TTMRN

A topological-geometric two-layer maritime route network modeling for ship intelligent navigation

Zhou, Chunhui; Xiang, Jiale; Huang, Hongxun; Yan, Yi; Huang, Liang; Wen, Yuanqiao; Xiao, Changshi

DOI
10.1016/j.oceaneng.2023.115884

Publication date
2023

Document Version
Final published version

Published in
Ocean Engineering

Ocean Engineering 287 (2023) 115884

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

TTMRN: A topological-geometric two-layer maritime route network modeling for ship intelligent navigation

Chunhui Zhou [a,b,e], Jiale Xiang [a,b], Hongxun Huang [a,b,*], Yi Yan [a,b], Liang Huang [c,f,g,**], Yuanqiao Wen [c,d], Changshi Xiao [a,b]

[a] *School of Navigation, Wuhan University of Technology, Wuhan, 430063, China*
[b] *Hubei Key Laboratory of Inland Shipping Technology, Wuhan, 430063, China*
[c] *Intelligent Transportation System Research Center, Wuhan University of Technology, Wuhan, China*
[d] *National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, China*
[e] *China Laboratory of Transport Pollution Control and Monitoring Technology, Beijing, China*
[f] *Safety and Security Science Group, Faculty of Technology, Policy and Management, Delft University of Technology, Delft, 2628BX, Netherlands*
[g] *Sanya Science and Education Innovation Park, Wuhan University of Technology, Sanya 572024, PR China*

ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

Keywords:
Trajectory feature points
Cluster analysis
Maritime route network
Intelligent navigation
AIS data

ABSTRACT

The construction of maritime route networks holds significant importance for autonomous navigation of vessels. In this study, a two-layer maritime route network modeling method based on huge amounts of ship trajectory data was proposed. Firstly, we introduce a novel method for extracting nodes of the marine route network, which identifies feature points in ship trajectories through clustering. Secondly, we use a spatial computing method to transform ship trajectory data into a sequence of waypoint regions and establish a node connection matrix to realize the nodes' connection of the topological layer route network. And routes are extracted between waypoint regions to characterize the connection relationship of the geometric layer network. Finally, by connecting nodes of the topological layer with the support of the connection matrix and waypoint regions of the geometric layer with the route, the two-layer maritime route network that combines topological and geometric layers is constructed. The proposed method was applied to the waters of Vancouver, successfully constructing a topological-geometric two-layer maritime route network. Overall, the proposed method is beneficial for improving the safety and efficiency of autonomous navigation of ships, and has a positive impact on the development of smart shipping industry.

1. Introduction

Autonomy and intelligence in ships have become hot research topics (de Vos et al., 2021; Liu et al., 2016). The initial task of intelligent ship operation is route design, or path planning(Yang et al., 2015; Zhou et al., 2020). For ships to follow the planned route, the ability to detect deviations from the route is crucial. Thus, intelligent navigation methods based on the designed route are necessary to ensure smooth navigation to the destination. Constructing a marine route network containing topological and geometric information can not only enhance intelligent ship route design but also provide geometric information about the designed route to prevent deviations from the intended path, just as road networks support car path planning and deviation detection. The construction of a marine route network has the potential to significantly contribute to intelligent ship navigation.

Using high-resolution remote sensing image data and image recognition tools to extract road networks is a common strategy(Lin et al., 2020). Techniques that employ vehicle movement data, especially road operation status data, have also been extensively utilized(Hashemi, 2019; Krumm, 2011; Yang et al., 2018). However, research regarding the extraction of marine route networks is comparatively scarce. This is mainly due to the indeterminate and ambiguous boundaries of maritime routes, which cannot be constructed with visualization methods. The massive quantity of ship trajectory data contains numerous patterns that are formed by ships navigating at sea(Karatas et al., 2021), which encodes relevant information about the route network. Therefore, mining

* Corresponding author. School of Navigation, Wuhan University of Technology, Wuhan, 430063, China.
** Corresponding author. Intelligent Transportation System Research Center, Wuhan University of Technology, Wuhan, China.
E-mail addresses: huanghx@whut.edu.cn (H. Huang), leung.huang@whut.edu.cn (L. Huang).

https://doi.org/10.1016/j.oceaneng.2023.115884
Received 27 April 2023; Received in revised form 8 August 2023; Accepted 17 September 2023
Available online 23 September 2023

data from the vast amount of ship track becomes a feasible approach for extracting marine route network information(Mazzarella et al., 2013, 2014; Yan et al., 2020; Zheng, 2015).

In this paper, we propose a two-layer maritime route network modeling and extracting method that incorporates topology and geometry for intelligent navigation of ships. The topological route network comprises topological nodes and edges that express the connectivity of the route network, which can be used for route design and path search of intelligent ships. In addition to that, the geometric route network provides further information regarding the geographical range of waypoints and routes, enabling the determination of the boundary of the waters available for safe navigation and preventing ships from yawing during the navigation of intelligent ships. Our approach offers a comprehensive solution to modeling and extracting maritime route networks for intelligent ship navigation.

The other chapters of this paper are organized as follows: Section 2 is a literature review; Section 3 introduces the topological-geometric two-layer maritime route network model and construction methods; Section 4 presents experimental cases and analysis; Section 5 discusses the method and Section 6 concludes.

2. Literature review

Navigational route planning plays a crucial role in navigation, where both safety and economy must be taken into account. Consequently, ships tend to follow established routes that lead to a clustering of ship trajectories. To extract these routes, Pallotta et al. (2013a) proposed a TREAD (Traffic Route Extraction for Anomaly Detection) framework that utilizes unsupervised and incremental learning methods to abstract the route network as a combination of nodes and edges. The TREAD method employs trajectory points of ships during port stays, offshore platform operations, and water entry/exit as nodes in the route network. Incremental DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is utilized to extend and update the nodes, while the regular routes taken by vessels between different route points are considered edges in the network. Then, he employed trajectory clustering to distinguish between ship trajectory clusters associated with different imports and exports in the study waters(Pallotta et al., 2013b). The resulting route network model offers topological information that can facilitate route design and path planning for intelligent ships.

Arguedas et al. (2018) further enhanced the method by defining trajectory points corresponding to turning behavior as route network nodes, thereby diversifying the types of nodes present in the network. Wu et al. (2017) and Lei et al. (2016) employed the gridding approach to partition the study waters and then utilized the grid as the fundamental unit for reorganizing the original traffic data. Vettor and Soares (2015) divided voluntary observing ship (VOS) data into density images based on geographic distribution and then extracted the main shipping routes in the North Atlantic to evaluate the relative strength and seasonal trends of maritime traffic.

By utilizing the traffic density statistics to identify grid areas that meet the threshold, these regions were used as route network nodes, while continuous grid areas between the nodes that did not meet the threshold were used as edges to build the route network. It is worth noting that in the study of the route network, nodes and edges are considered the fundamental components for route network construction (Varlamis et al., 2019), which permits the establishment of a route network structure with topological information that significantly simplifies the path planning and route design process for vessels. Although the grid-based approach is simple, it has a drawback that the grid size needs to be determined in advance.

Safe navigation of smart ships requires not only pre-designed routes but also the geometric information of each route segment, including boundaries and width of routes. To address this, scholars have focused on analyzing the edges of the route network to obtain boundary geometric features through ship trajectory data mining. For instance, Lee et al. (2007) proposed a segmented trajectory clustering method known as TRACLUS, which partitions trajectories into line segments and groups similar line segments into clusters to identify sub-trajectories. Similarly, Wang et al. (2019) used triangulation to segment the route and extract its centerline information. Another approach was taken by Lee et al. (2022) and his colleagues using 75% kernel density (KDE) as the width information of the route to regulate ship navigation. Gonzalez et al. (2014) used trajectory clustering to extract ship routes from segmented ship trajectory data. Li et al. (2018) proposed a method to measure the similarity between different trajectories using merge distance (MD) and constructed a suitable low-dimensional spatial expression of trajectory similarity using multidimensional scaling (MDS). Then, an improved DBSCAN algorithm was proposed to cluster spatial points for obtaining optimal clustering. These methods can provide ships with necessary segment geometry information to detect if they are sailing off-course and ensure safe navigation. However, the selection of clustering parameters still needs to be determined by human intervention, which leads to uncertainty in the clustering results. As well as the construction of a route network for large-scale marine areas using trajectory clustering methods is challenging due to the significant computational power required to handle the vast amount of trajectory data.

Although existing methods for maritime route network modeling and extraction have made significant progress in extracting either topological or geometric information, they still have limitations in meeting the requirements for intelligent navigation of ships. For instance, topology-based methods only extract the topological information of the route network and lack the geometric information needed for autonomous correction of deviation, while geometry-based methods only extract the geometric information of the route and cannot support path planning. To address these limitations, this paper proposes a novel two-layer approach for maritime route network modeling and extraction that integrates both topology and geometry information. The two-layer maritime route network enabling ships to safely navigate along the planned routes and autonomously correct any deviations that may occur during the voyage.

