Transportation Research Part C 164 (2024) 104670

Contents lists available at ScienceDirect

Transportation Research Part C

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

Maritime traffic partitioning: An adaptive semi-supervised spectral regularization approach for leveraging multi-graph evolutionary traffic interactions

Xuri Xin [a,b], Kezhong Liu [a], Huanhuan Li [b], Zaili Yang [b,*]

[a] School of Navigation, Wuhan University of Technology, Wuhan, Hubei, China
[b] Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK

ARTICLE INFO

Keywords:
Maritime safety
Intelligent transportation
Ship traffic partitioning
Multi-graph evolutionary traffic interactions
Spectral regularization

ABSTRACT

Maritime situational awareness (MSA) has long been a critical focus within the domain of maritime traffic surveillance and management. The increasing complexities of ship traffic, originating from sophisticated multi-attribute interactions among multiple ships, coupled with the continuous evolution of traffic dynamics, pose significant challenges in attaining accurate MSA, particularly in complex port waters. This study is dedicated to establishing an advanced methodology for partitioning maritime traffic, aimed at enhancing traffic pattern interpretability and strengthening ship anti-collision risk management. Specifically, three interaction measure metrics, including conflict criticality, spatial distance, and approaching rate, are initially introduced to quantify different aspects of spatiotemporal interactions among ships. Subsequently, a semi-supervised spectral regularization framework is devised to adeptly accommodate both multiple interaction information and prior knowledge derived from historic partitioning structures. This framework facilitates the segmentation of regional traffic into multiple clusters, wherein ships with the same cluster exhibit high temporal stability, conflict connectivity, spatial compactness, and convergent motion. Meanwhile, an adaptive hyperparameter selection model is engineered to seek optimal traffic partitioning outcomes across diverse scenarios, while also incorporating user preferences for specific interaction indicators. Comprehensive experiments using AIS data from Ningbo-Zhoushan Port are undertaken to thoroughly assess the models' efficacy. Research findings from case analyses and model comparisons distinctly showcase the capability of the proposed approach to successfully deconstruct the regional traffic complexity, capture high-risk zones, and strengthen strategic maritime safety measures. Consequently, the methodology holds significant promise for advancing the intelligence of maritime surveillance systems and facilitating the automation of maritime traffic management.

1. Introduction

Maritime traffic safety has consistently been a matter of widespread concern in the shipping industry due to the severe consequences of maritime traffic accidents, including personal casualties, environmental pollution, and economic losses. Therefore, advanced Maritime Situational Awareness (MSA), focusing on traffic situation perception and collision risk estimation, plays a pivotal

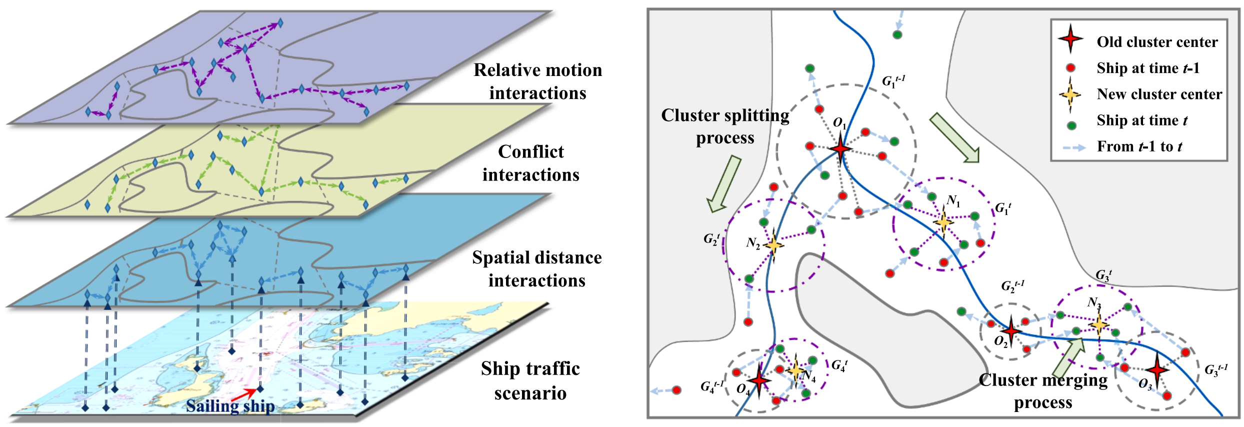
* Corresponding author.
E-mail address: Z.Yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.trc.2024.104670
Received 2 October 2023; Received in revised form 15 April 2024; Accepted 17 May 2024
Available online 27 May 2024

role in ensuring navigation safety and efficiency, economic benefits, and environmental preservation (Fang et al., 2018). The development of digitalization has facilitated the deployment and integration of various technological advances, such as Artificial Intelligence (AI), Internet of Things (IoT), Blockchain technology, Cloud Computing, and Big Data, into current maritime transportation systems at an unprecedented pace (Li et al., 2023; Li and Yang, 2023; Liu et al., 2022c). These intelligent and autonomous systems are anticipated to assist relevant users in proactive risk prediction and management, thereby mitigating accidents caused by human errors (Bakdi et al., 2021). Additionally, the elimination of unnecessary human operations will substantially reduce the staff's workload and associated resource costs. Driven by the potential to enhance safety and achieve attractive socioeconomic benefits, current developments are accelerating the revolutionization of the maritime industry toward perception automation and system intelligence. While practical applications of these advanced devices have begun in many regions, maritime navigation autonomy and intelligence are still in their initial development stage. As a result, there is a crucial need to develop advanced MSA techniques to support the growth of intelligent transportation service systems and pave the way for fully autonomous ships (Bogusławski et al., 2022; Montewka et al., 2022; Zhang et al., 2023a).

In respond to the increasing demand for enhancing MSA across relevant waters, numerous collision risk estimation and control approaches and models have been put forth to mitigate risks and enhance the efficiency of maritime transportation (Du et al., 2020; Huang et al., 2020; Xin et al., 2023a; Yu et al., 2023). Particularly, the widespread deployment of the Automatic Identification System (AIS) not only enhances safety management by providing early collision alerts and aiding anti-collision decision-making, but also facilitates trajectory behavior analysis and modeling due to the increasing availability of abundant AIS-based trajectory data (Li et al., 2024a,b; Xiao et al., 2019; Zhou et al., 2019). For instance, AIS data has found utility in diversified maritime safety-related investigations, supporting traffic characteristic statistics (Rong et al., 2021; Xin et al., 2019), traffic movement prediction (Li et al., 2023; Zhang et al., 2023b; Zhang et al., 2023c), multi-ship encounter detection (Wang et al., 2024a, 2024b), anomalous behavior detection (Rong et al., 2022), and route plan (Li and Yang, 2023; Zhang et al., 2022a), among others. However, the expanding traffic volume and the emergence of large, fast, and autonomous ships are amplifying traffic complexity and potential navigational hazards, especially within specialized water areas like congested ports and restricted waterways. These areas exhibit high traffic density, dynamic ship movements, and uneven traffic distribution, leading to an extremely high frequency of multi-ship conflicts over both space and time (Yu et al., 2021, 2019). This significantly impedes the practical effectiveness of existing state-of-the-art methods. In the meantime, the extensive surveillance areas and varied traffic movement activities present a challenge for maritime operational authorities seeking to comprehend traffic situations (Arguedas et al., 2017). Currently, maritime operators heavily rely on their intuition and experience to detect potential multi-ship encounters from a global surveillance perspective, which proves to be both unreliable and costly, further limiting their decision-making capabilities for risk mitigation. This concern becomes even more worrisome as traffic situations within port areas are projected to become increasingly complicated, given the possibility of hybrid encountering scenarios involving both manned and autonomous ships. Indeed, the ship traffic within high-density waters exhibits clustering characteristics, where ships within the same group engage in high spatiotemporal interactions, while ships between groups have relatively sparse spatiotemporal interactions. Therefore, the prospect of devising a practical traffic partitioning model to segment overall ship traffic into mutiple compact, scalable, and comprehensible clusters holds promise. Such a model could effectively break down regional traffic complexity, proactively capture high-risk multi-ship encounters, and facilitate anti-collision decision-making, thereby opening up new possibilities for tackling more challenging traffic scenarios.

Nonetheless, the unique characteristics of maritime traffic raise a significant challenge in developing a desirable traffic partitioning model, as outlined below.



(a) (b)

Fig. 1. Illustration of ship traffic encountering scenarios in high-density traffic areas. (a) Multi-attribute interactions among ships; (b) traffic evolution over time.

Problem 1: ***How to integrate different interaction indicators between ship pairs into the traffic partitioning model?***

The interactions between ship pairs can be characterized by various indicators, such as conflict criticality, spatial distance, and relative motion (Sun et al., 2024; Wen et al., 2015; Xin et al., 2021; Zhang et al., 2015). As shown in Fig. 1(a), each indicator highlights a different aspect of interactions among ship traffic, thereby contributing to a meaningful partitioning outcome from its distinct perspective. However, implementing traffic partitioning based solely on a single indicator fails to fully illuminate the underlying structures of traffic interactions due to the limited information each factor provides. It is plausible to surmise that an adept integration of these indicators will lead to improved traffic partitioning results, as it allows for efficient merging and fusion of the complementary and compatible information each indicator presents.

Problem 2: ***How to create a dynamic traffic partitioning model capable of adapting to the evolving temporal characteristics of maritime traffic?***

Maritime traffic constitutes a dynamic and time-variable system, where ships and their interactions undergo continuous changes influenced by short-term fluctuations and long-term trends (Li et al., 2022a; Rong et al., 2022). For example, the constant influx and departure of ships within the maritime regulatory region, and the emergence or termination of interactions among them over time, introduce dynamic variables (see Fig. 1(b)). These unpredictable dynamic features considerably impact the consistency and stability of traffic clusters over time, presenting substantial challenges to the uninterrupted and effective application of risk management strategies reliant on traffic clusters. In practical maritime scenarios, maneuvers like collision avoidance often entail a time-consuming process. Therefore, ensuring the temporal smoothness of executing traffic control measures necessitates dynamic traffic partitioning capable of accommodating the temporal change in maritime traffic. Essentially, an optimal traffic partitioning strategy should factor in current interactions among ships alongside recent historic partitioning results.

Unfortunately, maritime traffic partitioning remains an underdeveloped area. No existing work, to the authors' best knowledge, has successfully incorporated both the multi-attribute interactions among ships and the evolutionary patterns of ship traffic simultaneously. Hence, devising a comprehensive traffic partitioning approach that can address these two complexities is a valuable yet challenging task.

Motivated by the challenges highlighted above, this study seeks to develop a systematic maritime traffic partitioning approach that fulfills two primary objectives: 1) the production of traffic clusters with guaranteed conflict connectivity, spatial compactness, and motion convergence simultaneously, and 2) the generation of stable partitioning outcomes that exhibit reduced sensitivity to the temporal evolution of ship traffic. The key to accomplishing these goals lies in the reliable characterization of the complex spatiotemporal dependence among ships and the effective incorporation of the multi-attribute ship interactions and traffic evolution into the traffic partitioning process. To achieve these, three interaction measure indicators, including conflict criticality, spatial distance, and approaching rate, are collectively used. To ensure the adaptability of these interaction indicators to traffic scenarios in complex waters, the conflict detection model has to integrate ship motion dynamics and uncertainty, whereas spatial distance and approaching rate measurements should leverage maritime trajectory knowledge mining to account for water geographical characteristics. Based on this foundation, a three-layer graph is modeled, where each layer shares the same set of nodes (i.e., ships) but is associated with different edges (i.e., different interactions between ships). Extending the spectral clustering framework to traffic partitioning involves employing a jointly smooth spectrum to maximize the synthesis of interaction information from all layers and utilizing priori knowledge of historic partitioning structures to handle interaction variation among ships evolving with time. It is essential to note that the traffic partitioning framework involves several hyperparameters to control and balance preferences for different graph layers. These parameters need to be addressed rigorously. Consequently, an adaptive hyperparameter selection method is designed to assist in searching for the optimal traffic partitioning results under various traffic scenarios. The success of this work will provide new insights into simplifying the comprehension of the whole traffic situation and guiding ship collision risk management. It will make a significant contribution towards shaping operational services for intelligent traffic safety surveillance and collision control for future hybrid traffic involving autonomous ships. Specifically, the main contributions are summarised as follows:

1. Three interaction measure metrics, demonstrating distinct functionalities in supporting effective and reliable traffic partitioning (see Section 3.1), are collectively introduced. These indicators effectively characterize the relationship between ships from different angles and precisely quantify the multi-ship spatiotemporal interactions under the impact of ship motion dynamics, uncertainty, and limited water topography in complex waters.
2. A semi-supervised spectral regularization model is proposed for the first time to simultaneously accommodate the multiple interactions among ships and priori knowledge from historic partitioning structures. It can ensure both the multi-layer graph information fusion and temporal smoothness of traffic partitioning results, making it promise to support traffic situation interpretability and aiding the ongoing execution of risk control strategies.
3. An adaptive hyperparameter selection strategy is designed to automatically determine the optimal traffic partitioning results under diversified traffic scenarios. In comparison with traditional methods using globally fixed hyperparameters based on sensitivity analysis or grid search, the proposed strategy can explicitly account for the hierarchical priorities among different objectives based on the decision-makers' preferences, as well as adapt to the dynamic changes in traffic situations.

The remainder of this paper is organized as follows. Section 2 presents an overview of the research pertaining to maritime traffic safety. Section 3 explains the developed traffic partitioning methodology, including interaction measure indicators, semi-supervised spectral regularization algorithm, and hyperparameter selection strategy. The case analyses, model validation, and discussions and implications are demonstrated in Section 4. Section 5 outlines the conclusion and future directions.

2. Literature review

Maritime traffic safety encompasses a variety of specialized themes. Key areas of focus include ship collision risk assessment and estimation, ship collision avoidance, and ship traffic partitioning. A detailed overview of each topic follows.

2.1. Ship collision risk assessment and prediction

The assessment and prediction of ship collision risks are an integral aspect of MSA, given their crucial role in providing a quantitative basis for understanding traffic situations and delivering prompt collision alerts to enhance anti-collision decision-making. It is a well-established field backed by an extensive body of literature. Detailed reviews have been encapsulated in scholarly review papers (Cao et al., 2023; Du et al., 2020; Huang et al., 2020). Typically, these studies can be grouped into two predominant modeling paradigms: micro-level and macro-level.

