Li, H and Yang, Z

Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships

Li, H and Yang, Z (2023) Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships. Transportation Research Part E: Logistics and Transportation Review, 176. p. 103171. ISSN 1366-5545

Transportation Research Part E 176 (2023) 103171

Contents lists available at ScienceDirect

Transportation Research Part E

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

Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships

Huanhuan Li, Zaili Yang *

Liverpool Logistics Offshore and Marine Research Institute, Liverpool John Moores University, Liverpool, UK

ARTICLE INFO

Keywords:
Maritime Autonomous Surface Ships (MASS)
Feature measurement
Route planning
Pattern extraction
Maritime safety

ABSTRACT

Maritime Autonomous Surface Ships (MASS) are deemed as the future of maritime transport. Although showing attractiveness in terms of the solutions to emerging challenges such as carbon emission and insufficient labor caused by black swan events such as COVID-19, the applications of MASS have revealed problems in practice, among which MASS navigation safety presents a prioritized concern. To ensure safety, rational route planning for MASS is evident as the most critical step to avoiding any relevant collision accidents. This paper aims to develop a holistic framework for the unsupervised route planning of MASS using machine learning methods based on Automatic Identification System (AIS) data, including the coherent steps of new feature measurement, pattern extraction, and route planning algorithms. Historical AIS data from manned ships are trained to extract and generate movement patterns. The route planning for MASS is derived from the movement patterns according to a dynamic optimization method and a feature extraction algorithm. Numerical experiments are constructed on real AIS data to demonstrate the effectiveness of the proposed method in solving the route planning for different types of MASS.

1. Introduction

Shipping has undertaken more than 80% of global trade volume as a strong tie and primary indicator between the global economy and maritime trade (Bueger, 2015; Li et al., 2023b). Meanwhile, according to the global annual reports of marine casualties and accidents from 2014 to 2020 by the European Maritime Safety Agency (EMSA), the numbers of marine casualties and accidents (Fig. 1a) remain high, despite huge efforts in the past decades. The distribution of different ship types during the period of 2014 to 2020 (Fig. 1b) shows that cargo ships have the highest accident rate, indicated by the 24,772 relevant casualties and accidents (EMSA, 2022). Safety at sea has been at the top of the research agenda in maritime transport, and the emerging risks due to the fast development and possible uncertainties brought by the implementation of advanced technologies such as Maritime Autonomous Surface Ships (MASS) have even become more worrisome in recent years (Chang et al., 2021; Li and Fung, 2019; Zhang et al., 2021b).

Within the context of shipping automation and intelligence, the International Maritime Organisation (IMO) initially raised an E-Navigation concept to enhance navigation safety and facilitate the Maritime Intelligent Transportation System (M-ITS) in 2005. The modernization of the satellite-terrestrial is integrated with communication networks in M-ITS, enabling the collection of massive data in maritime traffic and promoting data mining technologies for maritime situational awareness and knowledge discovery (Li et al.,

* Corresponding author.
E-mail address: Z.Yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.tre.2023.103171
Received 9 June 2022; Received in revised form 19 April 2023; Accepted 16 May 2023
Available online 7 June 2023
1366-5545/

Nomenclature

Roman letters

Variable	**Definition**
AADTW	Automatic and Adaptive Dynamic Time Warping
ACDTW	Adaptive Constrained Dynamic Time Warping
ADP	Automatic Douglas and Peucker
AE	AutoEncoder
ARPA	Automatic Radar Plotting Aid
AIS	Automatic Identification System
ANN	Artificial Neural Network
ASV	Autonomous Surface Vehicles
CAE	Convolutional Auto-Encoder
CHS	Calinski-Harabasz Score
CNN	Convolutional Neural Network
COG	Course Over Ground
COLREG	Convention on the International Regulations for Preventing Collisions at Sea
CUSUM	CUmulative SUM
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DBSCANSD	Density-Based Spatial Clustering of Applications with Noise considering Speed and Direction
DDPG	Deep Deterministic Policy Gradient
DRL	Deep Reinforcement Learning
DTW	Dynamic Time Warping
ECDIS	Electronic Chart Display and Information System
EMSA	European Maritime Safety Agency
IMO	International Maritime Organisation
1NN	The Nearest Neighbor search
KNN	K-Nearest Neighbor
LNG	Liquefied Natural Gas
LPW	Local Planning Window
LRIT	Long-Range Identification and Tracking
LSTM	Long Short-Term Memory
MAS	Mayflower Autonomous Ship
MASS	Maritime Autonomous Surface Ships
M-ITS	Maritime Intelligent Transportation System
MMSI	Maritime Mobile Service Identity
PCA	Principal Component Analysis
RNN	Recurrent Neural Network
SAE	Sparse AutoEncoder
SCAF	Spectral Clustering with Affinity Feature
SC	Silhouette Coefficient
SOG	Speed Over Ground
USV	Unmanned Surface Vehicles
VAE	Variational Auto-Encoder
VG	Visibility Graph
VTS	Vessel Traffic Services

2023a; Zhang et al., 2011, 2022b). The global communication network of maritime vessels in M-ITS is shown in Fig. 2, which can help aid in realizing intelligent vessel traffic services according to the exchanged information by advanced techniques. Automatic Identification System (AIS) is a vessel self-reporting messaging system that can broadcast and provide a large amount of near-real static and dynamic information by combining with Radar and Vessel Traffic Services (VTS) (Goudossis and Katsikas, 2019; Liang et al., 2022; Liu et al., 2022b). The AIS data includes Maritime Mobile Service Identity (MMSI), position, Speed Over Ground (SOG), and Course Over Ground (COG), etc., which can provide navigation guidance and route reference for manned and unmanned ships. Facilitated by the widespread use of AIS, Automatic Radar Plotting Aid (ARPA), Long-Range Identification and Tracking (LRIT), and VTS equipments, the explosion of AIS data has made traffic pattern mining and route planning a new dimension for ensuring the safety of both manned ships and MASS and their hybrid traffic existence.

In recent years, shipping has been undergoing the digital technology revolution under the content of shipping 4.0 and the Maritime Internet of Things (Li et al., 2023a; Aiello et al., 2020). These technological and industrial advancements provide a solid foundation for developing autonomous shipping. The IMO defines MASS development at four levels according to the degree of human interaction to

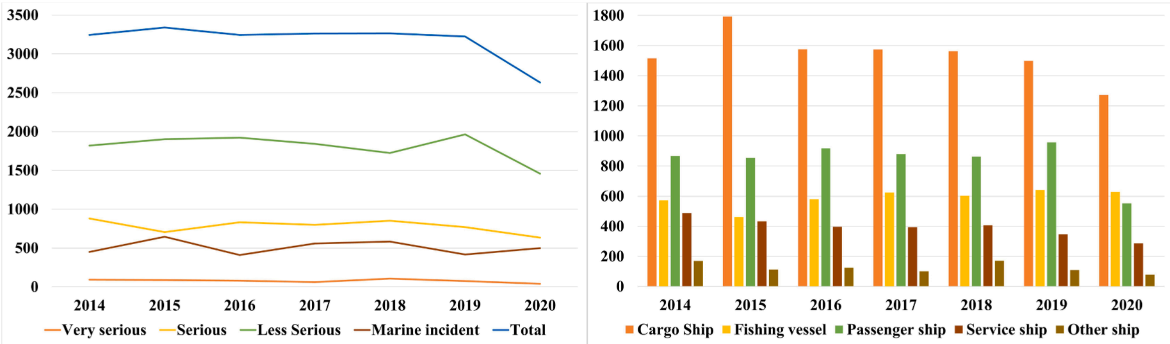

Fig. 1. The information on accidents from 2014 to 2020: the number of marine casualties and accidents (1a) and the distribution of different ship types (1b).

guide the development progress of MASS in the future (IMO, 2018). In industrial research, Rolls-Royce Marine believes that remotely controlled ships will initially appear in local maritime applications and move to the global area in the near future (Gu et al., 2021). Autonomous shipping research has been attracting growing interest and attention. While Mayflower Autonomous Ship (MAS), the world's first fully autonomous unmanned research ship, was launched to conduct scientific research tasks on Sep. 7, 2021. The Atlantic Crossing trip for MAS is planned for Spring 2022 (Mayflower Autonomous Ship, 2021). The critical development of MASS mainly relies on such areas as maritime situational awareness, system control, and independent decision-making to reduce human error-induced accidents and improve navigation safety (Ahvenjärvi, 2016; Burmeister et al., 2014b; Burmeister et al., 2014a). All their success lies in effective and safe route planning as an essential condition. Traditional route planning research methods are qualitative, subjective, and experience-oriented. For example, the route selection for cost and benefit optimization (Wu et al., 2021b), routing and scheduling for fuel supply vessels (Christiansen et al., 2017), routing for liner shipping (Lin and Chang, 2018), route planning for food delivery (Liu et al., 2020), routing for container transportation (Wang et al., 2021b), a location-inventory-routing model for managing supply chains (Liu et al., 2021), routing for shared autonomous electric vehicles (Ge et al., 2021; Zhang et al., 2022a), routing with inventory costs and emissions (Qi et al., 2022). Despite the fact that there are no short of routing studies in transportation, AIS data-driven ship trajectory research for MASS remains scanty. To date, the development of MASS is still at the ship model test and experimental study stages in both commercial and academic worlds (He et al., 2021; Zhang et al., 2021a). The hybrid traffic of unmanned and manned ships will be first appearing at the early stage. However, the real data from MASS is largely unavailable, which limits the development of situational awareness for hybrid traffic. M-ITS has enabled the collection of massive trajectory data from the AIS equipment in shipping. The explosion of AIS data will support the traffic pattern mining of classical manned ships and further aid the maritime situation awareness for MASS. After all, MASS route planning has to use the established shipping routes as the baseline to enable the incorporation of specific MASS features for better route planning in the future when more real-world MASS trajectories become available. The historical AIS data knowledge extraction will provide valuable navigation references for MASS in different

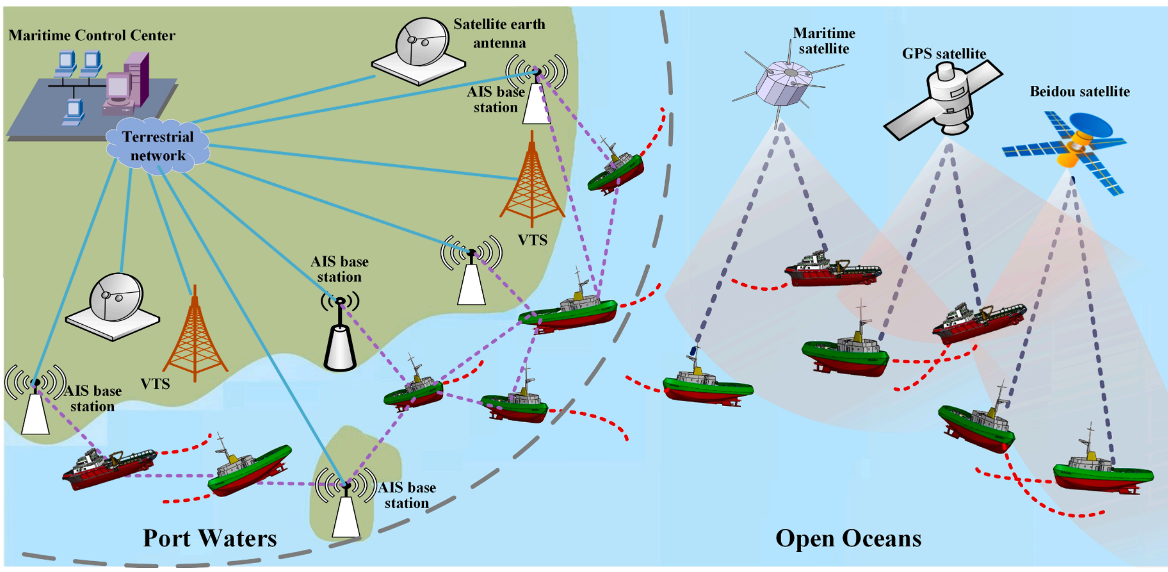

Fig. 2. The global communication network of maritime vessels in M-ITS.

areas. Meanwhile, the movement patterns from historical AIS data can provide effective sailing knowledge, research foundation, and automatic system design for MASS. Moreover, MASS navigation does not require human intervention. Therefore, using ship trajectories and new unsupervised routing algorithms for maritime traffic studies is imperative for the development of MASS.

The existing studies often rely on the optimization methods with specific conditions for tramp and liner ship routing, such as container routing and repositioning (Dong et al., 2015; Song and Dong, 2013, 2012) and tramp ship routing (Meng et al., 2015; Wang and Meng, 2012). It reveals some challenges to applying such methods from the existing literature to formulate the routing of MASS because MASS routing has special needs in the provision of route reference for different types of MASS. The development of data-driven knowledge extraction for an autonomous system looks promising. Moreover, it is challenging to directly conduct AIS data-based MASS route research because there are very few real trajectories by MASS in practice. A complex navigational environment further affects the applications of MASS in practice (Liu et al., 2022a; Öztürk et al., 2022; Shi and Liu, 2020). Meantime, MASS route planning is also related to communication, system control, traffic information exchange, and decision-making between manned and unmanned ships (Gao et al., 2022). This paper aims to conduct a deep analysis and classification exploration of historical data from manned ships to mine the movement patterns and develop a new and rational route planning methodology for MASS safe navigation.

The current route planning has three different definitions: 1) voyage planning (IMO Resolution, 1999), 2) weather routing (IMO Resolution, 1983), and 3) collision avoidance planning (Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs)). Our manuscript falls in the context of collision avoidance planning, especially in the water with traffic separation schemes to aid the realization of mixed traffic of manned ships and MASS in the same areas. Some existing studies (Cai et al., 2021; Chian Tan et al., 2021; Filipiak et al., 2020; Han and Yang, 2020; Jeong et al., 2019) already demonstrate the feasibility of using AIS data for developing MASS route planning. Although showing some attractiveness, these studies still reveal some theoretical challenges which have implications not being well dealt with in the current literature and cannot be easily solved without developing new methods. They are related to the issues of 1) how to make a thorough berth-to-berth guide; 2) how to optimize the routes based on the detailed and up-to-the-minute weather information adaptively; and 3) how to plan safe routes to avoid collisions based on AIS equipment, radar, Electronic Chart Display and Information System (ECDIS), LRIT, and APRA. All these issues become even more worrisome when mixed encounter situations happen in the waters of complex traffic (e.g. ports and canals). This paper attempts to address this research challenge by incorporating historical AIS data of manned ships into unsupervised route planning to guide safe and rational navigation of MASS in complex waters of mixed traffic of manned and unmanned ships. The route planning results can provide an effective solution to the collision avoidance of manned ships and MASS.

The remainder of this paper is organized as follows. The literature review of MASS route planning research is presented in Section 2. Problem statements are listed in Section 3. Section 4 describes the methodology, including the new movement pattern extraction methods and route planning algorithm. The experimental results and analysis are shown in Section 5. Section 6 concludes the paper with future exploration.

2. Literature review

In this section, a systematic review is first conducted to better understand the state of the art of route planning studies in maritime transport to extract the research with a focus on MASS (i.e. Section 2.1) and AIS-based route planning for manned ships and MASS (i.e. Section 2.2). It is followed by the critical analysis of movement pattern extraction methods (i.e. Section 2.3) and route planning methods (i.e. Section 2.4) from a methodological perspective. Finally, the state of the art and our contributions are listed in Section 2.5. The flowchart of the literature review is displayed in Fig. 3.

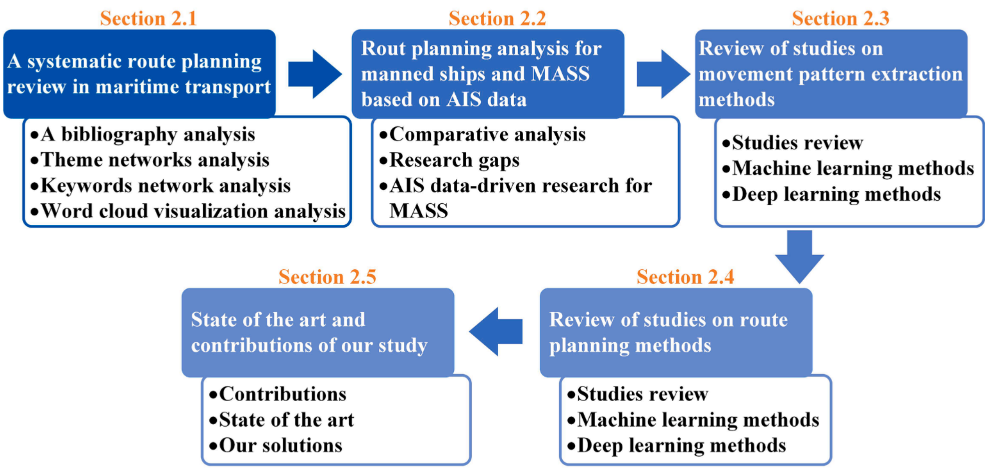

Fig. 3. The flowchart of the literature review.

2.1. Review of studies on route planning in maritime transport

To retrieve all the route planning research in maritime transport, a systematic review was conducted using the Web of Science (WoS) Core Collection in January 2022. The systematic review framework and content are shown in Fig. 4. It is evident that the route

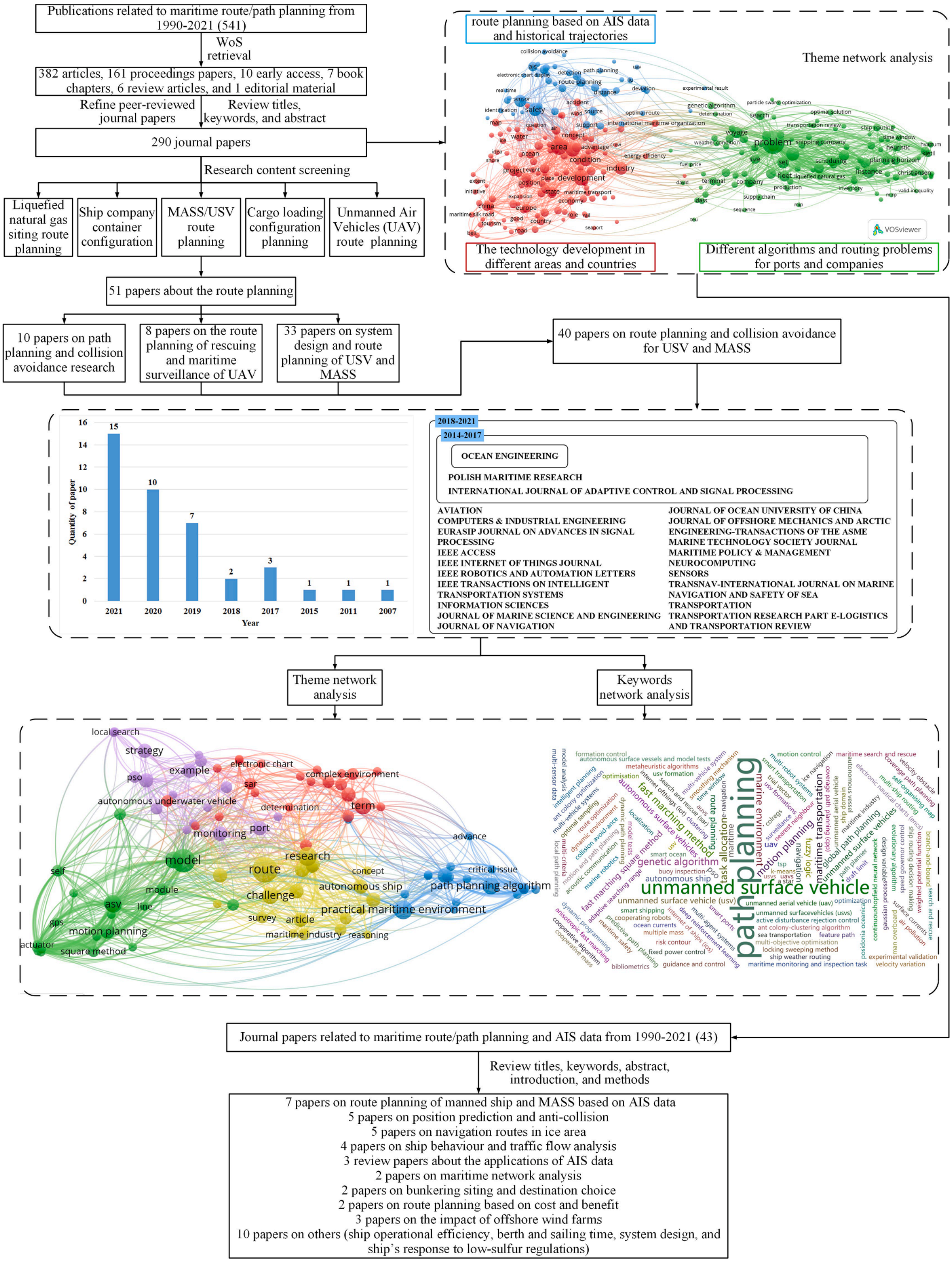

Fig. 4. The systematic review framework, content, and results.

planning for MASS and Unmanned Surface Vehicles (USV) is the foundation to help rescue, explore bathymetric study, and reduce the influence of human factors and carbon emission. A deep analysis of the route and path planning is carried out to explore the research theme and content distribution, especially for MASS and USV. First, it is obvious that the maritime route planning research tendency sharply increased from 2017 to 2021 and arrives at the highest level in 2022. Furthermore, a deeper understanding of the current landscape in different research areas is emphasized. The research focus was the route planning system control in 2014–2017 and collision avoidance and route planning in 2018–2021. Within 2018–2021, route planning based on AIS data has been initiated and gained preliminary developments. While revealing the importance of this research topic, the results clearly show the vast research challenges, such as the applications of historical trajectories in MASS, wanting for solutions to be found.

