Generation and Application of Maritime Route Networks
Overview and Future Research Directions

Huang, Liang; Wan, Chengpeng; Wen, Yuanqiao; Song, Rongxin; van Gelder, Pieter

DOI
10.1109/TITS.2024.3492060

Publication date
2024

Document Version
Final published version

Published in
IEEE Transactions on Intelligent Transportation Systems

Generation and Application of Maritime Route Networks: Overview and Future Research Directions

Liang Huang, Chengpeng Wan, Yuanqiao Wen, Rongxin Song, and Pieter van Gelder

***Abstract*— The development of advanced ship positioning and intelligent sensing technologies has transformed navigation at sea, moving beyond reliance on captains' experience and standard routes. The trajectories traversed by ships at sea contain valuable data that can be mined to map maritime transportation networks and inform intelligent navigation systems. Ship trajectory data at scale enables discovery of the underlying network of maritime routes, providing key insights for applications like intelligent navigation, abnormal behavior detection, trajectory prediction, and maritime traffic pattern analysis. This study reviews the development of research on maritime route networks (MRNs) derived from ship trajectory data. It summarizes the technical process to construct a MRN, contrasting approaches for identifying waypoints, extracting routes, and representing the overall maritime traffic network structure. Finally, this study explores potential applications of MRNs and anticipates promising future research directions in this domain.**

***Index Terms*— Intelligent navigation, maritime route network, maritime resilience, route extraction, ship trajectory, waypoint identification.**

I. Introduction

SHIP routing, which provides a suitable path between ports, is one of the key procedures in the marine industry [1]. Prior to navigation, the crews would establish a rough course using historical data from similar routes as well as the most recent navigation-related information given by various information systems. The path taken by ships is frequently a routine one that develops over time as a result of factors like sailing habits, travel expenses, navigational safety, and the ocean environment. Typically, these empirical pathways lack clear geographic boundaries and route details.

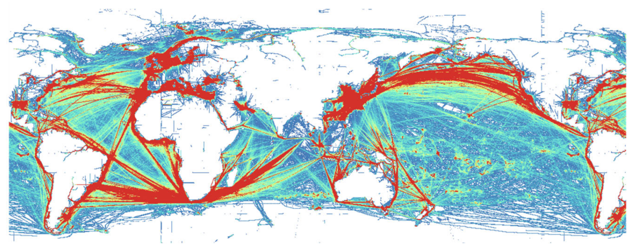


Fig. 1. Density map of maritime routes extracted from ship trajectories (Source: depicted by authors).

With the rapid development of sensors, spatial positioning, and artificial intelligence technologies, the capability of the ship to sense information about its own status, nearby targets, and the navigation environment is continuing to advance. Massive amounts of information about ship movement between ports can be recorded by various information systems, which can be utilized to help design and optimize navigation routes. Of these, The Automatic Identification System (AIS) has amassed a sizable amount of trajectory data that documents the ship's motion features in maritime navigation. These trajectory data include not only a space-time sampling of different navigation routes [2] but also a condensed summary of the sailing expertise of numerous ship drivers in different waters, as well as a substantial amount of customary route information [3]. As shown in Figure 1, the spatial aggregation of large-scale ship trajectories is rendered as a density map with multiple hues, ranging from blue (low density) to red (high density). The spatial distribution of maritime routes may be generally shown which serves a distinctive role in the network structure depiction of maritime transportation routes. It can also be seen that there are ever more alternatives for ship navigation routes between ports.

A data-driven approach of extracting routes and constructing a local or global route network from massive ship trajectory data has becomes popular. The general behavior and movement pattern of ships can be extracted from trajectories by spatiotemporal modeling of navigation routes, and functioned as a basis of the MRN construction. In terms of network flow, the MRN can assist in exploring marine vessel

Received 30 August 2023; revised 17 April 2024; accepted 1 October 2024. This work was supported in part by Hainan Provincial Joint Project of Sanya Yazhou Bay Science and Technology City under Grant 2021JJLH0012, in part by the National Natural Science Foundation of China under Grant 51909202, in part by Zhejiang Provincial Science and Technology Program under Grant 2021C01010, and in part by the Young Elite Scientist Sponsorship Program by the China Association for Science and Technology (CAST) under Grant 2021QNRC001. The Associate Editor for this article was E. Kaisar. *(Corresponding author: Chengpeng Wan.)*

Liang Huang is with the Intelligent Transportation Systems Research Centre, Wuhan University of Technology, Wuhan 430063, China, also with the Sanya Science and Education Innovation Park, Wuhan University of Technology, Sanya 572024, China, and also with the Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, The Netherlands (e-mail: L.Huang-3@tudelft.nl).

Chengpeng Wan and Yuanqiao Wen are with the State Key Laboratory of Maritime Technology and Safety, Wuhan University of Technology, Wuhan 430063, China, and also with the Intelligent Transportation Systems Research Centre, Wuhan University of Technology, Wuhan 430063, China (e-mail: cpwan@whut.edu.cn; yqwen@whut.edu.cn).

Rongxin Song and Pieter van Gelder are with the Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, The Netherlands (e-mail: R.Song@tudelft.nl; P.H.A.J.M.vanGelder@tudelft.nl).

Digital Object Identifier 10.1109/TITS.2024.3492060

activity rules and offer useful route details for navigational choices and maritime traffic management. Moreover, ship traffic characteristics of different routes in the MRN can also be statistically analyzed to identify historical navigation patterns, which can be utilized as reference data for route planning, destination prediction, and anomaly detection [4]. The accuracy of the navigational trajectories, however, has a significant impact on the outcomes of this data-driven method. Ship movements do not always exhibit regularity and clustering since navigable waters are less restricted in contrast to urban road transport network. The collected navigation trajectories from AIS data usually have the characteristics of sparse sampling, uneven distribution and high noise. These factors make the task of building a MRN more complicated.

Many studies have been conducted on this subject in the last decade and have developed a variety of advanced technologies. This article reviews and evaluates the current advancements in technologies and methods for maritime traffic network generation based on ship trajectory. The key steps in the processing of ship trajectory data are listed explicitly, which are maritime waypoint identification, maritime route extraction, and maritime route network construction. Popular approaches and models for every significant phase are compiled and compared to highlight the shortcomings and gaps of existing research. Prospective uses of the MRN are also proposed, including the identification of anomalous behavior, the forecast of ship destinations, and the planning of routes for navigation.

The remainder of the article is structured as follows: The review methodology is introduced in Section II. Following the technical process of MRN generation, the most recent developments in techniques for the MRN generation, including waypoint identification, route extraction, and network construction, are covered in Section III, along with the shortcomings and gaps of these studies. Potential applications of the MRN in maritime transportation are introduced in Section IV. The future research directions and key findings are concluded in Section V.

II. Methodology

This research aims to identify and summarize the available literature on the MRN generation based on ship trajectory data. The PRISMA flow diagram is used in this study to make the evaluation process more transparent and understandable. The whole procedure, which includes a literature search and document analysis, is shown in Figure 2.

Six researchers with diverse backgrounds who are all engaged in work and research pertaining to maritime traffic conducted the review. A few of the scholars' published works on this subject demonstrate their extensive research experience in marine traffic analysis. The review workflow consists of a total of four steps.

-STEP 1. Bibliometric search

The first stage in the review is to specify the parameters of the search query. The Web of Science (WoS) database was cited in the query as a primary source since it is the most major, all-encompassing academic information source,

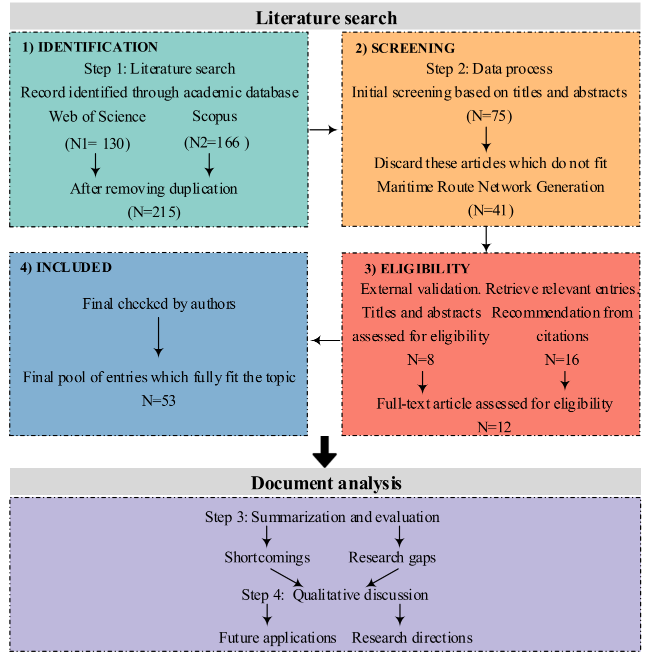


Fig. 2. The workflow for conducting the research.

covering the widest range of fields. The search had also been extended to Scopus in case any significant literary works had not been indexed in the WoS. As inputs, the following keywords were chosen.

(TS=Maritime)

AND

(TS=Traffic)

AND

((TS=Network) OR (TS=Route) OR (TS=waypoint) OR (TS= Node))

AND

((TS= extract*) OR (TS= mining)

Peer-reviewed journal publications and conference papers pertaining to the topic that were published in English between January 2010 and May 2023 were identified using keywords. To separate what is just conceptually relevant, two-step data processing is then used.

-STEP 2. Data process

In the end, 296 articles were found, including 166 from Scopus and 130 from the Web of Science. By comparing authors, titles, and keywords after converting the file to Endnote software, duplicate publications were found and eliminated. Then, to identify those papers that weren't relevant, a preliminary screening technique was used. As an illustration, several scholars provided a thorough analysis of maritime traffic as a complicated network. These publications did not, however, pay attention to extracting traffic routes or creating route networks; instead, they concentrated on mining topological characteristics of maritime logistics. Other studies largely focused on maritime spatial planning and management, ship target detection from images or videos with neural networks. These articles were consequently deleted.

After that, the remaining articles were evaluated for their research goals. A few studies have concentrated on anomaly identification, pattern extraction, and vessel prediction in maritime trajectories; nevertheless, the primary objective of these studies has little to do with the MRN generation. As a result, these items were also eliminated from the sample. 53 articles were ultimately left for additional examination.

-STEP 3. Construct the framework for the MRN generation

Based on the comprehensive review of the related work, the framework for the MRN generation is developed compositing of four steps, which are track data pre-processing, maritime waypoint identification, maritime route extraction, and multi-level network generation. This provides the main structure for the application analysis in the next step.

Based on a thorough analysis of the relevant literature, a framework for the creation of multi-level networks (MRNs) is designed, consisting of four steps: track data pre-processing, maritime waypoint identification, maritime route extraction, and multi-level network generation. This offers the primary framework for the application analysis in the following step.

-STEP 4. Qualitative discussion

A qualitative discussion was conducted to categorize the present research into different groups according the role they played during the MRN generation, to discuss potential applications, and to propose future research directions. The qualitative method used to generate the analysis results involved a thematic analysis of the content of the scientific publications collected. The text was manually coded based on themes, and the publications were categorized into different topics or fields. The approach enabled the identification of common patterns and trends in the research and provided insights into current research topics, future research directions, and applications. For example, the research directions and applications can be summarized from two aspects. On the one hand, research directions and applications from previous studies in related topics are collected and categorized. On the other hand, some new research directions and applications are also explored based on the shift of research trend in recent years identified from literature review.