Micro-level approaches encompass ship domain-based methods and Closest Point of Approach (CPA)-based methods as their primary components. Ship domain-based methods examine and predict collision risks by evaluating violations or overlaps of the encountering ships' domain areas. These approaches are particularly focused on identifying influencing factors (e.g., ship-related parameters, environmental situations, and condition and knowledge of ship drivers), defining domain shapes (e.g., circular, elliptical, quaternion, and risk-based domains), and specifying applied methodologies (e.g., knowledge-based, empirical, and analytical techniques) (Liu et al., 2021; Szlapczynski and Szlapczynska, 2017). Notably, advanced ship domain models have demonstrated their utility in various aspects, excelling in pinpointing high-risk collision zones, quantitatively assessing waterway capacities, and investigating the interplay between potential collision candidates and historic accident databases. However, the deployment of a domain model for collision risk prediction necessitates integration with trajectory prediction methods due to its inherent constraints in anticipating future ship movements.

Conversely, the CPA-based technique offers a straightforward approach to identifying imminent collision scenarios by leveraging parameters like Distance to CPA (DCPA) and Time to CPA (TCPA). These parameters, derived from the ships' current position, speed, and course with the assumption that ships will sail linearly during a foreseeable horizon, establish both the proximity and timing of the ships' convergence. In this realm of investigation, prevalent studies have concentrated on the synthesis of multiple indicators such as DCPA, TCPA, distance, relative bearing, environmental condition, and ship maneuverability (Fang et al., 2018; Gil et al., 2022; Zhao et al., 2016), of which the primary approaches involve Analytical Hierarchical Process (Zhao et al., 2016), Fuzzy theory (Lee and Rhee, 2001), Multilayer Perceptron (Ahn et al., 2012), and Support Vector Machine (Gang et al., 2016). Currently, the majority of commercial systems employ this type of model for collision detection due to its simplicity and pragmatic applicability when ships maintain a constant speed (Xiao et al., 2020). Nevertheless, its accuracy deteriorates when deployed in complex encountering scenarios involving dynamic ship maneuvers. Instances like turns and speed adjustments become imperative due to the constrained water topography. Therefore, advancements are requisite to enhance its practical utility, entailing the exploration of inherent spatiotemporal dynamic attributes ingrained within ship movements.

Diverging from micro-level risk modeling, the macro-level analysis of collision risk, approached from a global perspective, has received substantial attention in response to the escalating intricacies of contemporary traffic conditions. A multitude of studies have endeavored to craft sophisticated models aimed at comprehending the entirety of maritime traffic situations through the lens of "traffic complexity". Notable instances encompass introducing a traffic complexity model to quantify the congestion levels and collision hazards (Wen et al., 2015), designing a complexity predictive analytics method to unveil the holistic extent of traffic irregularity and unpredictability (Zhang et al., 2022b), and constructing a multi-stage and multi-topology complexity model to enable the exploration of the traffic topological and evolutionary characteristics (Xin et al., 2022b). Noteworthy is the application of complex network theory within these frameworks, which exhibits substantial prowess in elucidating the interplay and interdependence among different ships within a regional traffic milieu. While these models greatly enhance the perceptual capabilities of maritime operators regarding a given regional traffic scenario, they encounter challenges in capturing intricate local conflict patterns, such as the collision risk of multi-ship encounters within a sub-region. Hence, there is an escalating demand for advanced models capable of extracting and capturing high-risk multi-ship encounters across a whole regulatory area based on their complicated spatiotemporal interactions.

In summary, a substantial and continuous research emphasis remains directed towards the assessment and estimation of collision risks in maritime traffic. However, considerable efforts should be placed on two pivotal aspects. Firstly, the scarcity of studies addressing the impact of ship movement dynamics and uncertainties on collision risk prediction is evident. Most existing research relies on a stringent assumption that ships will adhere to linear trajectories or that future paths are definitively known, disregarding uncertainties. However, the accurate prediction of a ship's forthcoming trajectory is fraught with challenges arising from the ship's dynamic behavior, navigational intentions, and external factors like winds and ocean currents. In response, some studies have sought to construct robust and resilient risk evaluation models, such as probabilistic conflict detection methods to integrate the influence of trajectory dynamic and uncertainty (Park and Kim, 2016; Xin et al., 2021) and time-varying collision risk measurement models to account for the uncertain nature of ship maneuvers (Li et al., 2022b). These approaches show promise in enhancing the precision of spatiotemporal collision risk assessment. Secondly, encounters involving multiple ships are commonplace and inherently riskier within complex waters. The estimation of collision risk in maritime navigation scenarios is profoundly linked to interdependent conflicts among multiple ships, rather than just pairwise collision risks. Consequently, a heightened focus on the collective behaviour of multiple ships within their spatiotemporal interactions is paramount. Although numerous endeavours have been dedicated to estimating collision risks for multiple ships (Liu et al., 2019a; Zhang et al., 2019b), the pursuit of extracting complicated multi-ship encounters from a given regional scenario remains scanty. Thus, the incorporation of spatiotemporal dynamics and the interactive effects among multiple ships presents a pivotal and favorable direction to develop effective techniques to estimate collision risk within local traffic clusters.

2.2. Ship collision avoidance

Conflict resolution stands as the ultimate phase in the spectrum of collision prevention. In the past few decades, a multitude of collision avoidance methods have emerged from research efforts within the maritime traffic field (Huang et al., 2020; Wang et al., 2024; Zhang et al., 2021). These studies have exhibited remarkable efficacy in generating collision-free trajectories during encounters involving multiple ships, by taking into account a range of factors (e.g., COLREGs, ship maneuverability, and ship intentions (Liu et al., 2023a,b; Zhang et al., 2022a), employing diverse approaches (e.g., velocity obstacle (Cho et al., 2020), artificial potential field (Li et al., 2021), particle swarm optimization (Hu et al., 2019), and reinforcement learning (Wang et al., 2023, 2022)), and addressing varied optimization objectives (e.g., safety, path smoothness, time efficiency, and course/speed preference (Hu et al., 2019; Yu et al., 2021)). Despite the substantial volume of research dedicated to this topic, the predominant focus has been on developing collision avoidance models from a reactive and separate perspective. In high-density traffic regions, ship operators often necessitate continual adjustments to their avoidance strategies to navigate away from emerging collision risks posed by newly approaching ships, thereby elevating operational demands and navigational hazards. Within this context, extensions and refinements have emerged to address multi-ship collision avoidance from the perspective of distributed coordination. These models devise systematic strategies that consider interactions among all nearby ships, employing methodologies like the Alternative Direction Method of Multiplier (ADMM) (Chen et al., 2018), Asynchronous Forward Bounding (AFB) (Li et al., 2019), Distributed Local Search Algorithm (DLSA), and Distributed Tabu Search Algorithm (DTSA) (Kim et al., 2017). However, a notable deficiency within these endeavors emerges as none have delved into determining the involvement of specific nearby ships within the coordination framework based on their interaction degree. A traffic partitioning model that thoroughly integrates interaction effects among multiple ships, grouping those with significant interactions into one cluster, presents both a crucial and promising avenue. It can assist in establishing the constraint boundary, i.e., identifying neighboring ships that require consideration during conflict avoidance actions, for conflict resolution strategies, which in turn forms the foundation for more effective execution of distributed coordination tactics aimed at collision avoidance.

2.3. Ship traffic partitioning

As mentioned above, partitioning traffic to capture multi-ship encounters is pivotal in both enhancing regional traffic situation interpretability and directing ship collision risk control measures. Lately, there has been growing interest in identifying clusters of encountering ships using AIS-based trajectory data. Numerous studies (Xin et al., 2022a; Zhen et al., 2021, 2017) have utilized the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to identify clusters of encountering ships and exclude the less risky ones. Additionally, the study by Zhu et al. (2022) applied the complex network theory to delineate the spatiotemporal dynamics of multi-ship encounters. However, these approaches have faced criticism due to their inherent limitations, including overlooking or oversimplifying ship dynamics, focusing solely on traffic density, and struggling to identify traffic clusters with varied densities. Failing to rectify these shortcomings often undermines the accurate interpretation of genuine maritime traffic patterns and the reliable breakdown of regional collision risk. As such, current research on traffic cluster detection remains foundational, with numerous challenges yet to be addressed, including the following:

1. The impact of diverse traffic attributes, such as ship motion dynamics, uncertainty, and confined water topography, on the interrelationships between ships remains largely unexplored in traffic cluster detection. Clearly, neglecting these factors is problematic, as they are fundamental to investigating and characterizing the genuine interaction mechanisms between ships.
2. Many studies have focused solely on one form of interaction between ships, like distance. Such an approach provides an incomplete perspective on the interdependencies among encountering ships, failing to paint a holistic view. Drawing insights from multiple compatible and complementary interaction measurement indicators across multi-view feature spaces outperforms relying on just one, given that different indicators reflect distinct facets of ship-to-ship interactions.
3. Previous analyses of traffic clusters have concentrated on interactions between ships at singular moments, overlooking the critical traffic feature, i.e., temporal evolution. Such oversight can lead to the generation of fluctuating and inconsistent traffic clusters over time, impeding the sustained application of traffic cluster-based collision risk management strategies.
4. Determining the values of hyperparameters in clustering algorithms is an open and critical challenge in optimizing maritime traffic cluster detection. Adaptive hyperparameter search algorithms, rather than relying on empirical methods, show promise in facilitating effective traffic cluster detection across diverse traffic scenarios.

Recent progress in the road transportation area provides valuable insights for maritime traffic partitioning (Gu and Saberi, 2019; Saeedmanesh and Geroliminis, 2017, 2016). This is evident by their emphasis on the importance of simultaneously considering various interrelationships between network links and traffic evolutionary characteristics. As opposed to prevalent methodologies in the maritime traffic sector, these studies employed sophisticated graph-based clustering algorithms, such as Symmetric Nonnegative Matrix Factorization (SNMF), to analyze interconnections among different network links. Specifically, road network partitioning aims to segment the entire traffic network into spatially connected, compact-shaped, and homogeneous sub-areas based on metrics like density and link speed. For instance, Saeedmanesh and Geroliminis (2017) delved into dynamic road network partitioning, revealing its effectiveness in deconstructing traffic network complexity, pinpointing congestion zones, and tracing the emergence and resolution of congestion. However, there is ample room to more comprehensively integrate traffic characteristics into ship traffic partitioning, especially given the sector's traffic uniqueness. No existing literature presents a comprehensive approach considering the multi-attribute interactions between ships, the evolving characteristics of traffic, and the adaptive searching of hyperparameters in the clustering model simultaneously. Consequently, this study stands as a pioneering endeavour in maritime traffic partitioning, addressing all the aforementioned limitations to elevate intelligent traffic monitoring and management.

3. Maritime traffic partitioning methodology

Fig. 2 depicts the framework of the proposed approach for maritime traffic partitioning. This framework comprises three crucial modules. Initially, three measurement indicators, i.e., conflict criticality, spatial distance, and approaching rate, are introduced to measure the correlations and interactions among ships. These indicators encompass the impact of ship movement dynamics, uncertainty, and limited water topography to characterize distinct facets of a traffic scenario. Building upon this foundation, three similarity measure matrices are devised to facilitate the construction of graphs that elucidate the interrelationships between individual pairs of ships within a specified regional traffic scenario. Subsequently, a semi-supervised spectral regularization model is proposed to simultaneously incorporate both multi-layer graph information and historic priori knowledge. Within the algorithm, information from multiple interaction indicators is effectively fused using a spectral regularization process. This process ensures the smoothness of generated joint eigenvectors across all layer graphs. Meanwhile, prior knowledge derived from historic traffic partitioning structures is integrated into the algorithm to guide the generation of temporally stable traffic clusters. Finally, an adaptive hyperparameter search strategy is formulated to cater to the preferences of relevant users regarding different interaction indicators. This strategy dynamically seeks the Pareto front of optimal hyperparameters and allows the selection of the ultimate solution from this Pareto front under diverse traffic scenarios.

3.1. Similarity measure and models

Maritime traffic partitioning necessitates the design of similarity measures to describe interactions between pairs of ships within a designated water area. The interrelationships and interactions between ships can manifest in diverse manners. This study explores the utilization of the subsequent interaction indicators: conflict criticality, spatial distance, and approaching rate. These indicators unveil interactions between ship pairs from distinct perspectives, thereby aiding in grouping ships exhibiting high conflict connectivity, spatial compactness, and convergent motion into one cluster. The specifics of each indicator are expounded upon in the following:

1) Conflict criticality

Measuring conflict criticality stands as one of the most potent techniques for quantifying interactions between ships (Weng et al., 2012; Weng and Shan, 2015). Detecting traffic clusters based on this metric can unearth genuine traffic conflict patterns. A conflict is usually characterized as a situation wherein encountering ships would breach their minimum safety separation distance within a foreseeable time horizon (Hao et al., 2018). A probabilistic conflict detection model is employed in this study to quantify the conflict criticality between pairs of ships. This approach adapts to traffic scenarios encompassing ship spatiotemporal motion dynamics and uncertainty. The conflict criticality between ships is influenced by factors such as their ship domain size, spatial distance, and relative encountering angle. In line with the medium-term focus on conflict detection, a time horizon of 15 min is established, drawing from the research conducted by Bakdi et al. (2021). For a comprehensive understanding of the operational details of this employed approach, on can turn to the study conducted by Xin et al. (2021).

