Unexpected Trajectory Detection Based on the Geometrical Features of AIS-Generated Ship Tracks

Wayan Mahardhika Wijaya[1], Yasuhiro Nakamura[2]
Graduate School of Science and Engineering, National Defense Academy of Japan
1-10-20 Hashirimizu, Yokosuka-shi, Kanagawa-ken, Japan 239-0811[1]
National Defense Academy of Japan, 1-10-20 Hashirimizu, Yokosuka-shi, Kanagawa-ken, Japan 239-0811[2]

***Abstract*—Due to the efficiency and reliability of delivering goods by ships, maritime transport has been the backbone of global trade. In normal circumstances, a ship's voyage is expected to assure the safety of life at sea, efficient and safe navigation, and protection of the maritime environment. However, ships may demonstrate unexpected behavior due to certain situations, such as machinery malfunction, unexpected bad weather, and other emergencies, as well as involvement in illicit activities. These situations pose threats to the safety and security of maritime transport. The expansion of the threats makes manual surveillance inefficient, which involves extensive labor and is prone to oversight. Thus, automated surveillance systems are required. This paper proposes a method to detect the unexpected behavior of ships based on the Automatic Identification System (AIS) data. The method exploits the geometrical features of AIS-generated trajectories to identify *unexpected trajectory*, which could be a deviation from the common routes, loitering, or both deviating and loitering. It introduces novel formulas for calculating trajectory redundancy and curvature features. The DBSCAN clustering is applied based on the features to classify trajectories as *expected* or *unexpected*. Unlike existing methods, the proposed technique does not require trajectory-to-image conversion or training of labeled datasets. The technique was tested on real-world AIS data from the South China Sea, Western Indonesia, Singapore, and Malaysian waters between July 2021 and February 2022. The experimental results demonstrate the method's feasibility in detecting deviating and loitering behaviors. Evaluation on a labeled dataset shows superior performance compared to existing loitering detection methods across multiple metrics, with 99% accuracy and 100% precision in identifying loitering trajectories. The proposed method aims to provide maritime authorities and fleet owners with an efficient tool for monitoring ship behaviors in real time regarding safety, security, and economic concerns.**

Keywords*—Automatic identification system; vessel trajectory classification; unexpected behavior detection; data mining; data-driven decision support

I. Introduction

Global trade has been heavily reliant on maritime transport. More than 80 percent of the worldwide merchandise trade volume is delivered by ships, which are considered an economical, energy-efficient, and reliable long-distance means of transportation [1]. To ensure its economic and energy-efficient advantages, the voyage of a ship needs to be carefully planned.

According to the Guideline of Voyage Planning mandated by the International Maritime Organization (IMO), the voyage of a ship is a deliberately planned event that should assure the safety of life at sea, efficient and safe navigation, and protection of the marine environment [2]. For the sake of simplicity, this paper uses the term expected voyage to refer to such a voyage that is compliant with the guideline.

In normal circumstances, any ship will not make any maneuver that endangers people's lives at sea. They will navigate as efficiently as possible in a safe manner, which means that they are to take the shortest and fastest route whenever it is safe to do so. They will not deliberately conduct any activity that causes pollution or damage to the marine environment. Particularly for vessels of types of cargo and tanker, constrained by strict regulations, economic, and safety requirements, they should be the most likely to perform the expected voyage. However, ships might not follow the expected voyage due to certain situations, such as machinery malfunction, unexpected bad weather, and other emergencies, as well as involvement in illicit activities. These situations pose threats to maritime security, and the threats are continuously expanding, making automated surveillance systems critically required in the maritime domain [3]. In addition, the 12-month ship anomaly data provided by the Indonesian Coast Guard consists of nearly 400 ships that demonstrate anomalous behaviors such as loitering, deviation from common routes, and AIS on/off[1]. Roughly 97% of the ships are of types cargo and tanker, which are the core of the international maritime transport. The anomalous behaviors were identified manually by experts, which means the actual number of the anomalous ships could be higher due to the possibility of oversight. Thus, an automated means of monitoring and examining ship voyages is necessary to confirm compliance with the expected voyage and preserve the benefits of maritime transport.

Meanwhile, due to the worldwide adoption of the sea-borne Automatic Identification System (AIS) on seagoing vessels, AIS has emerged as a potential leading source of ship voyage data. AIS shares navigational data among vessels, terrestrial base stations, and/or satellites. The data consists of static, dynamic, and voyage-related information. The static information includes ship name, type, and MMSI. Ship position, position timestamp, speed over ground (SOG), course over ground (COG), heading, and navigational status are the dynamic information, while destination, estimated time of arrival, and draught are voyage-related. MMSI stands for Maritime Mobile Service Identity, a unique nine-digit number to uniquely identify a ship or a coast radio station [4]. AIS device transmits messages containing the information every 2 to 10 seconds for ships moving faster than 3 knots and every 3 minutes when they are at anchor or moored and not moving faster than 3 knots [5].

[1]The data were granted upon a formal request from the authors.

The real-time feature of AIS has made it possible for maritime stakeholders to utilize AIS as a monitoring tool. The abundant availability and straightforward accessibility of AIS data have facilitated researchers to develop methods for analyzing ships' tracks and trajectories to comprehend ship behavior. The methods are designated for various tasks, such as anomalous behavior detection [6], trajectory classification [7][8][9], construction of port performance indicators [10][11], and building real-time indicators of world seaborne trade [12].

In this paper, a method to detect ships that do not follow the expected voyage is proposed. The method exploits the geometrical features of the AIS-generated trajectories to identify unexpected trajectories, representing the routes of ships that do not comply with the expected voyage. Specifically, the unexpected trajectory could be a trajectory that deviates from the common routes, a loitering trajectory, or a trajectory that is both loitering and deviates from the common navigation routes.

Existing studies take two approaches to classifying vessels' trajectories based on AIS data: analyzing the spatiotemporal characteristics of the AIS tracks and finding patterns by studying the geometrical features of the AIS-generated trajectories. Although tankers and cargo vessels are the core of international maritime transport [13], few studies address these types of ships for trajectory classification [9]. Most works put their attention on the trajectory of fishing vessels, such as the works of references [14] and [15]. Furthermore, the approaches that examine the geometrical features of the AIS-generated vessels' trajectories often involve the conversion of trajectory data into digital images and rely on the trajectory classification on manually labeled datasets [7], [16], [17].