3. Methodology

3.1. Definition and construction of two-layer maritime route network model

As depicted in Fig. 1, this paper proposes a novel two-layer maritime route network modeling and extraction method to support the intelligent navigation of ships. The first layer represents the topological route network, which captures the connectivity between each topological node in the route network and enables functions such as path planning, network analysis, and anomaly detection. The second layer represents the geometric route network, which describes the geometric information of each waypoint area and the boundary of the route. This layer enables functions such as waypoint area analysis and yaw detection. The two layers of the route network model are intertwined and interconnected, creating a comprehensive model that effectively depicts the network characteristics of ship navigation in the study area. The proposed model provides multiple services to support intelligent ship navigation.

As depicted in Fig. 2(a), the first layer of the proposed maritime route network model is the topological route network, which comprises interconnected topological nodes and edges. Meanwhile, Fig. 2(b) shows the second layer, i.e., the geometric route network, which includes waypoint regions and geometric edges. The waypoint regions describe the geometric characteristics of each segment, such as its width and boundary, while the geometric edges represent the actual routes between the regions. By combining the topological and geometric layers, the two-layer route network in Fig. 2(c) can realize path planning, network analysis, and abnormality detection functions, and provide waypoint area analysis and yaw detection for intelligent ship navigation as well. Fig. 3 shows the process of the two-layer maritime route

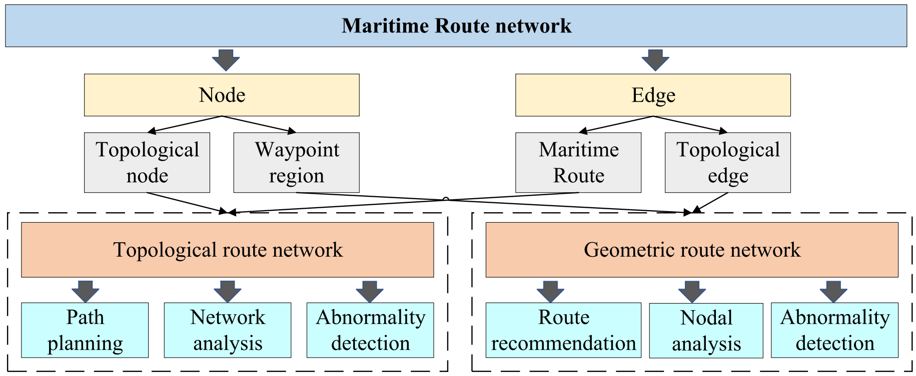


Fig. 1. Two-layer maritime route network model.

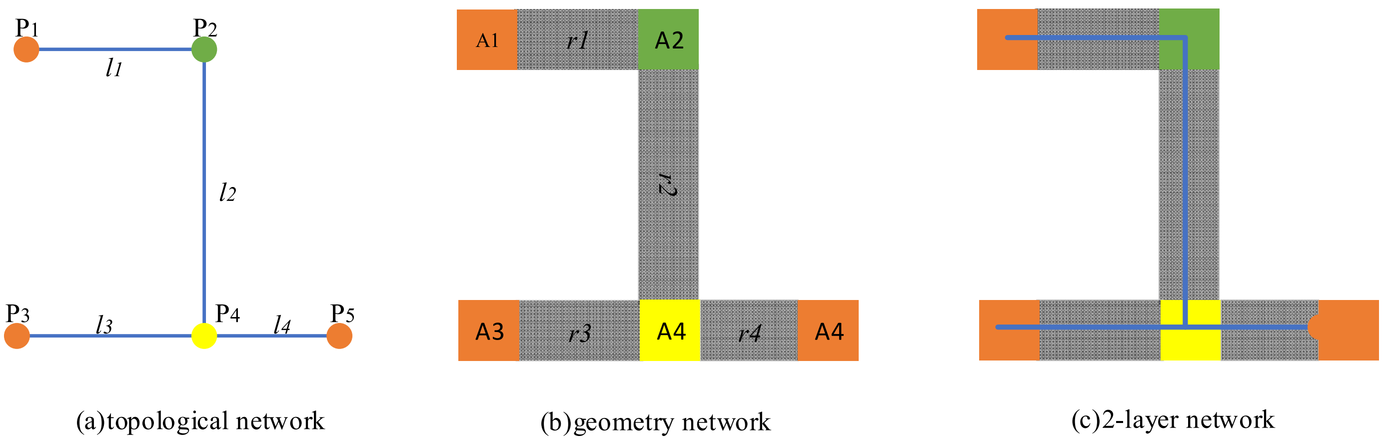


Fig. 2. The two-layer maritime route network.

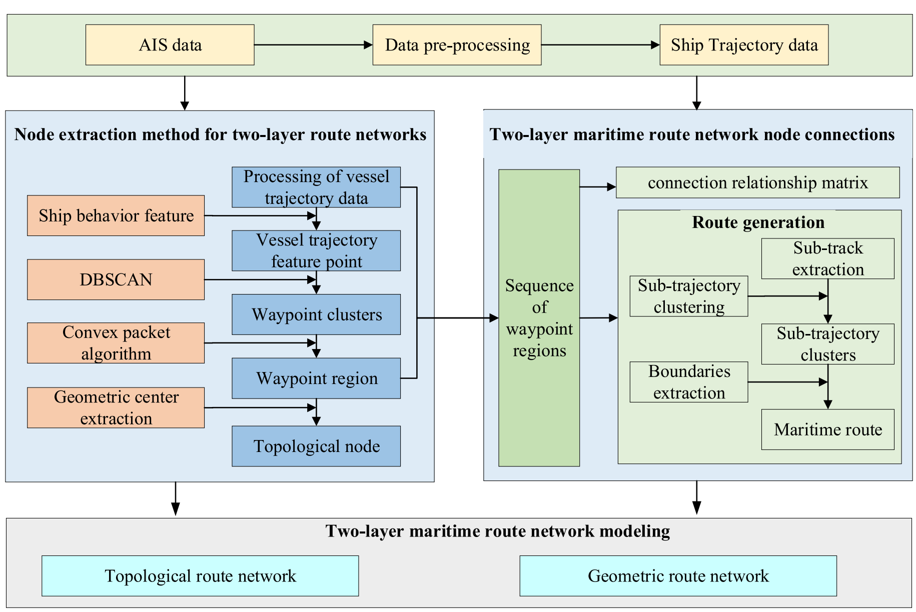


Fig. 3. Two-layer maritime route network construction process.

network.

The Automatic Identification System (AIS) data is a valuable source of ship information, providing essential details such as maritime mobile service identity (MMSI), timestamp, longitude, latitude, sailing speed, and course of ground (COG). In this study, we propose a methodology for constructing a two-layer maritime route network for intelligent navigation using pre-processed AIS data. The construction process is divided into three parts.

In the first step of constructing the two-layer maritime route network, ship behavior features are extracted from the AIS data based on attributes such as heading, speed, and time of the trajectory points. All ship trajectory feature points are then obtained from the pre-processed AIS data. To cluster the feature points, the density-based clustering algorithm is employed to obtain the waypoint clusters. The convex polygon of each waypoint cluster is obtained by applying the convex hull algorithm to form the waypoint region. And the geometric center of the waypoint region is extracted to serve as the topological node in the topological route network.

Each ship trajectory is transformed into a sequence of waypoint regions by utilizing a spatial intersection method. Then, the adjacent numbers in the sequence are extracted to create a connection matrix that serves as the foundation for linking the nodes of the route network. For connected topological nodes, the line segment direct connection is used. In the case of connected waypoint regions, the sub-trajectories between them are extracted to cluster them and obtain sub-trajectory clusters. These clusters' boundaries are then extracted to create polygonal regions, which are utilized as the route connecting waypoint regions.

Finally, the connection matrix is used to connect all the topological nodes and waypoint regions, resulting in the construction of a complete topological-geometric two-layer maritime route network. This network effectively combines the advantages of both topological and geometric models, enabling intelligent navigation of ships with functions such as path planning, network analysis, abnormality detection, waypoint area analysis, and yaw detection.

3.2. Key methods for modeling two-layer route networks

3.2.1. Node extraction

The two-layer maritime route network is composed of waypoint regions and topology nodes, which serve as the network nodes. Fig. 4 shows the extraction process of these nodes. The node extraction process can be divided into four steps: vessel trajectory feature point identification, waypoint clustering, waypoint region generation, and topology node generation. The topology nodes are obtained from the calculation center of the waypoint regions, and their generation method is relatively simple. Therefore, this section mainly focuses on the first three steps.

(1) Vessel trajectory feature point identification

In most cases, ships operate by maintaining a forward course and a steady speed, with changes in ship behavior occurring only in specific areas, such as turning zones or port anchorages. These ship behaviors mainly include berthing, anchoring, turning, entering, and leaving water boundaries. To address these specific ship behaviors, this study classifies the feature points in ship trajectory data into three categories: dwell feature points, turning feature points, and boundary feature points.