2) Spatial distance

Spatial distance is the most intuitive means of delineating interactions between ships (Cho et al., 2020; Zhang et al., 2015). Incorporating spatial distance into traffic partitioning allows the creation of spatially compact traffic clusters, easing the design and deployment of anti-collision management measures. However, conventional methods of calculating spatial compactness based on

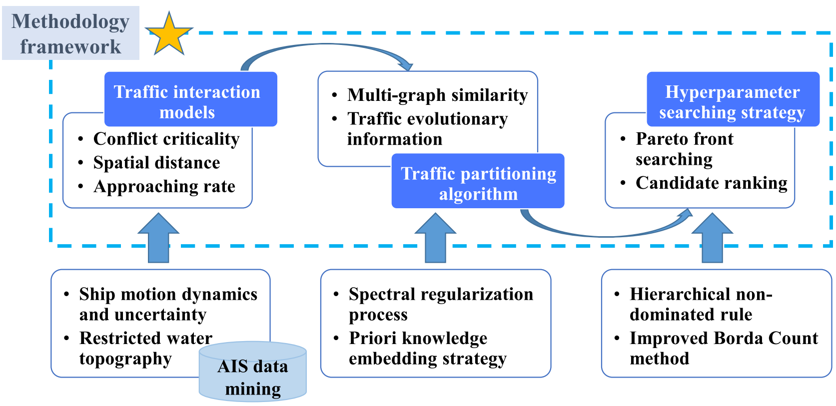


Fig. 2. Research framework.

Euclidean distance may fall short, as they disregard the influence of geographic water characteristics (Liu et al., 2019b; Zhen et al., 2017). For example, the adjacency of two ships in terms of Euclidean distance might not translate to accessibility in complex and restricted waterways. Obstacles such as small islands can impede direct paths, rendering them unattainable to each other. A practical solution to address this concern involves identifying the shortest path distance between pairs of ships under restricted water topography, achieved by referencing the maritime traffic route network to define their real spatial distance (Arguedas et al., 2017). In this context, a data-driven traffic route network developed by Xin et al., (2023b) is applied to ascertain the actual spatial distance between pairs of ships. When obstacles obstruct ship pairs, Dijkstra's algorithm is invoked to compute the shortest path distance based on the route network. Conversely, if unobstructed, the actual spatial distance is computed using the Euclidean metric.

3) Approaching rate

The approaching rate serves to unveil the relative motion between encountering ships. It plays a crucial role in understanding the potential dynamics of detected traffic clusters. In the study conducted by Wen et al. (2015), a relative motion factor was introduced to assess the complexity of traffic situations, which is subsequently widely utilized in the maritime realm (Sui et al., 2020; Zhang et al., 2019b). While this indicator performs admirably in open waters, it may yield inaccuracies in restricted water areas by neglecting the impact of water topography on actual spatial distances. Consequently, an indicator drawn from the air traffic filed (Wang et al., 2018), is designed to capture the changing rate of real spatial distances between ships. This is expressed in the following manner:

$$AR_t(i,j) = \frac{SD_t(i,j) - SD_{t-1}(i,j)}{SD_{t-1}(i,j)} \tag{1}$$

where $AR_t(i, j)$ signifies the approaching rate between ships i and j at time t. $AR_t(i, j) < 0$ suggests that the two ships are converging since their distance increases. Conversely, $AR_t(i, j) > 0$ indicates divergence between the two ships. It is noteworthy that the consideration of the approaching rate is relevant exclusively when these ships are engaged in an encountering situation, i.e., their distance is below a designated threshold (e.g., 6 nautical miles) (Cho et al., 2020).

Building upon the aforementioned indicators, the formulation of three distinct similarity measures becomes feasible. Since these indicators are not on the same scale, normalization becomes crucial to eliminate the impact of varying data magnitudes.

Regarding the similarity measure of conflict criticality between ships i and j (i.e., $W_t^{CC}(i,j)$), it can be directly equated to its corresponding conflict criticality $CC_t(i, j)$. This is due to the range of $CC_t(i, j)$ being confined within [0, 1]. A higher $CC_t(i, j)$ value implies a larger conflict criticality similarity, and vice versa.

Concerning the distance similarity measure, its definition unfolds as follows:

$$W_t^{SD}(i,j) = \begin{cases} \dfrac{SD_{thr} - SD_t(i,j)}{SD_{thr}}, & SD_t(i,j) < SD_{thr} \\ 0, & SD_t(i,j) \geq SD_{thr} \end{cases} \tag{2}$$

where SD_{thr} is denoted as a distance threshold beyond which a ship pair is deemed not involved in an encountered situation, and it is set at 6 nautical miles (Cho et al., 2020). As per Eq. (2), when the distance between ship pairs surpasses SD_{thr}, their distance similarity equates to 0. Additionally, a diminished spatial distance results in a heightened distance similarity, and vice versa.

Turning to the approaching rate similarity, its formulation assumes the subsequent expression:

$$W_t^{AR}(i,j) = \begin{cases} -AR_t(i,j), & AR_t(i,j) < 0, SD_t(i,j) < SD_{thr} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Table 1
Symbols and denotations.

Symbol	Definition and Explanation
G^m	the mth layer graph with vertex set V^m and edge set E^m
W^m	the adjacent similarity matrix for G^m
D^m	the degree matrix with $D^m(i,i) = \sum_{j=1}^{n} W_{ij}$
L^m	the Laplacian matrix with $L^m = D^m$-W^m
$W^m(i, j)$	the adjacent similarity between ships i and j for G^m
n	the ship count in G
k	the cluster count
L_{sym}	the symmetric graph Laplacian matrix
L_{rw}	the random walk graph Laplacian matrix
$\{C_i\}_{i=1}^k$	the cluster assignments
u_i	the ith eigenvector of L
U	the spectral embedding matrix
f_i	the scalar function associated with the graph G
Φ_{f_i}	the smooth function measuring the smoothness of the scalar function on G
$\overline{C}_i$	the complement of C_i in C
$Ncut_i$	the Normalised Cut of the ith similarity matrix
MD	the matching degree of partitioning results across successive moments

Eq. (3) resembles a classifier that segregates the relative motion of encountering ships into two categories: approaching and not approaching. When the two ships are in the process of moving apart, their approaching rate similarity stands at 0. Moreover, a larger value of $W_t^{AR}(i,j)$ signifies a more rapid convergence of the two ships.

The above-mentioned similarity measures encapsulate diverse dimensions of interaction information. The effective combination of these measures within the maritime traffic partitioning process undoubtedly contributes to enhanced traffic clustering. This will be expounded upon further below.

3.2. Semi-supervised spectral regularization model

Maritime traffic partitioning can be explained from a graph-cut perspective. It involves the creation of a graph where nodes denote individual ships and edges symbolize the neighboring relationships between ships. Given the multi-attribute interactions among ships, a multi-layer graph naturally takes form, where each layer shares the same node set while corresponding to distinct sets of edges (i.e., similarity matrices). Additionally, incorporating the evolutionary feature of traffic patterns into the graph-cut paradigm is pivotal to engendering clusters that are temporally consistent. Consequently, the traffic partitioning problem is transformed into grouping ships with high multi-graph interactions and high temporal consistency into one cluster.

A variety of symbols are employed to elucidate the traffic partitioning methodology. Therefore, a comprehensive compilation of notations and their respective meanings is presented in Table 1. Furthermore, symbols bearing a subscript 't' in the subsequent content signify the corresponding variable at the time 't'. This subsection begins with an introduction to the fundamental principles of spectral clustering, providing a well-defined mathematical framework adept at handling multi-layer graphs. Subsequently, an in-depth exploration of the spectral regularization process provides clarity on how interaction information from multi-layer graphs is harmonized. Following this, the elucidation pertains to the details of constructing and integrating prior knowledge concerning historic traffic partitioning structures into the traffic partitioning procedure. Finally, the hybrid traffic partitioning approach that seamlessly amalgamates both multi-graph interactions and historic priori information is unveiled.

3.2.1. Basic spectral clustering

Spectral clustering represents a widely embraced algorithm for addressing graph partitioning issues entailing optimization of diverse graph-based measures. In comparison to conventional clustering methodologies, it has gained substantial favor due to its straightforward implementation and commendable performance when partitioning graphs with complicated structures (Filippone et al., 2008; Shi and Malik, 2000). The foundational implementation of spectral clustering hinges on calculating the foremost k smallest eigenvalues and corresponding eigenvectors of the graph Laplacian matrix L. It is worth noting that the subsequent normalized versions of L find extensive application across various domains, which take the form as follows:

$$L_{sym} = D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}} \tag{4}$$

$$L_{rw} = D^{-1}(D-W) \tag{5}$$

where W denotes the adjacency matrix, D signifies the degree matrix with vertex degrees on the diagonal, L_{sym} retains the property of symmetry, while L_{rw} exhibits a close relationship with the random walk process on graphs (Von Luxburg, 2007). These two normalized versions are adopted in this study.

The process of implementing spectral clustering involves the following steps: 1) computing the eigenvectors of the normalized graph Laplacian matrix L_{sym} or L_{rw}; 2) constructing the spectral embedding matrix U; and 3) applying the k-means algorithm to the normalized U to determine cluster assignments. A concise outline of the algorithm is presented in Algorithm 1. This procedure has been substantiated as a potent mathematical framework (Ng et al., 2001). The inherent attributes of the graph Laplacian matrix enhance the underlying correlations and interactions among neighboring vertices. Consequently, it allows conventional clustering algorithms like the k-means algorithm to be effectively employed for clustering within the low-dimensional space demarcated by U.

Algorithm 1. Normalized spectral clustering.

Input: W: $n \times n$ weighted similarity matrix of graph G, k: the desired number of clusters.

Output: $\{C_i\}_{i=1}^{k}$: the clustering assignments.

1. Calculate the degree matrix D.
2. Calculate the normalized graph Laplacian matrix L_{sym} or L_{rw}.
3. Calculate the foremost k leading eigenvectors $u_1, u_2, \ldots, u_k$ of L_{sym} or L_{rw}.
4. Let $u_1, u_2, \ldots, u_k$ be the columns to form matrix $U \in \mathbb{R}^{n\times k}$.
5. Normalise the rows of matrix U to obtain the matrix U^*.
6. Let $y_i \in \mathbb{R}^k (i = 1, 2, \ldots, n)$ denotes the ith row of U^*, representing the ith vertex in the graph.
7. Apply the k-means algorithm to the y_i in $\mathbb{R}^k$ to obtain cluster assignments.

It is important to note that basic spectral clustering is applicable solely to single-layer graphs. However, the well-defined mathematical framework lends itself to flexible extensions for multi-layer networks. The subsequent subsection elaborates on the mechanism of generalizing basic spectral clustering, focusing on the creation of a unified spectral embedding matrix aimed at integrating information from multi-graph structures.

3.2.2. Multi-graph spectral clustering based on jointly smooth spectrum

In this study, a spectral regularization process is introduced to amalgamate information from multi-layer graphs. The fundamental concept underlying this process revolves around the creation of a series of joint eigenvectors exhibiting smoothness across all layer graphs, thereby preserving their distinct interrelation characteristics. Consider two-layer graphs, denoted as G^1 and G^2, sharing identical nodes. To find a jointly smooth spectrum across both graphs, an optimization model, as proposed by Dong et al. (2012), is employed. The optimization objective is detailed as follows:

$$\underset{f_i \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \|f_i - u_i\|_2^2 + \lambda \cdot \Phi_{f_i} \right\} \qquad j = 2, \cdots, n \tag{6}$$

where f_i: $V \to R$ represents a scalar function associated with the graphs, u_i is the ith eigenvector of the graph Laplacian matrix derived from G^1, the regularization parameter λ serves to balance the relative importance of the two layers, and $\Phi_{f_i} = f_i{}^T L_{sym}^2 f_i$ is a smooth function that measures the smoothness of the scalar function on G^2. The primary purpose of Eq. (6) is to determine a set of scalar functions f_i that are not only close to the u_i derived from G^1 but also exhibit sufficient smoothness on G^2. It aids in generating a set of f_i that display smoothness across both G^1 and G^2, forming what can be interpreted as a jointly smooth spectrum shared between the two graphs.

The analysis presented in Ng et al. (2001) indicates that the optimization problem outlined in Eq. (6) can be effectively addressed using the subsequent formula:

$$f_j^* = \frac{1}{\lambda} \left(L_{sym}^2 + \frac{1}{\lambda} I_n \right)^{-1} u_j \qquad j = 2, \cdots, n \tag{7}$$

It is important to note that the first eigenvector of the graph Laplacian matrix u_1 is a constant value (as detailed in Chung (1997)) and therefore does not present an optimization concern. Once a set of optimal joint eigenvectors f_j^* ($j = 2, 3, \cdots, n$) is derived using the formulation in Eq. (7), the foremost k leading eigenvectors are employed to construct a joint spectral embedding matrix. Subsequently, the k-means algorithm can be applied to attain a robust clustering solution across both layers.

Additionally, the above optimization framework doesn't treat G^1 and G^2 as equals due to their distinct roles. More specifically, G^1 is employed for eigen-decomposition to gain eigenvectors, while G^2 serves as a graph structure for regularization support. Typically, the layer graph bearing richer structural information is designated as G^1. Moreover, the optimization framework can be readily generalized from two-layer to multi-layer graphs. Specifically, it commences by seeking the optimal jointly smooth spectrum between G^1 and G^2, achieved by maximizing their mutual information. Subsequently, the jointly smooth spectrum of the initial two layers is combined with subsequent chosen layers, aiming to derive a new smooth spectrum. This iterative process continues until all layer graphs are integrated into the framework. A detailed algorithmic depiction of the spectral regularization process is outlined in Algorithm 2.

Algorithm 2. Spectral regularization-based clustering for multi-layer graphs

Input: W^i ($i = 1, 2, \cdots, m$): $n \times n$ weighted similarity matrix of the ith layer graph G^i, λ_i ($i = 1, 2, \cdots, m$-1): the regularization parameters, k: the desired number of clusters.

Output: $\{C\}_{c=1}^k$: the clustering assignments.

1. **For** G^1, calculate the degree matrix D^1 and graph Laplacian L_{rw}^1.
2. Calculate the first k leading eigenvectors $u_1, u_2, \cdots, u_k$ of L_{rw}^1.
3. Use $U \in \mathbb{R}^{n \times k}$ as the matrix including $u_1, u_2, \cdots, u_k$ as columns.
4. **For** $i = 2$ to m **do**
5. **For** $j = 2$ to k **do**
6. use Eq. (7) to address the spectral regularization issue of u_j related to L_{sym}^i.
7. replace u_j with f_j^* to form the new spectral embedding matrix U.
8. **End**
9. **End**
10. Normalise the rows of matrix U to obtain the matrix U^*.
11. Let $y_i \in \mathbb{R}^k (i = 1, 2, \ldots, n)$ denotes the ith row of U^*, representing the ith vertex in the graph.
12. Apply the k-means algorithm to the y_i in $\mathbb{R}^k$ to obtain cluster assignments.