This study takes the geometrical approach to classify ships trajectories. The approach calculates the rate of redundancy and curvature of all trajectories of interest. It proposes novel formulas for the calculation. Next, a weighted DBSCAN (Density-based Spatial Clustering of Application with Noise) clustering [18] is applied to the trajectories to classify trajectories that belong to the common voyages and those of the unexpected trajectory. The proposed method does not involve the conversion of trajectories into images and does not require any labeled datasets.

The purpose of this work is to provide a straightforward unexpected trajectory detection technique suitable to real-time surveillance systems for a designated maritime area.

It is intended to contribute in two ways: 1) help maritime authorities to efficiently identify unexpected behavior of vessels within their surveillance area, and 2) support merchant fleet owners with online monitoring tools to ensure that all of their ships are following the known efficient and safe voyage routes.

The remainder of this paper is organized as follows. Section II reviews relevant literature. Then, the proposed method is described in Section III followed by the presentation of the experiment results and evaluation in Section IV. Finally, Section V concludes this article and specifies the future tasks to further improve this work.

II. Related Works

Luo et al. classify ship trajectories into five types: 1) normal navigation trajectory, 2) anchoring or mooring trajectory, 3) navigation trajectory with deviation, 4) trajectory of missing AIS signal, and 5) irregular trajectory [9]. The normal navigation trajectory is defined as the trajectory of a ship traveling from the departure point to the destination port without redundancy, deviation of course, or loss of AIS transmission. In other words, any trajectories with redundancy, deviation, or loss of AIS transmission are deemed as not normal. Anchoring and mooring trajectories belonging to the ship-stopping behavior are discussed comprehensively in the work of Yan et al. [19]. In this paper, these stopping trajectories are removed by employing the method proposed in reference [11], [20], and only the trajectory between the start and end of a ship voyage is processed to identify unexpected trajectory. Navigation trajectory with deviation and trajectory of missing AIS signal refers to the types of anomalies in maritime traffic proposed by Lane et al. [21], whereas the irregular trajectory corresponds to the loitering trajectory discussed in references [16] and [22]. In this paper, any trajectories with loitering, or deviation, or having both of them are determined as unexpected trajectory.

Luo et al. employ an ensemble classifier, a combination of Naive Bayes and Random Forest classifiers, to classify the five types of vessel trajectories. The approach adopts the feature-extraction submodule of the Tsfresh package to automatically extract spatiotemporal features from vessel trajectories [23]. In their experiment, they rely on a labeled dataset to conduct the trajectory classification. However, their work does not provide the information on how and by who the dataset was labeled. In addition, each trajectory is given one label and grouped into one type of trajectory. In the real world, a trajectory may belong to more than one category, such as one that both deviates from the common routes and loiters. Thus, in this study, the unsupervised learning approach is selected, and each input trajectory is classified into an expected trajectory or an unexpected trajectory, where the unexpected trajectory may be deviating or loitering, or exhibiting both behaviors. In other words, this study does not involve the labor of dataset labeling and does not force the classification into the provided labels.

A technique to specifically detect loitering trajectory was proposed by Zhang et al. which introduce the concept of trajectory redundancy [16]. The formula to calculate the redundancy is as defined in Eq. 1. However, the method is designated to classify the trajectories belonging to the types of vessels that consider loitering a normal behavior in their nature of operation. These types of vessels include fishing ships, Search and Rescue (SAR) vessels, tug boats, survey ships, patrol boats, and ships of military operations. The method does not recognize loitering as an abnormal or unexpected behavior.

Identifying the gap, Wijaya and Nakamura proposed a loitering detection method targeting vessels that do not normally engage in loitering movement, such as tankers and cargo ships [22]. They define loitering as a type of anomaly in maritime traffic. The method exploits the spatiotemporal features of the AIS tracks, such as speed, course change, heading change, and the geodesic distance between two consecutive tracks. It identifies the loitering trajectory along with its score, which determines how severe the loitering is. The method's implementation in the maritime surveillance system will facilitate the operators to sort the detected loitering vessels based on their priority (loitering score). The evaluation experiment proves that the method outperforms the loitering detection technique proposed by Zhang et al. in all metrics (Table I). However, the method is specifically designated to detect loitering trajectory while the unexpected trajectory defined in this paper is not only about loitering. It has a wider scope to include loitering and deviating trakcs.

III. Proposed Methods

This study defines *unexpected trajectory* as a trajectory constructed with the AIS tracks of a ship voyage that does not follow the IMO's Guideline of Voyage Planning.

According to the guideline, the voyage of a ship should assure: 1) the safety of life at sea, 2) efficient and safe navigation, and 3) protection of the marine environment [2].

Considering the efficiency and safety of navigation, ships should take the most straightforward, shortest, and fastest route whenever it is safe to do so. This ensures efficiency in both cost and time. Thus, any ship trajectory that is not straightforward or demonstrates redundancy is considered an unexpected trajectory. It could be a trajectory that deviates from the common routes, a loitering trajectory, or a trajectory that is both loitering and deviates from the common navigation routes.

Zhang et al. introduce the concept of trajectory redundancy(TR) as a formula for detecting loitering trajectories from AIS tracks [16]. Eq. (1) represents the formula, where TR is denoted by ψ, D is the length of ship trajectory, and P is the perimeter of the minimum bounding rectangle of ship trajectory. The larger ψ is the greater the possibility of loitering, and the threshold is $\psi_{min} = 1$.

Fig. 1 shows three different trajectories in the same size of the spatial range (all have the same P) with $\psi < 1$, $\psi \approx 1$, and $\psi > 1$.

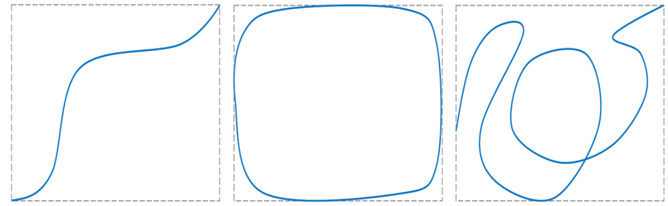
$$\psi = D/P \tag{1}$$

Since the TR calculates redundancy by comparing the length of trajectory with the perimeter of the trajectory's bounding box, a redundant trajectory along the diagonal of the bounding box may result in $\psi < 1$. In other words, it may not be considered as redundant or loitering.