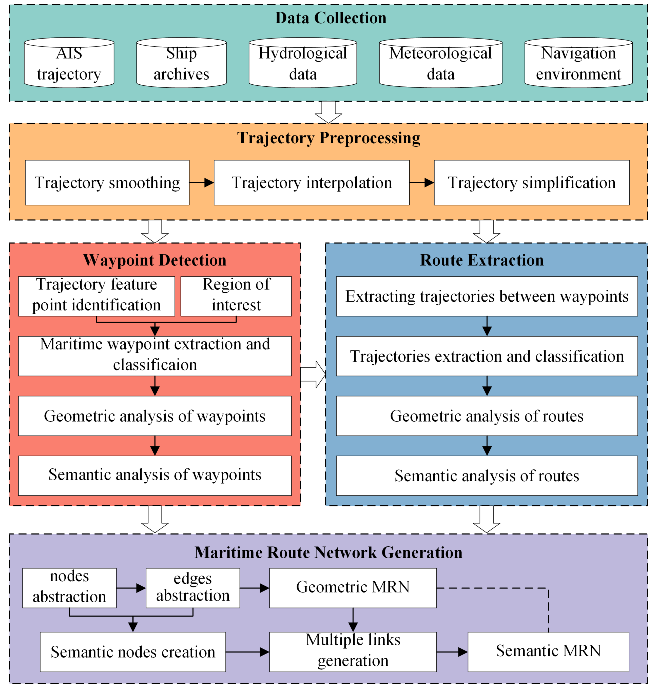


Fig. 3. Technical process of MRN generation.

TABLE I
WAYPOINT DEFINITION AND CLASSIFICATION

Waypoint type	Definition	References
Stationary Point (SP)	It describes the place where a ship berths or anchors in ports, anchorages, or offshore platforms when its speed is almost zero.	[1, 5-26]
Maneuver Point (MP)	It is where a ship's speed, course, or both are significantly changed while at sea, which also called turning point or breaking point.	[1, 3, 7, 11-14, 18-23, 27-32]
Entry/Exit Point (EP)	It is where a ship enters and leaves the defined area or region of interesting.	[5-12, 14-16, 22-24]
Junction Point (JP)	It denotes a location where one route branches out from another or where many routes meet together.	[24, 33]

III. STATE-OF-THE-ART TECHNIQUES FOR MRN GENERATION

This research aims to identify and summarize existing studies on building an MRN from massive ship trajectory data. According to a thorough literature review, a comprehensive framework of MRN generation is proposed in this study, as shown in Figure 3. The framework is generally composed of four steps, including data collection and preprocessing, waypoint identification, route extraction, and route network generation, which make up the main process of MRN generation.

Raw ship trajectory data has a few quality problems such as invalid data, errors, values missing, abnormal values and duplicate records, which can cause incomplete and inaccurate maritime route network. Ship trajectory data preprocessing such as trajectory smoothing, trajectory interpolation, trajectory segmentation, are important prerequisites for MRN generation. The purpose of trajectory smoothing is to filter the noise and invalid track points, which can filter out the small swing drift trajectory point caused by the positioning error of AIS data, and to increase the accuracy of waypoint identification. The trajectory data after smoothing may be discontinuous point sequence, it is necessary to perform a trajectory interpolation operation for obtaining more detailed ship movement. After that, ship trajectory data can be reconstructed to a simpler one that retains the morphology and movement characteristics of raw trajectory with fewer points. These pre-processes are not always performed and need to be selected according to the actual quality of trajectory data.

With simplified trajectory datasets, waypoints are first detected. This study considers that any critical points of particular significance to a ship's voyage can be waypoints. Both key feature points in historic trajectory data and region of interest defined by uses are the candidates for waypoint extraction. For each waypoint, geometric information (boundary and centroid) and semantic information (i.e., traffic characteristics, waypoint types, etc.) should be analyzed and extracted. Based on these waypoints with enriched information, various route extraction methods can be

TABLE II
METHODS OF WAYPOINT IDENTIFICATION

Usage	Methods	Features	References
Candidate extraction	Rule Detection	variations of motion features	[1, 5-13, 15, 19, 22, 29, 32, 34, 35]
	CUSUM (Cumulative Sum)	change rate and magnitude of motion features	[14, 28, 30, 36]
	DP (Douglas Peucker)	spatial shape features	[3, 18, 20, 21, 29, 37]
Waypoint identification	DBSCAN (Density-based Spatial Clustering of Applications with Noise)	Spatial aggregation and distribution	[1, 3, 5-11, 13-16, 18, 20-23, 29, 33, 38]
	OPTICS (Ordering Points to Identify the Clustering Structure)		[19, 21, 24, 35]
	GA (Genetic algorithm)		[33, 36, 39]
	Grid Statistic		[11, 13, 34]

used to detect the common routes and the corresponding traffic patterns as well as navigational characteristics. Similarly, each route is also associated with geometric information (boundaries and centerlines) and semantic information (i.e., average speed and direction as well as their standard deviations, number of vessels, vessel types, etc.).

Based on the results above, hierarchical maritime route networks can be constructed. The first level is geometric maritime route network consisting of waypoint areas (nodes) and sea lanes (edges). Geometric MRN is designed to represent the spatial distribution of significant water areas and bounded navigational channels between water areas with multiple representative routes. This network model can be mainly used for route planning, lane departure warning, destination prediction, etc. Then, waypoint areas, routes, vessels and their semantic information can be further mapped into nodes and links with specific functions that form a semantic MRN. The Semantic MRN is essentially a semantic graph that represents the semantic relationships between different concepts. It calls for the corresponding semantic graph mining methods to get unexpected routes and uncover potential maritime traffic patterns. It can be used for maritime knowledge discovery, route association mining, notable waypoint recognition, etc.

A. Maritime Waypoint Identification

In urban transportation, road intersections are frequently conceived of as waypoints in the road network. However, there are no corresponding natural waypoints in maritime transportation. Building a network of maritime routes therefore requires the setting of waypoints that could represent ports, anchorage areas, capes, offshore platforms, or even any areas where vessels change their course, speed and other dynamic features. Therefore, waypoints in the MRN can be functionally categorized as frequent regions, stationary points, maneuver points, entry/exit points and junction points based on an analysis of prior studies. The definition of each individual waypoint is shown in Table I.

Different kinds of waypoints will be classified as nodes in the MRN and have a significant impact on the MRN's structure and characteristics. There are four types of waypoints, which are Stationary Point (SP), Maneuver Point (MP), Entry/Exit Point (EP), and Junction Point (JP). Stationary points often refer to a ship's departure and arrival ports as well as intermediate stops in a global maritime route network. Entry and exit points can serve the same functions in a regional maritime route network, but they provide different semantic information for the nodes because they are important indicators of routes in the particular waters. Junction points don't exist in physical space, and are logically identified as places where traffic from various directions meets. Junction points show where different routes are correlated and where there may be a high risk of ship collision. Maneuver points can be used to pinpoint the locations on a route where a ship experiences cumulative variation or an abrupt change in speed or course. Particularly, trajectory points with a sharp change in course (also known as turning points or breaking points in some studies) are more useful in indicating route change information. In a word, stationary points and entry/exit points can generate a basic structure of the MRN, and other two types of waypoints benefit to build fine-grained and directional MRN which can be applied for route planning, navigation aid and anomaly detection.

SPs are suitable for maritime route networks (MRN) construction in almost all application scenarios, and they are the basic nodes for building an MRN. The choice of SPs will determine the scale and size of an MRN; MPs are also applicable to MRN construction in all scenarios. However, they are used more often in large-scale networks. This type of point does not affect the structure of a network, but it will influence the resolution of a network, or for specific, influence the richness of the information contained between two SPs. The selection of MPs helps can assist in ship route planning; Eps are mainly used for the construction of MRN in small-scale waters. Since the local MRN is usually incomplete (as a part of global MRN), the local network can be relatively closed by selecting suitable EPs, to carry out topological analysis; The last one, JP, is a special type of point, which belong to a kind of non-essential node during the construction of an MRN. It is unlike the node used in traditional topological analysis. JPs can be understood as the intersection of ship trajectories maybe at different times, which enhances the semantic information of an MRN. JPs can be used for the analysis of the ship traffic flow condition in maritime security supervision. In the future, it has the potential to be used for the identification of potential ship collision and conflict risk areas in intelligent ship sailing scenarios.

The procedure of waypoint identification has two stages: waypoint candidate extraction and waypoint candidate clustering. Waypoint candidate extraction is an essential pre-process step for obtaining significant points from ship trajectories as candidates before waypoint selection. Candidate clustering is then used to extract common waypoints, considering the

TABLE III
METHODS OF ROUTE EXTRACTION

Category	Methods	Features used	References
Cluster-based method	SNN (Shared Nearest Neighbor)	Position and bearing of points	[40]
	TREAD (Traffic Route Extraction for Anomaly Detection)	Static and dynamic information of points	[5-9, 15, 41]
	TRACLUS	Distance between segments	[27]
	SPTCLUST, SPTCLUST-II	Bearing between segments	[25, 26]
	Spectral Cluster	Minimum distances between key turning points	[37]
	MD-DBSCAN	spatial distance and COG (Course Over Ground) difference of turning points	[42]
Grid-based method	Prefixspan	Sequential pattern of frequent regions	[43, 44]
	Density analysis	Vessel position of grids	[22, 45, 46]
	Delaunay triangulation	Traffic density of grids	[47, 48]
Statistics-based method	KDE (Kernel Density Estimate)	Vessel position of grids	[10, 16, 20, 31, 49-51]
	Gaussian Process	Vessel position of grids and time	[24]
	GA	Vessel position of grids	[30, 36, 39]

variability of ship trajectories. Table II lists the most prevalent methods for these two stages.

1) Candidate Extraction Methods: Waypoint candidate extraction focuses on detecting significant changes in the motion characteristics of ship trajectory. Rule detection is the most used method for feature points extraction. Several threshold criteria will be set for ship speed and course in the rule detection technique. Track points whose steering angle or steering speed exceed the thresholds are considered maneuver points [6], [7], and stationary points are activated if the instantaneous or average speed, time interval and distance between adjacent points are less than the predefined threshold [5], [22]. But under the influence of varied environmental factors, ships of various sorts and sizes have distinct staying and turning properties. The selected thresholds for motion features vary in different studies [19], [29]. Absolute speed or course measurement leaves it open to abnormal records, which might produce false candidates.

Recently, several studies attempted to extract waypoint candidates by examining the change rate [11] and amplitude [7] of the course and speed. The CUSUM, a well-known algorithm used for change detection, has been used to detect significant changes in speed or course of ships in AIS data recently. The algorithm has a few implementations, such as one-sided algorithm for observations with the expected direction of the changes and two-sided algorithm that handles increases and decreases of the observed variable [30], [36]. The significant maneuvers detected are the preliminary waypoints, from which the final waypoints will be selected. The Ornstein-Uhlenbeck (OU) process, which is a mean-reverting stochastic process, is sometimes used to model the velocity of non-maneuvering vessels before the CUSUM test. It estimates the piecewise long-run velocity of a ship based on Page's test by assuming that the long-run mean velocity of the process can abruptly change at any instant. Then two CUSUM statistics are applied to detect possible positive and negative changes of velocity [28] and identify five different types of change points, including ports, navigational waypoints, entry, exit, and entry/exit points [14].

The DP-based compression algorithm [3], [20], [21], [29], which captures the ship directional changes based on spatial shape information of trajectory, is another frequently used method for identifying relevant turning points along the ship trajectory. The algorithm identifies the first and last points of a given trajectory and finds the farthest point from the baseline defined by these two points. If the transverse distance of the point to this line exceeds the distance threshold, then the point is retained, and the trajectory is split at that position. The algorithm is recursively applied to both sub-trajectories [18]. There are two drawbacks of the conventional DP method that it needs to pre-define the thresholds artificially and different trajectories share the same threshold. Following that, Li et al. [37] introduced the adaptive DP algorithm, which sets a different threshold for each trajectory based on its geometric features, speed variation rate and the distance of feature points.

2) Waypoint Identification Methods: The captured significant points are coarse and inevitably suffer from the undesirable biases and noise, which are only waypoint candidates. To eliminate these negative effects, it is essential to introduce clustering method to separate the outliers from normal feature points based on spatial aggregation and distribution of waypoint candidates. The refined feature points are then considered as the averages of normal ones and denoted as the common waypoints [19].