3.2.3. Prior information construction and embedding

One significant challenge within maritime traffic partitioning involves the limitation of solely conducting static clustering using traffic interaction information from a singular snapshot. It yields inconsistent and unstable clusters across consecutive time intervals, especially during pronounced fluctuations in traffic conditions. This inconsistency impedes the continuous implementation of collision mitigation strategies that rely on a continuous traffic cluster framework. An effective solution involves the application of semi-supervised based algorithms, which incorporate historic prior knowledge into the traffic partitioning procedure. This method ensures that the clustering structure remains relatively consistent with neighboring historical snapshots, averting excessive deviation.

Numerous studies have compellingly demonstrated that incorporating prior knowledge into clustering algorithms constitutes a powerful strategy for enhancing their performance. For example, prior knowledge has exhibited success in various domains, including community detection (Ma and Dong, 2017) and document clustering (Salah et al., 2017). However, no analogous efforts have been made in applying priori knowledge to the maritime traffic partitioning domain. This study harnesses priori knowledge to guarantee the temporal consistency of clustering structures over defined time intervals.

To establish the priori information, the historic traffic clusters $\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}$ that were executed conflict mitigation strategies from the preceding time moment t-1 are extracted, where k'_{t-1} represents the number of clusters within $\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}$. Subsequently, a weighting matrix Z_t with dimensions $n_t \times n_t$ is formulated, where n_t signifies the number of ships present at time t. The elements in Z_t are designed as follows:

$$Z_t(i,j) = \begin{cases} 1, & \text{if } x_i \text{ and } x_j \text{ have the same class label in } \left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}} \\ -1, & \text{otherwise} \end{cases} \tag{8}$$

The fundamental premise underlying this approach is that individuals having identical labels within $\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}$ are predisposed to cluster jointly at current time t. Conversely, those without the same labels are inclined to be allocated to distinct clusters. Furthermore, the constructed priori information matrix Z_t is incorporated into the similarity matrices through the inclusion of a regularization term, as illustrated below:

$$\widehat{W}_t^m = W_t^m + \gamma Z_t \tag{9}$$

where γ is a weighting parameter governing the relevant significance of the priori information. The incorporated priori information within Z_t serves to modify the weights of graph edges, so that guide the ship-pairs with the same class labels in $\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}$ to merge together and those with different labels to separate from each other. In this manner, the historic priori knowledge is explicitly enforced into the traffic partitioning process.

3.2.4. Hybrid traffic partition model

Based on the outlined models above, a pivotal concern arises: how to simultaneously integrate both the multi-graph structural information and historic priori information into the maritime traffic partitioning process. The semi-supervised model effectively incorporates priori knowledge by modifying the weights among graph edges, thereby amplifying or attenuating their interactions. Consequently, integrating priori information becomes a straightforward addition to the spectral regularization process outlined in Algorithm 2. Given the additional prior knowledge, the weighted adjacency matrices are initially updated using Eq. (9). Then, the spectral regularization-based clustering algorithm is executed to fuse the multi-graph information. This approach ensures the comprehensive incorporation of both elements, rendering the resultant amalgamation meaningful. Fig. 3 visually illustrates the hybrid traffic partitioning framework, elucidating how the amalgamation of multi-layer graph information and prior knowledge is seamlessly achieved.

Within the hybrid framework, careful consideration must be given to the prioritization of the three embedded similarity matrices within the traffic partitioning process, as this directly shapes user preferences across distinct layers. In this study, the "spatial distance" graph layer takes precedence in combining with the "conflict criticality" graph layer, as it contains richer structural information. Following the completion of this initial fusion, the third layer "approaching rate" is further integrated to yield the ultimate solution. Additionally, it is important to note that the hyperparameters within the hybrid model exert a substantial impact on partitioning performance, a question to be explored in the subsequent subsection.

3.3. Adaptive selection of hyperparameters

According to the above hybrid traffic partitioning model, there exist three hyperparameters that need to be determined, i.e., λ_1, λ_2, and γ. The proper employment of these hyperparameters holds paramount importance in enhancing the model's clustering performance. Traditional studies are prone to empirically specify globally fixed hyperparameters across all clustering scenarios. However,

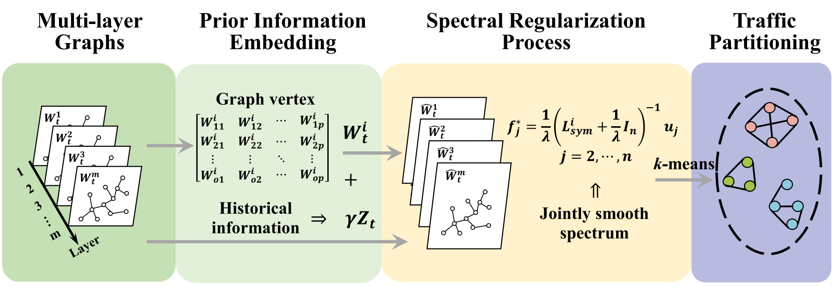


Fig. 3. Framework of the semi-supervised spectral regularization algorithm.

the realm of real-world maritime traffic scenarios exhibits notable variability, rendering constant hyperparameters inadequate for accommodating the changes in traffic interaction structures. A promising approach is to automatically and adaptively select the hyperparameters through an optimal mechanism, which however remains an open question. This study therefore addresses this concern by devising an adaptive searching procedure to dynamically fine-tune the hyperparameters.

In the pursuit of optimal hyperparameters across diverse traffic scenarios, an appropriate evaluation metric, i.e., optimization objective, is paramount to quantify the efficacy of traffic partitioning. In this context, the widely accepted graph-based metric, Normalised Cut (*Ncut*) (Shi and Malik, 2000), is utilized to evaluate the traffic partitioning performance. This metric takes into account both inter-cluster similarity and intra-cluster similarity, aiding in finding proper cluster results. Its formulation is encapsulated within the following equation:

$$Ncut = \sum_{o=1}^{k} \frac{assoc(C_o, \overline{C}_o)}{assoc(C_o, C)} \tag{10}$$

where $assoc(C_o, C) = \sum_{i \in C_o, j \in C} W(i,j)$, and $\overline{C}_o$ represents the complement of C_o. A smaller value of *Ncut* means that the clustering outcomes exhibit elevated intra-similarity alongside diminished inter-similarity, consequently indicating a superior partitioning quality.

Additionally, another important objective is to ensure the temporal smoothness of generated clusters across successive time snapshots. In this regard, the Jaccard coefficient (Greene et al., 2010) is utilized to measure the matching degree between traffic clusters across adjacent temporal intervals. Its formulation is as follows:

$$MD\left(\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}, \left\{C_{t,p}\right\}_{p=1}^{k_t}\right) = \frac{1}{k'_{t-1}} \sum_{o=1}^{k'_{t-1}} \frac{|C_{t-1,o} \cap C_{t,p}|}{|C_{t-1,o} \cup C_{t,p}|} \tag{11}$$

where k'_{t-1} represents the number of clusters that need to keep consistency in the subsequent time moments, and $C'_{t-1,o}$ and $C_{t,p}$ denote the *o*th clusters within $\left\{C'_{t-1,o}\right\}_{o=1}^{k'_{t-1}}$ at time *t*-1 and the *p*th cluster within $\left\{C_{t,p}\right\}_{p=1}^{k_t}$ at time *t*, respectively. This metric yields a score within the range of [0, 1]. A higher value of *MD* implies a more pronounced consistency between traffic clusters in successive time moments.

Optimal hyperparameter selection entails a typical multi-objective optimization problem, where numerous sub-objective functions often counteract one another. There exist no solutions that can universally minimize or maximize all objectives (i.e., $Ncut_{CC}$, $Ncut_{SD}$, $Ncut_{AR}$, and *MD* within a given traffic scenario) in unison. To tackle this complexity, a stochastic random sampling method is employed, birthing a candidate pool. Within this pool, the search for the Pareto front of optimal solutions is undertaken. This involves identifying a set of Pareto-efficient solutions that remain non-dominated by any other solution within the set. For instance, assume that there are two optimal solutions $P(x)$ and $P(y)$, if $P(x)$ surpasses $P(y)$ for a minimum of one objective, while $P(y)$ outperforms $P(x)$ for at least another objective, they are mutually non-dominated by each other. Importantly, it is worth noting that for the effective integration of prior information into the evolutionary traffic partitioning process, *MD* takes priority over the remaining objectives (i.e., $Ncut_{CC}$, $Ncut_{SD}$, and $Ncut_{AR}$). Consequentially, *MD* serves as the primary objective, with the other objectives considered as secondary equals. This hierarchical consideration yields the benefit of delivering augmented consistency and stability in clustering outcomes over time.

Referring to the set of Pareto optimal solutions generated, the ultimate traffic partitioning solution must undergo further selection. In real-world scenarios, maritime operators and ship navigators normally accord precedence to conflict criticality over spatial distance or approaching rate when evaluating ship interactions. Consequently, the choice of the ultimate solution from the Pareto optimal set should be made in a rigorous manner that aligns with the inclinations and preferences of the practitioners.

In light of this, a voting aggregation method known as the Borda Count method (Zwicker, 1991) is employed and enhanced for the purpose of ranking candidates within the Pareto optimal set. This method is viewed as a consensus-based method and is recognized for its ability to select a candidate with the broadest acceptance across all voters (Liu et al., 2018). The fundamental concept underlying this method is as follows: each candidate is assigned a specific number of points based on their position in the rankings provided by each voter. The candidate amassing the highest total points emerges as the winner. In this case, all solutions present in the Pareto optimal set are treated as candidates, and the *Ncuts* metrics serve as the voters. Assuming there are 10 candidates, if Candidate No. 2 is ranked first in terms of $Ncut_{CC}$ by a voter, it will accrue 10 points from that voter. Eventually, each candidate receives a cumulative score derived from all the voters' rankings.

In comparison to the original Borda Count method, two enhancements are undertaken. Firstly, in order to account for the preferences of decision-makers, the voters (*Ncuts*) are subjected to varying weights, departing from equal weighting. Specifically, $Ncut_{CC}$ is accorded greater significance compared to the others. In this scheme, the weights assigned to $Ncut_{CC}$, $Ncut_{SD}$, and $Ncut_{AR}$ are 3, 2, and 1, respectively. Secondly, the practice of assigning scores based solely on ranking positions proves inadequate in revealing the differences among distinct candidates concerning different objectives. To address this, a normalization function is employed to derive scores for each candidate concerning different objectives, as follows:

$$\delta(Ncut_i^s) = \frac{Ncut_i^{max} - Ncut_i^s}{Ncut_i^{max} - Ncut_i^{min}} \tag{12}$$

where $\delta\left(Ncut_i^s\right)$ represents the score assigned to the sth candidate in terms of its ith objective, $Ncut_i^{max}$ and $Ncut_i^{min}$ are the maximum and minimum values recorded for the ith objective on the Pareto front. Eq. (12) is used to eliminate the impact of variations in data dimensions. Subsequently, the aggregate score for the sth candidate is determined through the subsequent formula:

$$\delta^s = \sum_{i=1}^{3} w_i \cdot \delta\left(Ncut_i^s\right) \tag{13}$$

where w_i signifies the weight assigned to the ith objective. In this way, it provides a principled manner in aggregating the contributions of different interaction metrics, culminating in the determination of the ultimate optimal solution.

4. Applications and case study results

4.1. Research area and data preprocessing

To assess the viability and effectiveness of the proposed traffic partitioning methodology, the Hub areas of Ningbo-Zhoushan Port is selected as the testing site (refer to Fig. 4). This port stands out as an exceptional deep-water harbor, boasting one of the world's highest cargo throughputs, resulting in intensely concentrated maritime traffic. The port encompasses over 620 production berths, which include around 170 large-scale berths designed for ships exceeding 10,000 tons, along with more than 100 super large deep-water berths capable of accommodating ships over 50,000 tons. The constrained geographical characteristics, diverse ship types, wide-ranging ship movement patterns, and sophisticated environmental conditions prevalent in this region contribute to elevated risks and unpredictability in maritime navigation. Therefore, it is a typical complex water area that presents formidable challenges to maritime supervisors engaged in effective MSA.

For this study, the AIS messages spanning from November 1, 2018, to November 30, 2018, are collected from the AIS database provided by the Maritime Bureau in China, focusing on a geographical area bounded by longitudes 121°52′E-122°22′E and latitudes 29°43′N-30°02′N. Prior to conducting experimental analyses, it is imperative to preprocess the data to eliminate potential errors and noise. To achieve this, a two-step procedure is implemented. First, the noise removal method introduced by Kang et al. (2018) is employed, followed by the application of the trajectory consistency verification method proposed by Zhao et al. (2018), facilitating the reconstruction of accurate and clean traffic trajectories. Subsequently, to account for varying message transmission frequencies among different ships, a linear interpolation technique introduced by Zhang et al. (2019a) is employed. This facilitated the creation of snapshots of the maritime traffic situation, enabling the extraction of dependable and comprehensive information relevant to any given moment for the purposes of traffic scenario analysis and evaluation.

