Jiang, Z., Xiao, Z., Feng, Y., Cao, Y., Zhou, D., & Wang, X. (2024). Analysis of risk influential factors of marine pilots during the embarkation and disembarkation. Journal of Marine Engineering & Technology, 1–14. https://doi.org/10.1080/20464177.2024.2397847

Analysis of risk influential factors of marine pilots during the embarkation and disembarkation

Ziyi Jiang [a], Zhongming Xiao [a, *], Yinwei Feng [a, b], Yuhao Cao [b], Dongjian Zhou [e], and Xinjian Wang [a, b, c, d]

[a] *Navigation College, Dalian Maritime University, Dalian 116026, China*

[b] *Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool L3 3AF, UK*

[c] *Seafarers Research Institute, Dalian Maritime University, Dalian 116026, China*

[d] *Key Laboratory of International Shipping Development and Property Digitization of Hainan Free Trade Port, Hainan Vocational University of Science and Technology, Haikou 570100, China*

[e] *Ningbo Dagang Pilot Co., LTD, Ningbo 315000, China*

* Correspondence: xiaozhongming@dlmu.edu.cn (Z. Xiao)

Abstract: Due to the complex geographical conditions within the port waters, it is necessary to take pilotage operations. The embarkation and disembarkation (E&D) of marine pilots is the riskiest phase of pilotage work, with accidents occurring frequently. In order to examine the key Risk Influential Factors (RIFs) of E&D accident of marine pilots, in this study, a novel complex network model of RIFs of marine pilots during the E&D process is proposed. Firstly, based on the E&D accident investigation report of marine pilots, the causal chain of E&D accidents is extracted, and the RIFs leading to E&D accidents are identified. Secondly, a complex network model of RIFs of marine pilots during the E&D is constructed based on the complex networks theory, and the network reliability is verified. Finally, by comprehensive use of topology characteristics analysis, robustness analysis and module analysis, the key RIFs of pilot E&D accidents are identified. The research results show that pilot and management factors have important impact on pilot safety, "insufficient pilot ladder strength" is the most critical RIF. It should be focused on RIFs with high connectivity between modules and within modules, and the risk evolution can be blocked through the control of those RIFs.

Keywords: Maritime safety; Pilotage operations; Accident analysis; Machine learning; Complex network; Risk analysis

Abbreviations

BNs	Bayesian Networks
CDM	Critical Decision Method
DEMATEL	Decision-making Trial and Evaluation Laboratory
E&D	Embarkation and Disembarkation
FSA	Formal Safety Assessment
RIFs	Risk Influential Factors
RPD	Recognition-Primed Decision
SHERPA	Systematic Human Error Reduction and Prediction Approach
SOLAS	International Convention for the Safety of Life at Sea.
VTS	Vessel Traffic Services
TECHR	Technique for Early Consideration of Human Reliability

1. Introduction

As the main mode of global cargo transportation, maritime transportation occupies up to 80 percent of the market share and thus plays a pivotal role in the global supply chain and has an irreplaceable strategic position (Feng *et al.*, 2024c; Huang *et al.*, 2023; Lee and Yu, 2023). Hence, ships occupy a pivotal position among various modes of transportation, and the safety of ships will also affect the efficiency of cargo transport and the safety of people and property (Demirci *et al.*, 2022; Wang *et al.*, 2023b). Specifically, ports are considered to be one of the areas with a high frequency of maritime accidents (Cao *et al.*, 2023b; Wang *et al.*, 2021). Within port waters, where traffic is dense and certain areas are constrained by geography, relying solely on the bridge team to control the ship is highly risky. Therefore, it is necessary for the pilots to collaborate with the bridge team to coordinate the ship's handling during its entry and exit from the port, as well as its berthing and departure from the terminal (Ugurlu *et al.*, 2017). Consequently, most ports around the world require pilotage (Abreu *et al.*, 2022; Ugurlu *et al.*, 2017).

The job of a marine pilots carries a very high level of risk (Aydin *et al.*, 2022). On the one hand, in the process of ship pilotage, they need to face complex navigational situations, adverse sea conditions, weather conditions and other natural environments to ensure the safe entry and exit of ships to and from the port. On the other hand, the E&D process is the most dangerous phase in a pilot's pilotage operation (Tuncel *et al.*, 2022). Specially, the pilots E&D process is a rigorous and important operational process (Turna, 2024). The embarkation process begins when the pilot steps on a transfer arrangement or helicopter, and ends when the pilot has successfully onboard. Relatively, disembarking the ship is the opposite of the embarkation process. In this process, the pilot leaves the bridge and ends when the pilot arrives at the designated destination. Pilot transfer arrangements consist of helicopters, pilot boats and pilot ladders, gangways and other facilities required by the pilot (China Pilotage Association, 2022).

An in-depth study of the RIFs associated with marine pilots E&D accidents is crucial for improving maritime safety management, reducing accidents and ensuring pilots safety. It not only helps maritime authorities to propose more precise and effective management strategies, but also provides a solid scientific foundation for formulation and updating of industry regulations. By thoroughly understanding these RIFs, stakeholders can better anticipate potential risks and implement targeted measures to mitigate them, ultimately safeguarding the well-being of marine pilots and improving overall maritime operations. In this study, the RIFs of the marine pilots E&D process are analyzed by utilizing topology analysis, robustness analysis and module analysis of complex networks. The main contributions of this manuscript are as follows:

(1) This study collects accident investigation reports on E&D accidents of marine pilots and extracts the causal chain of these accidents. Based on the "human-ship-environment-management" framework and considering the actual working conditions of E&D of marine pilot, the RIFs of E&D accidents of marine pilots are classified into six types: pilot, crew, pilot ship, equipment, environment, and management.

(2) A complex network model of RIFs of marine pilots during the E&D is established by innovatively leveraging complex network theory, and a comprehensive investigation into the RIFs the safety of E&D of marine pilots is conducted by applying topological analysis and robustness analysis.

(3) Employing module analyzes, this study unveils the interconnected influences that hold significance across the entire network as well as within individual modules, shedding light on their association with the incidence of E&D accidents. Subsequently, suggestions and measures are proposed to enhance the safety of E&D of marine pilots, building upon the analytical findings.

2. Literature review

In the literature review section, this study focuses on two key aspects, the research of pilotage safety and the application of complex networks in safety analysis.

2.1. Pilotage safety

To ensure the safety of the ship during its berthing and departure from the ports, the ship should make use of pilotage services (Xi *et al.*, 2021). Nonetheless, pilotage accidents occur periodically, prompting numerous researchers to analyze the ship pilotage process. Aydogdu (2014) employed fuzzy hierarchical analysis to evaluate the risk perception of ship operators, including captains and pilots. Ernstsen and Nazir (2018) employed the SHERPA method to investigate errors and remedies in marine pilotage. The findings indicated that miscommunication and operational errors were the primary RIFs influencing pilotage operations. Atiyah *et al.* (2019) employed Analytical Hierarchy Process to predict pilot reliability. The outcomes of the prediction showed that selecting highly reliable pilots was important to ensure smooth pilotage. Abreu *et al.* (2022) utilized TECHR and BNs to conduct a risk assessment of the pilotage process, drawing upon expert opinion, AIS data, and accident reports. The outcomes demonstrated that the inclusion of an additional pilot in the pilotage process could effectively reduce the likelihood of accidents, thereby demonstrating the critical significance of the number of pilots in ensuring the safety of the pilotage process. Tuncel *et al.* (2022) conducted an analysis by scrutinizing the SOLAS Convention, the Safety Circulars of International Marine Pilotage Association, expert judgments and accident reports to ascertain the potential RIFs contributing to accidents during pilot transfer. The potential RIFs in the pilot transfer process underwent quantitative analysis using the fuzzy extended fault tree method.

In the existing studies, there is evidence suggesting a significant correlation between pilotage accidents and unsafe pilots behaviors (Xi *et al.*, 2019). For instance, Butler *et al.* (2022) conducted an analysis of the decision-making process of marine pilots within an integrated systems thinking framework, employing the RPD and CDM techniques. The results of this analysis proved instrumental in enhancing the safety of the pilotage process. Boudreau *et al.* (2018) conducted a thorough investigation into the factors contributing to diminished pilot alertness, utilizing both experiments and questionnaires. The study concluded that prolonged shifts resulting in irregular sleep patterns were the primary considerations for the decline in pilot alertness. Therefore, to ensure the safety of pilotage, the continuous working hours of the pilots should be reduced and the pilots should be guaranteed sufficient rest. Xi *et al.* (2019) delved into the risk tolerance of pilots through a comprehensive questionnaire, employing a fusion of structural equations and multi-level regression analysis. To delve deeper into the impact of a pilot's psychological safety on pilotage operations, Xi *et al.* (2021) utilized structural equations to evaluate the pilot's "risk tolerance", "risk perception", and "risk attitude" based on survey data. The study showed that hiring pilots with higher "risk perception" is more likely to improve the safety of pilotage operations. Park *et al.* (2019) studied the relationship between pilotage accidents and the age of the pilots, and the results of the study showed that prolonging the age of retirement increases the rate of accidents.