DBSCAN algorithm, which is the most used clustering technique, creates waypoint clusters from the area of high density [1], [3], [13], [16], [18], [20], [21], [22], [28], [34], including stationary area of interest or departure-arrival areas, turning sections, and entry/exit locations. DBSCAN forms clusters of elements based on the density of points in their neighborhood that can be measured by two parameters eps and minPts. Eps specifies how close points should be to each other to be considered a part of a cluster, while minPts represents the minimum neighbors n. The distance between two consecutive trajectory points is usually defined by the weighted calculation model of speed, course and location. The algorithm does not require the number of clusters to be given a priori and can manage noise while producing clusters of any shape [23], [33]. In most studies, original DBSCAN is applied to waypoint candidates directly, and there are also some improved versions of DBSCAN for different purposes. Since maritime traffic changes over time, Pallotta et al. proposed an incremental DBSCAN procedure [5], [6], [7], [8], [9], [10] to create, expand and merge stationary points, maneuver points and entry/exit points progressively for a regional area of interest. The clustering parameters are set based on specific traffic density, intensity and regularity in the area of interest. Zhang et al. [38] introduced a fuzzy adaptive DBSCAN technique known as FA-DBSCAN that employed fuzzy theory to limit its sensitivity to parameter settings (neighborhood radius and threshold) and ensured the effectiveness of the final clustering results. Ship turning regions can then be created and spread among the traffic-separation, channel, anchoring, port, and exit areas. To deal with the same problem, Onyango et al. [32] proposed a Hierarchical DBSCAN (HDBSCAN) algorithm to cluster the maneuver points with multiple densities. The HDBSCAN algorithm required the minimum size of clusters and the minimum samples as user input and then simplifies a complex single-linkage hierarchy to a small tree of candidate clusters. Each cluster has a measure of central tendency, the centroid which serve as the waypoint.

Another density-based clustering technique by the name of OPTICS [19], [21], [24] was recently used to identify turning sections and departure-arrival zones for marine traffic by grouping individual stationary locations, entry/exit sites, and maneuver points. The approach generates a data set ordering with each object's reachability distance serving as a representation of its density-based clustering structure. This information is sufficient to extract all density-based clusters with respect to any within-distance. The OPTICS algorithm produces a cluster-ordering consisting of the ordered points and the reachability distance of each ordered point, and makes clusters based on reachability distance which is the maximum of the actual distance and the core distance between two points. Like DBSCAN, OPTICS also requires two parameters: εmin, which describes the maximum distance to consider, and minPts, describing the number of points required to form a cluster. However, OPTICS can solve the sensitivity problem for input parameters.

GA is also used to identify waypoints and generate the mesh and recommended corridors for safe and efficient vessel operations.

GA is a type of optimization algorithm that uses selection and recombination operators to generate new sample points in a search space. The algorithm treats each waypoint as a gene, containing latitude, longitude, and radius values. These genes are grouped into sets forming chromosomes, which represent the waypoints set for the problem space. The fitness criterion used for the evaluation of each chromosome is the number of vessel points contained within all of its genes. The algorithm is run for each frame of AIS data, and the best fit chromosome is selected. The genetic algorithm is shown to be able to process streaming data in real-time, but the execution time increases with the increase of fitness values. The algorithm requires a balanced decision/trade-off between the number of waypoints and precision and the execution time [30], [33], [36], [39].

Besides three most used methods above, Huang et al. [40] proposed a k-nearest neighbor-based classification model of ship stay to distinguish berth from anchorage stopping, while Yan et al. [35] focused on the same problem using the random-forest-based algorithm. Dobrkovic et al. [33] compared the clustering results generated by DBSCAN, GA and modified ant colony optimization algorithm and explained that in a real case application algorithms have to cope with clustering maritime data points where density varies. This makes finding optimal parameters for all clustering algorithms challenging. The authors suggested a hybrid approach that combines the strengths of each algorithm to achieve better results.

B. Maritime Route Extraction

A maritime route describes the path taken by numerous vessels between two waypoints in the monitoring region, such as a stationary point and an entry point as well as two ports, and it occasionally includes a few intermediate waypoints like turning points or junction points. The techniques for maritime route extraction are generally divided into three classes, including vector-based methods, grid-based methods and statistics-based methods [19], as shown in Table III.

1) Cluster-Based Methods: By spatially grouping vast quantities of trajectory data into groups based on similarity measurements, mainly distance metrics, vector-based techniques attempt to find routes between waypoints. A synthetic or representative route in the form of polylines will be created as a compact depiction of trajectory cluster. Since the number of clusters need not be specified a priori and density-based clustering algorithms are resilient to outliers, they appear to be well-suited for the automatic identification of the main traffic routes. The studies about ship trajectory clustering can be broadly classified into three fields: point-based clustering, segment-based clustering and trajectory-based clustering.

The studies on ship trajectory points clustering mainly focus on the similarity measure between points and identify the movements between connected waypoints as the maritime routes. Santos et al. [41] considered the position and the bearing of trajectory point as a motion vector and refined the Shared Nearest Neighbor algorithm to identify clusters of motion vectors that are moving towards the same direction. Pallotta et al. [5], [6], [7], [8], [9] developed the TREAD method to identify spatial and temporal distributions of traffic routes. The algorithm defined a route object to connect the derived waypoints, including two stationary points, a stationary point and an entry/exit point, two entry/exit points, or a sequence of intermediate maneuvering points. The route objects synthesized both the static (e.g., vessel type) and dynamic (state vector observations) features inherited by the vessels that created or updated them. Each synthetic route summarizes the expected "central" behavior along the route by unfolding all route points on the temporal dimension for recursive position computation and prediction [15]. Similarly, Yan et al. [19] extracted navigation routes in the form of stationary point-maneuver point-stationary point by analyzing the navigation characteristics of ships and combining semantic modeling methods. After waypoint identification, these waypoints were connected to form navigation routes in the vast sea area by combining with graph theory.

Regarding ship trajectory segment clustering, overall ship trajectories are generally divided into several line segments for local features or waypoints, and then line segments are spatially clustered. A partition-and-group framework called the TRACLUS algorithm [27] is applied for the semi-automatic extraction of major ship lanes and associated corridors of ship movement. The algorithm first partitions trajectories into segments around the turning points and clusters these segments based on the sum of perpendicular, parallel, and angular distances. This technique has at least two threshold parameters that are highly sensitive to the dataset, making it difficult to reproduce or use in practice. In addition, the computations for the distance measure have quadratic complexity. Sheng and Yin [42] proposed a structure similarity distance measurement which consists of spatial distance, directional distance and speed distance between trajectory segments and revised the DBSCAN method to recognize clusters of different vessel densities based on synthetic similarity distances. The representative trajectory, which is a synthetic line sequence perpendicular to the clustered lines, will be extracted for each cluster to describe the overall movement without prior information. The quality of the clustering result is heavily dependent on extracting suitable features and specifying proper similarity measures. Anneken et al. [53] applied the Hausdorff distance to compare the trajectories with the same start and end areas, and create a synthetic route representation of the overall behavior in a trajectory cluster based on Gaussian Mixture Model (GMM). Later, Eljabu et al. [25] presented a novel spatial clustering approach called SPTCLUST for extracting spatial representations of sailing routes between ports using historical AIS data. SPRCLUST first divided the trajectory data into segments based on port information and grouped these segments based on their direction using a path finding method. All the segments are clustered by considering the pooled standard deviation between them. To extract shipping routes, the authors apply the RRoT (Reference Route of Trajectory) algorithm to each resulting cluster based on the arithmetic mean. However, this algorithm only generates a single reference route for segments with the same direction, which may lead to issues when obstacles such as islands or Marine Protected Areas (MPAs) separate trajectory segments. To address this, the authors incorporate port information as a semantic layer in the SPTCLUST-II method [26] to cluster the trajectory segments. Additionally, they employ Dynamic Time Warping (DTW) to quantify the similarity between trajectory segments, which enables the identification of global shape similarities in trajectories and facilitates the alignment of spatial features along the routes. This method overcomes challenges related to selecting optimal input parameters, processing large volumes of multidimensional data, and achieve accurate extraction of maritime traffic routes for tankers and cargoes, as demonstrated by high accuracy and f1-measure.

Both point-based clustering and segment-based clustering have the problem of computing efficiency since massive trajectory points are increasingly recorded in sea areas. Many studies explored the clustering method of overall ship trajectory after a compression. An unsupervised hierarchical methodology named ISCM (Improved Spectral Clustering with Mapping) was proposed by Li et al. [37] to extract vessel traffic routes for enhancing marine safety and situational awareness. The ISCM method also applied DTW algorithm to measure the similarity between trajectories by finding the optimal alignment between them, and then using the mapping function to map the original data to a new space for stable clustering since traditional spectral clustering algorithm is sensitive to the initial clustering center. However, the DTW distance between trajectories is the sum of the minimum distances between the points belonging to one trajectory and the points belonging to another. Thus, the accuracy may be affected by the large numerical difference between the points belonging to different trajectory. Following that, the Multi-dimensional DBSCAN (MD-DBSCAN) algorithm was proposed by Huang et al. [43] to cluster ship trajectories considering the similarity between the trajectories based on their spatial distance and COG difference of the key turning points. The limitation of this method is the setting work of multiple thresholds in the clustering process, which affect the effectiveness of clustering results.

2) Grid-Based Methods: The grid-based methods discretize the space into a series of equal-sized grids, and the traffic track data is then filled into different grids based on coordinates. The grids become the basic units for the multiple feature fusion of ship trajectories. Compared to vector clustering methods like DBSCAN, grid-based methods can adequately deal with the uneven density distribution of ship trajectories and contribute to the discovery of maritime routes at different granularities.

Grid density analysis is the most used method which can provide a rasterized route made up of a few grids with significant traffic flow. It first counts the ship traffic volume in each gird cell and colors each grid according to the ship traffic volume. The contour lines connected by similar color blocks are similar maritime routes. A trajectory density map with grading color is usually used to identify the main traffic routes in specific waters like the North Atlantic Vettor et al. [46] or global waters Wu et al. [47]. This method is simple to operate, but the size of the grid cell and the distribution of traffic volume directly affect the accuracy and number of routes. To solve the same problem, Liu et al. [22] improved the grid statistics analysis by establishing the cumulative grid importance function for gridded clusters of navigational trajectories in different departure-arrival areas, which were obtained from a semantic route extraction and decomposition algorithm with each grid storing the number of trajectories passing through the area grid. A boundary traffic density threshold was then used to locate the main route which is the primary line of each trajectory cluster, with a feature of relatively high traffic density. After that, a grid-based scarification model was proposed, with the importance function established to eliminate the grids outside the core area of the traffic group and roughly locate the blurred edges of the area. The grid importance was measured using grid values and the distances between the grid and the main route as metric factors, which can effectively deal with the density differences among traffic groups and identify the important area of the group. However, the grid density analysis is rather empirical, as these methods are effective for small-area surveillance applications but have the problem of heavy computation when the scale increases.

Another method, sequence pattern analysis, uses frequent regions instead of grids for route extraction by frequent sequential pattern-based trajectory clustering. Frequent regions are activated if the number of trajectories passed the grid cell has satisfied the user-defined minimum support threshold and raw trajectory data can be transformed into frequent region sequence. Based on frequent region, the Prefixspan algorithm is then used to mine the frequent sequential patterns (FSP) of region-based trajectories and prune them using the longest common subsequences [44]. In this pattern-aware trajectory cluster, the movement behavior in a region is represented by a mobility vector. The mobility vector consists of a major vector computed by the average of the moving vectors in the region and a crossing-section obtained by statistical analysis of the projected pins of points on the line perpendicular to the majority direction. Then, a maritime traffic route can be generated and represented by a sequence of ordered mobility vectors [45]. Since AIS data are large scale, high noisy, the density and quality are rather even in different areas, extracting a whole, continuous and smooth maritime route network based on frequent regions is still a challenging work. In order to integrate both local and global patterns, a time series clustering was later proposed to associate the spatial pattern with temporal pattern for shipping route mining [54]. The algorithm applied a locally weighted polynomial regression with the additive regression model to fit the historical trajectory point clouds with multiple predictors. The points of a possible route can then be derived from local polynomial regression with the K-nearest neighbors algorithm. However, the limitation of this method is that only major shipping routes can be detected. The frequent shipping activities that involve varied shipping routes may be missed. In addition, the route mining method using the nonparametric regression may also need good domain knowledge and empirical information for the complicated local route extraction and verification.