Thus, in this paper, the trajectory redundancy is calculated by comparing the length of trajectory, denoted by D, with the length of the diagonal of the trajectory's bounding box, denoted by L. Eq. 2 represents the comparison.

$$R = D/L \tag{2}$$

However, this study considers that merely calculating the redundancy of ship trajectory is not enough to detect unexpected trajectory of the ship's voyage. Hence, a formula to measure the curvature of vessel trajectory is proposed as in Eq. 4.



(a) (b) (c)

Fig. 1. Three different trajectories with the equivalent TR (ψ) values. The grey dashed lines represent the minimum bounding rectangle of each trajectory, while the blue solid lines depict vessel trajectories: (a), (b), and (c) are the trajectory with $\psi < 1$, $\psi \approx 1$, and $\psi > 1$, respectively.

For every voyage's trajectory $M = \{m_0, m_1, m_2, \ldots, m_n\}$ where $0 \leq i \leq n$, $T = \{t_0, t_1, t_2, \ldots, t_n\}$ is the corresponding timestamps of the trajectory M as to m_i is the track position at timestamp t_i. In other words, m_0 is the starting track, and m_n is the last track position of a ship's voyage. The trajectory curvature is defined as inversely proportional to the average Cartesian distance from the starting track position m_0 to each track position $m_1, m_2, m_3, \ldots, m_n$. Eq. 3 yields the average Cartesian distance $\overline{d}$, where $d(m_0, m_i)$ is the Cartesian distance between m_0 and m_i.

$$\overline{d} = \frac{1}{(n+1)} \sum_{i=1}^{n} d(m_0, m_i) \tag{3}$$

$$C = \frac{1}{\overline{d}} \tag{4}$$

The variables used in Eq. 2 and 4 indicate that this paper utilizes the geometrical features of vessels' trajectories to detect unexpected trajectories instead of exploiting the spatiotemporal characteristics as in the existing work of Wijaya and Nakamura [22].

The overall process of the unexpected trajectory detection method proposed in this paper is conducted in three steps: 1) AIS data preprocessing, 2) trajectory segmentation to split the stopping and underway trajectories of every ship's voyage, and 3) the implementation of Eq. 2 and 4 to detect unexpected trajectories. This paper employs the AIS data preprocessing and trajectory segmentation methods described in [20]. The preprocessing removes all invalid data, while the segmentation separates the stopping and underway segments of each trajectory representing a ship voyage. The validated underway segments are the input for the unexpected trajectory detection computation technique proposed in this paper.

Eq. 2 and 4 are applied to each underway segment of all trajectories of the ships of interest to calculate the rate of redundancy and curvature. Every ship's trajectory represents a ship's voyage from one endpoint to another. The starting point and the destination can be a port or a water area. For example, a container's voyage from Singapore port to the port of Jakarta, a tanker's voyage from the South China Sea to the Indian Ocean, and a cargo coming from the Indian Ocean to Singapore port. Due to the constraints of the geographical features, the typical characteristics of the waters, and the weather patterns between the two endpoints, voyages of the same ends should have at least one commonly traveled route, which has been proven to be efficient and safe for a known period of time. Normally, most ships are expected to follow the commonly known route(s) rather than taking the risk of navigating the unknown passage. However, under certain situations, a ship may take an unexpected path indicated by her trajectory.

To identify the unexpected trajectory(s) of the ships' trajectories belonging to voyages between the same endpoints, this paper applies the DBSCAN clustering algorithm to classify the trajectories by using the rate of redundancy (Eq. 2) and the trajectory curvature (Eq. 4) as the clustering features. Since the rate of redundancy R is the main parameter to detect loitering, and the curvature C is the extension, R is weighted twice the weight of C in the implementation of DBSCAN.

The DBSCAN algorithm is selected because it can detect noises, the ones whose features cannot be associated with any clusters. The noises are labeled '-1', which means they do not belong to any clusters. They are different and few in number. When there is no noise detected, none of the data will be labeled '-1'. Thus, if the unexpected trajectory exists, it should be classified as noise since it must have different features from the common trajectories and be much fewer in number.

For every voyage's trajectory $M = \{m_0, m_1, m_2, \ldots, m_n\}$ with timestamps $T = \{t_0, t_1, t_2, \ldots, t_n\}$, W is defined as a time window with an arbitrary duration k hours, where $W \subseteq T$ and the duration of W is less than the duration of T. The time window W is sliding from the start to the end of T while executing the calculation of the rate of redundancy R and the curvature C. The calculation results are compared and each maximum value of R and C is returned as in Eq. 5 and 6.

$$R_w = Max(R_{(t_0,t_k)}, R_{(t_{0+1},t_{k+1})}, \ldots, R_{(t_{n-k},t_n)}) \quad (5)$$

$$C_w = Max(C_{(t_0,t_k)}, C_{(t_{0+1},t_{k+1})}, \ldots, C_{(t_{n-k},t_n)}) \quad (6)$$

This process is to excerpt the segment of a voyage's trajectory with the maximum redundancy and curvature. To further precisely locate the segment, multiple time windows with different durations are applied. In this case, three time-windows, $W1 = 6$ hours, $W2 = 24$ hours, and $W3 = 48$ hours, are determined. The maximum R and C of each time window are compared to select the one final maximum R and C as expressed in Eq. 7 and 8.

$$R_{max} = Max(R_{w1}, R_{w2}, R_{w3}) \quad (7)$$

$$C_{max} = Max(C_{w1}, C_{w2}, C_{w3}) \quad (8)$$

The calculation of R and C with the sliding time windows is executed on every underway trajectory belonging to the voyages of the same endpoints. The computation results in a set of trajectory excerpts $E = \{e_0, e_1, e_2, \ldots, e_m\}$ having the R_{max} and C_{max} as their attributes, where $m+1$ is the number of trajectories belonging to the voyages of the same origin and destination. Thus, each trajectory excerpt e_j, for $0 \leq j \leq m$, can be expressed as a position on a two-dimensional space $e_j(x_j, y_j)$, where $x = R_{max}$ and $y = C_{max}$. Here, the DBSCAN algorithm is implemented to classify the set of trajectory excerpts E by taking R_{max} and C_{max} as the clustering features. The Euclidean distances amongst the set of E are calculated to determine the epsilon ε parameter of the DBSCAN algorithm. Every trajectory excerpt e_j with -1 label is classified as an excerpt of an unexpected trajectory, while the rest belong to the trajectories of the common routes. The whole process of the unexpected trajectory detection workflow is summarized in Fig. 2.

