Incorporating Haversine Formula into DBSCAN for AIS Data-Based Ship Encounter Detection

I Made Dwi Putra Asana, I Made Oka Widyantara, Linawati, Dewa Made Wiharta

Abstract—**Traditional Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method relying on Euclidean distance is often misrepresenting maritime spatial relationships, leading to inaccurate clustering. Therefore, this study aimed to integrate Haversine formula with DBSCAN algorithm to enhance ship encounter detection using Automatic Identification System (AIS) data. The application of Haversine formula ensured more precise geodesic distance measurement to improve clustering accuracy in ship encounter detection and collision risk assessment. The results showed that clustering evaluation using Silhouette Index and Davies-Bouldin Index (DBI) confirmed the superiority of DBSCAN-Haversine over DBSCAN-Euclidean in structuring well-separated clusters. Specifically, silhouette scores ranged from 0.240 to 0.310 and DBI values were between 1.037 and 1.335. An event-based evaluation also validated DBSCAN-Haversine by simulating ship movements and computing Collision Risk Index (CRI), showing the capability to detect high-risk encounter that DBSCAN-Euclidean misclassified as outliers. This study showed the importance of compliance with International Regulations for Preventing Collisions at Sea (COLREGs) in ship encounter scenarios. The results showed that DBSCAN-Haversine provided a more reliable method for early warning systems and maritime traffic management, ensuring safer navigation in dense ship traffic regions.**

Index Terms— **Automatic Identification System, Ship Encounter Detection, Haversine Formula, DBSCAN, Collision Risk Assessment**

I. INTRODUCTION

SHIP accidents in Indonesian waters are one of the main challenges in maintaining maritime safety. A significant effort that has been made to reduce these risks is through the mandatory use and activation of Automatic Identification System (AIS), as stipulated in the Minister of Transportation Regulation Number PM 7 of 2019. AIS plays a significant role in providing essential navigation data to monitor ship movements and support maritime safety control. The system automatically transmits dynamic data, such as ship location, speed, and course, as well as static data, including name and identification [1]. Additionally, the extensive coverage of AIS enables more comprehensive monitoring of ship activities in Indonesian waters. The real-time data transmission also enhances the effectiveness of maritime surveillance and safety control.

A critical element in ship collision risk analysis is Collision Risk Index (CRI). This index serves as an indicator of the potential for ship collision based on the interaction between two or more ship at sea. CRI ranging from 0 to 1 helps in evaluating the level of collision risk, where higher values indicate greater potential risk [2], [3]. The initial step during calculation is the detection of ship encounter, where traditional methods such as sliding windows are used. However, several challenges have been faced in determining the correct window size. Methods that are extremely small or large can cause inaccurate detection [4], [5]. In this context, Density-based clustering methods, such as Density-Based Spatial Clustering of Applications with Noise (DBSCAN), are a relevant option for addressing the issues. DBSCAN clusters ship that is in a certain distance from each other, showing the potential for encounter [6]. Many studies have used traditional DBSCAN methods with Euclidean distance metric, which may not be sufficiently effective in managing data with complex patterns such as AIS data [7], [8]. Regarding geospatial data such as latitude and longitude, Euclidean distance has limitations due to the inability to consider the curvature of the Earth, leading to inaccurate distance measurements [9].
To address the limitations of traditional methods, this study proposed the integration of Haversine formula into DBSCAN clustering process to improve the accuracy of ship encounter detection. Haversine formula has been proven effective in measuring geodesic distances in various geospatial applications, specifically in the context of navigation and accident detection [10]. This formula considers the curvature of the Earth, leading to more precise distance measurements, which are crucial in ship trajectory analysis and collision risk evaluation [11], [12]. Therefore, the application of Haversine formula in DBSCAN is expected to provide more accurate results in detecting ship encounter and improving collision risk evaluation based on AIS data.

II. RELATED WORK

The use of AIS data has revolutionized maritime safety and traffic management by enabling real-time tracking and analysis of ship positions, speeds, and courses. AIS provides

Manuscript received December 3, 2024; revised August 14, 2025.
This research was supported and funded by the Ministry of Education, Culture, Research and Technology and Udayana University with a contract number: B/519-20/UN14.4.A/PT.01.03/2024.
I Made Dwi Putra Asana is a doctoral student of Engineering Science Department, Udayana University, Badung, Indonesia (e-mail: dwiputraasana@student.unud.ac.id).
I Made Oka Widyantara is a professor of Electrical Engineering Department, Udayana University, Badung, Indonesia (corresponding author to provide phone: 62813-3849-9965; fax: 62361-701812; e-mail: oka.widyantara@unud.ac.id).
Linawati is a professor of Electrical Engineering Department, Udayana University, Badung, Indonesia (e-mail: linawati@unud.ac.id).
Dewa Made Wiharta is an associate professor of Electrical Engineering Department, Udayana University, Badung, Indonesia (e-mail: wiharta@unud.ac.id).

comprehensive coverage of maritime traffic, facilitating enhanced situational awareness and risk assessment. Furthermore, the integration of satellite AIS (SAIS) has extended ship tracking capabilities beyond coastal regions, offering critical insights into offshore collision risk, including marine wildlife [13]. These data are essential in collision risk modeling. For example, Pan [14] showed that using AIS data in assessing collision probabilities between ship and structures like bridges improved ship traffic management and waterway design. Bakdi et al. [15] also used AIS data to identify risk associated with ship interactions, including collision with fixed offshore platforms. This shows the significant role of AIS in developing robust models to predict and mitigate maritime accidents.

The application of DBSCAN algorithm in maritime navigation and ship encounter detection has gained significant traction. This is because of the ability to identify clusters of varying shapes and sizes without requiring a predefined number of clusters. The characteristics are particularly advantageous in analyzing complex maritime traffic patterns. In this context, several studies have reported the effectiveness of DBSCAN in identifying multiship encounter and detecting potential collision risk in dense maritime environments. Zhen and Shi [16] used DBSCAN to detect ship encounter in close proximity in surveillance waters, facilitating more accurate identification of potential collision scenarios. Additionally, DBSCAN has been used for trajectory clustering to analyze ship movement patterns, which is valuable for optimizing maritime traffic management. Widyantara et al. [17] showed the capability of DBSCAN to automatically determine the number of clusters based on data density, enhancing the analysis of ship behaviors. Algorithm resilience to noise and adaptability to irregular datasets, standard in AIS data, enhances the application in anomaly detection and situational awareness [18], [19], [20].

The integration of Haversine formula with DBSCAN algorithm represents a significant advancement in geospatial and maritime applications, particularly for clustering ship trajectories and detecting encounter. Haversine formula accurately calculates the great-circle distance between two points on the Earth's surface based on latitude and longitude. This analysis effectively accounts for the Earth's curvature, an aspect that traditional distance metrics such as Euclidean distance fail to consider. The accuracy is essential in maritime for safety, where ship travels over large distance. Sharmila and Sabarish [21] showed that using Haversine formula in DBSCAN significantly enhanced clustering performance for spatial trajectory data, providing more accurate results compared to Euclidean and Hausdorff methods. Several studies have shown the effectiveness of Haversine formula in improving the accuracy of distance-based clustering algorithm for geospatial applications. Roberts-Licklider [22] compared various distance metrics, including Haversine, Euclidean, and Manhattan methods, in optimizing treatment facility regions. The results showed the advantages of using Haversine formula for geographic data, indicating the effectiveness in enhancing clustering accuracy. Additionally, Wells and Kumar [23] applied Haversine formula in air traffic management to convert latitude and longitude points into straight-line distances, showing the versatility and importance of ensuring accurate distance calculations for safety across various domains.

Based on the literature, Haversine formula is proven to have the ability to measure distance considering the curvature of the Earth, serving as a relevant and essential tool in distance-based ship encounter detection as well as clustering using DBSCAN with AIS data. The ability of this formula to provide accurate distance calculations, specifically at long distances, ensures that the influence of the Earth's curvature is considered, which is particularly important in a maritime context. Therefore, the application of Haversine formula in DBSCAN offers an effective methodological solution for ship encounter detection and collision risk evaluation, using AIS data more accurately.

III. METHOD

A. Data

In this study, AIS data used were obtained by the remote-based station (RBS) at the Faculty of Engineering, Udayana University. This dataset included the maritime region of ALKI II, which comprised the Lombok Strait and the surroundings. Attributes extracted and analyzed from AIS data included Message Timestamp, MMSI, Latitude, Longitude, Heading, Course Over Ground (COG), and Speed Over Ground (SOG). The Message Timestamp showed the time at which AIS message was transmitted by ship. Latitude and Longitude provided the geographic coordinates of ship location, while Heading indicated directional orientation in degrees. COG represents the actual travel direction of ship, and SOG measures the ship's true movement speed over water in knots. This study also incorporated a comparative analysis [7], using haversine distance formula to enhance DBSCAN clustering results, compared with the use of Euclidean distance in the previous study. The comparative method served to validate the improvements achieved in clustering outcomes in AIS data analysis.

B. Proposed Method

The proposed method in this study incorporates the Haversine formula into the DBSCAN clustering algorithm to enhance the accuracy of ship encounter detection using AIS data, as illustrated in Fig. 1. The system, developed in Python, follows a structured sequence comprising AIS data acquisition, preprocessing, and geodesic distance calculation to ensure accurate spatial representation. By accounting for the Earth's curvature, the Haversine formula enables precise distance measurements, which are crucial in maritime navigation. These geodesic distances are then utilized in the DBSCAN algorithm to detect potential ship encounters based on spatial proximity. In the final stage, the CRI is computed to quantify the severity of each identified encounter. By integrating spatial accuracy with density-based clustering, the proposed method enables reliable detection of ship interactions and supports proactive maritime risk assessment, contributing to improved situational awareness and safety in high-traffic sea regions.

