how the travel duration varies in different road sections at different times of day and on weekends. It works by allowing the user to split a road into several segments, and for each of them, the trajectories are clustered based on travel duration, and overall rank is calculated for each trajectory. It also displays the distribution of travel time for each segment in a box-plot view; this visualization can be seen in Figure 3.10.

Similarly, the work propsed by Tominski et al. [50] also focus on understanding behavior on roads, however instead of splitting values into segments, it uses a 3D wall visualization (see Figure 3.11) to represent the change of attributes of multiple trajectories in a spatial way. It uses a time graph to show these attribute value variations by time. This visualization can be used to identify gradual or abrupt changes in space or time, trends, as well as finding local or global outliers. However, the wall can be hard to visualize paths that do not have the same geometry, and when

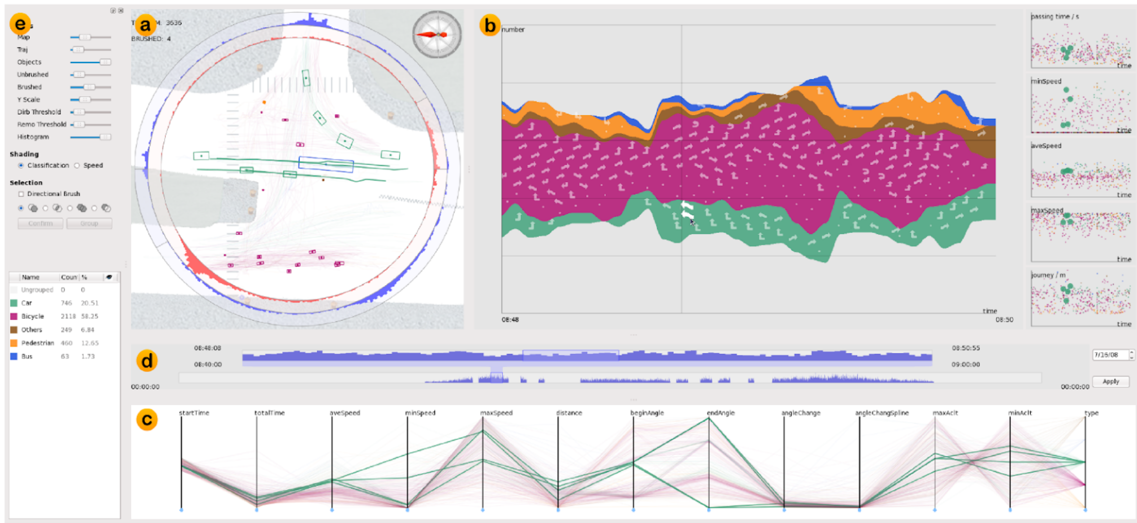

Figure 3.9. TripVista interface [13].

using many different attributes.

3.2.3 Comparative Analysis and Discussion

The aspects of each paper mentioned above is shown in Table 3.4. As can be seen, 6 out of the 9 papers studied work for finding anomalies in the maritime domain. The work done by Tominski et al. [50] is called generic in their own paper; however, it has a limitation that it requires trajectories to have somewhat the same geometry, which can be hard to be used in the maritime domain since there are no constraints such as "roads".

Concerning the anomaly scope, all papers but one allow for finding anomalies in the whole trajectory; and three papers also focus on analyzing and find anomalies on a local level.

Most of the works, 6 out of 9, use, or at least have the ability to use, three or more attributes to find and explore anomalies. This is due to the importance of using multiple attributes to find various possible anomalies rather than just positional anomalies, or speed. Wang et al. [54] is the only work that uses only the vessel coordinates to find anomalies.

From the works analyzed, only Lu et al. [29] use some sort of prioritization, which in their context is used to inform which trajectories are slow compared to others.

As for the interpolation factor aspect, no work has addressed the issue of taking

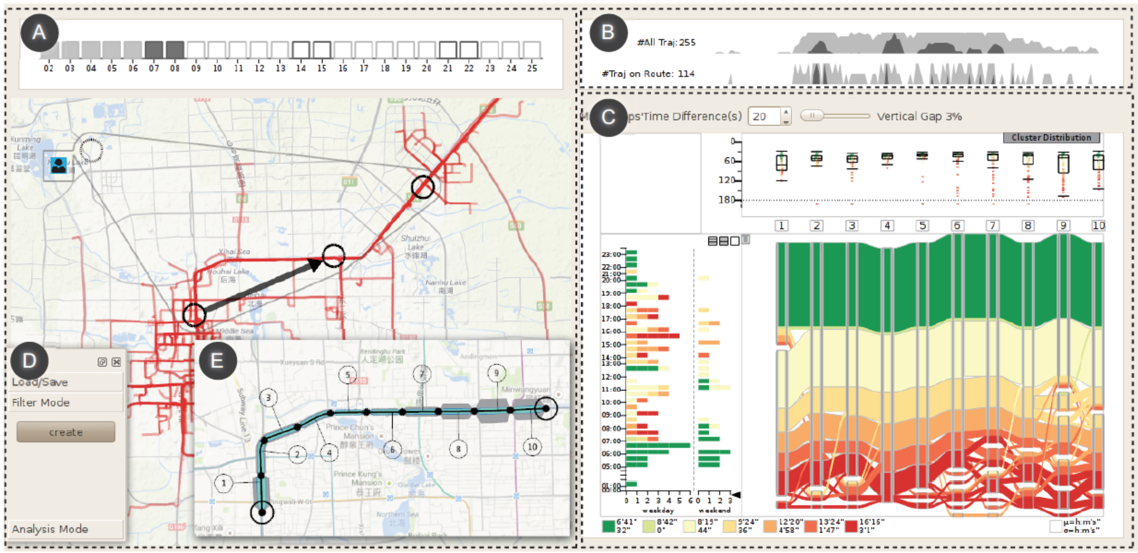

Figure 3.10. TrajRank interface [29].

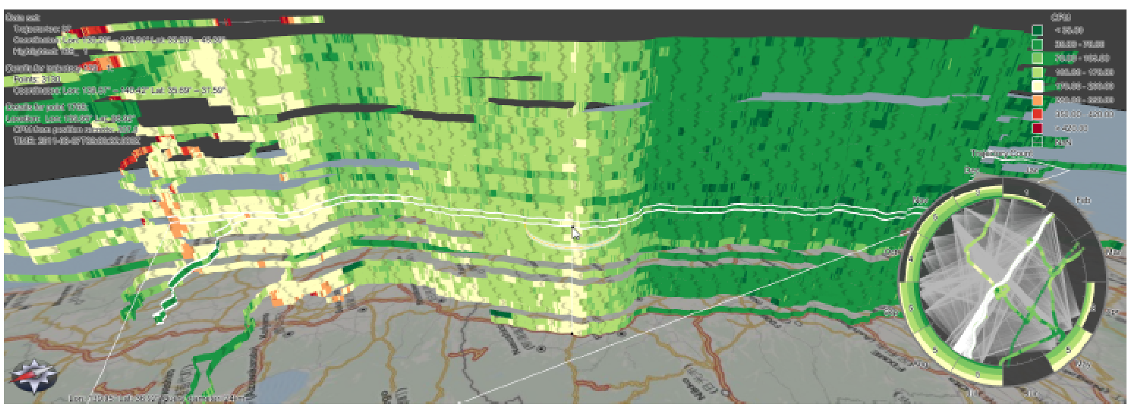

Figure 3.11. Tominski et al. trajectory wall [50].

interpolation into account in any way.

In summary, most of the works analyzed are set in the maritime domain, work on the whole trajectory, allow the usage of three or more attributes to explore trajectories, and don't use any sort of prioritization nor have any sort of visualization for interpolation.

Our work is also focused in the maritime domain, and it uses several AIS derived attributes, like position, speed, bearing, duration, etc, to find anomalies. We believe that using multiple attributes can help the user to get better insight on how trajectories may have deviated from normality. However, we plan to differentiate from the other works by focusing not only on analyzing the trajectory as a whole but also on different segments of a trajectory so that local anomalies may stand out. This is somewhat similar to what is proposed by Wang et al. in [54], but instead of comparing a single point of a trip against other trips, we aggregate all points inside a segment to calculate attribute values, like average speed, and then we give an score based on how this attribute deviate from the mean. We also take ship type into account when comparing trajectories, while [54] only used the AIS position to find anomalies. By using all relevant points of all trajectories which belong to the same vessel type, we will calculate a mean trajectory that will be used to compare against the other trajectories to show the correct path vessels should have used, similar to the one done by [22]; however, there the path is displayed as a simple straight line from the origin to destination, while we compute based on all other trajectories of the same group.

We are also be the only work in the maritime domain that, as far as we know, uses some sort of prioritization on the trajectories based on how anomalous they are, which is based on the work done by Lu et al. [29]. However, we will use multiple attributes to calculate the score, whereas [29] uses only the travel duration. And we allow users to select which attributes they want to use for the score calculation. Furthermore, we are also the only work which aims to help users make sense of the interpolated data in this domain.

Work	Domain	Anomaly scope	#Attributes	Prioritization	Interpolation Factor
Willems et al. (2009) [55]	maritime	global	2	no	no
Scheepens et al. (2011) [43]	maritime	global	3+	no	no
Willems et al. (2010) [56]	maritime	global	3+	no	no
Lavigne (2014) [22]	maritime	global	3+	no	no
Wang et al. (2017) [54]	maritime	local	1	no	no
Riviero et al. (2009) [41]	maritime	global	3+	no	no
Guo et al. (2011) [13]	road	global	3+	no	no
Lu et al. (2015) [29]	road	global and local	2	yes	no
Tominski et al. (2012) [50]	generic	global and local	3+	no	no
This work	maritime	global and local	3+	yes	yes

Table 3.4. Aspects values for each of the papers that use visual anomaly detection

Chapter 4

Methods

In this chapter we will give the requirements that our solution should support and then we will explain how we developed a tool that meet these requirements.

4.1 Requirements

As mentioned previously, this work aims to develop a tool for identifying local anomalies in trip trajectories while also providing user some information about the interpolation such as where and how it happened, and how much interpolation there is on the trajectory. Based on that, we came up with some high-level requirements:

- The tool should support the identification of trips which may have an anomalous behavior.
- The tool should support the identification of local anomalies.
- The tool should improve the user understanding where interpolation has happened in a trajectory and its impact, if any, on anomalies.
- The tool should support some sort of explanation of the cause of the anomaly.

There are some considerations that we need to take into account when developing a tool with the MSOCs personnel in mind. First, the tool should be easy to use and learn due to constant changes in the MSOCs personnel [22].

4.2 Tool Framework Overview

An overview of the framework created for this tool can be seen in Figure 4.1. It is composed of a preprocessing step that combines two sources of AIS data to get *trips* information. Then invalid trips are removed, and the remaining trips go through a cleaning process where invalid data is removed, and gaps are interpolated. We then create spatial segments that serve the purpose of partitioning each trip *trajectory* into *subtrajectories*. The *subtrajectories*' attributes, such as average speed, are given a score based on how much they deviate from the mean over all other trips attribute values; the combined final score for each *subtrajectory* is then displayed in a tabular visualization. Each trip is represented as a line in the table in which the first column may show the maximum or average score for a trip, depending on which option the user has selected, and the other columns show the *subtrajectory* scores, which are represented by a bar length, while the color of the bar shows the amount of interpolation in the *subtrajectory*.

Following the "Visual Information Seeking Mantra" [45], we first display an overview of the overall maritime situation in the table. The users can then use filters to remove uninteresting data, so it shows only trips of interest. They can hover or select an individual row to see the score and interpolation values. By clicking on a row, the trip trajectory will be displayed on the map. The user can then compare the trip trajectory against the mean trajectory to see if there were any deviations and if the interpolation was done correctly. The user can also choose which attributes and segments should be used during the score computation, which will update the *subtrajectory* score.

4.3 Rationale

Why segments?

In order to be able to expose local anomalies in trip trajectories we decided to use spatial segmentation on the trajectories. The reason for this is that it becomes visible for the user *where* the anomalies took place. There is also the potential for the user to define their segmentations which could be a certain area of interest [33] for the operator or, it could be be done automatically using strategies that try to divide a trajectory into multiple meaningful subtrajectories in a unsupervised [48, 11] or semi-supervised [17] way by applying Minimum Description Length (MDL) or Sliding Window Segmentation (SWS) techniques.