2.2. Complex networks applied in risk analysis

Maritime accidents are subject to multiple interrelated and interacting factors, and there are multiple methods to explore the causes of maritime accidents (Cao *et al.*, 2023a; Christensen *et al.*, 2022). Complex networks are more effective than traditional methods such as accident trees as well as BNs in analyzing the complex interactions between RIFs and identifying critical RIFs (Feng *et al.*, 2024a). Specifically, complex networks have the capacity to visually depict the intricate interconnections between accident RIFs and their subsequent outcomes. By employing this approach, researchers can effectively identify the factors that exert a more substantial impact on the outcomes, thus enabling them to take preventive measures and mitigate the probability of accidents, as shown in Table 1.

Table 1 Application of complex network in traffic and maritime related fields.

Categories	Literatures	Goal	Methods of analysis
Road transportation	(Dai *et al.*, 2016)	Assessing the impact of motorway accidents	Robustness analysis Topological properties
	(Sui *et al.*, 2021)	Assessment of factors affecting the safe operation of the metro	Topological properties
	(Hou *et al.*, 2019)	Assessment of factors affecting the metro construction environment	Robustness analysis
Maritime transportation	(Sui *et al.*, 2023)	Assessment of the temporal characteristics of marine accidents in Yangtze River waters from 2011 to 2020	Robustness analysis
	(Shi *et al.*, 2024)	Assessment of key risk influential factors involved in ship collision accidents	Topological properties & Robustness analysis
	(Lan *et al.*, 2023)	Assessment of the severity of ship collisions	Topological properties

As can be known from Table 1, the application of complex network theory to find the RIFs of accidents is first applied to road traffic accidents. However, in recent years, within the realm of maritime transportation, complex networks have been applied to find the correlations among maritime accident RIFs, thereby facilitating the attainment of effective control over the RIFs in maritime accidents. Sui *et al.* (2023) modelled maritime accidents based on data from 22 jurisdictions in the Yangtze River region, applying network topological features of temporal characteristics of maritime accidents. Shi *et al.* (2024) developed a quantitative risk analysis approach based on complex networks and DEMATEL, to study the evolution of ship collision accidents. Lan *et al.* (2023) integrated association rules, complex networks, and random forests to predict the severity of ship collisions. Their analysis demonstrated that the primary RIF determining the severity of ship collisions was team miscommunication.

To summarize, current pilotage research mainly focuses on risk analysis after pilot onboard and focuses on psychological factors such as pilots' risk tolerance, risk perception and unsafe behaviors (Abreu *et al.*, 2022; Aydin *et al.*, 2022; Xi *et al.*, 2019; Xi *et al.*, 2021). However, the systematic RIFs in the whole process of marine pilots E&D are less examined. In addition, the effectiveness of complex networks theory in identifying key RIFs in accident networks has been demonstrated in related studies (Dai *et al.*, 2016; Lam and Tai, 2020; Sui *et al.*, 2021), and it has been widely used in road and water transportation to reduce the probability of accidents. For this reason, this study extends the application of complex networks to the RIFs analysis of pilots E&D accidents.

3. Methodology

This study focuses on the risk analysis of RIFs during the E&D of marine pilots. In order to examine the key RIFs of E&D accidents of marine pilots, a novel complex network model of RIFs of marine pilots during the E&D is developed based on complex network theory. The modeling process is shown in Figure 1.

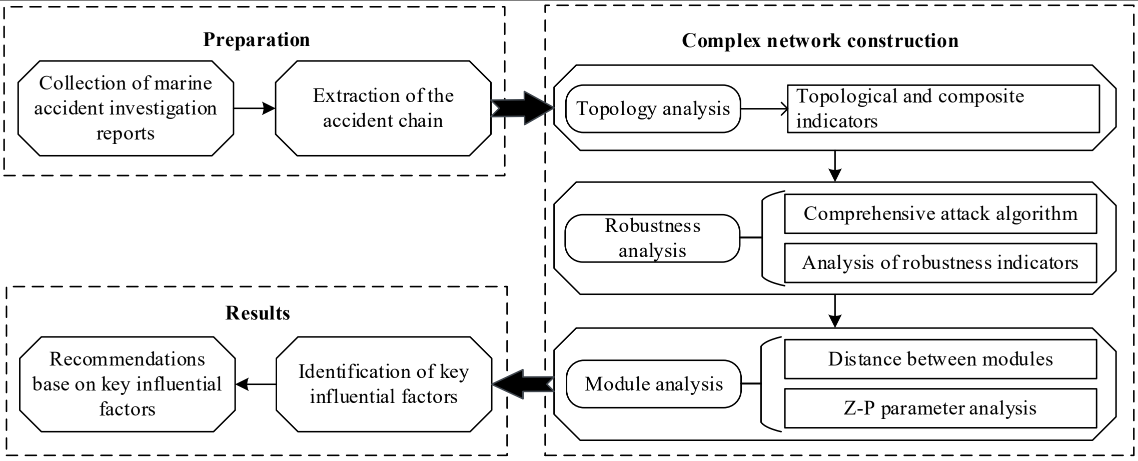

Figure 1 The flow chart of this study.

As shown in Figure 1, firstly, based on the literatures and the investigation reports of the E&D accident of marine pilots, the causal chains of the accidents are systematically analyzed and the key RIFs leading to the accidents are identified. Subsequently, a network model is constructed, which can reveal the complex interactions between RIFs by using the associations of the network nodes and edges, and then the reliability and validity of the model are verified. Further, some important RIFs are identified through diversified methods such as topological analysis, robustness analysis and modular analysis. Finally, countermeasures are proposed for the above key RIFs to improve the safety of marine pilots E&D process.

3.1. Theory of complex networks

Complex networks primarily originate from natural systems, and they have become more common in current society. Precisely, complex networks structure real-world phenomena by abstracting them into individuals and capturing the interactions between these individuals through networks (Feng *et al.*, 2024b; Shi *et al.*, 2024).

3.1.1. Statistical properties

The statistical characterization of a network involves a structural analysis based on the static information present in the network (Wang *et al.*, 2023a). In studies (Feng *et al.*, 2024a; Shi *et al.*, 2024) that employ complex networks to analyze the evolution of accidents, individual RIFs are represented as nodes in the network, and the interaction relationships between these nodes constitute the edges of the network. Through the analysis of the network's statistical properties, it becomes possible to effectively identify key nodes within the network, signifying the crucial RIFs that play a central role in the network's structure (Sui *et al.*, 2021).

(1) Topological indicator

The degree of connectivity and the interactions between nodes exhibited by a network are collectively referred to as the topological properties of the network (Tong *et al.*, 2023). The topological properties of a network include the degree of nodes, network diameter, clustering coefficient indicates, betweenness and closeness centrality (Feng *et al.*, 2024a). The degree of a node is the number of edges in the network that are connected to a node. A higher degree signifies easier direct interactions with more nodes and denotes a relatively higher importance. Network diameter is the maximum value of all distances between two points in the network. Average path length is the average of the sum of distances between all pairs of nodes in existence. The average path length reflects the number of influences that pass through the incident. The clustering coefficient, also known as the local clustering coefficient, indicates the degree of connectivity between nodes connected to a particular node and describes the extent to which the nodes are clustered together in a group. Betweenness Centrality is the number of shortest paths through a node. Higher values imply stronger betweenness centrality, indicating greater influence on neighboring nodes. Betweenness centrality emphasizes the node's ability to regulate and transit between other nodes. Closeness centrality indicates the inverse of the sum of distances from a node to all other nodes and reflects the closeness to other nodes. Smaller closeness centrality signifies a shorter distance from the node to others, placing it spatially at the center.

(2) Composite indicators

To compensate for the different focuses of single indicators in the assessment, the indicators are integrated together to propose a composite indicator that combines four indicators: degree, betweenness centrality, clustering coefficient, and closeness centrality. Three metrics such as global efficiency, transitivity, and maximum connected subgraphs are constructed into a comprehensive evaluation metric in robustness according to the methodology of constructing a comprehensive metric of statistical properties.

It can be assumed that there are *N* nodes in the network and each node feature indicator is *M*. The set of nodes in the network is $G=\{G_1,G_2,G_3,\cdots,G_N\}$, $b_{ij}=(i=j=1,2,3,\cdots,n)$ is the j^{th} feature indicator value of the i^{th} node. The characterization matrix of the network is shown in Equation (1):

$$x=\begin{array}{c c} & \begin{array}{cccc} G_1 & G_2 & \cdots & G_N \end{array} \\ \begin{array}{c} G_1 \\ G_2 \\ \vdots \\ G_N \end{array} & \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \end{array} \tag{1}$$

Each indicator in the network corresponds to a different feature matrix unit, and in order to aggregate the different indicators, it is necessary to unify, i.e., normalize, the metrics, as shown in Equation (2).

$$H_{(ij)}=\frac{b_{ij}-b_{ij,\min}}{b_{ij,\max}-b_{ij,\min}} \tag{2}$$

where H_{ij} in the above equation is the matrix after harmonization of the scales of all indicators. $b_{ij,\min}$ is the smallest eigenvalue in the eigenmatrix, $b_{ij,\max}$ is the largest eigenvalue in the eigenmatrix.