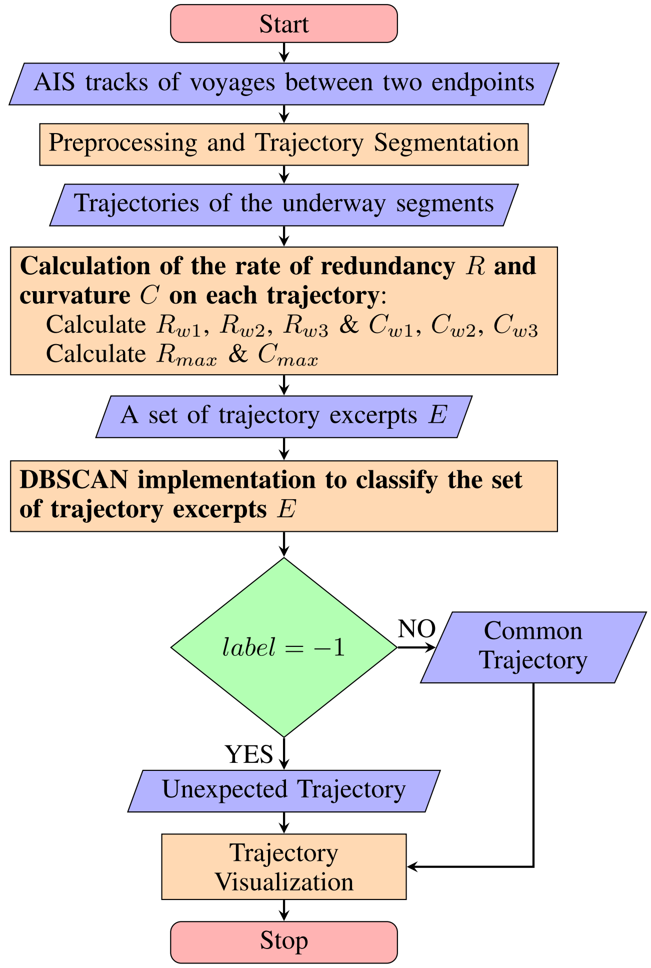


Fig. 2. The flowchart of the *Unexpected Trajectory* detection method.

IV. Experiment and Evaluation

The proposed unexpected trajectory detection technique is implemented on real-world historical AIS data of vessels navigating through the southern part of the South China Sea, western Indonesia, Singapore, and Malaysian waters. The area is roughly 3,230,663.98 km^2 depicted in Fig. 3. They were recorded between 1st July 2021 and 28th February 2022 within the area. The dataset is the same as the one used in [11].

A. AIS Data Preprocessing and Trajectory Segmentation

A one-month subset (1st - 31st July 2021) of the dataset is cleaned with the following filters:

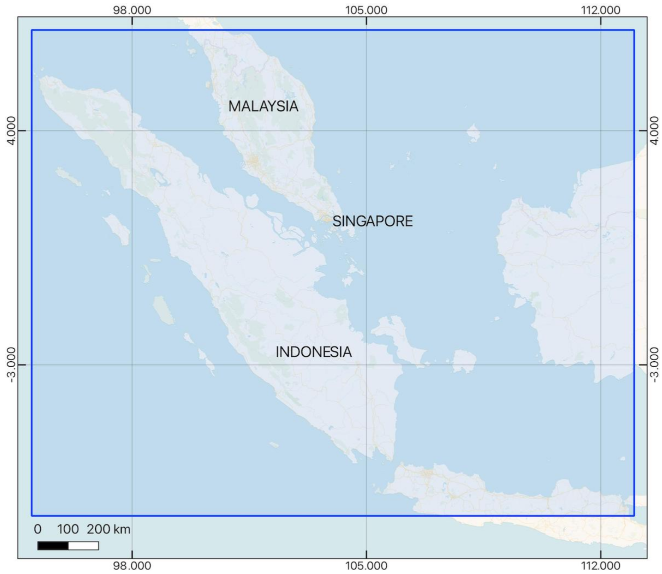


Fig. 3. The area within the blue rectangle is the experiment's area of interest. Basemap © CARTO © OpenMapTiles © OpenStreetMap contributors.

1) $length(MMSI) = 9$,
2) $-90 \leq Latitude \leq 90$,
3) $-180 \leq Longitude \leq 180$, and
4) $70 \leq vesselTypeCode \leq 89$

to retrieve all MMSIs of valid AIS messages of tankers and cargo ships. The $70 \leq vesselTypeCode \leq 89$ refers to the vessels of types cargo and tanker [24].

The filtering collects 6,950 unique MMSIs, of which 4,182 MMSIs are used as the keys to fetch one-month historical AIS data of 1st - 31st July 2021. It consists of 3,514,126 AIS records. The remaining 2,768 MMSIs are the keys to retrieve eight-month historical AIS data between 1st July 2021 and 28th February 2022 that contains 15,955,795 recorded AIS transmissions. Thus, this experiment processes 19,469,921 AIS messages in total.

Following the AIS data preprocessing, the trajectory segmentation procedure is executed on the one-month and eight-month AIS datasets to separate the stopping and underway segments of each ship's trajectory. This process adopts the trajectory segmentation technique presented in [20]. This paper's unexpected trajectory detection algorithm takes only the underway segments of each ship's trajectory as the input. The trajectories of the underway segments are grouped into four types of voyages as follows:

1) Voyages between two different ports: the voyages between the port of Jakarta and Singapore port, and between Port Klang and Singapore.
2) Voyages between a port and sea area: the voyages between Singapore port and the South China Sea.
3) Voyages between a sea area to another sea area: the voyages between the South China Sea and the Indian Ocean.
4) Voyages within a relatively wide area of the sea without stopping at any ports: the voyages within the western part of Indonesian archipelagic waters (Natuna Sea and Java Sea).

B. Detecting the Unexpected Trajectory

Before calculating the trajectory redundancy R and curvature C for the trajectories of the underway segments of each voyage group, the two endpoints (origin and destination point/area) of each voyage group need to be determined. The polygon defining the area of the Singapore Port is publicly available by the Maritime and Port Authority of Singapore [25], while the geometrical boundaries of the port of Jakarta and the Port Klang are defined in reference [11]. In the case of sea area, this experiment uses the geographic boundaries provided by MarineRegions.org [26].