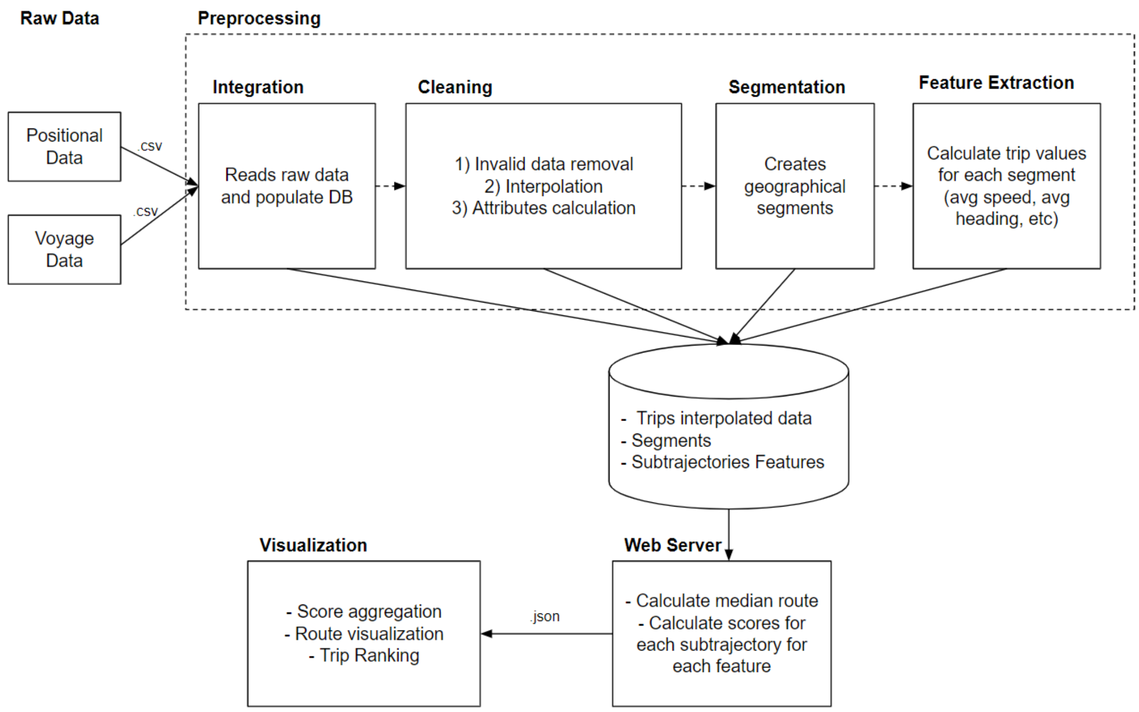

Figure 4.1. Overview of the framework of the Trip Outlier Scoring Tool

Why use mean and score?

Since we will be analyzing trajectories of the same type of vessel that goes in the same direction, the trajectories and attribute values are not likely to be very different. We compute the *z-score*, which gives the number of standard deviations a value is away from the mean, for the trajectory attributes as a distributional measure in respect to the other trips to see how much a trip is deviating from normality, considering the mean represents the normal behavior. Furthermore, by working with scores, it is fast to calculate and update based on weights, as opposed to working with machine learning models, which, as previously stated, cannot be updated on an everyday basis by an operator, reducing their ability to manipulate the output of a tool. Moreover, by using and combining scores, the operator can prioritize the anomalies based on what they find is important. This approach is different from automated approaches which use data mining techniques to simply output a label based on the previous data, and from the rule-based approaches which have a certain value to met to an alert to be fired. In our approach the user can look at just a subset of the vessels and then see for that particular group what look anomalous and what it does not.

Why show a map?

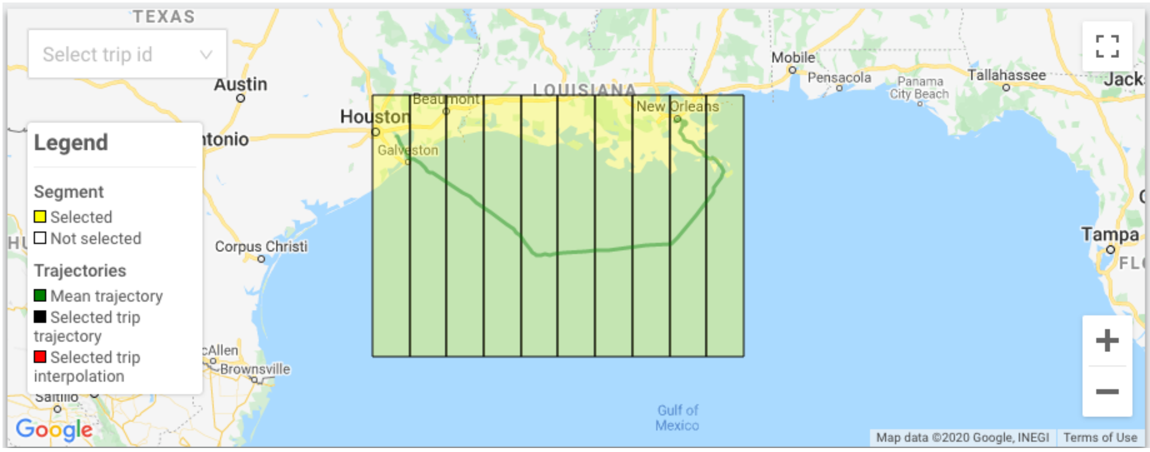

Figure 4.2. Map component.

A map is a crucial component for maritime operators to visualize how a trajectory occurred spatially and temporally. We use the map to display a selected trajectory in which original and interpolated points are plotted and differentiated by color. This way, the operator can have an idea if the interpolation looks correct, and if so, it may indicate that the score in that *subjtractory* is more reliable. We also plot what should be the ideal trajectory, so the user can estimate if a trip trajectory is anomalous. The map also allows the user to visualize where the segments are spatially located.

Why show scores in a table-like visualization?

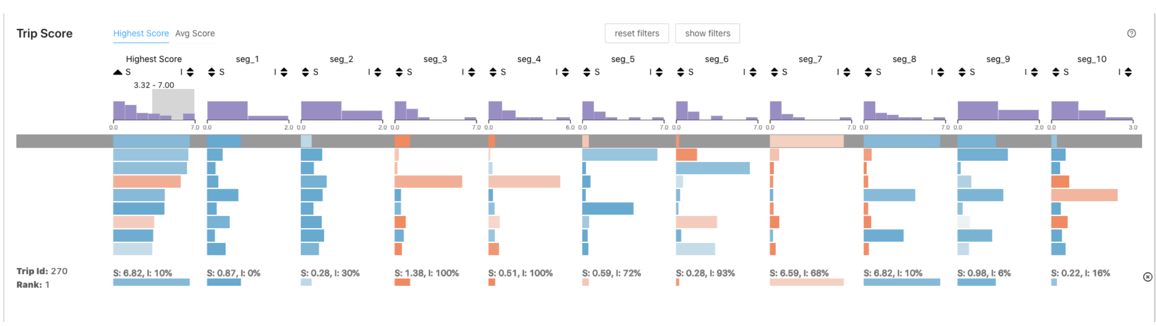

Figure 4.3. Trip Score component showing only trips that had a score above 3.32, ordered by highest score with the first line locked.

We use a tabular visualization based on based on Table lens [38] because it allows us to visualize two attributes for each *"cell"* easily. In our work, we want the user to have an overview of each trip's scores and how reliable they are in terms of the amount of interpolation there is in a *subtrajectory*. Thus, we can easily display these two attributes in our table, using the length of a bar as the score of a *subtrajectory* and the color as the amount of interpolation. Then if the user wants to see information about a single trip, they can just hover or click in the trip row to have the values score and interpolation values highlighted displayed.

4.4 Data source

The dataset used in this work is composed of trips between the ports of Houston and New Orleans between 2009 and 2014. This dataset is composed of two csv files. One contains a combination of static and positional data as described below:

- x,y - longitude and latitude positions
- basedatetime - UTC year-month-day hour:minute:second when data was generated
- mmsi - the vessel unique identifier
- cog - Course Over Ground (COG) in degrees
- sog - Speed Over Ground (SOG) in knots
- heading
- rot - Rate of Turn (ROT)
- voyageid - trip unique identifier
- zone
- year
- month
- new_voyageid - trip unique identifier to be linked with the second file
- med_length, med_width - vessel dimensions

- co_type - vessel type

The second file contains information about the vessel origin and destination and the planned time of arrival:

- x,y - longitude and latitude positions
- basedatetime - UTC year-month-day hour:minute:second when data was generated
- mmsi - the vessel unique identifier
- curr_dest - the name of the port of destination
- curr_eta - Estimated Time of Arrival (ETA)
- prev_eta - ETA of previous destination
- prev_dest - previous port of destination
- new_voyageid - an unique id for each trip in this file

We can see the number of trips made in this period for each type of vessel in this dataset in Table 4.1. The vast majority of the trips were made by cargo and tanker ships.

Vessel Type	Description[1]	Quantity
20	Wing in ground	1
31	Towing	60
32	Towing: length exceeds 200m or breadth exceeds 25m	11
33	Dredging or underwater ops	1
37	Pleasure Craft	1
52	Tug	15
60	Passenger	1
70	Cargo	2344
80	Tanker	702
90	Other Type	14
100	Reserved	39

Table 4.1. Quantity of trips by vessel type.

4.5 Pre-processing

In this section we will explain in detail how we cleaned our data and how we derived important information to be used by our tool.

4.5.1 Integration

The raw csv data is stored in a database, so it is easier and faster to query. We use the field *new_voyageid* to integrate the trip trajectory information with the origin destination. In this work we decided to use PostgresSQL[2] since it works well with spatial queries when the Postgis[3] extension is added. The remainder of the pre-processing stage is divided into processing the trips, creating segments and calculating scores.

4.5.2 Cleaning

Invalid data removal

The dataset used in this work has many issues that needed to be addressed before it could be used properly. As can be seen in Image 4.4, there are trips with positional

[1] `https://coast.noaa.gov/data/marinecadastre/ais/VesselTypeCodes2018.pdf`
[2] `https://www.postgresql.org/`
[3] `https://postgis.net/`

jumps, trips that don't start and end in the correct ports, and there are trips with incorrect AIS information, such as duplicated timestamps.

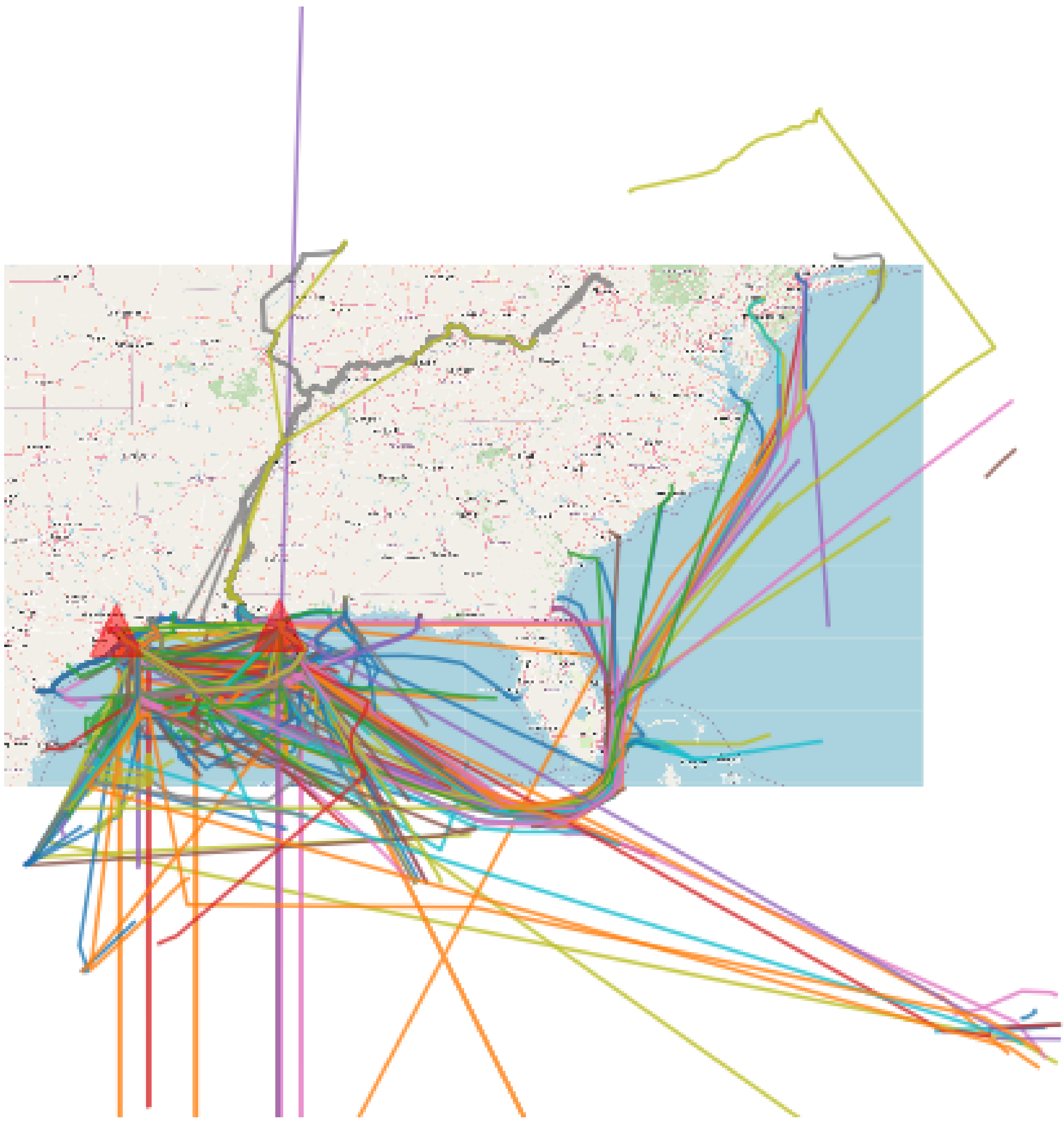

Figure 4.4. Raw AIS data. The two ports are represents by the red triangles.