Besides two methods above, a popular method in computational geometry called Delaunay Triangulation was also adopted to extract maritime routes in the form of polygons Wang et al. [48], [49]. A Quad-Tree data structure to store the grid data for efficiency, and the adaptive grid merging and filtering algorithm was then used to parallel extract high- and low-density areas with different extraction precisions. The grids of these areas with low overall average density were then merged and filtered for different windows using a sliding window filtering approach. Since the Delaunay triangulation algorithm can construct a triangulation of a set of points in a plane such that no point is inside the circumcircle of any triangle, it was applied on transforming the merged grids into triangles. All the triangles in lanes are filtered to generate smooth polygons as the lane boundaries and key points of these triangles are extracted and connected into lane centerlines. This method can accommodate different AIS data distribution densities but also requires significant computational resources to process large-scale data. In addition, abnormal data have a great influence on the results.

3) Statistics-Based Methods: The main idea of statistics-based methods is to perform a statistical analysis on AIS trajectory data with different models for obtaining traffic flow characteristics and then extract maritime routes with salient statistical features.

The mostly used statistical method is KDE, which is a non-parametric method for estimating the density functions of AIS data points in different water areas. In the early studies, the KDE method is mainly used to obtain the spatial extent of the extracted maritime routes. For instance, Tzavella et al. [50] first applied the Particle Filter to simulate sea lane traffic in different scenarios such as straight lanes, cross section, and turns, and then used Kernel methods to estimate densities of particles within a neighborhood around each output raster cell by integrating prior knowledge such as coastlines and bathymetry. The density threshold was finally selected to generate continuous vessel tracks indicating each vessel's route. Pallotta et al. [10] used the TREAD method to extract maritime routes and adopted the KDE method with a Gaussian kernel to probabilistically represent a route's spatial extent and certified that KDE has shown a superior ability to accurately model traffic routes, even in the case of a skewed distribution of vessel positions. In recent years, researchers begin to apply the KDE method on AIS data for creating traffic density map in advance. Route boundaries and centerlines can then be extracted from continuous density surface with different methods. Lee et al. [51] revealed attendant traffic density via kernel density estimation analysis of the AIS data and applied image processing techniques (binarization and edge extraction) and line smoothing to create and smooth the boundary of the novel route based on the density data. The centerline of the maritime route was also determined using the Delaunay triangulation algorithm and was compared with the original route in terms of sinuosity, intersection angle, and route change envelope for safety verification. By analyzing the density of various ships, the authors further categorized maritime traffic route into three types: the main route (high traffic density), inner branch route (connecting main route and outer branch route) and outer branch route (connecting main route and outside sea area) [52]. However, this method requires manual intervention and performs poorly in extracting routes at intersections. Similar work was also conducted by Lu et al. [16] in the areas of Shanghai port and Zhoushan port in China. The difference is that the KDE and the Delaunay triangulation were only used for an individual sub-route extracted by DBSCAN. The authors later improved the method and adopted sliding window algorithm to estimate the route centerlines from existing shipping route heat map [20]. In addition, it is also a common method to consider the application of neural network in the model to enhance the function of the KDE model. Rong et al. [31] applied improved kernel density estimation algorithm (KDE-T) and image processing method to mine the spatial-temporal characteristics of marine lanes.

Besides KDE, genetic algorithm was also frequently used to cluster ship trajectory data between waypoints for route extraction, the process of which begins when the GA completes the waypoint discovery [30], [36]. A directed graph was first created by checking the sequence of waypoints each ship traveled through and a series of intermediate operations including pruning and noise filtering as well as handling missing region data were applied on trajectory data between waypoints to reconstruct the route [39]. The GA approach could adapt to trajectory changes and handle imperfect input, which can discovery main route characteristics even without an ideal input.

Other studies focused on modeling ship trajectory data for a specific route and then delimited the route boundary and centerline. To model a ship route, Gaussian Process (GP) regression model which is a flexible and powerful Bayesian nonparametric approach for modelling time series, is applied to each observed ship trajectory separately by Rong et al. [24]. Then the Gaussian Process posterior distributions of the ship trajectories are aggregated into the ship route model. Benefiting from the inherent smoothing property of the Gaussian Process, the set of aggregated posterior distributions can easily estimate the route centerline and the route boundary at any location along the ship route.

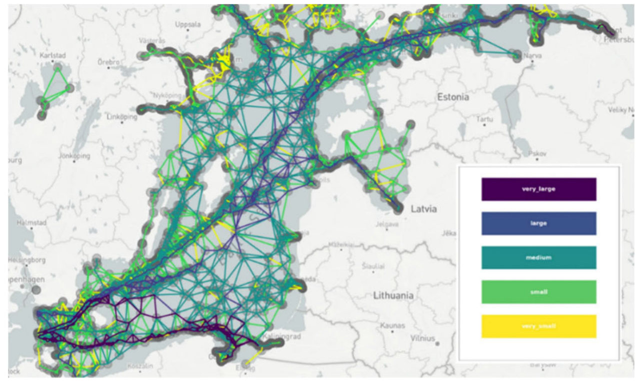


Fig. 4. Visualization of vectorized maritime route traffic network with different ship sizes [11].

C. Maritime Route Network Construction

After the waypoint identification and route extraction process, MRN can be constructed in three ways including topological networks, geometric networks, and semantic networks.

1) Vector-Based MRN: The vector-based methods pay attention to identify waypoints in ship trajectories and extract routes between waypoints to construct the vectorized MRN in the form of nodes-edges. The waypoints refer to nodes that could present ports, anchorage areas, capes, offshore platforms, or even any areas where vessels change their course and speed. The routes refer to edges that could characterize ship navigation characteristics between different nodes.

The vectorized MRN provides a light, structured and abstracted representation of maritime traffic flow, as shown in Figure 4 Varlamis et al. [11]. In the early studies, MRN is represented as a simple one-layer network with nodes and edges [7]. In order to cope with more complex traffic conditions such as uneven traffic density distribution in different water areas, a hierarchical MRN was further built upon two layers based on graph theory Arguedas et al. [8]. The external layer presented the visible elements of the maritime route network, including the previously detected routes and waypoints. It depicted the main characteristics of the route, such origin, destination or directionality, width and precision, as well as the relationship among various routes or the total maritime traffic. The internal layer synthetically and accurately represented individual route with a set of breakpoints which reflect the vessels constant and stable changes of behavior. It is meant to add granularity to the graph, adding detail and accuracy to the representation.

After that, followed studies paid major attention to enrich waypoints and maritime routes with traffic statistical information for comprehensive representation of MRN. A directed graph representation of maritime traffic was then proposed by Coscia et al. [28]. Each node represented a waypoint cluster in the form of convex region, and each edge was a navigational leg connecting the centroids of the waypoint clusters. The number of detected vessels transitioned between two clusters was used to calculate the orientation and weight of the navigational leg. Followed by that were navigational characteristics of maritime traffic, including the average travelling speed and heading of the vessels and their typical deviation, the average temporal distance of trajectory cluster on the route, which were added into the edges Kontopoulos et al. [13]. Rong et al. [3] further presented a probabilistic characterization of maritime traffic at the identified waypoints enriched with semantic information taken from the GeoNames ontology and Traffic Separation Schemes and route legs in terms of lateral distribution of the trajectories and speed profile, which allows the characterization of the typical behavior of a group of similar ships along a particular route [21]. Based on rich information above, some studies designed several rules to filter out invalid edges in the MRN [83], such as edges that are too long (e.g., distance>250km) or are very rare (e.g., followed by only a single vessel) Filipiak et al. [36]. Once the edges with low weight were turned off, some adjacent edges could be then merged to simplify the maritime route network Yan et al. [19]. The most visited edges were essentially retained as the traffic representation and highlighted the main traffic routes of ships Sing et al. [18]. Recently, more information has been considered to optimize the structure of MRN. Ship attributes were integrated with the edges in the MRN, and the weighted feature of an edge was computed from the cumulative count of ships by length, type, or both, which could provide a finer-grained depiction of quantitative traffic flow in the target waters Onyango et al. [32]. Several conditions, including boundary condition, continuity condition and monotonicity condition, were also put forward to obtain the optimal warping route between two connected nodes Liu et al. [1].

2) Grid-Based MRN: The grid-based methods focus on exploring the spatial extent of waypoints and ship navigation routes to construct the rasterized MRN in the form of route boundaries and centerlines. Route boundaries are a set of irregular polygons characterizing the spatial distribution of ship traffic, and centerlines are a set of segments representing the average routes or the representative routes of maritime traffic on the route [16], as shown Figure 5. Some of vessels in sea lane may move away from the centerline, but these deviations can still fall within a certain interval [10].

The rasterized MRN essentially sets up a zone graph representing the connection between important areas of interest. The connection between these zones was defined as a route, also called sea lane or sea leg. In general, the centerline could be first computed based on all trajectories in the route. For each trajectory, the distance to the centerline was calculated and a distance threshold (i.e., the ninth decile of these distances) was defined as an envelope around the route. A more thorough MRN was created by Wang et al. [48] using a 3-tuple graph with the three components nodes, segments, and segment curves. A node was a route centerline junction; A segment was an edge or arc indicated by two adjacent nodes where there was a centerline connected them. A segment curve was a curve made by discrete points, which are position points and grids it crossed. The main idea is to extracted different key points from various triangles generated from Delaunay triangulation algorithm separately and connect them into route centerlines, and then extract the information of nodes, segment and segment curves from the centerlines and add them to the MRN data structure. Followed by Lee et al. [51], a new method was further applied for verification of the safety of the novel maritime centerline compared with the original route quantitatively in terms of sinuosity, intersection angle and route change envelope. If the route stability of the novel maritime route, indicated by the vessel's historical track, is better, it will be widely used in the maritime traffic field. Motivated by the researches of using satellite image data to extract road networks, an automatic generation model based on the pix2pix network was put forward by Rong et al. [31] to reconstruct the maritime route network. Waypoint hotspot map was generated based on KDE-T analysis and then used to extract critical waypoints and traffic network boundaries based on conditional generative adversarial networks (pix2pix). Benefiting from the excellent expression ability of deep learning, the accuracy and robustness of the MRN were greatly improved. Recently, Liu et al. [22] further improved the effectiveness and the granularity of MRN generation based on grid important function. A boundary traffic density threshold was first used to locate the main route in the gridded clusters of navigational trajectories in different departure-arrival areas. it can avoid errors and minimize the interference of abnormal trajectories in precisely locating the main routes passing through the boundary of the water. The proposed cumulative grid importance function considering the grid value as well as the distance between the grid and the main route can effectively deal with the density differences among traffic groups, to identify the important area of the group.

3) Semantic-Based MRN: Topology is the most used form for network representation, and node-link model is the basic structure of topological network. Besides vectorized and rasterized MRN, a semantic network model in the form of topology was also proposed by Varlamis et al. to represent maritime traffic knowledge [11], which was called semantic traffic. Waypoints, routes, and ship entities were all abstracted as vertices of the network model, and directed edges with semantic annotations were used to link different types of vertices based on maritime traffic information. For instance, route vertices could be linked with waypoints using directed edges that denote the stay of the vessel within a waypoint (directed 'resides' label) or the travel direction of the moving vessel between the waypoint pair (directed 'voyage' edge).

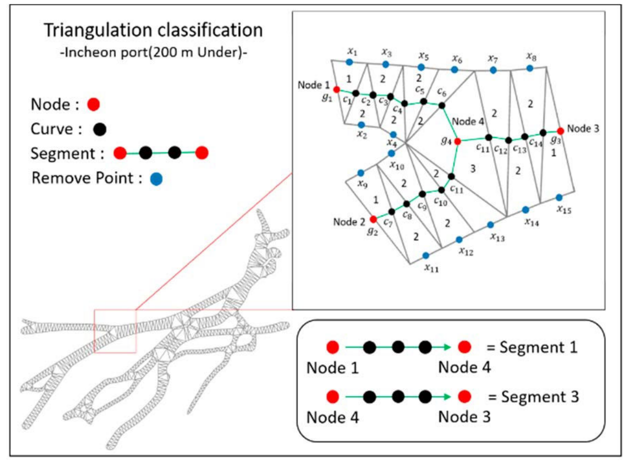


Fig. 5. Organization of the nodes, curves, and segments to create the centerlines [51].