This study identifies boundary feature points by analyzing the spatial relationship between ship trajectories and the boundary of the study area. To find the intersection point between the ship trajectory and the boundary, which helps to identify the boundary feature points, the ship trajectory data is spatially intersected with the polygon data of the study area boundary.

To identify the dwell feature points, this study sets a velocity threshold γ and a number threshold n. Specifically, when points with speeds below the set velocity threshold appear continuously in a single ship trajectory, and their count exceeds the set number threshold, these points are identified as dwell feature points.

For identification of turning points in ship trajectories, it is necessary to first determine whether the ship is navigating based on the velocity threshold γ and then calculate the course change α and course change rate β at each trajectory point. The α and β is calculated as shown in Equation (1) and Equation (2). A trajectory point is recognized as a turning feature point when α and β of a trajectory point exceeds the turning threshold μ and turning rate threshold ϑ while the ship is navigating.

$$\alpha_i = \left|cog_{t_{i+1}} - cog_{t_i}\right| \tag{1}$$

$$\beta_i = \left|\frac{cog_{t_{i+1}} - cog_{t_i}}{t_{i+1} - t_i}\right| \tag{2}$$

(2) Waypoint clustering

In this study, the DBSCAN algorithm is adopted to cluster and identify trajectory feature points to detect the clustering areas of trajectory feature points and exclude noise points. During the clustering process, two critical parameters, *eps* and *MinPts*, need to be determined, which directly affect the clustering effect of feature points. For each type of trajectory feature point, clustering is carried out according to the determined parameters, noise points are excluded, and each cluster is extracted and labeled with a clustering tag. This provides a basis for the

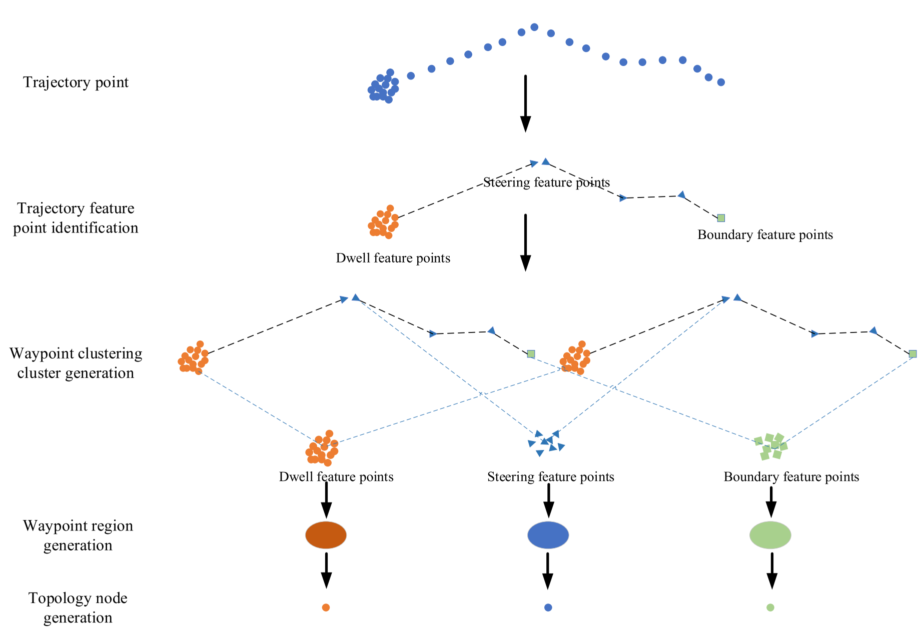


Figure 4. Waypoint region and topology node extraction process.

construction of nodes in the two-layer maritime route network.

(3) Waypoint region and topological node generation

In two-dimensional space, the convex hull is the smallest convex polygon that encloses all points in a point set. Its characteristic is that any point inside the convex polygon can be connected to a point on the convex polygon, and all points on the line segment connecting them lie inside the convex polygon, ensuring that the convex hull can contain all the points in the point set. Waypoint regions are obtained using the convex packet algorithm to calculate the convex polygons of the clusters of waypoints. Then, the geometric center of the convex hull polygon is calculated to obtain the nodes of the topological layer route network.

The nodes of the route network are fundamental elements of the route network model, and the edges are important means of connecting nodes. Nodes can only form a network structure by being correctly connected through edges. Therefore, two issues remain to be addressed. The first is how to correctly connect the nodes of the route network, i.e., the extraction of the connection between nodes. The second is how nodes are connected, i.e., the generation of edges in the route network.

3.2.2. Node connection method for topological maritime network construction

As a vessel travels through various waypoint regions, the connections of these regions are preserved in its trajectory. Therefore, by using spatial topology calculation, the vessel's trajectory can be transformed into a series of waypoint regions to acquire their connectivity. In space, there exist three topological relations between trajectory points and waypoint regions, including contains, intersects, and disjoint. The spatial relationship between the trajectory point and the waypoint region can be determined by the ray method.

Traverse the single ship trajectory sequence and all waypoint regions, and calculate the spatial relationship between trajectory points and waypoint regions. If the trajectory point intersects or is contained in the waypoint region, mark this trajectory point as the number of the waypoint region; if the trajectory point is disjoint from all waypoint regions, mark this trajectory point as 0 and generate the waypoint region sequence set of the trajectory.

Fig. 5 illustrates that the ship trajectory passes through a total of five waypoint regions six times in a sequence, and the trajectory points intersecting the waypoint regions are labeled as 1, 2, 3, 4, 5, and 4. By retaining only the first consecutive repetitive value in the sequence, the final waypoint sequence wp_seq = [1, 2, 3, 4, 5, 4] is obtained, indicating that two adjacent waypoint regions are connected and reachable to each other. The waypoint sequence includes five connected waypoint pairs, namely (1, 2), (2, 3), (3, 4), (4, 5), and (5, 4), which are stored in the corresponding positions in the matrix to construct the connection matrix of the two-layer waypoint network nodes.

The connection matrix is a data structure that stores the connection between nodes in the route network. After obtaining the waypoint sequences of all trajectories, adjacent pairs of points are extracted and stored in the corresponding positions in the connection matrix. The generation method of the connection matrix is as follows:

Firstly, an empty matrix D is constructed, where the rows and columns correspond to the serial number of the waypoint regions. Then, all the ship trajectory waypoint sequences are traversed, and for each sequence, all waypoint pairs (i,j) are identified. If (i,j) exists in the sequence, the value of D[i,j] is incremented by 1. The resulting matrix D contains the number of connections between waypoint regions, indicating the number of times that a ship travels between two regions. For example, D[i,j] represents the number of times a ship has traveled from waypoint region i to waypoint region j.

Through calculation of the connection matrix of the maritime network, the connection between nodes can be determined. Specifically, a non-zero value of matrix D[i,j] indicates a connection between node i and node j, which can be established by a topological edge. By traversing the entire connection matrix, all notes meeting the connection value can be connected through the use of topological edges, thus constructing a complete topological maritime network.

3.2.3. Route generation method for geometric maritime network construction

The procedure for generating waypoints is depicted in Fig. 6, which involves three steps: extracting sub-trajectories between waypoint regions, clustering sub-trajectories, and calculating the boundary of sub-trajectory clusters.

(1) Extracting sub-trajectories between waypoint regions

When converting ship trajectory sequences to waypoint region sequences, ship trajectory points intersect or are contained in the waypoint region and are labeled with the waypoint region number, while those disjointing the waypoint region are labeled as 0. This labeling enables the extraction of sub-trajectories between waypoint regions.

Firstly, for the waypoint region sequence of a certain ship trajectory, if the label of the i-th point in the trajectory is not equal to 0, and the label of the i-th point and the label of the (i+1)-th point in the trajectory is different, it indicates that the point intersects or is contained in the waypoint region corresponding to the label. Then, the index position of the point Pi in the vessel trajectory sequence and the corresponding waypoint region label is recorded as index_i. Similarly, the index position and corresponding waypoint region label of the next trajectory

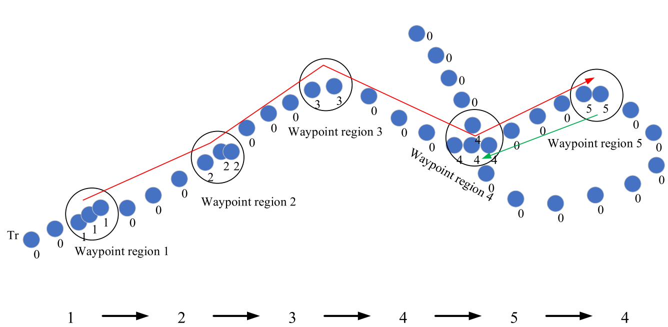


Fig. 5. Waypoint sequence generation.