The first step in this process is removing trips that don't start and end at the correct ports of origin and destination. We calculate the geodesic distance between the first and last points of a *trajectory* $<Longitude, Latitude>$ and the origin and destination ports. If any of those distances is higher than 10 nautical miles, we remove the trip from the dataset.

Then for each trip we look for rows with duplicated timestamp and remove all of them except for the first one. We could try to fix these rows timestamps based on the vessel's initial speed and terminal speed; however, since we will apply kinematic interpolation on the trips in a later stage, we decided to only remove them, the interpolation will be explained in more detail in the Section 4.5.2.

After removing the duplicated rows, we use a Hampel Filter to identify positional jumps. A Hampel Filter works by using a moving window. It then computes the median of this window and the standard deviation, if then the observation deviates from the window median by more than a predefined number of standard deviations, it is considered an outlier. For the two arguments that need to be chosen when using this filter, we selected 10 as the moving window size and 5 as the number of standard deviations, which were chosen empirically since it was simples and showed good results, and the number of standard deviations was set high so that no good points are removed. We could also potentially have tested with artificial outliers so we could see which values produced the best result. We apply this filter to a set made of the latitudes and longitudes of a trip separately, and then we removed all points that were returned as outliers.

Interpolation

As mentioned in the introduction, AIS data is often incomplete, and in addition to that, the *cleaning* step may leave more gaps in the dataset. These gaps may make it difficult for the user to analyze the trajectory, and they may also affect the model's accuracy [14], and in the our tool they would affect the values of the features extracted during the Feature Extraction phase (see Section 4.5.3). Thus, an interpolation process is used to fill gaps between data points that last for 6 minutes or longer.

The technique used to interpolate the *trajectory* data was kinematic interpolation [28], which works well for moving objects, which is the case for AIS trajectory data. Kinematic interpolation works by taking the speed at the last *point* <*Latitude, Longitude, Timestamp, Latitudinal Velocity, Longitudinal Velocity* > before the gap and the first *point* after the gap. It then calculates the acceleration between those two points to create the interpolations, which is modeled as a linear function of time. The velocities are represented as 2D vectors *(vy,vx)*, but we don't convert the latitude and longitude to x and y since the geographical error is small. We chose to generate one interpolate *point* <*Latitude, Longitude, Timestamp* > for every 3 minutes of gap.

Attribute calculation

After the previous step, we have a trajectory composed of original data *points* <*Trip Id, Timestamp, Latitude, Longitude, Heading, SOG, ROT, COG* > and interpolated data points <Trip Id, Timestamp, Latitude, Longitude, Interpolated >. Since we don't have SOG, ROT, COG, and Heading for the interpolated datapoints, we drop those values for the original data points. Then for every point, interpolated and original, we calculate *Speed, Bearing, and Distance Travelled* . To calculate the speed for point *pn* we divide geodesic distance between *pn-1* and *pn* by the time spent travelling between those two points. And for the bearing we use the *Forward Azimuth formula*[4].

4.5.3 Segmentation and Feature Extraction

Given all trajectory points, we use the minimum and maximum <*Latitude, Longitude* > to define a 2D bounding box. Then, we divide this bounding box into 10 segments that are orthogonal to the bounding box's longest side.

Once these segments are saved in the database each trip has its *trajectory* divided into *subtrajectories*, more specifically one *subtrajectory* for each *segment*. We then compute the features that will be used to compare against the normal behavior. For each *subtrajectory* we calculate:

- Minimum, average and maximum speed in knots
- Average heading in degrees
- Distance travelled in nautical miles
- Time travelled in seconds
- Interpolation percentage

[4]`https://www.movable-type.co.uk/scripts/latlong.html`

The reason why these attributes were chosen is that one of the kinematic anomalies that it is of interest to maritime operators is the vessel speed compared to the ship class [42]. We then use average speed to give a general idea of how fast a vessel traveled in an ocean section. We use maximum and minimum speed to get possible deviations that the average speed could not show. We use average heading to get maneuverability deviations [42] and deviations from normal routes without the need to plot all trajectories in the map, which could be very cluttered. It would be better to calculate the distance between a *subtrajectory* and the correct path for deviations from the normal route. Distance and time traveled are two pieces of information that are easy to compare between trajectories and may raise questions on why a trajectory took much longer than others. Finally, the interpolation of a subtrajectory will be used to indicate how many points of that trajectory are interpolated. In the future, it could be interesting to add the stopped duration, if there was any, to see if there were some vessels at anchor, and to add proximity between ships, which could indicate a *rendevouz*.

4.6 Backend

Our backend was created to serve the resources need by our frontend such as trip trajectories and trip scores. It was designed following REST architectural style, it was built using Python with Flask[5], we uses Psycopg library to communicate with PostgreSQL. Requests responses content are in JSON format.

[5] `https://flask.palletsprojects.com/en/1.1.x/`

4.7 Trip Outlier Scoring Tool (TOST)

Our tool has three main main components: the Score computation (A), a map (B), and Trip Score table (C), as shown in Figure 4.5.

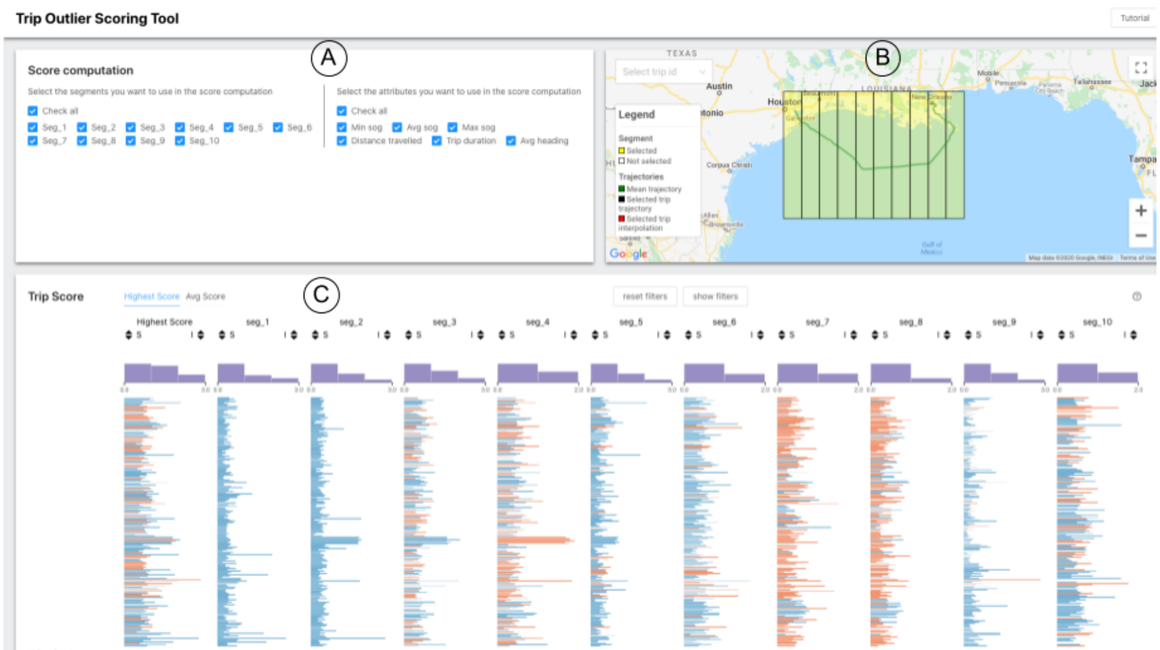

Figure 4.5. Overview of the Trip Outlier Scoring Tool (TOST). The user uses the Score computation component (A) to control which segments and attributes will be used in the score. The trip scores are visualizes in the Trip Score component (C) where the user can filter and sort the data, and select a trip trajectory to be displayed in the map (B).

4.7.1 Score computation

After the features have been computed for each *subtrajectory*, once the backend receives a request for trips' score, it calculates the *z-score* for each *subtrajectory* attribute. Then, on the frontend, for each *subtrajectory*, it averages the absolute values of the *z-scores*, which only use the attributes the user has selected, as seen in Figure 4.6. As an aggregate final score for each *trip*, we may show the highest score, which is the highest value amongst all trip subtrajectories, or it can show the average score of the trip subtrajectories.

Definition 3 (Subtrajectory Score) Given a set of *subtrajectectories* $ST = \{ st_1, {}_2, \ldots, st_n \}$ defined by a spatial segment S for a set of trips $T = \{ t_1, t_2, \ldots, t_n \}$, and the set of the *subtrajectory* attributes A = { a_1, a_2, ..., a_m }. Given that the set of values for attribute $a_k \in A$ can be represented as AV_k = { av_{1k} , av_{2k} , ..., av_{nk} }. The score of a *subtrajectory* st_i can be described as:

$$score(st_i) = \frac{1}{m}\sum_{j=1}^{m}|zscore(av_{ij}, AV_j)|$$

where *zscore* is a function which returns the z-score of a attribute value given a set a values.

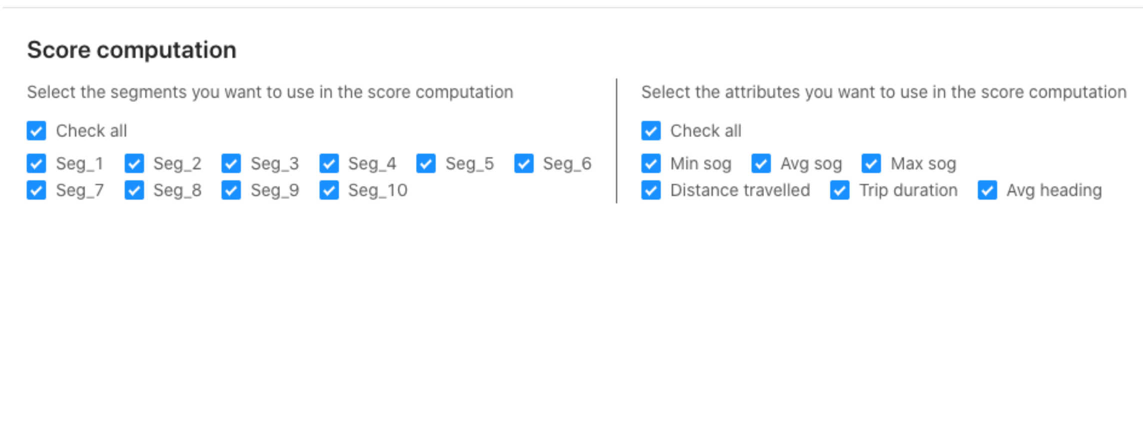

Figure 4.6. Score computation view.

4.7.2 Onboarding

An onboarding tutorial is provided in the system due to the fast rotation of maritime operators, as previously stated. It teaches the tool's main concepts while highlighting the components it refers to, as shown in Figure 4.7. The tutorial is always accessible through a button at the top right corner of the tool. The steps can be easily skipped so the users can go check only the information they need. The tutorial was built using the React Joyride library[6], which allows us to easily add new steps when news features are added to the system.

[6]`https://react-joyride.com/`

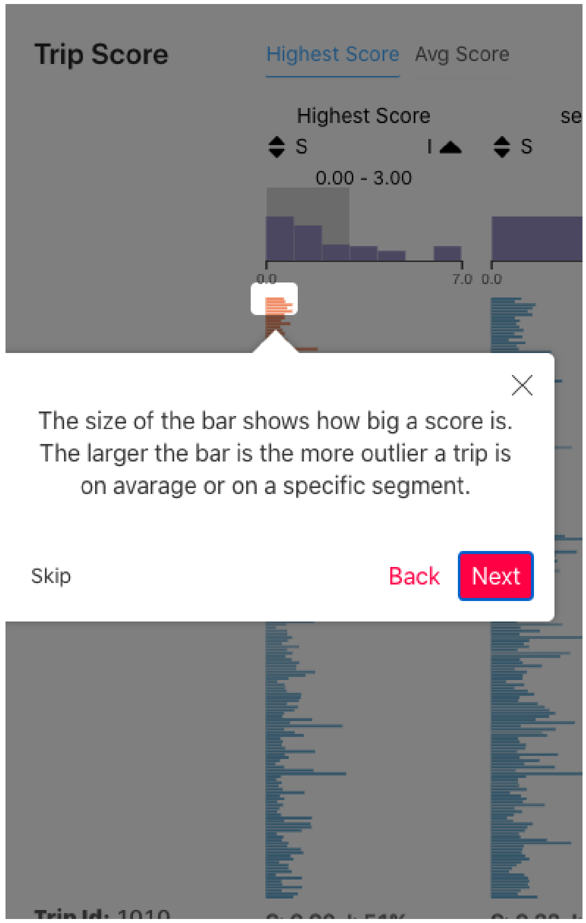

Figure 4.7. Example of a tutorial step.

4.7.3 Map

The map was created to display the previously created segments as well as trip trajectories. It is displayed with a zoom on the region containing the two ports, as seen in Figure 4.2, and the user is free to zoom-in or zoom-out the map. In the center of the map, the segments are displayed as polygons with a black border and a semi-transparent background so that the map underneath is still visible; the background color will only be displayed if that segment is being used in the score computation, otherwise it will have no background color. The user can hover a segment to see its name. A green trajectory is displayed on the map to show the normal path a vessel should do when traveling between those two ports.

