Received January 16, 2022, accepted February 16, 2022, date of publication February 24, 2022, date of current version March 7, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3154363

Maritime Traffic Route Detection Framework Based on Statistical Density Analysis From AIS Data Using a Clustering Algorithm

JEONG-SEOK LEE[1], **HYEONG-TAK LEE**[2], **AND IK-SOON CHO**[3]

[1]Graduate School, Korea Maritime & Ocean University, Busan 49112, South Korea
[2]Ocean Science and Technology School, Korea Maritime & Ocean University, Busan 49112, South Korea
[3]Division of Navigation Convergence Studies, Korea Maritime & Ocean University, Busan 49112, South Korea

Corresponding author: Ik-Soon Cho (ischo@kmou.ac.kr)

ABSTRACT Maritime traffic routes by ships navigation vary according to country and geographic characteristics, and they differ according to the characteristics of the ships. In ocean areas adjacent to coasts, regulated routes are present, e.g., traffic separation scheme for ships entering and leaving; however, most ocean areas do not have such routes. Maritime traffic route research has been conducted based on computer engineering to create routes; however, ship characteristics were not considered. Thus, this article proposes a framework to generate maritime traffic routes using statistical density analysis. Here, automatic identification system (AIS) data are used to derive quantitative traffic routes. Preprocessing is applied to the AIS data, and a similar ship trajectory pattern is decomposed into a matrix based on the Hausdorff–distance algorithm and then stored in a database. A similar pattern makes the AIS trajectory simple using the Douglas–Peucker algorithm. In addition, density-based spatial clustering of applications with noise (DBSCAN) is performed to identify the waypoints of vessels then create routes by connecting waypoints. The width of maritime routes created based on a similar ship trajectory is subjected to kernel density estimation analysis (KDE). Then, waypoints evaluation of the main route is performed from the results of KDE 75% and 90% considering the statistical in the total maritime traffic, and the results applied to the targeted ocean area are compared. Finally, the result of KDE 90% of maritime traffic with framework analyzed the safety route, which can be a basis for developing routes for maritime autonomous surface ships.

INDEX TERMS AIS data, clustering algorithm, DBSCAN, kernel density estimation, framework, maritime traffic route.

I. INTRODUCTION

As a result of developments in information and communication technologies and cloud computing technologies, the amount of real-time data is increasing rapidly. In the maritime transportation field, large amounts of automatic identification system (AIS) data are accumulated in real time, and research using such data is being conducted [1].

The shipping industry is the most effective means of transporting goods over long distances, and greater than 80% of the world's merchandize ships use the oceans [2], [3]. Therefore, with the expansion of the shipping industry, the demand for maritime transportation is increasing rapidly, which has resulted in increased ship sizes [4]. Thus, national organizations recommend or enforce route design standards for safe operation of ships and prevention of maritime accidents [5]. Maritime traffic routes for ships navigation vary by country and geographic characteristics. In ocean areas adjacent to the coast, there are prescribed routes that ships must use to enter and leave ports. In addition, the traffic separation scheme is used to guide ship traffic. Moreover, maritime autonomous surface ships (MASS) are expected in future [6]; thus, the need to create maritime traffic routes is increase, even in ocean areas where such routes do not currently exist.

A ship's trajectory data can identify the ship's operating environment from the ship's past motions and patterns. AIS data include various information, e.g., global positioning system information, speed over ground (SOG) information, course over ground (COG) information, ship type, length of all, breadth, time, and place of arrival. In addition, a ship's AIS data are exchanged data with adjacent ships and

The associate editor coordinating the review of this manuscript and approving it for publication was Haiyong Zheng.

AIS stations [7], [8]. The International Maritime Organization (IMO) requires cargo ships with gross tonnage of 300 tons or more engaged in international voyages, ships with gross tonnage of 500 tons or more not engaged in international voyages, and all passenger ships to be equipped with AIS [9]. These regulations form part of Chapter V of the Safety of Life at Sea (SOLAS) 3rd convention [10]. This AIS is connected to other navigation equipment, e.g., the electronic navigation chart display and information system and the automatic radar plotting aid, to prevent collisions and facilitate safe navigation [1]. In addition, according to the literature [9], the large-scale AIS data are also used to extract maritime routes and analyze ship motion. Thus, AIS data are used to research improved maritime routes by combining various ship information and computer technologies. Son *et al.* [11] explained the width of sailing ships from a reliability-based statistical perspective and conducted a study within a harbor. However, to the best of our knowledge, no previous study has considered density analysis based on statistics in maritime routes or maritime traffic networks using AIS data and route width information. Navigable areas differ depending on the size of the ship; however, this problem has not been considered in previous studies. Therefore, maritime traffic routes are extracted in wide ocean areas considering ship characteristics, and route width values and the main route are calculated according to statistical-based density analysis.

In this study, the AIS data of ships passing through the Southern Sea of Korea were targeted. The Korean Ministry of Oceans and Fisheries collects the AIS data of the ships passing through the ocean near Korea and provides these data for the General Information Center on Maritime Safety and Security data. The raw AIS data are preprocessed due to the data loss rate according to the distance and inaccuracy of the data [12]. Trajectory clustering differs from point clustering; thus, the trajectory should not deviate from generality and should include similar characteristics. Using the Hausdorff–distance algorithm, which considers the shape difference between two different sets of points, and the trajectory is clustered effectively [13]. Numerous course adjustment movements are performed in an actual ship trajectory to proceed to the course set by a vessel's autopilot, which increases the complexity of a ship's trajectory. Here, the Douglas–Peucker (DP) algorithm is used to simplify the trajectory [14]. A simple trajectory effectively represents waypoints that change a ship's course. In other words, it is possible to create a vertex point and clusters the vertices generated in a specific part to select a waypoint for a maritime traffic route. Here, density-based spatial clustering (DBSCAN) is applied to point clustering, and its use has been proven in the traffic route extraction and anomaly detection (TREAD) method [13]. A maritime traffic route in consideration of the characteristics of the trajectory's motion and ship size is created by connecting waypoints. Then, the kernel density estimation (KDE) of the line is analyzed to calculate the statistical route width and the calculated value of the main route [15]. KDE is limited in that it gives different results depending on the setting value and search range. However, KDE was used to quantitatively calculate the route width and was applied as a method to evaluate the performance of the maritime route. KDE identifies an efficient maritime traffic route by setting the search range from an arbitrary trajectory and analyzing routes as 75% and 90% of the total maritime traffic using a statistical method [4], [5]. Therefore, it is possible to create a maritime traffic route based on density analysis, and the result of density analysis can have a unique shipping area in a wide ocean area. This is expected to be applicable to autonomous ships based on a simplified main route, and it is applicable to anomaly detection based on routes [16]. In addition, create a maritime traffic route to protect the ocean environment, promote navigation safety of ships and crews, create marine spatial planning (MSP), and apply to the MASS [17]–[19].

II. RELATED STUDY AND MOTIVATION

As large amounts of AIS data are collected, the data should maximize public interest by linking methodologies in other research fields with the maritime industry. AIS data can be used to extract maritime traffic routes, and scientific techniques, e.g., deep learning, can be applied to real-time maritime traffic networks [1], [20]. Moreover, AIS data are applied to navigation technology for MASS by combining the collision risk of ships with the International Regulations for Preventing Collision at the Sea [21]. In addition, another study analyzed ship traffic in Singapore Strait using spatial-temporal analysis. Here, dense areas were divided into hotspots [22], and maritime traffic was measured by local hotspots and cold spots via spatial autocorrelation analysis [8]. In addition, AIS data are a key element in the analysis of a ship's operation pattern. Thus, AIS data analyzes the ship's COG, SOG, Rate of Turn data and Voyage Data Recorder information with a clustering algorithm. From the past data, it is possible to understand the ship's engine order and rudder use to prevent ship accidents [23]. Further, various machine learning techniques were used to predict the safety berthing velocity of ships. The algorithm with the best performance was derived through the receiver operating characteristic curve. This study aims to protect port facilities and the environment by presenting guidelines for large ships to dock safely at ports [24]. Therefore, AIS data can be fused with various algorithms and obtain desired results. In this study, a method for detecting main maritime traffic routes based on vast trajectory data is proposed. After detecting the maritime route, used the proposed a method to select the width of the route, and furthermore, propose a framework to represent the safety of the route.