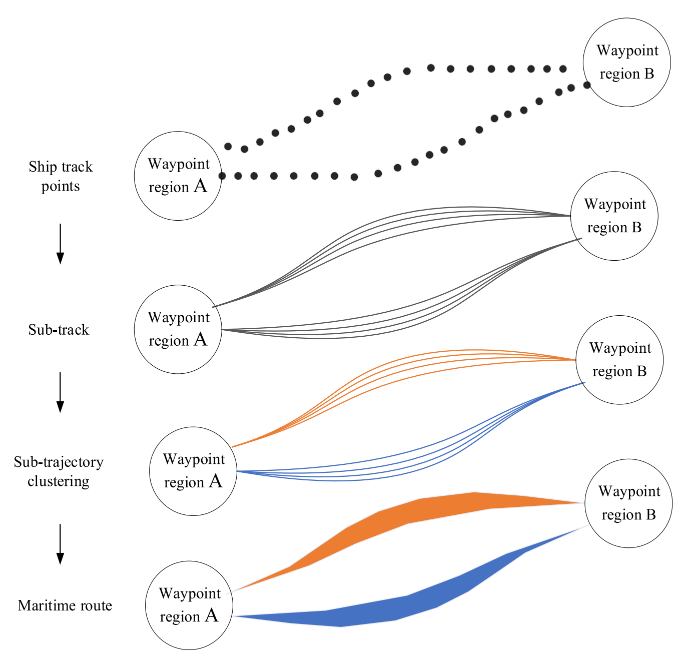


Figure 6. Diagram of the route generation process.

point that intersects or is contained in the other waypoint region are recorded as index_j. If the labels of these two waypoint regions are A and B, the trajectory points between index_i and index_j in the vessel trajectory sequence are extracted, which is the sub-trajectory between waypoint region A and waypoint region B. Finally, the waypoint region sequence set of all vessel trajectories is traversed to obtain sub-trajectories between waypoint region A and waypoint region B.

(2) Clustering sub-trajectory and calculating the boundary

In order to generate distinct routes between waypoint regions, it is essential to categorize sub-trajectories into various clusters through trajectory clustering, of which the essence is trajectory similarity measurement. In this study, an improved Hausdorff distance is used to measure the similarity between trajectories, and the DBSCAN clustering algorithm is employed to cluster the trajectories between waypoint regions.

The Hausdorff distance measurement method mainly calculates the similarity between trajectories based on the position features of trajectory points, without considering the heading characteristics of trajectories. However, the sub-trajectories between waypoint regions are usually formed by ships traveling back and forth between two places. Therefore, when clustering the trajectories between nodes, it is necessary to consider the heading characteristics of sub-trajectories to better distinguish all sub-trajectories with different directions. The starting and ending points of sub-trajectories between waypoint regions are mapped one-to-one with the waypoint regions, so the direction of the sub-trajectory can be described by the vector between the waypoint regions., as shown in Eq. (3).

$$C = \begin{cases} 1\left(\overrightarrow{AB}\right) \\ -1\left(\overrightarrow{BA}\right) \end{cases} \tag{3}$$

In this context, the notation $\overrightarrow{AB}$ indicates that the trajectory is from waypoint region A to waypoint region B, while $\overrightarrow{BA}$ indicates that the trajectory is from waypoint region B to waypoint region A.

When two sub-trajectories are in the same direction, their direction difference is 2 or -2; when two trajectories are in opposite directions, their direction difference is 0. Generally, trajectories in opposite directions should be classified into different sub-trajectories, and therefore, sub-trajectories in opposite directions should not have similarities. The similarity of sub-trajectories in the same direction is measured using the Hausdorff distance between trajectories, as shown in Fig. 7. The similarity measurement method of sub-trajectories can be defined as:

$$\eta = \begin{cases} +\infty \Delta C = 0 \\ H(A,B) \quad \Delta C = 2 \end{cases} \tag{4}$$

Where, ΔC is the directional difference of sub-trajectories, and $H(A,B)$ is the Hausdorff distance between sub-trajectories.

The proposed enhanced Hausdorff distance measurement algorithm can effectively capture the similarity of trajectories in different directions. The DBSCAN clustering method is utilized to cluster the trajectories between the waypoint regions by constructing the similarity matrix based on this algorithm.

Traverse all sub-trajectories between waypoint regions A and B, and calculate the direction of each trajectory. Then, the improved Hausdorff algorithm is used to calculate the distance between each trajectory and other trajectories. Finally, by inputting the corresponding parameters *eps* and *MinPts*, the DBSCAN clustering algorithm is used to cluster the sub-trajectories and obtain the trajectory clusters.

By clustering the sub-trajectories between waypoint regions, different directions of trajectories between nodes and trajectories with the same direction but different routes can be distinguished. By extracting the boundaries of each trajectory cluster using the Alpha-Shape contour line extraction method, the polygonal geometric region of the trajectory cluster can be obtained. This geometric region is called a route and is used to connect the waypoint regions in a geometric maritime network.

4. Case study and results

4.1. Study area and dataset

The study area is located in Vancouver waters, which is a coastal region in western Canada, as shown in Fig. 8. The area is bounded by longitudes 122.2°E to 126.0°E and latitudes 48°N to 49.4°N. It is a large and complex marine environment covering several important waterways, ports, and harbors, as well as a diverse range of marine habitats and ecosystems. The region is known for its high marine traffic volume, as it is a major gateway for international shipping and commercial vessels traveling to and from the Pacific Ocean.

To reduce the complexity of the experiment, 50 ships were randomly selected from a one-week Automatic Identification System (AIS) dataset in this study. The original dataset contains seven types of information on ship dynamics and statics, with a total of 152,894 records.

The original AIS data may exhibit various errors, including position anomalies, velocity anomalies, and trajectory position anomalies. Position anomalies refer to instances where ship data is labeled on wrong locations, such as land. Velocity anomalies indicate cases where ship speed exceeds its maximum navigational speed value. Trajectory position anomalies occur when ship trajectory points deviate, resulting in noticeable differences compared to adjacent trajectory points. To ensure experimental accuracy, this study employed a deletion strategy to handle these abnormal data instances. Detailed explanations of the specific data preprocessing methods can be found in Reference(Zhao et al., 2018). After data preprocessing, 151,348 ship data records were retained.

4.2. Results analysis

4.2.1. Results of node extraction

In this study, a speed threshold of $\gamma = 0.2$ nmile/h and a quantity threshold of $n = 20$ were set to identify dwell feature points These parameter values are widely used in maritime research and are considered reasonable because they reflect the slow speed of a vessel

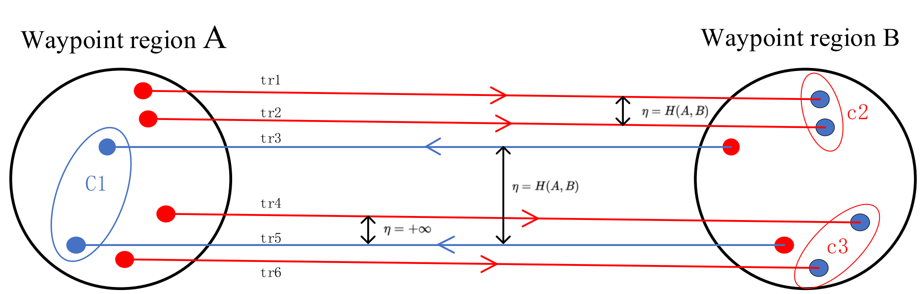


Fig. 7. Sub-trajectory clustering based on an improved trajectory similarity.

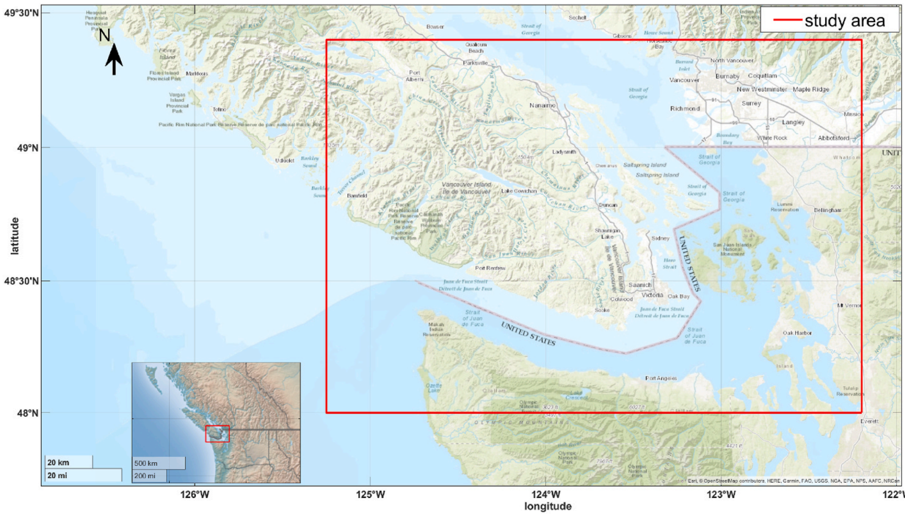


Fig. 8. Overview map of the study area.

when it is in a docking or anchoring state(Yan et al., 2022). The turning threshold for identifying turning feature points was set to 10°, while the turning rate threshold was set to 5°/min. These values are commonly used in trajectory analysis and are effective in identifying changes in vessel direction(Zhang et al., 2021). The boundary feature points were identified based on the predefined study area. Three kinds of trajectory feature point identification results are shown in Fig. 9(a).