4.2. Application case analysis

Within this subsection, two practical experimental analyses are undertaken to vividly illustrate the prowess and practicality of the proposed methodology. Firstly, the methodology's adeptness in integrating multi-attribute interactions among ships and decomposing

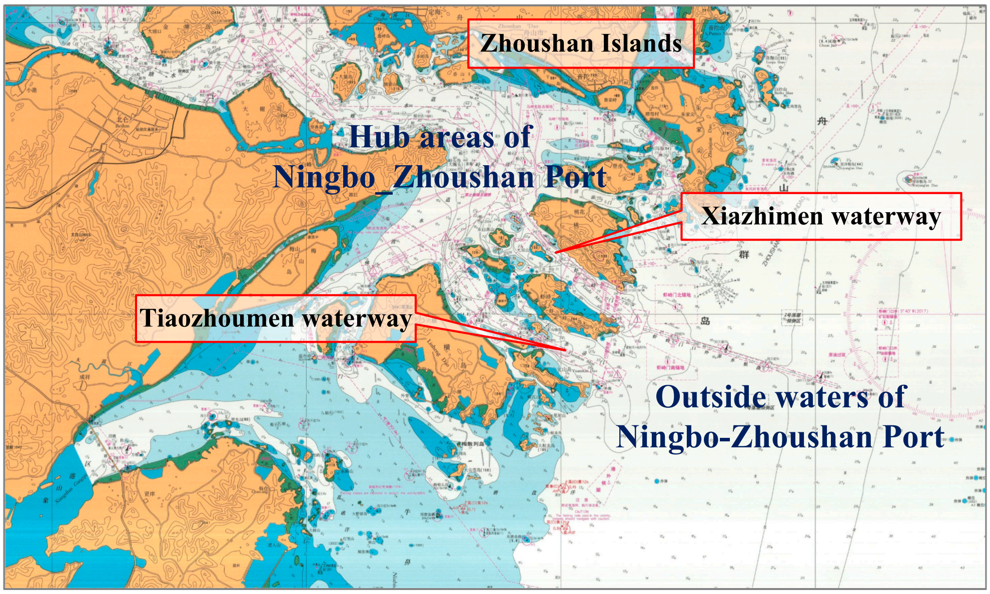


Fig. 4. Illustration of Ningbo-Zhoushan Port, China.

the regional traffic complexity is showcased via traffic partitioning for a specific traffic scenario at a given instant. Secondly, the methodology's capacity to leverage prior knowledge from historic traffic clustering structures is analyzed through a traffic evolution scheme, which highlights how it assists in guiding ship collision risk control.

4.2.1. Case analysis without considering traffic evolutionary characteristics

Fig. 5 provides an illustrative depiction of the ship traffic partitioning results at a specific snapshot, without considering traffic evolutionary characteristics. In Fig. 5(a), the visual representation of ship interactions in the form of graphs is exhibited. In this representation, ships are denoted by red points, and interactions between ships are depicted by connecting lines. The figure underscores the presence of nested interactions among multiple ships in terms of each metric. Hence, the need to perform traffic partitioning becomes paramount to enable an in-depth analysis of the complex spatiotemporal interactions among multiple ships, shifting the focus from interactions between individual ship pairs. Fig. 5(b) provides an insight into the *Ncut* values stemming from the traffic partitioning results when the number of input clusters *k* equates to 5, 10, and 15. Within this representation, red points and blue points correspond respectively to the highest and lowest values within each objective on the Pareto frontier generated. Notably, the proximity of each *Ncut* value to its associated lower bound is evident, signifying the successful trade-off and fusion of multi-attribute interaction information among ships. The visualized outcomes of traffic partitioning are unveiled in Fig. 5(c)-(e). An important observation emerges: ships within the same cluster exhibit dense interconnectivity, while those across different clusters demonstrate sparse connections. This holds true across different interaction metrics and the cluster count *k*. The robustness of the model in effectively partitioning the entirety of ship traffic into clusters that exhibit concurrent traits of high conflict connectivity, spatial compactness, and relative convergence is thus convincingly demonstrated.

Additionally, Fig. 6 presents the statistical attributes of each generated cluster, including number of ships, cumulative conflict criticality, spatial density, and the count of converging ship pairs within each cluster. Based on these statistics, it becomes straightforward to discern clusters marked by high conflict criticality, spatial density, or motion convergence. For instance, as depicted in Fig. 6(b), Clusters 4 and 8 exhibit high cumulative conflict criticality, while Clusters 2 and 5 in Fig. 6(f) are characterized by high spatial density. This approach stands in contrast to conventional methods of implementing MSA from a broader regional perspective. The proposed model effectively decomposes the whole traffic complexity, enabling maritime operators to pinpoint sub-regions warranting intensified surveillance. This, in turn, supports intelligent traffic monitoring and management, offering a sophisticated means of enhancing maritime safety and control.

To enhance comprehension of the adaptive hyperparameter search strategy, Fig. 7 presents the Pareto front, comprising a set of feasible solutions, alongside the ultimate optimal solution chosen when $k = 15$. This visualization highlights the inherent trade-offs among the three objectives, as there are no solutions that can simultaneously achieve minimal values for all three objectives. Furthermore, the selected optimal solution exhibits outstanding performance across the three objectives, particularly excelling in the

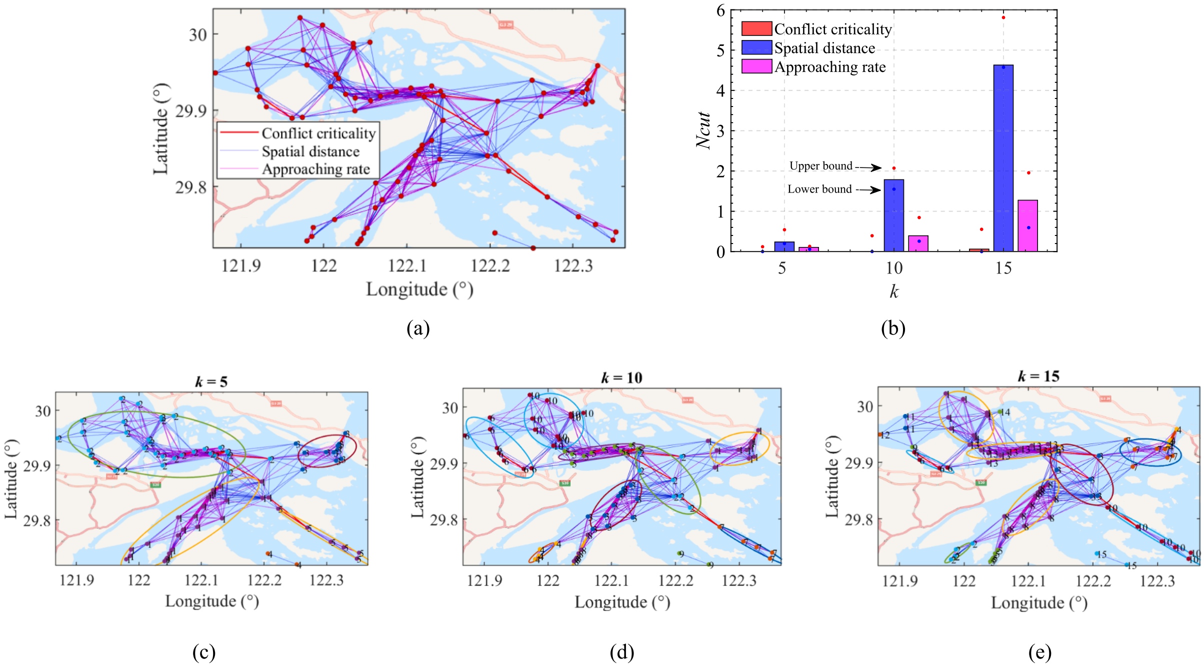


Fig. 5. Ship traffic partitioning results without prior information. (a) visualization of ship traffic network; (b) *Ncut* values across varying input cluster numbers; (c)-(f) visualization results of traffic partitioning for cluster numbers 5, 10, and 15.

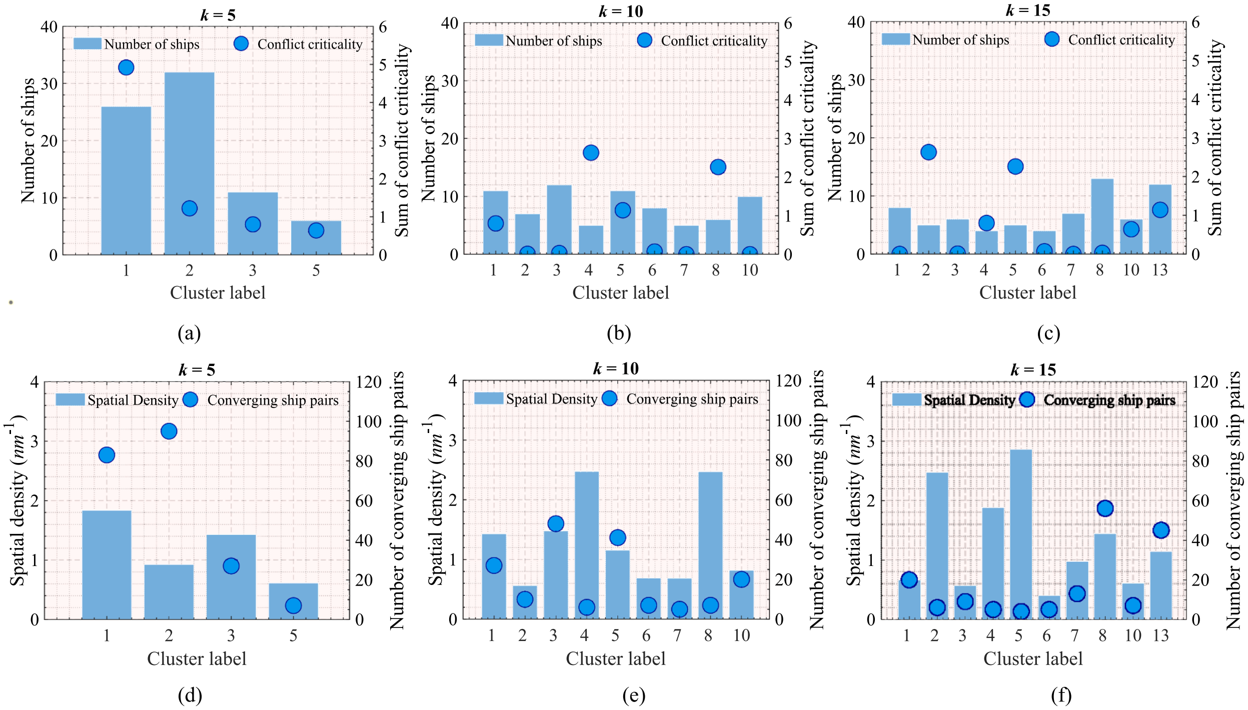


Fig. 6. Feature statistics of each cluster, including the number of ships, cumulative conflict criticality, spatial distance, and number of converging ship pairs. Note that only clusters with more than 2 ships are retained for analysis.

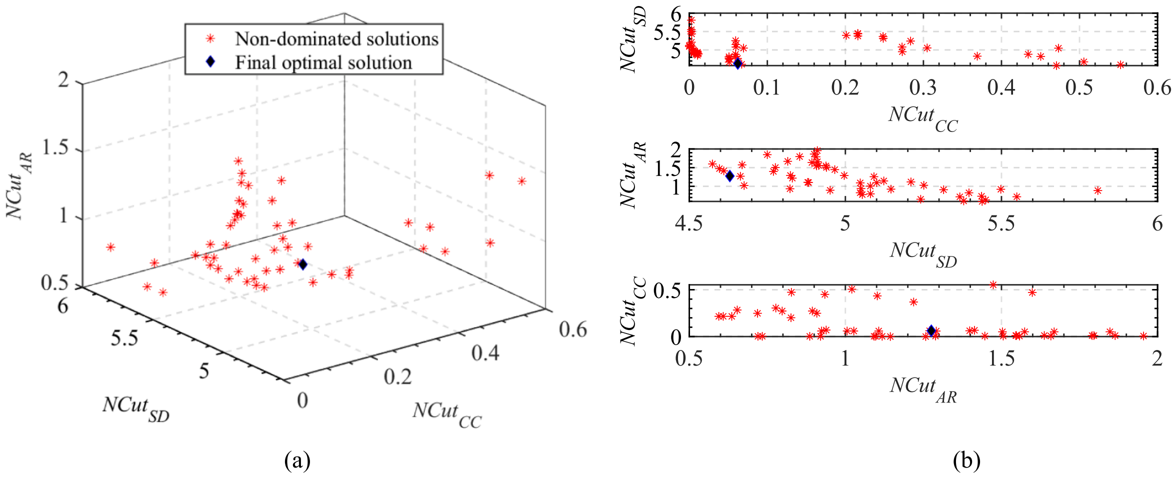


Fig. 7. Pareto front of optimal objectives in three traffic interaction graphs. (a) A three-dimensional plot of the Pareto front; (b) a two-dimensional plot of the Pareto front.

first and second objectives. This underscores the effectiveness of the hyperparameter search strategy in not only achieving a favorable balance among these objectives but also in catering to users' distinct preferences for different objectives.