The network centrality indicator C_i is a composite of four indicators, namely, the degree of analysis, betweenness centrality coefficient, clustering coefficient, and closeness centrality, as shown in equation (3).

$$C_i = \sqrt[4]{\prod_{i=1}^{4} H_i} \tag{3}$$

where H_1 is the identity matrix of degree, H_2 is the identity matrix of betweenness centrality coefficient , H_3 is the identity matrix of clustering coefficient, H_4 is the identity matrix of closeness centrality .

3.1.2. Structural properties

In real networks, most networks are large, but a short path exists between any two nodes in the network. Through the study, it has been observed that small-world networks possess both large clustering coefficients and small mean paths, and these characteristics serve as distinctive features of small-world networks (Shi *et al.*, 2024). Therefore, a network is considered small-world if the shortest path between any two nodes requires passing through six or fewer intermediate nodes (Sui *et al.*, 2020).

In small-world networks, the scale-free property implies that the majority of nodes in the network have only a few connections, while a small number of nodes have a significant number of connections (Wang *et al.*, 2023a). The scale-free property manifests itself in the form of a power law in the degree distribution and is mostly characterized by a long-tailed distribution. The small-world and scale-free nature of networks can be demonstrated as essential characteristics of real networks (Dai *et al.*, 2016).

3.2. Robustness analysis of networks

The robustness of a network refers to its capacity to maintain connectivity even in the event of one or more nodes experiencing failure (Feng *et al.*, 2024a). Thus, reducing the connectivity of the network is crucial in preventing further escalation and lowering the probability of accidents. Therefore, this study employs robustness analysis to identify the significance of nodes in the network's evolutionary process.

(1) Global efficiency

The global efficiency Eg reflects the rate at which the risk evolves and propagates through the network. In this study, the network robustness is assessed through the investigation of the global network efficiency, represented by the indicator shown in Equation (4):

$$Eg=\frac{1}{M(M-1)}\sum_{i\neq j}\frac{1}{D_{ij}} \tag{4}$$

where M is the total number of nodes in the network.

(2) Transitivity

Transitivity is defined as the ratio of the number of triangles in the network to the number of connected triples. Its object is the entire network rather than a specific region. The higher the connectivity between nodes after an attack, the more adaptive and robust the network will be, as indicated in Equation (5).

$$T = \frac{\Delta}{V/3} \tag{5}$$

where Δ is the number of triangles in the topology graph of the complex network and $V/3$ is the number of connected triples in the topology graph.

(3) Maximum connected subgraph

The maximum connected subgraph of a network is the subgraph that connects all nodes with the least number of edges. The application of the maximum connected subgraph can visualize the movement of topology change before and after the attack on the network, as shown in Equation (6):

$$\text{M}=\frac{|v_d^{'}|}{|v_d|} \tag{6}$$

where $|v_d^{'}|$ represents the number of nodes in the maximum connected subgraph of the network after performing the attack, $|v_d|$ represents the number of nodes in the maximum connected subgraph in the initial network.

3.3. Module analysis of networks

In modular analyze, the goal is to identify and analyzes groups of nodes in a network that are structurally or functionally closer together than other parts of the network (Guimera and Nunes Amaral, 2005). Dividing the complex network into modules can better analyze the interactions between different modules and the interactions of nodes within the modules, and identify the important nodes connecting different modules as a way to mine the key nodes in the network. Following this, modular analysis is used in this study to reveal the important role of nodes in the internal structure, organization and function of the network.

(1) Within-module connectivity

Within-module connectivity, Z_i is the degree of connectivity of the node in this module and can be used to identify the central node within the module, as shown in Equation (7). If a node has a large Z_i value, it presents that the factor is located in a central position in the factor module and is a core influential factor.

$$Z_i = \frac{K_i - \overline{K_{Si}}}{\delta_{K_{Si}}} \tag{7}$$

where K_i is the number of edges of node i with other nodes in module S_i, $\overline{K_{S_i}}$ is the average value of K for all nodes in module S_i, $\delta_{K_{S_i}}$ is the standard deviation of K values for all nodes in module S_i.

(2) Among-module connectivity

The P_i indicates the degree to which a node is connected to nodes in other modules, and can be used to identify nodes that exchange information in a module, as shown in Equation (8). The closer the value is to 1, the greater the impact on the stability of the network.

$$P_i = 1 - \sum_{s=1}^{N_m} \left(\frac{K_{is}}{K_i} \right) \tag{8}$$

where K_{is} is the number of edges between node i and the nodes in module S, K_i is the degree of node i. M for modules, N_m for all modules.

4. Case study

The primary objective of this study is to examine the correlation existing between RIFs of E&D accidents of marine pilots. This study utilizes complex network theory as its analytical framework, with the specific aim of identifying key RIFs. In complex network theory, directed complex networks can determine if there are interactions between nodes to find nodes that are important in influencing the outcome. Therefore, this study proposes a directed complex network based on accident chains to assess the interactions and correlations between the RIFs.

4.1. Establishment of the network model

The construction process of the network model covers two core aspects: the construction of the complex network model based on the accident chain, and the validation process of the model. During the model construction process, accident reports are utilized to form the accident chain, and the complex network model is further constructed based on the internal logic of the accident chain. In the model validation stage, the authenticity of the network model is verified based on the small-world and scale-free characteristics of complex networks to ensure its scientific and accuracy.

4.1.1. RIFs of marine pilots during the E&D

In this study, first, based on an in-depth review of relevant literatures, the potential range of RIFs is initially identified, and a corresponding list of RIFs is created. Second, a panel of five experts in the field of navigational and pilotage safety is then organized to discuss the list. Semantically similar terms, such as "general decline in pilot ladder rope performance" and "significant decrease in rope strength," are merged into the category "insufficient strength of the pilot ladder rope," thereby finalizing the RIFs. Finally, the experts discuss and evaluate the classification of the RIFs. The basic profiles of these experts are shown in Table 2.

Table 2 The background information of the employed experts.

Expert No.	Age	Time at	Job Title	Field and Experience

	(year)	sea (year)		
Expert A	45	10	Officer of the Maritime Administration	Engaged in maritime supervision for 15 years
Expert B	35	3	Pilot assistant	Engaged in operation practice related to pilotage for 3 years
Expert C	48	10	Senior pilot	Engaged in operation practice related to pilotage for 10 years
Expert D	63	10	Chief officer and professor	Engaged in research related to maritime safety for 35 years
Expert E	49	16	Captain and associate professor	Engaged in theory and practice related to marine navigation for 21 years

This study synthesizes expert opinion on factors affecting pilots E&D and relevant studies to determine the final RIFs classification. By integrating the distinguishing characteristics of the four aspects, namely human, ship, environment, and management, in the maritime domain, the accident factors have been systematically categorized into four layers. These layers represent the underlying factors associated with human involvement, ship-related aspects, environmental conditions, and management practices (Wang *et al.*, 2023c). Pilots, despite being part of the crew, serve as specialized personnel responsible for offering safety advice to the captain, thus warranting a separate categorization within the human factor. As a result, the human factor is further divided into the pilots factor and the crew factor (Ugurlu *et al.*, 2017). Furthermore, the pilot transfer process necessitates the collaborative involvement of both the piloted ship and the pilot ship (Xi *et al.*, 2019). The piloted ship mainly provides the required equipment in this process. Thus, the ship factor is further divided into the pilot ship factor and the equipment factor. Detailed descriptions of the six categories of factors are shown in Table 3. Environmental factors include wind, waves, visibility, etc. Equipment factors include damage to pilot ladders, handrail ropes, and other equipment, etc. When an E&D accident occurs on ships, it frequently results in equipment damage to the ship, such as the dislodging of the pilot ladder. Furthermore, the pilots may sustain injuries or, in severe cases, face fatalities. In order to systematically classify the influence of such accidents, the consequences are categorized into two main groups: those affecting the ship and those impacting the pilots. The specific details of these consequences are outlined in Table 4.

This study explains the terms or some of the difficult nodes involving the pilotage profession as follows. As shown in Table 3, among the human factors, Node 1 (Poor timing of ED) refers to the pilot's failure to thoroughly consider the influences of wind, waves, currents, and other factors, resulting in the selection of an unsuitable time for E&D the ship. Node 5 (Poor choice of E&D positions) refers to the pilot did not take into account factors such as the distance between the pilot ladder and the pilotage ladder, physical fitness, and the weight of the load on his body when E&D and chose an inappropriate position for the E&D operation. Node 8 (Inappropriate use of trap-doors) refers to a passageway located between a vertically suspended pilot ladder and a metal gangway, which allows the pilot to pass through an opening in the lower part of the gangway. This behavior requires good coordination between the pilot's body and limbs.

Table 3 The RIFs affecting accidents involving E&D of marine pilots.