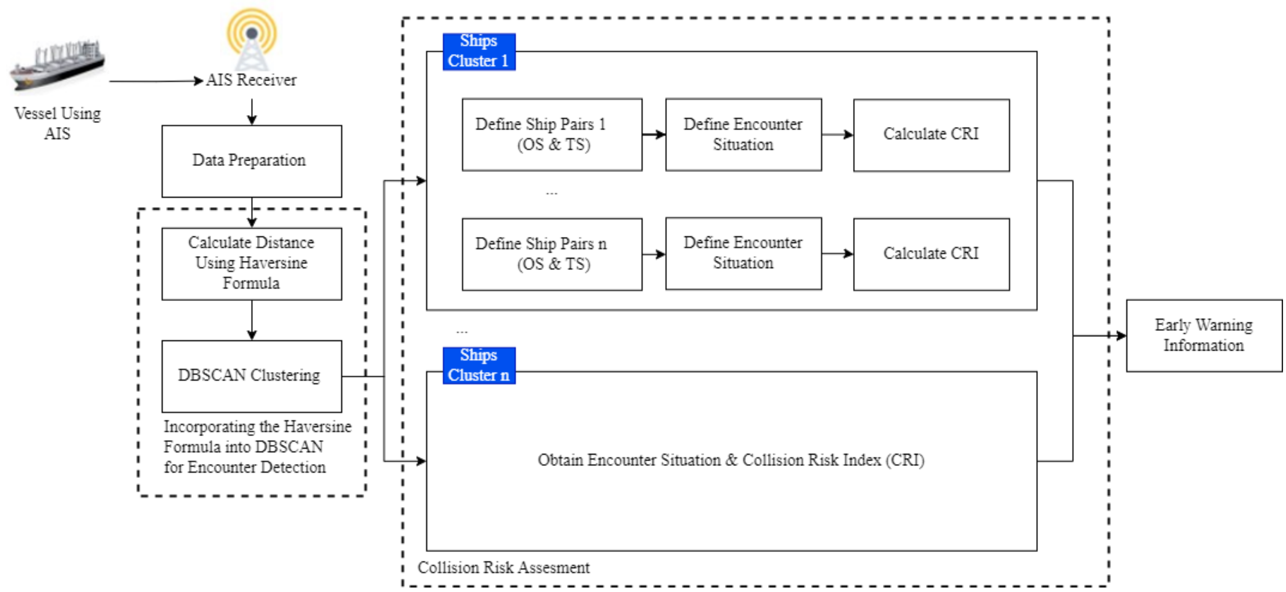


Fig. 1. Proposed method.

The proposed method in Fig. 1 shows a structured framework for maritime collision risk assessment using AIS data, integrating ship encounter detection with DBSCAN clustering and Collision Risk Index (CRI) computation. AIS data are collected in real time from Remote-Based Station (RBS) devices, with NMEA message decoding to extract static and dynamic ship attributes. In the encounter detection stage, inter-ship distances are computed using the Haversine formula, ensuring accurate estimation by considering the Earth's curvature. These geodesic distances are then utilized in DBSCAN clustering to group nearby ships as potential encounters. The main contribution of this study is the integration of the Haversine formula into DBSCAN to enhance detection accuracy. During risk assessment, ships in each cluster are designated as own ship (OS) and target ship (TS), and encounter situations are determined based on relative bearings. CRI values are subsequently calculated for each OS–TS pair, incorporating relative distance, Time to Closest Point of Approach (TCPA), and Distance to Closest Point of Approach (DCPA). This systematic approach enables early warning issuance when collision threats are detected, supporting proactive navigation and improving maritime safety through precise spatial analysis and quantitative risk evaluation.

C. Incorporating Haversine Formula to DBSCAN for Encounter Detection

DBSCAN algorithm is a robust, density-based clustering method commonly used to identify arbitrary-shaped clusters in spatial data [24], [25]. It is particularly effective in applications where dataset contains noise and the number of clusters is previously determined. The algorithm operates by exploring dataset, categorizing each point as a core, border, or noise point. In comparison, core point has a minimum number of other points ( ) in a specified radius ( ). Border point has fewer in but is in the neighborhood of a core point. Meanwhile, noise point is neither core nor border [26]. The core mechanism of DBSCAN can be described through Equations (1) and (2).

$$|N_\varepsilon(p)| \geq MinPts \tag{1}$$

Where $N_\varepsilon(p)$ is the ε-neighborhood of p, defined as :

$$N_\varepsilon(p) = \{q \in D | distance(p,q) \leq \varepsilon \} \tag{2}$$

D is the dataset containing all points. A point q is a border point when it is not a core point but is in the neighborhood of a core point. Point that is neither core nor border is classified as noise, which does not belong to any cluster and is typically considered outliers in the dataset.

Integrating the Haversine formula into DBSCAN is crucial for accurate maritime geospatial distance measurement. The formula computes great-circle distances between two points on a sphere using their longitudes and latitudes, ensuring precise geodesic calculations [27]. This adaptation is vital for defining the ε parameter in DBSCAN when applied to geographic coordinates. Equation (3) illustrates the Haversine formula's application in determining distances across the Earth's curved surface with improved spatial accuracy [28].

$$d = 2r\, arcsin\left(\sqrt{sin^2\left(\frac{x_1-x_2}{2}\right) + cos(x_1)\, cos(x_1)\, sin^2\left(\frac{y_1+y_2}{2}\right)}\right) \tag{3}$$

Where x_1,y_1 and x_2,y_2 are the latitudes and longitudes of two points in radians, and r is the Earth's radius, the Haversine formula accurately calculates the shortest distance over the Earth's surface. Its simplicity and precision make it essential for geographic distance estimation. Integrating Haversine into DBSCAN replaces the distance calculation in Equation (2) with the Haversine distance in Equation (4), redefining the ε-neighborhood for geospatial datasets. Algorithm 1 outlines this integration, detailing steps to detect ship encounters using spatial proximity and geodesic distance within AIS-derived datasets, thereby enhancing maritime encounter detection accuracy.

$$N_\varepsilon(p) = \{q \in D | Haversine(p,q) \leq \varepsilon \} \tag{4}$$

Algorithm 1. DBSCAN Clustering with Haversine Distance.

1 **Input**:
X: Array of geographic coordinates $[latitude, longitude]$
eps: Radius of neighborhood (in nautical mile)
$min\_samples$: Minimum number of points required to form a dense region
Output:
$clusters$: Array of cluster labels for each point in X
Definitions:
$Haversine\_Distance(lat1, lon1, lat2, lon2)$: Computes the great-circle distance between two points on the Earth's surface via Eq. (9).
Procedure:
Convert Coordinates: For each point in X, convert latitude and longitude from degrees to radians.
2 Initialize Distance Matrix:
Let num_points be the number of points in X.
Initialize $dist\_matrix$ as num_points X num_points matrix filled with zeros.
3 Compute Distances:
For $i = 0$ to $num\_points - 1$
  For $j = i + 1$ to num_points
    $dist \leftarrow Haversine\_Distance(X[i,1], X[i,0], X[j,1], X[j,0])$.
    $dist\_matrix[i,j] = dist\_matrix[i,j] = dist$.
  End For
End For
4 Apply DBSCAN:
Initialize all points as 'unclassified'.
For each point i:
  If point i is 'unclassified':
    $Neighbors \leftarrow$ retrieve all points within eps from point i in $dist\_matrix$
    If $|Neighbors| < min\_samples$:
      Label point i as 'noise'.
    Else:
      Label point i with the new cluster ID.
      For each point j in $Neighbors$:
        If point j in 'noise':
          Label point j with the new cluster ID.
        If point j is 'unclassified':
          Label point j with the new cluster ID.
          $Neighbors_j \leftarrow$ retrieve all points within eps from point j.
          If $|Neighbors_j| \geq min\_samples$:
            Add $Neighbors_j$ to $Neighbors$.
          End If
        End If
      End For
    End If
  End If
  Increment the cluster ID.
End For
5 **Return** $clusters$

D. *Encounter Situation*

International Regulations for Preventing Collision at Sea (COLREGs), established by the International Maritime Organization (IMO) in 1972, are designed to enhance maritime safety by delineating protocols to mitigate ship collisions. These comprehensive regulations specify navigational guidelines, including protocols for ship right of way, overtaking procedures, as well as the roles and responsibilities of ship in various maritime scenarios [29]. The five main rules of COLREGs related to collision avoidance are Rule 13 – Overtaking, Rule 14 - Head-on Situation, Rule 15 - Crossing Situation, Rule 16 - Action by Give-way Ship, Rule 17 - Action by Stand-on Ship [30]. Generally, the calculation of encounter situation includes determining the relative bearing angle (φ) between OS and TS to assess the risk of collision and plan avoidance maneuvers [31]. Ship encounter is typically classified into three distinct types, namely head-on, overtaking, and crossing, which are differentiated by the relative directions. Specifically, crossing encounter is further categorized into two subtypes based on the position from which the give-way ship methods, either crossing behind (from the stern) or in front of (from the bow) [32], [33]. Table 1 shows ship encounter categories and relative direction criteria.