A. RESEARCH ON MARITIME ROUTE EXTRACTION

Three methods are used to extract maritime routes, i.e., density-based analysis, clustering-based route extraction, and TREAD. Lee *et al.* [5] extracted density-based routes by analyzing 90% of maritime traffic with KDE based on a

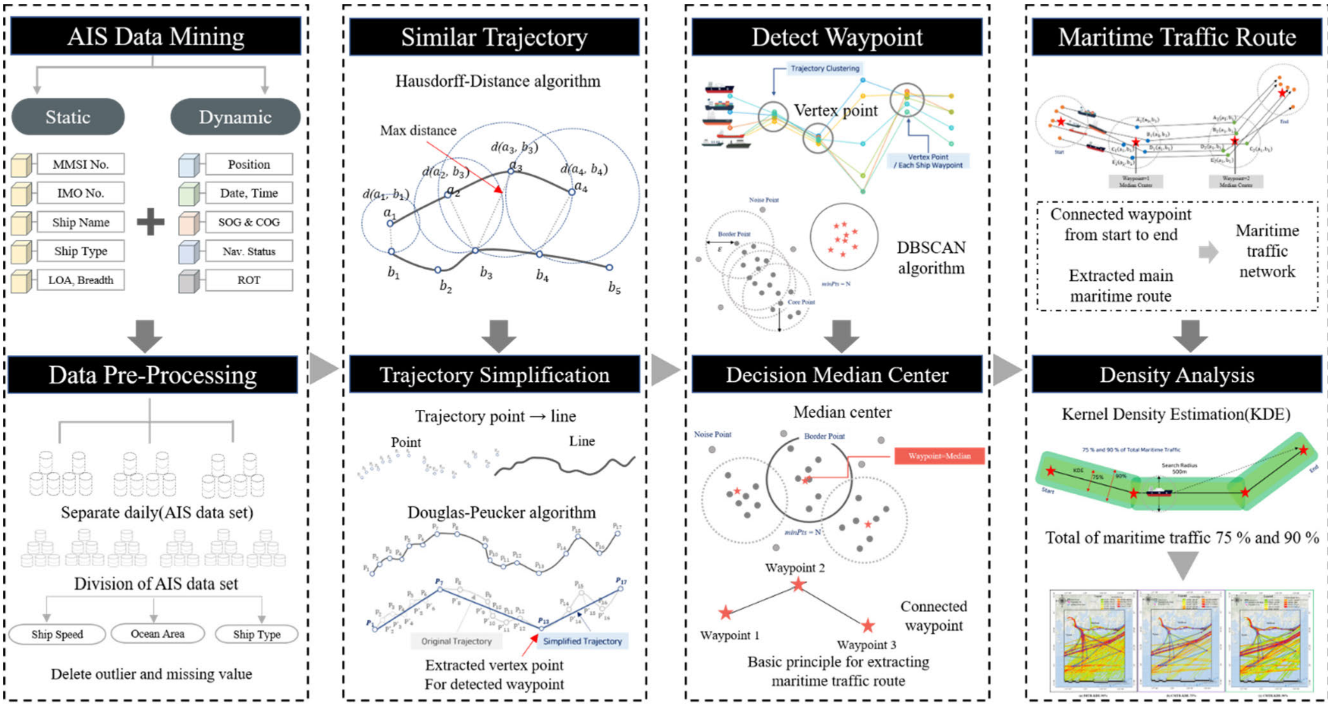


FIGURE 1. **Maritime traffic route framework using the clustering algorithm and statistical density analysis.**

geographic information system (GIS). The Maritime route boundary was extracted using the Otsu binary analysis method and the Canny algorithm, which is an edge extraction method, for the density analysis results [25], [26]. The superiority of the previously set route was explained through image processing.

Liang et al [27] aimed to achieve reliable mining results for massive vessel trajectories for efficiently computing the similarities between different vessel trajectories. Liu et al [28] predicted the vessel trajectory for improving smart maritime traffic services with deep learning. This work proposed an AIS data-driven trajectory prediction framework based long short term memory network.

Wang et al [29] used the Douglas–Peucker algorithm to simplify the ship trajectory, and the Hausdorff–distance algorithm was used in the matrix. In addition, hierarchical density-based spatial clustering of applications with noise (HDBSCAN) was used to cluster trajectory according to hierarchical. Wang et al [29] compared results of the K-means, spectral, DBSCAN, and HDBSCAN clustering analysis methods and found that HDBSCAN had the best clustering performance.

TREAD which is commonly used for the maritime traffic extraction task, was proposed by Pallotta et al [13]. The AIS data were used to understand, classify, and predict maritime traffic patterns in order to automatically detect anomalies in vessel operation patterns.

B. MARITIME TRAFFIC NETWORK

A method to create a maritime traffic network using the TREAD method was proposed by Arguedas *et al.* [30]. This method automatically generates a synthetic maritime traffic network based on the historical positioning data of ships. Here, waypoint and route detections are based on the TREAD method and DBSCAN algorithm. In addition, TREAD uses the Hausdorff–distance algorithm for route decomposition and the Douglas–Peucker algorithm to detect breakpoint detection. The created route is unsupervised, and the learning result is displayed as performance evaluation. A maritime traffic network includes connected nodes that are created by extracting routes. Here, a node represents a ship's navigation waypoint, and the method used to extract the node is approached in a scientific way.

Forti *et al.* [31] used AIS data to generate graph-based maritime traffic network, and they proposed an unsupervised approach to extract the patterns of ships [32], [33]. Here, the Ornstein–Uhlenbeck mean-reverting stochastic process was applied to describe the application of graphs to detect and connect waypoints via ship motion modeling [34], [35]. This creates a network and realizes maritime traffic pattern detection and statistical characterization.

Yan *et al.* [9] proposed a maritime traffic network based on the framework of the ship trip semantic object (STSO). Here, STSO describe a network creation methodology that finds and connects stop point and waypoint using the ordering points to identify the clustering structure clustering algorithm [36].

Based on the study by Wang *et al.* [37], this study extracted a maritime traffic network and proposed a different maritime traffic network creation method. The method included how to create the lane boundaries, lane centerlines, and junctions. The route was created by applying the image processing technique based on the KDE density analysis. The outside of the route was extracted as a lane, and the bumpy lane was smoothed out. Further, by applying the Delaunay triangulation model, attribute values according to adjacent triangles were assigned. Finally, nodes and segments were created and the centerline connected to them was presented as a maritime traffic network.

III. MATERIALS AND METHODS

To generate a maritime traffic route, it is necessary to understand ships' movement patterns and operational factors. Generally, ship trajectories are created to reduce travel distance and maintain safety; thus, most merchant ships have similar trajectories. This means that AIS data can be used to generate maritime traffic routes [38]. In this study, route width and the main route are also considered by applying statistical density analysis to differentiate existing maritime traffic networks. Therefore detecting the maritime traffic route in the vast trajectory data is crucial [27]. In addition, for a vessel to sail safely, an appropriate width of route is an essential condition. The selecting appropriate width of route was presented as a new method by performing a simplified trajectory-based density analysis. Furthermore, safety of the route is a key factor. To execute the sailing plan without the ship grounding, the reliability of the detected maritime traffic route must be guaranteed. To solve this problem, a novel statistical density analysis was applied. Further, the median center was used to express one waypoint that appears as clustering to detect maritime traffic routes. As a whole, a vast amount of trajectory data was concisely collated and efficiently analyzed. The proposed maritime traffic route framework is illustrated in Fig. 1.

A. DATA MINING AND PREPROCESSING AIS DATA

Currently, the AIS database (approximately 950 GB) stores data for all ships that passed through waters near Korea from 2018 to 2020. However, the raw AIS data are unsuitable for analysis. The AIS data comprise both static and dynamic data. Thus, to understand and interpret the data accurately, the data must be combined. Combining the data was is performed based on the maritime mobile service identifier (MMSI), which has the same information. After being combined, the AIS data include static data comprising the MMSI, name, type, IMO number, draft, tons, and length, and the dynamic data include date (year-month-day-time), latitude, longitude, SOG, ship's heading, and COG. The MMSI comprises nine numbers and attempts to make it easy to identify a vessel. In addition, the ship type consists of numbers from 1 to 99 [39].

After data combining, data preprocessing was performed. To perform data analysis efficiently, the data were divided into daily data, and then the analysis area was divided into 16 areas and stored in the database. In addition, data analysis may yield different results depending on a vessel's type and characteristics. Therefore, the types of ships were divided into cargo, tanker, passenger, and towing ships, and the length of the ship was divided as > 60 m.