Type	Index number	Details	Type	Index number	Details
Pilot factors (12)	1	Poor timing of E&D	Management factors (6)	23	Improperly maintained and stowed of pilot ladders
	2	Failure to adequately inspect the pilot ladder prior to E&D		24	Inadequate training of pilots
	3	Poor health status		25	Inadequate emergency response
	4	Failure to stand firmly during E&D		26	Inadequate training of relevant personnel to assist in the transfer of pilots
	5	Poor choice of E&D positions		27	Inadequate supervision of ladder placement measures
	6	Poor concentration		28	Inadequate pre-boarding physical examination of pilots
	7	Excessive complacency	Pilot ship factors (7)	29	Improper positioning of pilot ship
	8	Inappropriate use of trap-doors		30	Pilot ship breakdown
	9	Failure to use protective measures		31	Improper operation by support personnel of pilot boat
	10	Working under the influence of alcohol (drugs)		32	Failure to communicate adequately between the pilot ship and the ship being piloted
	11	Overweight (baggage)		33	Inadequate ship illumination on pilot boats
	12	Improper cooperation between ship's officer, pilot, and winchman		34	Improper installation of pilot ladder
Crew factors (5)	13	Inadequate maintenance of pilot ladder		35	Inadequate life-saving equipment on ships employing pilots
	14	Poor storage of pilot ladder	Equipment factors (7)	36	Deterioration of the pilot ladder
	15	Failure to assist the pilot in a timely manner		37	Insufficient pilot ladder strength
	16	Use of substandard pilot ladders		38	Man rope not installed
	17	Improper cooperation between pilot and crew		39	Reel cart of ladder malfunctions
Environment factors (5)	18	Strong wind		40	Poorly designed pilot ladder
	19	Very rough sea		41	Improperly set of pilot ladder
	20	Water temperature too low		42	Improperly set of deck doors
	21	Darkness			
	22	Poor visibility			

Table 4 Results of E&D accidents of marine pilot.

Types	Results
Loss of ships	A: Equipment damage
Loss of personnel	B: Fall down of marine pilot C: Injuries to personnel D: Fatality

Among the ship factors, Node 14 (Poor storage of pilot ladders) denotes the improper storage of removable pilot ladders in environments exposed to humidity, direct sunlight, and other factors, leading to natural erosion and a decrease in ladder strength. Node 15 (Failure to assist the pilot in a timely manner) indicates that the crew should offer prompt assistance to the pilot, involving three essential aspects: escorting the pilot, keeping the pilot ladder close to the ship during E&D, and timely reminders to the pilot regarding the number of remaining treads.

Node 23 (Improperly maintained and stowed of pilot ladders) denotes the occurrence of accidents when the piloted ship fails to conduct sufficient inspections and maintenance of the pilot ladders, resulting in substandard quality of the pilot ladders and accidents during E&D of marine pilot. Node 24 (Inadequate training for pilot) is the result of inadequate supervision of pilot training at the pilotage station and the lack of professional knowledge of the pilots. Node 25 (Inadequate emergency response) pertains to the inadequacy of emergency response capabilities among personnel due to insufficient emergency response drills conducted by relevant agencies in response to pilot hazards. Node 27 (Inadequate supervision of ladder placement measures) indicates insufficient oversight by relevant authorities in inspecting pilot ladders, resulting in the improper installation of the ladder according to regulations, thereby impacting the strength and safety of the pilot ladder.

Node 40 (Poorly designed pilot ladder) denotes the imprudent design of the pilot ladder by the relevant ship design department, leading to inadequate ladder strength when subjected to environmental factors. Node 41 (Improperly set of pilot ladder) is where the pilot ladder is not installed according to the regulations, thus affecting the strength.

4.1.2. E&D accidents chain of marine pilots

This study searches for accident reports from more than 30 maritime accident investigation agencies in coastal countries around the world. However, due to reasons such as confidentiality of pilots E&D accidents in some countries, 50 marine pilots E&D accident reports are obtained from relevant safety agencies. For example, 24 marine investigation reports of E&D accidents are obtained from these U.S. Coast Guard, 12 from the European Maritime Safety Agency, and 14 from websites such as the Australian Maritime Safety Authority, China Maritime Safety Administration, etc. An analysis of the accident investigation report shows that the RIFs of the pilots E&D process are not isolated, but are interconnected. There is information dissemination, interaction or other forms of interaction between these RIFs.

Upon analyzing the E&D accident report, the progression of the accident evolution process is captured through the construction of an accident chain. This construction adheres to three fundamental principles governing accident evolution: 1) The sequential unfolding of events; 2) The logical interconnection between contributing factors; 3) The cascading relationships among these factors. To encapsulate this, complex network theory is applied to the analysis of E&D accidents involving marine pilots. Within this framework, RIFs are represented as nodes, the connections between these RIFs as edges, and the culmination of the collective impact of multiple RIFs signifies the endpoint of event node propagation.

Utilizing the information gleaned from the accident investigation report, accidental RIFs are identified and amalgamated, and requisite accident causal chains are merged. As an illustrative instance, the February 2004 accident involving the pilots of the vessel "Alesia" falling down the gangway is examined. First, six accident chains are extracted based on the accident report, i.e., the graphical process. In addition, the six accident causal chains are transformed into an adjacency matrix. In order to enhance the clarity of this study, only two states are used to characterize the correlation between RIFs when identifying RIFs in the accident chain. As shown in Figure 2, if there is a direct relationship between two RIFs, it is indicated by the number "1", and if there is no direct relationship between two RIFs, it is indicated by the number "0". Ultimately, this matrix is transformed into a network representation to elucidate the complex network model of RIFs of marine pilots during the E&D.

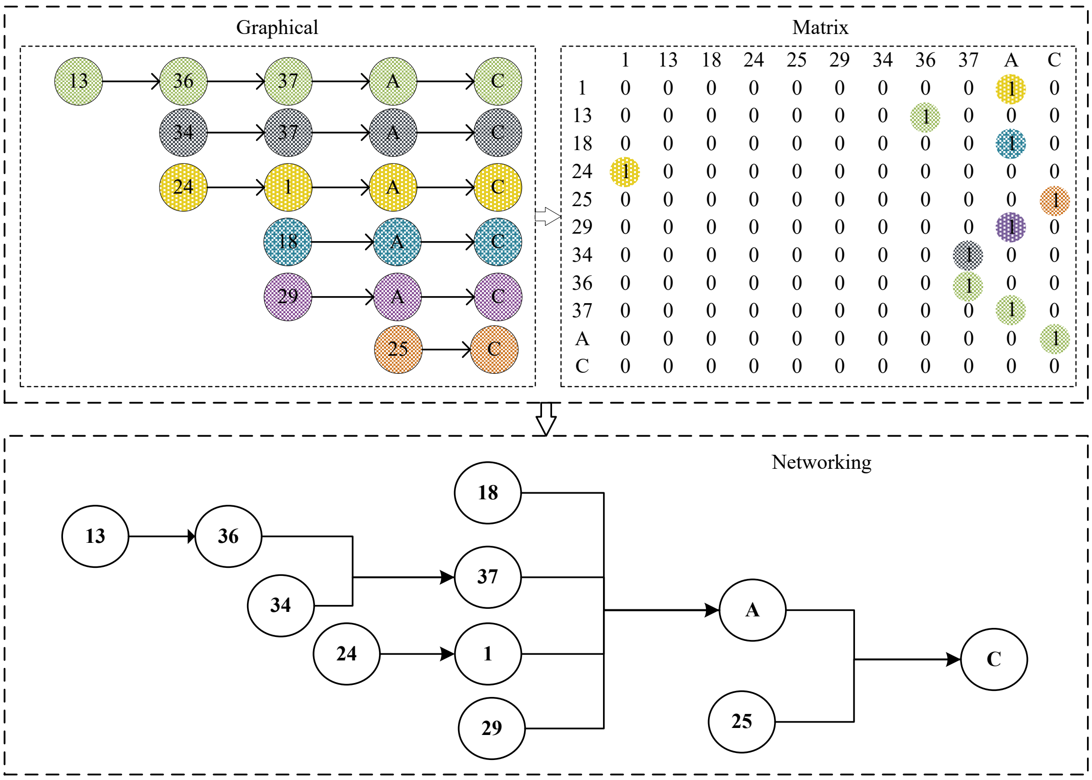

Figure 2 The merging diagram of E&D accidents causal chain of marine pilots.

4.1.3. Network modeling

In this study, a topology structural model is developed to illustrate a complex network of RIFs for accidents based on the logical relationships of the accident chain. The accident chain combination process as shown in Figure 2 is applied to 50 accident reports, then all the accident chains are aggregated and the node impact matrix of the accident network is obtained. The accident impact matrix is input into Python tool, and the network algorithm and drawing function of Python are used to visualize the matrix to obtain Figure 3. Therefore, the network of RIFs of marine pilots during the E&D is a directed network containing 46 nodes, 81 connected edges, and a network diameter of 5, as shown in Figure 3.