TABLE I
SHIP ENCOUNTER SITUATION CATEGORIES

Encounter situation	Criteria
Overtaking	$\varphi \leq 25$
Head On	$165 \leq \varphi \leq 195$
Crossing give way ship passing at bow	$25 < \varphi < 165$ or $195 < \varphi < 335$ $\alpha \leq 90$ or $\alpha \geq 270$
Crossing give way ship passing at stern	$25 < \varphi < 165$ or $195 < \varphi < 335$ $\alpha \leq 90$ or $\alpha \geq 270$

E. *Collision Risk Index*

CRI quantifies the likelihood of maritime collisions, assigning values between 0 and 1, where higher values signify increased risk [34]. This index is integral to collision avoidance system, facilitating timely interventions when thresholds are exceeded [35]. Additionally, CRI incorporates multiple parameters such as Distance to Closest Point of Approach (DCPA), Time to Closest Point of Approach (TCPA), relative distance, bearing angles, and speed ratios of included ship. In maritime safety assessments, DCPA and TCPA are integral in the evaluation of potential collision risks between ships. DCPA represents the minimum predicted distance between two converging ships, while TCPA indicates the estimated time until closest method occurs [36]. In this study, the parameters used are relative distance (D), DCPA, and TCPA referring to the basic CRI formula [37], [38] as shown in Equation (5).

$$CRI = \left(a_1 \left(\frac{DCPA}{Ds}\right)^2 + a_2 \left(\frac{TCPA}{Ts}\right)^2 + a_3 \left(\frac{D}{Ds}\right)^2 \right)^{-\frac{1}{2}} \tag{5}$$

Where Ds denotes the minimal safe distance considered necessary for proactive navigation, while Ts encapsulates the time required for detecting collision risks, making decisive avoidance maneuvers, and steering adjustments. The coefficients a_1, a_2, and a_3 serve as weighted factors, which prioritize the significance of environmental visibility, ship dimensions, and the prevailing conditions of the navigational waters. The relative distance between ship at any given moment is determined using Equation (6), while DCPA and TCPA are quantified by Equations (7) and (8), respectively [39], [40].

$$D = \sqrt{(xTS - xOS)^2 - (yTS - yOS)^2} \tag{6}$$

Where x and y denote the coordinates of the target ship (TS) and own ship (OS), respectively.

$$DCPA = |R \sin(\varphi_R - \alpha_T - \pi)| \tag{7}$$

$$TCPA = \left| \frac{R \cos(\varphi_R - \alpha_T - \pi)}{v_R} \right| \tag{8}$$

Where φ_R represents the relative bearing, α_T the target ship's course angle, and v_R is the relative speed.

F. Encounter Detection Evaluation

Cluster evaluation metrics are essential in assessing the quality of clustering results [41], particularly in applications including ship encounter detection, where the accuracy of groupings has direct implications for maritime safety. In this study, two widely recognized cluster validation indexes, Silhouette Index and Davies-Bouldin Index (DBI), are used to measure clustering effectiveness. These indexes assess both cluster cohesion and separation, which are fundamental in determining whether ship clusters represent actual navigational interactions. High cohesion and well-defined separation suggest that ship encounter is correctly identified, enabling more precise collision risk analysis and supporting reliable early-warning system in maritime traffic monitoring.

Silhouette Index

Silhouette Index measures how well each data point fits in the assigned cluster by evaluating cohesion (similarity of data points in cluster) and separation (distinction between clusters) [42], [43]. In this study, Silhouette coefficient for a data point i is defined in Equation (9).

$$S(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{9}$$

Where $a(i)$ is the average distance between data point i and all other points in the same cluster, showing cohesion. $b(i)$ denotes the average distance from point i to all points in the nearest neighboring cluster, indicating separation.

Silhouette score, defined in Equation (9), evaluates the compactness and separation of clusters, with values ranging from -1 to 1. Values close to 1 indicate that data points are well-matched to their cluster and poorly related to neighboring cluster. Meanwhile, values close to 0 suggest overlapping cluster and negative scores signal potential misclassification. This suggests that a higher Silhouette score represents a stronger and more coherent clustering structure.

Davies-Bouldin Index (DBI)

DBI evaluates clustering performance based on the compactness of each cluster and the degree of separation [44]. Meanwhile, DBI formula is shown in Equation (10).

$$DBI = \frac{1}{N} \sum_{i=1}^{N} \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d_{ij}} \right) \tag{10}$$

Where N is the number of clusters. σ_i and σ_j represent the average intra-cluster distances (compactness) for cluster i and j, respectively. d_ij is the distance between cluster centroids i and j (separation). A lower DBI value suggests better clustering, indicating compactness and well-separation.

Regarding ship encounter detection, Silhouette Index and DBI are essential in determining the reliability of clustering structure. A high Silhouette score ensures that ship encounter is well-grouped and distinct from non-encountering types. Meanwhile, a low DBI value confirms that identified cluster is compact and well-separated, reducing noise and false positives in ship encounter analysis.

IV. Result and Discussion

A. Ship Encounter Detection Using DBSCAN with Haversine Formula

This section presents the results of ship encounter detection using the DBSCAN clustering algorithm, enhanced by the Haversine formula to account for geodesic distance. AIS data from vessels navigating the Lombok Strait—a region characterized by high-density maritime traffic—was selected as the experimental dataset. As illustrated in Fig. 2, the proposed detection framework consists of several key stages: AIS data preprocessing, distance computation using the Haversine formula, spatial clustering using DBSCAN, and post-clustering evaluation. This approach enables the identification of meaningful ship encounter groups with improved accuracy, facilitating more reliable maritime collision risk assessments.

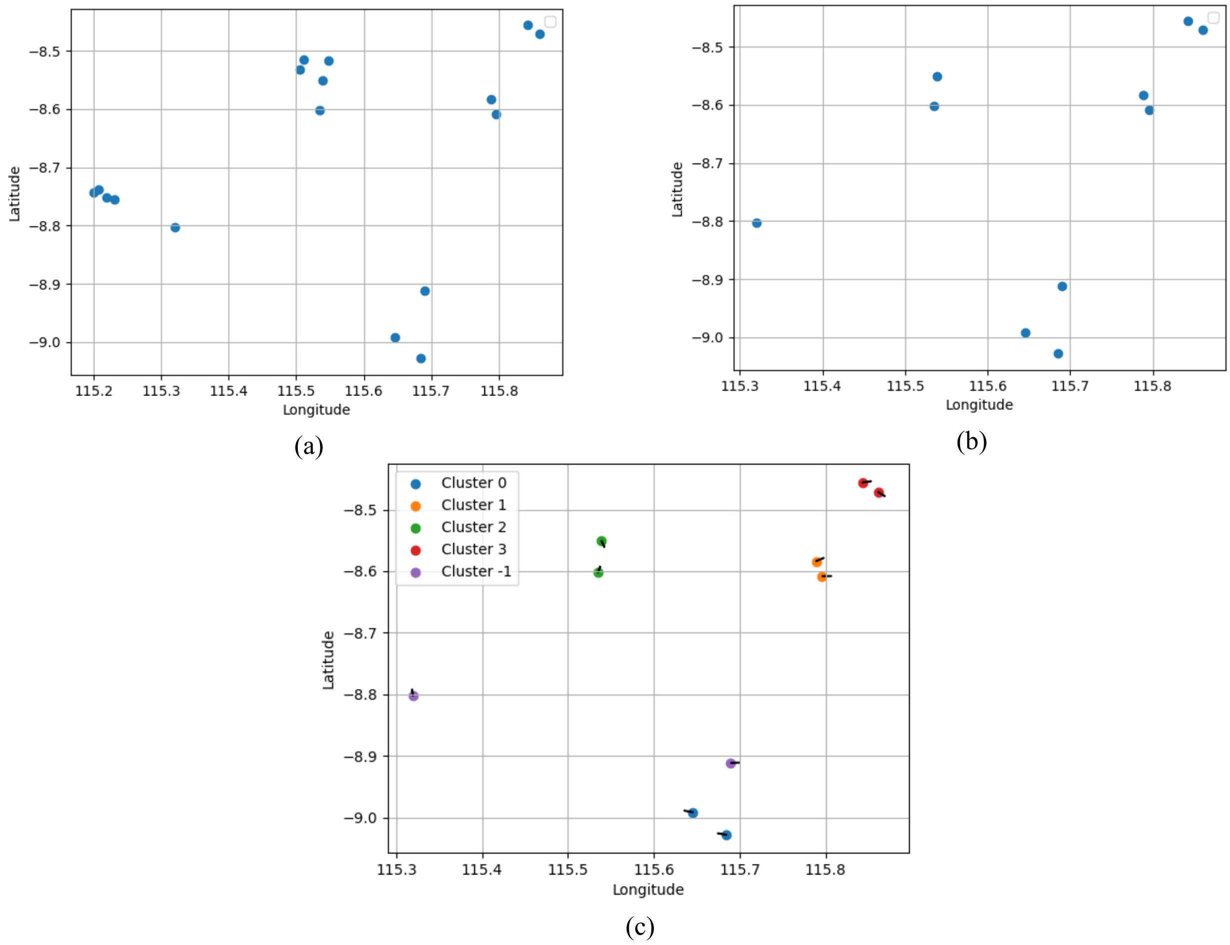


Fig. 2. Ship encounter detection results using DBSCAN with Haversine formula (a) AIS data received (b) AIS data after cleaning (c) clustering result using DBSCAN-Haversine.

Fig. 2 shows the results of decoding AIS data received at RBS (Fig. 2 (a)), cleaning the data (Fig. 2 (b)), and clustering using DBSCAN and Haversine formula (Fig. 2 (c)). The data cleaning stage included removing records of ship with a speed of zero and port regions, thereby focusing the analysis exclusively on active ship navigating in open waters. The clustering of AIS data was conducted after cleaning, leading to four distinct clusters. These included cluster 0, 1, 2, and 3, as shown in Fig. 2(c). Ship that did not meet the clustering criteria were classified as outlier and grouped into cluster "-1". Each identified cluster represented encounter scenario, where ship is grouped based on closeness as a predefined distance threshold. By measuring in nautical miles, this threshold ensures that ship in proximity is flagged as potential encounter, facilitating the detection of potential maritime interactions.