To construct the nodes of the two-layer maritime route network, waypoint regions need to be identified using trajectory feature points. The waypoint region extraction stage in this study employs a clustering algorithm, of which initial step is to set the parameters and they have a significant impact on the final extraction results. Appendix A explains the method for determining clustering parameters during the process of clustering trajectory feature points to form waypoint regions. The determination of three kinds of trajectory feature point clustering parameters and the corresponding overall average silhouette coefficients are shown in Table 1.

By clustering three kinds of trajectory feature points, a total of 25 waypoint clusters are identified, and 25 waypoint regions are obtained after calculating the convex polygons of all waypoint clusters using the

Table 1
The setting of trajectory feature point clustering parameters.

	MinPts	*eps*/nmile	Number of clusters	Silhouette coefficients
Turning points	15	1.5	20	0.8605
Dwell points	10	2.4	3	0.9558
Boundary points	10	0.9	2	0.9480

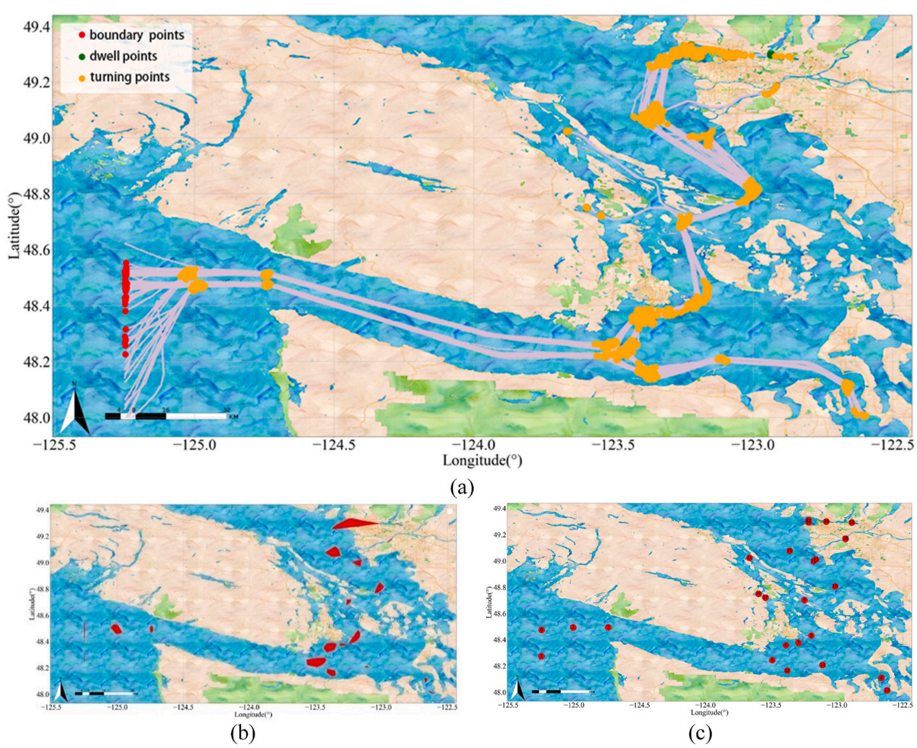


Fig. 9. Results of node extraction.

convex packet algorithm, as depicted in Fig. 9(b). The obtained waypoint regions, which indicate the nodes of the geometric maritime network, are shown as the red area in the figure. To determine the nodes of the topological maritime network, we calculate the geometric centroids of the respective waypoint regions, as illustrated in Fig. 9(c).

4.2.2. Results of node connection matrix and route generation

The construction of the connection matrix involves transforming the ship track sequence into a sequence of waypoint regions. Fig. 10 illustrates the ship's trajectory points marked in green as it passes through the corresponding waypoint regions. The ship's trajectory sequence is transformed into a sequence of waypoint regions, namely [1,4,5,6,7,8,9,10,21,10,9,8,7,6,5,4,1].

The conversion of all ship trajectory sequences to waypoint region sequences enabled the creation of a two-layer route network node connection matrix D, which is presented in Table 2. In the table, the horizontal and vertical labels, denoted as 0–6 correspond to the maritime route network node numbers 0–6, respectively. The numerical values in the table represent the quantity of ship trajectories existing between nodes. For instance, in Table 2, the value at position (2,3) is 44, indicating that there are a total of 44 ship trajectories traveling between network node 2 and network node 3.

The method of extracting sub-trajectories between waypoint regions can effectively extract sub-trajectories between two waypoint regions. By using the improved Hausdorff distance to calculate the similarity between trajectories, different directions of sub-trajectories can be effectively distinguished. The method of DBSCAN to cluster sub-trajectories based on the similarity between trajectories can eliminate noise trajectory data. Fig. 11(a) shows the clustering results of these sub-trajectories between waypoint region 9 and waypoint region 10. The yellow trajectory represents the ship trajectory from waypoint region 9 to waypoint region 10, and the green trajectory represents the ship trajectory from waypoint region 10 to waypoint region 9.

By using the contour line extraction algorithm to extract the boundaries of various directional sub-trajectories, the route between waypoint regions can be determined. As illustrated in Fig. 11(b), the orange area represents the route from waypoint region 9 to waypoint region 10, while the blue area represents the route from waypoint region 10 to waypoint region 9.

4.2.3. Results of two-layer maritime route network construction

The two-layer route network's node connection matrix are used to determine the connectivity between topological nodes. If a non-zero value are found in D[i,j] of the matrix, it indicate that there is a connection between topological node i and topological node j, and they can be connected by topological edges. The resulting topological route network is depicted in Fig. 12, where red points represent topological nodes and black lines reflect topological edges connecting each topological node.

Table 2
Matrix of node connection (partial).

Node number	0	1	2	3	4	5	6
0	0	0	32	0	0	0	0
1	0	0	0	0	0	0	0
2	25	1	0	44	0	0	0
3	2	0	34	0	0	40	0
4	0	0	0	0	0	0	0
5	0	0	0	37	0	0	30
6	0	0	0	0	0	27	0

The geometric network is a spatial representation of the maritime network, which includes the shape, location, boundary, and other geometric elements of the waypoint regions and routes. To construct the route of a geometric maritime network, the connection matrix is filtered by a connection threshold. The connection threshold defines the minimum number of trajectories required to establish a connection between two waypoint regions in the geometric network. If the number of trajectories passing between two regions is below the connection threshold, it indicates that ships rarely navigate this route, and thus, the connection between the two regions is considered invalid. On the other hand, if the number of trajectories between two regions exceeds the connection threshold, it indicates that the route is navigated frequently, and the connection between the two regions is valid.

In this study, the connection threshold is set to 5, as it is determined that when the number of tracks passing through a waypoint region is less than 5, it becomes difficult to extract valuable route information. Fig. 13 (a) shows the trajectories between two waypoint regions that satisfy the connection threshold and are considered valid connections, while the trajectories between the two waypoint regions in Fig. 13(b) do not meet the threshold and are considered an invalid connection. If two waypoint regions meet the connection threshold, they can be connected through routes, while they are not connected if they fail to meet the connection threshold. The resulting geometric route network is depicted in Fig. 13 (c), where the red areas represent the waypoint regions, and the blue and orange areas represent two routes in different directions.

5. Discussion

This paper investigates a method for constructing a two-layer maritime route network model and then constructs a network using AIS data in a real study water area to verify the feasibility of the model. To substantiate the superiority of the proposed two-layer route network

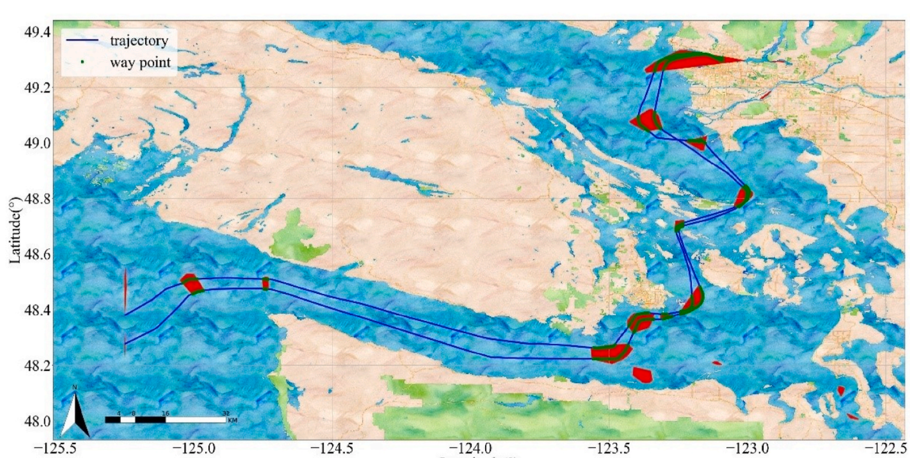


Fig. 10. The process of converting a vessel trajectory sequence to a waypoint region sequence.