On the top left of the map, the user has a select input where a trip id can be selected or input, and the corresponding trajectory will be displayed. Since we want the user to be able to differentiate the original *points* and from the ones that were created after the interpolation, we distinguish them by color. The black portion of the trajectory was created from the original data points, while the red portion was interpolated.

The map was created using google maps api[7] and the React Google Maps (react-google-maps) library[8] which works as a wrapper to the google maps api for React.

Mean trajectory

In order to calculate the mean trajectory, we use a function of the tool created by Erland et al. [9], which we pass as part of the arguments <*Latitude, Longitude, Distance Travelled in Percentage* > and 200 as the number of points to be used to create the mean trajectory. Then, for each of the trajectories it takes 200 equidistant points to compute the average x, y. It is worth mentioning that although averaging points is a simple solution, the points generated may not represent the reality and in some cases since it may generate points in impossible locations, for example, where there is the option to go around an island through the left or the right, the average point may end up being on top of the island. Thus, a better approach would be to use a medoid trajectory in the future. The issue then would be deciding on the best method to calculate the distance between trajectories, which is well analyzed in the work done by [58]. And an interesting solution was developed by Kevin et al. [4], it combines segments of different trajectories to create a representing trajectory. Due to time constraints, we chose to use solution that was already available.

4.7.4 Score Table

The Score Table, displayed in Figure 4.3, was based on Table lens [38] work, a zoomed version of part of this table can be seen in Figure 4.8. Each line in this table represents a trip, and for each column, there is a bar in which its length represents the *subtrajectory* aggregated score and the color represents the percentage of interpolated points. This table has a fixed height so the user can look at the table and have a general idea of all scores without scrolling. So the bar's height are dynamic; they

[7] `https://developers.google.com/maps/documentation/javascript/overview`
[8] `https://tomchentw.github.io/react-google-maps/`

change based on how many trips are being displayed at a given time.

A longer bar may indicate a higher deviation from normality since our score is derived from the z-score. Longer bars also stand out in comparison to smaller bars. And the interpolation is displayed as a gradient from blue when there is 0 interpolation to white when there is 50% interpolation and then to red when there is 100 percent interpolation. We use the Data-Driven Documents (D3)[9] scaleLinear function to get the correct color given the percentage of interpolation in a subtrajectory.

In addition to the columns for each segment, we added a column that shows the trip's highest score or the average score depending on which option the user chose to see. This was created with the intent to help the user to find trips that may have a good score overall, but had a bad score in a specific segment more easily.

The exact scores and interpolation values for a trip, as well as the trip id, can be seen at the bottom of a table when a user hovers over a row with the mouse. It displays the trip id, its rank on the left, and then for the additional column and all other ones it shows the score and interpolation values. The initial idea was that on hover, the row would increase its height to show the bars' score, but due to performance issues when many lines are displayed, we opted to show it at the bottom. The user can also click in a row to lock it, so the values don't change when moving the mouse around. Clicking on a row also displays the trip trajectory on the map.

At the top of the table, we have a purple bar, which shows vessels' distribution by score. The bar height represents the number of vessels in log scale so that scores intervals with a lower number of vessels are still visible to the user, and each bar represents a score in intervals of one. This visualization has two purposes: first, the user can brush the region to filter out uninteresting vessels, and so decreasing the number of vessels displayed at the table which could improve the table visibility. Second, since each segment is a spatial region, showing the distribution may reveal regions with a higher number of outliers than others or a region where the outliers have a much higher score. For example, there could be a region where vessels speed much more than others. At the bottom of the bar, we display an axis with the minimum and maximum score for those segments, which help the user have an idea of the bars score.

[9] `https://d3js.org/`

The user can also change the order in which the trips are displayed by clicking in the sort icon in one of the columns, which will sort the trip by score and change the rank that is displayed when the user hovers a row. This was created so that the user is able to find the top outlier trips in specific segments without changing the overall score.

This table was created initially using <table >, <tr >, <td >html tags; however due to high number of trips being displayed the table became very unresponsive after 200 trips. For this reason, we changed the implementation to be built entirely using Data-Driven Documents (D3), which works more efficiently when rendering large amounts of data.

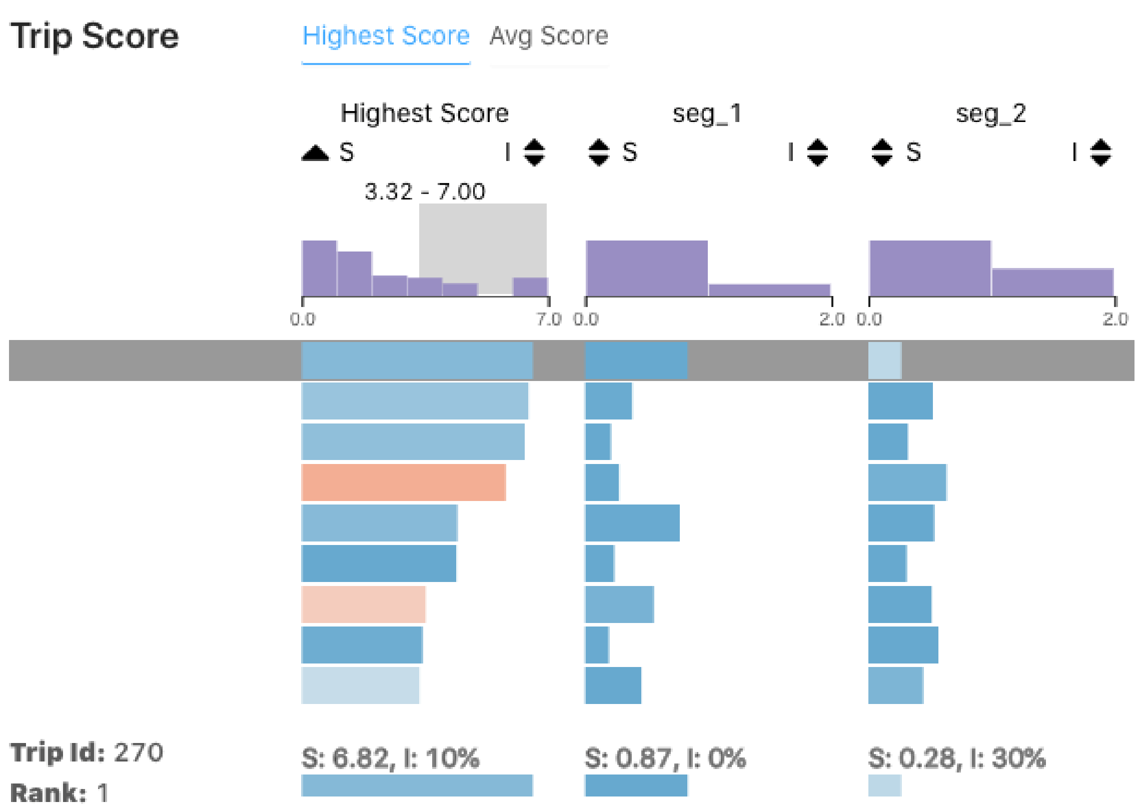

Figure 4.8. Zoom on part of the Figure 4.3

.

4.8 Use case

In this use case will be analysing all trips made by cargo ships that travelled from Houston to New Orleans that we have in our dataset. We can use the filter to display only trips that have score above 3 in any *subtrajectory* using the filter on the Highest Score column which leads us to 31 potential anomalous trips, as shown in Figure 4.9. After sorting by highest score, we can lock the first row to see the top outlier trip, which is the trip with id 2276. At a glance we can see that this trip had overall good scores in all segments except in segments 7 and 8. We can then look at the bottom of the table to see that this trip had a score 9.40 on segment 7 and 6.90 in segments 8, both with a very few interpolated data points.

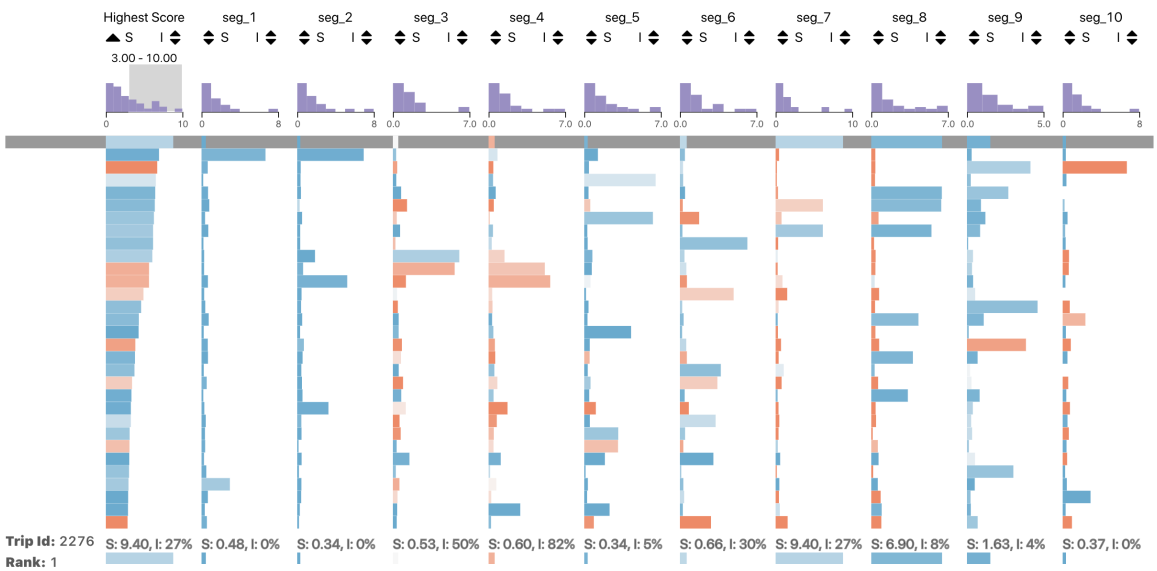

Figure 4.9. Part of Score Table displaying trips with score above 3 with the first line locked

.

Now, when we analyse trip 1102, which ranks 13 on the trips with highest score, as shown in Figure 4.10, we can see that it had a very bad score on segment 6, however, the color indicates there was a lot of interpolated datapoints in this *subtrajectory*, which may indicate that this score is not reliable. By looking at the bottom of the table we can see that 69% of the datapoints have been artificially created, and by looking at the map we can see that the trajectory created does not seem reasonable as seen in Figure 4.11.

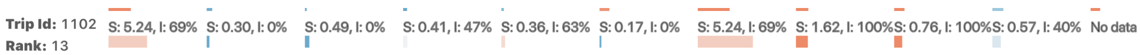

Figure 4.10. Trip 1102 scores

.

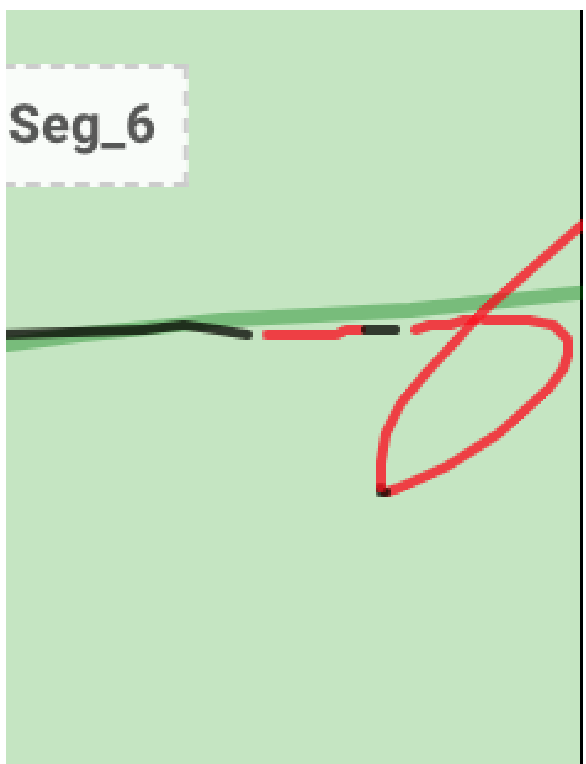

Figure 4.11. Trip 1102 trajectory on segment 6

.

Chapter 5

Evaluation

In order to evaluate the software usability and possible improvements, we conducted a user study. The study was done individually with each participant, and it was conducted online due to in-person restrictions. During the study, the participants received a short tutorial on how to use the tool. Then they had to interact with the TOST to answer a few scenario-based questions. After that, they had to complete a small demographic questionnaire and answer a few closed and open-ended questions about the tool. The whole session took between 45 and 60 minutes.

5.1 Participant Selection

Maritime operators would be the ideal users to test this tool, due to time constraints on our part, we opted to invite computer science students from Dalhousie University. The decision to invite only computer science students is that students in this field usually have some knowledge working with computers and some familiarity with statistics, which can help them to better understand what the *subtrajectory* score (see Section 4.7.1) represents. We then sent an open invitation by email to two mailing lists that all Computer Science students are subscribed by default. Then we picked the first 10 potential participants that replied to our email. Most of them were undergrad students, while one student was doing Masters and another doing Ph.D.