For the first voyage group, the trajectories are filtered to select those that start and end at either the Jakarta or Singapore ports, and those that start and end either at Port Klang or Singapore port. Further, this experiment selects only the trajectories whose track interval $\leq$ 6 hours to avoid processing truncated trajectories. This filter is applied to all voyage groups. The rate of redundancy R and curvature C are calculated within three time-windows W on every trajectory. The time windows are $W1 = 6$ hours, $W2 = 24$ hours, and $W3 = 48$ hours. This process returns a set of trajectory excerpts E of which each excerpt has two attributes: R_{max} and C_{max}. The DBSCAN clustering algorithm is applied to the set of excerpts E. The result is visualized as depicted in Fig. 4 and 5. The experiment produces two unexpected trajectories from the voyages between Jakarta and Singapore ports. The trajectory's excerpt near Singapore Port, labeled Ship-AL, belongs to a container ship with a gross tonnage of 66,280 tons and a dimension of 276 x 40 meters, while the excerpt near the port of Jakarta is of a container ship measuring 161.85 x 25.6 meters.

The same processing procedure is applied to the remaining three voyage groups. The results are visualized in Fig. 6, 7, and 8.

This experiment confirms that DBSCAN clustering does not forcibly classify the dataset into cluster and noise. When noise does not exist, none would be labeled as one. It is observable in the visualization of the trajectories between the South China Sea and the Indian Ocean (Fig. 7). None of the trajectories is classified as noise as all of them seem to follow the common routes.

The implementation of the unexpected trajectory detection algorithm reveals the same ship, labeled Ship-AL, shows unexpected movement on the voyage between Jakarta and Singapore port and the voyage within the Western part of Indonesian archipelagic waters. When the unexpected trajectory algorithm is applied to the ship trajectories individually, it confirms that Ship-AL frequently demonstrates unexpected behaviors during her voyages between July 1st, 2021 to February 28th, 2022. Fig. 9 depicts Ship-AL's trajectories with excerpts indicating the unexpected behaviors. Ship-AL is a container ship of Portuguese nationality measuring 66,280 tons of gross tonnage and 275 x 40 (meters) in dimension. Considering the type and size of the ship, her behavior is definitely not normal. Another finding is the trajectory labeled Ship-MA in the voyage within the Western Indonesian Archipelagic waters. The ship, a crude oil tanker of 105,484 tons (deadweight), was loitering at sea for

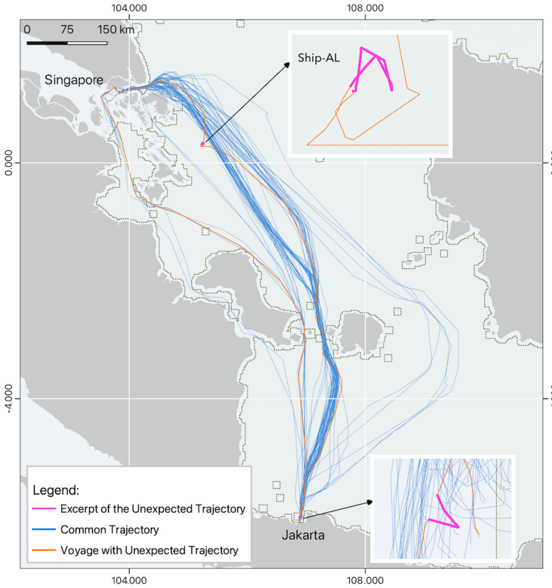


Fig. 4. The trajectories of voyages between Jakarta and Singapore Port. The orange lines illustrate voyages with *unexpected trajectory* whose excerpts are depicted by the magenta lines. The common trajectories of the voyages are indicated with blue lines. The trajectory excerpt labeled Ship-AL belongs to a container ship with a gross tonnage of 66,280 tons measuring 276 x 40 meters in length and width.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

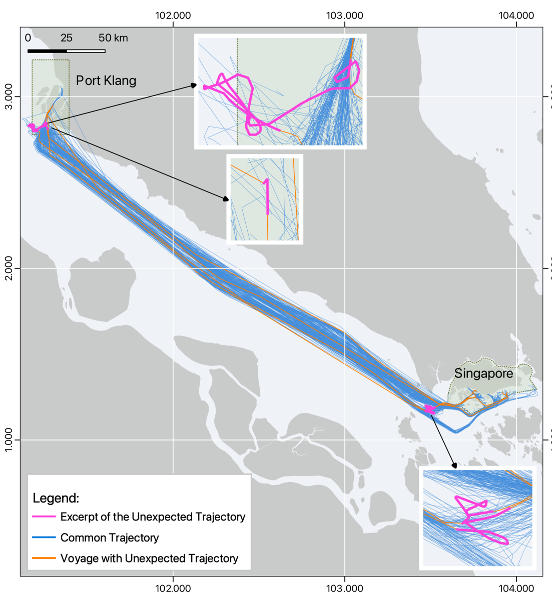


Fig. 5. The trajectories of voyages between Port Klang and Singapore Port. The orange lines illustrate voyages with *unexpected trajectory* whose excerpts are depicted by the magenta lines. The common trajectories are drawn in blue lines.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

155 hours. This type of ship, with a size of 243 x 42 (meters), is not normally engaged in loitering movement [22], which is normal for other types of vessels such as fishing boats, patrol vessels, and SAR (Search and Rescue) ships[16].