To assess how the research focus of route planning of MASS and USV have developed and evolved, the theme network, keywords network, and overlay visualization are analyzed using the VOSviewer software. The themes of the retrieval results in titles and abstracts are further analyzed to discover the different theme clusters and allow researchers to gain insights into different research networks. There are three clusters in the visualization result about maritime route planning. The green cluster represents the different algorithms and routing problems for ports and companies, such as the planning problem in supply chains, liquid natural gas, fuel consumption, and cargo planning with high profits and low costs. The red cluster is the technology development in different areas and countries. Risk and technology are the two important themes in this network. The route planning based on AIS data and historical trajectories is shown in the blue cluster, while data and knowledge can provide support for rational decision-making. Collision avoidance is also a small part related to the application of route planning. The line indicates their relationships across different themes. The bubbles show their frequency. The bigger the circles, the more they occur. The three theme clusters from Fig. 4 clearly indicate that the route planning research based on AIS data is far less than the other two parts (e.g. routing problems for ports and companies and the technology development in different areas and countries), despite their fast growth in recent years.

Furthermore, the network visualization of research themes for MASS and USV reveals the five relevant clusters, including 1) route planning based on the dynamic obstacle and case study, 2) the motion planning and control of Autonomous Surface Vehicles (ASV), 3) path planning algorithm, 4) the route planning and challenges analysis, and 5) the optimization methods of route planning in USVs. The word cloud visualization shows the keyword analysis results, such as path planning problems, optimation algorithms, systems, and collision avoidance applications.

Despite its high attractiveness, there is little evidence of the use of AIS data for MASS route planning research in the theme analysis. The vast volume of positioning data contained in AIS makes it impractical to determine routes for MASS and USV in the maritime environment. Moreover, it's indispensable to investigate and extract rational routes by historical AIS data because route planning is dynamic and changeable in MASS navigation under complicated environments and traffic conditions. Given the facts that 1) there is little real data available from MASS trajectories and 2) MASS routing will at large follow the established manned ship trajectories, it is significant and insightful to extract and design the safe routes for MASS based on the historical AIS data from manned ships (Chian Tan et al., 2021; Han and Yang, 2020).

2.2. Review of studies on route planning based on AIS data

To explore the development of maritime route planning based on AIS data, we retrieved all the relevant publications and found 290 journal papers from 1990 to 2021. Through the title, abstract, introduction, and methods screening in Fig. 4, seven papers have shown their direct association with the route planning of manned ships and MASS based on AIS data. The research content and method

Table 1
Ship route planning comparative analysis based on AIS data.

Research paper	Method	Application	Strength	Weakness
(Cai et al., 2021)	Speed-weighted geolocation method, DBSCAN, and Kmeans method	Route selection in tramp shipping	Propose a practical data-driven methodology to select routes	Expert/manual intervention, connection points identification is not accurate
(Filipiak et al., 2020)	Cumulative sum (CUSUM) algorithm, a genetic algorithm with spatial partitioning	Generate maritime traffic network and plan routes	Find the route based on the graph creation	A zig-zag route in historical trajectories
(Guo et al., 2020)	Deep deterministic policy gradient (DDPG) algorithm and improved deep reinforcement learning (DRL) method	Interaction with the environment and historical data	Better learning and fitting ability	Environment information and motion model are ignored, and complex design
(Jeong et al., 2019)	Risk contour map, multi-criteria model	Integrate the safety, efficiency,convenience, and ability of navigation	Propose a multi-criteria route planning technique	Content only for liquefied natural gas (LNG) ship type, subjective factor evaluation
(Han and Yang, 2020)	DBSCAN, central line extraction	Integrate empirical navigation information	Construct ship historical route network topology	Simple patterns, only explore the historical trajectories without different ship types
(Chian Tan et al., 2021)	The nearest neighbor search (1NN), feature representation	Estimate risk based on route and 1NN distance	Select routes for autonomous ships	Insufficient and clear experiments
(Naus, 2020)	GRID reference systems (density, speed, course)	Generate route planning templates	Take into account average speed and course	High computational complexity, big data preprocessing

comparison are listed in Table 1. It can be seen that the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method is the commonly used pattern extraction method. The pattern extraction method can effectively mine the hidden information to aid route planning in both manned and unmanned ships. The comprehensive comparative results in Table 1 show that the automatic route generation methods include 1) electronic charts-based routes, 2) turning points identification-based routes by clustering in historical trajectories, and 3) routes in specific sea areas from historical trajectories. These methods generate ship routes from local features, turning points, and lines based on historical AIS data. From this state-of-the-art analysis, it is evident that MASS route planning has theoretical implications that are not being dealt with holistically in the current literature. It urgently needs powerful algorithms using big AIS data for an overall effective solution that can fit different ship types for MASS.

2.3. Review of studies on movement pattern extraction methods

The patterns are extracted from the historical trajectories by the trajectory data mining methods, including machine learning and deep learning methods.

2.3.1. Movement pattern extraction based on machine learning methods

Machine learning methods are widely used in exploring and mining ships' movement patterns. Data preprocessing, speed, and Gaussian Mixture Model were applied to track the fishing footprints and patterns (Vespe et al., 2016). This method can discover fishing knowledge and avoid steaming errors; however, it only relies on speed anomaly. Principal Component Analysis (PCA) and multi-step trajectory clustering method were put forward to discover movement patterns and discern the course (Li et al., 2017). Li et al. (2018a) developed the DBSCAN method to mine the hidden patterns in historical trajectories based on merge distance. Liu et al. (2019) proposed an Automatic Douglas and Peucker (ADP) compression method to simplify the data set and reserve the critical features, then simplify the data set and extract the critical features. However, the previous three studies are not enough to verify the effectiveness with small datasets, and they are also difficult to be applied to large datasets directly. Wei et al. (2020) combined the clustering algorithm and the sliding window method to mine moving features and spatial information. Zhao and Shi (2019) developed a new clustering method, Density-Based Spatial Clustering of Applications with Noise considering Speed and Direction (DBSCANSD), to extract the distribution information and discern the normal behavior patterns. Li et al. (2020) proposed an Adaptive Constrained Dynamic Time Warping (ACDTW) to calculate the distance more accurately and aid the pattern extraction. A novel trajectory similarity measurement method, hierarchical clustering, and a mixed regression model were combined to discover the hidden information and discover the hidden information (Zhen et al., 2017). A spatio-temporal trajectory clustering method was proposed to discover intricate features and patterns (Nanni and Pedreschi, 2006). A novel Bayesian network learning model was developed to learn the patterns of moving vessels and detect abnormal behavior (Mascaro et al., 2014). Ordering Points To Identify the Clustering Structure (OPTICS), Gaussian process, and graph theory methods were applied to extract ship movement patterns and aid maritime traffic prediction (Rong et al., 2022). Pallotta et al. (2013) proposed a Traffic Route Extraction and Anomaly Detection (TREAD) method to discover traffic knowledge and understand traffic patterns.

From the comparison of the previous studies, it is evident that the similarity measurement and feature extraction methods are the common ones to mine similar patterns and find the optimal path, such as Dynamic Time Warping (DTW) and its improvement methods. Compared with Euclidean Distance (Chan et al., 2003), Hidden Markov Model (Krogh et al., 2001), Hausdorff Distance (Huttenlocher et al., 1993), Fréchet distance (Alt and Godau, 1995), Longest Common Subsequence (Hirschberg, 1977), the strengths of DTW and the improvement methods lie in the capacity to find the shape similarity of the investigated trajectories and warping an optimal route from feature to feature. Meanwhile, the chosen clustering methods are density-based (i.e. DBSCAN and OPTICS) since they do not require the number of cluster centers to be defined. However, the radius and the number contained in the circle should be set in advance. Therefore, the optimization methods of these two parameters (i.e. radius and number) are deemed the main innovations in the improved methods in the current literature.

2.3.2. Movement pattern extraction based on deep learning methods

In recent years, deep learning methods such as Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) have been attracting great attention in maritime safety research. The ship AIS trajectories are first encoded and decoded into feature vectors (e.g. encoder-decoder, AutoEncoder (AE), Variational Auto-Encoder (VAE), and Sparse AutoEncoder (SAE)). Then, different point clustering methods can be conducted to cluster the trajectories and mine useful information. A Seq2Seq learning (t2vec) method was proposed to extract the features and compute the similarity of ship trajectories, and then realize the representations of trajectories (Li et al., 2018b). Yao et al. (2017) combined a sliding window and RNN-based Seq2Seq autoencoder model to achieve deep representations of moving trajectories in a small AIS dataset. Furthermore, they applied a fixed-length sliding window and a sequence-to-sequence auto-encoder to extract moving features and realize the deep representation of the trajectories (Yao et al., 2018). Zhang et al. (2019a) considered an attention mechanism and an auto-encoder model to realize the feature representations of noisy vessel trajectories in a low dimensional space in a small AIS dataset. Zhang et al. (2019b) put forward a new Activity trajectory to a vector (At2vec) model to measure trajectory similarity with semantic information. Liang et al. (2021a) proposed a convolutional auto-encoder (CAE) feature extraction method to calculate the similarity between vessel trajectories accurately and realize the deep representations of trajectories. Deep learning methods have strong learning abilities; however, their poor interpretability and parameter setting are the main problems in the current literature. Overall, it is found that there is no uniform way to set all parameters, and the interpretability of the results cannot be effectively matched to the models.

Although deep learning methods have shown strong feature learning performance, it is still difficult to determine the parameter setting in the learning rate, batch size, gride size, epoch size, iteration number, and loss function, hence leading to limited applications.

2.4. Review of studies on route planning methods

2.4.1. Route planning based on machine learning methods

Heuristic algorithms are the common methods in ship route design and planning, including the genetic algorithm (Zhao et al., 2021), the Dijkstra algorithm (Dijkstra, 1959), the ant colony algorithm (Tsou and Cheng, 2013), the A* algorithm (Wang et al., 2021a), and sampling-based approach (Devaurs et al., 2015). However, the associated local optimal solution problem and high computational cost are often the encountered difficulties influencing their applications. A genetic algorithm and cell-free method were proposed to determine optimal shipping routes and speed simultaneously (Lee et al., 2018). However, it reveals that complex input parameters and settings are difficult to be solved. A three-dimensional Dijkstra algorithm was developed to receive a global optimal ship route, and then reduce fuel consumption (Wang et al., 2019). Despite the advantages, the weight setting of different variables becomes problematic. A route planning method based on A* was proposed to avoid collisions and support autonomous navigation (Larson et al., 2006). A multi-criteria route planning method was applied with an ECDIS to design routes according to the navigational traffic risk and risk contour map (Jeong et al., 2019). It is, however, evident that it is difficult to set the weight of different factors. Dijkstra, the ant colony algorithm, and the DBSCAN method were combined to extract and optimize shipping routes (He et al., 2019). However, the parameter setting and the optimal method selection are identified as the main problems. A rapidly-exploring random tree (Véras et al., 2019) is a common method in sampling-based approaches and can find optimal routes in constructing a graph. It is criticized as always finding non-optimal solutions. A fuzzy genetic algorithm and fuzzy set theory were jointly proposed to aid shipping planning with market demand based on the simulation experiments for container ships (Chuang et al., 2010). However, the relevant parameters are still difficult to set optimally. A route prediction method was combined to extract historical ship routes and predict the future position based on the k-nearest neighbor (KNN) classification (Duca et al., 2017). DBSCAN and grid methods were combined to generate the topological structure and shipping routes from the historical AIS data (Han and Yang, 2020). A regular square reference system (GRID) traffic intensity was proposed to integrate speed and course into route plan templates (Naus, 2020). However, this method has exposed high computational complexity.

2.4.2. Route planning based on deep learning methods

Deep learning methods have gradually emerged to support route planning in recent years. Heuristic planning, CNN, and Long short-term memory (LSTM) are combined in one way or another to optimize route planning and avoid collision (Li and Zheng, 2021). Deep reinforcement learning and artificial potential field are applied to generate safe routes for USV (Li et al., 2021). Xue et al. (2011) proposed a potential field method and dynamic route planning to extract the turning area and avoid collision for autonomous ships. Wen et al. (2020) combined the DBSCAN and Artificial Neural Network (ANN) methods to design the routes for autonomous ships. Wu et al. (2021a) put forward the Visibility Graph (VG) and Local Planning Window (LPW) to improve route accuracy based on small, middle, and large scales. The route planning research shows that deep learning has exposed the difficulties of addressing hyper-parameter selection problems. Its inherent black-box mechanism results in that the results are often unexplainable.

2.5. State of the art and contributions of our study

From both the systematic analysis of MASS routing studies and the critical review of generic routing research using machine learning and deep learning methods, it reveals that it is necessary to develop new research using historical AIS data to address the two major weaknesses (i.e. inference invisible and result unexplainable) when using the deep learning methods in MASS routing. In practice, the development and implementation of MASS require insightful information on the implications of the inference process and the result to be observable to a certain extent. Furthermore, MASS routing has to evolve from a local to global level, starting with its pilot in domestic waters (Cheng and Ouyang, 2021). From this perspective, the MASS route planning for autonomous navigation in a local area will provide useful insights for the global shipping of MASS. This paper uses historical AIS data in a domestic waterway with traffic separation schemes to develop a new optimal function. The model is tested at a local-level navigation situation to demonstrate its feasibility for route planning of MASS. It can be easily popularized as a globally optimal solution when more AIS data becomes available and applicable.

More specifically, following the above critical analysis, our research falls into the route planning for MASS based on AIS data and enriches the existing literature from the following three parts. The state-of-the-art against the three new contributions (i.e. N1-N3) is summarised below.

N1. A new feature measurement method in movement pattern extraction of historical trajectories from manned ships for MASS.

State of the art: The traditional similarity measurement method, DTW, always finds a local optimization and results in over-stretching and over-compression of features during the distance calculation of different trajectories (Li et al., 2020; Liu et al., 2019). The similarities among different trajectories are critical for maritime movement pattern extraction, anti-collision, and route planning. Therefore, it is challenging to address the issue of how to find the optimal global similarity between trajectories.

Our solution: This paper presents a pioneering attempt to propose and apply an Automatic and Adaptive Dynamic Time Warping (AADTW) algorithm to aid situational awareness in modeling the MASS route planning problem. The weight and range are taken into account in the AADTW method to find the global optimization similarity value in the warping routes. The new feature method can extract effective features based on historical AIS data and be used for finding feature centers as route guidance in the route planning of MASS.

N2. An unsupervised clustering method in movement pattern mining of historical trajectories from manned ships for MASS.

State of the art: The traditional point clustering methods cannot aid in achieving any improved performance when handling large-scale ship trajectories (Li et al., 2017; Zhao and Shi, 2019). Furthermore, some parameters need to be set manually, such as density threshold, radius threshold, clustering centers, and the number of clustering centers (Li et al., 2018a; Liang et al., 2021a). Obviously, human intervention is not suitable for the pattern mining of a MASS navigation system. Therefore, the concern on how to eliminate the influence of the setting values of the parameters has not yet been well addressed.

Our solution: An unsupervised pattern extraction method, Spectral Clustering with Affinity Feature (SCAF), is newly proposed to extract the movement patterns and is combined with the AADTW method to make the route planning robustly based on the graph theory and affinity features for supporting MASS. Meanwhile, a novel clustering index function is developed in this process to enable the fast identification of the number of clustering centers to speed up the pattern extraction. The SCAF method can mine the hidden patterns from historical trajectory data, which provides the pattern reference for MASS navigation. The dynamic programming method, AADTW, can further extract the pattern centers for MASS route selection.

N3. The classification pattern mining and feature centers generation for the route planning of MASS with multiple purposes.

State of the art: The current route planning methods, including machine and deep learning methods, are based on different optimization methods to find the optimal routes (Tam et al., 2009; Wang et al., 2019; Wu et al., 2021a). They focus on efficiency-oriented, obstacle avoidance, and the shortest routes to make plans for a specific ship type. However, different ship types are not taken into account in route planning (Feng and Zhu, 2016; Liang et al., 2021b; Sheng et al., 2018). On the other hand, the exposed local optimal solutions and high time complexity problems have not been well addressed, wanting effective solutions to be found.

Our solution: It is among the first investigations to incorporate the historical trajectories of different types of ships to generate routes for MASS with multiple purposes. The main ship types are extracted and classified to analyze the navigation routes separately from historical AIS data to aid the MASS route planning, projecting to support MASS involving various navigation purposes holistically. A new feature center generation method is proposed to find the optimal global route based on the AADTW dynamic programming method for designing routes for MASS in future route planning. It is crucial from a theoretical perspective to incorporate more specific parameters into MASS route planning formation, as MASS route planning needs to be dynamic and optimal subject to different parameters (e.g. ship types and seasons). It pioneers the new incorporation of such influential parameters into MASS route planning.

3. Definitions and problem statements

In this research, the historical AIS data of manned ships is incorporated into unsupervised route planning of MASS to aid the realization of safe and rational navigation in complex waters of hybrid traffic of manned and unmanned ships. The methodology consists of two main integrated parts: a movement pattern extraction method and a route planning method. These two parts are combined to create a holistic framework for solving the research problems defined in Sections 3.1 and 3.2 that are associated with the hybrid traffic of manned and unmanned ships. Fig. 5 illustrates the proposed framework and its corresponding methods through a

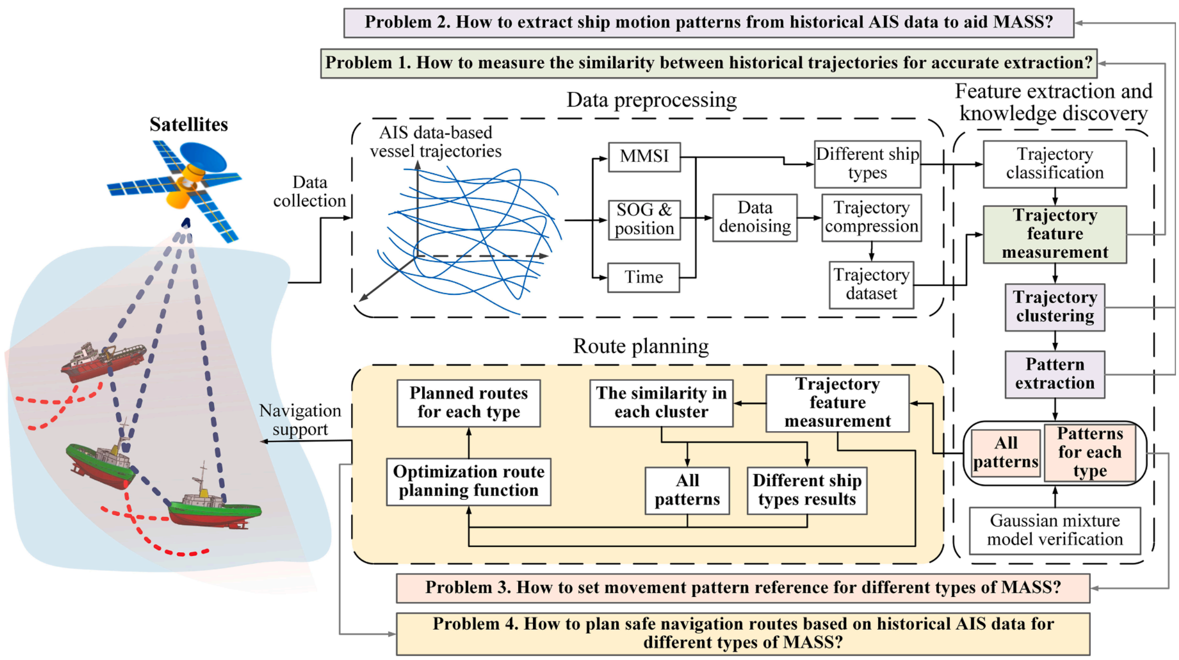

Fig. 5. The research problems and the proposed framework.

flowchart, depicting the flow of information among all the components.