4.2.2. Case analysis considering traffic evolutionary characteristics

Maritime traffic partitioning to identify high-risk traffic clusters constitutes the initial phase of risk management. Subsequently, measures to mitigate collision risks within these identified clusters become imperative. Fig. 8 elucidates the process of generating temporally stable traffic clusters through the utilization of prior knowledge, thereby facilitating cluster-based conflict resolution. In Fig. 8(a)-(c), the visual representations of static traffic partitioning at three different time instances, t, $t+1$, and $t+2$, are presented. Notably, Cluster 1 in Fig. 8(a) exhibits the highest cumulative conflict criticality. It is unveiled that the static partitioning approach

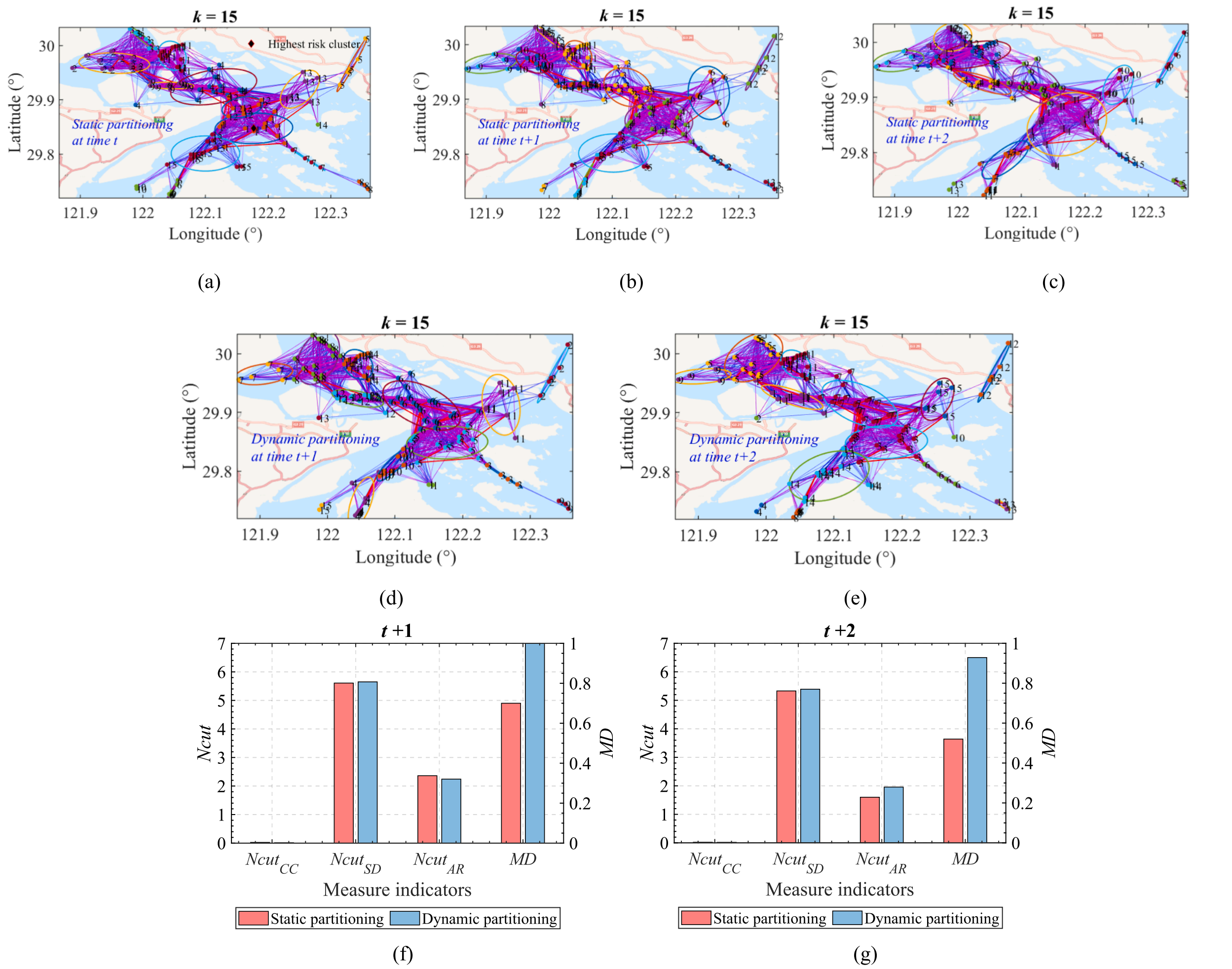


Fig. 8. Ship traffic partitioning results when incorporating prior knowledge derived from historic traffic partitioning structures. (a)-(c) Visualisation of static traffic partitioning at time t, $t+1$, and $t+2$; (d)-(e) visualisation of dynamic traffic partitioning at time $t+1$ and $t+2$; (f)-(g) metric comparisons between static and dynamic traffic partitioning at time $t+1$ and $t+2$.

inadequately preserves the stability of Cluster 1 over time (see Fig. 8(b)-(c)). This phenomenon primarily stems from the dynamic nature of ship traffic, i.e., the evolving interactions among ships. By contrast, imposing priori knowledge constructed based on the historic Cluster 1 in Fig. 8(a) into the traffic partitioning process guides the identification of temporally consistent traffic clusters, exemplified by Cluster 5 in Fig. 8(d) and Cluster 3 in Fig. 8(e). Statistical analysis in Fig. 8(f)-(g) demonstrates that dynamic traffic partitioning yields a higher *MD* value, offset by a negligible increase in *Ncuts*. The resultant stable traffic clusters, exhibiting minimal temporal drift, form a foundational requirement for the uninterrupted execution of cluster-based collision risk control strategies, playing a crucial role in enhancing maritime traffic safety. In practical applications of dynamic traffic partitioning, the determination of execution time steps should be contingent on the real-time decision-making needs of ship navigators or traffic operators. For instance, this determination could be based on the time required for a ship to perform an evasive maneuver within a traffic cluster. Thus, the model presented herein lays a robust groundwork for future research endeavors focused on devising pragmatic control strategies to effectively mitigate risks in densely trafficked water areas.

4.3. Methodological comparison and validation

Comparison and validation of methodologies are pivotal for the pragmatic implementation of the proposed approach. As such, three distinct performance comparisons have been conducted to thoroughly assess and ascertain the efficacy of the pivotal modules within the proposed approach. Initially, the focus is on demonstrating the approach's superiority in balancing and utilizing the contemporary information extracted from multiple graphs, in contrast to classical graph-based clustering algorithms. Subsequently, the examination extends to evaluating the approach's effectiveness in incorporating historic priori knowledge to yield traffic clusters that remain temporally stable. Lastly, validation is centered on the approach's robustness in flexibly searching for optimal hyper-parameters across diverse traffic scenarios.

4.3.1. Performance evaluation in incorporating multi-graph information

To examine the effectiveness of the proposed method in effectively integrating information across multiple graph layers, a comparative analysis is conducted against alternative graph-based clustering algorithms. Specifically, two commonly adopted graph-based clustering algorithms and their respective variations are employed for this comparative evaluation. These competing algorithms, along with certain implementation details, are summarised in the following:

- SC: The Spectral Clustering (SC) algorithm serves as a fundamental baseline approach. It has found extensive applications in extracting maritime traffic patterns and detecting abnormal behaviors (Li et al., 2022a; Liu et al., 2022a). However, its utilization in the context of maritime traffic cluster detection remains relatively less explored.
- SNMF: Symmetric Non-negative Matrix Factorization (SNMF) stands as a resilient matrix factorization technique tailored for non-negative and symmetric matrices. It accomplishes clustering assignments by approximating the nonnegative lower rank of a graph adjacent matrix. Its adeptness in adapting to various similarity matrices and delivering superior graph cuts has led to its successful application in road transportation network partitioning (Saeedmanesh and Geroliminis, 2016).

The aforementioned two algorithms typically utilize individual similarity matrices as inputs, limiting their ability to perform clustering with multiple graphs. Consequently, their respective variations are subsequently introduced:

- SC-AVE: SC takes the average of multiple similarity matrices as input.
- SNMF-AVE: SNMF takes the average of multiple similarity matrices as input.

It should be noted that other categories of clustering algorithms, such as density-based and prototype-based clustering, are omitted from this comparison due to their inherent incompetence in handling graph-based networks. Despite being employed for detecting traffic clusters within the maritime domain (Liu et al., 2019b; Xin et al., 2022a), these algorithmic types predominantly concentrate on the dataset itself rather than the interactions between data points.

Fig. 9 presents an average performance comparison among various clustering algorithms when applied to various traffic scenarios. Among these competitors, SC and SNMF excel in achieving high-quality graph cuts for individual graph layers but falter in simultaneously considering other interaction information. For instance, employing SC with WW_{cc} as input yields $Ncut_{cc}$ values of 0, but considerably high values for $Ncut_{SD}$ and $Ncut_{AP}$. Conversely, SC-AVE and SNMF-AVE, by combining multiple similarity matrices rather than relying solely on each layer's graph, manage to achieve a certain level of trade-off across multiple interactions. However, their enhancements are modest in contrast to the proposed model. This can be primarily attributed to the straightforward averaging of information from diverse graphs in SC-AVE and SNMF-AVE, which fails to harness the full potential of the rich and heterogeneous information within each layer.

In more detail, the proposed approach maintains its superiority over SC-AVE across all three metrics. When compared to SNMF-AVE, it demonstrates superior clustering outcomes concerning $Ncut_{cc}$ and $Ncut_{AP}$, whilst exhibiting comparatively weaker performance in $Ncut_{SD}$. This can be attributed to the chosen weight preferences for different objectives in the proposed approach. For instance, when assigning weights of 1, 2, and 1 to $Ncut_{CC}$, $Ncut_{SD}$, and $Ncut_{AR,}$ respectively, the resultant optimal values for these indicators in the proposed model are 0.027, 2.64, and 0.80. Similarly, when weights of 1, 3, and 1 are assigned, the optimal values

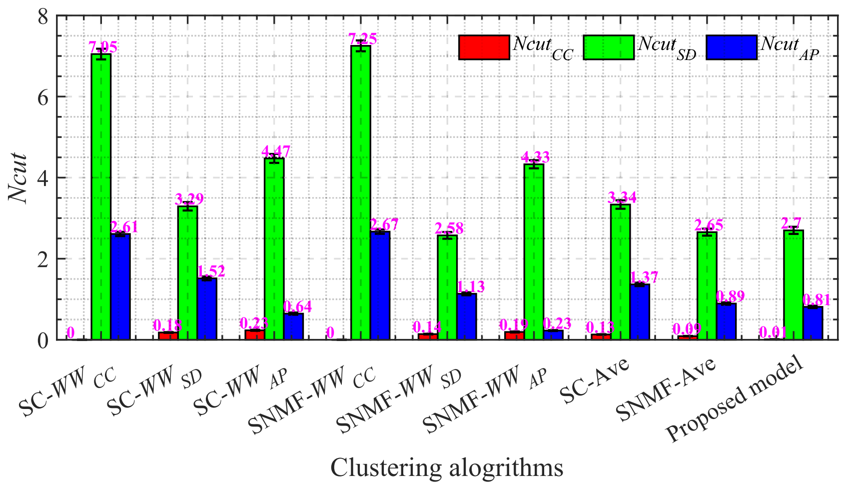


Fig. 9. Comparisons between the proposed model and other graph-based clustering algorithms. The error bars represent the 5% significant interval.

become 0.036, 2.59, and 0.88, respectively. Under these weight configurations, the proposed method overwhelmingly outperforms SNMF-AVE across all three metrics. Moreover, it is noteworthy that each *Ncut* value within the proposed model closely approximates the best results attained through clustering of each individual layer graph. For example, the $Ncut_{CC}$ value of 0.1 in the proposed model closely mirrors the 0 in SC-WW_{CC}. This phenomenon arises from the proposed approach of fusing multiple graph layers to maximize compatibility and complementary information. As a result, the proposed model demonstrates a superior and competitive edge in terms of comprehensively considering all three interaction indicators.

4.3.2. Performance evaluation in incorporating priori evolutionary information

Fig. 10 presents an in-depth comparison between static and dynamic traffic partitioning. The dynamic method integrates traffic partitioning structures from earlier moments, specifically 1, 3, and 5 min prior. Two primary insights are evident. Firstly, the dynamic model exhibits a slight decline in clustering quality (i.e., *NCuts*) when compared to the static model. However, it showcases a marked enhancement in temporal smoothness (i.e., *MD*). This highlights its ability to remain faithful to both the current traffic characteristics and the historic traffic structures. Secondly, when integrating older historic traffic partitioning structures, the dynamic partitioning approach gains a more notable edge in temporal smoothness. Remarkably, the *MD* values remain close to 1, irrespective of how dated the referenced historic traffic structures are. This consistency can be credited to the adaptive hyperparameter selection strategy, wherein *MD* is prioritized over *NCuts* to ensure that prior knowledge is optimally utilized during the evolutionary traffic partitioning process. These observations show the dynamic model's prowess in adeptly leveraging prior knowledge derived from historic traffic partitioning structures.

4.3.3. Performance evaluation in adaptive selection of hyperparameters

Fig. 11 offers a detailed comparison between the proposed adaptive hyperparameter selection strategy and the globally fixed hyperparameter strategy. Within the figure, red points denote the optimal traffic partitioning results derived from 50 sets of randomly chosen globally fixed hyperparameters. In contrast, the blue points illustrate the partitioning outcomes when different objective weights are applied in the adaptive hyperparameter selection approach. It is evident that compared to the globally fixed hyper-parameter strategy, all *Ncuts* from the adaptive model reside in the bottom-left quadrant of the coordinate system. Particularly, the adaptive approach exhibits substantial improvements in $Ncut_{cc}$. This performance boost arises from the model's ability to autono-mously determine optimal hyperparameters across diverse traffic conditions. Conversely, the globally fixed hyperparameter strategy often falls short in accurately considering the distinct traffic features of varying scenarios. These observations emphasize the critical need for the flexible adjustment of hyperparameters in response to changing traffic situations. Furthermore, by rigorously designating different weights for each *Ncut* within the adaptive strategy, a trade-off solution reflecting user preferences becomes readily attainable. Thus, the experimental results demonstrate the strategy's resilience and versatility in adapting to a range of traffic situations and accounting for the hierarchical priorities among different objectives.

4.4. Discussions and implications

Maritime traffic partitioning represents an emerging area of research fraught with complex challenges. This study pioneers the development of an evolutionary multi-graph traffic partitioning methodology designed to identify temporally stable, conflict-connected, spatially compact, and motion-convergent traffic clusters in complex waters. In essence, 'complex waters' typically denote ports or channels confronted with complex traffic scenarios marked by dynamic ship motions, high density, and limited geographical features, among other elements (Fang et al., 2018; Yu et al., 2019). The proposed methodology showcases robust adaptability to such contexts through: 1) the integration of a probabilistic conflict detection model addressing traffic dynamics and uncertainty; 2) the utilization of a traffic route network for precise spatial distance measurement, even in the presence of obstacles like islands; and 3) the formulation of a traffic partitioning model that prioritizes multi-ship interactions in heavily trafficked areas. These advantageous attributes of the proposed methodology facilitate its seamless customization and deployment across diverse port and waterway environments.

Comprehensive experiments are undertaken using AIS data from one of the largest ports in the world, Ningbo-Zhoushan Port, to thoroughly assess the methodology's efficacy. Throughout this process, a meticulous AIS data preprocessing procedure is employed,

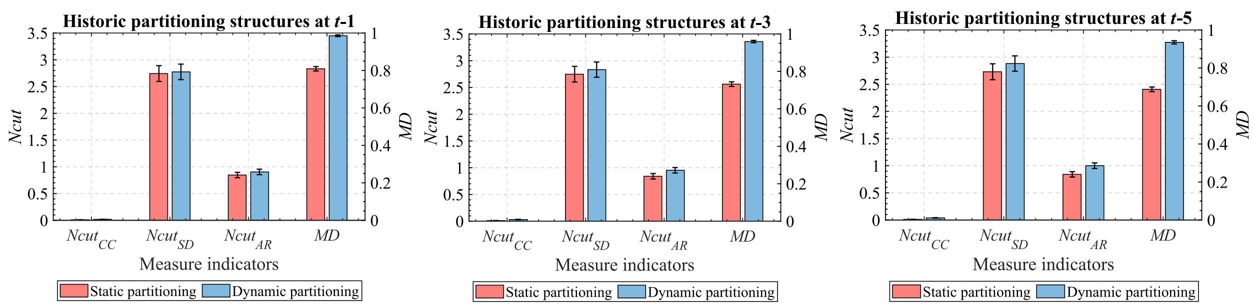


Fig. 10. Comparisons between static traffic partitioning and dynamic traffic partitioning. The error bars represent the 5% significant interval.