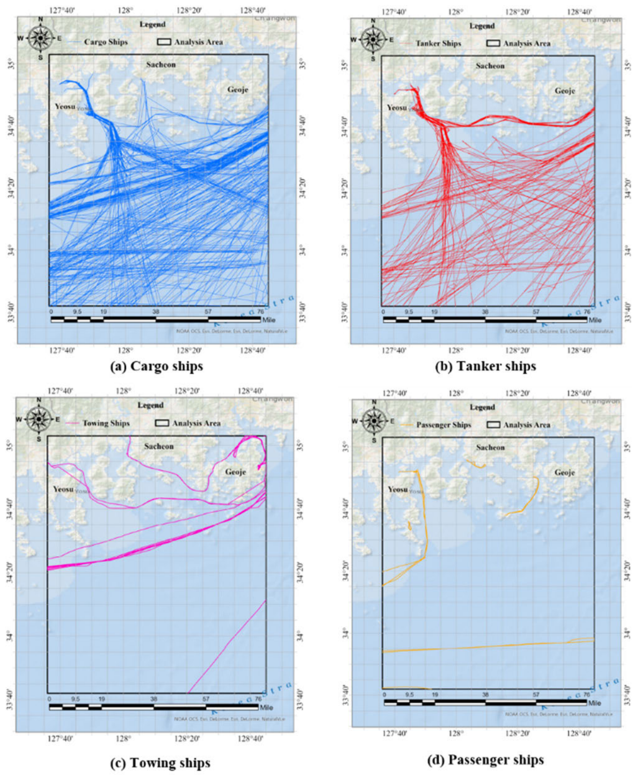


FIGURE 2. AIS data plot from March 1, 2018 to March 7, 2018, (a) cargo ships and (b) tanker ships, completely different from tanker ship (c) towing ships and (d) passenger ships.

During data preprocessing, missing values were removed, and all inaccurate characters and symbols were deleted. In addition, due to the characteristics of the ships, the SOG is nearly zero during berthing, anchoring, and drifting conditions; thus, SOG is unsuitable for determining maritime traffic routes. Due to the characteristics of merchant ships, there are no data with SOG of 25 kn or greater; thus, SOG was selected as 3 kn or more and 25 kn or less. For COG, which indicates the ship's moving direction, all values other than 0 ° – 360 ° were deleted. As shown in Fig. 2, the operation pattern differs depending on the type, length, and breadth of the vessel, and the route appears differently. Therefore, the AIS data should be analyzed by categorizing ship types.

B. CLUSTERING ALGORITHM

Ship trajectories are clustered to identified similar trajectories. Merchant ships exhibit similar trajectories according to the route selection purpose, and it is necessary to cluster appropriately [38]. The Hausdorff–distance algorithm can represent a similar ship trajectory as a matrix [29]. If the Hausdorff–distance is a subset of a non-empty metric space, where T_a represents trajectory a, and T_b represents trajectory b, $d_H\left(T_a, T_b\right)$ in the Hausdorff space is expressed as follows.

$$d_H\left(T_a, T_b\right) = \max\left\{d\left(T_a, T_b\right), d\left(T_b, T_a\right)\right\} \quad (1)$$

Here, T denotes the trajectory, $T_a = [a_1, a_2, \cdots, a_{n-1}, a_n]$ denotes the trajectory of ship a, and $T_b = [b_1, b_2, \cdots, b_{n-1}, b_n]$ denotes the trajectory of ship b. Given

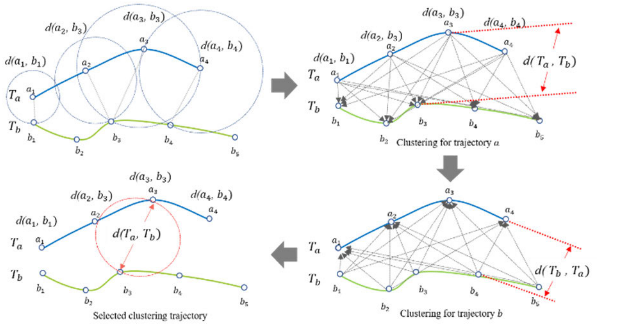


FIGURE 3. **Conceptual diagram of the Hausdorff–distance and process to calculate the similar trajectory.**

two sets of points T_a and T_b, the Hausdorff–distance in one direction is obtained as follows.

$$d(T_a, T_b) = \max\{\min\{\|a_i - b_j\|\}\} \tag{2}$$

$$d(T_b, T_a) = \max\{\min\{\|b_j - a_i\|\}\} \tag{3}$$

Here, $\|a_i - b_j\|$ and $\|b_j - a_i\|$ use the Euclidean–distance to define two trajectory points as the Hausdorff–distance [40]. An example of the Hausdorff–distance based on the ship trajectory is shown in Fig. 3.

Fig. 3 shows that the shortest distance from T_a to T_b is $d\,(a_3, b_3)$, and the shortest distance from T_b to T_a is $d\,(b_4, a_4)$. In other words, the Hausdorff–distance calculates the largest distance among all distances from a point on one trajectory to the nearest point on another trajectory.

These connected distances can be clustered by the matrix of the past trajectory. This study approaches from a statistical perspective and extract the main route in consideration of route width. Therefore, to cluster similar trajectories, a value of 0.1 or less for each MMSI is used as one clustering. After clustering, the vessel changes points to lines in over time based on the MMSI number. Thus, the trajectory line is performed as a line-based analysis rather than a point-based analysis.

C. SIMPLIFICATION OF SHIP TRAJECTORY

Owing to the complexity of a ship's operation, the ships do not continuously maintain a certain course. Ships operation includes continuously resisting the external force conditions of the ship, e.g., wind, waves, and dynamic factors, and continuously steadying the course based on the course setting value set by the autopilot system. Therefore, clustering trajectories that maintaining natural motion is performed first, and trajectory simplification is performed on the clustered result. Here, the Douglas–Peucker algorithm is used [29], [30], [41]. This algorithm maintains the threshold points to maintain the essential shape of the line.

The Douglas–Peucker algorithm begins by connecting one end of the line to the trend line, and the distance between the trend line and the trajectory is measured vertically. Here, vertices that are closer to the trend line than the threshold are removed. This procedure is performed until all vertices are removed while dividing the trend line [42]. In this study,

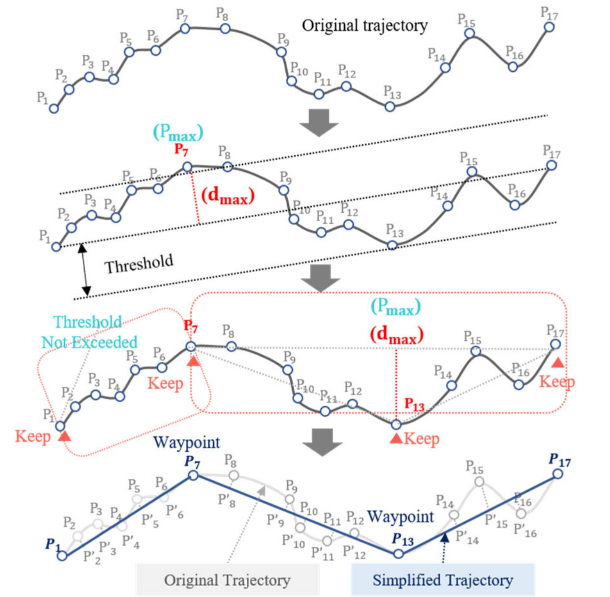


FIGURE 4. **Simplifying the ship trajectory using the Douglas–Peucker algorithm.**

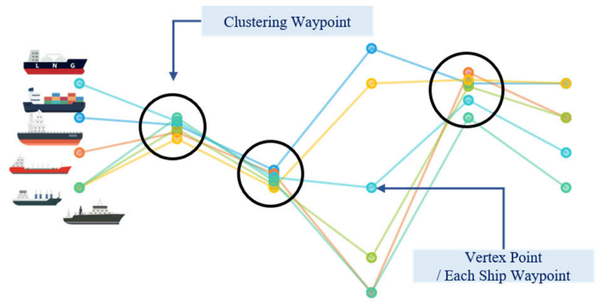


FIGURE 5. **Extracted to the vertex point of the ships trajectory.**

the Douglas–Peucker algorithm first creates trend lines of $P_1, P_2, \cdots, P_{n-1}, P_n$ and sets the threshold value d. Then, d_{max} exceeding threshold d is selected, and the maximum value is expressed as P_{max}. Therefore, when $d_{max} > d$, the procedure is divided into two parts with the corresponding boundary point P_{max} and then repeated (Fig. 4).

D. WAYPOINT DETECTION IN MARITIME TRAFFIC ROUTE

A clear waypoint can be calculated by simplifying the trajectory of each ship. However, even if the trajectories of all ships are similar, the waypoints will differ slightly. Therefore, to extract the waypoints of all trajectory lines, all vertex points (where one trajectory line meets another trajectory line and creates a certain angle) are extracted. An example is shown Fig. 5.