To clarify the content, the notations in Section 3 are summarized in Table 2.

3.1. Definitions

Definition 1. ***Ship trajectory dataset and trajectory.***

Each trajectory includes a series of time series points with time and location (i.e. longitude and latitude) that are collected from AIS equipment. The m-th trajectory Tra_m with N_m points in a ship trajectory dataset Tra is shown below.

$$Tra_m = \left[P_1^m, \cdots, P_i^m, \cdots, P_{N_m}^m\right], \quad i = 1, \cdots, N_m; \quad m = 1, \cdots, M \tag{1}$$

with the i-th point $P_i^m = (t_i^m, lon_i^m, lat_i^m)$.

Then, the m-th ship trajectory can be expressed as

$$Tra_m = \left[\left(t_1^m, lon_1^m, lat_1^m\right), \cdots, \left(t_i^m, lon_i^m, lat_i^m\right), \cdots, \left(t_{N_m}^m, lon_{N_m}^m, lat_{N_m}^m\right)\right] \tag{2}$$

Definition 2. ***Trajectory similarity matrix.***

The similarity between ship trajectories is measured by the proposed AADTW method, and a detailed description of AADTW is introduced in Sec 4.1.2.

The similarity matrix of M trajectories in the whole trajectory dataset can be expressed as

$$AADTW_{M\times M} = \begin{bmatrix} AADTW(Tra_1, Tra_1) & \cdots & AADTW(Tra_1, Tra_n) & \cdots & AADTW(Tra_1, Tra_M) \\ \vdots & & & & \vdots \\ AADTW(Tra_m, Tra_1) & & \ddots & & AADTW(Tra_m, Tra_M) \\ \vdots & & & & \vdots \\ AADTW(Tra_M, Tra_1) & \cdots & AADTW(Tra_M, Tra_i) & \cdots & AADTW(Tra_M, Tra_M) \end{bmatrix} \tag{3}$$

Definition 3. ***Ship movement patterns.***

Different ship types are identified by the MMSI in AIS data, and the main ship types include cargo ships, tankers, and container ships. The SCAF method is proposed to extract the ship movement patterns by historical trajectories from the whole dataset, cargo ships, tankers, and container ships, respectively. Therefore, their extracted movement patterns are shown below.

$$\begin{aligned} MP &= C_1, C_2, \cdots, C_k \\ MP1 &= C_1^1, C_2^1, \cdots, C_{k1}^1, \quad k1 \leqslant k \\ MP2 &= C_1^2, C_2^2, \cdots, C_{k2}^2, \quad k2 \leqslant k \\ MP3 &= C_1^3, C_2^3, \cdots, C_{k3}^3, \quad k3 \leqslant k \end{aligned} \tag{4}$$

The number of clusters (i.e. k, $k1$, $k2$, and $k3$) and the clusters (i.e. $C_1, C_2, \cdots, C_k$, $C_1^1, C_2^1, \cdots, C_{k1}^1$, $C_1^2, C_2^2, \cdots, C_{k2}^2$, and $C_1^3, C_2^3, \cdots, C_{k3}^3$) are generated from the trajectories of the whole dataset, cargo ships, tankers, and container ships based on the SCAF method for the associated route planning accordingly.

Table 2
List of the notations used in Section 3.

Notations	Definition	Notations	Definition
Tra	A ship trajectory dataset	$MP1$	The movement patterns of cargo ships
M	The number of trajectories in the ship trajectory dataset	$MP2$	The movement patterns of tankers
Tra_m,Tra_n	The m-th and n-th trajectories	$MP3$	The movement patterns of container ships
N_m,N_n	The length of the m-th and n-th trajectories (i.e. the number of points in the m-th and n-th trajectories)	k	The number of movement patterns in MP (i.e. the number of cluster centers)
P_i^m	The i-th point of the m-th trajectory	$k1$	The number of movement patterns in $MP1$
t_i^m	The time of the i-th point in the m-th trajectory	$k2$	The number of movement patterns in $MP2$
lon_i^m	The longitude of the i-th point in the m-th trajectory	$k3$	The number of movement patterns in $MP3$
lat_i^m	The latitude of the i-th point in the m-th trajectory	$C_1, C_2, \cdots, C_k$	k movement patterns of the whole dataset
$AADTW(Tra_m, Tra_n)$	The similarity between the m-th trajectory and n-th trajectory (i.e. the minimum value of the warping path based on the proposed AADTW method)	$C_1^1, C_2^1, \cdots, C_{k1}^1$	$k1$ movement patterns of cargo ships
$AADTW_{M\times M}$	The similarity matrix of M trajectories in the whole trajectory dataset	$C_1^2, C_2^2, \cdots, C_{k2}^2$	$k2$ movement patterns of tankers
MP	The movement patterns of the whole trajectory dataset	$C_1^3, C_2^3, \cdots, C_{k3}^3$	$k3$ movement patterns of container ships

Table 3
List of the notations for the solutions in Section 4.

Notations	Definition	Notations	Definition
$P_j^n = (t_j^n, lon_j^n, lat_j^n)$	The j-th point of the n-th trajectory	$b(x)$	The separation
$a_{i,\ j} = d(P_i^m, P_j^n)$	The Euclidean distance between points P_i^m and P_j^n	$Tr(B_{k_c})$	The trace of the between-cluster scatter matrix
$A_{N_m \times N_n} = \{a_{i,\ j}\}$	The distance between all points in the m-th and n-th trajectories	$Tr(W_{k_c})$	The trace of the within-cluster scatter matrix
$W = \{W_1, \cdots, W_K\}$	A warping path with the length K	$vol(C_l)$	The total degree of the vertices in the cluster C_l
$DTW(Tra_m, Tra_n)$	The minimum value of the warping path based on the DTW method	$Ncut(C_1, C_2, \cdots, C_k)$	The Normalized Cut
$D(i,j)$	The distance from the previous step to step (i, j) by the DTW method (i.e. i denotes the i-th point in the m-th trajectory and j indicates the j-th point in the n-th trajectory)	$AD(i,j)$	The distance from the previous step to step (i, j) by the proposed AADTW method (i.e. i denotes the i-th point in the m-th trajectory and j indicates the j-th point in the n-th trajectory)
$N(i,j)$	The usage counts of one step (i, j) in a warping process (i.e. the i-th point in the m-th trajectory and j-th point in the n-th trajectory)	$w_{i,j}$	The dynamic weight in one step (i, j) of a warping process (i.e. the i-th point in the m-th trajectory and j-th point in the n-th trajectory)
L	The Laplacian matrix	Q	The similarity matrix
D	The degree matrix	L_{norm}	The Laplacian regularization
T_{N_m}, T_{N_n}	The usage counts of points in the trajectories Tra_m and Tra_n	H	The membership matrix
C^{k_c}	The current cluster	$V = \{V_1, \cdots, V_M\}$	Vertex (i.e. the points mapped by the trajectories)
$\lvert C^{k_c} \rvert$	The number of the current cluster	E	Edge (i.e. the distance between points)
C_h, C_l	The h-th and l-th clusters	$\mu_r, r = 1, \cdots, k$	The clustering centres
$\lvert C_h \rvert$	The number of the cluster C_h	$R(V_p, \mu_r)$	The responsibility matrix
$\overline{C}_l$	The complement of C_l (i.e. the other clusters except for the l-th cluster)	$A(V_p, \mu_r)$	The availability matrix
x, x', x''	Sample points in the mapping point dataset	$R_1, R_2, \cdots, R_k$	The optimal routes of the whole dataset
$\overline{c}$	The mean of the target sample set	$R_1, R_2, \cdots, R_{k1}$	The optimal routes of cargo ships
$\overline{c}_o$	The mean of the other clusters	$R_1, R_2, \cdots, R_{k2}$	The optimal routes of tankers
$a(x)$	The cohesion	$R_1, R_2, \cdots, R_{k3}$	The optimal routes of container ships

3.2. Problem statements

As shown in Fig. 5, Problem 1 aims to solve the trajectory feature measurement and calculate accurate similarity values, which serves as the foundation of the following three problems. Therefore, the accuracy of feature measurement can determine the performance of knowledge discovery and route planning. Problem 2 involves pattern extraction based on the proposed SCAF method to obtain effective patterns from historical AIS data. It can provide the pattern library for the entire dataset to aid the pattern mining of ship classification in Problem 3. Problem 3 is to design routes for different types of MASS based on the results from Problem 2. Lastly, Problem 4 is the safe navigation route reference for different types of MASS based on the results from Problem 3. These four research problems are interconnected and indispensable.

Problem 1. ***How to measure the similarity between historical trajectories for accurate feature extraction*?**

The similarity between historical trajectories is measured by their distances. The larger the distance between two trajectories, the smaller their similarity, and vice versa. A similarity measurement method, AADTW, is proposed to calculate the distance between trajectories. The solving process involves finding the optimal warping route between any two trajectories based on dynamic programming. The optimal warping route is the most similar correspondence of points between the two trajectories, i.e. the shortest route without over-stretching and over-compression.

The challenge of feature extraction is to identify useful features that can aid in the subsequent pattern extraction contents. Given that pattern extraction relies on the similarity (i.e. distance) between trajectories, accurate measurement of the similarity between trajectories is crucial for feature extraction. Furthermore, the similarities between all trajectories (i.e. the optimal similarity value $AADTW(Tra_m, Tra_n),\ m = 1, \cdots, M, n = 1, \cdots, M$) are taken as an effective basis for ship movement pattern extraction. Therefore, the feature extraction is in nature to solve the optimal similarity value $AADTW(Tra_m, Tra_n),\ m = 1, \cdots, M, n = 1, \cdots, M$ between the m-th trajectory Tra_m and the n-th trajectory Tra_n.

Problem 2. ***How to extract ship movement patterns from historical AIS data to aid the safe navigation of MASS*?**

There is a lack of real trajectory data of autonomous ships to aid the route planning research for MASS in the current literature. The pattern extraction of historical AIS data from manned ships will have to be used to provide the baseline support for MASS (Chen et al., 2021). With the growth of MASS, more trajectory data involving MASS could be obtained and then used to support the updated analysis of MASS route planning. MASS should be categorized based on the cargo they carry, hence following the same/similar routes of the same types of manned cargo ships currently used. Consequently, it is beneficial to analyze the clustering results based on the SCAF method for the whole fleets and different ship types.

The pattern extraction problem is to solve the number of clusters k and the optimal clusters $C_1, C_2, \cdots, C_k$ from the historical AIS data based on the SCAF method as the pattern library.

Problem 3. ***How to set the movement pattern reference for different types of MASS*?**

Different types of ships have significant diversity in their navigational characteristics and patterns. It is therefore essential to classify the historical trajectory data to design the routes for different types of MASS. The trajectories of cargo ships, tankers, and container ships are identified using MMSI from the historical AIS data, which serves as the datasets. The navigation route reference problem of different types of MASS is to mine the patterns $C_1^1, C_2^1, \cdots, C_{k1}^1$ for cargo ships, $C_1^2, C_2^2, \cdots, C_{k2}^2$ for tankers, and $C_1^3, C_2^3, \cdots, C_{k3}^3$ for container ships based on the SCAF method, respectively. Meantime, the number of clusters $k1$, $k2$, and $k3$ can be calculated in the solving process.

Problem 4. ***How to plan safe navigation routes based on historical AIS data for different types of MASS*?**

The configured route that contains the maximum similarity among the trajectories has been used to indicate an acceptable safe route (Sarraf and McGuire, 2020; Yu et al., 2023). Within this context, the representation feature centers and route generation algorithm are proposed to plan the safe routes for different types of ships based on the feature extraction method and dynamic programming. The safe navigation route planning problem is to excavate safe routes with maximum similarity among trajectories in each pattern of $C_1, C_2, \cdots, C_k$ as the representation feature routes in the investigated water area. In particular, the planned representation feature routes for cargo ships, tankers, and container ships are solved from the patterns $C_1^1, C_2^1, \cdots, C_{k1}^1$, patterns $C_1^2, C_2^2, \cdots, C_{k2}^2$, and patterns $C_1^3, C_2^3, \cdots, C_{k3}^3$ based on the route generation method, respectively.

4. Methodology

This section outlines a new methodology of integrating a new movement pattern extraction method and a machine learning-based route planning using a global optimization method for MASS route planning. As shown in Fig. 5, the movement pattern extraction method includes data preprocessing, a new feature measurement algorithm, a novel trajectory clustering algorithm, and trajectory classification to discover ship moving patterns. The new route planning method contains a route optimization method and a feature center generation algorithm to automatically make the route planning for MASS. Along with the notations used for problem definitions in Section 3, this section employs new notations for the solutions listed in Table 3.

4.1. The movement pattern extraction method

A new movement pattern extraction method has been developed, leveraging exploratory data analysis and innovative data measurement from an unsupervised perspective. It depends on a feature extraction method and a novel trajectory clustering method. The feature extraction method enables the development of the trajectory similarity measurement to increase the similarity and enlarge dissimilarity. The proposed AADTW method is used to alleviate the pathological correspondence between trajectories and accurately calculate the distance between trajectories. Furthermore, the SCAF method is put forward to extract the movement patterns automatically based on the trajectory features. The SCAF algorithm integrates a new clustering index function and the affinity feature into an improved spectral clustering method to achieve unsupervised pattern extraction with better performance.

4.1.1. The original feature extraction method

The original DTW can minimize the cumulative distance between two trajectories with local optimization. The dynamic warping route of the trajectories Tra_m and Tra_n is selected from the distance matrix $A_{N_m \times N_n}$, and the distance between points P_i^m and P_j^n is $d(P_i^m, P_j^n) = \sqrt{(P_i^m - P_j^n)^2} \in A_{N_m \times N_n} = \{a_{i,\ j}\}$. Then the distance between all points in the two trajectories Tra_m and Tra_n is expressed as

$$A_{N_m \times N_n} = \begin{bmatrix} a_{1,\ 1} & a_{1,\ 2} & \cdots & a_{1,\ (N_n-1)} & a_{1,\ N_n} \\ a_{2,\ 1} & a_{2,\ 2} & \cdots & a_{2,\ (N_n-1)} & a_{2,\ N_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N_m-1,\ 1} & a_{N_m-1,\ 2} & \cdots & a_{(N_m-1),\ (N_n-1)} & a_{(N_m-1),\ N_n} \\ a_{N_m,\ 1} & a_{N_m,\ 2} & \cdots & a_{N_m,\ (N_n-1)} & a_{N_m,\ N_n} \end{bmatrix} = \begin{bmatrix} d(P_1^m, P_1^n) & d(P_1^m, P_2^n) & \cdots & d(P_1^m, P_{N_n-1}^n) & d(P_1^m, P_{N_n}^n) \\ d(P_2^m, P_1^n) & d(P_2^m, P_2^n) & \cdots & d(P_2^m, P_{N_n-1}^n) & d(P_2^m, P_{N_n}^n) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ d(P_{N_m-1}^m, P_1^n) & d(P_{N_m-1}^m, P_2^n) & \cdots & d(P_{N_m-1}^m, P_{N_n-1}^n) & d(P_{N_m-1}^m, P_{N_n}^n) \\ d(P_{N_m}^m, P_1^n) & d(P_{N_m}^m, P_2^n) & \cdots & d(P_{N_m}^m, P_{N_n-1}^n) & d(P_{N_m}^m, P_{N_n}^n) \end{bmatrix} \tag{5}$$

The essence of the DTW is to calculate the distance matrix between two trajectories to find the optimal warping path. The warping path $W = \{W_1, \cdots, W_t, \cdots, W_K\}$ with K denotes the length of the warping path and the normalizing factor (Kwong et al., 1998; Rashidi et al., 2015), $\max\{N_m, N_n\} \langle K \leqslant N_m + N_n - 1$.

Every warping path has to meet the following conditions (Keogh and Ratanamahatana, 2005; Petitjean et al., 2011):

(1) Boundary condition: $W_1 = a_{1,\ 1}, W_K = a_{N_m,\ N_n}$;

(2) Continuity: if $W_{t-1} = a_{i',\ j'}, W_t = a_{i,\ j}$, then $i - i' \leqslant 1,\ j - j' \leqslant 1$;

(3) Monotonicity: if $W_{t-1} = a_{i',\ j'}, W_t = a_{i,\ j}$, then $i - i' \geqslant 0,\ j - j' \geqslant 0$, the time at each point is also monotonic in W.

The schematic diagram of the distance matrix for trajectories Tra_m and Tra_n is displayed in Fig. 6 (a), while the green route is the optimal warping route received by the DTW method. Take the largest and smallest K as examples to explain the warping path W in Fig. 6 (b). The light blue route is the warping path with the minimum K when $N_m > N_n$. The warping path with the length $K = N_m$ is $W = \{a_{1,\ 1}, a_{2,\ 1}, \cdots, a_{(N_m-N_n-1),\ 1}, a_{(N_m-N_n),\ 1}, \cdots, a_{(N_m-3),\ (N_n-3)}, a_{(N_m-2),\ (N_n-2)}, a_{(N_m-1),\ (N_n-1)}, a_{N_m,\ N_n}\}$. While the dark blue route is the warping path with the maximum $K = N_m + N_n - 1, N_m > N_n$ and the warping path is $W = \{a_{1,\ 1}, a_{1,\ 2}, \cdots, a_{1,\ (N_n-1)}, a_{1,\ N_n}, a_{2,\ N_n}, \cdots, a_{(N_m-2),\ N_n}, a_{(N_m-1),\ N_n}, a_{N_m,\ N_n}\}$.

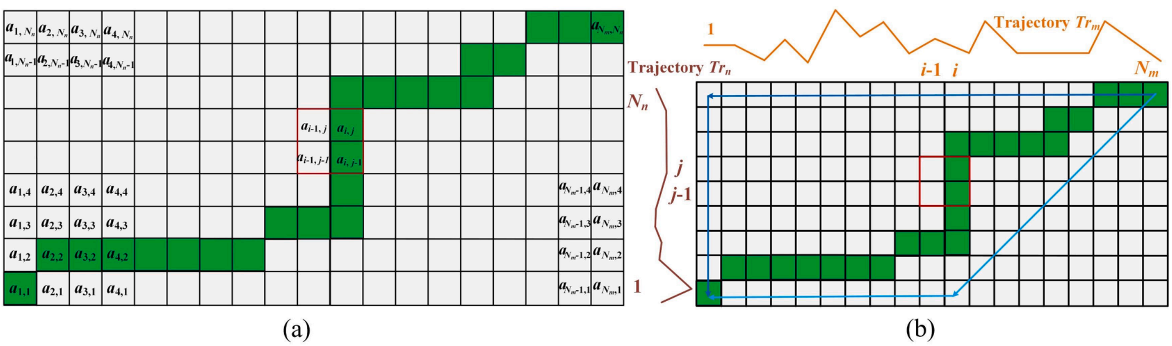

Fig. 6. The schematic diagram of the DTW algorithm, (a) the schematic diagram of the distance matrix for trajectories Tra_m and Tra_n, and (b) the schematic diagram of the warping paths based on the DTW method.