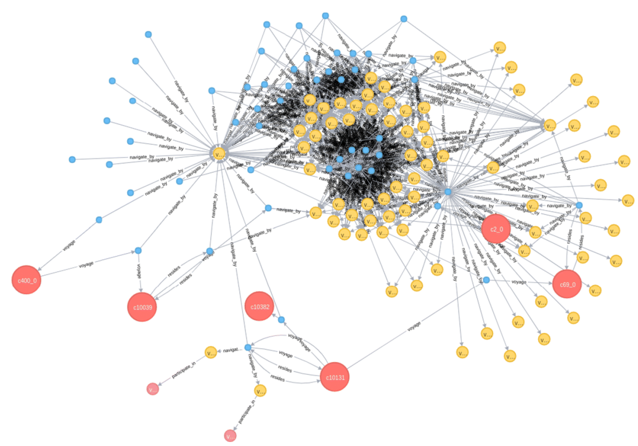


Fig. 6. Semantic network model of maritime traffic knowledge [11].

Vessel vertex could be connected with route vertex using directed 'navigate_by' edges that represent the movement of the vessel along the route, as shown in Figure 6.

Besides the directed edge, vertex in the semantic network model could also be semantically enriched during the creation and enrichment of the network abstraction. Any additional traffic information knowledge, such as average speed and standard deviation, direction, number of vessels that contributed to the route, area covered by a waypoint, and etc., can be attached to the network nodes. In addition, more types of nodes can be created to represent the navigation characteristics of vessels along different routes or across various waypoints. At present, the research on semantic route network is still in the initial stage, and there is still need to further study how to abstract the maritime traffic knowledge hidden in ship tracks and other maritime datasets into semantic network, expand and enrich it to facilitate maritime traffic management.

D. Discussions on Keys Steps of MRN Generation

Waypoint identification is a critical step of maritime route network construction. Due to the low quality of AIS data, many pre-process methods have been applied to extract feature points from raw ship trajectories for further waypoint identification. The rule detection method is easy to operate, but its parameter set is difficult since the navigational characteristics of various ships vary in different water areas. The Cumulative Sum method taking account into continuous motion features has a better effect on feature point detection, but it is vulnerable to extreme abnormal value of trajectory point data. From a perspective of spatio-temporal analysis, the DP method can extract the morphological key points of ship trajectory. It is difficult to set the optimal distance threshold of key point selection for all ship trajectories since there are obvious differences in ship movements. In addition, there is large computation cost when the water area is large scale. In recent study, various interpolation-based methods are compared to address the problem of detecting waypoints in a single representative trajectory with insufficient data, which can improve the spatial-temporal resolution of ship trajectory data for waypoint detection in critical areas [55]. Spatial-temporal features of ship trajectory data can also be analyzed and integrated to identify different types of waypoints, i.e., anchoring points and berthing points [35], [40].

After feature point extraction, it is necessary to further identify the common areas of these feature points as waypoints. Grid statistics methods are easy to operate, but the models are difficult to construct when processing large-scale data. Evolutionary algorithms (i.e., GA) try to extract the optimal route point group from massive trajectories by multiple iterations. These methods have to balance the accuracy and the completeness of key route point detection in large-scale water areas, and also have the problems of computing time and parameter determination. Comparatively, cluster methods such as DBSCAN, OPTICS, and their variations, facilitate the identification of irregular shapes of waypoint clusters and can effectively relieve the computational burden. However, the accuracy of clustering results is deeply affected by the uneven ship trajectory density in different areas, since adaptive parameter selection is still difficult. Therefore, several spatial partition strategies such as Quadtrees, K-D Trees and B-trees [36] can be used to improve the effectiveness of waypoint identification methods algorithms with different parameters. Meanwhile, distributed computing techniques such as Apache Spark are required to improve the processing speed and scalability [11], [28].

Regarding maritime route extraction, the cluster-based methods focus on calculating the similarities between different trajectories and classifying them into various maritime traffic patterns as routes between waypoints. These methods need a large number of ship tracks and excellent trajectory similarity evaluation model to ensure the effect of clustering. Moreover, the problem will be trickier when the ship behaviors are complex, especially in uncontrolled waters [22]. To deal with the problem, an incremental clustering method for hierarchically creating, merging and updating routes becomes more and more popular [38]. Besides, the basic unit for clustering can also affect the results of route extraction. From trajectory point to trajectory segment, local similarity is the major measurement of trajectory clustering. However, it is difficult to ensure the integrity of the extracted route in the region. Recent studies begin to explore global similarity model based on the complete trajectories for route extraction. Meanwhile, some time series analysis methods like Dynamic Time Warping are integrated for similarity evaluation considering computational efficiency. The grid-based route extraction methods pay attention to identifying the spatial extent and the representative movement of maritime routes from massive ship trajectory data. The selection of grid size will deeply affect the accuracy of route extraction due to uneven trajectory density distribution. The computation burden of grid-based methods incrementally increases as the expansion of the study area. Recent trend is to implement grid-based route extraction in a parallel way based on space partition, but an adaptive parameter determination method is still lacking. In addition, subsequent restoration treatments after parallel computation are also essential to the completeness of generated routes. The statistical-based route extraction methods allow quantitative modeling of maritime traffic features, which can help researchers determine important parameters or distributions related to the extraction of traffic patterns. These methods are non-sensitive to the maritime traffic density and support route extraction in different granularity. However, the statistical model is difficult to construct and other auxiliary techniques like edge smoothing and Delaunay triangulation are also essential to extract accurate route, which leads to high computational cost to obtain the optimal results when processing large-scale data. Overall, Existing methods still have various problems in adaptability to large-scale sea areas, complexity of methods, and compliance with real-world maritime routes, so there is an urgent need for a simple, feasible, and realistic method to extract regional or even global maritime traffic routes.

In terms of network construction, three types of MRN have their distinct advantages and disadvantages. Vectorized MRN can provide a multiscale skeleton of road network in sea, which mainly describe the connectivity and the movement direction between water areas. The navigational characteristics of maritime traffic are statistically analyzed and associated to different waypoints and routes. The vector-based methods of MRN construction rely heavily on the characteristics of data distribution, and there may be many overlapping, missing, and abnormal edges in the generated route network. The generalization ability of these methods is not strong, resulting in the problem of low completeness and accuracy of the maritime route network. Rasterized MRN delimits the spatial boundaries and the representative movements of connected maritime routes. It can provide clear distribution of maritime traffic but less navigational characteristics information of different routes. The grid-based methods of MRN construction requires a longer computing time and a higher capacity for computing resources and limits routes away from fairways and conventional routes [1]. Besides, abnormal data has a great influence on the results. The topological MRN is good at representing enrich information on nodes (vessels, waypoints, routes, etc.) and edges (their relations). However, the spatial and navigational characteristics of maritime routes are hidden in the nodes, and only relations are significantly expressed. The semantic-based methods of MRN construction

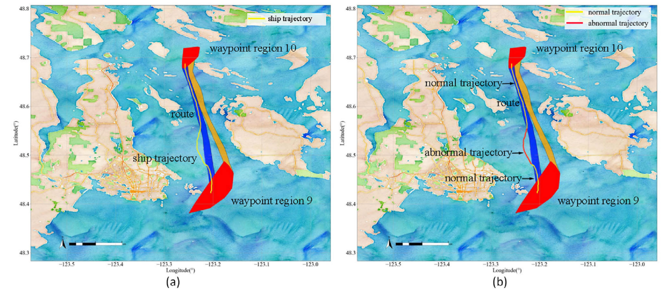


Fig. 7. Detection of vessel track deviation from planned route [91].

are essentially a semantic mapping process of waypoints and routes extracted by other methods. It provides a flexible and expandable data structure for constructing customed route network. More information including navigation rules, vessel behavior, logistical information, etc., can be integrated to construct a comprehensive MRN for applications. The difficulty is the structure of semantic MRN will become large and complex as the node information increases. The corresponding search and calculation methods of semantic MRN are in urgent need for enhancing the usability of the network model.

IV. Applications of MRN in the Maritime Industry

The construction of sea route networks can excavate channel information and discover the characteristic law of traffic flow. Improving the situational awareness in traffic safety management is conducive to strengthening the supervision of ships, reducing improper navigation pattern in complex water areas, identifying abnormal trajectory of ships and avoiding collision of ships. The construction of maritime traffic network can effectively understand the dynamic traffic and environmental data collection at sea. Through the analysis of the obtained traffic data, the ship track can be summarized to analyze the potential collision risk, the abnormal track to identify, and the future state of the ship. It plays a vital role in the maritime traffic safety management. This section introduces the typical application methods of ship trajectory anomaly identification, trajectory prediction and ship collision warning in the study area based on the extraction of maritime traffic network.

Several maritime situational awareness applications could benefit from a network representation of maritime traffic, presented as the following.

A. Detection of Ship Abnormal Behavior

Ship abnormal behavior detection is an important component in marine safety management. Through abnormal behavior detection, not only can it provide macroscopic support for ship navigation, it can also help maritime authorities monitor potential unconventional behavior and provide timely early warnings, supporting subsequent safety decisions. In anomaly detection, not only the normal behavior of ships should be modeled, but also the abnormal behavior of ships should be described by AIS trajectory data. Kontopoulos et al. [13] detected abnormal behavior of the ship based on the route boundary, which considers the trajectory that deviates from the polygonal area as abnormal trajectory, while the trajectory that navigates inside the polygonal area is considered normal trajectory. By using a clustering algorithm to extend the network edge, the algorithm can recognize different route motion patterns among the nodes of the same route network, and build motion models for different regions and ship types based on the clustering information, and use them to detect abnormal behaviors that deviate from the model.

Cluster analysis method is commonly used in ship anomaly recognition. Knorr et al. [56] constructed the recognition distance matrix between tracks by using the attributes of speed, direction and number of tracks in ship trajectory, and finally distinguished the abnormal path of ships based on the clustering results. Xiao et al. [57] based on the extracted traffic patterns, compared and analyzed different ship behavior patterns in the waters of Sutong Bridge and Rotterdam Port from multiple angles such as trajectory, bow direction and speed, and identified abnormal trajectory by cluster analysis. Boer et al. [58] analyze ship behavior characteristics in the study water area through AIS data. Based on the unsupervised DBSCAN algorithm, it effectively distinguished the distribution difference of trajectory in space position, and realized the classification and identification of different behavior patterns and abnormal behavior tracks. Some scholars proposed that a two-layer network can be used for abnormal behavior detection. The inner layer identifies details such as network nodes, and the outer layer connects nodes to form a track segment, forming not only a traffic network to represent navigation patterns. Arguedas et al. [8] proposed a two-layer network model, based on TREAD analysis method and DP algorithm, in the extraction of traffic network, the trajectory segment that deviates from the specified maximum distance difference and steering angle difference is stored, so as to realize abnormality detection.

With the deepening of the research on ship anomaly monitoring, scholars put forward the combination of different algorithms based on clustering method to improve the efficiency of abnormal behavior monitoring. Ristic et al. [59] proposed incremental DBSCAN algorithm to calculate ship behavior, aiming at the deficiency of traditional DBSCAN clustering algorithm. Then, the ship behavior pattern was mined by using kernel density estimation, and the algorithm for detecting abnormal ship behavior is constructed. Many methods based on machine learning, such as adaptive model of neural network, can obtain the result of abnormal identification by training the obtained data. Nguyen et al. [60] constructed GeoTrackNet to represent marine traffic patterns, recognize abnormal ship behavior and realize real-time situation awareness. First, the data representing the AIS timestamp are transformed into "four heat vectors", and then the trajectory distribution is obtained by variational recursive neural network (VRNN), that is, probabilistic traffic network. Then, the geospatial inverse detector is used to evaluate the possibility of AIS trajectory and realize the trajectory anomaly identification. Similarly, Liu et al. [61] proposed a two-stage computing framework. In the first stage, the DBSCAN analysis method was used to cluster trajectory data to identify unexpected outliers. In the second stage, a supervised learning technology based on bidirectional long-term memory (BLSTM) was proposed to reconstruct the location of identified outliers. Osekowska et al. [62] proposed that the study area be defined as a charge potential field based on grid statistics, and that the traffic pattern is represented by modeling marine traffic data based on the potential field. By analyzing the potential distribution charge and accumulation mode of AIS data ship position, the “charge” in the grid that does not meet the standard position is marked, so as to realize the identification of abnormal behavior.