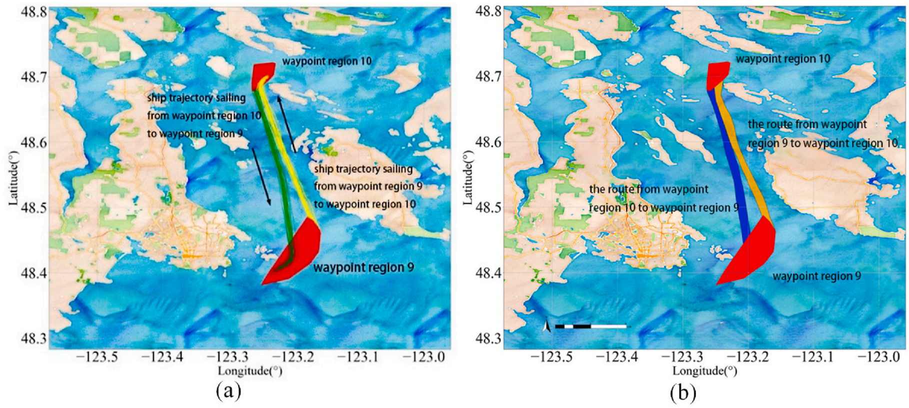


Fig. 11. Extraction results of sub-trajectory clusters and routes between waypoint regions.

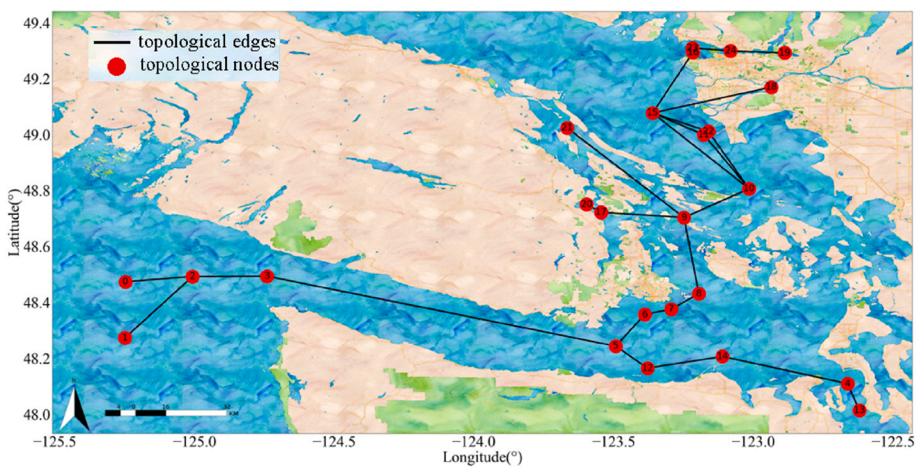


Fig. 12. The results of the topological route network construction.

model, a comparative analysis is conducted against previously conducted research, assessing node extraction techniques and the number of network layers employed in the model's construction. Additionally, experimental validation is performed to verify the practical viability of the proposed two-layer maritime route network within real-world applications.

Table 3 intuitively illustrates the advantages of the proposed two-layer maritime route network model in comparison with other studies. In terms of node extraction methods, prior research(Lei et al., 2016; Wu et al., 2017) utilized grid partitioning techniques to extract key nodes, designating grid regions with larger data volumes as pivotal areas and smaller regions as non-key points. However, such node extraction approaches overlook the motion attributes of ship trajectories embedded within the network nodes. Other studies (Arguedas et al., 2018; Filipiak et al., 2020; Rong et al., 2020) focused on specific categories of network nodes, such as turning feature points or dwell feature points. These node construction methods can only encapsulate certain attributes of ship trajectories, such as directional characteristics or speed attributes. Varlamis et al. (2021) introduced a method that, akin to this study, considers several features of ship trajectory, including speed, direction, and trajectory point clustering, to ensure that network nodes embody a wider range of ship trajectory motion characteristics.

Meanwhile, in terms of network construction, the maritime route network models proposed by Yan et al. (2020); Wang et al. (2019) and Lee et al. (2022) are single-layer models that solely incorporate either the topological or geometric layer of the maritime route network. In comparison, the maritime route network constructed in this study encompasses both geometric and topological layers. This multi-layer approach imparts a more comprehensive set of network information compared to single-layer counterparts.

To verify the feasibility of the proposed two-layer maritime route network model in assisting intelligent vessel navigation, we applied the topological layer network from the model for path planning during the intelligent vessel navigation process. Additionally, we utilized the geometric layer network to detect deviations in vessel trajectories during navigation.

Path planning is a crucial task in intelligent vessel navigation as it directly impacts the efficiency and safety of maritime navigation. The topological layer network within the proposed two-layer maritime route network model authentically reflects the structure and characteristics of the maritime route network, thereby providing more precise guidance for intelligent vessel path planning. For instance, in Fig. 14 considering a scenario with the starting node p_0 and destination node p_{18}, experimental results yield a planned path node sequence of [0, 2, 3, 5, 6, 7, 8, 9, 10, 15, 18]. This signifies that vessels will navigate along the path delineated by the yellow solid line in Fig. 14.

In the experiment, distances between adjacent nodes along the planned route were calculated, resulting in Table 4. This planned route encompasses 11 nodes and 10 segments, providing accurate distance references for intelligent vessel path planning. For instance, between node 10 and node 15, there exist two selectable routes indicated by the node sequence as [10, 15] and [10, 14, 22, 15]. However, upon computation, the total length of the [10, 14, 22, 15] route is 30.45 km, whereas the planned route [10, 15] is notably more economically efficient with a length of only 21.21 km.

In the process of intelligent ship navigation, timely monitoring of

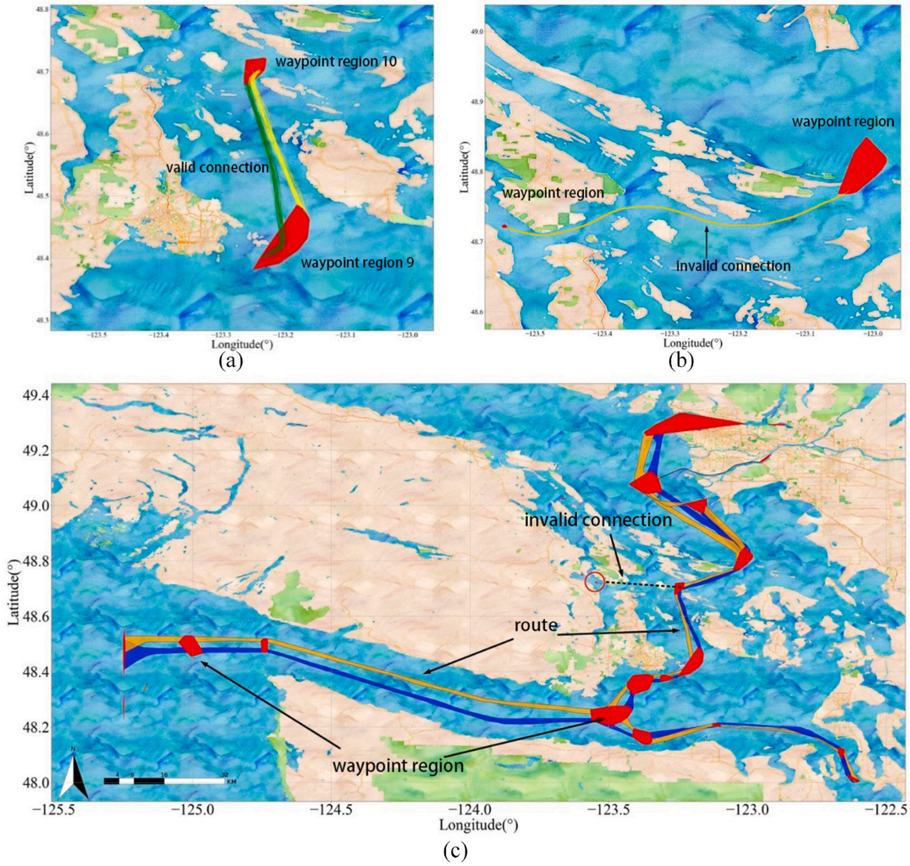


Fig. 13. Representation of results of the geometric route network construction.

Table 3
Comparison between the proposed model and relevant studies.

Research paper	Characteristics of network nodes				Layers in maritime route network			Application	
	Speed	Direction	Trajectory point clustering	Semantic knowledge	Topological layer network	Geometric layer network	Semantic layer network	Route planning	Abnormal route conditions
Filipiak et al. (2020)		✓			✓			✓	
Arguedas et al. (2018)	✓		✓			✓	✓		
Rong et al. (2020)		✓	✓		✓				✓
Varlamis et al. (2021)	✓	✓	✓		✓				✓
Wang et al. (2019)					✓				
Lei et al. (2016)			✓		✓				
Wu et al. (2017)			✓						
Yan et al. (2020)	✓	✓		✓		✓			
Lee et al. (2022)			✓		✓				
This study	✓	✓	✓		✓	✓		✓	✓

vessel trajectory deviations is also a crucial task. Due to the unclear boundaries of maritime routes or navigable areas, visual methods are ineffective in detecting vessel trajectory deviations. However, in the constructed two-layer maritime network model of this study, the geometric layer network is derived from a substantial historical dataset of vessel trajectories, enabling the extraction of navigable water regions. Leveraging the geometric layer maritime network allows for a swift and convenient detection of vessel trajectory deviations, facilitating real-time deviation alerts during ship navigation. This capability ensures vessels adhere to their designated routes, thereby enhancing navigation safety.