As the ships have a similar route owing to the operational characteristics of the ship, the DBSCAN algorithm is used to find the position where many vertex points (representing a waypoint) are derived. The DBSCAN algorithm is an unsupervised machine learning technique that is widely used in point information-based research [23]. For waypoint detection in a maritime traffic network, waypoints have been

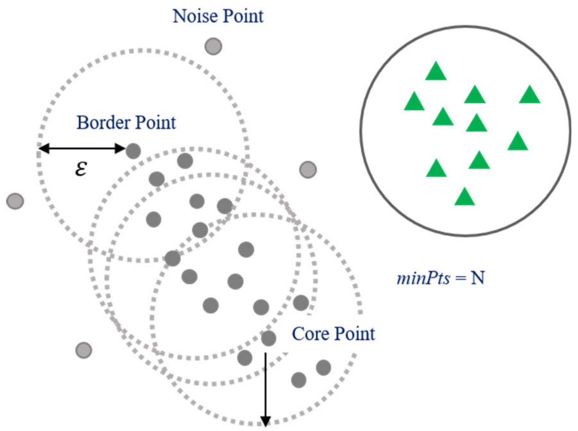


FIGURE 6. DBSCAN works process and create the cluster.

extracted using DBSCAN from TREAD and the STSO, which is also considered in this this study [8], [14]. Two parameters are required to use DBSCAN, i.e., epsilon (ε) is a factor that determines the minimum distance between all data points, and the minimum point (*minPts*), which is the minimum number of data that must be included within the ε minimum distance [43]. As shown in Fig. 6, when there are many *minPts* inside ε, the DBSCAN algorithm is initiated and expands in the direction of the data that satisfies the same conditions adjacent to each other.

Here, if the dataset is defined as D, the vertex point is represented by p and the ε neighborhood area of point is defined as N and follows Eq. (4).

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

where p follows Eq. (5), which constitutes the *minPts* of the DBSCAN algorithm.

$$\left|\boldsymbol{N}(\boldsymbol{p})\right| \geq \boldsymbol{minPts} \tag{5}$$

In addition, a boundary point is a point that is close to the core point but less than *minPts* from the ε neighborhood, and noise is defined as a point other than the core and boundary points. In waypoint detection using the DBSCAN algorithm, several points appear as individual clusters, and the exact center point must be found. As a measure of the central tendency of the dataset, DBSCAN is suitable for detecting outliers and has the characteristic of minimizing multiple waypoints to only one waypoint. Therefore, an example of extracting the DBSCAN median center is shown in Fig. 7, and the equation is given as follows.

$$d_i^t = \sqrt{(X_i - X^t)^2 + (Y_i - Y^t)^2 + (Z_i - Z^t)^2} \tag{6}$$

Here, X_i, Y_i, and Z_i are the coordinates for feature i, and n is equal to the total number of features. The median center is used to search for effective and integrated waypoints of different merchant ships, and weights can be assigned according to the attribute values of these points. Therefore, when waypoints 1 and 2 are connected, a leg of a single maritime route is created. In addition, the start and end

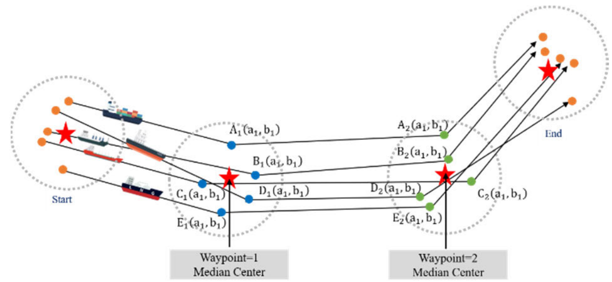


FIGURE 7. Median center selection based on the DBSCAN cluster results of vertex point.

points can be extracted using the same method, and the final maritime traffic route is created when the start point, end point, and all waypoints are connected.

E. KERNEL DENSITY ESTIMATION OF MARITIME TRAFFIC ROUTE WIDTH

When the extraction of maritime traffic routes is completed by connecting the waypoints to each other, KDE analysis is performed on the trajectory the vessel has followed to create the route width. Lee *et al.* [5] analyzed the KDE analysis to calculate route width with trajectory point data. However, this study is based on the trajectory line. Because when the ships were docked in a port for a long period of time, there is a limit to an accurate analysis because there are many trajectory point data. The density analysis method estimates the characteristics of random data variables. The most effective method to achieve this involves drawing a histogram. A histogram is parametric and suitable for visualizing only simple frequencies. However, such a method is discontinuous at the boundary of each histogram bin and has different disadvantages depending on the width of the bin. In addition, KDE is effective for estimating nonparametric density functions for data following nonparametric distributions. The KDE equation is given in Eq. (7).

$$\hat{f}(x, \hbar) = \frac{1}{n}\sum\nolimits_{i=1}^{n} K_{\hbar}(x - x_i) \tag{7}$$

Here, x is a random variable, and x_i is an observation. In addition, $\hbar$ means bandwidth and kernel width; smoothing parameter as a parameter for adjusting the kernel. The kurtosis of the kernel is determined according to the size of $\hbar$. Here, K is divided by the sum of the observations with the kernel function applied to the number of observations [44].

In this study, trajectory line data were used to perform KDE analysis. The grid cell visually representing the analysis result was set to 100 m. The search radius used to calculate the density of adjacent lines in one random trajectory line was set to 500 m. According to Article 60, paragraph 5 of the United Nations Convention on the Law of the Sea (UNCLOS), an artificial island and a maritime facility installed in the exclusive economic zone is determined as a safety zone at least 500 m from the maritime structure [45]. Finally, as a determining factor of route width, 75% and 90% of the total maritime traffic are represented using the quantile

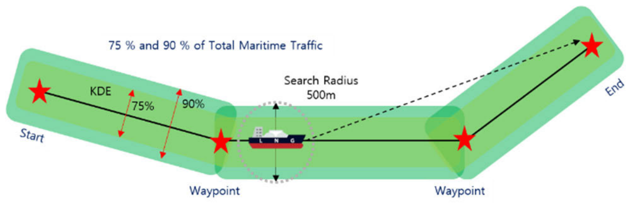


FIGURE 8. KDE analysis conceptual diagram applying 75% and 90% of the total maritime traffic.

method. Due to recent developments in offshore wind farms, research is being conducted to define the minimum separation distance to ensure the safe passage of ships. Here, 90% of the maritime traffic is arranged by configuring the safety distance between the offshore wind farms and ships [46]. In addition to ocean energy development, 90% of marine traffic is applied to MSP, which considers various entities using different ocean spaces [47]. However, density analysis when 90% of maritime traffic is applied represented that even if only one vessel passes through the space of the ocean, there is a limit to representing all the space of the ocean as the passage of ships, and to make other marine uses sustainable to reduce the width of the route. Therefore, in this study, 75%, i.e., range of 2σ, was also analyzed according to Chebyshev's inequality. Chebyshev's inequality is more general, stating that a minimum of 75% of values must lie within two standard deviations of the mean and 88.89% within three standard deviations for a broad range of probability distributions [48]. The route width created in this study from a statistical perspective is shown in Fig. 8.

IV. CASE STUDY

A. DATA MINING AND PREPROCESSING FOR AIS

In this study, the analyzed area was the Southern Sea of Korea, where many cargo and tanker ships pass through. In addition, many passenger ships are present in these waters due to the nearby islands. The analysis area is from latitude 33.70 ° north to 35.05 ° north, and longitude 127.60 ° east to 128.75 ° east. The analysis area is shown in Fig. 9.

The data used for analysis were for cargo ships and tanker ships that have passed through the target area. It is impossible to target all ship types to select a maritime traffic route because many types of ship, e.g., towing ships, passenger ships, and fishing ships, exhibit irregular and nonuniform traffic patterns. Thus, towing ships, passenger ships and fishing ships were omitted from the analysis because of industrial activities, e.g., maritime work, fishing activity, and cable installation, rather than maintaining a constant track according to vessel characteristics. Table 1 shows the data collected from March 1, 2018 to March 7, 2018, with 319 cargo ships and 136 tanker ships passing through the target area.

In addition, the detailed data statistics of the ships are shown in Tables 2 and 3. Maritime traffic routes were created based on the data of 319 cargo ships and 136 tanker ships.