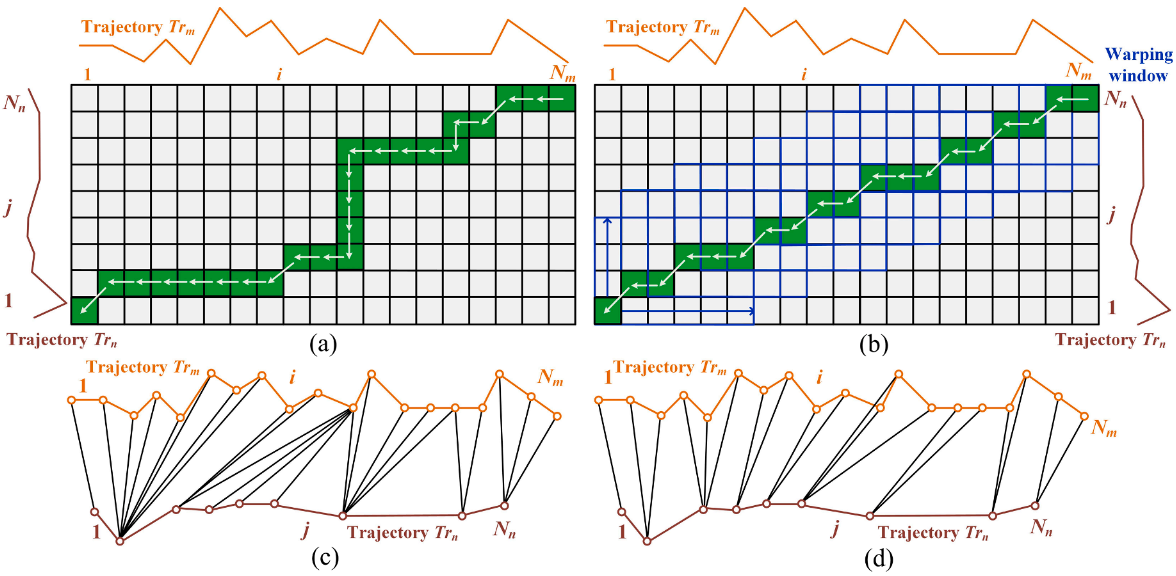

Fig. 7. The schematic diagram and results of the DTW and proposed AADTW methods, (a) the warping path of DTW, (b) the warping path of AADTW, (c) the corresponding result based on DTW, and (d) the corresponding result based on AADTW.

The DTW method aims to find the optimal warping path, i.e. the minimum warping path. The path with the lowest warping cost in the original DTW (Berndt and Clifford, 1994) is given by

$$DTW(Tra_m, Tra_n) = \min\left\{\frac{1}{K}\sum_{t=1}^{K} W_t\right\} \tag{6}$$

s.t.

$$D(1,1) = d\left(P_1^m, P_1^n\right) \tag{7}$$

$$\begin{gathered} D(i,j) = d\left(P_i^m, P_j^n\right) + \min\{D(i,j-1), D(i-1,j-1), D(i-1,j)\} \\ i \geqslant 2, \quad i = 1,\cdots,N_m; \quad j \geqslant 2, j = 1,\cdots,N_n \end{gathered} \tag{8}$$

The objective function (i.e. Eq. (6)) minimizes the warping route, where $\sum_{t=1}^{K} W_t$ is the sum of the elements in the warping route. Constraint (7) mandates that to determine the correspondence between the points, the warping path needs to commence from the first point of the two trajectories. Constraint (8) indicates that there are three different choices from the previous step (i.e. $(i, j-1)$, $(i\text{-}1, j-1)$, and $(i\text{-}1, j)$) to reach the point (i, j). The range of i and j implies that once the mapping relationship between the initial points of the two trajectories has been established, the subsequent corresponding relationships are determined by starting from the second point, and the minimum distance is guaranteed.

4.1.2. A new feature extraction method

To find the optimal global similarity between trajectories and avoid their overstretch, a new AADTW method is developed based on the optimal warping routes and the dynamic weights to different steps. The proposed AADTW method includes two innovative features: limiting the searching range and implementing dynamic weights based on the usage counts of each point.

Given the usage counts of each point in a warping process $N(i,j)$, $i = 1,\cdots,N_m;\quad j = 1,\cdots,N_n$ and the dynamic weight $w_{i,j} = \frac{N_m+N_n}{2\max(N_m,N_n)} \bullet N(i,j)$, which can be adaptively controlled based on a cumulative usage factor. The problem of the optimal global warping path based on the proposed AADTW can therefore be formulated by a mathematical model as follows:

$$AADTW(Tra_m, Tra_n) = \min\left\{\frac{1}{K}\sum_{t=1}^{K} W_t\right\} \tag{9}$$

s.t.

$$AD(1,1) = d\left(P_1^m, P_1^n\right) \tag{10}$$

$$AD(i,1) = AD(1,j) = +\infty \tag{11}$$

$$AD(i,j) = \min\begin{cases} (1+w_{i-1,j}) \bullet d\left(P_i^m, P_j^n\right) + AD(i-1,j), \\ d\left(P_i^m, P_j^n\right) + AD(i-1,j-1), \\ (1+w_{i,j-1}) \bullet d\left(P_i^m, P_j^n\right) + AD(i,j-1), \end{cases} \tag{12}$$

$$w_{i,j} = \frac{N_m + N_n}{2\max(N_m, N_n)} \bullet N(i,j) \tag{13}$$

$$\begin{aligned} &N(i,j) \leqslant N\left(i+\frac{N_m}{3}, j\right), \quad N(i,j) \leqslant N\left(i, j+\frac{N_n}{3}\right) \\ &\quad i \geqslant 2, \quad i = 1, \cdots, N_m; \quad j \geqslant 2, j = 1, \cdots, N_n \\ &\quad\quad m = 1, \cdots, M; n = 1, \cdots, M \end{aligned} \tag{14}$$

The objective function (i.e. Eq. (9)) minimizes the warping route, where $\sum_{t=1}^{K} W_t$ is the sum of the elements in the warping route. Similar to constraint (7), constraint (10) also denotes that the warping path needs to start from the first point of the two trajectories. Constraint (11) indicates the avoidance of all points in one trajectory corresponding to a point in another trajectory. Constraint (12) shows that there are three different choices from the previous step (i.e. $(i, j-1)$, $(i\text{-}1, j-1)$, and $(i\text{-}1, j)$) to step (i, j) in the proposed AADTW method. As seen in Eq. (12), the j-th point in the Tra_n is used again if the path is from $(i-1, j)$ to (i, j), then the weight $w_{i-1,j}$ will increase to constrain the overstretching. Similarly, the i-th point in the Tra_m is reused if the path is from $(i, j-1)$ to (i, j), then the weight $w_{i,j-1}$ will increase to constrain the overstretching. There is no weight if the path is from $(i\text{-}1, j-1)$ to (i, j). Constraint (13) expresses the dynamic weights based on the usage counts of each point and the length of two trajectories Tra_m and Tra_n. Constraint (14) denotes the limitation of usage counts of each point, defined by $\frac{N_m}{3}$ and $\frac{N_n}{3}$, representing the length and width of the search space. i and j indicate the range of points in two trajectories Tra_m and Tra_n, respectively.

The distance between any two trajectories can be calculated by Eq. (9) to form the final similarity matrix of M trajectories (i.e. $AADTW_{M\times M}$) in the whole dataset. The AADTW method is to find the most similar route and eliminate the overstretching within a certain boundary by Eq. (14).

The pseudocode of the AADTW algorithm is shown in Algorithm 1.

Algorithm 1: AADTW algorithm.

Input: The trajectories Tra_m with length N_m and Tra_n with length N_n.
Output: $AADTW(Tra_m, Tra_n)$.
1. Initialize: Zero matrices T_{N_m} and T_{N_n} denote the usage counts of points in the trajectories Tra_m and Tra_n.
$T_{N_m}(1,1) = T_{N_n}(1,1) = 1$ and $T_{N_m}(i,j) = T_{N_n}(i,j) = 0$.
$AD(1,1) = d(P_1^m, P_1^n)$ indicates the calculation process starts at the first point in trajectories Tr_m and Tr_n.
2. $AD(i,1) = AD(1,j) = +\infty$.
3. ***for*** i=1: N_m **do**
4. ***for*** j=1: N_n **do**
5. $AD1 = d_{ij} + w_{i-1,j} \bullet d_{ij} + AD(i-1,j)$;
6. $AD2 = d_{ij} + AD(i-1,j-1)$;
7. $AD3 = d_{ij} + w_{i,j-1} \bullet d_{ij} + AD(i,j-1)$;
8. $AD(i,j) = \min(AD1, AD2, AD3)$;
9. $W(i,j) = min\_index[(i-1,j),(i-1,j-1),(i,j-1)]$;
10. $N(i,j) \leqslant N(i+\frac{N_m}{3}, j)$, $N(i,j) \leqslant N\left(i, j+\frac{N_n}{3}\right)$;
11. ***if*** $AD(i,j) == AD1$;
12 then $T_{N_m}(i,j) = 1$, $T_{N_n}(i,j) = T_{N_n}(i-1,j) + 1$;
13. ***else if*** $AD(i,j) == AD2$;
14. then $T_{N_m}(i,j) = 1$, $T_{N_n}(i,j) = 1$;
15. ***else*** $AD(i,j) == AD3$;
16. then $T_{N_m}(i,j) = T_{N_m}(i,j-1) + 1$, $T_{N_n}(i,j) = 1$.
17. $AD_{Tr_m,Tr_n} = AD(N_m, N_n)$.
18. **end** ***if***
19. **end** ***for***
20. **end** ***for***

Fig. 7 illustrates the schematic diagrams and corresponding results of the DTW and AADTW methods, both of which aim to identify the optimal warping path with the lowest warping cost. The warping cost is inversely proportional to the similarity between the two trajectories. The warping window, represented by the blue box in Fig. 7, is computed using Eq. (14) to restrict the scope of optimal feature extraction. The green squares in Fig. 7 (a) and (b) represent the optimal warping path and the best features between two trajectories identified by the traditional DTW and the newly proposed AADTW methods, respectively. The globally optimal path is taken as the sum of all green routes. The correspondence between the two trajectories is depicted in Fig. 7 (c) and (d). The results of the comparison in Fig. 7 (c) and (d) clearly indicate that the proposed AADTW method does not suffer from the issue of overstretching that is present in the original DTW method.

4.1.3. The proposed feature identification method

A new unsupervised SCAF algorithm is designed for conducting trajectory clustering and support pattern extraction. It is unsuitable to use traditional point clustering methods such as K-means, hierarchical-based, and density-based methods to directly undertake trajectory clustering due to the large data volume and lack of global features (Li et al., 2017, 2018a). To achieve accurate features and mine the complete patterns of ship trajectories, an unsupervised SCAF algorithm is developed in this paper by integrating two parts: a new clustering index function and an improved spectral clustering with affinity features. The proposed clustering index function can automatically receive the number of clustering centers in the process of graph optimization. Then, the SCAF, integrating the essence of affinity propagation, can automatically be applied to extract the movement patterns. The unsupervised SCAF algorithm can be used for the situational awareness and route planning system design of MASS.

Traditional index functions are used to evaluate the clustering performance separately. For instance, the Silhouette Coefficient (SC) index (Layton et al., 2013) can help measure the similarity between the clusters by comparing the degree of compactness and separation for different clusters, while the Calinski-Harabasz Score (CHS) index (Chen et al., 2022) can compare the similarity of within-cluster and between-cluster. The larger these two indexes, the better the clustering performance. However, no single index can capture all aspects of clustering quality, and a combination of multiple indexes can therefore provide more information in evaluating clustering performance (Jaskowiak et al., 2016).

To address this challenge, a new optimal function is formulated by combining the maximum similarity ratio of the within-cluster and the between-cluster, as well as the maximum ratio of the between-cluster scatter and within-cluster scatter. This hybrid model provides a more accurate and comprehensive measure of the clustering quality than the case of using individual measures, and can help researchers determine the appropriate number of clusters for their data. A new clustering index function is developed based on the degree of compactness and separation and the distance within and between clusters. The larger the new index, the better the clustering performance.

The ship trajectories can be mapped into points based on the graph theory. Then, the distance between the points can be measured by the similarity matrix between trajectories. More specifically, the hybrid model of combining the maximum similarity ratio of the within-cluster and the between-cluster and the maximum ratio of the between-cluster scatter and within-cluster scatter is formulated as the new optimal function to calculate the number of clustering centres. The optimization goal is formulated as

$$f(k) = \max_{k} \sum_{x=1}^{|C_h|} s(x) + \max_{k} \frac{Tr(B_{k_c})}{Tr(W_{k_c})} \times \frac{(k - k_c)}{(k_c - 1)} \tag{15}$$

s.t.

$$\sum_{x=1}^{|C_h|} s(x) = \sum_{x=1}^{|C_h|} \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} = \begin{cases} 1 - \dfrac{a(x)}{b(x)}, a(x) < b(x) \\ 0 \\ \dfrac{b(x)}{a(x)} - 1, a(x) > b(x) \end{cases} \tag{16}$$

$$a(x) = \frac{\sum_{x' \in C_{k_c}, x \neq x'} d(x, x')}{|C_{k_c}| - 1} \tag{17}$$

$$b(x) = \min_{1 \leqslant h \leqslant k, h \neq k_c} \frac{\sum_{x'' \in C_h} d(x, x'')}{|C_h|}, h = 1, \cdots, k \tag{18}$$

$$Tr(B_{k_c}) = \sum_{o=1}^{k_c} \|\bar{c}_o - \bar{c}\|^2 \tag{19}$$

$$Tr(W_{k_c}) = \sum_{o=1}^{k_c} \sum \|x - \bar{c}_o\|^{2^2}, o = 1, \cdots, k_c \tag{20}$$

The objective function (i.e. Eq. (15)) indicates the result of maximizing intra-class similarity and inter-class dispersion. $a(x)$ in constraint (17) expresses the average distances between a sample x and all the other samples x' in the same cluster. $b(x)$ in constraint (18) shows the minimum average distances between a sample x and all the other samples x'' in other clusters. $Tr(B_{k_c})$ and $Tr(W_{k_c})$ in constraints (19) and (20) denote the trace of the between-cluster and the within-cluster scatter matrix, respectively.

The number of clustering centers is determined by the new clustering index function from Eq. (15). The proposed SCAF algorithm and the solution method are presented as follows.

Problem modeling. Given a ship trajectory dataset $Tra = \{Tra_1, \cdots, Tra_M\}$, the result of the proposed SCAF is to solve k clusters C_1, $C_2, \cdots, C_k$, where $C_1 \cup C_2 \cup \cdots \cup C_k = V$, $C_1 \cap C_2 \cap \cdots \cap C_k = \varnothing$ (Li et al., 2022). The trajectories are mapped into vertexes based on the graph theory and the similarity matrix. Then the ship trajectory dataset Tra is converted into the point dataset $V = \{V_1, \cdots V_p, \cdots, V_M\}$, $p = 1, \cdots, M$ with the edge weight $E = ACDTW(Tr_m, Tr_n)$. The dataset is formed into an undirected graph $G(V, E)$ based on the result of the AADTW method.

Suppose the clustering centers are $\mu_r, r = 1, \cdots, k$. The responsibility matrix $R(V_p, \mu_r)$ is a square matrix that represents the responsibility of each data point in a cluster to serve as an exemplar for another data point. It measures the availability of a candidate exemplar to be selected by other data points. The responsibility matrix can update the current estimates of the exemplars during each iteration of the algorithm (Sun and Guo, 2014). The availability matrix, $A(V_p,\mu_r)$, is another square matrix that measures the suitability of each candidate exemplar to serve as an exemplar for other data points. It represents the ability of a candidate exemplar to be selected by other data points as their exemplar. The availability matrix is utilized to update the current estimates of the cluster centers during each iteration of the algorithm (Sun et al., 2017). In essence, the responsibility matrix and the availability matrix work together to iteratively update the estimates of the exemplars and the cluster centers until the convergence in the original affinity propagation algorithm is reached.

The edge weight of the m-th trajectory Tra_m and n-th trajectory Tra_n is $AADTW(Tra_m, Tra_n)$, and the graph cut (i.e. the sum of connection weights) is calculated by Eq. (21),

$$Cut(C_l, \overline{C_l}) = \frac{1}{2} \sum_{m \in C_l, n \notin C_l} AADTW(Tra_m, Tra_n), \qquad l = 1, \cdots, k \tag{21}$$

The total degrees of V in the cluster C_l is $vol(C_l) = \sum_{V_m \in C_l} AADTW(Tr_m, Tr_m)$, and the Normalized Cut is $Ncut(C_1, C_2, \cdots, C_k) = Tr(Y^T LY)$ (Von Luxburg, 2007). Therefore, the goal of the SCAF clustering method is to solve the Normalized Cut objective function through Eq. (22):

$$Ncut(C_1, C_2, \cdots, C_k) = \sum_{l=1}^{k} \frac{Cut(C_l, \overline{C_l})}{vol(C_l)} = \sum_{l=1}^{k} \frac{y_l^T L y_l}{y_l^T D y_l} = Tr\left(Y^T LY\right) \tag{22}$$

Problem-solving. The objective function in Eq. (22) can be converted into

$$\min_{Y^T DY = I} Tr\left(Y^T LY\right) \tag{23}$$

Suppose the similarity matrix is $Q = AADTW_{M \times M}$ and the degree matrix is $D = diag(D_{11}, D_{22}, \ldots, D_{MM})$, then the Laplacian matrix is $L = D - Q$. Let us normalize the Laplacian matrix, and the property of the Laplacian matrix is obtained as follows.

$$L_{norm} = D^{-1/2} L D^{-1/2} \tag{24}$$

with $H = D^{1/2} Y$, and then $Y = D^{-1/2} H$. Therefore, the problem in Eq. (23) can be rewritten based on the graph Laplacian.

$$\begin{aligned} &\min_{H \in \mathbb{R}^{m \times k}} Tr\left(H^T L_{norm} H\right) \\ &\quad s.t.\ H^T H = I_k \end{aligned} \tag{25}$$

Then the first k smallest eigenvalues of the normalized Laplacian matrix L_{norm} can be calculated as follows.

$$\begin{aligned} L_{norm} H_l &= \lambda H_l \\ D^{-1/2} L D^{-1/2} H_l &= \lambda H_l \\ D^{-1/2} L D^{-1/2} D^{1/2} Y_l &= \lambda D^{1/2} Y_l \\ D^{-1} L &= Y_l \end{aligned} \tag{26}$$

Finally, Y has the k eigenvectors of $D^{-1}L$ corresponding to its k smallest eigenvalues, while k is solved by Eq. (15).

While the similarity matrix is $Q(V_m, V_n) = AADTW(Tra_m, Tra_n)$ and the cluster number is k, the clustering centers are $\mu_r, r = 1, \cdots, k$.

Initialize the responsibility matrix and availability matrix $R(V_p, \mu_r) = A(V_p, \mu_r) = 0$. $R(V_p, \mu_r)$ and $A(V_p, \mu_r)$ can be iteratively calculated based on each row in Y and the similarity matrix W.

$$\begin{aligned} &R\left(V_p, \mu_r\right) \leftarrow s\left(V_p, \mu_r\right) - \max_{r' \neq k}\left(s\left(V_p, \mu_{r'}\right) + A\left(V_p, \mu_{r'}\right)\right), \\ &A\left(V_p, \mu_r\right) \leftarrow \min\left\{0, R(V_{r'}, \mu_{r'}) + \sum_{n' \neq m,n} \max(0, R(V_i, \mu_{r'}))\right\} \end{aligned} \tag{27}$$

With $p = 1, \cdots, M; \quad r, r' = 1, \cdots, k.$

The update process is shown as follows:

$$\begin{aligned} R_{tt+1}\left(V_p, \mu_r\right) &= 0.5 \times \left[R_{tt+1}\left(V_p, \mu_r\right) + R_{tt}\left(V_p, \mu_r\right)\right] \\ A_{tt+1}\left(V_p, \mu_r\right) &= 0.5 \times \left[A_{tt+1}\left(V_p, \mu_r\right) + A_{tt}\left(V_p, \mu_r\right)\right] \end{aligned} \tag{28}$$

with $p = 1, \cdots, M; \quad r = 1, \cdots, k; \quad tt = 1, \cdots, M/3.$

Finally, the update process is ended as the clustering results.

To have a clear understanding, the flowchart of the SCAF method is illustrated in Fig. 8. The algorithm flow is shown in Algorithm 2.