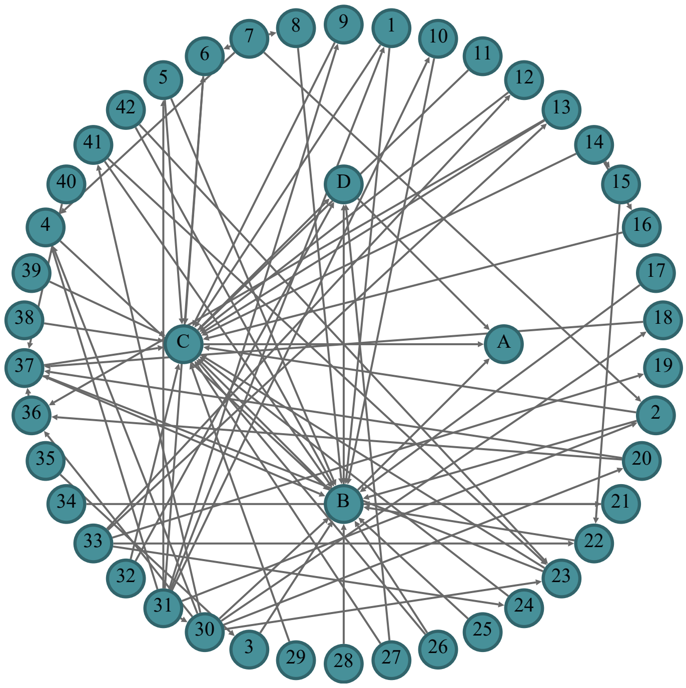

Figure 3 Schematic diagram of a complex network model of RIFs of marine pilots during the E&D.

After Python analysis, it can be observed that the diameter of the complex network model of RIFs of marine pilots during the E&D constructed in this study is 5. This implies that the two most distant points in the network can interact with each other within only 5 steps. Therefore, the network depicting RIFs influencing E&D accidents of marine pilots exhibits characteristics of a small-world network. Further analysis reveals that the regression model is carried out with an accuracy of 0.95515. The degree distribution within the network conforms to the attributes of a power-law distribution, exhibiting a pattern where only a small number of nodes possess substantial degree values, leading to a distribution characterized by a prolonged tail. This distinctive feature is visually represented in Figure 4. This indicates that the network constructed in this study conforms to the characteristics of a scale-free network. Hence, the network of RIFs for E&D accidents of marine pilots developed in this study can be analyzed using the method of complex networks.

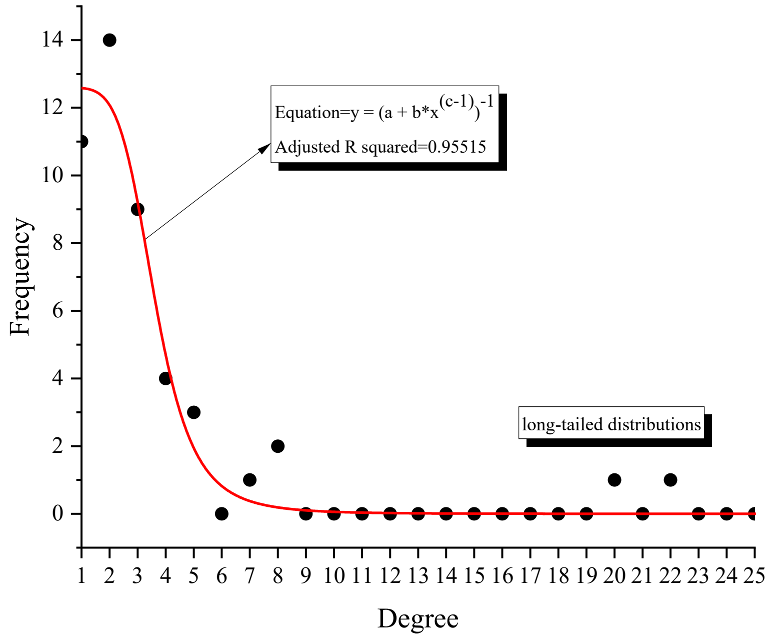

Figure 4 Degree of network model nodes.

Only a select few nodes in the network of marine pilots E&D accident RIFs of marine pilots hold significant positions within the network. Consequently, by controlling only a few key factors, it is possible to effectively halt the evolution of accident network risks and thereby reduce the occurrence of accidents.

4.2. Analysis of the network model

In order to make the analysis of research findings more specific and comprehensive, this study uses the methods of topological analysis, robustness analysis and modular analysis. The topological analysis is conducted based on the structural characteristics of the network. The robustness analysis analyzes the structure and function to verify and supplement the results of the topological analysis. The module analysis analyzes the integrated module "Pilot-Crew-Ship-Equipment-Environment-Management", which can help to identify the key nodes within the module more accurately.

4.2.1. Topology analysis

In this study, the node degree, betweenness, clustering coefficient, and Closeness centrality of the network of marine pilots E&D accident RIFs are statistically analyzed to determine the order of importance of these factors, as depicted in Figure 5.

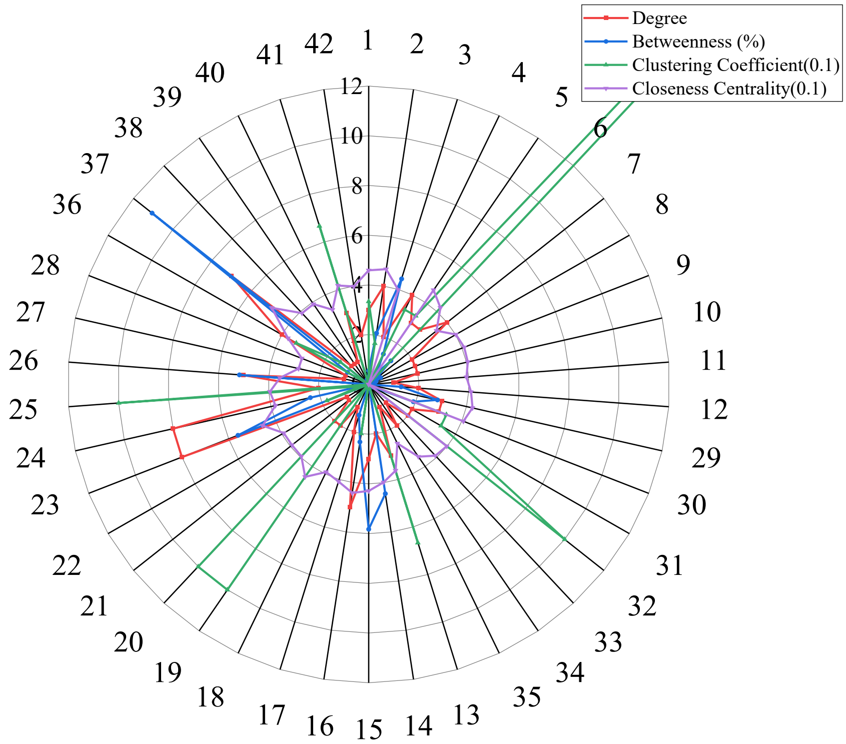

Figure 5 The topological values of nodes based on topology analysis

(1) Nodal degree

As shown in Figure 5, a small number of nodes in the network show large degree values. Specifically, as shown in Table 5, Node 23 (Improperly maintained and stowed of pilot ladders) and Node 24 (Inadequate training of pilots) have a degree value of 8, and Node 37 (Insufficient pilot ladder strength) has a degree value of 7. These nodes are characterized by relatively high degree values and are highly valued as RIFs. The fact that these RIFs are more active in the network indicates that these RIFs are prone to interact with other RIFs and are important nodes in the network of RIFs for E&D accidents of marine pilots.

Table 5 The top 5 nodes in terms of degree

Ranking	Node number	Node names	Nodal degree
1	23	Improperly maintained and stowed of pilot ladders	8
2	24	Inadequate training of pilots	8
3	37	Insufficient pilot ladder strength	7
4	16	Use of substandard pilot ladders	5
5	26	Inadequate training of relevant personnel to assist in the transfer of pilots	5

These RIFs, are characterized by their improved degree values, marking them as hub nodes in the network. Consequently, prioritizing and reinforcing protective and control measures for these key nodes proves to be a more effective strategy.

(2) Betweenness centrality

As shown in Figure 5 and Table 6, the betweenness centrality value of Node 37 (Insufficient pilot ladder strength) is 11.05, which is much higher than the other nodes. The description indicates that this node plays a mediating role in information interaction to a greater extent.

Table 6 The top 5 nodes in terms of Betweenness centrality.

Ranking	Node number	Node names	Betweenness
1	37	Insufficient pilot ladder strength	11.05
2	15	Failure to assist the pilot in a timely manner	5.82
3	23	Improperly maintained and stowed of pilot ladders	5.59
4	3	Poor health status	4.44
5	2	Failure to adequately inspect the pilot ladder prior to E&D	2.08

Furthermore, the higher betweenness centrality of Node 37 (Insufficient pilot ladder strength), node 23 (Improperly maintained and stowed of pilot ladders), Node 15 (Failure to assist the pilot in a timely manner), and Node 3 (Poor health status) in the network indicates that such nodes act to a greater extent as mediators of information interactions. They act as key hubs in E&D accidents of marine pilots. Changes in these nodes are prone to slow down the network information interaction, so controlling such nodes avoids further escalation of the danger. The strength of the pilot ladder is a prerequisite for the safe working of the pilot during the actual E&D of pilot operations.

(3) Clustering coefficient

The average clustering coefficient of the E&D accidents network is 0.203, and some nodes exhibit a clustering coefficient greater than 1. This indicates that these nodes have a high level of interconnectedness with other nodes, leading to the rapid evolution of accidents within the network.