As shown in Fig. 15(a), a maritime route is established between navigational waypoint regions 9 and 10 based on historical vessel AIS trajectories. Vessels navigating normally between waypoint regions 9 and 10 should have their trajectories within the designated maritime route range between these two regions. In the case where portions of a

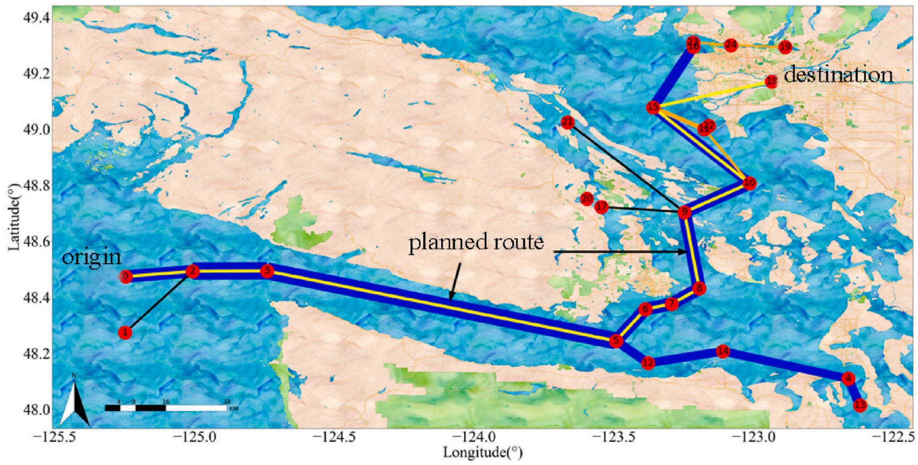


Fig. 14. Path planning using the topological layer network.

Table 4
Length of segments in planned route (km).

Segment numbering	Distance	Segment numbering	Distance
0–2	9.63	7–8	5.14
2–3	1.57	8–9	16.52
3–5	51.98	9–10	11.09
5–6	7.97	10–15	21.21
6–7	3.92	15–18	17.58

vessel's trajectory in Fig. 15(a) extend beyond the boundaries of the designated route area, it is deduced that the vessel's navigation trajectory has deviated from the intended route range.

Fig. 15(b) illustrates the results of vessel track deviation detection. In the figure, the yellow tracks represent the vessel's trajectories during normal navigation within the planned route, while the red tracks depict the trajectories indicating vessel deviation from the intended route. The results demonstrate that the utilization of the geometric layer of the two-layer maritime route network model enables effective vessel track deviation detection. This allows vessels to maintain navigation within the designated route during their voyage, thereby avoiding the potential increase in time and fuel consumption resulting from vessel route deviation during navigation.

This study does, however, have some restrictions. In the feature point recognition stage, only three types of ship trajectory feature points were considered: dwell points, boundary points, and turning points. While other types of feature points, such as dividing dwell points into anchorage and berthing ones, were not considered. In addition, when constructing the network model, differences in ship types were not taken into account, and dedicated route network models were not established for different types of vessels.

6. Conclusion

In this paper, a novel approach is proposed for the construction of a topological-geometric two-layer marine route network to support ship intelligent navigation. The method uses ship behavior traits to extract trajectory feature points from ship trajectory data. The density clustering algorithm combined with the convex packet algorithm is used to obtain the two-layer route network nodes, which are made up of waypoint regions and topological nodes. By using the spatial intersection approach, the ship's trajectory is transformed into a series of waypoint regions, and a two-layer route network nodes connection matrix is built to provide the framework for node connection. To create route-connected waypoint areas, the method employs sub-trajectory extraction and clustering algorithms. The topological-geometric two-layer maritime route network is completed by connecting all topology nodes and waypoint regions according to the relation displayed in the connection matrix. The proposed approach provides ship operators with route planning and yaw detection functionalities, thus ensuring safe and efficient navigation. The study results demonstrate the effectiveness and feasibility of the proposed approach.

In future work, it is necessary to further expand the types of ship characteristic points and incorporate more types of characteristic points into construction of ship route network models. Also, it is important to consider categorizing ships according to their features and creating specialized route network models for different ship types, which can

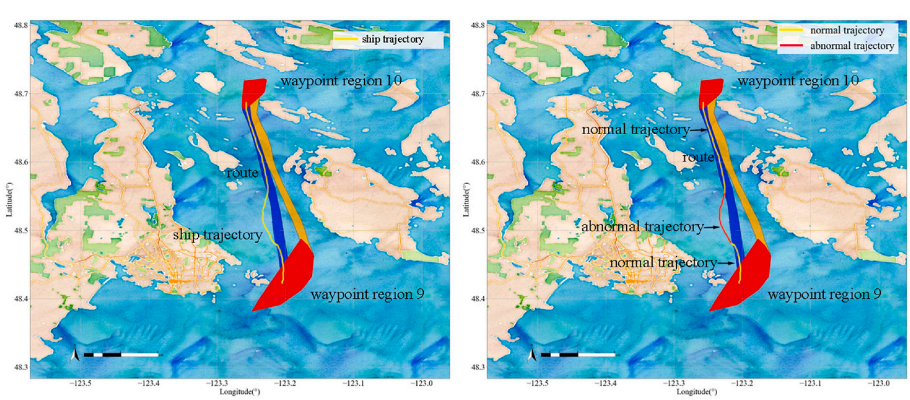


Fig. 15. Detection of vessel track deviation from planned route.

better meet the needs of different ships and improve the safety and efficiency of ship navigation.

CRediT authorship contribution statement

Chunhui Zhou: Conceptualization, Methodology. **Jiale Xiang:** Data curation, Writing – original draft. **Hongxun Huang:** Data curation, Methodology, Software, Visualization, Writing – review & editing. **Yi Yan:** Methodology, Visualization. **Liang Huang:** Methodology, Writing – review & editing. **Yuanqiao Wen:** Funding acquisition, Investigation. **Changshi Xiao:** Methodology, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We sincerely thank the editor and reviewers for their kind and helpful comments on this manuscript. This work was supported by the Hainan Provincial Joint Project of Sanya Yazhou Bay Science and Technology City (Grant No.2021JJLH0012), the National Science Foundation of China (Grant No.52171349), the Laboratory of Transport Pollution Control and Monitoring Technology (Grant No.2022JH-F038) and the Key Research Plan of Zhejiang Provincial Department of Science and Technology, China (Grant No.2021C01010).

Appendix A

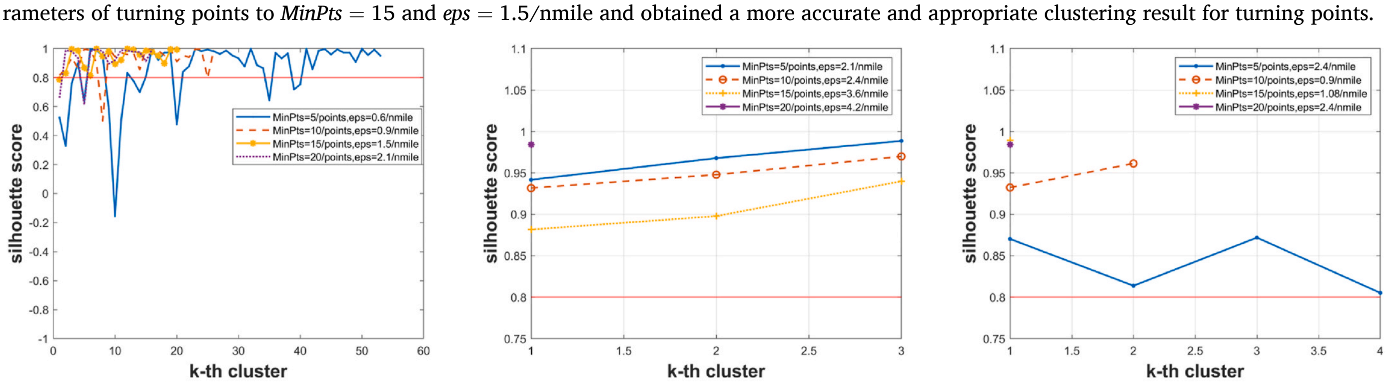
The method of using the k-nearest algorithm to determine the parameters *eps* and *MinPts* of the DBSCAN algorithm has been extensively investigated. It is generally recommended that the value of *MinPts* should be greater than or equal to the number of dimensions of the dataset plus one. To determine the value of *eps*, we defined the k-nearest function to calculate the distance between each feature point and its k-th nearest neighbor (i.e., k-dist), sorted the k-dist of the dataset, and plotted the distance graph. The distance corresponding to the inflection point of the curve in the distance graph is considered as the optimal value of *eps*. By adopting this method, the subjectivity and complexity associated with manually setting parameters can be avoided, thereby improving the accuracy and efficiency of route point area recognition. To objectively evaluate the clustering results under different parameter settings, the silhouette coefficient of clusters was calculated.