The pseudocode of the unsupervised SCAF algorithm is shown below. Algorithm 2 can be utilized to generate the movement patterns $C_1, C_2, \cdots, C_k$ for the entire dataset. Furthermore, movement patterns for specific types of ships, such as cargo ships (i.e. C_1, C_2,

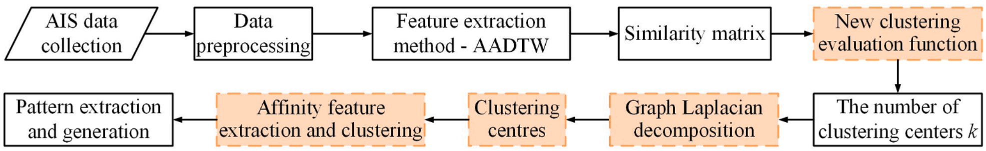

Fig. 8. The flowchart of the proposed SCAF method.

$\cdots, C_{k1}$), tankers(i.e. $C_1, C_2, \cdots, C_{k2}$), and container ships (i.e. $C_1, C_2, \cdots, C_{k3}$), can also be calculated using the unsupervised SCAF algorithm.

Algorithm 2: The unsupervised SCAF algorithm.

Input: The similarity matrix $Q = AADTW_{M\times M}$.
Output: The movement patterns $C_1, C_2, \cdots, C_k$.

```
Initialize: zero matrix R(V_p, μ_r) = A(V_p, μ_r)
Graph network ← Q
Degree matrix D ← Graph network
L = D − Q
L_norm ← D^{-1/2} L D^{-1/2}
Y_l = D^{-1} L
The number of clustering centers k ← f(k)
  k ← max f(k)
  μ_r; r = 1, ⋯, k ← select the k cluster centers from Q
   for p=1: M
     for tt=1: M/3
       R(V_p, μ_r),  A(V_p, μ_r) ← Y, Q
       Update R_{tt+1}(V_p, μ_r) and A_{tt+1}(V_p, μ_r)
        Until C_1, C_2, ⋯, C_k are received
     end for
   end for
```

4.2. Route planning algorithm for MASS

To address the aforementioned optimal local solution and high time complexity in traditional route planning methods, a new route planning algorithm is developed based on the similarity and optimization algorithm in this paper. Leveraging the hidden patterns revealed by the trajectory clustering, the represented route in each cluster is extracted by the proposed AADTW method and the route planning method. The route optimization problem is transformed into a maximum similarity measurement problem, which is essentially a minimum distance calculation problem. The essence of the optimization in the proposed AADTW method is to calculate the similarity and find the best similarity value in each part, i.e., the shortest path, to warp the route based on the Dijkstra algorithm.

Problem modeling. In order to navigate safely and efficiently, historical AIS data is utilized to determine the optimal route between ports or within waterways. By analyzing customary routes from historical AIS data, ships can be guided towards a safe and effective path at sea. It is also important to take into account the specific navigation characteristics and customary paths of different types of ships, such as cargo ships, tankers, and container ships. Therefore, it is crucial to establish distinct routes for each type of vessel to ensure safe navigation. The feature trajectory can be defined as the minimum distance from one trajectory to all other trajectories in each cluster of the whole dataset, cargo ship dataset, tanker dataset, and container ship dataset, and find the specific trajectory as the safe trajectory. Then a new route planning optimization model is formulated by

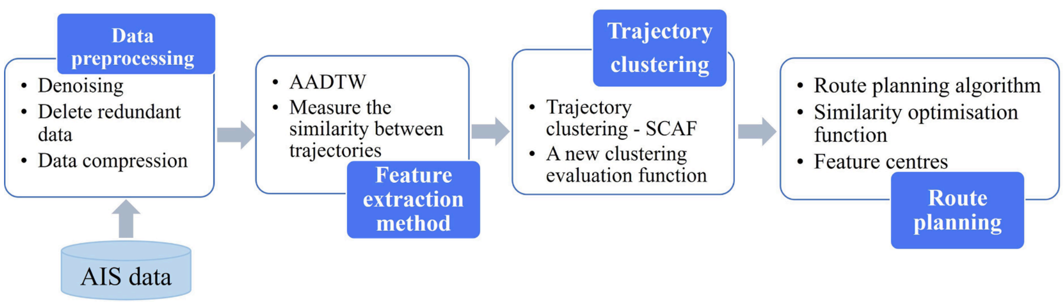

Fig. 9. The experimental flowchart.

$$\begin{aligned} f(n) &= \min_n \sum_{m=1}^{|C_h|} \min\left\{\frac{1}{K}\sum_{t=1}^{K} W_t\right\} \\ &= \min_n \sum_{m=1}^{|C_h|} AADTW(Tr_m, Tr_n) \end{aligned} \tag{29}$$

s.t.

$$AD(1,1) = d\left(P_1^m, P_1^n\right) \tag{30}$$

$$AD(i,1) = AD(1,j) = +\infty \tag{31}$$

$$AD(i,j) = \min\begin{cases} \left(1 + w_{i-1,j}\right) \bullet d\left(P_i^m, P_j^n\right) + AD(i-1,j), \\ d\left(P_i^m, P_j^n\right) + AD(i-1,j-1), \\ \left(1 + w_{i,j-1}\right) \bullet d\left(P_i^m, P_j^n\right) + AD(i,j-1), \end{cases} \tag{32}$$

$$AADTW(Tr_m, Tr_n) = AADTW(Tr_n, Tr_m) \tag{33}$$

$$w_{i,j} = \frac{N_m + N_n}{2\max(N_m, N_n)} \bullet N\left(x_{i,j}\right) \tag{34}$$

$$\begin{aligned} & N\left(x_{i,j}\right) \leqslant \frac{N_m}{3}, \quad N\left(x_{i,j}\right) \leqslant \frac{N_n}{3} \\ i \geqslant 2, \quad & i = 1, \cdots, N_m; \quad j \geqslant 2, j = 1, \cdots, N_n \\ & C_h = 1, 2, \cdots, k \end{aligned} \tag{35}$$

The objective function (i.e. Eq. (29)) aims to minimize the distance from one trajectory to all other trajectories, ultimately identifying the feature center in each cluster. Constraint (30) ensures that the warping path starts from the first point of both trajectories. Constraint (31) restricts all points from one trajectory to be matched with a point in another trajectory. Constraint (32) is to find the minimum path in the three choices from the step (i.e. (i, $j-1$), (i-1, $j-1$), and (i-1, j)) to (i, j) in the proposed AADTW method. Constraint (33) requires the distance matrix is a real symmetric matrix. Constraint (34) expresses the dynamic weights, and constraint (35) limits the usage counts of each point. Similarly, this optimization model can also be used to generate the optimal routes for cargo ships, tankers, and container ships when the input patterns are $C_1, C_2, \cdots, C_{k1}$, $C_1, C_2, \cdots, C_{k2}$, and $C_1, C_2, \cdots, C_{k3}$.

Problem-solving. Given $V \leftarrow (Tra_1, \cdots, Tra_M)$ and $E \leftarrow \{ACDTW_{M\times M}\}$. The goal is to find the minimum distance from one trajectory to all other trajectories in the same cluster C_h for the whole dataset with $h = 1, \cdots, k$.

The trajectory with the maximum similarity with all the other trajectories in each cluster will be determined by Eq. (29). Therefore, $R_1, R_2, \cdots, R_k$ are calculated and defined as the optimal routes for different patterns.

The pseudocode of the proposed route planning method is outlined in Algorithm 3. Step 5 in Algorithm 3 is solved by the objective function specified in Eq. (9). Moreover, the same model can be utilized to calculate the optimal routes $R_1, \cdots, R_{k1}$, $R_1, \cdots, R_{k2}$, and $R_1, \cdots, R_{k3}$ for cargo ships, tankers, and container ships, respectively. The final optimal routes $R_1, R_2, \cdots, R_k$, $R_1, R_2, \cdots, R_{k1}$, $R_1, R_2, \cdots, R_{k2}$, and $R_1, R_2, \cdots, R_{k3}$ are generated to support the navigation of MASS.

Algorithm 3. Route planning algorithm.

Input: Cluster $C_1, C_2, \cdots, C_k$
Output: The safe routes $R_1, R_2, \cdots, R_k$
1. Initialize: The similarity matrix $AD = \varnothing$
2. ***for*** $h = 1 : k$
3. ***for*** $m = 1 : |C_h|$
4. ***for*** $n = 1 : |C_h|$
5. $AD(m,n) \leftarrow AADTW(Tra_m, Tra_n)$
6. $AD \leftarrow AD(m,n)$
7. $f(n) \leftarrow \sum_{n=1}^{n_i} AD(m,n)$
8. $R_n^i \leftarrow \min_n f(n)$
9. $R_i \leftarrow R_n^i$
10. **end *for***
11. **end *for***
12. **end *for***

5. Experimental results and analysis

5.1. Experimental setting

The autonomous navigation of unmanned ships needs experience learned from manned ships. Therefore, the research of movement patterns based on historical AIS data has important guidance for MASS in the waters with a traffic separation scheme. A traffic separation scheme, governing the ships of opposite navigational directions in two separate lanes, is often established in areas with complex waterways to ensure the safe navigation of ships in maritime traffic management. In this paper, the Chengshan Jiao Promontory (CJP) area, one of the complex shipping traffic areas with a traffic separation scheme, is chosen to analyze the movement patterns and hence develop the route planning for MASS. As one of the busiest coastal waters in the world, the CJP water area has a complicated traffic flow and diverse natural environments (wind, wave, current, fog, etc.). Many shipping routes traverse this water area, involving geographical regions such as South Korea, Japan, Taiwan, Bohai Bay, Dalian, Dandong Port, and the Shandong Peninsula.

All numerical experiments are performed using 64-bit Windows 10 on a 3.60 GHz Intel Core i9-11900U CPU, 1080 Ti GPU with 32 GB memory. The proposed algorithms are programmed in MATLAB R2020a and Python. The experimental flowchart is displayed in Fig. 9.

5.2. Data collection and preprocessing

There are many noise points, missing data, incomplete data, and redundant data in the raw AIS dataset. Data processing is indispensable in feature engineering and further guarantees the accuracy of feature mining, such as denoising and compression (Huang et al., 2020). Table 4 lists the statistical and geometrical information in the investigated CJP water. The investigated water area is defined by a longitudinal range of [122.58, 123.17] and a latitudinal range of [37.16, 37.76]. The MMSI, a unique nine-digit number, is selected to identify the different trajectories and classify the ship types in the collected AIS data. The trajectory preprocessing can help find the noise and missing data during the data acquisition and transmission processes. Trajectory compression can significantly simplify ship trajectories while maintaining the main geometric structure and features. The adaptive trajectory compression method is applied to improve the accuracy of feature extraction for MASS route planning (Li et al., 2022; Liu et al., 2019). After the data preprocessing, the dataset comprises 5944 trajectories containing 288,045 points, as displayed in Fig. 10 (a). The investigated CJP area can be defined by the four vertices (122.58, 37.16), (122.58, 37.76), (123.17, 37.16), and (123.17, 37.76). Fig. 10 (b) presents the ship traffic separation scheme in the investigated CJP water area. The black solid lines depict the navigation directions of various ships, while the irregular pink areas are the defined separation zones in different areas.

The statistical information of different types of ships is shown in Table 5, with cargo ships ranking first (73.46%), followed by tankers (15.36%) and container ships (8.8%). The pattern extraction based on ship type classification is carried out in Section 5.5, discovering the hidden features and pattern information to generate navigation knowledge.

5.3. Feature visualization

The features are extracted by the proposed AADTW method, and the similarity between trajectories is measured to discern the patterns for the following MASS route planning. The visualization results of two-dimensional (2D) and three-dimensional (3D) features are displayed in Fig. 11. The 2D result can clearly show the differences in the similarity between trajectories, while the 3D image can further reveal the differences between different patterns. Comparing the results in Fig. 11 (a) and (b), the similarities among trajectories are highlighted to emphasize the high similarity. Furthermore, the comparison results from Fig. 11 also aid in verifying the effectiveness of the proposed AADTW method.

5.4. Pattern extraction results and analysis

To mine and discover the movement patterns quickly and accurately, the 5944 trajectories with 288,045 points are mapped into 5944 points based on feature isometric transformation using the proposed SCAF method in Section 4.2. The optimal k is calculated as the number of clustering centers according to Eq. (15) based on the proposed internal evaluation index function. The optimal value is determined to be the best when $k = 11$. The pattern extraction is carried out after setting the number of clustering to 11 in the abstract graph space based on the SCAF method. The final movement pattern extraction results are displayed in Fig. 12.

The 11 different movement patterns are discovered from the historical AIS trajectory data that can be viewed as a reference for abnormal behaviors. The difference between the routing scheme and the actual routing can also be found by comparing the routing scheme of vessels in the CJP water with Fig. 12. The movement pattern 3 is prohibited in the routing scheme, while it can still be extracted in the historical AIS trajectories data. A thorough analysis is conducted to find the reason for this pattern. The ship types information indicates that the trajectories of pattern 3 are derived from the patrol vessel, rescue ship, and others.

The pattern findings can be used to discern abnormal routes and behaviors when comparing the pattern results with the routes in the traffic separation scheme. Finally, the movement patterns provide navigation support to the route planning for MASS.

To further analyze the information on different patterns, the pattern descriptions are listed to explore the navigation purposes from an overall perspective in Table 4. It shows the moving characteristics and directions in the investigated CJP area. The number of trajectories in each pattern is presented in Table 4, which helps uncover how busy different routes are. It also indicates the traffic situations and trade flows of different waterways. Maritime authorities can therefore increase their capacity to supervise busy waterways or channels to ensure navigation safety. From the number and percentage comparison in Table 6, the top five patterns are 11, 7, 4, 10, and 6, indicating the high traffic density in navigation networks.

To verify the accuracy and effectiveness of pattern extraction, we select the Gaussian mixture model to fit the similarity distribution in each cluster (Li et al., 2022). The features of the within-cluster are extracted based on the AADTW algorithm. Then the Gaussian fitting function is applied to measure the degree of compactness in different clusters. The fitting results of different patterns are shown in Table 7. The fitting performance evaluation based on R^2, *Adjusted* R^2 and Root Mean Square Error (*RMSE*). The closer both R^2 and *Adjusted* R^2 are to 1, the better fitting result is. As shown in Table 11, the values of R^2, *Adjusted* R^2, and *RMSE* all fall in a reasonable range. The ranges of R^2 and *Adjusted* R^2 belong to $[0.9793, 0.9999]$ in each cluster. Therefore, the Gaussian fitting results further verify the effectiveness and accuracy of pattern extraction based on the SCAF method.

Table 4
The dataset information in the CJP water area.

Date	Dataset	Number of ships	Number of points	Boundary points	Longitude (°)	Latitude (°)
March 2018	Original	9148	6,931,737	Left top	122.58	37.16
	After cleaning	5944	288,045	Right bottom	123.17	37.76

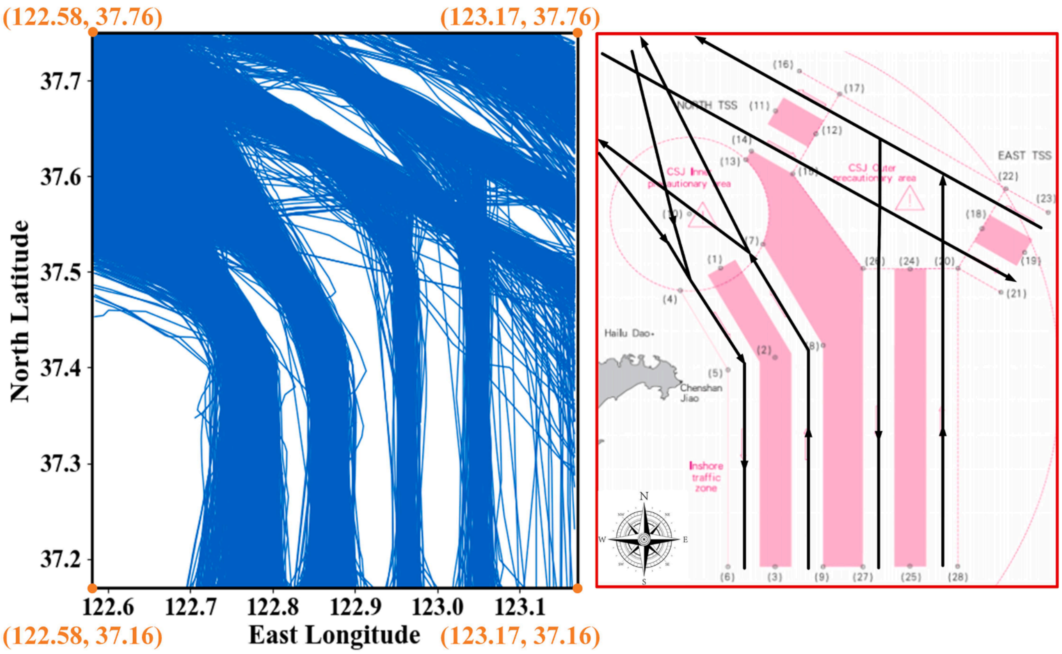

Fig. 10. Visualisation of the ship traffic separation scheme and the dataset after cleaning in the CJP water area.

Table 5
Different types of ships.

Type	Raw dataset			After data preprocessing		
	Number of ships	Percentage (%)	Number of points	Number of ships	Percentage (%)	Number of points
Cargo ship	6720	73.46	5,101,617	4299	72.33	187,912
Tanker	1405	15.36	1,098,657	1033	17.38	66,569
Container ship	805	8.80	577,092	501	8.43	29,879
Fishing vessel	64	0.70	59,630	21	0.35	235
Tugboat	40	0.44	57,679	24	0.40	379
Passenger ship	35	0.38	15,946	20	0.34	628
Ground-effect ship	14	0.15	7580	7	0.12	592
Patrol vessel	8	0.09	5203	7	0.12	655
Dredging vessel	7	0.08	4349	7	0.12	286
Pilot vessel	5	0.05	1564	4	0.07	214
Rescue ship	5	0.05	2420	3	0.05	496
Others	40	0.44	31,113	18	0.30	434
All	9148	1	6,931,737	5944	1	288,045

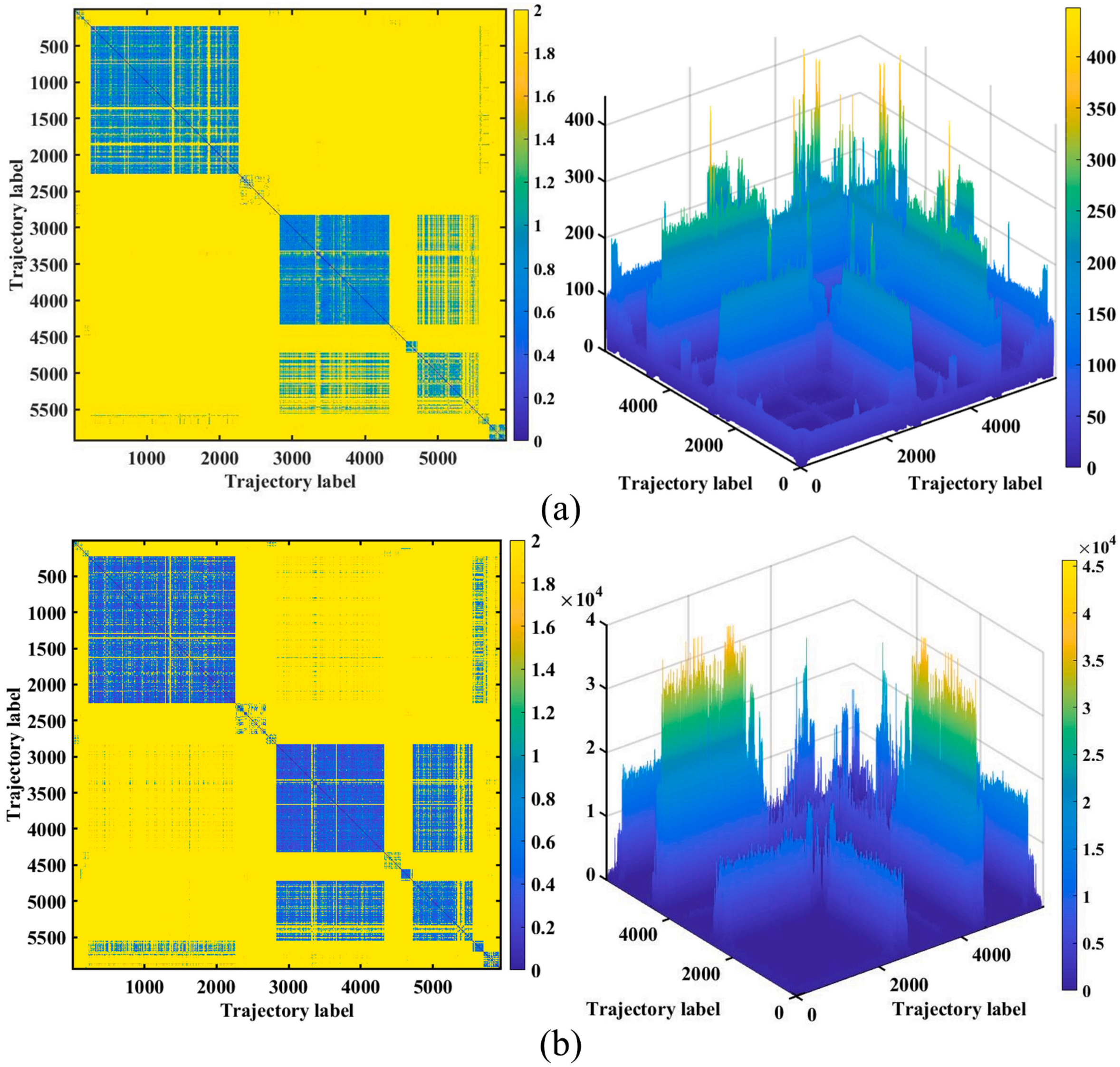

Fig. 11. Visualization of 2D and 3D results of original DTW and the proposed AADTW feature extraction method, (a) the results of the original feature extraction method by DTW, and (b) the results of the new feature extraction method by the proposed AADTW.