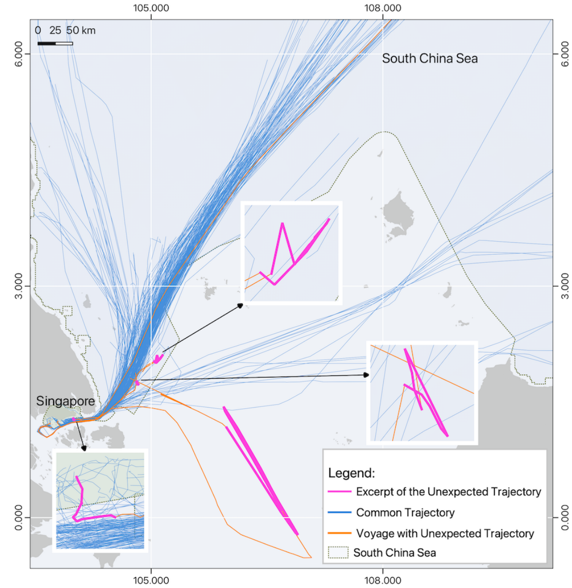


Fig. 6. The trajectories of voyages between the South China Sea and Singapore Port. The orange lines indicate voyages with *unexpected trajectory* whose excerpts are colored magenta. The common trajectories are drawn in blue lines.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

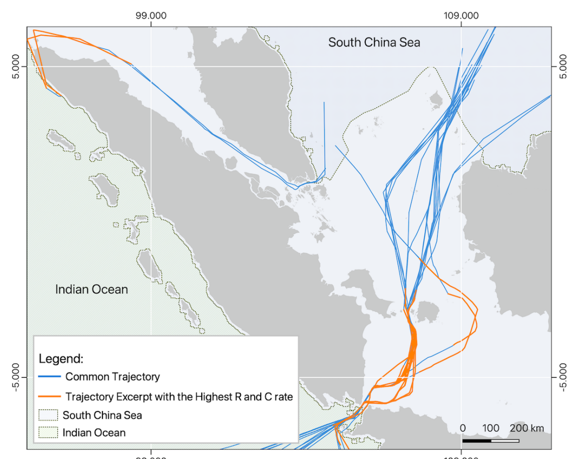


Fig. 7. The trajectories of voyages between the South China Sea and the Indian Ocean. The common trajectories are depicted with blue lines, while the orange lines indicate the excerpt of the trajectories with the highest rate of redundancy R and curvature C. In this case, the DBSCAN clustering returns no trajectory excerpt with a -1 label, meaning that the dataset has no *unexpected trajectory*. The visualization confirms that all voyages seem to follow the common routes.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

C. Evaluation

The experiment results prove the capability of the proposed method to detect unexpected trajectories of vessels navigating through the sea area of interest. To measure the efficacy of the technique, an evaluation is conducted on the same dataset as the evaluation section of reference [22]. The dataset consists of 137 labeled trajectories of vessels navigating through the West Coast of North America. It comprises 24 loitering (anomalous) and 113 normal trajectories. Since the unexpected trajectory

TABLE I. EVALUATION METRICS COMPARISON WITH EXISTING METHODS

Method	***Accuracy***	***Specificity***	***Precision***	***F-score***	***Undetected*****	***False Negative***
TR	0.87	0.84	0.60	0.70	4	14
$F(c)$	0.95	1.0	0.78	0.88	0	7
$F(c,h,d)$	0.93	0.96	0.75	0.84	1	8
$Integrated^{*}$	0.97	0.96	0.89	0.92	1	3
$ProposedMethod$	0.99	0.92	1.00	0.96	2	0

*Weighted integration of $F(c)$ and $F(c,h,d)$.
**The number of undetected loitering ships.

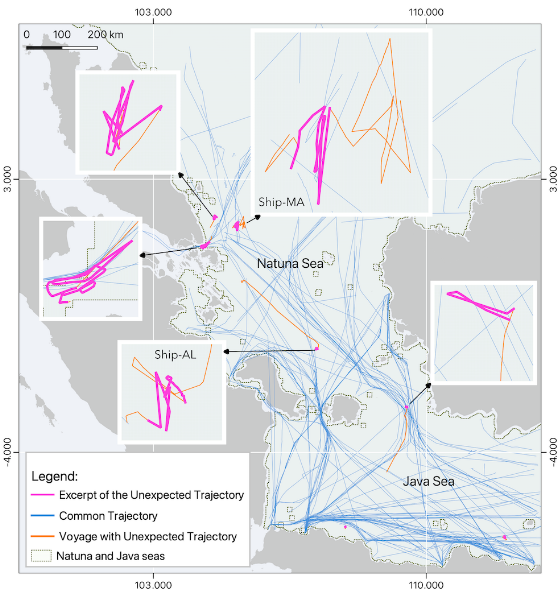


Fig. 8. The trajectories of voyages within the Western Indonesian Archipelagic Waters. The blue lines indicate the common trajectories, while those in orange are the *unexpected trajectories* whose excerpt is in magenta. The ship with an *unexpected trajectory* labeled Ship-AL is also revealed in the voyages between Jakarta and Singapore port. The trajectory labeled Ship-MA is a crude oil tanker loitering for 155 hours.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

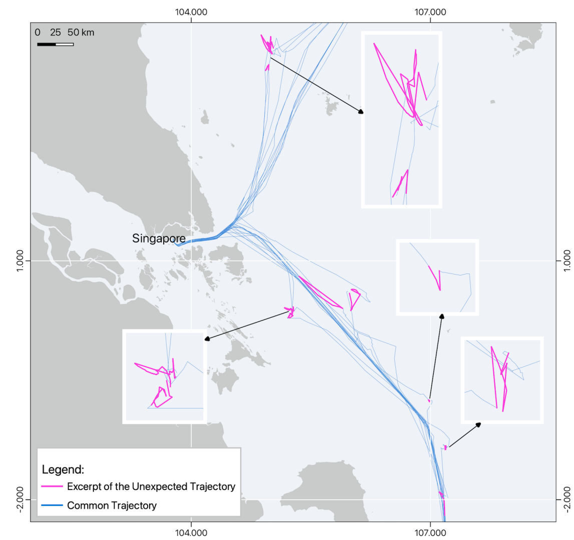


Fig. 9. The trajectories of the Ship-AL between 1st July 2021 and 28th February 2022 that are revealed in the Jakarta-Singapore voyages and the voyages within the Western Indonesian Archipelagic Waters. It is a Portuguese container ship measuring 66,280 tons of gross tonnage and 275 x 40 (meters) dimension.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

could be a trajectory that deviates from the standard routes or that of a loitering behavior, the dataset is valid for this evaluation. The results are compared with the Trajectory Redundancy (TR) calculation formula proposed by Zhang et al. [16] and the loitering detection method of Wijaya and Nakamura [22]. Table I presents the comparison. The proposed unexpected trajectory detection technique outperforms all of the existing methods with 0.99 accuracies, 0.92 specificities, 1.00 precision, and 0.96 F-score. The technique produces no false negatives and merely two undetected loitering trajectories. The prediction results are visualized in Fig. 10. The undetected loitering trajectories possess loitering movements that last less than an hour. It seems that the detection fails because the duration of the loitering movement is too short.