The 319 cargo ships that passed through the target area were navigating at an mean of 11.85 kn, and the maximum

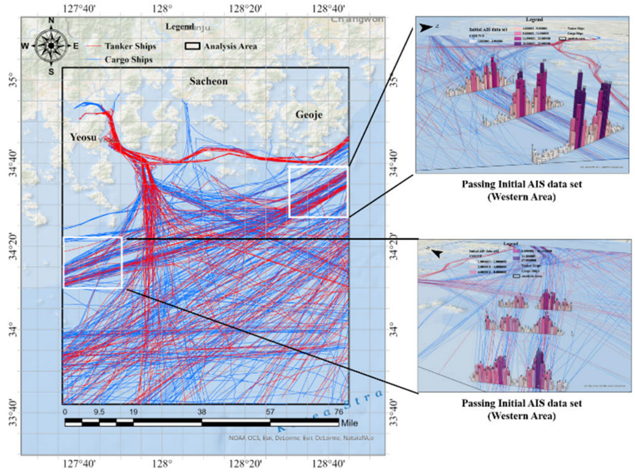


FIGURE 9. Area of analysis; Southern Sea of Korea with 3D result of passing ships.

TABLE 1. Counting analysis point data and traffic data according to ship type.

Ship type	Total points	Total ships
Cargo ship	545,779	319
Tanker ship	236,866	136
Total	782,645	457

TABLE 2. Statistical analysis result of cargo ships.

Cargo ship	SOG (kn)	COG (°)	Length (m)	Breadth (m)
Mean	11.85	167.96	168.28	23.72
Std.	3.48	98.53	83.20	11.68
Min.	3.00	0.00	64.00	10.00
25%	9.40	68.90	104.00	15.00
50%	11.60	176.00	148.00	20.00
75%	14.30	244.00	199.00	32.00
Max.	22.30	359.90	399.00	62.00

TABLE 3. Statistical analysis results of tanker ships.

Tanker ship	SOG (kn)	COG (°)	Length (m)	Breadth (m)
Mean	10.96	178.10	141.42	25.71
Std.	2.89	98.99	67.61	10.76
Min.	3.00	0.00	60.00	9.00
25%	9.10	77.00	90.00	17.00
50%	11.30	184.00	119.00	24.00
75%	13.20	254.80	183.00	32.00
Max.	20.10	359.90	336.00	60.00

SOG was 22.30 kn. The mean vessel length was 168.28 m, and the longest vessel length was 399.00 m. The mean breadth of the vessel was 23.72 m, with a maximum breadth of 62.00 m.

The 36 tanker ships passed through the mean SOG at 10.96 kn, and the maximum was 20.10 kn. The mean vessel length was 141.42 m, and the largest vessel was 336.00 m. In addition, the mean breadth of the vessels was 25.71 m, and the maximum breadth was 60.00 m. There was a slight difference in COG between cargo and tanker ships because the purpose of operation is different according to the type of ships.

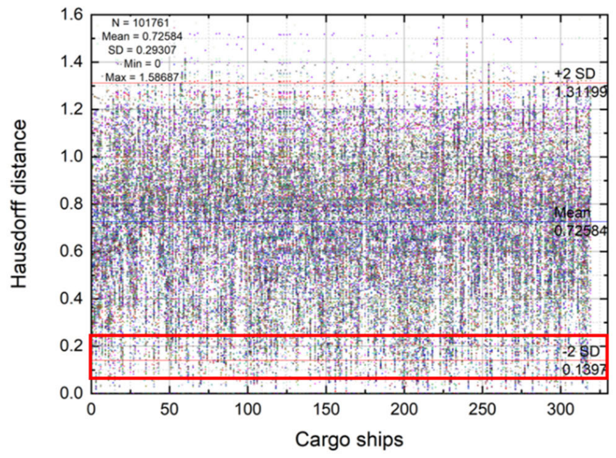


FIGURE 10. Hausdorff-distance calculation result of cargo ships.

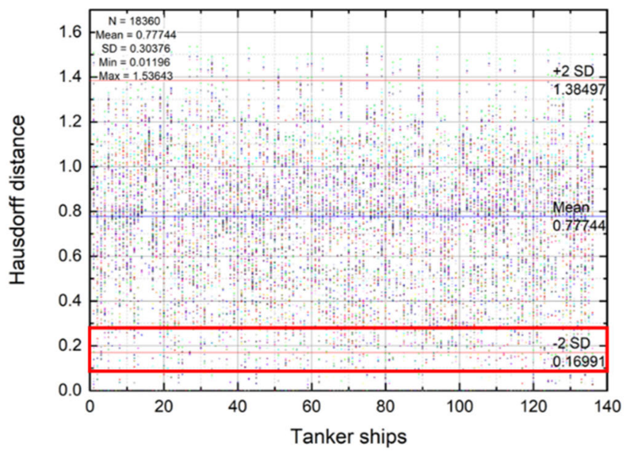


FIGURE 11. Hausdorff-distance calculation result for tanker ships.

B. RESULTS OF SHIP TRAJECTORY CLUSTERING AND WAYPOINT DETECTION

Clustering similar ship trajectories was selected according to the Hausdorff–distance algorithm. It calculates the distance between ships MMSI, and as the numerical value of distance appears as closer to zero, each trajectory means similar. Conversely, dissimilar trajectories have large numerical value.

In this study, the range of $\pm 2\sigma$ was calculated from the mean value, and a similar trajectory was selected as 0.1, which is less than -2σ, as shown in Fig. 10 and Fig. 11. Fig. 10 and Fig. 11 shows the Hausdorff–distance calculations of a cargo and tanker ships, respectively.

The calculations average values of cargo and tanker ships are 0.72 and 0.77, respectively, and the -2σ values are 0.13 and 0.15, respectively. Thus, a value of 0.1 was selected as a reference value by rounding down. Fig. 12 shows the clustering of similar trajectories of cargo and tanker ships, where the number of clusters of similar tracks is 23.

Fig. 13 shows the process of extracting the waypoints of similar ship trajectories. Here, Fig. 13(a) shows the ship trajectory clustering results. The Douglas–Peucker algorithm,

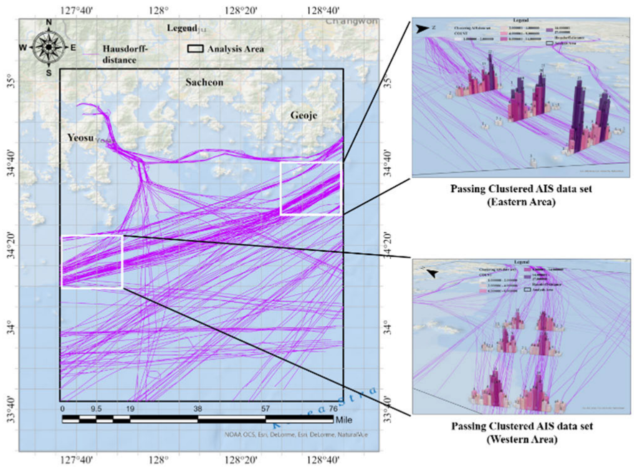


FIGURE 12. Similar ship trajectory clustering results (23 clusters) with 3D result of passing ships.

which simplifies complex trajectories, was applied to the trajectory created by clustering, and the results are shown in Fig. 13(b). Fig. 13(c) shows the detection of vertex points where one trajectory line meets another trajectory line, and Fig. 13(d) illustrated the results of DBSCAN algorithm based on the extracted vertex points. Here, the parameter values of DBSCAN were $\varepsilon = 1000\ m$ and $minPts = 50$ pieces.

The vertex points of individual ships represent waypoints; thus, the median center value was extracted (Fig. 13(e)) to extract a single accurate waypoint. If these waypoints are connected with lines, they form a maritime traffic route. As a result, as the data before applying the proposed framework, 871 waypoints were extracted, but performing the proposed method, 43 waypoints were extracted. In addition, as a result of analysis based on the framework, it was reduced from 457 ships to 195 ships, and the number of vertex points decreased from 271,626 to 8,683, as shown in Table 4.

A large amount of data went through the proposed framework, and the result show that the number of vertex points was reduced to 96.9%. These points are selected as waypoints by DBSCAN, and the waypoints can be decreased up to 95.2%. Based on these results, KDE analysis can be performed, and the route width can be accessed from a statistical perspective. Thus, the KDE analysis result of the main route can be obtained.