Table 2 presents the experimental results of ship encounter detection using DBSCAN-Haversine with varying minimum epsilon values, expressed in nautical miles (nm). A nautical mile—equivalent to one minute of latitude or approximately 1.852 kilometers—is the standard unit for maritime distance measurement due to its alignment with the Earth's spherical geometry. This unit ensures geodetic consistency in evaluating spatial proximity between ship. The use of multiple epsilon thresholds allows for the assessment of clustering sensitivity and outlier identification.

The experimental results in Table 2 show the effects of varying minimum epsilon values on the number of cluster, outliers, and cluster membership size in ship encounter detection. According to maritime safety principles [7], the clustering process must meet certain conditions, (1) each cluster must contain a minimum of two ships, (2) the number of noise points (outliers) should be minimized, and (3) the number of ships in each cluster should remain in a reasonable limit to avoid misidentifying potential encounter.

At a minimum epsilon of 3 nm, the number of cluster ranges from 2 to 4, with each containing two to three ships. This suggests that a 3 nm epsilon effectively separates ship into distinct clusters, showing realistic encounter scenarios of close proximity. When the epsilon increases to 4 or 5 nm, the number of clusters decreases while the size of the largest cluster increases. The result indicates that ship at greater distances is grouped together as part of the same encounter. However, there is a reduction in the granularity needed for detecting smaller-scale encounter. Studies in maritime safety emphasize the importance of maintaining a distance of 1 to 2 nautical miles between ships to prevent collisions, particularly for smaller sizes used during fishing activities [40]. In adverse weather or poor visibility, a minimum distance of 2 nautical miles is recommended to ensure safe maneuvering [41]. Therefore, the selection of 3 nautical miles as the minimum epsilon is both practical and in line with the established safety standards, ensuring encounter detection is both accurate and adheres to recognized safe distance guidelines.

TABLE II
CLUSTERING RESULTS USING DBSCAN-HAVERSINE WITH VARIOUS MINIMUM EPSILON VALUES

Minimum epsilon / Data	3 nm			4 nm			5 nm		
	Number of clusters	Number of most members in the cluster	Number of outliers	Number of clusters	Number of most members in the cluster	Number of outliers	Number of Cluster	Number of most members in the cluster	Number of Outlier
2023-02-05 00-08-30	4	2	2	4	3	1	3	4	1
2023-02-05 00-10-13	4	2	2	4	2	2	3	5	1
2023-02-03 00-21-22	3	2	1	2	5	0	2	5	0
2023-02-05 02-32-36	3	3	2	4	3	0	3	5	0
2023-02-02 20-17-04	2	2	3	2	2	3	2	2	3
2023-02-03 00-00-03	2	3	1	2	4	0	2	4	0
2023-02-04 08-00-04	2	2	3	2	3	2	2	4	1

B. Collision Risk Assessment

Collision risk assessment is the stage conducted after ship encounter detection. At this stage, risk of a potential collision is measured using CRI. Beyond CRI value, additional essential information is also provided, including details of the ship encounter situation, relative distances between ship, and DCPA and TCPA. These metrics provide a comprehensive understanding of collision and help quantify risk associated with each situation. In this study, experiment was carried out to evaluate how effectively the proposed framework correlated with ship distance-based risk assessment.

Fig. 4 shows the application of collision risk assessment based on the ship encounter detection results. Each cluster from Fig. 2 is analyzed by forming OS (own ship) and TS (target ship) pairs. For clusters with more than two ships, all possible pairwise evaluations are performed—for example, ships 1, 2, and 3 are assessed as pairs (1-2), (2-3), and (1-3). In Cluster 0 (Fig. 4a), ships are spaced apart with a CRI of 0.302, indicating a low-risk overtaking scenario. Cluster 1 (Fig. 4b) shows a passing encounter with a CRI of 0.410, where the give-way ship passes astern appropriately. In Cluster 2 (Fig. 4c), a short TCPA (0.008) and negative DCPA (-3.070) point to an imminent collision threat, even though the CRI is moderate (0.195). Cluster 3 (Fig. 6d) reveals a moderate CRI of 0.513, requiring navigational caution. This cluster-based approach enables thorough risk evaluation in multi-ship encounters.

C. Comparison of Ship Encounter Detection Method

This section presents a comparative analysis of ship encounter detection using DBSCAN with two different distance metrics: the Haversine formula and traditional Euclidean distance. The objective is to assess each method's effectiveness in identifying ship clusters based on AIS data from the study by Liu et al. [7]. Each figure (Fig. 5 to Fig. 8) shows clustering outcomes of both methods across different experimental datasets. The visual comparison emphasizes how the distance metric influences the formation of ship clusters and the detection of encounter patterns. The analysis reveals that DBSCAN with the Haversine formula consistently captures more relevant encounter scenarios, particularly in spatial configurations where accurate distance representation is essential. In contrast, the Euclidean-based approach exhibits limitations in these maritime datasets, leading to less representative clustering outcomes. This comparative evaluation highlights the advantage of using geodesic-based clustering for improving encounter detection fidelity in maritime traffic analysis.

Fig. 5 shows the clustering results from an experiment where two ships are in encounter situation, with a relative distance of 2.420 nautical miles. DBSCAN-Haversine method successfully classifies ship as a cluster, showing a realistic overtaking encounter. In comparison, Euclidean-based DBSCAN fails to cluster, incorrectly treating ship as outliers. This shows the advantage of Haversine in capturing maritime spatial relationships more accurately.

Fig. 6 shows another overtaking scenario with a relative distance of 2.612 nautical miles. DBSCAN-Haversine detects a valid cluster, where DBSCAN-Euclidean misclassifies one ship as an outlier. This supports the results that Euclidean metric may underestimate the spatial proximity of ships on curved geographic coordinates, which can lead to missed encounter detections.

In Fig. 7, clustering structure differs more significantly. DBSCAN-Haversine groups four ships into a single cluster with a maximum inter-ship distance of 3.851 nautical miles, identifying a broader encounter event. Meanwhile, DBSCAN-Euclidean splits the group into two smaller clusters, each comprising two members. These results technically meet encounter detection criteria, although Haversine-based allows for a more comprehensive risk assessment by including all related ship in a single interaction group.

Fig. 8 presents a scenario where both methods produce similar results, forming comparable clusters. The results suggest that under certain spatial configurations, Euclidean distance can suffice. However, the case is the exception rather than the norm across the evaluated datasets.

In Euclidean-based experiment conducted in this study, the minimum epsilon was tested between 0 and 1 to determine an optimal value for ship encounter detection. The closest value to meeting the detection requirements was 0.05, which was used for this study. This experiment showed the challenge of selecting an appropriate epsilon in Euclidean-based methods, requiring significant tuning to correlate with the needs of maritime navigation, compared to Haversine that directly showed geodesic distances.

Haversine formula shows clear advantages in determining the minimum epsilon according to actual geodesic distances, serving as a more accurate and robust method for distance-based ship collision risk analysis using AIS data. Although Euclidean distance remains a viable method due to the wide application in other studies, application in ship encounter detection requires further optimization. This includes careful tuning of the epsilon parameter to adequately meet the specific demands of maritime safety applications.

D. Cluster Index Evaluation of DBSCAN-Haversine vs. DBSCAN-Euclidean for Ship Encounter Detection

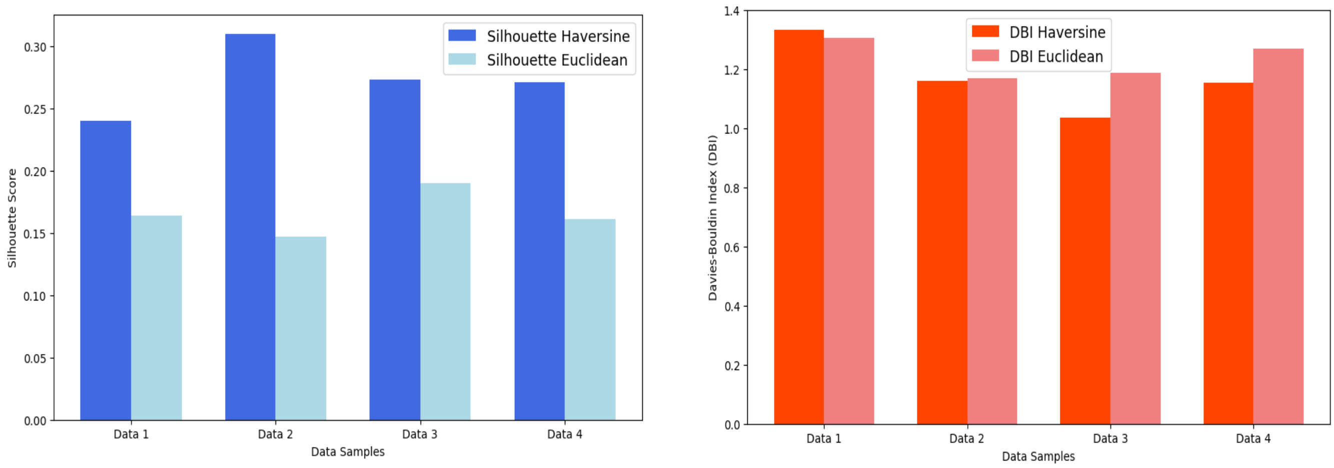
The evaluation of clustering results between DBSCAN using Haversine formula and Euclidean distance aims to determine the effectiveness of each method in producing structured and well-separated cluster. In the context of encounter detection, well-defined cluster ensures that ship included in potential encounter is correctly grouped. Meanwhile, ship that is not part of significant interactions remain separate. This evaluation is essential for collision risk assessment, as the accuracy of clustering directly impacts the detection of ship encounter and risk predictions. Table IV presents the clustering performance based on Silhouette score and DBI, comparing both method. The minimum epsilon value for DBSCAN-Haversine is set to 3 nm, while DBSCAN-Euclidean is 0.05. These values represent the optimal distance parameters for each method to detect ship encounter effectively. Silhouette score evaluates the proper formation of cluster, where higher values indicate better-defined cluster. DBI measures cluster compactness and separation, with lower values showing better clustering performance.