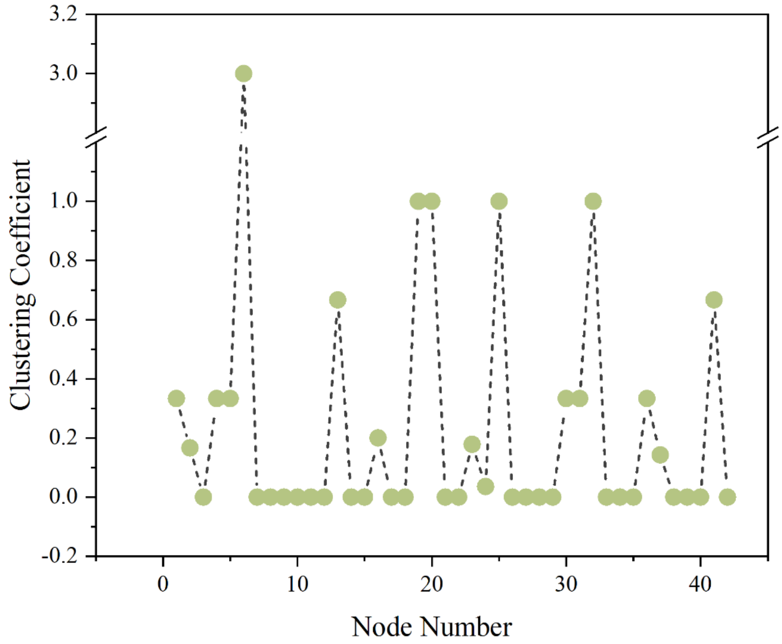

Figure 6 Ranking results of clustering coefficient.

For example, as shown in Figure 6, Nodes 7, 8 and 9 have clustering coefficients of 0. This is because they have only one connected node. The clustering coefficient of Node 6 is the most prominent, indicating that this node is prone to interact with other nodes. The control of node 6 can effectively prevent further evolution of the accident and reduce the probability of accident in the network. Thus, controlling Node 6 can effectively impede the further evolution of accidents and alleviate the challenges in accident control within the network.

(4) Closeness centrality

The overall closeness centrality of the nodes is more evenly distributed, as shown in Figure 5. Specifically, through the presentation in Figure 7, it can be observed that the closeness centrality value of Node 37 (Insufficient pilot ladder strength) is 4.96, the closeness centrality value of Node 2 (Failure to adequately inspect the pilot ladder prior to E&D) is 4.68, and the closeness centrality value of Node 5 (Poor choice of E&D positions) is 4.60. These values significantly indicate that such nodes have prominent closeness centrality characteristics. These nodes have relatively short distances from other nodes and are located in the central region of the network, thus occupying a key position in the entire network.

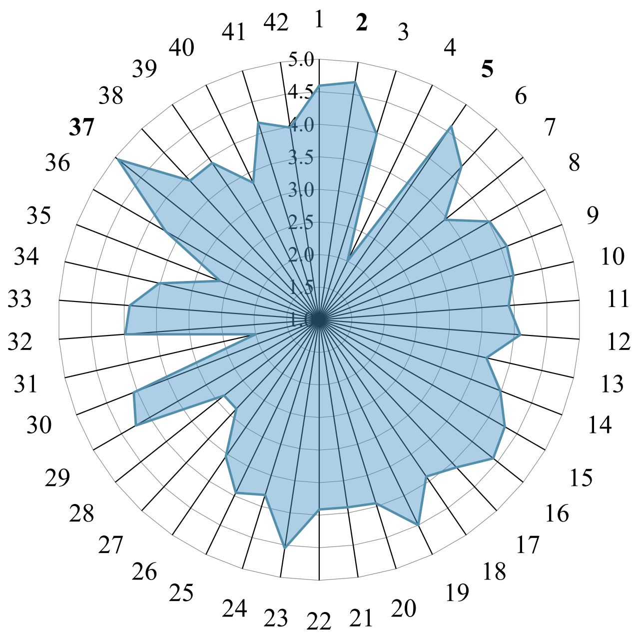

Figure 7 Ranking results of closeness centrality

Using node 5 as an illustration, the poor choice of E&D refers to the pilot's unfamiliarity with the characteristics of wind and waves, and the inability to comprehend the pattern of substitution between big waves and small waves. Typically, a few big waves will be followed by a few small waves, forming an alternating pattern known as the "three big and eight small" sequence. Pilots should master the laws of wind and waves, and choose the right time for E&D operations, to avoid putting themselves in danger.

(5) Composite indicators

In the network of factors influencing E&D accidents of marine pilots, the application of topological characterization of the network can identify important RIFs. However, the focus of different indicators in the statistical characterization is different, and the corresponding ranking results are also different. This study proposes a comprehensive ranking in order to facilitate more accurate and comprehensive mining of nodes that have a high impact on the network. The comprehensive sorting results are obtained by dimensionless and standardized degree, betweenness centrality, clustering coefficient, and closeness centrality using Equations (7) to (9), as shown in Figure 8.

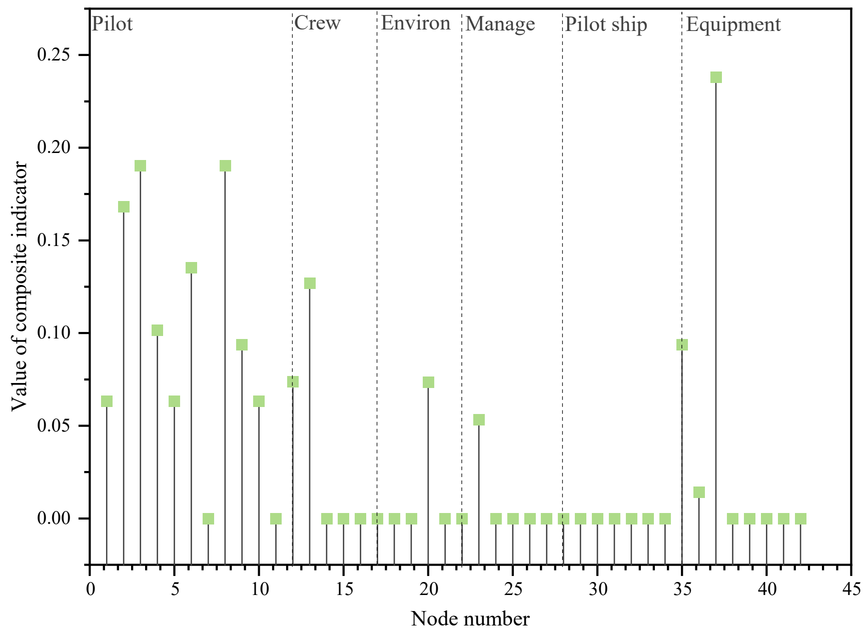

Figure 8 Distribution of nodes based on comprehensive indicators.

As illustrated in Figure 8, Node 37 (Insufficient pilot ladder strength) attains the highest composite indicator value. This highlights the critical concern regarding the inadequacy of pilot ladder strength, necessitating immediate attention. Notably, 80 percent of the nodes associated with pilot-related factors occupy the top positions in the ranking results. This prevalence underscores the significance of pilot factors as pivotal influencers. The prominence of pilot factors implies the imperative need to focus on elements tied to the pilots themselves, as this holds substantial importance in steering the control of situational evolution.

4.2.2. Robustness analysis

Traditionally, "robustness" refers to the ability of a system, algorithm or model to maintain stability and reliability in the face of disturbances. In this study, the network modelled is the network of the accident evolution process. The higher the robustness of the network is, the better the connectivity of the accident network is, the smoother the accident evolution process is, and the less conducive to accident prevention or safety management is. To better characterize the changes in network robustness under different attacks, the specific changes are shown in Figure 9. The specific details selected in the figure are demonstrated in Figure 10.

As depicted in Figure 9, the overall metrics of the network exhibit a decreasing trend with an increasing number of attacking nodes. The smoother decline observed under random attacks can be attributed to the utilization of the average of 1000 simulations as the final attack effect when nodes are randomly selected. The effectiveness of the attack is quantified by the robustness composite indicator, denoted as "robustness." Robustness demonstrates a decreasing trend under both deliberate and random attacks. However, relying on a single robustness indicator may portray an upward trend. For instance, using only the network's transmissibility as an indicator might suggest that the network spreads information or influence more rapidly when nodes distant from the center are removed. Therefore, network robustness is comprehensively assessed through three key metrics: global efficiency, transitivity, and maximum connected subgraph.