C. APPLYING KDE ANALYSIS RESULT AND CALCULATING ROUTE WIDTH BASED ON STATISTICS

Similar tracks were collected by clustering ship trajectories, and waypoints were extracted by clustering vertex points. If data preprocessing and clustering of similar tracks are not performed, detected routes can be in the all ocean space. Note that merchant ships sail to the next port, and related economic and safety factors are included; thus, they have similar appearances. Therefore, a high density (red area) is shown in a lot of traffic, and the low density (green area) is shown with little traffic. Here, Initial maritime traffic route (IMTR)-KDE-90% means maritime traffic route of initial

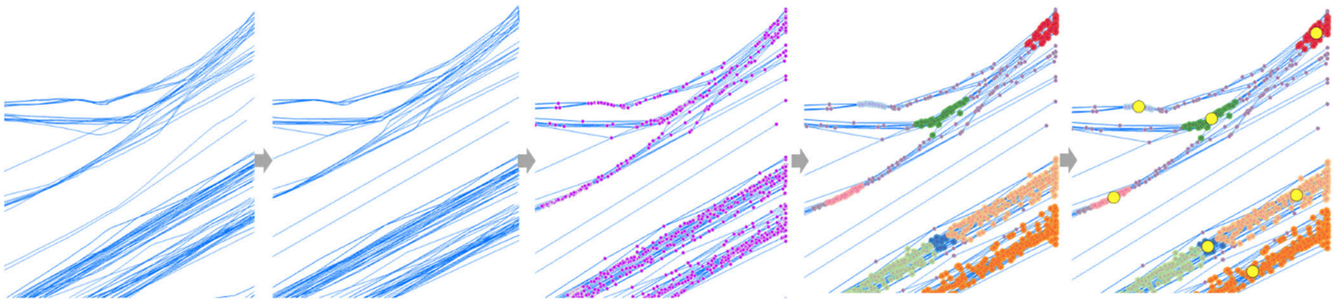


FIGURE 13. Extracting waypoints from similar ship trajectory clusters.

TABLE 4. The compared results for number of waypoints, total ships, and vertex points.

Categorization	Waypoint	Total ships	Vertex points
Initial AIS data set	871	457	271,626
Clustered AIS data set	43	195	8,683
Performance of reduction rate	95.2%	70.0%	96.9%

TABLE 5. Results of ship's route width calculation.

Maritime Traffic Route	Way-point 5 (m)	Way-point 4 (m)	Way-point 3 (m)	Way-point 2 (m)	Way-point 10 (m)	Mean (m)
IMTR-KDE-90%	4,005	2,114	3,711	3,438	3,080	3,269
CMTR-KDE-75%	1,980	1,276	2544	2,616	1,892	2,061
CMTR-KDE-90%	1,832	1,118	2,015	2,554	1,064	1,716

AIS data set. In the initial AIS data, where the clustering algorithm is not applied, 90% of the total maritime traffic volume is expressed by KDE. Clustered maritime traffic route (CMTR)-KDE-75% represents the maritime traffic route of clustered AIS data set. Based on the data applied with the framework of this study, 75% of the total maritime traffic volume is expressed as KDE. In addition, CMTR-KDE-90% is the result of KDE representing 90% of the total maritime traffic volume based on the data applied with the framework of this study.

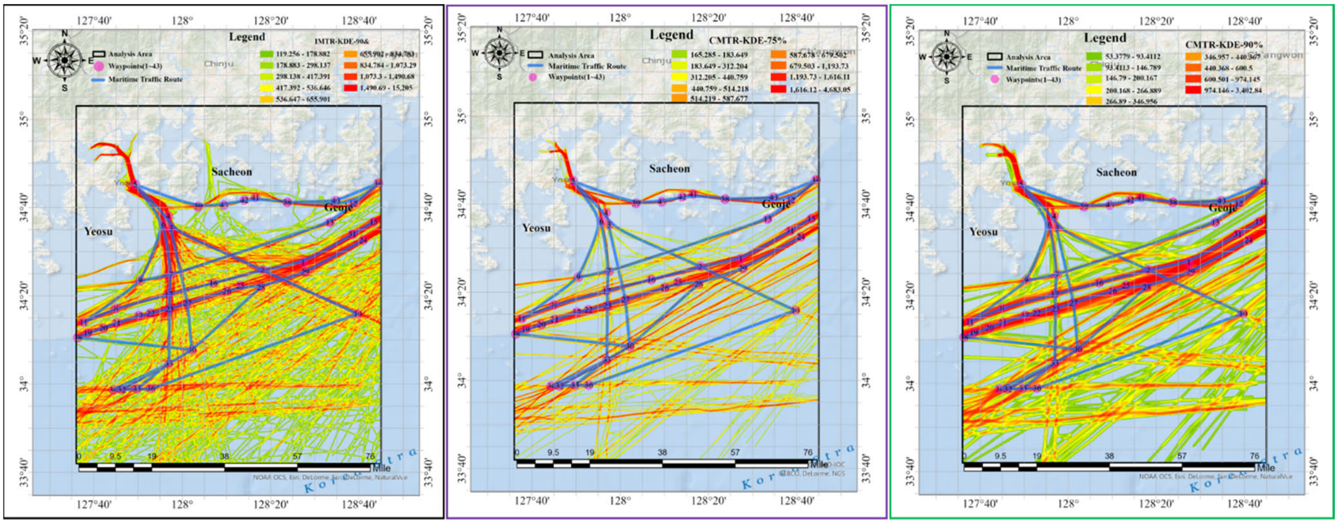
Fig. 14 shows the KDE analysis results obtained with and without the proposed framework, where KDE analysis was performed based on the 90 percentage of total maritime traffic. Fig. 14(a) shows the KDE results for 90% total maritime traffic without the proposed framework, and it appears that most of the analyzed area was traversed by ships. Fig. 14(b) shows the KDE results for 75% total maritime traffic with the proposed framework, and Fig. 14(c) shows the KDE results applied to 90% of the total maritime traffic volume. In addition, the total length of the maritime traffic route is approximately 930 km, and the total number of legs is 17. The biggest reason for applying density analysis is that the route width can be calculated quantitatively from a statistical perspective; therefore, to calculate a navigable area, a random route is selected. The result of the route width in south-eastern bound Way$Point = 5 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 10$ is shown in Table 5.

The south-eastern route has the mean width of 2,061.6 m when the proposed framework was applied with 90% of the total maritime traffic. This demonstrate a lot of difference compared to the when the proposed framework was not used. This method can prevent excessive navigation area in ocean space and avoid friction with other maritime users. In addition, this proposed framework can be applied to the waypoint and cross track limit of MASS because the framework has the advantage of being clearly extracted as a maritime traffic route.

D. EVALUATION BASED ON KDE ANALYSIS

The density-based traffic route creation can visually identify areas where maritime traffic is and is not concentrated. The area with the highest density, i.e., the largest legend (red area), is the main traffic route traveled by many ships. However, it is difficult to understand the main flow because the passage pattern is broken or dispersed in various parts of the main route.

If the ocean area set as a waypoint shows a large difference according to the KDE calculation result, the reliability of the route of a vessel moving in a certain direction will be reduced. Therefore, the KDE calculation result can be described as a main traffic route when it is kept constant without significant differences depending on the route. In other words, if the KDE calculation results with waypoints show a lot of difference, these results should not be used to create a route. Therefore, the density calculation results of the waypoint ocean area were compared and evaluated. Fig. 15 shows the KDE calculation result of waypoint–detection. IMTR-KDE-90% showed the largest value at 8,584.6 (trajectory line/km^2) and the smallest value at 588.54, with an average of 1,893.83. CMTR-KDE-75% was 3,037.1 at maximum, 714.45 at minimum, and 1,656.24 on average. In addition, CMTR-KDE-90% was 2,698.76 at maximum, 740.5 at minimum, and 1,382.30 on average. Therefore, the KDE analysis result of 90% of maritime traffic



(a) IMTR-KDE-90% (b) CMTR-KDE-75% (c) CMTR-KDE-90%

FIGURE 14. Comparison of the general KDE analysis results without the proposed framework.