Half of the users had no familiarity with data, and only 3 felt that they were somewhat familiar as can be seen in Figure 5.1.

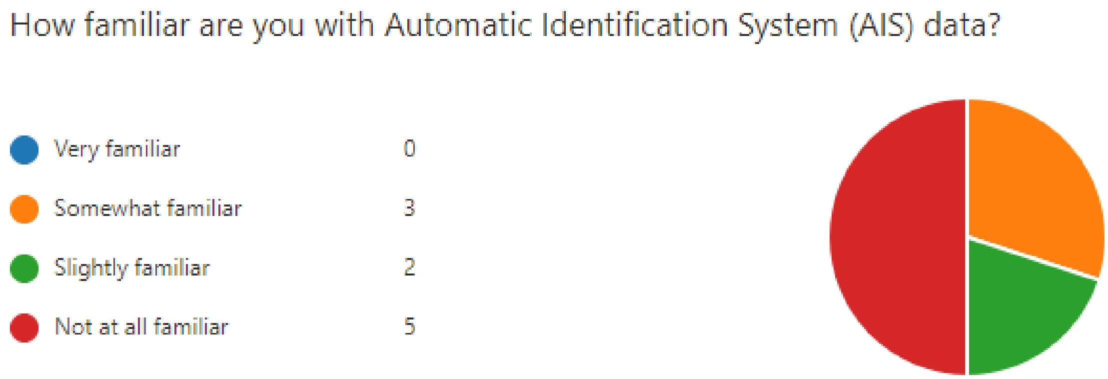

Figure 5.1. Users familiarity with AIS data.

5.2 Experiment Setup

For this experiment a picker component was added to the tool to allow users to select specific scenarios as requested during the study. We also added an option to sort trips by the amount of interpolation.

We recruited participants by sending a recruitment email to the email list *csjobs@cs.dal.ca.*. The fist participants that responded our email were sent the consent form so they could read it and then decide if they still want to participate. During the study the participants also had access to the consent form through a link where they were given time to read before consenting in participating in the study, the consent form can be seen in Appendix A.

The meeting with the participant was conducted online through Microsoft Teams, and the participant had access to the web tool through a link that was shared with them.

5.3 Training

The training was given to each participant on the day of the study to teach them essential concepts about the tool and how the it works, so that no previous knowledge about the maritime domain or about AIS was required. During the training, I shared my screen and used the previously created tutorial to highlight the explained component. After that, I showed a use case of tool based on different data from what they would be using during the study. In the use case I showed users how to use the filtering to display only potential outlier trips, and how to sort based on the score, how to visualize a trip scores and interpolation information, how to find in which segment a trip had an outlier behaviour, how to see which segments had more outliers than others, and how to display trip trajectory on the map. The whole tutorial took about 5 minutes.

5.4 Scenario exercises

Before the experiment, the tool had been slightly modified to display a dropdown component containing different scenario options for the participant to chose from. When a scenario was selected, the data displayed to the user changed; this was done so that we could make the same question for different data in order to evaluate whether the user was able to use the tool in different settings, an example of scenarios is shown in Figures 5.2 and 5.3

The participants then received an online questionnaire that was divided into sections. At the beginning of each section, the participants were instructed to select a specific scenario and then answer a few questions that require the operator to use the tool.

The scenarios were all presented in the same order to the participants, which could have introduced some ordering effects on your data.

For the whole exercise, we defined that any trip with a subtrajectory score above 3 should be considered an outlier, except for questions 19, 20 and 21 where the users needed to take the interpolation into account. It is worth mentioning that throughout the whole study, we used the term outlier instead of anomaly since it is a common term in statistics.

5.4.1 Exercise rationale

- How many trips are outliers? - we want to validate if the participants can identify which trips are outliers. They will have to filter the data either by brushing or typing it directly on the filter component. Since asking for several id's can be time-consuming and prone to errors, we ask the number of trips that are outliers.

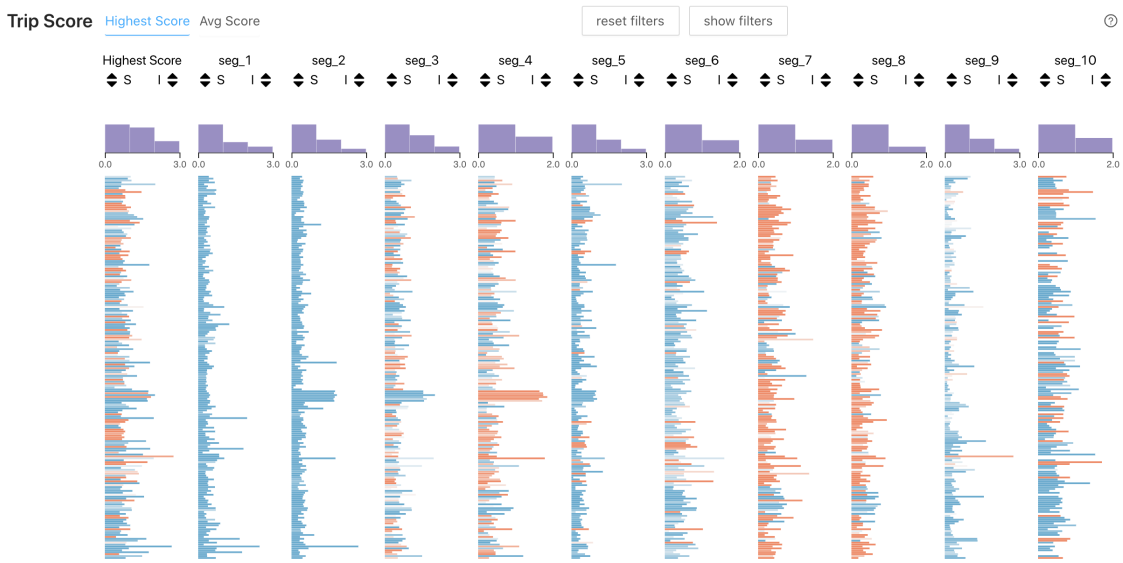

Figure 5.2. Scenario 1 data displayed on Score Table

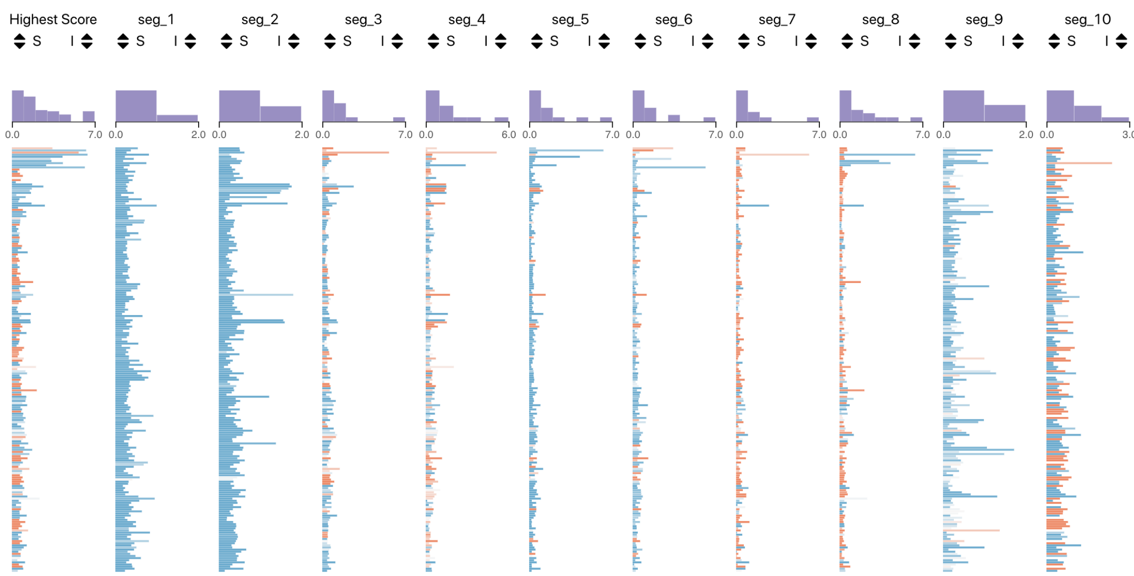

Figure 5.3. Scenario 2 data displayed on Score Table.

- What is the Id of the trip with the highest score? - this question tries to see if the participants understood how to sort trips and the ranking concept.
- Which segments have more outliers than others? - this question tries to check if participants can make use of the score distribution to visualize segments with more anomalies
- In which segments did trip X have an outlier behavior? - in this question

we wanted to see if the participants understood the score concept and how to visualize it, which can be either be by hovering over a row and seeing the score at the bottom of the table or by looking at the axis at the top of the table.

- Ideally, we would like to have the dataset with very few interpolations. Based on this information, and without using any type of sorting, how much interpolation do you think there is in this dataset? - this question tries to access if using color to interpret the interpolation gives an overall idea of the amount of interpolation used in the dataset.

- How many trips have, on AVERAGE, ABOVE 50% interpolation? - this question tries to check if the participant understood how the interpolation concept is displayed.

- Given trip X, choose the most appropriate option - In this question, we put together the concepts of score, interpolation and trajectory together. The user then has to choose one of the following options:

  – It is not an outlier, it has a good score and good interpolation

  – It is an outlier, it has bad score and bad interpolation

  – I can't say, there is too much interpolation, or the interpolation seems incorrect

5.4.2 Results

Scenario	Question	Correct Answer	Percentage of correct responses
1	1) How many trips are outliers?	0	100%
	2) What is the Id of the trip with the highest score?	542	80%
	3) Which segments have more outliers than others?	None	50%
2	4) How many trips are outliers?	10	70%
	5) What is the Id of the trip with the highest score?	270	90%
	6) Which segments have more outliers than others?	3;4;5;6;7;8	30%
3	7) How many trips are outliers?	25	60%
	8) What is the Id of the trip with the highest score?	2276	90%
4	10) In which segments the trip 1006 had an outlier behaviour?	4	70%
	11) In which segments the trip 1059 had an outlier behaviour?	6	80%
	12) In which segments the trip 1079 had an outlier behaviour?	9	80%
5	13) How much interpolation do you think there is in this dataset	-	-
	14) How many trips have, on average, above 50% interpolation?	14	80%
6	15) How much interpolation do you think there is in this dataset	-	-
	16) How many trips have, on average, above 50% interpolation?	21	70%
7	17) How much interpolation do you think there is in this dataset	-	-
	18) How many trips have, on average, above 50% interpolation?	32	70%
8	19) Given the trip 2276 choose the most appropriate option	It is an outlier	20%
	20) Given the trip 1963 choose the most appropriate option	It is not an outlier	100%
	21) Given the trip 3062 choose the most appropriate option	I can't say	80%

Table 5.1. Scenario exercises responses

We show a summary of how many participants got each question correct in Table 5.1. We can see the participants had no issues in identifying when there were no outliers in the dataset; however, as the number of the outliers increased, the number of correct answers decreased, and the answers were more diverse, as we can see in Figures 5.4 and 5.5. A possible reason may be that the users did not understand how to use the filter properly, or in which columns they should apply the filter to; it is hard to explain why some users chose 0 or 1 as the number of outliers in question 7.

Figure 5.4. Number of responses to the available options for question 4: "How many trips are outliers?".

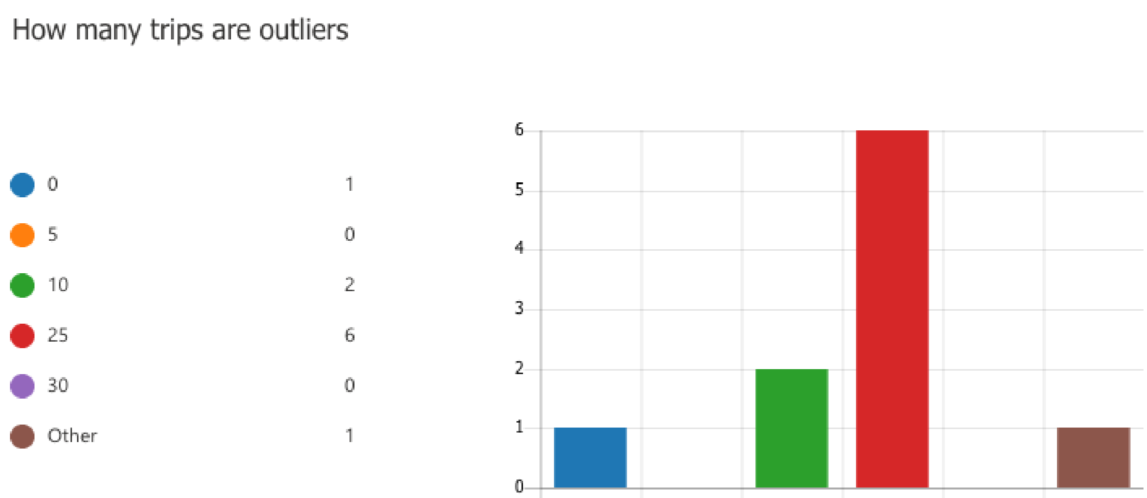

Figure 5.5. Number of responses to the available options for question 7: "How many trips are outliers?".