TABLE IV
CLUSTER INDEX EVALUATION BETWEEN DBSCAN USING HAVERSINE FORMULA AND DBSCAN USING EUCLIDEAN DISTANCE

Data	Silhouette Score		Davis Boudin Index	
	Haversine Formula	Euclidean Distance	Haversine Formula	Euclidean Distance
1	**0.240**	0.164	1.335	**1.308**
2	**0.310**	0.147	**1.160**	1.170
3	**0.273**	0.190	**1.037**	1.190
4	**0.271**	0.161	**1.155**	1.269

Fig. 3 presents a comparative visual analysis of the clustering performance achieved by DBSCAN using the Haversine formula and traditional Euclidean distance, corresponding to the evaluation metrics summarized in Table IV. The visualizations depict clustering quality through two widely recognized indices: the Silhouette Score and the DBI. The Silhouette Score assesses cluster cohesion and separation, while the DBI evaluates intra-cluster compactness and inter-cluster distinction. Higher Silhouette Scores and lower DBI values indicate better clustering outcomes. Fig. 3 clearly illustrates that the DBSCAN-Haversine approach consistently outperforms the Euclidean-based variant, exhibiting improved spatial clustering quality.

The evaluation results in Table IV and Fig. 3(a) demonstrate that the DBSCAN algorithm incorporating the Haversine formula consistently outperforms its Euclidean-based counterpart in clustering ship encounters. DBSCAN-Haversine achieves higher Silhouette scores, ranging from 0.240 to 0.310, compared to 0.147 to 0.164 for DBSCAN-Euclidean. These findings indicate that the Haversine distance metric more accurately represents geodesic spatial relationships, leading to better-defined and more cohesive clusters. In contrast, the Euclidean distance often underestimates proximity on a curved surface, leading to inaccurate cluster boundaries and frequent misclassification of encounter ships. As a result, clustering outcomes based on Haversine distance offer improved precision for ship encounter detection in maritime spatial analysis.



(a) (b)

Fig. 3. Cluster index evaluation between DBSCAN using Haversine Formula and DBSCAN using Euclidean Distance: (a) Silhouette Score comparison, (b) Davies-Bouldin Index (DBI) comparison.

As shown in Table IV and Fig. 3(b), DBI results demonstrate that DBSCAN employing the Haversine distance generally produces more compact and clearly separated clusters than its Euclidean-based counterpart. While a few instances display marginally higher DBI values for the Haversine method, the overall pattern consistently reflects greater stability and reliability in clustering performance. The minimum DBI attained with the Haversine formulation was 1.037, which surpasses the best Euclidean outcome of 1.170. Only in a single case did the Euclidean method achieve a slightly lower DBI (1.308 versus 1.335), though this exception is irregular and most likely a result of arbitrary parameter tuning for epsilon. By incorporating geodesic distance, the Haversine formula substantially improves the spatial accuracy of clustering, thereby reducing the incidence of noise points and misclassified encounters. These findings emphasize the robustness of DBSCAN-Haversine, reinforcing its suitability for precise ship encounter detection and reliable collision risk assessment in maritime contexts.

E. Event-Based Evaluation of DBSCAN-Haversine for Practical Ship Encounter Detection

To validate DBSCAN-Haversine in detecting practical ship encounter, an event-based evaluation is conducted using real-world AIS data. This simulation assesses ship movements in detected cluster to evaluate collision risks. The objective is to show that ship cluster identified by DBSCAN-Haversine represents meaningful early warnings for collision risk. The study focuses on two clusters from Fig. 6, which DBSCAN-Haversine identified as encounter but are considered outliers by DBSCAN-Euclidean, showing the importance of geodesic distance in maritime spatial analysis. The simulation generates new coordinate points by maintaining the same velocity and direction from the initial position, indicating realistic ship movement over time.

The event-based evaluation in Fig. 9 simulates the movement of two vessels (MMSI: 412175000 and 477899700) engaged in a head-on encounter. The simulation assumes both ships maintain constant COG and SOG derived from real AIS data. Ship 1 (blue) proceeds northwest at 12.4 knots with a COG of 299.5°, while Ship 2 (red) moves southeast at 10.6 knots with a COG of 118.94°. Successive coordinate points are generated to calculate CRI dynamically as the vessels converge. The results reveal a steady increase in CRI values, culminating at 0.72, which indicates a high-risk situation. The nearly symmetrical progression of CRI trends highlights that both vessels simultaneously perceive escalating collision risk. In accordance with COLREGs Rule 14, both ships are required to alter course to starboard to avoid collision. Any delay in taking corrective action significantly increases the probability of an incident, particularly in congested waterways where traffic density amplifies navigational complexity and reduces reaction time.

The event-based evaluation in Fig. 10 presents a crossing situation between Ship 1 (MMSI: 412175000) and Ship 2 (MMSI: 477899700). Ship 1 (blue) travels northwest (COG: 299.5°) at 12.4 knots, while Ship 2 (red) moves eastward (COG: 118.94°) at 10.6 knots. The CRI initially registers a low value of 0.38 when ships are distant, but steadily increases as proximity grows, reaching a critical peak of 0.90 at the Closest Point of Approach (CPA). This trend highlights the escalating danger and the necessity for early evasive maneuvers. The CRI progression also confirms that both vessels simultaneously perceive the rising collision threat. According to COLREGs Rule 15, in crossing situations, the vessel with the other on its starboard side must give way. In this case, Ship 2 (red) must take prompt action—such as altering course to starboard or reducing speed—to avoid crossing ahead of Ship 1 (blue). Failure to yield could result in a serious near-collision scenario, emphasizing the critical role of accurate encounter detection.

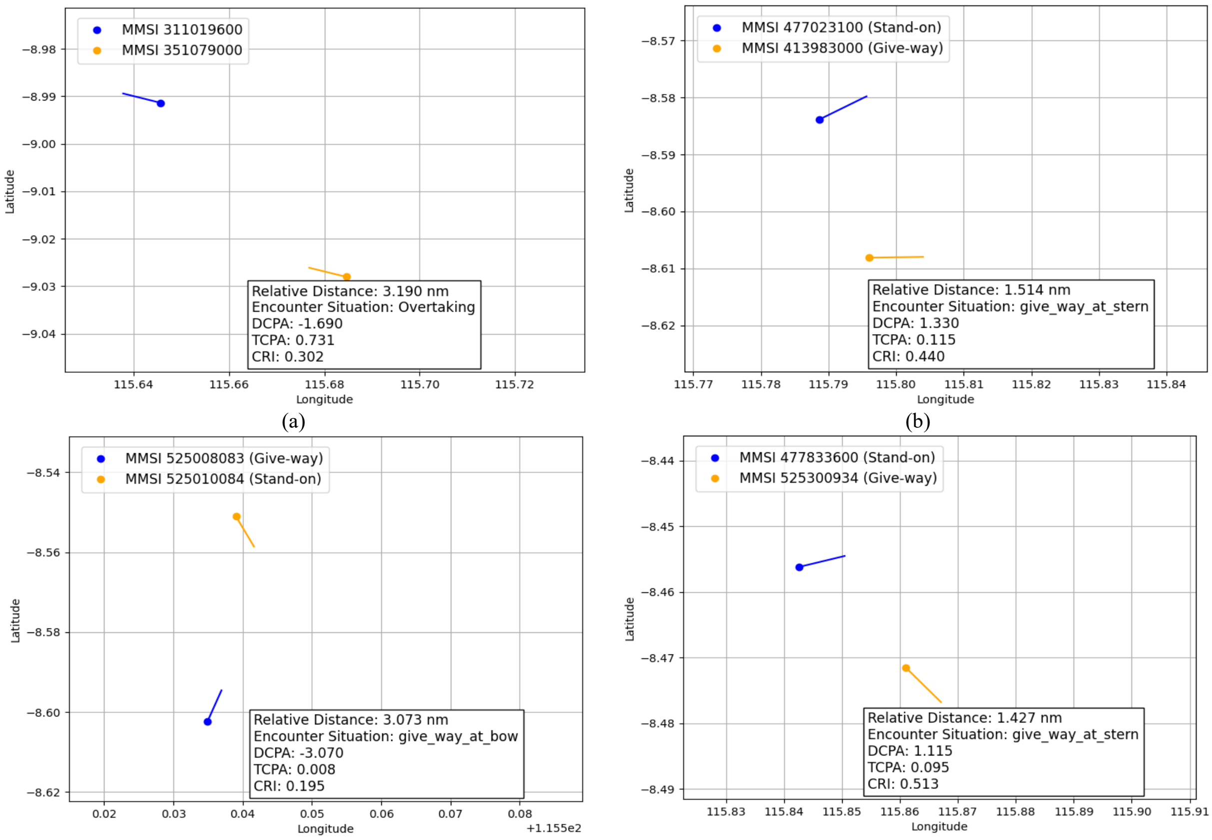
This study confirms that DBSCAN-Haversine significantly enhances ship encounter detection by correctly clustering maritime encounter misclassified by DBSCAN-Euclidean. The event-based evaluation validates real-world applicability, as ship trajectories and CRI measurements show a clear escalation of collision risk in detected cluster. The ability to define critical encounter using geodesic distance makes DBSCAN-Haversine a superior method for maritime safety, providing a more reliable foundation for collision risk assessment and early warning system.