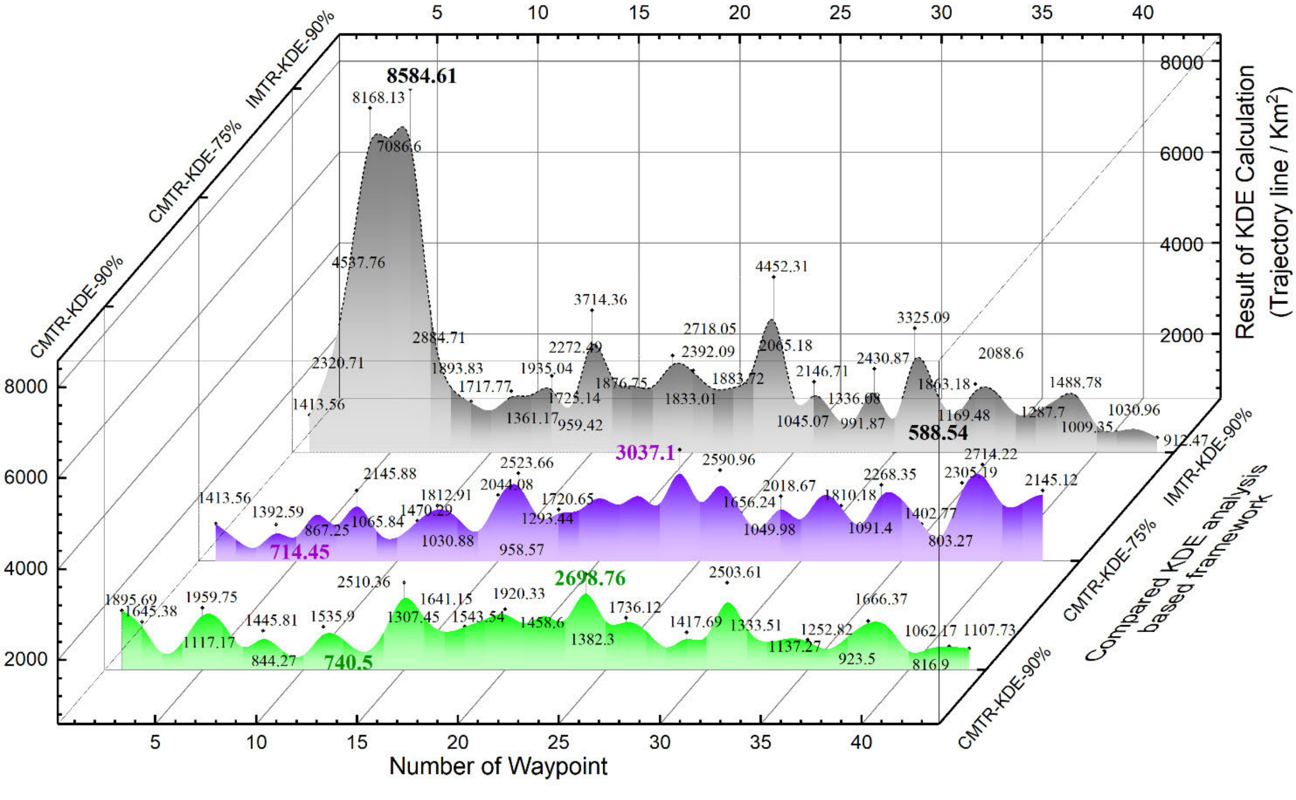


FIGURE 15. Waterfall graph visualization results according to the KDE calculation result obtained with the proposed framework.

obtained using the proposed framework with the smallest difference between the maximum and minimum values was found to be the best. The density-based evaluation was performed to select the main route. Generally, this area was investigated to prove that it is unique in terms of ship passage. In addition, there is a limit to selecting potentials all areas as traffic routes; thus, 75% and 90% of the total maritime traffic volume were applied, and the route width was indicated.

V. DISCUSSION

Unlike urban road traffic, maritime traffic involves larger spaces, various environmental factors, and irregular maritime traffic characteristics. To solve this problem, the ship's operation pattern should be considered for extracting maritime traffic routes based on ship type and geographic characteristics. Density-based extraction analysis to detect maritime traffic routes and networks, automatic anomaly detection in ship navigation patterns, graph-based extraction analysis, application of the Ornstein–Uhlenbeck process, method of STSO, and detected centerline [8], [15], [31], [38]. In common, the studies related to maritime route network were a method that uses a vast amount of AIS data from large ocean space, detects maritime routes, finds waypoints, and finds networks that connect them. The maritime route connected by lines plays an important role in the design of routes and safety diagnosis by finding networks. However, as ships increase in size, they cannot use all routes and similar routes. Therefore, the safety route should be based on the appropriate water depth and route width, and the routes connected by waypoints should be connected without a large difference in the KDE calculated values.

This study deviated from the method of detecting an overall network in large ocean space and focused on detecting a traffic route near a specific large port. In addition, it evaluated whether the extraction routes were well connected. As shown in Fig. 14(a), the KDE analysis of initial AIS data appears to have almost passed in the ocean area. Generally, it is possible to determine which ocean area ships can pass; however, it is difficult to find the main route between dispersed routes. Therefore, Fig. 14(b) shows the KDE analysis of 75% of maritime traffic with the proposed framework, and Fig. 14(c) is the KDE analysis result of 90% of the maritime traffic. The main route was extracted by clustering similar tracks, and waypoint detection via DBCSAN obtained a more concise and accurate route. From a route width perspective, the larger the width, the better; however, in coastal waters, if the route width is excessive, it may overlap with navigational obstacles, which induce unsafe navigation.

In addition, if the routes connecting waypoints are broken or disconnected, route reliability will be reduced. Therefore, as shown in Fig. 15, the most stable route is evaluated as the difference between the maximum and minimum density values is smaller based on the KDE calculation results.

Therefore, this study selected the result of CMTR-KDE-90% with the smallest difference. The ocean space is used by not only merchant ships, but also various entities, e.g., fishing activities, resource extraction, military districts, and marine energy development. If an accurate shipping area is selected first, operation safety and marine environment protection can be promoted, and if a waypoint is created, it will be applied effectively to MASS. As a result, future maritime routes can be created automatically based on highly reliable waypoints.

VI. CONCLUSION AND FUTURE WORK

In this study, the maritime traffic route framework has been proposed. The proposed framework represents a new route detection method that considers both total maritime traffic and statistics to calculate ship routes, including route width. Based on a vessel's movement pattern, the vertex points are extracted to identify the waypoints of a vessel. In addition, the route evaluation through the KDE analysis result indicates that the created route is considered according to the type and size of the ship. To represent the main traffic route, proposed framework can effectively represent the navigational area of a ship, and it can be applied to route design and safety diagnosis. In addition, the framework can build a maritime transport network at the national level beyond a specific port. In future, when generating a maritime traffic network, it will be necessary to subdivide the network according to the purpose, type, and size of each vessel, and it will be necessary to create guidelines for selecting the correct route width. Further, combining the extraction of maritime traffic route and density analysis will contribute to the development of the MASS technology.

REFERENCES

[1] B. Murray and L. P. Perera, "An AIS-based deep learning framework for regional ship behavior prediction," *Rel. Eng. Syst. Saf.*, vol. 215, Nov. 2021, Art. no. 107819.

[2] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G.-B. Huang, "Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1559–1582, May 2018.

[3] *UNCTAD, Review of Maritime Transport 2019*, United Nation Publication, New York, NY, USA, 2019.

[4] J. Weng and D. Yang, "Investigation of shipping accident injury severity and mortality," *Accident Anal. Prevention*, vol. 76, pp. 92–101, Mar. 2015.

[5] J.-S. Lee, W.-J. Son, H.-T. Lee, and I.-S. Cho, "Verification of novel maritime route extraction using kernel density estimation analysis with automatic identification system data," *J. Mar. Sci. Eng.*, vol. 8, no. 5, p. 375, May 2020.

[6] J.-S. Lee and I.-S. Cho, "Extracting the maritime traffic route in Korea based on probabilistic approach using automatic identification system big data," *Appl. Sci.*, vol. 12, no. 2, p. 635, Jan. 2022.

[7] C. B. Moorthy, M. K. Deshmukh, and D. Mukherejee, "Optimal location of wind turbines in a wind farm using genetic algorithm," *TELKOMNIKA Indonesian J. Electr. Eng.*, vol. 12, no. 8, pp. 658–665, Aug. 2014.

[8] J.-S. Lee, W.-J. Son, H.-T. Lee, and I.-S. Cho, "A study on the factors affecting optimal site of offshore wind farm from the perspective of maritime traffic using spatial analysis," *Korea Soc. Coastal Disaster Prevention*, vol. 7, no. 2, pp. 85–96, Apr. 2020.

[9] Z. Yan, Y. Xiao, L. Cheng, R. He, X. Ruan, X. Zhou, M. Li, and R. Bin, "Exploring AIS data for intelligent maritime routes extraction," *Appl. Ocean Res.*, vol. 101, Aug. 2020, Art. no. 102271.

[10] *International Convention for the Safety of Life at Sea (SOLAS)*. (Nov. 1974). [Online]. Available: http://www.imo.org/About/Conventions/ListOfConventions/Pages/International-Convention-for-the-Safety-of-Life-at-Sea-%28SOLAS%29,-1974.aspx

[11] W.-J. Son, J.-S. Lee, H.-T. Lee, and I.-S. Cho, "An investigation of the ship safety distance for bridges across waterways based on traffic distribution," *J. Mar. Sci. Eng.*, vol. 8, no. 5, p. 331, May 2020.

[12] W. Hasbi, "Tracking capability and detection probability assessment of space-based automatic identification system (AIS) from equatorial and polar orbiting satellites constellation," *IEEE Access*, vol. 8, pp. 184120–184136, 2020.