The proposed unexpected trajectory detection technique performs remarkably better in all measurement metrics compared with the existing loitering detection methods. However, the approaches with $F(c)$ and $F(c,h,d)$ formulas return loitering scores. Each detected loitering trajectory is given a loitering score that indicates the severity of the loitering movement. The approach is intended to help maritime authorities to achieve better efficiency in conducting maritime surveillance. It does not only automatically detect loitering ships but also suggests their priority so that the officer in charge can decide which ship to handle first, second, and soon. On the other hand, the unexpected trajectory detection approach proposed in this paper is intended to provide a high-accuracy detection tool without considering the priority of the detected vessels. The result is binary, either normal or unexpected trajectory.

V. CONCLUSION

This paper presents a novel method for detecting unexpected trajectories of vessels based on AIS tracks. The proposed approach leverages the geometrical features of ship trajectories, specifically the rate of redundancy and curvature. It is to identify voyages that deviate from the expected voyage. By applying DBSCAN clustering based on these geometrical features, the method can effectively distinguish between trajectories that follow the common routes and that of the unexpected trajectory. The classification is accomplished without relying on labeled training data or image conversion techniques.

The experimental results demonstrate the efficacy of the proposed method across various types of maritime voyages,

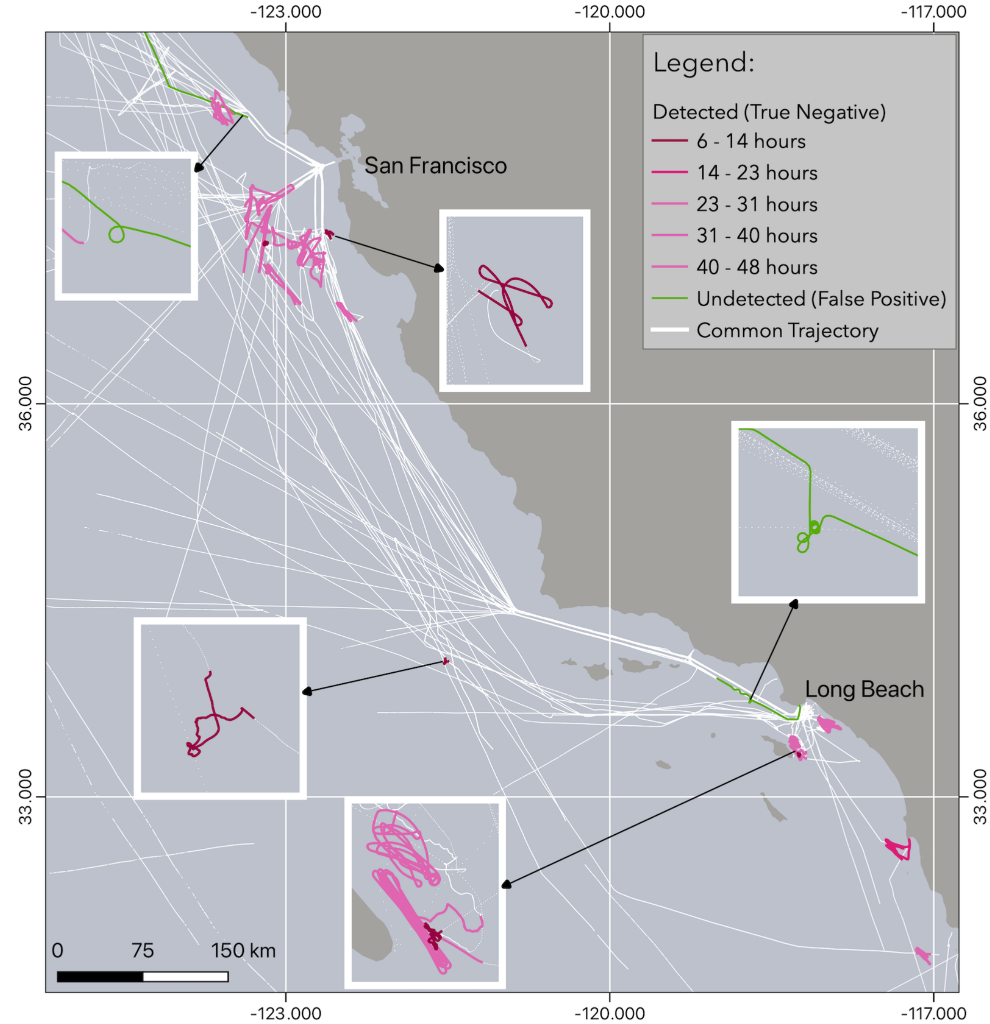


Fig. 10. The purple-red lines indicate the prediction of loitering that matched the actual loitering trajectories (true negative), while the green lines represent the undetected loiterings (false positive). The white lines illustrate the trajectories of the common routes. The prediction achieves 0.99 accuracy, 1.00 precision, and 0.96 F-score with two false positives and zero false negatives.
Basemap adopts geoBoundaries by D. Runfola et al. [27]

including port-to-port, port-to-sea area, and open-water routes. The technique successfully identified several instances of unexpected behavior, including a container ship exhibiting frequent unexpected trajectory and a large oil tanker engaged in prolonged loitering. These findings highlight the method's potential to detect behaviors that may need further investigation by maritime authorities.

The Comparative evaluation against existing approaches shows that the proposed method achieves superior performance across multiple metrics, including accuracy, precision, and F-score. This indicates that the technique offers a robust and reliable means of identifying unexpected trajectory in maritime traffic. The evaluation result confirms that the proposed method is not region-dependent. The evaluation dataset is of the West Coast of North America, while the experiment dataset covers the archipelagos of Indonesia, Malaysia, and Singapore. Despite the proven performance and versatility, the proposed unexpected trajectory detection method possesses several limitations. When it is applied to detect loitering movement, the detection fails if the loitering duration is too short, such as less than an hour. The method is also unable to determine the magnitude of the detected unexpected trajectory, whether it is a slight track deviation due to an instantaneous unplanned maneuver to evade danger or a redundant deviation because of a deliberately planned maneuver.

To further enhance and extend the proposed approach, this study considers the following future works: 1) combining both geometrical and spatiotemporal features to potentially improve detection accuracy and provide a more nuanced characterization of unexpected behaviors, 2) integrating the unexpected trajectory detection method into real-time maritime surveillance systems to evaluate its performance in operational scenarios, and 3) investigating the potential of the approach to detect other types of maritime anomalies.