V. Conclusion

In conclusion, this study demonstrated the effectiveness of incorporating the Haversine formula into DBSCAN for ship encounter detection using AIS data. The results showed that DBSCAN-Haversine consistently outperformed the Euclidean-based approach by addressing the inherent limitations of Euclidean distance in geospatial maritime contexts. Cluster quality assessments using Silhouette and DBI further validated its superiority, with higher Silhouette scores indicating more distinct encounter groupings and lower DBI values reflecting greater compactness and separation. Collectively, these outcomes highlight the robustness of DBSCAN-Haversine as a reliable method for maritime collision risk assessment and accurate encounter detection.

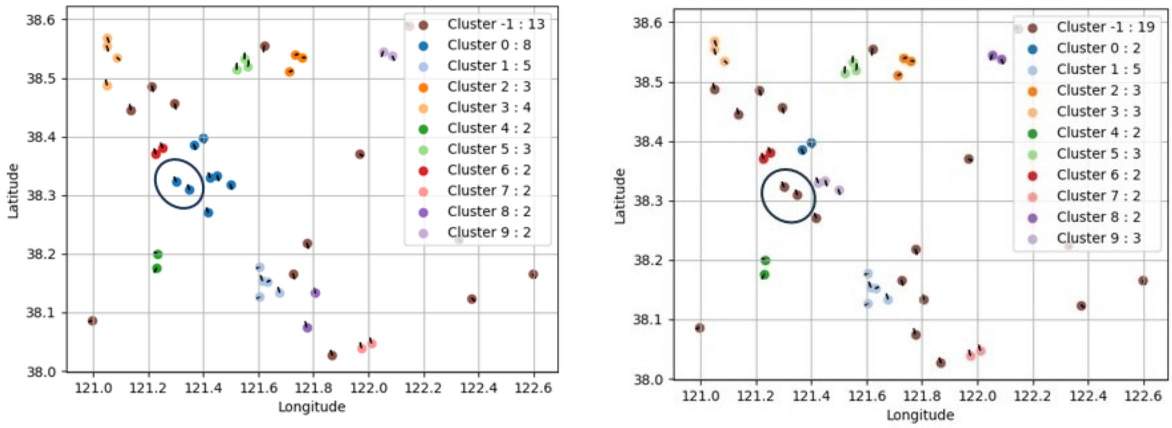
The event-based evaluation highlighted the practical significance of DBSCAN-Haversine for maritime traffic analysis. The movement simulations of detected ship clusters validated its ability to provide reliable early warnings of collision risk, a critical requirement for maritime safety. The observed CRI trends demonstrated that DBSCAN-Haversine effectively identified high-risk encounter scenarios that DBSCAN-Euclidean frequently misclassified as outliers. Moreover, the incorporation of COLREGs compliance in the detected encounters underscored the necessity of accurate clustering methods to support real-time decision-making in navigation and collision avoidance. These results affirm the robustness of the proposed approach in operational contexts where timely detection is vital. Looking forward, further research should investigate adaptive parameter tuning of DBSCAN-Haversine under diverse navigational conditions and explore its integration with machine learning techniques to develop predictive frameworks for collision risk modeling and enhanced maritime traffic management.

APPENDIX



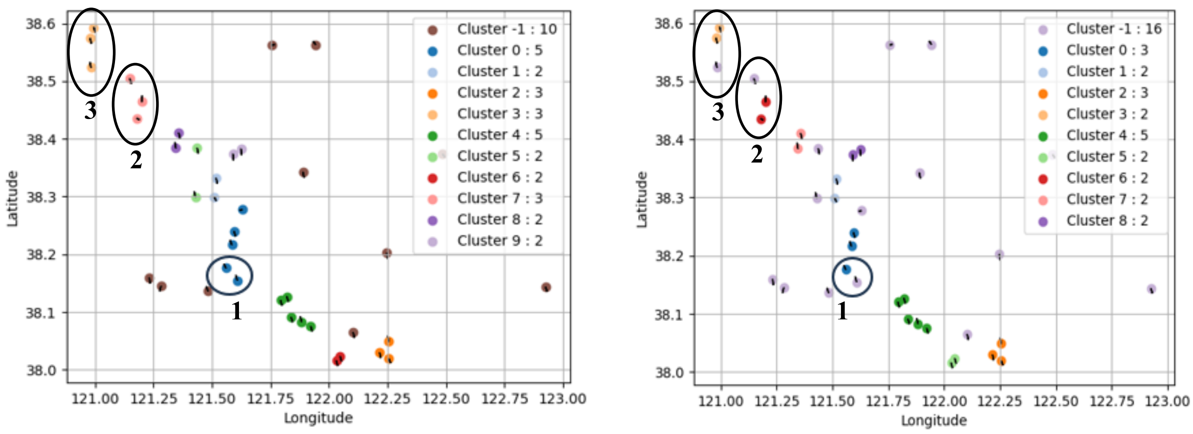
(a) (b) (c) (d)

Fig. 4. Ship collision risk assessment for each cluster result (a) cluster label 0 (b) cluster label 1 (c) cluster label 2 (d) cluster label 3.



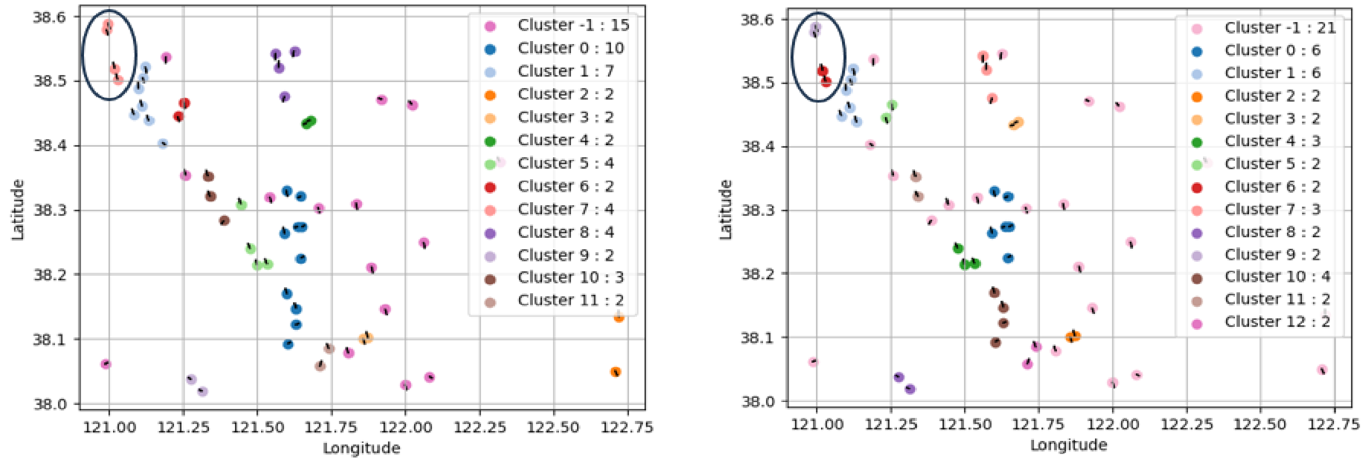
(a) (b)

Fig. 5. Comparison of ship encounter detection results on Dataset 1: (a) DBSCAN-Haversine, (b) DBSCAN-Euclidean.



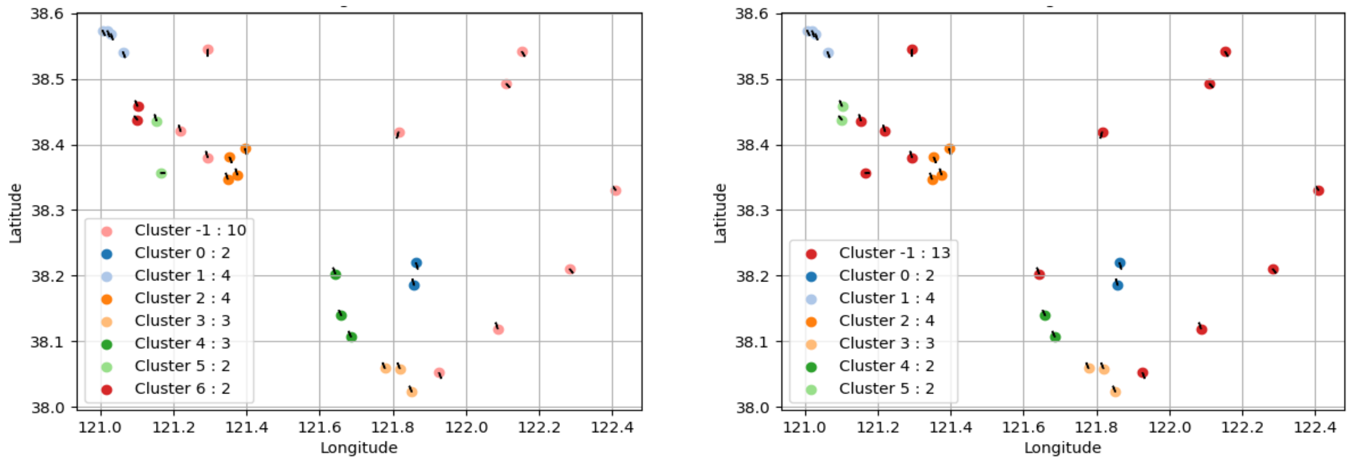
(a) (b)

Fig. 6. Comparison of ship encounter detection results on Dataset 2: (a) DBSCAN-Haversine, (b) DBSCAN-Euclidean.