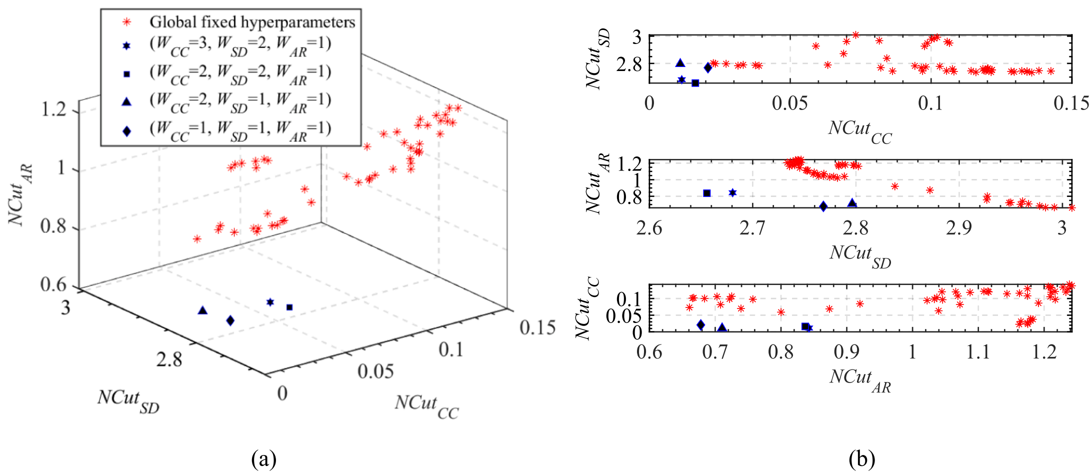


Fig. 11. Comparisons between the globally fixed hyperparameter strategy and the adaptive hyperparameter selection strategy. (a) Three-dimensional plot of the *Ncuts*; (b) two-dimensional plot of the *Ncuts*.

encompassing data cleaning, denoising, and interpolation, aimed at ensuring data reliability and experimental validity. This step is pivotal, as lower-quality AIS data or the loss of essential information during preprocessing may introduce biases or inaccuracies into subsequent analyses, potentially compromising the credibility of the research findings. With well-preprocessed data, various interactions among ships, including conflict criticality, spatial distance, and approaching rate, can be accurately determined, thereby facilitating reliable and effective traffic partitioning thereafter. Thorough case analyses and rigorous model validations demonstrate the effectiveness of the proposed methodology in overcoming the limitations outlined in Section 2.3. This includes achieving a balance and utilization of contemporary information extracted from multiple interactions among ships, integrating historical priors to establish temporally stable traffic clusters, and dynamically searching for optimal hyperparameters across various traffic scenarios. These findings underscore the significant theoretical and practical contributions of the proposed methodology. Key insights and implications gleaned from experimental outcomes and analytical discussions are drawn as follows:

This study offers valuable insights for enhancing intelligent maritime surveillance. While conventional research has generally centered on collision risks between individual ship pairs or within an entire regional maritime area, the proposed methodology excels in breaking down regional traffic complexity. This is crucial for identifying high-risk sub-regions, a contribution significantly elevating maritime intelligent perception capabilities. Specifically, the proposed approach enhances the comprehension of traffic patterns by dividing maritime traffic into several distinct sub-clusters. Based on this, maritime authorities can acquire comprehensive insights into the spatial distribution of collision risks and identify specific sub-regions necessitating prompt risk mitigation measures. Moreover, the developed models function as reliable tools for tactical traffic management. For example, in high-risk traffic conditions, the complex behaviors among ships can lead to a surge in potential conflicts. Maritime surveillance operators, under such pressure, often face daunting challenges in monitoring tasks and devising effective measures to alleviate traffic complexity. A common but suboptimal solution is to limit the number of ships entering the surveillance area, which compromises traffic efficiency and can inadvertently overlook localized conflicts, endangering navigational safety. In contrast, the proposed methodology aids in the identification of key influential traffic clusters, thereby enabling strategic and proactive risk control measures. Targeted strategy deployment and maneuvering guidelines for these pivotal clusters can significantly reduce the traffic complexity from a regional perspective. This boosts the operational capabilities of maritime operators dealing with high-complexity scenarios and opens up new avenues for balancing complexities across adjacent sectors. Consequently, this work supports maritime operators in advancing operational safety management without requiring either a compromise on traffic flow efficiency or additional investment in infrastructure upgrades. It holds the promise of integration into intelligent maritime surveillance systems, thereby contributing to the development of smart ports.

This study holds significant implications for advancing maritime navigation autonomy, a field where collision avoidance is a cornerstone for realizing a robust autonomous shipping agenda. Traditional collision avoidance research has predominantly focused on localized traffic analysis, often overlooking the complexities of maritime traffic in broader regional contexts. Meanwhile, ships predominantly employ reactive, individualized evasive maneuvers to avoid collisions. As maritime activities in complex port waters become increasingly complicated and autonomous ship development advances, this reactive approach to collision risk is becoming progressively inadequate. In today's complex port waters, ships often engage in multi-conflict situations. The decisions made by one ship not only affect nearby ships but can also influence the actions of distant ships entangled in nested conflicts. The high-complexity traffic situation creates challenges for ship navigators to understand, adapt, and eventually devise a resilient anti-collision measure. Accordingly, it will force ship navigators to take continuous maneuvers to avoid potential new collisions, which increases the likelihood of making erroneous decisions and even endangers maritime traffic safety. As a recommendation, this study offers valuable insights by examining potential conflicts among ships in various adjoining waters from a global traffic network perspective. By shifting the focus from localized traffic analysis to a global/regional approach, it proposes a proactive strategy to manage ship anti-collision practices, enabling coordinated efforts across multiple vessels to control traffic risks effectively. Additionally, the innovative traffic partitioning methodology segregates highly spatially correlated ships into cohesive clusters. This clustering facilitates more precise, boundary-defined collision avoidance planning. Importantly, the new dynamic traffic partitioning strategy takes into account evolving maritime traffic characteristics, offering a resilient foundation for the ongoing implementation of cluster-based risk control measures. Furthermore, the developed dynamic traffic partitioning approach illuminates the exploration of co-evolving behaviors among multiple ships by identifying temporally stable traffic clusters across consecutive time steps from historical AIS trajectory data. It can generate a wealth of complex, time-evolving multi-ship encounter scenarios grounded in real-world conditions. This serves as an invaluable resource for rigorously testing and validating emerging intelligent traffic management strategies, including autonomous decision-making and integrated assessments of physical and cyber risks. Such extensive testing and validation are crucial milestones prior to the practical implementation of these innovative techniques. Hence, the methodology introduced demonstrates significant applicability within the realm of autonomous maritime collision risk control, laying a robust groundwork for the continued coexistence of both manned and autonomous vessels.

This study offers considerable significance for improving both the reliability and competitiveness of ports. Implementation of the proposed methodology within real-world maritime network systems has the potential to substantially mitigate collision risks and traffic incidents, alleviate port congestion and delays, and bolster overall port traffic resilience. Clearly, safety and efficiency are paramount concerns for port stakeholders, including ship owners, port operators, and shippers. Ports offering top-tier safety and efficiency services are likely to attract increased direct investment. Moreover, the enhanced maritime efficiency that the proposed approach offers could lead to notable reductions in fuel costs and emissions, contributing to the development of environmentally responsible, modern, and intelligent ports. Therefore, the developed methodology stands as a foundational instrument for elevating a port's competitiveness and sustainability.

While the proposed approach has shown superior performance in addressing traffic partitioning issues compared to traditional models, it still possesses limitations that require careful consideration. One such limitation arises from the necessity to adequately consider the influence of traffic topological attributes on collision risk evaluation for traffic clusters. This study identifies and monitors high-risk traffic clusters in terms of metrics such as traffic density and cumulative conflict criticality. However, relying solely on these indicators may not fully identify the intricate interactions among multiple ships. Additional metrics like clustering coefficients and motif-based measures are indispensable for comprehensively assessing the collective interactions among multiple vessels. Hence, there is a pressing need for the development of advanced models capable of thoroughly evaluating multiple ship risks from various topological perspectives. Such models would enable maritime operators to prioritize attention to traffic clusters more accurately.

5. Conclusions and future research

Advancing maritime surveillance and management is a critical element in the rise of smart ports and autonomous ships. This study introduces a comprehensive approach for partitioning regional ship traffic to improve MSA and mitigate collision risks. The methodology offers several key advantages: 1) it incorporates three distinct interaction metrics to quantitatively assess the spatiotemporal relationships between ships; 2) it utilizes a semi-supervised spectral regularization model to process both multi-graph information and traffic evolutionary information; and 3) it features an adaptive hyperparameter selection model to adapt to the changeable traffic situations and accommodate maritime operators' preferences for different interaction indicators. Thorough case analyses validate the methodology's effectiveness in simplifying regional traffic complexity, identifying high-risk sub-regions, and enhancing strategic maritime safety measures. Furthermore, rigorous model comparisons and validations indicate the superior or competitive edge of the proposed methodology over widely accepted baseline methods and state-of-the-art techniques. Hence, this study offers valuable insights into real-time maritime traffic monitoring and management. It also promises to significantly impact the intelligent coordination of future mixed-traffic scenarios involving autonomous ships.

Future research could delve into the following promising avenues. Firstly, the collective effect of multiple independent conflicts on the overall collision risk within traffic clusters deserves to be investigated. This exploration will enable the accurate identification of critical traffic clusters that require prioritized attention in risk control measures. Secondly, the extraction of stably dynamic traffic clusters across consecutive time snapshots from historical AIS data could provide invaluable insights into the evolving interactions among multiple ships. This would facilitate a deeper understanding of cooperative behaviors in multi-ship encountering scenarios over time. Thirdly, the development of a conflict resolution framework that adeptly balances both intra-cluster and inter-cluster collision risks is another area deserving in-depth study. Such a framework could equip maritime surveillance operators with the means to craft multi-level strategies specifically aimed at hierarchical risk mitigation.

CRediT authorship contribution statement

Xuri Xin: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Kezhong Liu:** Conceptualization, Writing – review & editing, Project administration, Resources, Supervision. **Huanhuan Li:** Formal analysis, Writing – review & editing. **Zaili Yang:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Data availability

The authors do not have permission to share data.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant No. 52031009) and a European Research Council project under the European Union's Horizon 2020 research and innovation program (TRUST CoG 2019 864724).