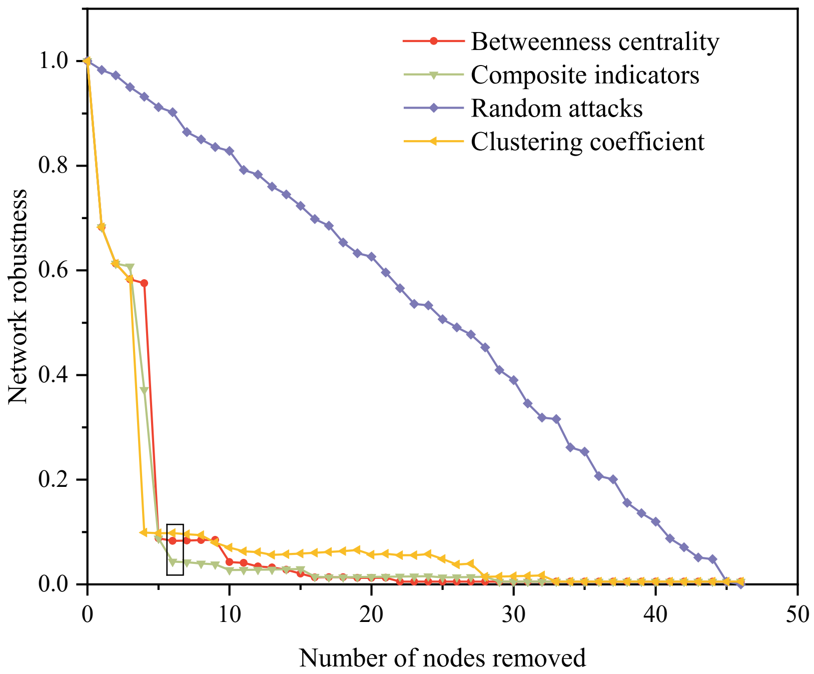

Figure 9 The network robustness of post-attack.

As shown in Figure 9, it can be seen that the network collapses when the number of attacking nodes is sufficiently high. Therefore, this study judges the important RIFs on the basis of the speed of the decrease of the network robustness. These targeted attacks prove to be more effective in disrupting the network structure and isolating critical nodes. This underscores the importance of employing comprehensive metrics to identify key RIFs. The analysis clearly indicates that the network depicting factors influencing E&D accidents of marine pilots is more vulnerable to deliberate attacks, given that such attacks specifically target critical nodes, resulting in a swift deterioration of network. On the other hand, random attacks demonstrate that the network possesses a certain degree of fault tolerance. The impact on the network is less severe, allowing it to maintain a relatively high level of efficiency even with random node removals.

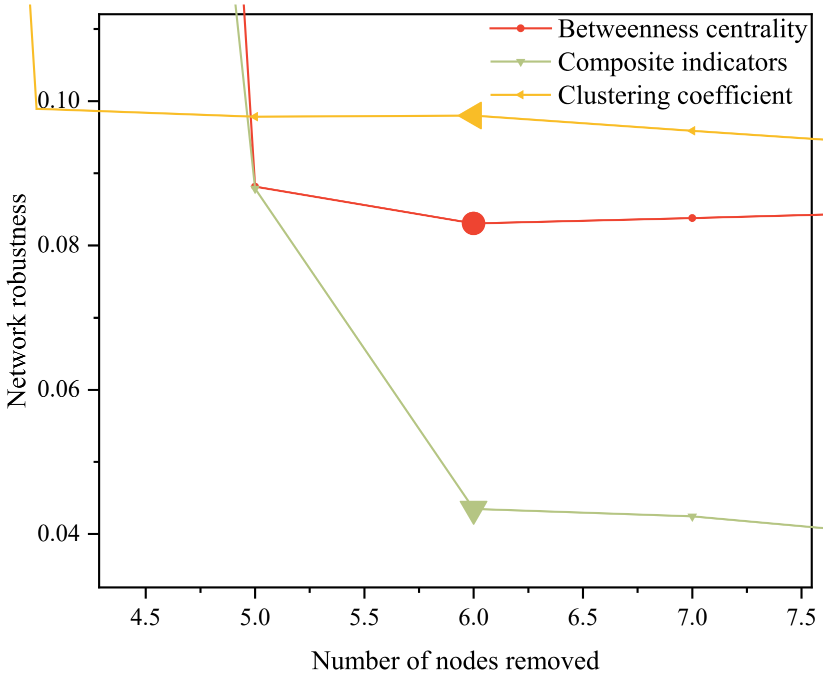

Figure 10 The details of network robustness after removing 6 nodes.

As shown in Figure 10, both the targeted attack based on betweenness centrality and the targeted attack based on composite indicator are significantly faster than the targeted attack based on clustering coefficients. In the case of removing six nodes, the robustness of the composite indicators-based attack is 0.004. The robustness of the betweenness centrality-based attack network is 0.008, and the robustness of the clustering coefficient-based attack is 0.009. Thus, the composite indicator-based attack exhibits a more significant reduction trend compared to the other two types of targeted attacks. This observation suggests that isolating key nodes in the risk network and severing connections between these key nodes and others can effectively impede the evolution of risk from triggering events to accident occurrence in the network, thereby reducing the incidence of accidents.

4.2.3. Module analysis

In order to further explore the relationship between the various RIFs affecting the E&D of marine pilots, the network is divided into modules. The network is categorized into six modules according to the maritime accidents "Pilot-Crew-Ship-Equipment-Environment-Management". Further explore the role of each RIF in the module, in the network as a whole, and in the results. The average path length from each module to the result determines the location of each module, and the module division is shown in Figure 11.

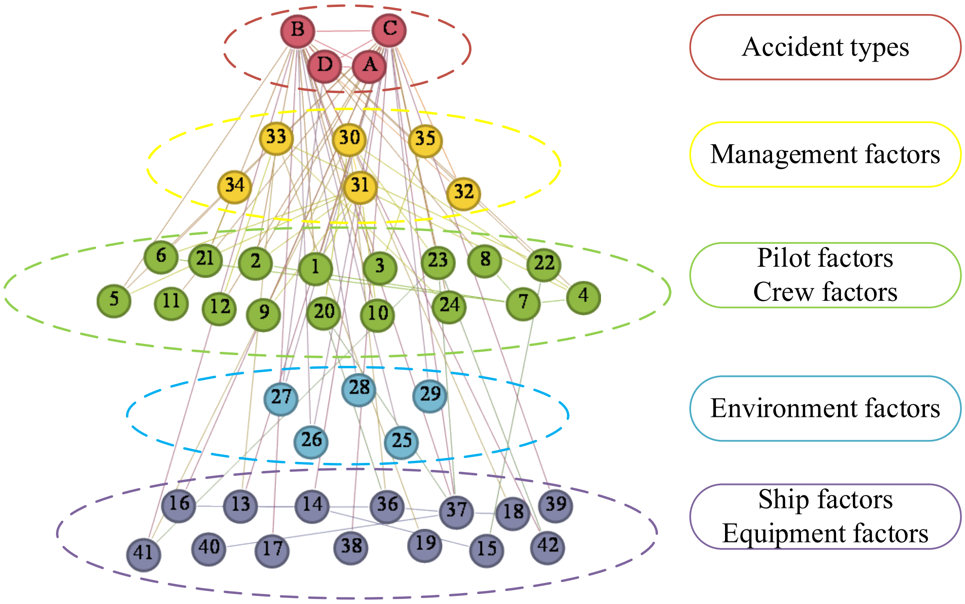

Figure 11 The results of layered visualization.

As shown in Figure 11, the management factor module has the shortest distance to reach the result module, indicating that this type of factor has a higher priority. Management factors primarily encompass the oversight and supervision of the pilot by the pilot station, as well as the supervision of the pilot ship by the relevant ship flag authorities. Conversely, the average path length from the environmental factors layer to each node representing accident outcomes is comparatively lengthier, suggesting a lower priority for environmental factors in influencing the occurrence of accidents. High-priority factors need to be controlled more effectively to reduce the rate of risk evolution and prevent accidents from occurring.

In this study, based on the assessment of the location of each module, the Among-module connectivity (P_i) and Within-module connectivity (Z_i) are used to identify the importance of each RIF in the whole network, as shown in Figure 12. In the established network, there are some RIFs (nodes) that have the same values of within-module connectivity and intra-module connectivity, for example, Node 2 and Node 4 overlap, which are expressed as 4(2). Exploring the key nodes of the network is based on the sub-module of "Human-Ship-Pilot-Crew-Ship-Equipment-Environment-Management". As can be seen in Figure 12, Node 7, Node 14, Node 30, Node 36, Node 37 have higher Z_i values, indicating that such nodes are core nodes within the module. For example, Node 7 (Excessive complacency) is the key node of the "human" module, which may cause Node 2 (Failure to adequately inspect the pilot ladder prior to E&D), Node 5 (Poor choice of E&D positions), Node 6 (Poor concentration). Node 7, Node 30 and Node 37 have higher Among-module connectivity, indicating that such nodes are more connected to nodes in other modules and exchange information more frequently. Controlling such nodes in the network depicting factors influencing E&D accidents of marine pilots reduces the speed of the accident evolution process.

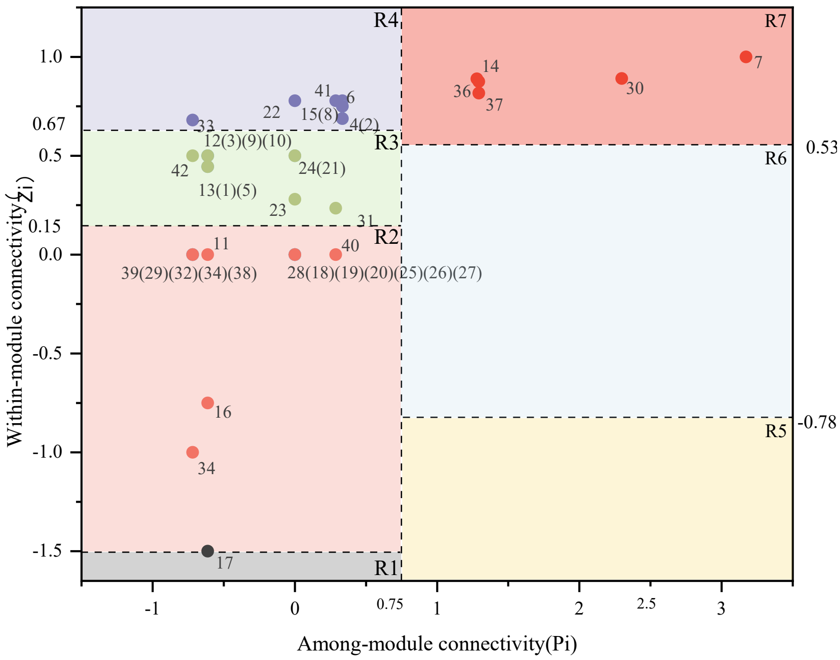

Figure 12 The analysis of the Z-P parameter.