(a) (b)

Fig. 7. Comparison of ship encounter detection results on Dataset 3: (a) DBSCAN-Haversine, (b) DBSCAN-Euclidean.



(a) (b)

Fig. 8. Comparison of ship encounter detection results on Dataset 3: (a) DBSCAN-Haversine, (b) DBSCAN-Euclidean.

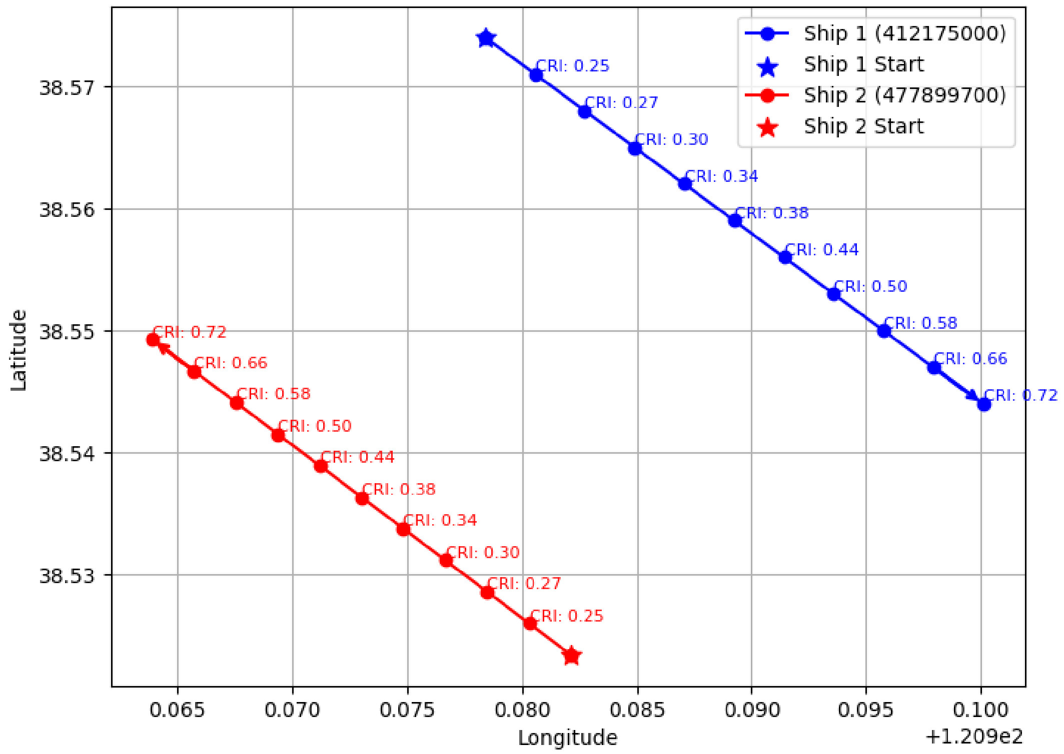


Fig. 9. Simulated ship trajectories and Collision Risk Index (CRI) for head-on encounter scenario detected by DBSCAN-Haversine.

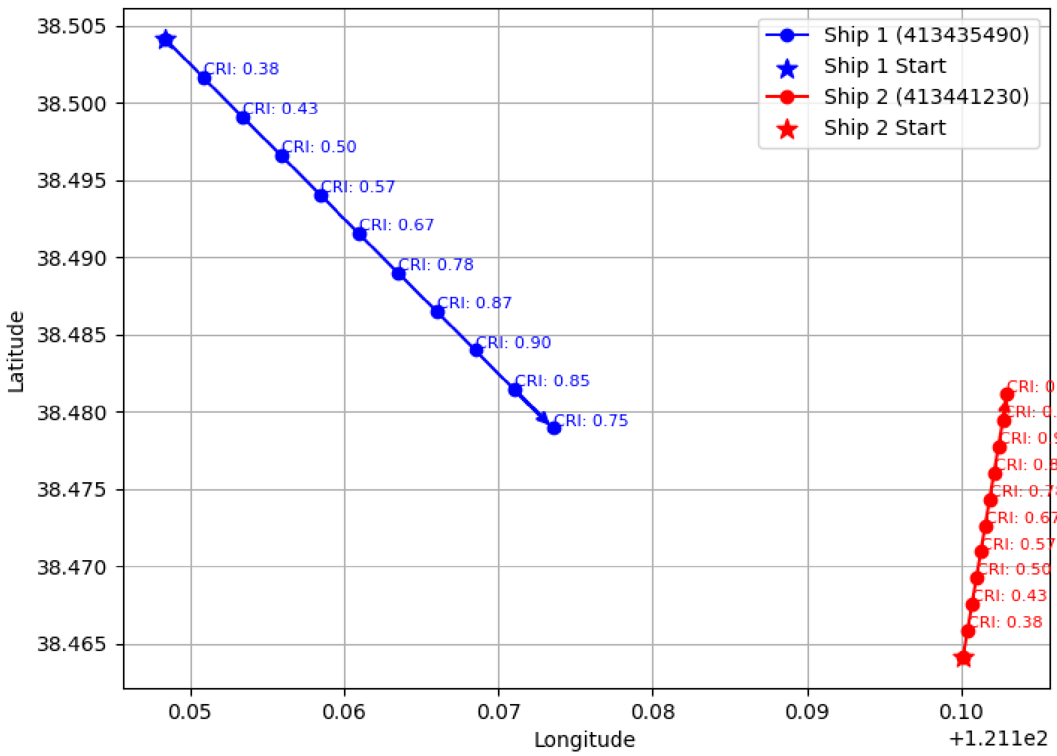


Fig. 10. Simulated ship trajectories and Collision Risk Index (CRI) for crossing encounter scenario detected by DBSCAN-Haversine.

REFERENCES

[1] S. Arasteh et al., "Fishing Vessels Activity Detection from Longitudinal AIS Data," in Proceedings of the 28th International Conference on Advances in Geographic Information Systems, New York, NY, USA: ACM, Nov. 2020, pp. 347–356. doi: 10.1145/3397536.3422267.

[2] D. Zhang, J. Li, Q. Wu, X. Liu, X. Chu, and W. He, "Enhance the AIS data availability by screening and interpolation," 2017 4th International Conference on Transportation Information and Safety, ICTIS 2017 - Proceedings, pp. 981–986, 2017, doi: 10.1109/ICTIS.2017.8047888.

[3] B. Tian, W. J. Gao, and Q. Wang, "Studies on Ships Collision Avoidance with SAPSO Arithmetic," Adv Mat Res, vol. 971–973, pp. 1338–1342, Jun. 2014, doi: 10.4028/www.scientific.net/AMR.971-973.1338.

[4] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Ship collision avoidance behaviour recognition and analysis based on AIS data," Ocean Engineering, vol. 245, p. 110479, Feb. 2022, doi: 10.1016/j.oceaneng.2021.110479.

[5] A. Vestre, A. Bakdi, E. Vanem, and Ø. Engelhardtsen, "AIS-based near-collision database generation and analysis of real collision avoidance manoeuvres," Journal of Navigation, vol. 74, no. 5, pp. 985–1008, Sep. 2021, doi: 10.1017/S0373463321000357.

[6] D. Deng, "DBSCAN Clustering Algorithm Based on Density," in 2020 7th International Forum on Electrical Engineering and Automation (IFEEA), IEEE, Sep. 2020, pp. 949–953. doi: 10.1109/IFEEA51475.2020.00199.

[7] Z. Liu, Z. Wu, and Z. Zheng, "A novel framework for regional collision risk identification based on AIS data," Applied Ocean Research, vol. 89, pp. 261–272, Aug. 2019, doi: 10.1016/j.apor.2019.05.020.

[8] H. Namgung and J.-S. Kim, "Regional Collision Risk Prediction System at a Collision Area Considering Spatial Pattern," J Mar Sci Eng, vol. 9, no. 12, p. 1365, Dec. 2021, doi: 10.3390/jmse9121365.

[9] M. Sangeetha, V. Padikkaramu, and R. Chellan, "A Novel Density Based Clustering Algorithm by Incorporating Mahalanobis Distance," International Journal of Intelligent Engineering and Systems, vol. 11, no. 3, pp. 121–129, Jun. 2018, doi: 10.22266/ijies2018.0630.13.

[10] T. Jabid et al., "An intelligent road accident reduction system using device-to-device communication," Indonesian Journal of Electrical Engineering and Computer Science, vol. 30, no. 3, p. 1729, Jun. 2023, doi: 10.11591/ijeecs.v30.i3.pp1729-1739.

[11] Y. Liu, J. Liu, Q. Zhang, Y. Liu, and Y. Wang, "Effect of dynamic safety distance of heterogeneous traffic flows on ship traffic efficiency: A prediction and simulation approach," Ocean Engineering, vol. 294, p. 116660, Feb. 2024, doi: 10.1016/j.oceaneng.2023.116660.

[12] M. M. Alam, G. Spadon, M. Etemad, L. Torgo, and E. Milios, "Enhancing short-term vessel trajectory prediction with clustering for heterogeneous and multi-modal movement patterns," Ocean Engineering, vol. 308, p. 118303, Sep. 2024, doi: 10.1016/j.oceaneng.2024.118303.