5.5. Ship classification analysis

5.5.1. Pattern extraction result and verification for cargo ship

The number of cargo ships accounts for 72.33% of the ship trajectory dataset. The features of cargo ship trajectories are extracted based on the proposed AADTW method. The 2D and 3D visualization results are displayed in Fig. 13, revealing that the difference in similarity between the trajectories is evident.

The number of clustering centers is obtained after the similarity feature extraction is calculated based on Eq. (15). The optimal k is 10 from the function. The patterns of cargo ships are discerned based on the proposed SCAF method, shown in Fig. 14. There are ten patterns in the cargo ship trajectories. Although the ten patterns are included in the eleven patterns in the whole dataset, the features of the ten patterns are not the same as those in the eleven patterns. The essence of the difference is determined by the inherent navigation characteristic. The results also demonstrate the necessity of ship classification analysis, which provides more accurate route planning results for different types of ships.

The detailed pattern descriptions are listed in Table 8. The number and percentage of each cluster are compared to show the difference in different patterns. Table 8 depicts the moving characteristics and directions of cargo ships in the investigated CJP area. The first four patterns are 6, 10, 7, and 9, illustrating that the main navigation areas of cargo ships are between the southeast and south areas of the investigated waters.

Furthermore, the Gaussian mixture model based on AADTW is applied to fit the similarity distribution within each cluster, which

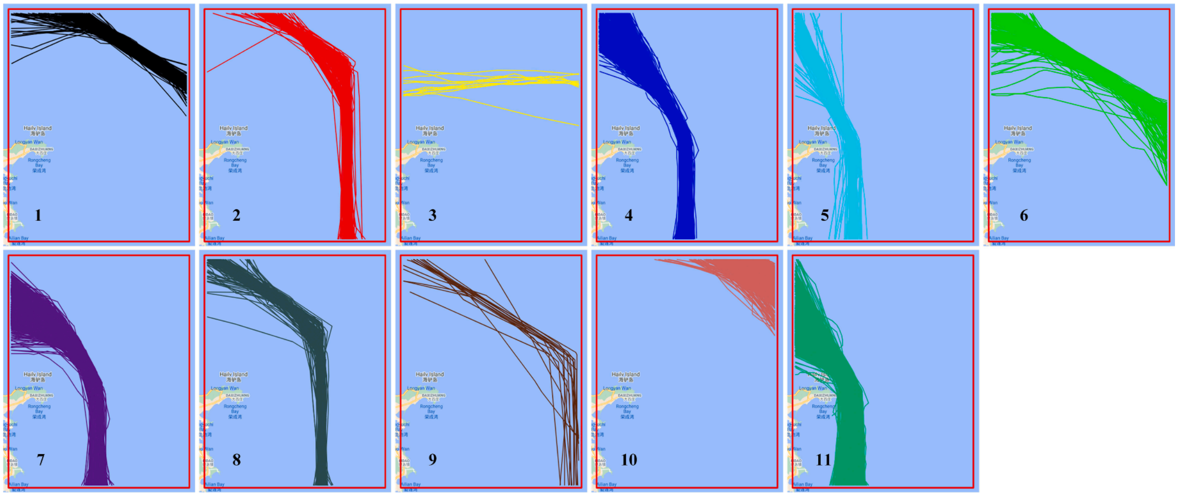

Fig. 12. The results of movement pattern extraction in the investigated CJP water area.

Table 6
Pattern description in the whole AIS dataset.

No.	Pattern description	Number	Percentage (%)	No.	Pattern description	Number	Percentage (%)
1	Southeast-Northwest	281	4.73	**7**	**South-North-West**	**1505**	**25.32**
2	South-North-Northwest (right)	161	2.71	8	Northwest-Southeast-South (middle)	215	3.62
3	East-West	15	0.25	9	Northwest-Southeast-South (right)	17	0.29
4	**South-North-Northwest (middle)**	**836**	**14.06**	**10**	**Northwest-Southeast**	**425**	**7.15**
5	Northwest-Southeast-South (left)	147	2.47	**11**	**West-Southeast-South**	**2031**	**34.17**
6	**Northwest-Southeast (center)**	**311**	**5.23**				

Table 7
The fitting performance evaluation.

No.	R^2	*Adjusted* R^2	*RMSE*	No.	R^2	*Adjusted* R^2	*RMSE*
1	0.9980	0.9980	76.94	7	0.9998	0.9998	838.70
2	0.9939	0.9938	69.18	8	0.9996	0.9996	34.39
3	0.9913	0.9803	5.74	9	0.9988	0.9983	1.96
4	0.9801	0.9830	377	10	0.9998	0.9998	114.90
5	0.9795	0.9793	37.89	11	0.9999	0.9999	881.10
6	0.9999	0.9999	36.31				

enables the exploration of the within-cluster characteristics. The fitting results of the three indexes in the different patterns are shown in Table 9, indicating that the three evaluation indexes further verify the effectiveness of clustering results.

5.5.2. Pattern extraction result and verification for tanker

The number of tankers accounts for 17.38% of the dataset. Similarly to the analysis of cargo ships in Section 5.5.1, the 2D and 3D feature visualization results of tanker trajectories are illustrated in Fig. 15. The features are better extracted to aid the pattern extraction.

The optimal number of clustering centers is calculated as 9 based on Eq. (15). The pattern extraction results of tankers are displayed in Fig. 16, and the nine patterns also show differences from the patterns in the whole dataset. The pattern extraction results of tankers can help make route planning more accurate according to the inherent navigation characteristic. Therefore, the pattern extraction based on the ship classification is useful for planning routes for MASS with different purposes. As shown in Table 10, patterns 2, 3, and 8 occupy the top three seats, accounting for more than 15% of the tanker trajectories. The results of the pattern extraction and their corresponding descriptions reveal that the well-established routes lie in the west, southeast, and south areas of the investigated waters, as shown in Fig. 16 and Table 10. These findings are useful for providing references for planning routes, managing traffic, and guiding the voyage.

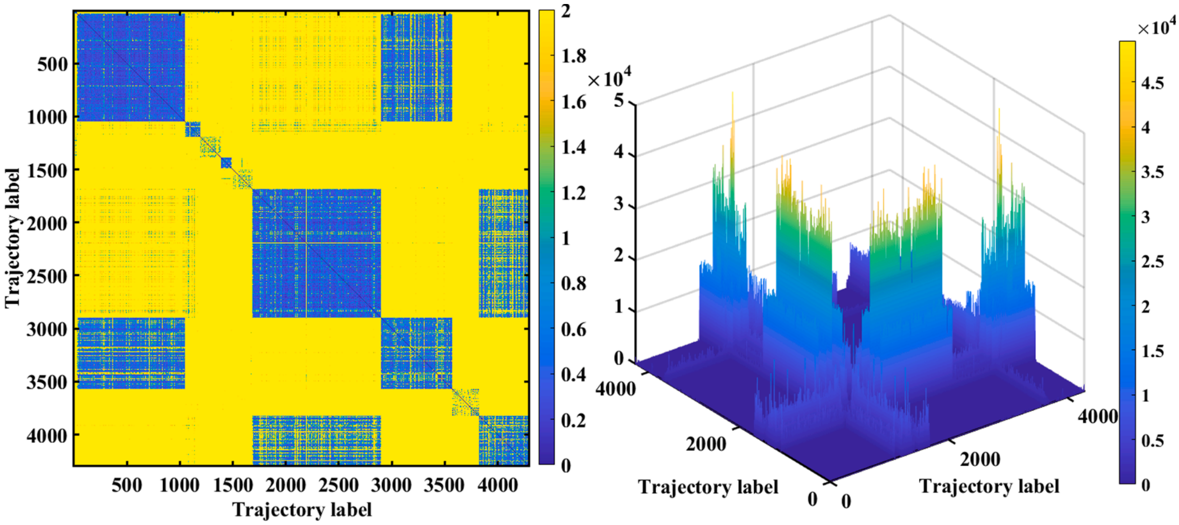

Fig. 13. Visualization of 2D and 3D results of the AADTW feature extraction method.

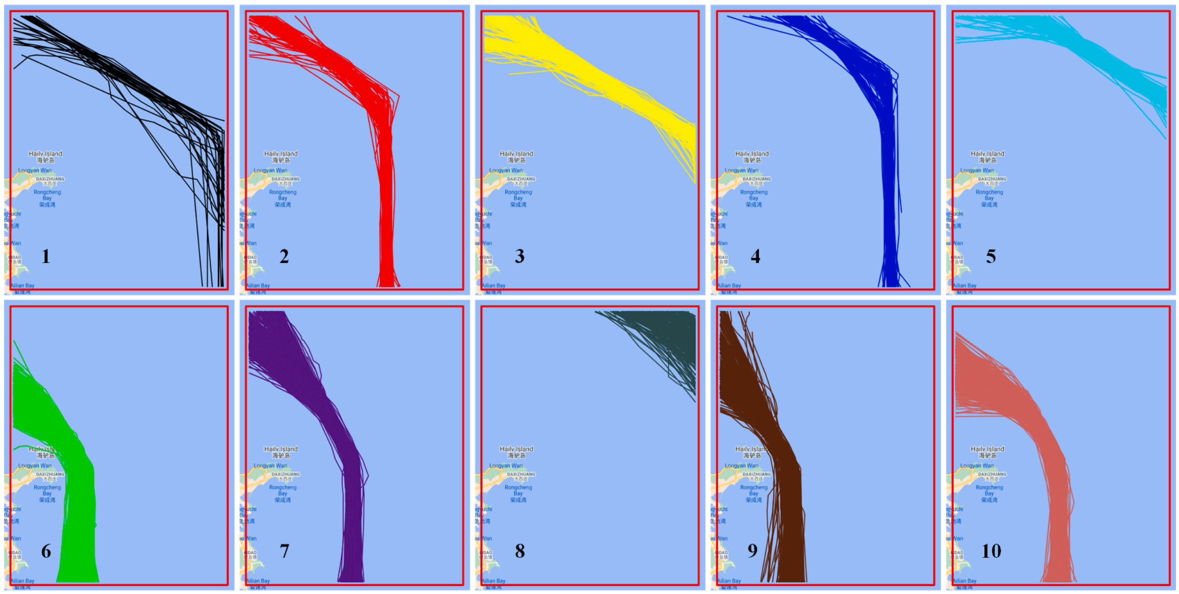

Fig. 14. The results of movement pattern extraction of cargo ships.

Table 8
Pattern description of cargo ship trajectories.

No.	Pattern description	Number	Percentage (%)	No.	Pattern description	Number	Percentage (%)
1	Northwest-Southeast-South (right)	30	0.70	**6**	**West-Southeast-South**	**1214**	**28.24**
2	Northwest-Southeast-South (middle)	144	3.35	**7**	**South-North-Northwest (middle)**	**673**	**15.65**
3	Northwest-Southeast (center)	195	4.54	8	Northwest-Southeast	252	5.86
4	South-North-Northwest (right)	106	2.47	**9**	**Northwest-Southeast-South (left)**	**478**	**11.12**
5	Southeast-Northwest	187	4.35	**10**	**South-North-West**	**1020**	**23.73**

The within-cluster similarity distribution is further analyzed to discover the intrinsic rules and verify the accuracy of pattern extraction. The proposed AADTW algorithm is applied to calculate the similarity matrix in each cluster, and then the Gaussian mixture model is used to fit the similarity distribution. Table 11 displays the results of three evaluation indexes in each cluster. The results of both R^2 and *Adjusted* R^2 in six clusters are close to 1, while the *RMSE* values are close to 0. The results therefore verify the effectiveness

Table 9
The fitting performance evaluation.

No.	R^2	*Adjusted* R^2	*RMSE*	No.	R^2	*Adjusted* R^2	*RMSE*
1	0.9993	0.9991	4.27	6	0.981	0.9808	2099
2	1	1	1.30	7	0.9932	0.9931	175.10
3	1	1	2.84	8	0.9997	0.9997	28.07
4	0.9892	0.989	37.01	9	0.9685	0.9682	448.10
5	0.9893	0.9891	77.41	10	0.987	0.9869	739.90

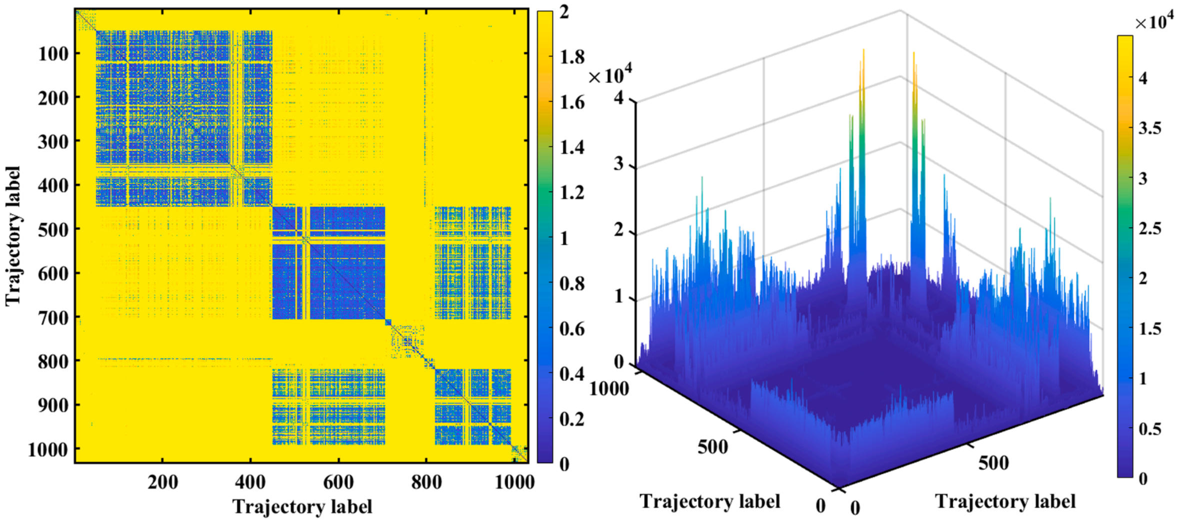

Fig. 15. Visualization of 2D and 3D results of the AADTW feature extraction method.

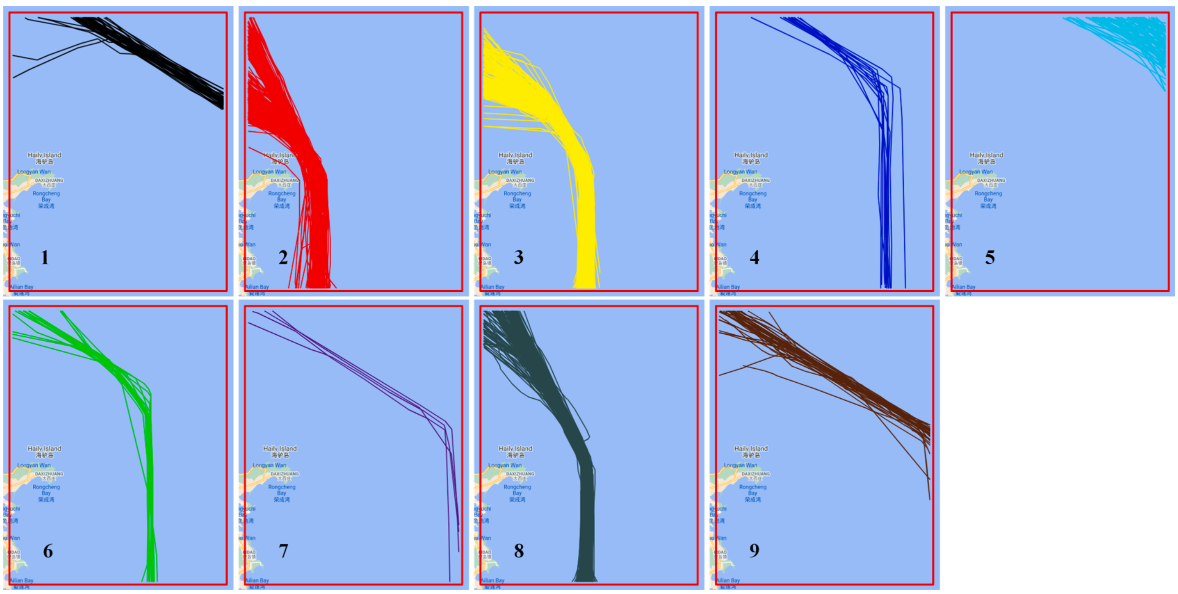

Fig. 16. The results of movement pattern extraction of tankers.

of pattern extraction based on the SCAF method.

5.5.3. *Pattern extraction result and verification for container ship*

The container ship is one of the important ship types in the dataset, accounting for 8.43% of the total. The 2D and 3D feature visualization results of container ship trajectories are shown in Fig. 17 based on the proposed AADTW algorithm. It is evident that the features are more obvious to aid the pattern extraction.

Table 10
Pattern description of tanker trajectories.

No.	Pattern description	Number	Percentage (%)	No.	Pattern description	Number	Percentage (%)
1	Southeast-Northwest	48	4.65	6	Northwest-Southeast-South (middle)	20	1.94
2	**West-Southeast-South**	**401**	**38.82**	7	Northwest-Southeast-South (right)	4	0.39
3	**South-North-West**	**256**	**24.78**	**8**	**South-North-Northwest (middle)**	**173**	**16.75**
4	South-North-Northwest (right)	14	1.36	9	Northwest-Southeast	42	4.07
5	Northwest-Southeast	75	7.26				

Table 11
The fitting performance evaluation.

No.	R^2	*Adjusted* R^2	*RMSE*	No.	R^2	*Adjusted* R^2	*RMSE*
1	1	1	1.200	6	0.9994	0.9991	2.641
2	1	1	37.140	7	1	1	0.001
3	1	1	13.890	8	0.9999	0.9999	10.460
4	0.9938	0.986	1.950	9	1	1	0.914
5	1	1	4.679				

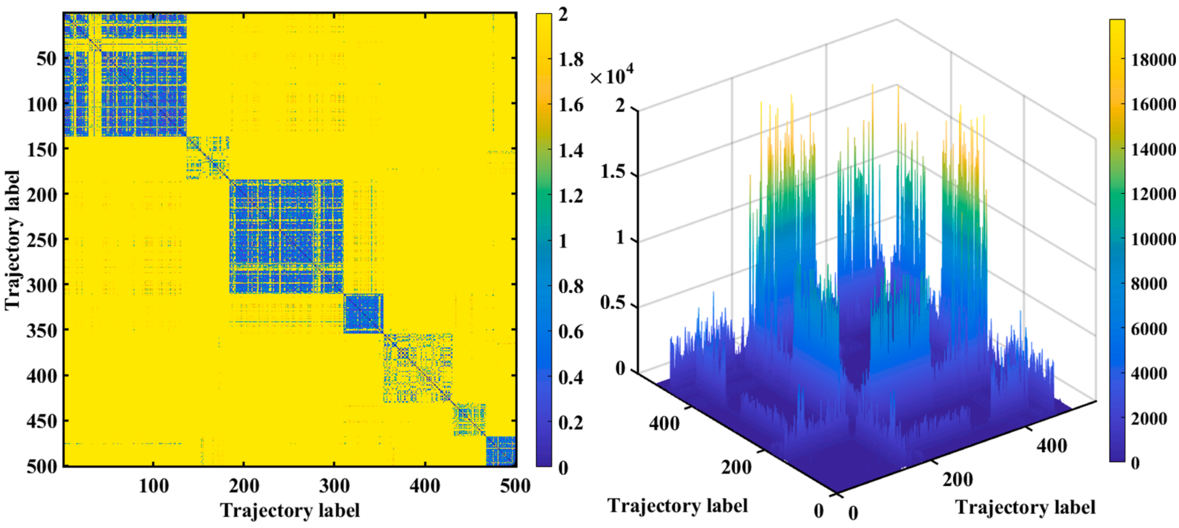

Fig. 17. Visualization of 2D and 3D results of the AADTW feature extraction method.