In conclusion, this study offers a feasible approach for maritime authorities and fleet operators to efficiently monitor vessel voyages and identify potential security, safety, or efficiency concerns. It is the answer to the need for automated surveillance systems because of the increasing threats to maritime security. The technique could substantially contribute to the overall safety and efficiency of maritime transportation.

REFERENCES

[1] The World Bank and S&P Global Market Intelligence, "The Container Port Performance Index 2022: A Comparable Assessment of Performance based on Vessel Time in Port(Fine)," *World Bank Group*, Washington, DC, 2023.

[2] International Maritime Organization (IMO), *Resolution A.893(21) guidelines for voyage planning*, London: International Maritime Organization, Feb 2000.

[3] C. Gamage, R. Dinalankara, J. Samarabandu, A. Subasinghe, "A comprehensive survey on the applications of machine learning techniques on maritime surveillance to detect abnormal maritime vessel behaviors," *WMU Journal of Maritime Affairs*, vol. 22, pp. 447–477, Jun 2023.

[4] Federal Communications Commission (2022, Oct. 4). *Maritime Mobile Service Identities - MMSI*[Online]. Available: https://www.fcc.gov/wireless/bureau-divisions/mobility-division/maritime-mobile/ship-radio-stations/maritime-mobile

[5] International Maritime Organization (2015, Dec. 15). *esolution A.1106(29) Revised Guidelines for the Onboard Operational Use of Shipborne Automatic Identification Systems(AIS)*[Online]. Available: https://wwwcdn.imo.org/localresources/en/OurWork/Safety/Documents/AIS/Resolution%20A.1106(29).pdf

[6] C.V. Ribeiro, A. Paes, D.d. Oliveira, "AIS-based maritime anomaly traffic detection: A review," *Expert System with Applications*, vol. 231, 120561, Nov 2023.

[7] I. Kontopoulos, A. Makris, K. Tserpes, "A Deep Learning Streaming Methodology for Trajectory Classification," *ISPRS Int. J. Geo-Inf.* vol 10, 4, Apr 2021.

[8] C. Zhang, S. Liu, M. Guo, Y. Liu, "A novel ship trajectory clustering analysis and anomaly detection method based on AIS data," *Ocean Engineering*, vol. 288, 2, 116082, Nov 2023.

[9] D. Luo, P. Chen, J. Yang, X. Li, Y. Zhao, "A New Classification Method for Ship Trajectories Based on AIS Data," *J. Mar. Sci. Eng.*, vol. 11, 9, 1646, Aug 2023.

[10] L. Chen, D. Zhang, X. Ma, L. Wang, S. Li, Z. Wu, G. Pan, "Container Port Performance Measurement and Comparison Leveraging Ship GPS Traces and Maritime Open Data," *IEEE Transactions on Intelligent Transportation System*, vol. 17, 5, pp. 1227-1242, May 2016.

[11] W. M. Wijaya, Y. Nakamura, "Port Performance Indicators Construction based on the AIS-generated Trajectory Segmentation and Classification," Apr 2024, PREPRINT (Version 1) available at Research Square https://doi.org/10.21203/rs.3.rs-4195656/v1.

[12] D.A. Cerderio, A. Komaromi, Y. Liu, M. Saeed, "World Seaborne Trade in Real Time: A Proof of Concept for Building AIS-based Nowcasts from Scratch," *IMF Working Paper*, May 2020.

[13] UNCTAD, "Review of Maritime Transport 2023 - Towards a green and just transition," *United Nations Publications*, New York, 2023.

[14] J.H. Ford, D. Peel, B.D. Hardesty, U. Rosebrock, C. Wilcox, "Loitering with intent–Catching the outlier vessels at sea," *PLoS ONE*, vol. 13, 7, Jul 2018.

[15] E.N. de Souza, K. Boerder, S. Matwin, B. Worm, "Improving Fishing Pattern Detection from Satellite AIS Using Data Mining and Machine Learning," *PLoS ONE*, vol. 11, 9, Jul 2016.

[16] Z. Zhang, L. Huang, X. Peng, Y. Wen, L. Song, "Loitering behavior detection and classification of vessel movements based on trajectory shape and convolutional neural networks," *Ocean Engineering*, vol. 258, 111852, Aug 2022.

[17] T. Guo, L. Xie, "Research on Ship Trajectory Classification Based on a Deep Convolutional Neural Network," *J. Mar. Sci. Eng.*, vol 10, 5, Apr 2022.

[18] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 226–231. AAAI Press, Washington, D.C., Aug 1996.

[19] Z. Yan, L. Cheng, R. He, H. Yang, "Extracting ship stopping information from AIS data," *Ocean Engineering*, vol. 250, 111004, Apr 2022.

[20] W.M. Wijaya, Y. Nakamura, "Ship Navigational Status Classification based on the Geometrical and Spatiotemporal Features of the AIS-generated Trajectory," *2024 the 9th International Conference on Big Data Analytics*, pp. 103-112, Mar 2024.

[21] R.O. Lane, D.A. Nevell, S.D. Hayward, T.W. Beaney, "Maritime anomaly detection and threat assessment," *2010 13th International Conference on Information Fusion*, 2010.

[22] W.M. Wijaya, Y. Nakamura, "Loitering behavior detection by spatiotemporal characteristics quantification based on the dynamic features of Automatic Identification System (AIS) messages," *PeerJ Computer Science*, 9:e1572, Sep 2023.

[23] M. Christ, N. Braun, J. Neuffer, A.W. Kempa-Liehr, "Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh - A Python package)," *Neurocomputing*, vol. 307, pp. 72-77, Sep 2018.

[24] Spire. *Ship Type mappings*[Online]. Available: https://documentation.spire.com/ais-fundamentals/ship-type-mappings/

[25] Maritime and Port Authority of Singapore, "Maritime and Port Authority of Singapore (Port) Regulations, pp. 76-79," *Singapore Statutes Online*, Singapore, 2024.

[26] VLIZ, "Marine Gazetteer geographic name search," *MarineRegions.Org*, 2024.

[27] D. Runfola et al., "geoBoundaries: A global database of political administrative boundaries," *PLoS ONE*, vol. 15, 4, e0231866, Apr 2020.