In the clustering process, this study first determined the parameter *MinPts* and used the method mentioned in the previous paragraph to determine the value of *eps*. To objectively evaluate the clustering effect of turning points under different parameters, we calculated the silhouette coefficient of all points under different clustering parameters and calculated the average values, as shown in Tab.A1:

TableA1
Clustering results of feature points under different parameter settings

	Number	*MinPts*/points	*eps*/nmile	Number of clusters	Silhouette coefficient
Turning point	1	5	0.6	53	0.6624
	2	10	0.9	26	0.7784
	3	15	1.5	20	0.8605
	4	20	2.1	16	0.7856
Dwell point	1	5	2.1	3	0.9603
	2	10	2.4	3	0.9558
	3	15	3.6	3	0.9127
	4	20	4.2	1	0.9844
Boundary point	1	5	2.4	4	0.8325
	2	10	0.9	2	0.9480
	3	15	1.08	1	0.9896
	4	20	2.4	1	0.9814

Calculating the average silhouette distance of different clusters under different clustering parameters, and visualization of the results, as shown in Fig.A1. When the number of clusters is 20 (*MinPts* = 15, *eps* = 1.5/nmile) of turning points clustering, the silhouette coefficient of clustering result was the highest, reaching 0.8605, and the average silhouette distance of each cluster was greater than 0.8. Therefore, this study set the clustering parameters of turning points to *MinPts* = 15 and *eps* = 1.5/nmile and obtained a more accurate and appropriate clustering result for turning points.



(a)The average silhouette coefficient for each cluster of turning feature points under different clustering parameters

(b)The average silhouette coefficient for each cluster of dwell feature points under different clustering parameters

(c)The average silhouette coefficient for each cluster of boundary feature points under different clustering parameters

Fig. A1. The average silhouette distance of each cluster under different clustering parameters.

For dwell points clustering, although the overall silhouette coefficient of clustering is higher in the fourth case, due to the selection of an overly large *eps* value, the noise data could not be accurately identified, and only one cluster was generated. Therefore, this study set the clustering parameters for dwell point detection as *MinPts* $= 10$ and *eps* $= 2.4$/nmile.

In the third and fourth cases of boundary points clustering, too large *MinPts* values resulted in excessive normal data being identified as noise points, leading to only one cluster being generated. Therefore, the boundary point clustering parameters were set as *MinPts* of 10 and *eps* of 0.9/nmile.

References

Arguedas, V.F., Pallotta, G., Vespe, M., 2018. Maritime traffic networks: from historical positioning data to unsupervised maritime traffic monitoring. IEEE Trans. Intell. Transport. Syst. 19 (3), 722–732.

de Vos, J., Hekkenberg, R.G., Banda, O.A.V., 2021. The Impact of Autonomous Ships on Safety at Sea - A Statistical Analysis. Reliability Engineering & System Safety 210.

Filipiak, D., Wecel, K., Strozyna, M., Michalak, M., Abramowicz, W., 2020. Extracting maritime traffic networks from AIS data using evolutionary algorithm. Bus.; Inform. Sys. Eng. 62 (5), 435–450.

Gonzalez, J., Battistello, G., Schmiegelt, P., Biermann, J., 2014. Semi-automatic extraction of ship lanes and movement corridors from AIS data. In: 2014 IEEE Geoscience and Remote Sensing Symposium. IEEE, pp. 1847–1850.

Hashemi, M., 2019. Automatic inference of road and pedestrian networks from spatial-temporal trajectories. IEEE Trans. Intell. Transport. Syst. 20 (12), 4604–4620.

Karatas, G.B., Karagoz, P., Ayran, O., 2021. Trajectory pattern extraction and anomaly detection for maritime vessels. Internet Things 16.

Krumm, J., 2011. Trajectory Analysis for Driving, Computing with Spatial Trajectories. Springer, pp. 213–241.

Lee, J.-G., Han, J., Whang, K.-Y., 2007. Trajectory clustering: a partition-and-group framework. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, pp. 593–604.

Lee, J.S., Lee, H.T., Cho, I.S., 2022. Maritime traffic route detection framework based on statistical density analysis from AIS data using a clustering algorithm. IEEE Access 10, 23355–23366.

Lei, P.R., Tsai, T.H., Peng, W.C., 2016. Ieee. In: Discovering Maritime Traffic Route from AIS Network, 18th Asia-Pacific Network Operations and Management Symposium (APNOMS). Kanazawa Chamber Commerce & Ind, Kanazawa, JAPAN.

Li, H., Liu, J., Wu, K., Yang, Z., Liu, R.W., Xiong, N., 2018. Spatio-temporal vessel trajectory clustering based on data mapping and density. IEEE Access 6, 58939–58954.

Lin, Y., Xu, D., Wang, N., Shi, Z., Chen, Q., 2020. Road extraction from very-high-resolution remote sensing images via a nested SE-Deeplab model. Rem. Sens. 12 (18), 2985.

Liu, Z., Zhang, Y., Yu, X., Yuan, C., 2016. Unmanned surface vehicles: an overview of developments and challenges. Annu. Rev. Control 41, 71–93.

Mazzarella, F., Alessandrini, A., Greidanus, H., Alvarez, M., Argentieri, P., Nappo, D., Ziemba, L., 2013. Data Fusion for Wide-Area Maritime Surveillance, Workshop on Moving Objects at Sea.

Mazzarella, F., Vespe, M., Damalas, D., Osio, G., 2014. Discovering vessel activities at sea using AIS data: mapping of fishing footprints. In: 17th International Conference on Information Fusion (FUSION). IEEE, pp. 1–7.

Pallotta, G., Vespe, M., Bryan, K., 2013a. Vessel pattern knowledge discovery from AIS data: a framework for anomaly detection and route prediction. Entropy 15 (6), 2218–2245.

Pallotta, G., Vespe, M., Bryan, K., 2013b. Ieee. In: Traffic Knowledge Discovery from AIS Data, 16th International Conference on Information Fusion (FUSION). Istanbul, TURKEY, pp. 1996–2003.

Rong, H., Teixeira, A.P., Soares, C.G., 2020. Data mining approach to shipping route characterization and anomaly detection based on MS data. Ocean. Eng. 198.

Varlamis, I., Kontopoulos, I., Tserpes, K., Etemad, M., Soares, A., Matwin, S., 2021. Building navigation networks from multi-vessel trajectory data. GeoInformatica 25, 69–97.

Varlamis, I., Tserpes, K., Etemad, M., Júnior, A.S., Matwin, S., 2019. A Network Abstraction of Multi-Vessel Trajectory Data for Detecting Anomalies. EDBT/ICDT Workshops.

Vettor, R., Soares, C.G., 2015. Detection and analysis of the main routes of voluntary observing ships in the North Atlantic. J. Navig. 68 (2), 397–410.

Wang, G.L., Meng, J.L., Han, Y.B., 2019. Extraction of maritime road networks from large-scale AIS data. IEEE Access 7, 123035–123048.

Wu, L., Xu, Y.J., Wang, Q., Wang, F., Xu, Z.W., 2017. Mapping global shipping density from AIS data. J. Navig. 70 (1), 67–81.

Yan, Z., Cheng, L., He, R., Yang, H., 2022. Extracting ship stopping information from AIS data. Ocean. Eng. 250.

Yan, Z.J., Xiao, Y.J., Cheng, L., He, R., Ruan, X.G., Zhou, X., Li, M.C., Bin, R., 2020. Exploring AIS data for intelligent maritime routes extraction. Appl. Ocean Res. 101.

Yang, J.-M., Tseng, C.-M., Tseng, P.S., 2015. Path planning on satellite images for unmanned surface vehicles. Int. J. Nav. Archit. Ocean Eng. 7 (1), 87–99.

Yang, W., Ai, T.H., Lu, W., 2018. A method for extracting road boundary information from crowdsourcing vehicle GPS trajectories. Sensors 18 (4).

Zhang, D.H., Zhang, Y.J., Zhang, C., 2021. Data mining approach for automatic ship-route design for coastal seas using AIS trajectory clustering analysis. Ocean. Eng. 236.

Zhao, L.B., Shi, G.Y., Yang, J.X., 2018. Ship trajectories pre-processing based on AIS data. J. Navig. 71 (5), 1210–1230.

Zheng, Y., 2015. Trajectory data mining: an overview. Acm Trans. Intell. Sys. Technol. 6 (3).

Zhou, C.H., Gu, S.D., Wen, Y.Q., Du, Z., Xiao, C.S., Huang, L., Zhu, M., 2020. The review unmanned surface vehicle path planning: based on multi-modality constraint. Ocean. Eng. 200.