References

Ahn, J.-H., Rhee, K.-P., You, Y.-J., 2012. A study on the collision avoidance of a ship using neural networks and fuzzy logic. Appl. Ocean Res. 37, 162–173.
Arguedas, V.F., Pallotta, G., Vespe, M., 2017. Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring. IEEE Trans. Intell. Transp. Syst. 19, 722–732.
Bakdi, A., Glad, I.K., Vanem, E., 2021. Testbed scenario design exploiting traffic big data for autonomous ship trials under multiple conflicts with collision/grounding risks and spatio-temporal dependencies. IEEE Trans. Intell. Transp. Syst. 22, 7914–7930.
Bogusławski, K., Nasur, J., Li, J., Gil, M., Wróbel, K., Goerlandts, F., 2022. A cross-domain scientometric analysis of situational awareness of autonomous vehicles with focus on the maritime domain. IEEE Access 10, 50047–50061.
Cao, Y., Wang, X., Yang, Z., Wang, J., Wang, H., Liu, Z., 2023. Research in marine accidents: A bibliometric analysis, systematic review and future directions. Ocean Eng. 284, 115048.
Chen, L., Hopman, H., Negenborn, R.R., 2018. Distributed model predictive control for vessel train formations of cooperative multi-vessel systems. Transp. Res. Part C Emerg. Technol. 92, 101–118.
Cho, Y., Han, J., Kim, J., 2020. Efficient COLREG-compliant collision avoidance in multi-ship encounter situations. IEEE Trans. Intell. Transp. Syst. 23, 1899–1910.
Chung, F.R.K., 1997. Spectral graph theory. American Mathematical Soc.
Dong, X., Frossard, P., Vandergheynst, P., Nefedov, N., 2012. Clustering with multi-layer graphs: a spectral perspective. IEEE Trans. Signal Process. 60, 5820–5831.
Du, L., Goerlandt, F., Kujala, P., 2020. Review and analysis of methods for assessing maritime waterway risk based on non-accident critical events detected from AIS data. Reliab. Eng. Syst. Saf. 106933.
Fang, Z., Yu, H., Ke, R., Shaw, S.-L., Peng, G., 2018. Automatic identification system-based approach for assessing the near-miss collision risk dynamics of ships in ports. IEEE Trans. Intell. Transp. Syst. 20, 534–543.
Filippone, M., Camastra, F., Masulli, F., Rovetta, S., 2008. A survey of kernel and spectral methods for clustering. Pattern Recogn. 41, 176–190.
Gang, L., Wang, Y., Sun, Y., Zhou, L., Zhang, M., 2016. Estimation of vessel collision risk index based on support vector machine. Adv. Mech. Eng. 8, 1687814016671250.
Gil, M., Kozioł, P., Wróbel, K., Montewka, J., 2022. Know your safety indicator–A determination of merchant vessels Bow Crossing Range based on big data analytics. Reliab. Eng. Syst. Saf. 220, 108311.
Greene, D., Doyle, D., Cunningham, P., 2010. Tracking the evolution of communities in dynamic social networks, in: 2010 International Conference on Advances in Social Networks Analysis and Mining. IEEE, pp. 176–183.
Gu, Z., Saberi, M., 2019. A bi-partitioning approach to congestion pattern recognition in a congested monocentric city. Transp. Res. Part C Emerg. Technol. 109, 305–320.
Hao, S., Zhang, Y., Cheng, S., Liu, R., Xing, Z., 2018. Probabilistic multi-aircraft conflict detection approach for trajectory-based operation. Transp. Res. Part C Emerg. Technol. 95, 698–712.
Hu, L., Naeem, W., Rajabally, E., Watson, G., Mills, T., Bhuiyan, Z., Raeburn, C., Salter, I., Pekcan, C., 2019. A multiobjective optimization approach for COLREGs-compliant path planning of autonomous surface vehicles verified on networked bridge simulators. IEEE Trans. Intell. Transp. Syst. 21, 1167–1179.
Huang, Y., Chen, L., Chen, P., Negenborn, R.R., van Gelder, P., 2020. Ship collision avoidance methods: state-of-the-art. Saf. Sci. 121, 451–473.
Kang, L., Meng, Q., Liu, Q., 2018. Fundamental diagram of ship traffic in the Singapore Strait. Ocean Eng. 147, 340–354.
Kim, D., Hirayama, K., Okimoto, T., 2017. Distributed stochastic search algorithm for multi-ship encounter situations. J. Navig. 70, 699–718.
Lee, H.-J., Rhee, K.P., 2001. Development of collision avoidance system by using expert system and search algorithm. Int. Shipbuild. Prog. 48, 197–212.
Li, H., Jiao, H., Yang, Z., 2023. 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.
Li, S., Liu, J., Negenborn, R.R., 2019. Distributed coordination for collision avoidance of multiple ships considering ship maneuverability. Ocean Eng. 181, 212–226.
Li, H., Lam, J.S.L., Yang, Z., Liu, J., Liu, R.W., Liang, M., Li, Y., 2022a. Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. Transp. Res. Part C Emerg. Technol. 143, 103856.
Li, M., Mou, J., Chen, P., Rong, H., Chen, L., van Gelder, P., 2022b. Towards real-time ship collision risk analysis: an improved R-TCR model considering target ship motion uncertainty. Reliab. Eng. Syst. Saf. 226, 108650.
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.
Li, Y., Li, H., Zhang, C., Zhao, Y., Yang, Z., 2024a. Incorporation of adaptive compression into a GPU parallel computing framework for analyzing large-scale vessel trajectories. Transp. Res. Part C Emerg. Technol. 163, 104648.
Li, H., Xing, W., Jiao, H., Yang, Z., Li, Y., 2024b. Deep bi-directional information-empowered ship trajectory prediction for maritime autonomous surface ships. Transp. Res. Part E Logist. Transp. Rev. 181, 103367.
Li, H., Yang, Z., 2023. Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships. Transp. Res. Part E Logist. Transp. Rev. 176, 103171.
Liu, H., Tian, Z., Huang, A., Yang, Z., 2018. Analysis of vulnerabilities in maritime supply chains. Reliab. Eng. Syst. Saf. 169, 475–484.
Liu, Z., Wu, Z., Zheng, Z., 2019a. A cooperative game approach for assessing the collision risk in multi-vessel encountering. Ocean Eng. 187, 106175.
Liu, Z., Wu, Z., Zheng, Z., 2019b. A novel framework for regional collision risk identification based on AIS data. Appl. Ocean Res. 89, 261–272.
Liu, K., Yuan, Z., Xin, X., Zhang, J., Wang, W., 2021. Conflict detection method based on dynamic ship domain model for visualization of collision risk Hot-Spots. Ocean Eng. 242, 110143.
Liu, H., Liu, Y., Li, B., Qi, Z., 2022a. Research on detection mechanism of vessel abnormal behavior based on immune genetic spectral clustering algorithm. Ocean Eng. 266, 113099.
Liu, J., Zhang, J., Yan, X., Soares, C.G., 2022b. Multi-ship collision avoidance decision-making and coordination mechanism in mixed navigation scenarios. Ocean Eng. 257, 111666.
Liu, R.W., Liang, M., Nie, J., Lim, W.Y.B., Zhang, Y., Guizani, M., 2022c. Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime Internet of Things. IEEE Trans. Netw. Sci. Eng.
Liu, K., Wu, X., Zhou, Y., Yuan, Z., Xin, X., Zhang, J., 2023a. Coordinated multi-stage and multi-objective optimization approach for ship collision avoidance decision-making. Ocean Eng. 287, 115888.
Liu, K., Wu, X., Zhou, Y., Yuan, Z., Yang, X., Xin, X., Zhuang, S., 2023b. A conflict cluster-based method for collision avoidance decision-making in multi-ship encounter situations. Ocean Eng. 288, 116038.
Ma, X., Dong, D., 2017. Evolutionary nonnegative matrix factorization algorithms for community detection in dynamic networks. IEEE Trans. Knowl. Data Eng. 29, 1045–1058.
Montewka, J., Manderbacka, T., Ruponen, P., Tompuri, M., Gil, M., Hirdaris, S., 2022. Accident susceptibility index for a passenger ship-a framework and case study. Reliab. Eng. Syst. Saf. 218, 108145.
Ng, A., Jordan, M., Weiss, Y., 2001. On spectral clustering: Analysis and an algorithm. Adv. Neural Inf. Process. Syst. 14.

Park, J., Kim, J., 2016. Predictive evaluation of ship collision risk using the concept of probability flow. IEEE J. Ocean. Eng. 42, 836–845.
Rong, H., Teixeira, A.P., Soares, C.G., 2021. Spatial correlation analysis of near ship collision hotspots with local maritime traffic characteristics. Reliab. Eng. Syst. Saf. 209, 107463.
Rong, H., Teixeira, A.P., Soares, C.G., 2022. Maritime traffic probabilistic prediction based on ship motion pattern extraction. Reliab. Eng. Syst. Saf. 217, 108061.
Saeedmanesh, M., Geroliminis, N., 2016. Clustering of heterogeneous networks with directional flows based on "Snake" similarities. Transp. Res. Part B Methodol. 91, 250–269.
Saeedmanesh, M., Geroliminis, N., 2017. Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. Transp. Res. Procedia 23, 962–979.
Salah, A., Ailem, M., Nadif, M., 2017. A way to boost semi-NMF for document clustering, in. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 2275–2278.
Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22, 888–905.
Sui, Z., Wen, Y., Huang, Y., Zhou, C., Xiao, C., Chen, H., 2020. Empirical analysis of complex network for marine traffic situation. Ocean Eng. 214, 107848.
Sun, Y., Yang, J., Song, Q., 2024. Modeling and application of ship traffic conflicts in nearshore waters with heavy traffic. Ocean Eng. 291, 116347.
Szlapczynski, R., Szlapczynska, J., 2017. Review of ship safety domains: Models and applications. Ocean Eng. 145C, 277–289.
Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416.
Wang, H., Song, Z., Wen, R., 2018. Modeling air traffic situation complexity with a dynamic weighted network approach. J. Adv. Transp. 2018.
Wang, C., Zhang, X., Yang, Z., Bashir, M., Lee, K., 2023. Collision avoidance for autonomous ship using deep reinforcement learning and prior-knowledge-based approximate representation. Front. Mar. Sci. 9.
Wang, W., Huang, L., Liu, K., Wu, X., Wang, J., 2022. A COLREGs-compliant collision avoidance decision approach based on deep reinforcement learning. J. Mar. Sci. Eng. 10, 944.
Wang, W., Huang, L., Liu, K., Zhou, Y., Yuan, Z., Xin, X., Wu, X., 2024a. Ship behavior pattern analysis based on multiship encounter detection. ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng. 10, 4023045.
Wang, W., Huang, L., Liu, K., Zhou, Y., Yuan, Z., Xin, X., Wu, X., 2024b. Ship encounter scenario generation for collision avoidance algorithm testing based on AIS data. Ocean Eng. 291, 116436.
Wang, C., Zhang, X., Gao, H., Bashir, M., Li, H., Yang, Z., 2024. Optimizing anti-collision strategy for MASS: A safe reinforcement learning approach to improve maritime traffic safety. Ocean Coast. Manag. 253, 107161.
Wen, Y., Huang, Y., Zhou, C., Yang, J., Xiao, C., Wu, X., 2015. Modelling of marine traffic flow complexity. Ocean Eng. 104, 500–510.
Weng, J., Meng, Q., Qu, X., 2012. Vessel collision frequency estimation in the singapore strait. J. Navig. 65, 207–221.
Weng, J., Shan, X., 2015. Ship collision frequency estimation in port fairways: a case study. J. Navig. 68, 602–618.
Xiao, Z., Fu, X., Zhang, L., Goh, R.S.M., 2019. Traffic pattern mining and forecasting technologies in maritime traffic service networks: a comprehensive survey. IEEE Trans. Intell. Transp. Syst. 21, 1796–1825.
Xiao, Z., Fu, X., Zhang, L., Zhang, W., Liu, R.W., Liu, Z., Goh, R.S.M., 2020. Big data driven vessel trajectory and navigating state prediction with adaptive learning, motion modeling and particle filtering techniques. IEEE Trans. Intell. Transp. Syst. 23, 3696–3709.
Xin, X., Liu, K., Yang, Z., Zhang, J., Wu, X., 2021. A probabilistic risk approach for the collision detection of multi-ships under spatiotemporal movement uncertainty. Reliab. Eng. Syst. Saf. 107772.
Xin, X., Liu, K., Loughney, S., Wang, J., Yang, Z., 2022a. Maritime traffic clustering to capture high-risk multi-ship encounters in complex waters. Reliab. Eng. Syst. Saf. 108936.
Xin, X., Yang, Z., Liu, K., Zhang, J., Wu, X., 2022b. Multi-stage and multi-topology analysis of ship traffic complexity for probabilistic collision detection. Expert Syst. Appl. 118890.
Xin, X., Liu, K., Loughney, S., Wang, J., Li, H., Ekere, N., Yang, Z., 2023a. Multi-Scale Collision Risk Estimation for Maritime Traffic in Complex Port Waters. Reliab. Eng. Syst. Saf. 109554.
Xin, X., Liu, K., Loughney, S., Wang, J., Li, H., Yang, Z., 2023b. Graph-based ship traffic partitioning for intelligent maritime surveillance in complex port waters. Expert Syst. Appl. 120825.
Xin, X., Liu, K., Yang, X., Yuan, Z., Zhang, J., 2019. A simulation model for ship navigation in the "Xiazhimen" waterway based on statistical analysis of AIS data. Ocean Eng. 180, 279–289.
Yu, H., Fang, Z., Murray, A.T., Peng, G., 2019. A direction-constrained space-time prism-based approach for quantifying possible multi-ship collision risks. IEEE Trans. Intell. Transp. Syst. 22, 131–141.
Yu, H., Murray, A.T., Fang, Z., Liu, J., Peng, G., Solgi, M., Zhang, W., 2021. Ship path optimization that accounts for geographical traffic characteristics to increase maritime port safety. IEEE Trans. Intell. Transp. Syst. 23, 5765–5776.
Yu, H., Meng, Q., Fang, Z., Liu, J., Xu, L., 2023. A review of ship collision risk assessment, hotspot detection and path planning for maritime traffic control in restricted waters. J. Navig. 1–27.
Zhang, J., Liu, J., Hirdaris, S., Zhang, M., Tian, W., 2022a. An interpretable knowledge-based decision support method for ship collision avoidance using AIS data. Reliab. Eng. Syst. Saf. 108919.
Zhang, M., Zhang, D., Fu, S., Kujala, P., Hirdaris, S., 2022b. A Predictive Analytics Method for Maritime Traffic Flow Complexity Estimation in Inland Waterways. Reliab. Eng. Syst. Saf. 108317.
Zhang, W., Goerlandt, F., Montewka, J., Kujala, P., 2015. A method for detecting possible near miss ship collisions from AIS data. Ocean Eng. 107, 60–69.
Zhang, L., Meng, Q., Fwa, T.F., 2019a. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. Transp. Res. Part E Logist. Transp. Rev. 129, 287–304.
Zhang, W., Feng, X., Qi, Y., Feng, S., Zhang, Y., Wang, Y., 2019b. Towards a model of regional vessel near-miss collision risk assessment for open waters based on AIS data. J. Navig. 72, 1449–1468.
Zhang, F., Liu, Y., Du, L., Goerlandt, F., Sui, Z., Wen, Y., 2023a. A rule-based maritime traffic situation complex network approach for enhancing situation awareness of vessel traffic service operators. Ocean Eng. 284, 115203.
Zhang, M., Kujala, P., Musharraf, M., Zhang, J., Hirdaris, S., 2023b. A machine learning method for the prediction of ship motion trajectories in real operational conditions. Ocean Eng. 283, 114905.
Zhang, M., Taimuri, G., Zhang, J., Hirdaris, S., 2023c. A deep learning method for the prediction of 6-DoF ship motions in real conditions. Proc. Inst. Mech. Eng. Part M J. Eng. Marit. Environ. 14750902231157852.
Zhang, X., Wang, C., Jiang, L., An, L., Yang, R., 2021. Collision-avoidance navigation systems for maritime autonomous surface ships: a state of the art survey. Ocean Eng. 235, 109380.
Zhao, Y., Li, W., Shi, P., 2016. A real-time collision avoidance learning system for unmanned surface vessels. Neurocomputing 182, 255–266.
Zhao, L., Shi, G., Yang, J., 2018. Ship trajectories pre-processing based on AIS data. J. Navig. 71, 1210–1230.
Zhen, R., Riveiro, M., Jin, Y., 2017. A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance. Ocean Eng. 145, 492–501.
Zhen, R., Shi, Z., Shao, Z., Liu, J., 2021. A novel regional collision risk assessment method considering aggregation density under multi-ship encounter situations. J. Navig. 1–19.
Zhou, Y., Daamen, W., Vellinga, T., Hoogendoorn, S.P., 2019. Ship classification based on ship behavior clustering from AIS data. Ocean Eng. 175, 176–187.
Zhu, J., Gao, M., Zhang, A., Hu, Y., Zeng, X., 2022. Multi-ship encounter situation identification and analysis based on AIS data and graph complex network theory. J. Mar. Sci. Eng. 10, 1536.
Zwicker, W.S., 1991. The voters' paradox, spin, and the Borda count. Math. Soc. Sci. 22, 187–227.