A case of detection of ship abnormal behavior can be found in [91], as shown in Fig. 7(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 vessel’s trajectory in Fig. 7(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. 7(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.

B. Ship Trajectory Prediction

Ship traffic prediction is an important application of situational awareness. Ship traffic prediction is based on the acquired marine traffic pattern, and then the collected data are converted into the current and future conditions [88], [89]. Traffic prediction based on the constructed maritime traffic network can not only understand the navigation state of ships, but also identify the traffic density and risk hotspots by the output of the future position and trajectory prediction of ships, so as to realize the rational planning of navigation routes. At present, the trajectory prediction methods based on maritime traffic network can be divided into linear and nonlinear basic models, and the prediction results can be applied to applications such as collision risk assessment and traffic hotspot detection to enhance maritime situation awareness and traffic planning management.

The linear prediction model is also called the constant velocity model [63]. This method aims to use the time-stamped AIS trajectory data as a time-stamped data set and use the spatiotemporal information of the ship to represent the predicted trajectory. Starting from a characteristic time point, the average speed of the ship is constant, and along the projected route, the calculation formula is constructed to calculate the ship trajectory at a specific time interval. For example, Millefiori et al. proposed using a linear model based on a constant velocity model combined with an Ornstein Uhlenbeck (OU) mean regression stochastic process as a linear model for ship trajectory prediction [64]. With the modified target ship state equation, the calculation equation and the time scale rate of the associated uncertainty. Suo et al. proposed an automatic identification system using mean theorem and improved DP algorithm to learn the target ship model, obtain the ship position data and forecast route, and also help to analyze port operation [65]. Ristic et al. put forward a statistical model of ship motion pattern using adaptive kernel density estimation (KDE) to describe the distribution of position and velocity vectors in AIS data, and to predict the marine traffic with “nearly uniform motion model”. The constant speed model can deal with the trajectory prediction of short-term channel on straight track, because the ship navigation state change is generally small on straight track, so the prediction process is constant depending on course and speed, which makes the model simple and enhances the operability and practicability of the model. However, the linear model relies heavily on the input model trajectory data and the quality of traffic network. For the studied maritime traffic network, if the ship navigates along the turn section and the intersection area of the route, the constant speed and course assumed by the model can not meet the need of prediction effectively. In order to make up for the lack of precision of the constant model, many scholars put forward nonlinear prediction models in the follow-up research, including ship motion model, statistical model, machine learning algorithm, clustering algorithm and so on.

The most widely used method in nonlinear models is the neural network model, which passes through neural networks, such as artificial neural network (ANN) [66], backpropagation network (BP) [67], recurrent neural network (GRNN) [68] and long-term memory network (LSTM). Based on neural network technology, researchers focus on how to reasonably input and output the extracted traffic pattern data, and build adaptive training program through neural network to realize the model of ship position and track prediction. Specht et al. proposed to build a regression model based on GRNN to approximate nonlinear functions [69]. The prediction model is based on multiple GRNNs to train different length data sets and build multiple different regression models, so as to learn the long-term stable state of traffic network and realize the prediction task based on different traffic patterns. Liu et al. proposed a data-driven trajectory prediction framework based on AIS trajectory data mining [70]. The author proposes to embed the social force model (SFM) into the LSTM network to consider the interaction between ships, which can not only improve the accuracy of trajectory prediction, but also predict collision according to the relationship between ships, to warn the navigation risk of ships. Similarly, Yang et al. proposed the combination of DBSCAN and LSTM to realize the ship track recognition and behavior prediction [71].

The research on adaptive prediction model is to ensure the latest navigation state can be updated for online dynamic prediction output, for which the BP network training process is very standard, Xu et al. put forward the method based on backpropagation (BP) network. The difference between the course, speed and longitude and latitude of the input trajectory ship is obtained as the predicted longitude and latitude as the output. A BP network is used to obtain the calculated output from the longitude and latitude inputted by the current trajectory parameters, and then the loss is compared with the target output to obtain the loss, and finally the loss is back propagated to iteratively adjust the maritime traffic network parameters. Lacki [72] proposed a method based on artificial neural ANN to predict ship maneuvering, and based on the constructed topological network, it continuously adjusts the ship's travel route by means of Evolutionary Algorithm (EA), so as to provide the most appropriate traffic network for different types of ship's travel tasks. The model can also provide the evaluation result of speed prediction for route planning.

In addition to neural network, nonlinear model can predict ship trajectory by other methods. For example, Zhang et al. [73] proposed an improved k-nearest neighbor (KNN) method for short-term prediction of target ship position by calculating the similarity between neighboring points. Zandipour et al. [74] proposed a method based on association learning to predict the navigation state of ships, and uniformly dispersed the processed AIS trajectory in the study area based on grid, so as to complete the spatial matching. Then, the trajectory data is mined by association learning, and the correlation between target objects is found, and the ship trajectory prediction result is obtained. Similarly, Xiao et al. [75] also used DBSCAN algorithm to extract the ship traffic pattern based on the network, and uses KDE to study the distance change between routes to generate the travel distance distribution within a specific time window, according to which the prediction model is constructed.

C. Ship Route Planning

Ensuring the safe operation of ships on the sea is the core purpose of studying the traffic pattern extraction of maritime traffic. Ships are easily affected by weather, speed and other complex and changeable factors, which leads to unsafe conditions when ships are navigating in a channel. So how to use AIS trajectory data mining to abstract maritime traffic network for ship safe driving has been a focus of ship field. Ship safety planning is usually composed of short-term traffic prediction, collision detection and route optimization. Different from ship trajectory prediction, traffic prediction oriented to ship safety planning usually focuses on the prediction of ship position, allowing planners to know the location relationship between target ships and surrounding ships in a specific time, to detect collisions [76], to identify hot spots in a traffic network and to monitor ship navigation in real time. This enables risk warnings and route planning.

Zheng et al. [77] forecasted the future COG and SOG values of ships by exponential smoothing model. Then the prediction results are output, and the ship speed change and steering rate are calculated, and the ship route encounter situation is forecasted, and the ship collision is early-warned. Wen et al. [54] put forward a spatiotemporal mining algorithm, which uses non-parametric regression-local weighted scattergram to smooth record the position of the ship on the trajectory, obtains the fitted trajectory point cloud, and performs trajectory modeling through the historical fitting collection point cloud. Time-space ship collision behavior excavation for unsafe driving conditions to warn. Silveira et al. carried out a statistical analysis of maritime traffic off the coast of Portugal [78] and proposed a method for assessing collision risk to measure the risk status of the traffic network and the importance of the route of ships. Collision candidate sets are determined by predicting the ship position and the distance relation between the ship and the route, and compared with the collision threshold, and collision risk detection and warning are carried out.

How to plan ship route reasonably based on the extracted traffic pattern is also the key point of traffic safety planning. Unlike ship trajectory prediction, scholars mostly identify traffic density hotspots and risk hotspots by focusing on the potential location of ships. These hot spots are the key to traffic safety planning, as the input of route planning to optimize traffic routes and take relevant measures to mitigate collision risks. Young [79] proposed a ship position prediction method based on random forest. This paper established an ensemble learning method based on the original neural network to train the input forecast variables and output the predicted ship position based on the model. The distribution and hot spot area of the ship traffic volume in the traffic network can be obtained by using the estimated future position of the ship, to make the route planning. Wen et al. proposed the next turning position of the ship based on artificial neural network (ANN) [80], first using DBSCAN to obtain the turning region, then using artificial neural network to model, through iterative training to obtain the next turning point, And the optimal steering area is matched to optimize the navigation route automatically.

Generally, ship route planning is formulated as an optimization problem considering the navigation cost or emission reduction; thus, operational research methods were the primary solution. However, these methods relied heavily on the information from charts and paid little attention to the ship's seaworthiness. Because the planned routes based on historical navigation records are more likely to be consistent with the regular navigation pattern, some researchers have recently proposed data-driven route planning methods with historical navigation information. The core idea is to find the waypoints from massive navigation data, followed by linking turning areas. The linkage can be generated by considering the sailing distance or automatically learning from historical AIS data [85], [86], [87].

D. Resilience-Based Risk Management

The concept of resilience was first introduced by Holling [81], who defined resilience as "the ability to recover from wild shocks or the unexpected." Subsequently, researchers gradually applied resilience to other disciplines, including engineering resilience, economic resilience, and social resilience. Resilience concept offers a new perspective for risk management of the maritime shipping industry, and the connotations and characteristics of transport resilience can be found in a previous study by Wan et al. [82]. Generally, the resilience of an MRN can be measured from two aspects, which are vulnerability-oriented and recovery-oriented resilience analysis. The former paid attention to the vulnerability analysis (usually from a topological perspective) of main components consisting of an MRN such as ports, shipping routes, and key maritime channels. While, the latter mainly focused on the recovery stage of an MRN within certain time and costs after being affected by disruptions. Resilience-based risk management of MRNs is suggested at least considering the following aspects. First, the resilience of MRN in the post-epidemic era. Over the past three years, the port and shipping industry has experienced unprecedentedly significant market volatility and a volatile environment, with increased congestion, soaring freight rates, loading delays, and disruptions to maritime supply chains [84], [90]. COVID-19 and its associated restrictions have also caused severe disruptions to ports, and port-level risks will, in turn, continue to be passed on and amplified across global supply chains and cross-border trade. In addition to the COVID-19 pandemic, escalating geopolitical conflicts, increasing trends of counter-globalization, and new technological changes have put enormous pressure on port and shipping management. Therefore, in a volatile and uncertain environment [92], it is essential to strengthen risk management in ports and enhance their resilience management capabilities in order to ensure the resilience of global maritime supply chains. Second, the resilience of MRN under the Russian-Ukrainian Conflict. As Russia is a major global oil and gas exporter, the Russian-Ukrainian war has borne the brunt of the impact on international commodities. With the disruption of trade between Russia and Europe and the escalation of sanctions against Russia in Europe and the US, the supply of commodities, including crude oil, gas, and food, has become tight, and prices have soared. In addition, the military conflict has halted operations in the Black Sea and Ukrainian ports, bringing bulk cargoes to a near halt and hitting the European maritime supply chain hard. Third, the impact of climate change and carbon emission policies on the resilience of MRN. Climate change is a challenging global issue, impacting countries politically, economically, and culturally. The international shipping industry has significantly contributed to global greenhouse gas emissions, accounting for approximately 3% of total global greenhouse gas emissions yearly. The long-term challenge for shipping companies is how to cope with increasingly stringent policies to reduce emissions from shipping.

V. Conclusion

This paper systematically reviews and summarizes the construction methods and applications of MRN, in the aspects of waypoint identification, route extraction, network expression and network application based on MRN. New methods and technologies are being actively studied, and great research results and progress have been made. The paper also introduces the key content of different research and compares the advantages and disadvantages of different methods.