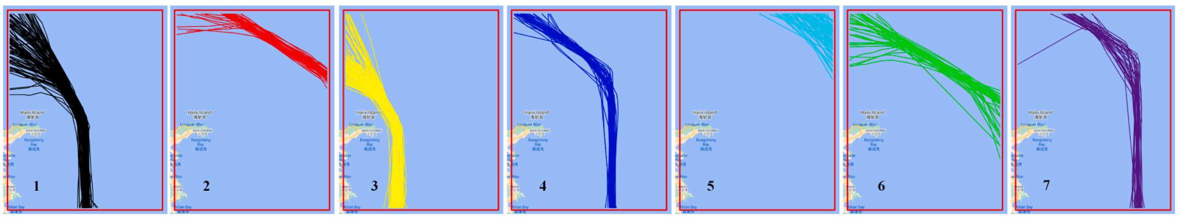

Fig. 18. The results of movement pattern extraction of container ships.

Table 12
Pattern description of container ship trajectories.

No.	Pattern description	Number	Percentage (%)	No.	Pattern description	Number	Percentage (%)
1	**South-North-Northwest (middle)**	**136**	**27.15**	**5**	**Northwest-Southeast**	**76**	**15.17**
2	Southeast-Northwest	48	9.58	6	Northwest-Southeast (Reader)	37	7.39
3	**Northwest-Southeast-South (left)**	**126**	**25.15**	7	South-North-Northwest (right)	34	6.79
4	Northwest-Southeast-South (middle)	44	8.78				

Table 13
The fitting performance evaluation.

No.	R^2	*Adjusted* R^2	*RMSE*	No.	R^2	*Adjusted* R^2	*RMSE*
1	0.9999	0.9999	5.829	5	1	1	3.275
2	0.9999	0.9999	3.115	6	1	1	0.785
3	0.9992	0.9992	7.991	7	0.9996	0.9994	3.777
4	0.9979	0.9971	11.540				

The number of clustering centers in container ship trajectories is defined as 7 by Eq. (15). The pattern extraction results are illustrated in Fig. 18 to visualize the distribution of the moving features. The detailed pattern descriptions are listed in Table 12, revealing that patterns 1, 3, and 5 (which account for more than 15%) are at the top of the list in the container ship trajectories. The best-fit historical routes in container ship trajectories are Northwest-South, South-Northwest, and Northwest-Southeast. The comparison results among Figs. 14, 16, and 18 show that different types of ships have distinct navigation characteristics and preferred routes. Therefore, it is necessary and indispensable to conduct pattern extraction based on ship classification for the route planning of MASS with different purposes.

The intrinsic rules of within-cluster distribution are explored to verify the accuracy of pattern extraction based on the proposed AADTW algorithm and the Gaussian mixture model. The fitting performance evaluation index results are presented in Table 13, demonstrating the effectiveness of pattern extraction based on the SCAF method.

5.5.4. Pattern comparison

The comparison results of different ship types, including cargo ships, tankers, and container ships, are displayed in Fig. 19, revealing their historical trajectory patterns. Cargo ships display the main patterns in clusters 6, 7, 8, and 10, while tankers belong to clusters 2, 3, and 8, and container ships belong to clusters 1, 3, and 5. It is noteworthy that the patterns used for different ship types are significantly different. The patterns are labeled with clusters 1–11, and the movement patterns and customary routes vary across different ship types. Therefore, route planning based on the results of pattern mining is proved to be a rational and effective approach to managing MASS with different purposes.

5.6. Route planning for MASS

The route planning function for MASS is solved from a new data-driven perspective by the AADTW algorithm among the different movement patterns, as described in Section 4.2. A globally optimal solution is found by identifying the similar features in each pattern. The final patterns and the planning routes in the target CJP area are presented in Fig. 20. All the patterns are discovered by the SCAF method from historical AIS trajectory data, as shown in Fig. 20 (a) – (d). The route planning results of different types of ships in the investigated waters are displayed in Fig. 20 (e) – (h). The whole dataset includes 11 different routes in the target area based on different trading routes, which can provide multiple selections for the autonomous navigation of MASS. Different types of ships have different navigational characteristics and modes. Therefore, it is vital to conduct trajectory classification research for the knowledge discovery to support route planning of MASS with different purposes.

From the route planning comparison results in Fig. 20 (e) – (h), cargo ships are the main type of ships in the investigated CJP water area. There are ten patterns in the whole dataset, which can provide references for the manned ships, the mixture of manned and unmanned ships, and MASS. Tankers present the second main type in the CJP waters, and the nine routes are derived from the movement patterns and provide plans based on a ship’s original and destination ports. Container ships have seven alternative routes in this area. The results can be used to develop an automatic system to guide MASS depending on the pair of original and destination ports. Meantime, the seven routes provide sufficient resilient alternatives for tackling the occurrence of any unexpected accident in this area that influences the initially planned route.

Traditional route planning is to design the routes between different ports for ships to find the shortest and/or the lowest-cost routes. The proposed method aids in planning different routes for MASS in the same water according to their best-established navigation behavior and navigation purposes. All these results help build a foundation to support the realization of MASS navigation.

6. Conclusion

In this paper, an AIS data-based machine learning method is developed to conduct feature extraction and unsupervised route planning for MASS using the AADTW, SCAF, and a new route optimization algorithm. Unlike classical ship route planning methods, the newly developed data-driven machine learning method simulates the best-established shipping routes for different types of ships in waters of complex traffic. These routes can then be used to guide both MASS and/or manned ships separately or jointly. The proposed method makes significant contributions to route planning involving real-time collision avoidance in the sense that key hotspots of intersections of the developed routes will be paid more safety attention from a temporal analysis perspective. The proposed AADTW method can extract valuable and accurate features to aid pattern mining, while the unsupervised machine learning method SCAF helps obtain the movement patterns without parameter intervention and address an optimal local problem that existing methods fail to solve. Historical AIS data-based analysis and mining can further discover actual navigation routes and aid route planning for MASS effectively.

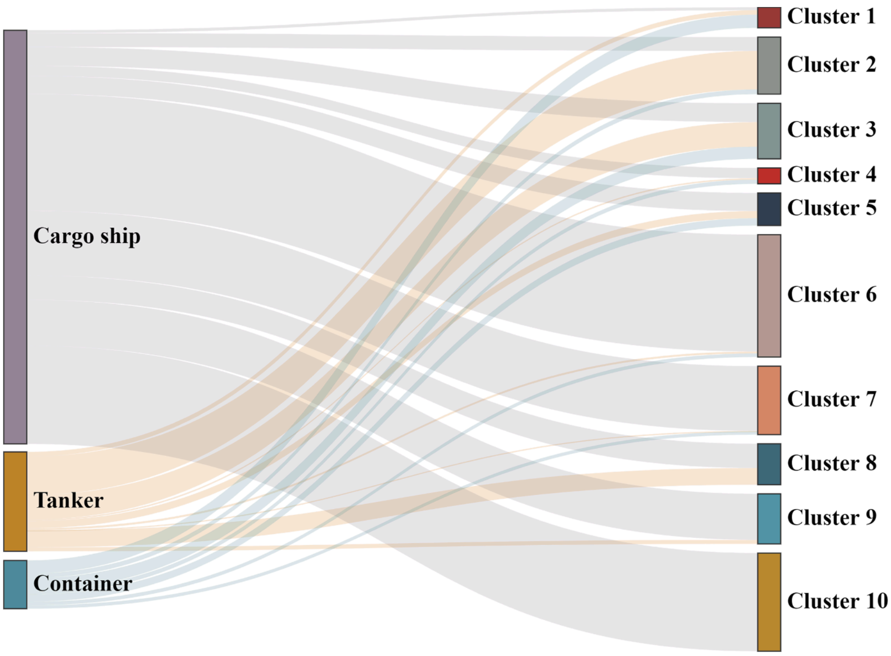

Fig. 19. The comparison result of trajectory data flow.

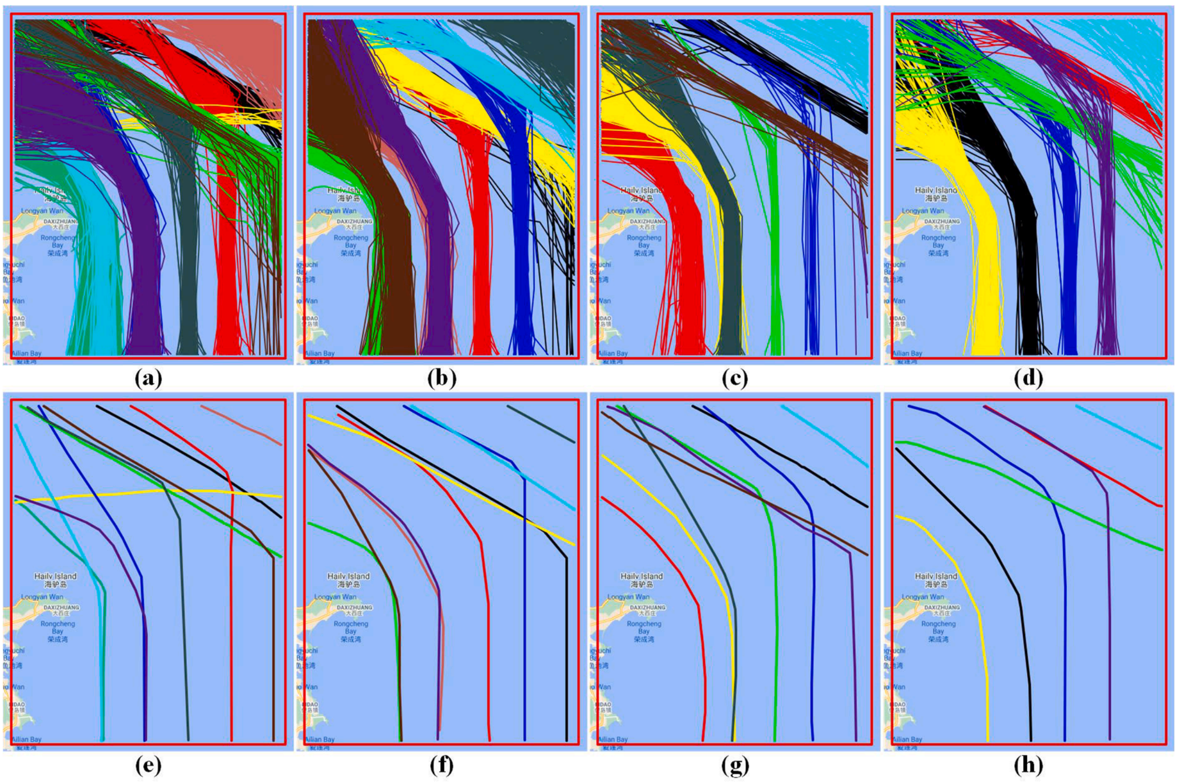

Fig. 20. The final pattern libraries and the route planning results, (a) the pattern results in the investigated waters, (b) the pattern results of cargo ships, (c) the pattern results of tankers, (d) the pattern results of container ships, (e) the planned routes in the investigated waters, (f) the planned routes for cargo ships, (g) the planned routes for tankers, and (h) the planned routes for container ships.

The trajectory classification results provide support for different types of MASS. Due to the joint impact of multiple factors influencing safe navigation in complex waters, the trajectory classification-based route planning method in this study reveals a few advantages over the state-of-the-art methods. For instance, it can effectively solve the local-level optimal problem by integrating the movement patterns from historical AIS data and features obtained by the proposed AADTW method into route planning. Furthermore, the proposed method can find different routes for different types of MASS in the same water to best fit their navigation purposes. Traditional collision avoidance routing methods require setting assumptions due to the high number of influential factors, such as the target ship's tentative behaviors. The newly proposed method can make effective route recommendations based on big data, inherently taking into account the impact of the hidden factors.

The proposed method has significant managerial implications and benefits, including the provision of safe routes in specific areas for MASS to ensure the safety of hybrid traffic, guiding different ship types to navigate through areas of complicated traffic safely, hence reducing the burdens for port management officers. Moreover, the generated safe routes for MASS can provide references for crews in manned ships to guarantee the safety of mixed traffic in the near future. The findings of the route planning for MASS can also reduce human intention and improve the reliability of MASS remote controllers. Finally, the proposed methodology allows MASS system software engineers to embed more parameters influencing safe routing into their design and manufacturing.

Future work could focus on the combination of route prediction, route planning, and GPU-accelerated ensemble algorithms for MASS realization. Meanwhile, the feature extraction of deep clustering without parameters is another promising research direction to provide seamless support for the route design and update for MASS.

CRediT authorship contribution statement

Huanhuan Li: Conceptualization, Methodology, Validation, Investigation, Resources, Project administration, Data curation, Software, Visualization, Writing – original draft, Writing – review & editing. **Zaili Yang:** Validation, Resources, Supervision, Project administration, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by the European Research Council project (TRUST CoG 2019 864724) and Royal Society International Exchanges 2021 Cost Share (NSFC) (IEC\NSFC\211211).

References

Ahvenjärvi, S., 2016. The human element and autonomous ships. TransNav Int. J. Mar. Navig. Saf. Sea Transp. 10 https://doi.org/10.12716/1001.10.03.18.

Aiello, G., Giallanza, A., Mascarella, G., 2020. Towards Shipping 4.0. A preliminary gap analysis. Procedia Manuf., International Conference on Industry 4.0 and Smart Manufacturing (ISM 2019) 42, 24–29. 10.1016/j.promfg.2020.02.019.

Alt, H., Godau, M., 1995. Computing the Fréchet distance between two polygonal curves. Int. J. Comput. Geom. Appl. 5, 75–91.

Berndt, D.J., Clifford, J., 1994. Using Dynamic Time Warping to Find Patterns in Time Series 10, 359–370.

Bueger, C., 2015. What is maritime security? Mar. Policy 53, 159–164.

Burmeister, H.-C., Bruhn, W., Rødseth, Ø.J., Porathe, T., 2014b. Autonomous Unmanned Merchant Vessel and its Contribution towards the e-Navigation Implementation: the MUNIN Perspective. Int. J. E-Navig. Marit. Econ. 1, 1–13. https://doi.org/10.1016/j.enavi.2014.12.002.

Burmeister, H., Bruhn, W., Rødseth, Ø., 2014. Can unmanned ships improve navigational safety? Presented at the Proceedings of the Transport Research Arena, TRA 2014, Paris, pp. 1–10.

Cai, J., Chen, G., Lützen, M., Rytter, N.G.M., 2021. A practical AIS-based route library for voyage planning at the pre-fixture stage. Ocean Eng. 236, 109478 https://doi.org/10.1016/j.oceaneng.2021.109478.

Chan, F.-K.-P., Fu, A.-W.-C., Yu, C., 2003. Haar wavelets for efficient similarity search of time-series: with and without time warping. IEEE Trans. Knowl. Data Eng. 15, 686–705. https://doi.org/10.1109/TKDE.2003.1198399.

Chang, C.-H., Kontovas, C., Yu, Q., Yang, Z., 2021. Risk assessment of the operations of maritime autonomous surface ships. Reliab. Eng. Syst. Saf. 207, 107324 https://doi.org/10.1016/j.ress.2020.107324.

Chen, G., He, C., Wang, T., Zhu, K., Liao, P., Zhang, X., 2022. A superpixel-guided unsupervised fast semantic segmentation method of remote sensing images. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2022.3198065.

Chen, X., Liu, Y., Achuthan, K., Zhang, X., Chen, J., 2021. A semi-supervised deep learning model for ship encounter situation classification. Ocean Eng. 239, 109824 https://doi.org/10.1016/j.oceaneng.2021.109824.

Cheng, H.-H., Ouyang, K., 2021. Development of a strategic policy for unmanned autonomous ships: a study on Taiwan. Marit. Policy Manag. 48, 316–330. https://doi.org/10.1080/03088839.2020.1768315.

Chian Tan, W., Hian Chua, K., Wu, Y., 2021. The automated risk estimation for the navigation of autonomous ships by learning with navigation feature. Int. J. Comput. Methods 18, 2041003. https://doi.org/10.1142/S0219876220410030.

Christiansen, M., Fagerholt, K., Rachaniotis, N.P., Stålhane, M., 2017. Operational planning of routes and schedules for a fleet of fuel supply vessels. Transp. Res. Part E Logist. Transp. Rev. 105, 163–175. https://doi.org/10.1016/j.tre.2016.07.009.

Chuang, T.-N., Lin, C.-T., Kung, J.-Y., Lin, M.-D., 2010. Planning the route of container ships: a fuzzy genetic approach. Expert Syst. Appl. 37, 2948–2956. https://doi.org/10.1016/j.eswa.2009.09.040.

Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs) [WWW Document], n.d. URL https://www.imo.org/en/About/Conventions/Pages/COLREG.aspx (accessed 6.6.22).

Devaurs, D., Siméon, T., Cortés, J., 2015. Efficient sampling-based approaches to optimal path planning in complex cost spaces. In: Akin, H.L., Amato, N.M., Isler, V., van der Stappen, A.F. (Eds.), Algorithmic Foundations of Robotics XI: Selected Contributions of the Eleventh International Workshop on the Algorithmic Foundations of Robotics, Springer Tracts in Advanced Robotics. Springer International Publishing, Cham, pp. 143–159. https://doi.org/10.1007/978-3-319-16595-0_9.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271.

Dong, J.-X., Lee, C.-Y., Song, D.-P., 2015. Joint service capacity planning and dynamic container routing in shipping network with uncertain demands. Transp. Res. Part B Methodol. 78, 404–421. https://doi.org/10.1016/j.trb.2015.05.005.

Duca, A.L., Bacciu, C., Marchetti, A., 2017. A K-nearest neighbor classifier for ship route prediction, in: OCEANS 2017-Aberdeen. IEEE, pp. 1–6.

Emsa, 2022. Preliminary annual overview of marine casualties and incidents 2014–2020 [WWW Document]. accessed 5.30.22. https://www.emsa.europa.eu/publications/item/4378-preliminary-annual-overview-of-marine-casualties-and-incidents-2014-2020.html.

Feng, Z., Zhu, Y., 2016. A survey on trajectory data mining: techniques and applications. IEEE Access 4, 2056–2067. https://doi.org/10.1109/ACCESS.2016.2553681.

Filipiak, D., Węcel, K., Stróżyna, M., Michalak, M., Abramowicz, W., 2020. Extracting maritime traffic networks from AIS data using evolutionary algorithm. Bus. Inf. Syst. Eng. 62, 435–450. https://doi.org/10.1007/s12599-020-00661-0.

Gao, M., Kang, Z., Zhang, A., Liu, J., Zhao, F., 2022. MASS autonomous navigation system based on AIS big data with dueling deep Q networks prioritized replay reinforcement learning. Ocean Eng. 249, 110834 https://doi.org/10.1016/j.oceaneng.2022.110834.

Ge, Q., Han, K., Liu, X., 2021. Matching and routing for shared autonomous vehicles in congestible network. Transp. Res. Part E Logist. Transp. Rev. 156, 102513 https://doi.org/10.1016/j.tre.2021.102513.

Goudossis, A., Katsikas, S.K., 2019. Towards a secure automatic identification system (AIS). J. Mar. Sci. Technol. 24, 410–423. https://doi.org/10.1007/s00773-018-0561-3.