We can see that most users were able to properly sort by score and select the trips that had the highest outlier score when we see the results of questions 2, 7 and 8. However, questions 3 and 6 did not have a good result: only 5 percent and 30 percent of the participants chose all the correct options, and we can see the responses in more detail for these questions in Figure 5.6 and 5.7. In question 6 we can see that although the number of total correct responses were low, the most the segments that were chosen by the participants were correct, except for 2 participants that thought no segment was more outlier than others. Even though all participants correctly answered question 1, a possible reason for them to have selected some segments as having more outliers than others could be that in some segments the score was higher than in others, but as seen in Figure 5.6, some participants had chosen segments 7 and 8 as having more outliers even though there was no *subtrajectory* with a score above 2. This could be the result of the question not being well formulated, or the participants didn't understand this functionality. A possible revision to this study design would include a follow-up discussion to provide some explanation for this behaviour.

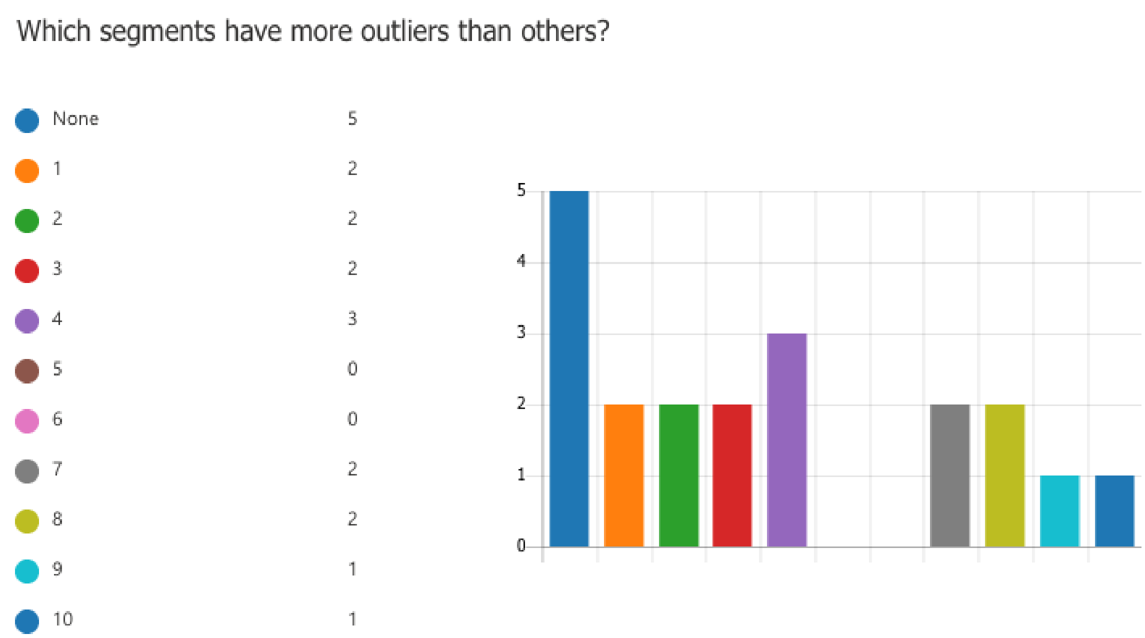

Figure 5.6. Number of responses to the available options for question 3: "Which segments have more outliers than others ?".

Most of the users were also able to correctly answer questions 10, 11, and 12, which shows that they were able to identify which subtrajectories contributed to the trip to be considered an outlier. This means that they correctly understood how a bigger score or larger bar correlates to a trip being more of an outlier, and that they were able to either correctly use the bar width to get this information or they were able to hover over a row and check for the score at the bottom of the table.

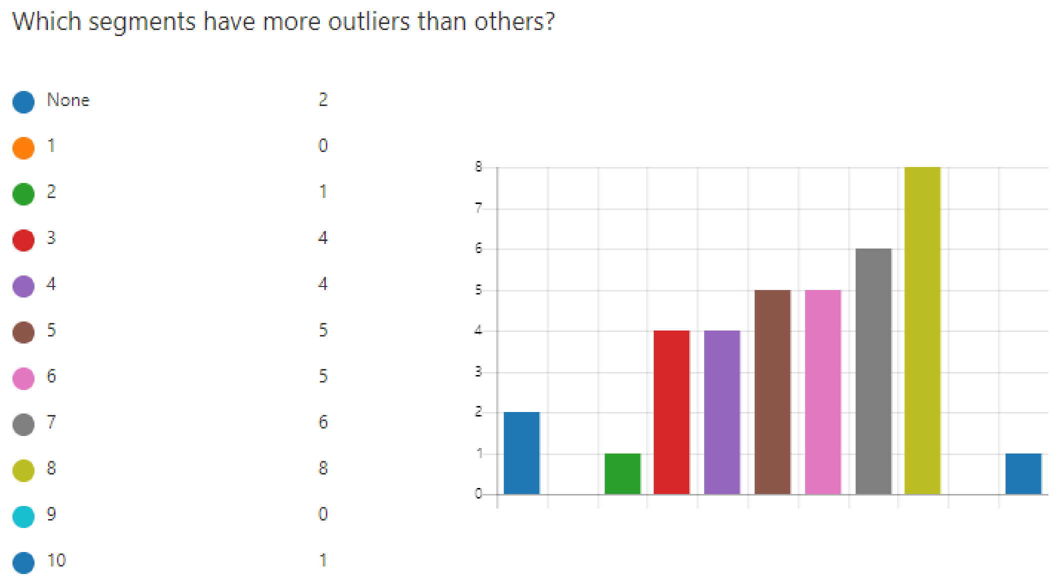

Figure 5.7. Number of responses to the available options for question 6: "Which segments have more outliers than others ?".

Questions 13, 15, and 17 don't necessarily have a correct answer, we wanted to understand how the users feel when they see the bar colors representing the interpolation, and we expected that none of them would choose the option "There seems to be almost no interpolation" which is only selected by one participant in all three questions as seen in Figure 5.8. Most of the time, participants felt that the amount of interpolation was reasonable, which is understandable, although there were too many gaps in this dataset. However, when asked about the number of trips that had interpolation above 50 per cent in questions 14, 16 and 18, most participants got it correct.

For us, the most important questions were 19, 20, and 21 since they put together essential concepts used in this tool. And most of the users correctly identified that trip 1963, in question 20, was not an outlier, and most of them understood that the interpolation affected the bad score of trip 3062 in question 21. However, most of them incorrectly said that trip 2276 was not an outlier, and this could be because the correct answer had a typo: "It is an outlier, it has bad score and bad interpolation" should be "It is an outlier, it has bad score and good interpolation".

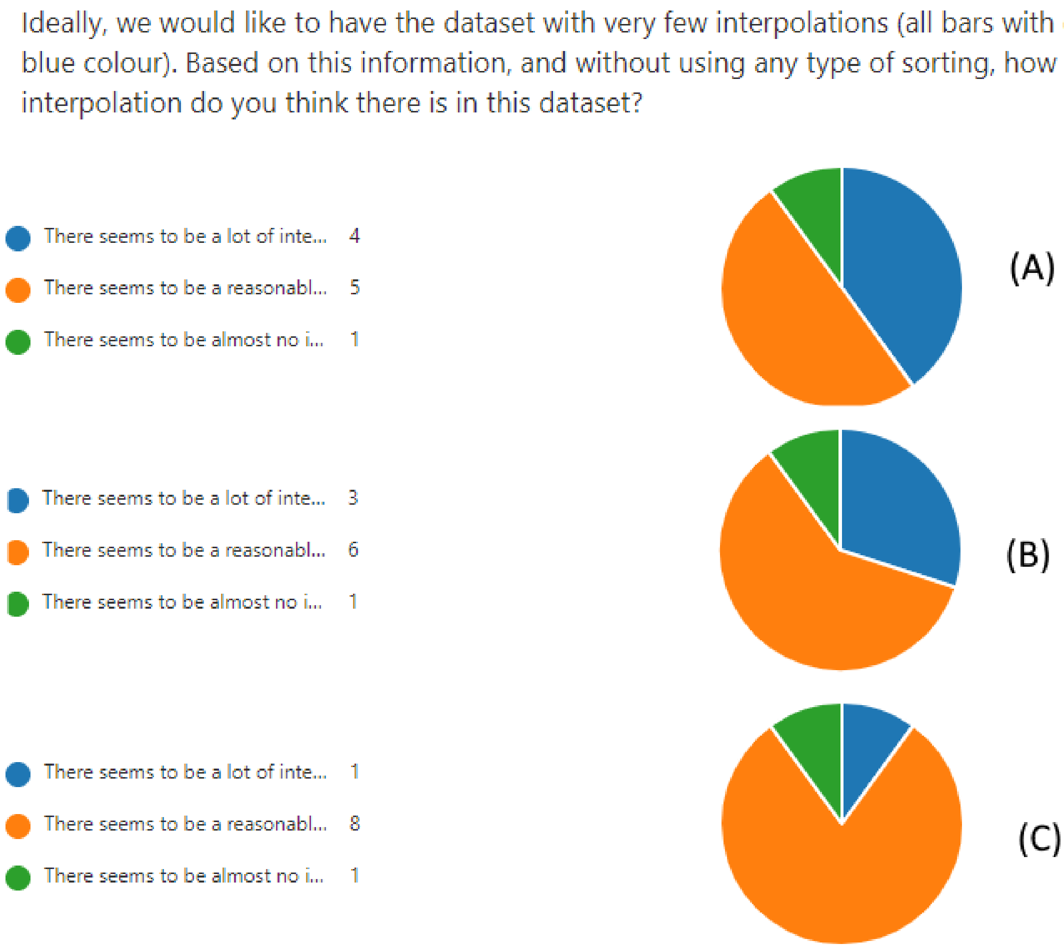

Figure 5.8. Responses for the questions 13 (A), 15 (B) and 17 (C).

5.5 Questionnaire

After the task exercises were completed we sent to the participants a small demographic questionnaire and a survey about the tool usability using 5-point Likert scale questions. After that they had to answer the following open ended questions:

- Please give us more comments about the system, especially things that you liked/disliked

- Is there any functionality that you wish was included?

5.5.1 Results

An overview of the answers for the questionnaire can be seen in Figure 5.9, and overall the result seems promising with most of the participants having a promising outlook towards the tool. The exception is the participants' feeling towards plotting the trajectory, with 4 participants being neutral about it being easy to plot, which is understandable since almost no exercise required plotting trajectories except for exercise 21, although we don't know why one participant somewhat disagreed with this statement. It is also interesting to note that 30 percent of the participants were neutral about noticing that some trips were more anomalous in specific segments. However, most participants correctly answered questions 11, 12 and 13, indicating that this neutral feeling could be when they were having an overview of all trips.

We got very positive feedback for the open-ended questions, such as: "The System is very interactive", "I like using filters to find outliers for each segment" and "The interface was sleek and intuitive and uncluttered". But we also received some feedback about improving the filters and some users talked about the confusion between colors and bar length for representing the score, such as this answer "I liked it , it was a good one indeed, found it little confusing figuring out the outliers and stuff but as it went on got comfortable using it." and "...although I may have gotten mixed up in the beginning with identifying bars that were orange with outliers. In the end, it all made sense, and I understood that longer bars mean high scores, which means something is an outlier.". This confusion was also noticed during the study.

5.6 Discussion

Overall, users were able to find anomalies using the tool and find in which subtrajectory the anomaly took place. The users were also able to make sense of the interpolation and also able to decide how it affected the score of a subtrajectory. We also found that most participants liked the usability of our tool in general. However, some of the functionalities we envisioned for our tool did not work as expected.