[13] N. C. Greig, E. M. Hines, S. Cope, and X. Liu, "Using Satellite AIS to Analyze Vessel Speeds Off the Coast of Washington State, U.S., as a Risk Analysis for Cetacean-Vessel Collisions," Front Mar Sci, vol. 7, Feb. 2020, doi: 10.3389/fmars.2020.00109.

[14] J. Pan, Y. Wang, T. Wang, and M. Xu, "Study on Assessment of Collision Probability between Ship and Bridge Based on Automatic Identify System Data," J Mar Sci Eng, vol. 12, no. 3, p. 452, Mar. 2024, doi: 10.3390/jmse12030452.

[15] A. Bakdi, I. K. Glad, E. Vanem, and Ø. Engelhardtsen, "AIS-Based Multiple Vessel Collision and Grounding Risk Identification based on Adaptive Safety Domain," J Mar Sci Eng, vol. 8, no. 1, p. 5, Dec. 2019, doi: 10.3390/jmse8010005.

[16] R. Zhen and Z. Shi, "A Novel Adaptive Visual Analytics Framework for Multiship Encounter Identification," J Adv Transp, vol. 2021, pp. 1–11, Jul. 2021, doi: 10.1155/2021/3063957.

[17] I. M. Oka Widyantara, I. P. Noven Hartawan, A. A. I. Ngurah Eka Karyawati, N. I. Er, and K. B. Artana, "Automatic identification system-based trajectory clustering framework to identify vessel movement pattern," IAES International Journal of Artificial Intelligence (IJ-AI), vol. 12, no. 1, p. 1, Mar. 2023, doi: 10.11591/ijai.v12.i1.pp1-11.

[18] Y. Zhang, X. Yuan, M. Li, G. Zhao, and H. Wang, "Multi-Density Adaptive Trajectory Clustering Algorithm for Ships Based on AIS Data," IEEE Access, vol. 11, pp. 108198–108210, 2023, doi: 10.1109/ACCESS.2023.3321270.

[19] J.-S. Lee, H.-T. Lee, and I.-S. Cho, "Maritime Traffic Route Detection Framework Based on Statistical Density Analysis From AIS Data Using a Clustering Algorithm," IEEE Access, vol. 10, pp. 23355–23366, 2022, doi: 10.1109/ACCESS.2022.3154363.

[20] L. Zhao and G. Shi, "Maritime Anomaly Detection using Density-based Clustering and Recurrent Neural Network," Journal of Navigation, vol. 72, no. 04, pp. 894–916, Jul. 2019, doi: 10.1017/S0373463319000031.

[21] S. Sharmila and B. A. Sabarish, "Analysis of distance measures in spatial trajectory data clustering," IOP Conf Ser Mater Sci Eng, vol. 1085, no. 1, p. 012021, Feb. 2021, doi: 10.1088/1757-899X/1085/1/012021.

[22] K. Roberts-licklider and T. Trafalis, "Optimizing Treatment Facility Locations in Oklahoma Using Haversine , Euclidean , Manhattan , and Chebyshev Distance Optimization Optimizing Treatment Facility Locations in Oklahoma Using Haversine , Euclidean ," pp. 0–18, 2024, doi: https://doi.org/10.21203/rs.3.rs-4530065/v1.

[23] J. Z. Wells and M. Kumar, "Predicting sUAS conflicts in the national airspace with interacting multiple models and Haversine-based conflict detection system," Frontiers in Aerospace Engineering, vol. 2, May 2023, doi: 10.3389/fpace.2023.1184094.

[24] X. Han, C. Armenakis, and M. Jadidi, "Modeling Vessel Behaviours by Clustering AIS Data Using Optimized DBSCAN," Sustainability, vol. 13, no. 15, p. 8162, Jul. 2021, doi: 10.3390/su13158162.

[25] X. Duan et al., "Cleaning of Transient Fault Data in Distribution Network Based on Clustering by Fast Search and Find of Density Peaks," Engineering Letters, vol. 31, no. 4, pp. 1348–1358, 2023.

[26] A. A. Bushra and G. Yi, "Comparative Analysis Review of Pioneering DBSCAN and Successive Density-Based Clustering Algorithms," IEEE Access, vol. 9, pp. 87918–87935, 2021, doi: 10.1109/ACCESS.2021.3089036.

[27] D. A. Prasetya, P. T. Nguyen, R. Faizullin, I. Iswanto, and E. F. Armay, "Resolving the shortest path problem using the haversine algorithm," Journal of Critical Reviews, vol. 7, no. 1, pp. 62–64, 2020, doi: 10.22159/jcr.07.01.11.

[28] M. Amiruzzaman, R. Rahman, Md. R. Islam, and R. M. Nor, "Logical analysis of built-in DBSCAN Functions in Popular Data Science Programming Languages," MIST INTERNATIONAL JOURNAL OF SCIENCE AND TECHNOLOGY, vol. 10, pp. 25–32, Jun. 2022, doi: 10.47981/j.mijst.10(01)2022.349(25-32).

[29] R. Zaccone, "COLREG-Compliant Optimal Path Planning for Real-Time Guidance and Control of Autonomous Ships," J Mar Sci Eng, vol. 9, no. 4, p. 405, Apr. 2021, doi: 10.3390/jmse9040405.

[30] Y.-Y. Chen, M.-Z. Ellis-Tiew, W.-C. Chen, and C.-Z. Wang, "Fuzzy Risk Evaluation and Collision Avoidance Control of Unmanned Surface Vessels," Applied Sciences, vol. 11, no. 14, p. 6338, Jul. 2021, doi: 10.3390/app11146338.

[31] F. Deng, L. Jin, X. Hou, L. Wang, B. Li, and H. Yang, "COLREGs: Compliant Dynamic Obstacle Avoidance of USVs Based on the Dynamic Navigation Ship Domain," J Mar Sci Eng, vol. 9, no. 8, p. 837, Aug. 2021, doi: 10.3390/jmse9080837.

[32] E. van Iperen, "Classifying Ship Encounters to Monitor Traffic Safety on the North Sea from AIS Data," TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation, vol. 9, no. 1, pp. 51–58, 2015, doi: 10.12716/1001.09.01.06.

[33] M. Hassel, A. L. Aalberg, and H. A. Nordkvist, "An Advanced Method for Detecting Exceptional Vessel Encounters in Open Waters from High Resolution AIS Data," in Proceedings of the 29th European Safety and Reliability Conference (ESREL), Singapore: Research Publishing Services, 2019, pp. 1709–1714. doi: 10.3850/978-981-11-2724-3_0038-cd.

[34] H. Liu, R. Deng, and L. Zhang, "The application research for ship collision avoidance with hybrid optimization algorithm," in 2016 IEEE International Conference on Information and Automation (ICIA), IEEE, Aug. 2016, pp. 760–767. doi: 10.1109/ICInfA.2016.7831921.

[35] J. Zhou, F. Ding, J. Yang, Z. Pei, C. Wang, and A. Zhang, "Navigation safety domain and collision risk index for decision support of collision avoidance of USVs," International Journal of Naval Architecture and Ocean Engineering, vol. 13, pp. 340–350, 2021, doi: 10.1016/j.ijnaoe.2021.03.001.

[36] M. Cai, J. Zhang, D. Zhang, X. Yuan, and C. G. Soares, "Collision risk analysis on ferry ships in Jiangsu Section of the Yangtze River based on AIS data," Reliab Eng Syst Saf, vol. 215, p. 107901, Nov. 2021, doi: 10.1016/j.ress.2021.107901.

[37] J. Lisowski, "Determining the optimal ship trajectory in a collision situation," in Proceedings of the IX International Scientific and Technical Conference on Marine Traffic Engineering, Szczecin, Poland, 2001, pp. 192–201.

[38] W. Li, L. Zhong, Y. Xu, and G. Shi, "Collision Risk Index Calculation Based on an Improved Ship Domain Model," J Mar Sci Eng, vol. 10, no. 12, p. 2016, Dec. 2022, doi: 10.3390/jmse10122016.

[39] D. Chen, C. Dai, X. Wan, and J. Mou, "A research on AIS-based embedded system for ship collision avoidance," in 2015 International Conference on Transportation Information and Safety (ICTIS), IEEE, Jun. 2015, pp. 512–517. doi: 10.1109/ICTIS.2015.7232141.

[40] M. Nguyen, S. Zhang, and X. Wang, "A Novel Method for Risk Assessment and Simulation of Collision Avoidance for Vessels based on AIS," Algorithms, vol. 11, no. 12, p. 204, Dec. 2018, doi: 10.3390/a11120204.

[41] G. Wang, C. Xing, J.-S. Wang, H.-Y. Wang, and J.-X. Liu, "Clustering Validity Evaluation Functions of Fuzzy C-means Clustering Algorithm," IAENG Int J Comput Sci, vol. 49, no. 2, pp. 356–363, 2022.

[42] N. E. Ivari, M. Wachowicz, P. A. H. Williams, and T. Agnew, "DISCOVERING CLUSTERING PATTERNS FROM E-COUNTER DATA STREAMS," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. X-4/W3-202, pp. 81–88, Oct. 2022, doi: 10.5194/isprs-annals-X-4-W3-2022-81-2022.

[43] S. Adinugroho, M. A. Rahman, and D. Syauqy, "Newsgroup Topic Extraction using Probabilistic Inverse Cluster Frequency Term-Cluster Weighting and Growing Neural Gas Clustering," vol. 48, no. 1, 2021.

[44] Y. Arie Wijaya, D. Achmad Kurniady, E. Setyanto, W. Sanur Tarihoran, D. Rusmana, and R. Rahim, "Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities," TEM Journal, pp. 1099–1103, Aug. 2021, doi: 10.18421/TEM103-13.