[13] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.

[14] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," in *Classics Cartography: Reflections Influential Articles From Cartographica*, vol. 2. Toronto, ON, Canada: Univ. Toronto Press, 1973.

[15] O. Daae Lampe and H. Hauser, "Interactive visualization of streaming data with kernel density estimation," in *Proc. IEEE Pacific Visualizat. Symp.*, Mar. 2011, pp. 1–8.

[16] E. d'Afflisio, P. Braca, L. M. Millefiori, and P. Willett, "Maritime anomaly detection based on mean-reverting stochastic processes applied to a real-world scenario," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 1171–1177.

[17] J. Montewka, S. Ehlers, F. Goerlandt, T. Hinz, K. Tabri, and P. Kujala, "A framework for risk assessment for maritime transportation systems—A case study for open sea collisions involving RoPax vessels," *Rel. Eng. Syst. Saf.*, vol. 124, pp. 142–157, Apr. 2014.

[18] F. Goerlandt and J. Montewka, "Maritime transportation risk analysis: Review and analysis in light of some foundational issues," *Rel. Eng. Syst. Saf.*, vol. 138, pp. 115–134, Jun. 2015.
[19] K. D. Schwehr and P. A. McGillivary, "Marine ship automatic identification system (AIS) for enhanced coastal security capabilities: An oil spill tracking application," in *Proc. OCEANS*, Sep. 2007, pp. 1–9.
[20] A. Serry, *The Automatic Identification System (AIS): A Data Source for Studying Maritime Traffic*. Barcelona, Spain: Maritime Transport, 2018. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01724104/document
[21] H. Namgung and J.-S. Kim, "Collision risk inference system for maritime autonomous surface ships using COLREGs rules compliant collision avoidance," *IEEE Access*, vol. 9, pp. 7823–7835, 2021.
[22] L. Zhang, Q. Meng, and T. Fang Fwa, "Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters," *Transp. Res. E, Logistics Transp. Rev.*, vol. 129, pp. 287–304, Sep. 2019.
[23] H.-T. Lee, J.-S. Lee, H. Yang, and I.-S. Cho, "An AIS data-driven approach to analyze the pattern of ship trajectories in ports using the DBSCAN algorithm," *Appl. Sci.*, vol. 11, no. 2, p. 799, Jan. 2021.
[24] H.-T. Lee, J.-S. Lee, W.-J. Son, and I.-S. Cho, "Development of machine learning strategy for predicting the risk range of Ship's berthing velocity," *J. Mar. Sci. Eng.*, vol. 8, no. 5, p. 376, May 2020.
[25] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
[26] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[27] M. Liang, R. W. Liu, S. Li, Z. Xiao, X. Liu, and F. Lu, "An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation," *Ocean Eng.*, vol. 225, Apr. 2021, Art. no. 108803.
[28] R. W. Liu, M. Liang, J. Nie, W. Y. B. Lim, Y. Zhang, and M. Guizani, "Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime Internet of Things," *IEEE Trans. Net. Sci. Eng.*, vol. 14, no. 8, pp. 2–3, Aug. 2015.
[29] L. Wang, P. Chen, L. Chen, and J. Mou, "Ship AIS trajectory clustering: An HDBSCAN-based approach," *J. Mar. Sci. Eng.*, vol. 9, no. 6, p. 566, May 2021.
[30] V. Fernandez Arguedas, G. Pallotta, and M. Vespe, "Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 722–732, Mar. 2018.
[31] N. Forti, L. M. Millefiori, and P. Braca, "Unsupervised extraction of maritime patterns of life from automatic identification system data," in *Proc. OCEANS Marseille*, Jun. 2019, pp. 1–5.
[32] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2313–2330, Oct. 2016.
[33] L. M. Millefiori, P. Braca, and G. Arcieri, "Scalable distributed change detection and its application to maritime traffic," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 1650–1657.
[34] L. M. Millefiori, P. Braca, and P. Willett, "Consistent estimation of randomly sampled Ornstein–Uhlenbeck process long-run mean for long-term target state prediction," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1562–1566, Nov. 2016.
[35] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based al- gorithm for discovering clusters in large spatial databases with noise," in *Int. Conf. Data Mining Knowl. Discovery*, vol. 4, pp. 226–231, 1996.
[36] R. Xu and D. C. Wunsch II, *Sequential Data Clustering*. Hoboken, NJ, USA: Wiley, 2008, ch. 7, pp. 179–212.
[37] G. Wang, J. Meng, and Y. Han, "Extraction of maritime road networks from large-scale AIS data," *IEEE Access*, vol. 7, pp. 123035–123048, 2019.
[38] Z. Yan, Y. Xiao, L. Cheng, S. Chen, X. Zhou, X. Ruan, M. Li, R. He, and B. Ran, "Analysis of global marine oil trade based on automatic identification system (AIS) data," *J. Transp. Geography*, vol. 83, Feb. 2020, Art. no. 102637.
[39] *Technical Characteristics for an Automatic Identification System Using Time Division Multiple Access in the VHF Maritime Mobile Frequency Band*, document ITU-R M. 1371-5, Radiocommunication Sector of International Telecommunication Union (ITU-R), 2014, p. 114.
[40] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.
[41] L. Zhao and G. Shi, "A trajectory clustering method based on Douglas-Peucker compression and density for marine traffic pattern recognition," *Ocean Eng.*, vol. 172, pp. 167–456, Jan. 2019.
[42] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica, Int. J. Geographic Inf. Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.
[43] A. Ram, S. Jalal, A. S. Jalal, and M. Kumar, "A density based algorithm for discovering density varied clusters in large spatial databases," *Int. J. Comput. Appl.*, vol. 3, no. 6, pp. 1–4, Jun. 2010.
[44] N. Lu, M. Liang, R. Zheng, and R. W. Liu, "Historical AIS data-driven unsupervised automatic extraction of directional maritime traffic networks," in *Proc. IEEE 5th Int. Conf. Cloud Comput. Big Data Analytics (ICCCBDA)*, Apr. 2020, pp. 7–12.
[45] *United Nations Convention on the Law of the Sea*, UNCLOS, UN (United Nations), New York, NY, USA, 1982, p. 45.
[46] M. Maritime and C. Agency, *MGN 543(M+F) Safety of Navigation: Offshore Renewable Energy Installations (ORELs)*. Guidance on UK Navigation Practice, Safety and Emergency Response, MCA, Southampton, U.K., 2016, p. 18.
[47] *Marine Spatial Plan for the Belgian Part of the North Sea*, Federal Public Service, Brussels, Belgium, 2016, pp. 42–43. [Online]. Available: https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/19094275/Summary%20Marine%20Spatial%20Plan.pdf
[48] A. H. Kvanli, R. J. Pavur, and K. B. Keeling, *Concise Managerial Statistics*. Boston, MA, USA: Cengage Learning, 2006, pp. 81–82.

JEONG-SEOK LEE received the B.S. degree from the Division of Navigation Science, Korea Maritime and Ocean University, Busan, South Korea, in 2014. From 2014 to 2017, he worked as a Navigator for the third and second officers at SK Shipping Company, with experience in bulk carriers, product oil-tanker carriers, and very large crude oil carriers. In 2018, he performed related works as a Maritime Simulator Operator and is conducting research using maritime traffic data. Conducted research using maritime data-based computing algorithm. In particular, uses the GIS program in maritime. His research interest includes fusion research using artificial intelligence techniques.

HYEONG-TAK LEE received the B.S. degree in nautical science course from the Department of Maritime Transportation Sciences, Korea Maritime & Ocean University (KMOU), Busan, South Korea, in 2013, the M.E. degree from the Graduate School, KMOU, in 2019, and the Ph.D. degree from the Ocean Science and Technology (OST) School, KMOU, in 2022. He worked at Hanjin Shipping Company Ltd., as a Navigating Officer and worked for four years on container vessels. He also worked at Training Ship HANBADA as an Instructor and a Chief Officer for two and a half years. His research interests include maritime artificial intelligence and data science.

IK-SOON CHO received the B.Sc. degree in nautical science course, in 1996, the M.E. degree in maritime transportation science from Korea Maritime & Ocean University (KMOU), Busan, South Korea, in 2000, and the Ph.D. degree from the Graduate School of Science and Technology, Kobe University, Kobe, Japan, in 2005. He has worked with the Traffic Safety Audit Center, KMOU based on the Korea Maritime Safety Act, where he is currently working as a Professor with the Division of Navigation Convergences Studies, College of Maritime Sciences. His research interests include maritime safety evaluation, ship operation, and mooring safety analysis based on big-data using the machine learning and deep learning algorithm.