One of the problems found with this study is that users can get confused between color or bar length representing the score of a subtrajectory. We tried to solve this issue prior to our experiment by emphasizing this difference during the tutorial and

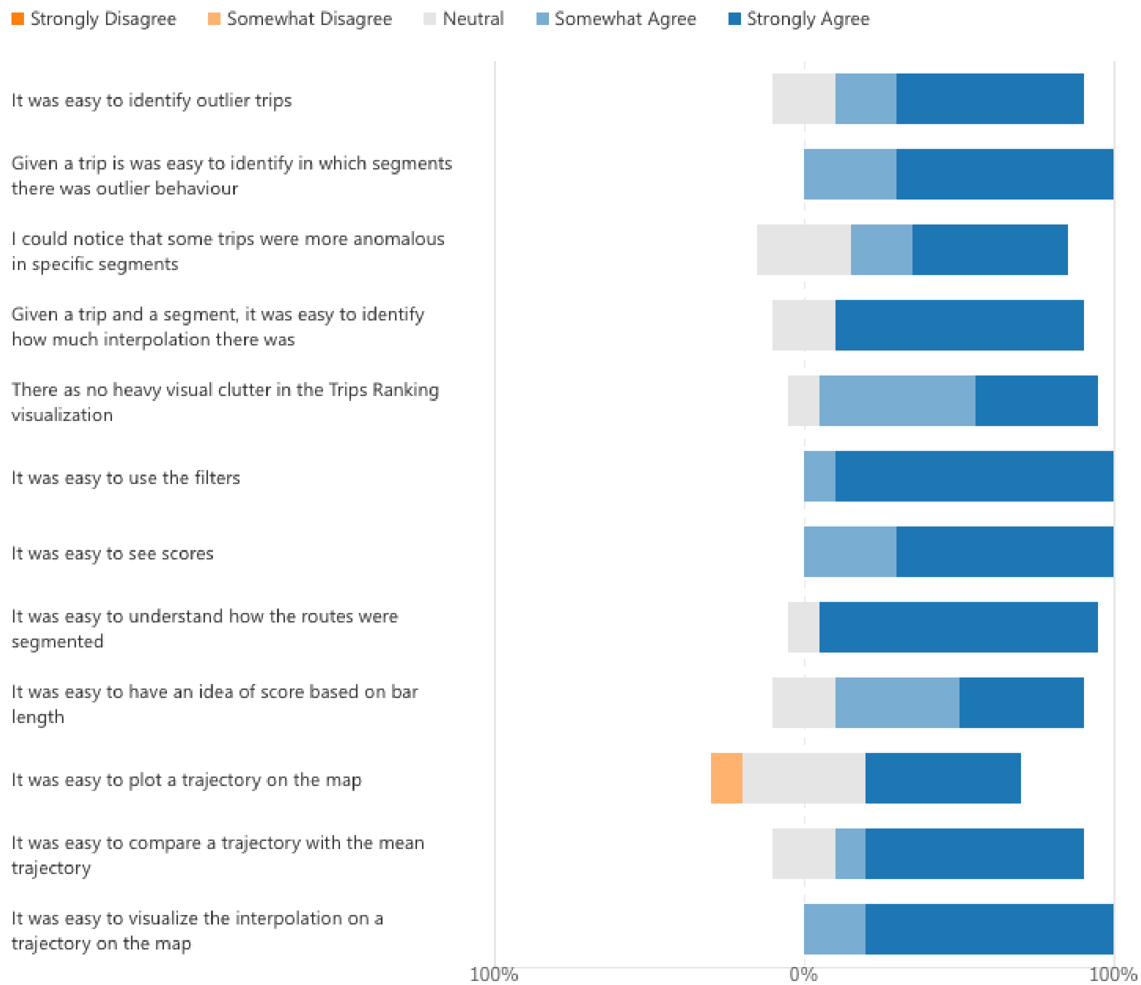

Figure 5.9. Usability questions using 5-point Likert scale and percentage of answers for each of the possible options.

adding a question mark in the tool, which also explained the difference between them. One of the reasons for that is that the color stands out more than the bar length even though the score is more important than the interpolation. However, it seems from the open questionnaire that the users started getting used to it after using it for a while. Still, this is something we should take into account and maybe we should allow the user to choose if they want the color or the length of the bar to represent the score. Another thing that needs to be improved is the anomaly distribution; we need to either improve the explanation or how we display it to the user.

This study was conducted online, however I forgot to add that we would need participants share their screen with me. For this reason, I could not see the users interacting with the tool, which makes it hard to identify why some users got some questions wrong on the scenario exercises. We also felt that a post-exercise interview could give more in-depth feedback about our tool, especially about things that did not work as expected.

Chapter 6

Conclusions

In this work, we have investigated the current works that focus on finding *kinematic* anomalies in the maritime domain, and we found a lack of visual analytics tools that focus on finding local anomalies and that take trajectories interpolation into account when displaying anomalies to users (described in Section 3). We then proposed and developed a web tool that segmenting trip trajectories and giving a score to each subtrajectory, users are then able to interact with this tool through filtering, sorting to find trips that have local anomalies. The users can also plot trip trajectories in the map and identify which portions of that trajectory were interpolated. A significant part of this work was done in the preprocessing step, where raw AIS data is cleaned, and trip trajectories were interpolated, then segments and substrajectories are created before the user can interact with the system. We then evaluated our tool with users and we found that overall users were able to find trips with outlier behaviour and identify in which spatial segment the anomaly took place, and users were also able to use the interpolation as a way to increase or decrease their confidence in a score. However, we also found some limitations and a lot of space for improvement which will be discussed in the next section.

6.1 Discussions

In this section we will discuss some of the limitations we found in our work and how we plan to address them.

User study

Ideally we would like to have followed a User-Centered Design process, starting with getting requirements from maritime security personnel, identifying which metaphors would work the best, creating some proof-of-concepts and improve on it and finally develop a working tool. However, due to the fact that we didn't have time to go though the process during a Master Thesis, our work was based on paper in the field. In the future, we want to be able to talk with possible users, and identify some problems our solutions is missing and how we can improve our tool.

Score calculation

One of the main limitations in this work is the way we calculate the score. We make the assumption that the subtrajectories values follow a single normal distribution, with most data being represented by non-anomalous trips. We believe that the second assumption should be valid in most cases; however, even when comparing the same class of vessels, some abnormal conditions, such as windy weather, may affect vessel speed and trajectory, causing them to be perceived as anomalous in our system. In order to solve this limitation we plan to use a clustering algorithm, such as k-means or DBSCAN, to group trips with similar trajectories. Then we could extract the normal behaviour and give a score for each of these groups.

Segmentation and local anomaly

Our local anomaly detection only works with well-segmented subtrajectories. In cases where the spatial segmentation is too large, it may miss some anomalies. Another limitation is that for each trip we only create one subtrajectory per segment; this means that it won't work well for trajectories that pass through the same segment more than once, which would be the case for fishing trajectories or trips that start and end at the same port.

We plan to address some of these issues by adding a page that allows the users to choose between creating the segments automatically or manually. If the user chooses to create manually, the user should be able to draw spatial segments on a map using drawing tools in the map. Otherwise, we will create segments based on trajectory patterns, such as straight lines, loops, etc. We will also change how subtrajectories are created so that a segment may create multiple subtrajectories for the same trip. This solution may still not work for fishing vessels since they have a much more complicated pattern, but it is not something we plan to address in the near future.

Mean trajectory

As discussed previously, we use mean trajectory to display the correct path a trip should follow, and one of the issues in using this approach is that generated points may not represent a real trajectory. In some cases, the trajectory may even be located in impossible places, such as in the middle of an island. Another limitation is that we assume there is one correct path, which may often be the case, especially when in the open sea due to sea lanes regulations. However, there may be other correct paths, especially in regions close to ports which are not covered by our solution.

For this reason, we plan to use a medoid trajectory instead of the mean, which will result in always having a valid trajectory to represent the correct path. When the user selects a trip the correct path will change based on which cluster the selected trip belongs to.

Exploration and visualization

We know it is essential for maritime operators to identify anomalies and understand what causes them, for example, understanding if the deviation is related to a very low speed or deviation from the path. In our work, this is still limited. The only way the user can identify which attribute contributed to the anomaly is by recomputing the score using only a single attribute.

A limitation we have with the current visualization is that the width of the columns is dynamic to the user's screen, which may limit the number of segments we are able to show to the user without affecting the readability of the table. And if there are too many trips to be displayed the lines become too small.

We could use the current tabular metaphor for exploration in a way that users could see attribute values and distribution for each attribute, this way the user will be able to see which attribute contributes to the deviation for a specific segment. But we are also considering using using a pixel oriented visualization [19] for that, which works well for high amounts of data, which may improve some of the visual clutter the users find and also we may use it to show the scores if we find that the trajectories need many segments. Another metaphor we are considering adopting to reduce visual clutter is parallel sets [2] where we could group trips by similar score over segments, and then the user could see more details by clicking on a group.

6.1.1 User input on score calculation

Right now the ways the user can change how the score is computed is limited to choosing segments or attributes. It may be interesting to allow more ways to user affect the score. One way we think may be interesting to consider is that, if the user knows a trip that has a good pattern they may choose it as the normal behaviour and then other trips could be scored in comparison to it.

6.1.2 Interpolation

A great deal of our work aims to show the impact of the interpolation on the score; however different interpolation techniques may produce very different results, and our tool is limited by the technique and parameters we chose. It may be interesting to give the option for the user to change the interpolation technique used, especially for trips where the user noticed the interpolation was done incorrectly.

Bibliography

[1] Amina Adadi and Mohammed Berrada. Peeking inside the black-box: A survey on explainable artificial intelligence (xai). *IEEE Access*, 6:52138–52160, 2018.

[2] Fabian Bendix, Robert Kosara, and Helwig Hauser. Parallel sets: visual analysis of categorical data. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, pages 133–140. IEEE, 2005.

[3] Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. Lof: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pages 93–104, 2000.

[4] Kevin Buchin, Maike Buchin, Marc Van Kreveld, Maarten Löffler, Rodrigo I Silveira, Carola Wenk, and Lionov Wiratma. Median trajectories. *Algorithmica*, 66(3):595–614, 2013.

[5] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.

[6] Tatyana Dimitrova, Aris Tsois, and Elena Camossi. Development of a web-based geographical information system for interactive visualization and analysis of container itineraries. *Int. J. Comput. Inf. Technol*, 3(1), 2014.

[7] Renata Dividino, Amilcar Soares, Stan Matwin, Anthony W Isenor, Sean Webb, and Matthew Brousseau. Semantic integration of real-time heterogeneous data streams for ocean-related decision making. In *Big Data and Artificial Intelligence for Military Decision Making*. STO, 2018.

[8] Enrica d'Afflisio, Paolo Braca, Leonardo M Millefiori, and Peter Willett. Detecting anomalous deviations from standard maritime routes using the ornstein–uhlenbeck process. *IEEE Transactions on Signal Processing*, 66(24):6474–6487, 2018.

[9] Willem Eerland, Simon Box, Hans Fangohr, and András Sóbester. Teetool–a probabilistic trajectory analysis tool. *Journal of Open Research Software*, 5(1), 2017.

[10] Torkild Eriksen, Gudrun Høye, Bjørn Narheim, and Bente Jensløkken Meland. Maritime traffic monitoring using a space-based ais receiver. *Acta Astronautica*, 58(10):537–549, 2006.

[11] Mohammad Etemad, Amilcar Soares, Elham Etemad, Jordan Rose, Luis Torgo, and Stan Matwin. SWS: An unsupervised trajectory segmentation algorithm

based on change detection with interpolation kernels. *GeoInformatica*, pages 1–21, 2020.

[12] Michele Fiorini, Andrea Capata, and Domenico D Bloisi. Ais data visualization for maritime spatial planning (msp). *International Journal of e-Navigation and Maritime Economy*, 5:45–60, 2016.

[13] Hanqi Guo, Zuchao Wang, Bowen Yu, Huijing Zhao, and Xiaoru Yuan. Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In *2011 IEEE Pacific Visualization Symposium*, pages 163–170. IEEE, 2011.

[14] Dini Oktarina Dwi Handayani, Wahju Sediono, and Asadullah Shah. Anomaly detection in vessel tracking using support vector machines (svms). In *2013 International Conference on Advanced Computer Science Applications and Technologies*, pages 213–217. IEEE, 2013.

[15] Bilal Idiri and Aldo Napoli. The automatic identification system of maritime accident risk using rule-based reasoning. In *2012 7th International Conference on System of Systems Engineering (SoSE)*, pages 125–130. IEEE, 2012.

[16] Amílcar Soares Júnior, Chiara Renso, and Stan Matwin. Analytic: An active learning system for trajectory classification. *IEEE computer graphics and applications*, 37(5):28–39, 2017.

[17] Amilcar Soares Junior, Valeria Cesario Times, Chiara Renso, Stan Matwin, and Lucidio AF Cabral. A semi-supervised approach for the semantic segmentation of trajectories. In *2018 19th IEEE International Conference on Mobile Data Management (MDM)*, pages 145–154. IEEE, 2018.

[18] Samira Kazemi, Shahrooz Abghari, Niklas Lavesson, Henric Johnson, and Peter Ryman. Open data for anomaly detection in maritime surveillance. *Expert Systems with Applications*, 40(14):5719–5729, 2013.

[19] Daniel A Keim. Pixel-oriented visualization techniques for exploring very large data bases. *Journal of Computational and Graphical Statistics*, 5(1):58–77, 1996.

[20] Daniel A Keim, Florian Mansmann, and Jim Thomas. Visual analytics: how much visualization and how much analytics? *ACM SIGKDD Explorations Newsletter*, 11(2):5–8, 2010.

[21] Kwang-Il Kim and Keon Myung Lee. Deep learning-based caution area traffic prediction with automatic identification system sensor data. *Sensors*, 18(9):3172, 2018.

[22] Valérie Lavigne. Interactive visualization applications for maritime anomaly detection and analysis. In *ACM SIGKDD Workshop on Interactive Data Exploration and Analytics*, page 75, 2014.

[23] Rikard Laxhammar. *Anomaly detection in trajectory data for surveillance applications*. PhD thesis, Örebro universitet, 2011.

[24] Rikard Laxhammar. *Conformal anomaly detection: Detecting abnormal trajectories in surveillance applications*. PhD thesis, University of Skövde, 2014.

[25] Rikard Laxhammar and Göran Falkman. Conformal prediction for distribution-independent anomaly detection in streaming vessel data. In *Proceedings of the first international workshop on novel data stream pattern mining techniques*, pages 47–55, 2010.