Overall, there are still many deficiencies in the methods of extracting traffic patterns. For example, the accuracy of route network construction depends largely on the accuracy of waypoint recognition. However, when the existing waypoint recognition method faces large-scale, high-noise and unevenly distributed ship track data, there are still some problems such as insufficient accuracy, low efficiency, and not being able to extract the route point clusters with different traffic flow density adaptively. The extracted traffic patterns are usually applied to simple routes, which cannot reflect the real complex traffic conditions completely. Moreover, the existing research tends to explore how to extract traffic models and design the method, but there are few considerations about the precision, storage and query methods of the extracted traffic features. During route extraction, most studies only extract main routes, which leads to incomplete route network. In the generation and expression of ship route networks, the current research lacks further mining and semantic modeling of ship behavior pattern in route network. Additionally, most methods are still experimentally verified in limited cases or specific scenarios, and the application scenarios of route networks remain relatively narrow. With research based on just a few special cases or scenarios, the varying navigation states across different areas require further investigation. Therefore, future work should focus on optimizing the construction and expression methods of MRNS, as well as expanding the applications of MRNs. Potential research directions include:

1) The MRN should be constructed multi-scale, multi-level, and multi-type manner to accommodate different application scenarios. For example, from a geospatial perspective, route networks can be constructed at the global scale between major ports or port groups, at the regional scale for coastal areas, at the local scale for port waters or intersection waters, and so on. This enables constructing the route network from course to fine resolutions, with different levels and types at each level. Route networks can also be tailored for different ship types and sizes, reflecting the varying navigation modes and movement patterns of distinct vessel categories.
2) It is necessary to extract the route points and routes adaptively based on varying traffic density, and generate route networks that capture different density flows. Research should explore route network expression modes and visualization methods to enable comprehensive, intuitive, accurate, real-time, multi-level representations of marine traffic patterns. As maritime trade expands globally, effective storage of route network data (nodes and edges) for mapping, querying and searching will facilitate utilization of maritime traffic knowledge - this is an important research need.
3) MRNs exhibit unbalanced structures. Studying how local network changes influence the overall network, and evaluating transportation efficiency, are critical research directions. Drawing from graph theory and complex network knowledge, the stability and transportation efficiency of MRN can be analyzed.
4) As marine traffic service networks advance, further research can explore ship navigation and movement patterns within node and route areas. This enables capabilities like automatic route planning, intelligent

navigation [93], trajectory prediction, and abnormality detection based on the route network. Additionally, regional marine traffic management can be strengthened using the sea route network. For example, by monitoring real-time traffic flow changes, local network carrying capacity can be considered holistically to rationally allocate ships and further improve transportation efficiency and economics.

5) In future research, a three-dimensional transportation network connecting sea, land, and air could be constructed. Analyzing the characteristics and correlations of networks across these spatial domains, coordinating their various functions, and enabling seamless transitions would maximize convenience for passengers and freight.

Acknowledgment

The authors would like to express their sincere gratitude to Prof. Ming Yang and Di Xiang for their contribution to this work.

References

[1] L. Liu, R. Shibasaki, Y. Zhang, N. Kosuge, M. Zhang, and Y. Hu, "Data-driven framework for extracting global maritime shipping networks by machine learning," *Ocean Eng.*, vol. 269, Feb. 2023, Art. no. 113494.

[2] L. Huang et al., "Mobility pattern analysis of ship trajectories based on semantic transformation and topic model," *Ocean Eng.*, vol. 201, Apr. 2020, Art. no. 107092.

[3] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Data mining approach to shipping route characterization and anomaly detection based on AIS data," *Ocean Eng.*, vol. 198, Feb. 2020, Art. no. 106936.

[4] D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, "A multi-task deep learning architecture for maritime surveillance using AIS data streams," in *Proc. IEEE 5th Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Oct. 2018, pp. 331–340.

[5] G. Pallotta, M. Vespe, and K. Bryan, "Traffic knowledge discovery from AIS data," in *Proc. 16th Int. Conf. Inf. Fusion*, Jul. 2013, pp. 1996–2003.

[6] 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. 12, pp. 2218–2245, Jun. 2013.

[7] V. F. Arguedas, G. Pallotta, and M. Vespe, "Automatic generation of geographical networks for maritime traffic surveillance," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–8.

[8] 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.

[9] G. Pallotta, S. Horn, P. Braca, and K. Bryan, "Context-enhanced vessel prediction based on Ornstein–Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–7.

[10] G. Pallotta and A.-L. Jousselme, "Data-driven detection and context-based classification of maritime anomalies," in *Proc. 18th Int. Conf. Inf. Fusion (Fusion)*, Jul. 2015, pp. 1152–1159.

[11] I. Varlamis, I. Kontopoulos, K. Tserpes, M. Etemad, A. Soares, and S. Matwin, "Building navigation networks from multi-vessel trajectory data," *GeoInformatica*, vol. 25, pp. 69–97, Aug. 2021.

[12] P. Prasad, V. Vatsal, and R. R. Chowdhury, "Maritime vessel route extraction and automatic information system (AIS) spoofing detection," in *Proc. Int. Conf. Adv. Elect., Comput., Commun. Sustain. Technol. (ICAECT)*, Feb. 2021, pp. 1–11.

[13] I. Kontopoulos, I. Varlamis, and K. Tserpes, "Uncovering hidden concepts from AIS data: A network abstraction of maritime traffic for anomaly detection," in *Multiple-Aspect Analysis of Semantic Trajectories* (Lecture Notes in Artificial Intelligence). Würzburg, Germany, 2020, pp. 6–20.

[14] N. Forti, L. M. Millefiori, and P. Braca, "Unsupervised extraction of maritime patterns of life from automatic identification system data," in *Proc. OCEANS*, Jun. 2019, pp. 1–5.

[15] L. N. Tafa, X. Su, J. Hong, and C. Choi, "Automatic Maritime traffic synthetic route: A framework for route prediction," in *Pervasive Systems, Algorithms and Networks* (Communications in Computer and Information Science). Naples, Italy: Springer, 2019, pp. 3–14.

[16] 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 Anal. (ICCCBDA)*, Apr. 2020, pp. 7–12.

[17] W. Yan, R. Wen, A. N. Zhang, and D. Yang, "Vessel movement analysis and pattern discovery using density-based clustering approach," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2016, pp. 3798–3806.

[18] S. K. Singh, J. S. Fowdur, J. Gawlikowski, and D. Medina, "Leveraging graph and deep learning uncertainties to detect anomalous maritime trajectories," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 23488–23502, Dec. 2022.

[19] Z. Yan et al., "Exploring AIS data for intelligent maritime routes extraction," *Appl. Ocean Res.*, vol. 101, Aug. 2020, Art. no. 102271.

[20] R. W. Liu, M. Liang, J. Nie, S. Garg, Y. Zhang, and Z. Xiong, "Extraction of hottest shipping routes: From positioning data to intelligent surveillance," in *Proc. IEEE 22nd Int. Conf. Inf. Reuse Integr. Data Sci. (IRI)*, Aug. 2021, pp. 255–262.

[21] H. Rong, Â. P. Teixeira, and C. G. Soares, "Maritime traffic network extraction and application based on AIS data," in *Proc. 6th Int. Conf. Transp. Inf. Saf. (ICTIS)*, Oct. 2021, pp. 1244–1252.

[22] Z. Liu, H. Gao, M. Zhang, R. Yan, and J. Liu, "A data mining method to extract traffic network for maritime transport management," *Ocean Coastal Manage.*, vol. 239, May 2023, Art. no. 106622.

[23] M. Liang, R. W. Liu, Y. Zhan, H. Li, F. Zhu, and F.-Y. Wang, "Fine-grained vessel traffic flow prediction with a spatio-temporal multigraph convolutional network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 23694–23707, Dec. 2022.

[24] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Maritime traffic probabilistic prediction based on ship motion pattern extraction," *Rel. Eng. Syst. Saf.*, vol. 217, Jan. 2022, Art. no. 108061.

[25] L. Eljabu, M. Etemad, and S. Matwin, "Spatial clustering model of vessel trajectory to extract sailing routes based on AIS data," *Int. J. Comput. Syst. Eng.*, vol. 16, no. 10, pp. 491–501, 2022.

[26] L. Eljabu, M. Etemad, and S. Matwin, "Spatial clustering method of historical AIS data for maritime traffic routes extraction," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2022, pp. 893–902.

[27] J. Gonzalez, G. Battistello, P. Schmiegelt, and J. Biermann, "Semi-automatic extraction of ship lanes and movement corridors from AIS data," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 1847–1850.

[28] P. Coscia, P. Braca, L. M. Millefiori, F. A. N. Palmieri, and P. Willett, "Multiple Ornstein–Uhlenbeck processes for maritime traffic graph representation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 5, pp. 2158–2170, Oct. 2018.

[29] S.-K. Zhang, G.-Y. Shi, Z.-J. Liu, Z.-W. Zhao, and Z.-L. Wu, "Data-driven based automatic maritime routing from massive AIS trajectories in the face of disparity," *Ocean Eng.*, vol. 155, pp. 240–250, May 2018.

[30] D. Filipiak, M. Stróżyna, K. Wecel, and W. Abramowicz, "Application of AI and in-memory computing for extracting vessel movement patterns from historical data," in *Proc. 14th NATO Oper. Res. Anal. (OR&A) Conf., Emerg. Disruptive Technol.*, 2021, pp. 6–10.

[31] Y. Rong, Z. Zhuang, Z. He, and X. Wang, "A maritime traffic network mining method based on massive trajectory data," *Electronics*, vol. 11, no. 7, p. 987, Mar. 2022.

[32] S. O. Onyango, S. A. Owiredu, K.-I. Kim, and S.-L. Yoo, "A quasi-intelligent maritime route extraction from AIS data," *Sensors*, vol. 22, no. 22, p. 8639, Nov. 2022.

[33] A. Dobrkovic, M. E. Iacob, and J. V. Hillegersberg, "Using machine learning for unsupervised maritime waypoint discovery from streaming AIS data," in *Proc. 15th Int. Conf. Knowl. Technol. Data-Driven Business*, 2015, pp. 1–8.

[34] I. Kontopoulos, I. Varlamis, and K. Tserpes, "A distributed framework for extracting maritime traffic patterns," *Int. J. Geographical Inf. Sci.*, vol. 35, no. 4, pp. 767–792, Apr. 2021.

[35] Z. Yan, L. Cheng, R. He, and H. Yang, "Extracting ship stopping information from AIS data," *Ocean Eng.*, vol. 250, Apr. 2022, Art. no. 111004.

[36] D. Filipiak, K. Węcel, M. Stróżyna, M. Michalak, and W. Abramowicz, "Extracting maritime traffic networks from AIS data using evolutionary algorithm," *Bus. Inf. Syst. Eng.*, vol. 62, no. 5, pp. 435–450, Oct. 2020.

[37] H. Li et al., "Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery," *Transp. Res. C, Emerg. Technol.*, vol. 143, Oct. 2022, Art. no. 103856.

[38] D. Zhang, Y. Zhang, and C. Zhang, "Data mining approach for automatic ship-route design for coastal seas using AIS trajectory clustering analysis," *Ocean Eng.*, vol. 236, Sep. 2021, Art. no. 109535.

[39] A. Dobrkovic, M. E. Iacob, and J. van Hillegersberg, "Maritime pattern extraction and route reconstruction from incomplete AIS data," *Int. J. Data Sci. Anal.*, vol. 5, nos. 2–3, pp. 111–136, Mar. 2018.

[40] L. Huang et al., "Stopping behavior recognition and classification of ship based on trajectory characteristics," *J. Traffic Transp. Eng*, vol. 21, no. 5, pp. 189–198, 2021.

[41] M. Y. Santos, J. P. Silva, J. Moura-Pires, and M. Wachowicz, "Automated traffic route identification through the ared nearest neighbour algorithm," in *Bridging the Geographic Information Sciences* (Lecture Notes in Geoinformation and Cartography). Avignon, France: Springer, 2012, pp. 231–248.

[42] P. Sheng and J. Yin, "Extracting shipping route patterns by trajectory clustering model based on automatic identification system data," *Sustainability*, vol. 10, no. 7, p. 2327, Jul. 2018.

[43] C. Huang, X. Qi, J. Zheng, R. Zhu, and J. Shen, "A maritime traffic route extraction method based on density-based spatial clustering of applications with noise for multi-dimensional data," *Ocean Eng.*, vol. 268, Jan. 2023, Art. no. 113036.

[44] Y. T. Wen, C. H. Lai, P. R. Lei, and W. C. Peng, "RouteMiner: Mining ship routes from a massive maritime trajectories," in *Proc. IEEE 15th Int. Conf. Mobile Data Manage.*, Jul. 2014, pp. 353–356.

[45] P. R. Lei, T. H. Tsai, and W. C. Peng, "Discovering Maritime Traffic Route from AIS network," in *Proc. 18th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Oct. 2016, pp. 1–6.