Gu, Y., Goez, J.C., Guajardo, M., Wallace, S.W., 2021. Autonomous vessels: state of the art and potential opportunities in logistics. Int. Trans. Oper. Res. 28, 1706–1739. https://doi.org/10.1111/itor.12785.

Guo, S., Zhang, X., Zheng, Y., Du, Y., 2020. An autonomous path planning model for unmanned ships based on deep reinforcement learning. Sensors 20, 426. https://doi.org/10.3390/s20020426.

Han, P., Yang, X., 2020. Big data-driven automatic generation of ship route planning in complex maritime environments. Acta Oceanol. Sin. 39, 113–120. https://doi.org/10.1007/s13131-020-1638-5.

He, Y., Mou, J., Chen, L., Zeng, Q., Chen, P., Zhang, S., 2021. Survey on hydrodynamic effects on cooperative control of Maritime Autonomous Surface Ships. Ocean Eng. 235, 109300 https://doi.org/10.1016/j.oceaneng.2021.109300.

He, Y.K., Zhang, D., Zhang, J.F., Zhang, M.Y., Li, T.W., 2019. Ship route planning using historical trajectories derived from AIS Data. TransNav Int. J. Mar. Navig. Saf. Sea Transp. 13 https://doi.org/10.12716/1001.13.01.06.

Hirschberg, D.S., 1977. Algorithms for the longest common subsequence problem. J. ACM JACM 24, 664–675. https://doi.org/10.1145/322033.322044.

Huang, Y., Li, Y., Zhang, Z., Liu, R.W., 2020. GPU-Accelerated compression and visualization of large-scale vessel trajectories in maritime IoT industries. IEEE Internet Things J. 7, 10794–10812. https://doi.org/10.1109/JIOT.2020.2989398.

Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 15, 850–863. https://doi.org/10.1109/34.232073.

IMO takes first steps to address autonomous ships [WWW Document], 2018. URL https://www.imo.org/en/MediaCentre/PressBriefings/Pages/08-MSC-99-MASS-scoping.aspx (accessed 12.14.21).

Jaskowiak, P.A., Moulavi, D., Furtado, A.C.S., Campello, R.J.G.B., Zimek, A., Sander, J., 2016. On strategies for building effective ensembles of relative clustering validity criteria. Knowl. Inf. Syst. 47, 329–354. https://doi.org/10.1007/s10115-015-0851-6.

Jeong, M.-G., Lee, E.-B., Lee, M., Jung, J.-Y., 2019. Multi-criteria route planning with risk contour map for smart navigation. Ocean Eng. 172, 72–85. https://doi.org/10.1016/j.oceaneng.2018.11.050.

Keogh, E., Ratanamahatana, C.A., 2005. Exact indexing of dynamic time warping. Knowl. Inf. Syst. 7, 358–386. https://doi.org/10.1007/s10115-004-0154-9.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. J. Mol. Biol. 305, 567–580. https://doi.org/10.1006/jmbi.2000.4315.

Kwong, S., He, Q.H., Man, K.F., Tang, K.S., Chau, C.W., 1998. Parallel genetic-based hybrid pattern matching algorithm for isolated word recognition. Int. J. Pattern Recognit. Artif. Intell. 12, 573–594. https://doi.org/10.1142/S0218001498000348.

Larson, J., Bruch, M., Ebken, J., 2006. Autonomous navigation and obstacle avoidance for unmanned surface vehicles, in: Unmanned Systems Technology VIII. Presented at the Unmanned Systems Technology VIII, SPIE, pp. 53–64. 10.1117/12.663798.

Layton, R., Watters, P., Dazeley, R., 2013. Evaluating authorship distance methods using the positive Silhouette coefficient. Nat. Lang. Eng. 19, 517–535. https://doi.org/10.1017/S1351324912000241.

Lee, S.-M., Roh, M.-I., Kim, K.-S., Jung, H., Park, J.J., 2018. Method for a simultaneous determination of the path and the speed for ship route planning problems. Ocean Eng. 157, 301–312. https://doi.org/10.1016/j.oceaneng.2018.03.068.

Li, X., Zhao, K., Cong, G., Jensen, C.S., Wei, W., 2018. Deep Representation Learning for Trajectory Similarity Computation, in: 2018 IEEE 34th International Conference on Data Engineering (ICDE). Presented at the 2018 IEEE 34th International Conference on Data Engineering (ICDE), pp. 617–628. 10.1109/ICDE.2018.00062.

Li, S., Fung, K.S., 2019. Maritime autonomous surface ships (MASS): implementation and legal issues. Marit. Bus. Rev. 4, 330–339. https://doi.org/10.1108/MABR-01-2019-0006.

Li, H., Liu, J., Liu, R.W., Xiong, N., Wu, K., Kim, T., 2017. A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis. Sensors 17, 1792. https://doi.org/10.3390/s17081792.

Li, H., Liu, J., Wu, K., Yang, Z., Liu, R.W., Xiong, N., 2018a. Spatio-temporal vessel trajectory clustering based on data mapping and density. IEEE Access 6, 58939–58954. https://doi.org/10.1109/ACCESS.2018.2866364.

Li, H., Liu, J., Yang, Z., Liu, R.W., Wu, K., Wan, Y., 2020. Adaptively constrained dynamic time warping for time series classification and clustering. Inf. Sci. 534, 97–116. https://doi.org/10.1016/j.ins.2020.04.009.

Li, H., Jiao, H., Yang, Z., 2023a. AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods. Transp. Res. Part E Logist. Transp. Rev. 175, 103152. https://doi.org/10.1016/j.tre.2023.103152.

Li, H., Lam, J.S.L., Yang, Z., Liu, J., Liu, R.W., Liang, M., Li, Y., 2022. Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. Transp. Res. Part C Emerg. Technol. 143, 103856 https://doi.org/10.1016/j.trc.2022.103856.

Li, H., Ren, X., Yang, Z., 2023b. Data-driven Bayesian network for risk analysis of global maritime accidents. Reliab. Eng. Syst. Saf. 230, 108938 https://doi.org/10.1016/j.ress.2022.108938.

Li, L., Wu, D., Huang, Y., Yuan, Z.-M., 2021. A path planning strategy unified with a COLREGS collision avoidance function based on deep reinforcement learning and artificial potential field. Appl. Ocean Res. 113, 102759 https://doi.org/10.1016/j.apor.2021.102759.

Li, Y., Zheng, J., 2021. Deep learning structure for collision avoidance planning of unmanned surface vessel. Proc. Inst. Mech. Eng. Part M J. Eng. Marit. Environ. 235, 511–520. https://doi.org/10.1177/1475090220970102.

Liang, M., Liu, R.W., Li, S., Xiao, Z., Liu, X., Lu, F., 2021a. An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation. Ocean Eng. 225, 108803 https://doi.org/10.1016/j.oceaneng.2021.108803.

Liang, M., Zhan, Y., Liu, R.W., 2021b. MVFFNet: Multi-view feature fusion network for imbalanced ship classification. Pattern Recognit. Lett. 151, 26–32. https://doi.org/10.1016/j.patrec.2021.07.024.

Liang, M., Liu, R.W., Zhan, Y., Li, H., Zhu, F., Wang, F.-Y., 2022. Fine-grained vessel traffic flow prediction with a spatio-temporal multigraph convolutional network. IEEE Trans. Intell. Transp. Syst. 1–14 https://doi.org/10.1109/TITS.2022.3199160.

Lin, D.-Y., Chang, Y.-T., 2018. Ship routing and freight assignment problem for liner shipping: Application to the Northern Sea Route planning problem. Transp. Res. Part E Logist. Transp. Rev. 110, 47–70. https://doi.org/10.1016/j.tre.2017.12.003.

Liu, C., Chu, X., Wu, W., Li, S., He, Z., Zheng, M., Zhou, H., Li, Z., 2022a. Human–machine cooperation research for navigation of maritime autonomous surface ships: a review and consideration. Ocean Eng. 246, 110555 https://doi.org/10.1016/j.oceaneng.2022.110555.

Liu, S., Jiang, H., Chen, S., Ye, J., He, R., Sun, Z., 2020. Integrating Dijkstra's algorithm into deep inverse reinforcement learning for food delivery route planning. Transp. Res. Part E Logist. Transp. Rev. 142, 102070 https://doi.org/10.1016/j.tre.2020.102070.

Liu, J., Li, H., Yang, Z., Wu, K., Liu, Y., Liu, R.W., 2019. Adaptive douglas-peucker algorithm with automatic thresholding for AIS-based vessel trajectory compression. IEEE Access 7, 150677–150692. https://doi.org/10.1109/ACCESS.2019.2947111.

Liu, R.W., Liang, M., Nie, J., Yuan, Y., Xiong, Z., Yu, H., Guizani, N., 2022b. STMGCN: mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multi-graph convolutional network. IEEE Trans. Ind. Inform. 18, 7977–7987. https://doi.org/10.1109/TII.2022.3165886.

Liu, A., Zhu, Q., Xu, L., Lu, Q., Fan, Y., 2021. Sustainable supply chain management for perishable products in emerging markets: an integrated location-inventory-routing model. Transp. Res. Part E Logist. Transp. Rev. 150, 102319 https://doi.org/10.1016/j.tre.2021.102319.

Mayflower Autonomous Ship [WWW Document], 2021. URL https://mas400.com/dashboard#currentLocation (accessed 12.14.21).

Mascaro, S., Nicholso, A.E., Korb, K.B., 2014. Anomaly detection in vessel tracks using Bayesian networks. Int. J. Approx. Reason. 55, 84–98.

Meng, Q., Wang, S., Lee, C.-Y., 2015. A tailored branch-and-price approach for a joint tramp ship routing and bunkering problem. Transp. Res. Part B Methodol. 72, 1–19. https://doi.org/10.1016/j.trb.2014.11.008.

Nanni, M., Pedreschi, D., 2006. Time-focused clustering of trajectories of moving objects. J. Intell. Inf. Syst. 27, 267–289. https://doi.org/10.1007/s10844-006-9953-7.

Naus, K., 2020. Drafting route plan templates for ships on the basis of AIS historical data. J. Navig. 73, 726–745.

Öztürk, Ü., Akdağ, M., Ayabakan, T., 2022. A review of path planning algorithms in maritime autonomous surface ships: navigation safety perspective. Ocean Eng. 251, 111010 https://doi.org/10.1016/j.oceaneng.2022.111010.

Pallotta, G., Vespe, M., Bryan, K., 2013. Vessel pattern knowledge discovery from AIS Data: a framework for anomaly detection and route prediction. Entropy 15, 2218–2245. https://doi.org/10.3390/e15062218.

Petitjean, F., Ketterlin, A., Gançarski, P., 2011. A global averaging method for dynamic time warping, with applications to clustering. Pattern Recognit. 44, 678–693. https://doi.org/10.1016/j.patcog.2010.09.013.

Qi, Y., Harrod, S., Psaraftis, H.N., Lang, M., 2022. Transport service selection and routing with carbon emissions and inventory costs consideration in the context of the Belt and Road Initiative. Transp. Res. Part E Logist. Transp. Rev. 159, 102630 https://doi.org/10.1016/j.tre.2022.102630.

Rashidi, S., Fallah, A., Towhidkhah, F., 2015. Local features analysis of on-line signature using modified distance of DTW. Int. J. Comput. Methods 12, 1550016. https://doi.org/10.1142/S0219876215500164.

IMO Resolution A.528(13) – Recommendation on weather routing – (Adopted on 17 November 1983) [WWW Document], 1983. . topic. URL https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/AssemblyDocuments/A.528(13).pdf (accessed 6.6.22).

IMO Resolution A.893(21) – Guidelines for Voyage Planning – (Adopted on 25 November 1999) [WWW Document], n.d. URL https://imorules.com/IMORES_A893.21.html (accessed 6.6.22).

Rong, H., Teixeira, A.P., Guedes Soares, C., 2022. Maritime traffic probabilistic prediction based on ship motion pattern extraction. Reliab. Eng. Syst. Saf. 217, 108061 https://doi.org/10.1016/j.ress.2021.108061.

Sarraf, R., McGuire, M.P., 2020. Integration and comparison of multi-criteria decision making methods in safe route planner. Expert Syst. Appl. 154, 113399 https://doi.org/10.1016/j.eswa.2020.113399.

Sheng, K., Liu, Z., Zhou, D., He, A., Feng, C., 2018. Research on ship classification based on trajectory features. J. Navig. 71, 100–116. https://doi.org/10.1017/S0373463317000546.

Shi, J., Liu, Z., 2020. Deep learning in unmanned surface vehicles collision-avoidance pattern based on AIS Big Data with Double GRU-RNN. J. Mar. Sci. Eng. 8, 682. https://doi.org/10.3390/jmse8090682.

Song, D.-P., Dong, J.-X., 2012. Cargo routing and empty container repositioning in multiple shipping service routes. Transp. Res. Part B Methodol. 46, 1556–1575. https://doi.org/10.1016/j.trb.2012.08.003.

Song, D.-P., Dong, J.-X., 2013. Long-haul liner service route design with ship deployment and empty container repositioning. Transp. Res. Part B Methodol. 55, 188–211. https://doi.org/10.1016/j.trb.2013.06.012.

Sun, L., Guo, C., 2014. Incremental affinity propagation clustering based on message passing. IEEE Trans. Knowl. Data Eng. 26, 2731–2744. https://doi.org/10.1109/TKDE.2014.2310215.

Sun, L., Guo, C., Liu, C., Xiong, H., 2017. Fast affinity propagation clustering based on incomplete similarity matrix. Knowl. Inf. Syst. 51, 941–963. https://doi.org/10.1007/s10115-016-0996-y.

Tam, C., Bucknall, R., Greig, A., 2009. Review of collision avoidance and path planning methods for ships in close range encounters. J. Navig. 62, 455–476. https://doi.org/10.1017/S0373463308005134.

Tsou, M.-C., Cheng, H.-C., 2013. An ant colony algorithm for efficient ship routing. Pol. Marit. Res. 3, 28–38.

Véras, L.G.D.O., Medeiros, F.L.L., Guimaráes, L.N.F., 2019. Systematic literature review of sampling process in rapidly-exploring random trees. IEEE Access 7, 50933–50953. https://doi.org/10.1109/ACCESS.2019.2908100.

Vespe, M., Gibin, M., Alessandrini, A., Natale, F., Mazzarella, F., Osio, G.C., 2016. Mapping EU fishing activities using ship tracking data. J. Maps 12, 520–525. https://doi.org/10.1080/17445647.2016.1195299.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416. https://doi.org/10.1007/s11222-007-9033-z.

Wang, Y., Liu, K., Zhang, R., Qian, L., Shan, Y., 2021b. Feasibility of the Northeast Passage: the role of vessel speed, route planning, and icebreaking assistance determined by sea-ice conditions for the container shipping market during 2020–2030. Transp. Res. Part E Logist. Transp. Rev. 149, 102235 https://doi.org/10.1016/j.tre.2021.102235.

Wang, H., Mao, W., Eriksson, L., 2019. A three-dimensional dijkstra's algorithm for multi-objective ship voyage optimization. Ocean Eng. 186, 106131 https://doi.org/10.1016/j.oceaneng.2019.106131.

Wang, S., Meng, Q., 2012. Liner ship route schedule design with sea contingency time and port time uncertainty. Transp. Res. Part B Methodol. 46, 615–633. https://doi.org/10.1016/j.trb.2012.01.003.

Wang, Q., Zhong, M., Shi, G., Zhao, J., Bai, C., 2021a. Route planning and tracking for ships based on the ECDIS platform. IEEE Access 9, 71754–71762. https://doi.org/10.1109/ACCESS.2021.3078899.

Wei, Z., Xie, X., Zhang, X., 2020. AIS trajectory simplification algorithm considering ship behaviours. Ocean Eng. 216, 108086 https://doi.org/10.1016/j.oceaneng.2020.108086.

Wen, Y., Sui, Z., Zhou, C., Xiao, C., Chen, Q., Han, D., Zhang, Y., 2020. Automatic ship route design between two ports: a data-driven method. Appl. Ocean Res. 96, 102049 https://doi.org/10.1016/j.apor.2019.102049.

Wu, G., Atilla, I., Tahsin, T., Terziev, M., Wang, L., 2021a. Long-voyage route planning method based on multi-scale visibility graph for autonomous ships. Ocean Eng. 219, 108242 https://doi.org/10.1016/j.oceaneng.2020.108242.

Wu, L., Wang, S., Laporte, G., 2021b. The robust bulk ship routing problem with batched cargo selection. Transp. Res. Part B Methodol. 143, 124–159. https://doi.org/10.1016/j.trb.2020.11.003.

Xue, Y., Clelland, D., Lee, B.S., Han, D., 2011. Automatic simulation of ship navigation. Ocean Eng. 38, 2290–2305. https://doi.org/10.1016/j.oceaneng.2011.10.011.

Yao, D., Zhang, C., Zhu, Z., Huang, J., Bi, J., 2017. Trajectory clustering via deep representation learning. In: in: 2017 International Joint Conference on Neural Networks (IJCNN). Presented at the 2017 International Joint Conference on Neural Networks (IJCNN), pp. 3880–3887. https://doi.org/10.1109/IJCNN.2017.7966345.
Yao, D., Zhang, C., Zhu, Z., Hu, Q., Wang, Z., Huang, J., Bi, J., 2018. Learning deep representation for trajectory clustering. Expert Syst. 35, e12252.
Yu, G.-F., Lin, Y.-J., Luo, X.-M., 2023. A novel variable weight VIKOR grade assessment method for waterway navigation safe routes selection. J. Mar. Sci. Eng. 11, 347. https://doi.org/10.3390/jmse11020347.
Zhang, M., Kujala, P., Hirdaris, S., 2022b. A machine learning method for the evaluation of ship grounding risk in real operational conditions. Reliab. Eng. Syst. Saf. 226, 108697 https://doi.org/10.1016/j.ress.2022.108697.
Zhang, Y., Liu, A., Liu, G., Li, Z., Li, Q., 2019b. Deep representation learning of activity trajectory similarity computation. In: in: 2019 IEEE International Conference on Web Services (ICWS). Presented at the 2019 IEEE International Conference on Web Services (ICWS), pp. 312–319. https://doi.org/10.1109/ICWS.2019.00059.
Zhang, L., Liu, Z., Yu, L., Fang, K., Yao, B., Yu, B., 2022a. Routing optimization of shared autonomous electric vehicles under uncertain travel time and uncertain service time. Transp. Res. Part E Logist. Transp. Rev. 157, 102548 https://doi.org/10.1016/j.tre.2021.102548.
Zhang, Y., Sun, X., Chen, J., Cheng, C., 2021b. Spatial patterns and characteristics of global maritime accidents. Reliab. Eng. Syst. Saf. 206, 107310 https://doi.org/10.1016/j.ress.2020.107310.
Zhang, J., Wang, F.-Y., Wang, K., Lin, W.-H., Xu, X., Chen, C., 2011. Data-driven intelligent transportation systems: a survey. IEEE Trans. Intell. Transp. Syst. 12, 1624–1639. https://doi.org/10.1109/TITS.2011.2158001.
Zhang, X., Wang, C., Jiang, L., An, L., Yang, R., 2021a. Collision-avoidance navigation systems for maritime autonomous surface ships: a state of the art survey. Ocean Eng. 235, 109380 https://doi.org/10.1016/j.oceaneng.2021.109380.
Zhang, R., Xie, P., Jiang, H., Xiao, Z., Wang, C., Liu, L., 2019a. Clustering Noisy Trajectories via Robust Deep Attention Auto-Encoders. In: IEEE International Conference on Mobile Data Management (MDM). Presented at the 2019 20th IEEE International Conference on Mobile Data Management (MDM), pp. 63–71. https://doi.org/10.1109/MDM.2019.00-73.
Zhao, L., Shi, G., 2019. A novel similarity measure for clustering vessel trajectories based on dynamic time warping. J. Navig. 72, 290–306.
Zhao, W., Wang, Y., Zhang, Z., Wang, H., 2021. Multicriteria ship route planning method based on improved particle swarm optimization-genetic algorithm. J. Mar. Sci. Eng. 9, 357. https://doi.org/10.3390/jmse9040357.
Zhen, R., Jin, Y., Hu, Q., Shao, Z., Nikitakos, N., 2017. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. J. Navig. 70, 648–670.