[26] Rikard Laxhammar and Göran Falkman. Online detection of anomalous sub-trajectories: A sliding window approach based on conformal anomaly detection and local outlier factor. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 192–202. Springer, 2012.

[27] Changqing Liu and Xiaoqian Chen. Inference of single vessel behaviour with incomplete satellite-based ais data. *The Journal of navigation*, 66(6):813, 2013.

[28] Jed A Long. Kinematic interpolation of movement data. *International Journal of Geographical Information Science*, 30(5):854–868, 2016.

[29] Min Lu, Zuchao Wang, and Xiaoru Yuan. Trajrank: Exploring travel behaviour on a route by trajectory ranking. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 311–318. IEEE, 2015.

[30] Etienne Martineau and Jean Roy. Maritime anomaly detection: Domain introduction and review of selected literature. Technical report, DEFENCE RESEARCH AND DEVELOPMENT CANADA VALCARTIER (QUEBEC), 2011.

[31] Steven Mascaro, Ann E Nicholso, and Kevin B Korb. Anomaly detection in vessel tracks using bayesian networks. *International Journal of Approximate Reasoning*, 55(1):84–98, 2014.

[32] Lucas May Petry, Amilcar Soares, Vania Bogorny, Bruno Brandoli, and Stan Matwin. Challenges in vessel behavior and anomaly detection: From classical machine learning to deep learning. In Cyril Goutte and Xiaodan Zhu, editors, *Advances in Artificial Intelligence*, pages 401–407, Cham, 2020. Springer International Publishing.

[33] Fabio Mazzarella, Alfredo Alessandrini, Harm Greidanus, Marlene Alvarez, Pietro Argentieri, Domenico Nappo, and Lukasz Ziemba. Data fusion for wide-area maritime surveillance. In *Workshop on Moving objects at Sea*, 2013.

[34] Fabio Mazzarella, Michele Vespe, Alfredo Alessandrini, Dario Tarchi, Giuseppe Aulicino, and Antonio Vollero. A novel anomaly detection approach to identify intentional ais on-off switching. *Expert Systems with Applications*, 78:110–123, 2017.

[35] Van-Suong Nguyen, Nam-kyun Im, and Sang-min Lee. The interpolation method for the missing ais data of ship. *Journal of Navigation and Port Research*, 39(5):377–384, 2015.

[36] Giuliana Pallotta, Michele Vespe, and Karna Bryan. Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction. *Entropy*, 15(6):2218–2245, 2013.

[37] Animesh Patcha and Jung-Min Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks*, 51(12):3448–3470, 2007.

[38] Peter Pirolli and Ramana Rao. Table lens as a tool for making sense of data. In *Proceedings of the workshop on Advanced visual interfaces*, pages 67–80, 1996.

[39] Maria Riveiro and Göran Falkman. The role of visualization and interaction in maritime anomaly detection. In *Visualization and Data Analysis 2011*, volume 7868, page 78680M. International Society for Optics and Photonics, 2011.

[40] Maria Riveiro, Goran Falkman, and Tom Ziemke. Improving maritime anomaly detection and situation awareness through interactive visualization. In *2008 11th International Conference on Information Fusion*, pages 1–8. IEEE, 2008.

[41] Maria Riveiro, Göran Falkman, Tom Ziemke, and Håkan Warston. Visad: an interactive and visual analytical tool for the detection of behavioral anomalies in maritime traffic data. In *Visual Analytics for Homeland Defense and Security*, volume 7346, page 734607. International Society for Optics and Photonics, 2009.

[42] Jean Roy. Anomaly detection in the maritime domain. In *Optics and Photonics in Global Homeland Security IV*, volume 6945, page 69450W. International Society for Optics and Photonics, 2008.

[43] Roeland Scheepens, Niels Willems, Huub van de Wetering, and Jarke J Van Wijk. Interactive visualization of multivariate trajectory data with density maps. In *2011 IEEE pacific visualization symposium*, pages 147–154. IEEE, 2011.

[44] Pan Sheng and Jingbo Yin. Extracting shipping route patterns by trajectory clustering model based on automatic identification system data. *Sustainability*, 10(7):2327, 2018.

[45] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE symposium on visual languages*, pages 336–343. IEEE, 1996.

[46] Amílcar Soares, Renata Dividino, Fernando Abreu, Matthew Brousseau, Anthony W Isenor, Sean Webb, and Stan Matwin. Crisis: integrating ais and ocean data streams using semantic web standards for event detection. In *2019 International Conference on Military Communications and Information Systems (ICMCIS)*, pages 1–7. IEEE, 2019.

[47] Amílcar Soares, Jordan Rose, Mohammad Etemad, Chiara Renso, and Stan Matwin. Vista: A visual analytics platform for semantic annotation of trajectories. In *EDBT*, pages 570–573, 2019.

[48] Amílcar Soares Júnior, Bruno Neiva Moreno, Valéria Cesário Times, Stan Matwin, and Lucídio dos Anjos Formiga Cabral. Grasp-uts: an algorithm for unsupervised trajectory segmentation. *International Journal of Geographical Information Science*, 29(1):46–68, 2015.

[49] J Thomas and K Cook. Illuminating the path: Research and development agenda for visual analytics. national visualization and analytics center; ieee: 2005.

[50] Christian Tominski, Heidrun Schumann, Gennady Andrienko, and Natalia Andrienko. Stacking-based visualization of trajectory attribute data. *IEEE Transactions on visualization and Computer Graphics*, 18(12):2565–2574, 2012.

[51] Joeri Van Laere and Maria Nilsson. Evaluation of a workshop to capture knowledge from subject matter experts in maritime surveillance. In *2009 12th International Conference on Information Fusion*, pages 171–178. IEEE, 2009.

[52] Iraklis Varlamis, Ioannis Kontopoulos, Konstantinos Tserpes, Mohammad Etemad, Amilcar Soares, and Stan Matwin. Building navigation networks from multi-vessel trajectory data. *GeoInformatica*, 2020.

[53] Iraklis Varlamis, Konstantinos Tserpes, Mohammad Etemad, Amílcar Soares Júnior, and Stan Matwin. A network abstraction of multi-vessel trajectory data for detecting anomalies. In *EDBT/ICDT Workshops*, volume 2019, 2019.

[54] Guizhen Wang, Abish Malik, Calvin Yau, Chittayong Surakitbanharn, and David S Ebert. Traseer: A visual analytics tool for vessel movements in the coastal areas. In *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, pages 1–6. IEEE, 2017.

[55] Niels Willems, Huub Van De Wetering, and Jarke J Van Wijk. Visualization of vessel movements. In *Computer Graphics Forum*, volume 28, pages 959–966. Wiley Online Library, 2009.

[56] Niels Willems, Willem Robert van Hage, Gerben de Vries, Jeroen HM Janssens, and Véronique Malaisé. An integrated approach for visual analysis of a multisource moving objects knowledge base. *International Journal of Geographical Information Science*, 24(10):1543–1558, 2010.

[57] Xing Wu, Afifa Rahman, and Victor A Zaloom. Study of travel behavior of vessels in narrow waterways using ais data–a case study in sabine-neches waterways. *Ocean Engineering*, 147:399–413, 2018.

[58] Mingyue Xie. Trajectories medoid and clustering. *Computer Science*, 2019.

[59] Wanqi Yang, Yang Gao, and Longbing Cao. Trasmil: A local anomaly detection framework based on trajectory segmentation and multi-instance learning. *Computer Vision and Image Understanding*, 117(10):1273–1286, 2013.

[60] Daiyong Zhang, Jia Li, Qing Wu, Xinglong Liu, Xiumin Chu, and Wei He. Enhance the ais data availability by screening and interpolation. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)*, pages 981–986. IEEE, 2017.

[61] Rong Zhen, Yongxing Jin, Qinyou Hu, Zheping Shao, and Nikitas Nikitakos. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and naïve bayes classifier. *The Journal of Navigation*, 70(3):648, 2017.

[62] Dimitrios Zissis, Elias K Xidias, and Dimitrios Lekkas. A cloud based architecture capable of perceiving and predicting multiple vessel behaviour. *Applied Soft Computing*, 35:652–661, 2015.

Appendix A

Consent Form

DALHOUSIE
UNIVERSITY

CONSENT FORM

Project title:
User evaluation of Trip Outlier Scoring Tool

Lead researcher
Fernando Henrique Oliveira Abreu,
Faculty of Computer Science, Dalhousie University, 6050 University Ave., PO Box 15000, Halifax, NS, B3H 4R2, Canada
Phone 902-880-9634, Email fernando.abreu@dal.ca

Supervisor:
Dr. Stan Matwin
Faculty of Computer Science, Dalhousie University, 6050 University Ave., PO Box 15000, Halifax, NS, B3H 4R2, Canada
Phone 902-494-4320, Email stan@cs.dal.ca

Introduction
We invite you to take part in a research study being conducted by Fernando Henrique Oliveira Abreu, who is a Master of Computer Science student at the Faculty of Computer Science, Dalhousie University. Choosing whether or not to take part in this research is entirely your choice. There will be no impact on your studies/work if you decide not to participate in the research. The information below tells you about what is involved in the research, what you will be asked to do and about any benefit, risk, inconvenience or discomfort that you might experience.
You should discuss any questions you have during or after this study with Fernando Henrique Oliveira Abreu. Please ask as many questions as you like.

Purpose and Outline of the Research Study
The goal of this study is to evaluate a tool created to help maritime operators to identify vessels that may show some signs of abnormality during their trip (e.g. a vessel that speeds too much compared to others). We want to validate how easy it is to use.

Who Can Take Part in the Research Study
You may participate in this study if you are a Dalhousie University student/staff/faculty. In case this study is conducted online the participant will need a computer with a browser installed with javascript enabled and internet access, and an equipment that allows us to communicate (e.g. headset).

What You Will Be Asked to Do
If you decide to participate in this research you will be asked to either attend one visit to the lead researcher lab located at the Playground 441 in the Faculty of Computer Science, Dalhousie University, 6050 University Ave., PO Box 15000, Halifax, NS, B3H 4R2, Canada, or access a link to an online meeting through Microsoft Teams. The study will take approximately one hour. During the study, you will be doing the following:

- You will sign the consent form.
- You will complete a demographic questionnaire.
- You will be given a tutorial on how to use the software.
- You will be given a randomly generated ID and the evaluation (post-condition) questionnaire.
- You will perform tasks on the proposed tool.
- You will submit the post-study questionnaire and comments

Possible Benefits, Risks, and Discomforts
Benefits: You will be given a $20 CAD egift card for compensation. In addition, your participation will be greatly appreciated, and we expect that it will help us to learn about the effectiveness and usability of our tool.

Risks: No extraordinary risks are anticipated in the present study. The only risk that could happen is the participant's fatigue. Your name will not be connected to the data collected from you.

Discomforts: If participation in the study brings you any discomfort, please do not hesitate to contact the lead researcher, Fernando Henrique Oliveira Abreu by email at **fernando.abreu@dal.ca**.

Compensation / Reimbursement
To thank you for your time, we will give you a $20 CAD egift card at the end of the study even if you do not complete this study. You will be asked to send an email confirming that you have received the compensation.

How your information will be protected:
Confidentiality: Name and email address will be collected however these data will not have any direct link to your responses. The participant will receive a randomly generated number (not your name) in written records so that the research information I have about you contains no names, and so there will be no link between your code and your personal information. All paper records will be kept secure in a locked filing cabinet located in the researcher’s desk. In case the study is conducted online, the data will be stored on Microsoft Forms on my password-protected Dalhousie account, and we would use Microsoft Teams which is Dalhousie’s approved video conferencing tool, no video conference data will be stored. All data gathered from this study may potentially be used in publications and in the researcher master thesis. The quantitative data will be reported as grouped results and the qualitative data collected from the questionnaire will have an alphabet letter assigned to it with no meaning. This means that you will not be identified in any way in our reports. The only person who will conduct and have access to the participant responses data will be the Lead Researcher, Fernando Henrique Oliveira Abreu. David Langstroth will be forwarded to the participant’s egift receipt without any link to the participants’ response. Your email will only be stored in the consent form in case you wish to receive updates about this study.

Data retention: The data will be retained for five years after publication and then destroyed.

Data repositories: Microsoft Forms from the Lead Researcher account may be used in case of the study being performed online, and all data will be stored in a password-protected account created only for this study.

If You Decide to Stop Participating

You are free to leave the study at any time. If you decide to do so all the information that you have provided up to that point will be removed. After participating in the study, you can decide for up to one week if you want us to remove your data, to do so send an email to fernando.abreu@dal.ca with your random generated participant id. After that time, it will become impossible for us to remove it because it will already be analyzed. You still will receive full compensation even if you don't complete the study.

How to Obtain Results

If you would like to receive the study results you can add your email at the end of this form. If you do so the Lead Researcher will provide you with a short description of the study results by email when the study is finished if you would like.

Questions

We are happy to talk with you about any questions or concerns you may have about your participation in this research study. Please contact Fernando Henrique Oliveira Abreu at 902 880-9634, or email: fernando.abreu@dal.ca

If you have any complaints about the experiment you may contact the Ethics department:

Research Ethics Office of Research Services

P.O Box 15000 Halifax, NS B3H 4R2 Canada

Phone 902-494-3423, Email ethics@dal.ca