[46] R. Vettor and C. G. Soares, "Detection and analysis of the main routes of voluntary observing ships in the North Atlantic," *J. Navig.*, vol. 68, no. 2, pp. 397–410, 2015.

[47] L. Wu, Y. Xu, Q. Wang, F. Wang, and Z. Xu, "Mapping global shipping density from AIS data," *J. Navig.*, vol. 70, no. 1, pp. 67–81, Jan. 2017.

[48] 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.

[49] G. Wang, M. Zhang, J. Gao, and Y. Han, "Where is the next path? A deep learning approach to path prediction without prior road networks," in *Collaborative Computing: Networking, Applications and Worksharing* (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering). Shanghai, China: Springer, 2021, pp. 219–235.

[50] K. Tzavella and M. Ulmke, "Semi-automatic sea lane extraction combining particle filtering (PF) and geographic information systems (GIS)," in *Proc. GI_Forum Geospatial Innov. Soc.*, 2014, pp. 2–11.

[51] 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.

[52] 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.

[53] M. Anneken, A.-L. Jousselme, S. Robert, and J. Beyerer, "Synthetic trajectory extraction for maritime anomaly detection," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2018, pp. 1048–1053.

[54] R. Wen, W. Yan, A. N. Zhang, N. Q. Chinh, and O. Akcan, "Spatio-temporal route mining and visualization for busy waterways," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 849–854.

[55] B. Zaman, D. Marijan, and T. Kholodna, "Interpolation-based inference of vessel trajectory waypoints from sparse AIS data in maritime," *J. Mar. Sci. Eng.*, vol. 11, no. 3, p. 615, Mar. 2023.

[56] E. M. Knorr and N. V. Tucakov, "Distance-based outliers: Algorithms and applications," *VLDB J.*, vol. 8, pp. 237–253, Feb. 2000.

[57] F. Xiao, H. Ligteringen, C. van Gulijk, and B. Ale, "Comparison study on AIS data of ship traffic behavior," *Ocean Eng.*, vol. 95, pp. 84–93, Feb. 2015.

[58] T. M. D. Boer, "An analysis of vessel behaviour based on AIS data," M.S. thesis, Dept. Civil Eng. Geosci., Delft Univ. Technol., Delft, The Netherlands, 2010.

[59] B. Ristic, "Detecting anomalies from a multitarget tracking output," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 798–803, Jan. 2014.

[60] D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, "GeoTrackNet—A maritime anomaly detector using probabilistic neural network representation of AIS tracks and a contrario detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5655–5667, Jun. 2022.

[61] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, and M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5374–5385, Apr. 2021.

[62] E. Osekowska, H. Johnson, and B. Carlsson, "Grid size optimization for potential field based maritime anomaly detection," *Transp. Res. Proc.*, vol. 3, pp. 720–729, Jan. 2014.

[63] M. Posada, H. Greidanus, M. Alvarez, M. Vespe, T. Cokacar, and S. Falchetti, "Maritime awareness for counter-piracy in the Gulf of Aden," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Vancouver, BC, Canada, Jul. 2011, pp. 249–252.

[64] 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.

[65] Y. Suo, W. Chen, C. Claramunt, and S. Yang, "A ship trajectory prediction framework based on a recurrent neural network," *Sensors*, vol. 20, no. 18, p. 5133, Sep. 2020.

[66] A. Daranda, "Neural network approach to predict marine traffic," *Trans. Baltic J. Mod. Comput.*, vol. 4, no. 3, p. 483, 2016.

[67] T. Xu, X. Liu, and X. Yang, "A novel approach for ship trajectory online prediction using BP neural network algorithm," *Int. J. Adv. Inf. Sci. Service Sci.*, vol. 4, no. 11, pp. 271–277, 2012.

[68] P. Borkowski, "The ship movement trajectory prediction algorithm using navigational data fusion," *Sensors*, vol. 17, no. 6, p. 1432, Jun. 2017.

[69] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Netw.*, vol. 2, no. 6, pp. 568–576, Nov. 1991.

[70] 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. Netw. Sci. Eng.*, vol. 9, no. 5, pp. 3080–3094, Sep. 2022.

[71] C.-H. Yang, G.-C. Lin, C.-H. Wu, Y.-H. Liu, Y.-C. Wang, and K.-C. Chen, "Deep learning for vessel trajectory prediction using clustered AIS data," *Mathematics*, vol. 10, no. 16, p. 2936, Aug. 2022.

[72] M. Łącki, "Intelligent prediction of ship maneuvering," *TransNav: Int. J. Mar. Navig. Saf. Sea Transp.*, vol. 10, no. 3, pp. 1–6, 2016.

[73] M. Zhang, L. Huang, Y. Wen, J. Zhang, Y. Huang, and M. Zhu, "Short-term trajectory prediction of maritime vessel using k-nearest neighbor points," *J. Mar. Sci. Eng.*, vol. 10, no. 12, p. 1939, Dec. 2022.

[74] M. Zandipour, B. J. Rhodes, and N. A. Bomberger, "Probabilistic prediction of vessel motion at multiple spatial scales for maritime situation awareness," in *Proc. 11th Int. Conf. Inf. Fusion*, Jun./Jul. 2008, pp. 1–6.

[75] Z. Xiao, L. Ponnambalam, X. Fu, and W. Zhang, "Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3122–3134, Nov. 2017.

[76] H.-C. Burmeister, W. Bruhn, Ø. J. Rødseth, and T. Porathe, "Autonomous unmanned merchant vessel and its contribution towards the e-Navigation implementation: The MUNIN perspective," *Int. J. e-Navig. Maritime Economy*, vol. 1, pp. 1–13, Dec. 2014.

[77] L. Qi and Z. Zheng, "Trajectory prediction of vessels based on data mining and machine learning," *J. Digit. Inf. Manage.*, vol. 14, no. 1, pp. 33–40, 2016.

[78] P. Silveira, A. Teixeira, and C. G. Soares, "Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal," *J. Navigat.*, vol. 66, no. 6, pp. 879–898, 2013.

[79] B. L. Young, "Predicting vessel trajectories from AIS data using R," M.S. thesis, Nav. Postgraduate School, Monterey, CA, USA, 2017.

[80] Y. Wen et al., "Automatic ship route design between two ports: A data-driven method," *Appl. Ocean Res.*, vol. 96, Mar. 2020, Art. no. 102049.

[81] C. S. Holling, "Resilience and stability of ecological systems," *Annu. Rev. Ecol. Syst.*, vol. 4, no. 1, pp. 1–23, 1973.

[82] C. Wan, Z. Yang, D. Zhang, X. Yan, and S. Fan, "Resilience in transportation systems: A systematic review and future directions," *Transp. Rev.*, vol. 38, no. 4, pp. 479–498, 2018.

[83] I. L. Ansorena, "On the liner shipping network design: The maritime silk route case study," *Int. J. Logistics Syst. Manage.*, vol. 44, no. 1, pp. 132–146, 2023.

[84] X. Xin, K. Liu, S. Loughney, J. Wang, H. Li, and Z. Yang, "Graph-based ship traffic partitioning for intelligent maritime surveillance in complex port waters," *Expert Syst. Appl.*, vol. 231, Nov. 2023, Art. no. 120825.

[85] C. Liu, Y. Ma, C. Cao, and X. Yan, "Ship route planning in the pirate area via hybrid probabilistic roadmap algorithm within the context of the maritime silk road," *Ocean Coastal Manage.*, vol. 238, May 2023, Art. no. 106585.

[86] D. Liu, H. Rong, and C. Guedes Soares, "Shipping route modelling of AIS maritime traffic data at the approach to ports," *Ocean Eng.*, vol. 289, Dec. 2023, Art. no. 115868.

[87] Z. Guo, H. Qiang, S. Xie, and X. Peng, "Unsupervised knowledge discovery framework: From AIS data processing to maritime traffic networks generating," *Appl. Ocean Res.*, vol. 146, May 2024, Art. no. 103924.

[88] Y. Shin, N. Kim, H. Lee, S. Y. In, M. Hansen, and Y. Yoon, "Deep learning framework for vessel trajectory prediction using auxiliary tasks and convolutional networks," *Eng. Appl. Artif. Intell.*, vol. 132, Jun. 2024, Art. no. 107936.

[89] D. Kaklis, I. Kontopoulos, I. Varlamis, I. Z. Emiris, and T. Varelas, "Trajectory mining and routing: A cross-sectoral approach," *J. Mar. Sci. Eng.*, vol. 12, no. 1, p. 157, Jan. 2024.

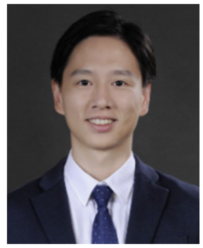
[90] L. Cong, H. Zhang, P. Wang, C. Chu, and J. Wang, "Impact of the Russia–Ukraine conflict on global marine network based on massive vessel trajectories," *Remote Sens.*, vol. 16, no. 8, p. 1329, Apr. 2024.

[91] C. Zhou et al., "TTMRN: A topological-geometric two-layer maritime route network modeling for ship intelligent navigation," *Ocean Eng.*, vol. 287, Nov. 2023, Art. no. 115884.

[92] C. Wan, J. Yuan, D. Cao, T. Wang, and A. K. Ng, "A fuzzy evidential reasoning-based model for evaluating resilience of ports to typhoons," *Transp. Res. Part D, Transp. Environ.*, vol. 133, Aug. 2024, Art. no. 104228.

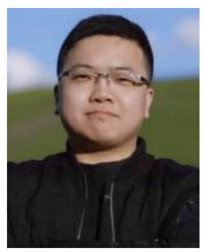
[93] C. Wan, Y. Zhao, D. Zhang, and L. Fan, "A system dynamics-based approach for risk analysis of waterway transportation in a mixed traffic environment," *Maritime Policy Manage.*, vol. 51, no. 6, pp. 1147–1169, 2023.

Liang Huang received the Ph.D. degree in cartography and geographical information engineering from Wuhan University, China, in 2014. From June 2015 to June 2018, he conducted Post-Doctoral Research at Wuhan University of Technology (WUT), focusing on maritime big data mining. He is currently an Associate Professor at the Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan. In 2023, he was a visiting scholar at Delft University of Technology, The Netherlands. His research interests include maritime safety and environmental monitoring, maritime big data analytics, and intelligent navigation systems.



Chengpeng Wan received the B.S degree in marine engineering from Wuhan University of Technology (WUT), Wuhan, China, in 2012, and two Ph.D degrees with one in transportation engineering from WUT, in 2018, while the other one in maritime technology from Liverpool John Moores University of Technology, Liverpool, U.K.,in 2019. He is currently an Associate Professor with the Intelligent Transportation Systems Research Center, Wuhan University of Technology, China. He is the author of three book chapters and more than 30 journal articles. His research interests include transportation system resilience, transport network modelling and analysis, and risk management of intelligent maritime transport.

Yuanqiao Wen received the doctorate degree from Huazhong University of Science and Technology in 2006. He is a Vice Director of the National Engineering Research Center for Water Transportation Safety and Vice Director of Intelligent Traffic System Research Center (Wuhan University of Technology). His research centers on waterborne transportation safety and environment, green and sustainable shipping technologies. He has hosted or participated in six National Natural Science Foundation projects and National Major Science and Technology Support Plan projects and published more than 140 research papers.



Rongxin Song is a Ph.D. candidate in the decision-making of Maritime Autonomous Surface Ships (MASS) with the Section of Safety and Security Science, TU Delft. His research focuses on the build-up of situational awareness for MASS, proactive collision avoidance strategy design for MASS, and trust-aware navigation by MASS.

Pieter van Gelder is currently a Professor with the Safety Science in the Safety and Security Science Section, TU Delft. He has been involved in research and education on safety and reliability since 1991. He teaches 4th and 5th year courses at TU Delft and conducts research on new methods and techniques in risk analysis, in particular on decompositioning techniques and probabilistic analytics. He has published a large number of highly cited articles in the field of infrastructure safety, the statistical modelling of high impact Low Probability (HILP) events and the consequent decision-making processes, with applications to natural hazards and infrastructures. His research interests are in risk analysis and optimisation of systems, processes and structures.