Based on the theory of module analysis (Guimera and Nunes Amaral, 2005), the nodes are divided into regions proportionally. The nodes in the regions of R1, R2, R3 and R4 have less information transfer within the modules. The nodes in the region of R5 and R6 are mostly connected within the modules and less connected between the modules. The nodes in the region of R7 are the central nodes within the modules as well as between the modules, and this kind of nodes has an important impact on the stability of the network. Therefore, the control of the nodes in the R7 region should be strengthened. In the network of factors affecting pilot E&D, Node 7 (Excessive complacency) and Node 14 (Poor storage of pilot ladder) are the central nodes in the "human" factor module. Node 36 (Aging of the pilot ladder) and Node 37 (Insufficient strength of the pilot ladder) are the central nodes of the "management" factor module. At the same time, these nodes have a high degree of connectivity between modules. It indicates that such RIFs play an important role in the whole E&D accident network, and depict factors influencing E&D accidents of marine pilots and each factor module. Effective control of such RIFs can reduce the risk of pilots E&D.

4.3. Discussion and implication

In summary, the accidents that occur during the E&D of marine pilots are not random events but can be prevented to a certain extent by identifying the RIFs and implementing appropriate measures. Hence, based on the results obtained from topological analysis, robustness analysis, and modular analysis, this study proposes comprehensive and targeted accident prevention measures to enhance safety and mitigate risks.

According to the results of the topological analysis, node 37 (insufficient strength of the pilot ladder) has the highest value of the composite indicator, indicating that the insufficient strength of the pilot ladder has a significantly higher impact on the network than other nodes. Indeed, the reasons behind the insufficient strength of the pilot ladder can be primarily categorized as follows. 1) Inadequate design: This can lead to malfunction during pilot transfers, as the design may not adequately support the operational demands. 2) Non-compliance with Current Standards: The pilot ladder may not be suitable for conducting personnel transfer operations, as evidenced by recent regulations that prohibit the use of mechanical pilot lifts on ships. 3) Insufficient Maintenance: Factors such as wear on the side cables, tilting of the treads, presence of oil stains, and rotting of the brown ropes due to improper storage contribute to the ladder's diminished integrity. To ensure the strength and safety of the pilot ladder, it is crucial for the crew to meticulously maintain it. This includes regular inspections, immediate repairs, and careful storage. Additionally, installation of the pilot ladder must be executed with precision, using only qualified ladders that have passed thorough testing prior to use in E&D processes. This stringent approach is essential to guarantee the utmost safety for pilots during these critical operations.

According to the results of the robustness analysis, attacks using betweenness centrality and composite indicators clearly lead to the most significant changes in the network structure. It shows that these metrics have a profound effect on the structure and function of the network. As depicted in Figure 13, Node 37 (Insufficient strength of the pilot ladder) exhibits the highest values for both the betweenness centrality and the composite indicators, indicating that node 37 plays a pivotal role in determining the robustness of the network structure. Node 37 (Insufficient strength of the pilot ladder) ranks high in the topological analysis and has high values in the robustness analysis under both the betweenness centrality and composite indicators. This phenomenon suggests that the RIFs of insufficient strength of the pilot ladder play a key role in the whole network of factors influencing the pilot B&D network accidents, both in terms of structure and function. In fact, the E&D accident rate of marine pilots is closely related to the strength of the pilot ladder. Considering this, it is recommended to design suitable pilot ladders, strengthen the inspection and maintenance of pilot ladders, and ensure that the pilot ladders have sufficient strength for personnel transfer operations.

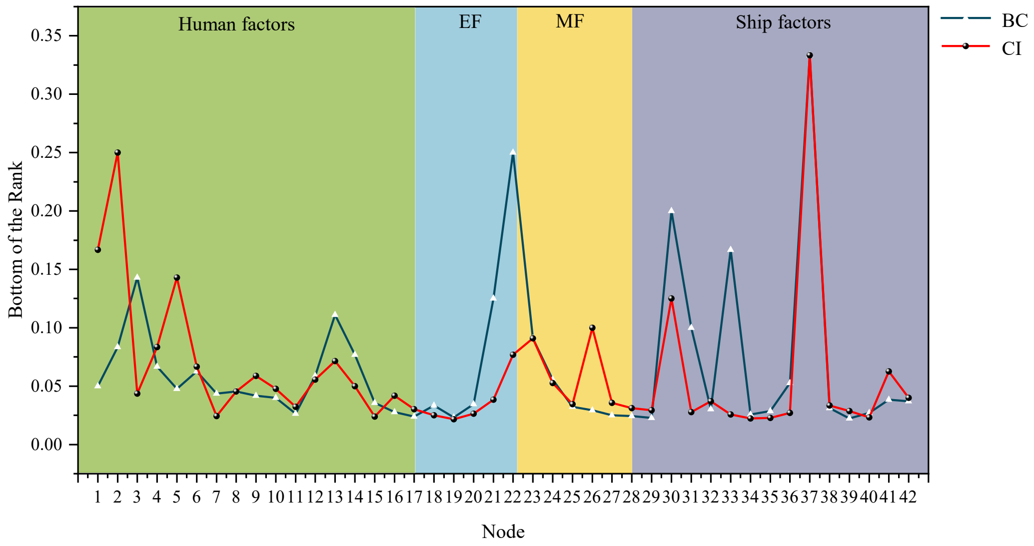

Figure 13 The number of intermediaries and the ranking of composite indicators.

Based on the outcomes of the modules analysis, the management factor layer exhibits the shortest average distance from the nodes of the outcome modules, signifying its pivotal role in the risk evolution process. Consequently, there is a pressing need to reinforce management measures to obstruct the evolutionary process of accidents. This encompasses bolstering the management of E&D operations by ships, shipping companies, and pertinent pilotage stations. In addition, the focus should be on nodes with high within-module connectivity and intra-module connectivity, and the control of such nodes can block the evolution of risks. For example, Node 7, associated with behaviors such as poor concentration, inadequate inspection of the pilot ladder, and unstable footing, significantly increases the probability of E&D accidents. Regarding Node 4 (Excessive Complacency), marine pilot stations are required to implement a regular psychological assessment and training program for pilots to enhance their situational awareness and reduce complacency. Effective control of such nodes can cause the decomposition of modules and networks, affecting the overall functioning of the network, preventing the evolutionary process of risk, and reducing the probability of E&D accidents of marine pilots.

5. Conclusions

Based on the publicly available pilots E&D accident reports, this study extracts and establishes the corresponding accident chains according to the evolution process of the accidents. Then, a complex network model of RIFs of marine pilots during the E&D is constructed using complex network theory to integrate the accident chain. In this model, the "human-ship-environment-management" framework is subdivided into six key modules, pilot, crew, ship, equipment, environment and management. Then, 42 RIFs are identified, and the reliability of the network is further verified based on the scale-free and small-world characteristics of the network. In addition, a combination of topological characterization, robustness analysis and module analysis are used to identify the key RIFs affecting marine pilots E&D accidents.

Although this study offers valuable insights to stakeholders in the shipping industry, there are some limitations need attentions. On the one hand, the relatively limited size of the dataset may limit the depth and comprehensiveness of the findings to some extent. On the other hand, when applying complex network theory to analyze the RIFs of accidents, it is limited by the framework of local concepts, and thus the potential influence of specific nodes in accidents may be neglected. In order to effectively reduce the risk of maritime pilotage, further research can be conducted in the following two core aspects in the future. First, in-depth research on machine learning algorithms to predict the likelihood of E&D accidents, provides solid theoretical support for the development of forward-looking safety measures. Second, considering the application of dynamic modeling in network models, effectively alleviates the inherent limitations of existing network models.

Acknowledgments:

The authors gratefully acknowledge support from the National Natural Science Foundation of China [Grant No. 52101399], National Key Research and Development Plan of China (Grant No. 2019YFE0111600), Bolian Research Funds of Dalian Maritime University (Grant No. 3132023617). This work is also supported by the EU H2020 ERC Consolidator Grant program (TRUST Grant No. 864724).

Conflicts of Interest:

The authors declare no conflict of interest.

Appendices

Appendix A. Pseudo-codes of the algorithms of this study

Table A1 Pseudo-codes of Topological analysis.

Algorithm 1 Topological analysis

Require: Pilot E&D Incident Chain Data *data*

Initialize an ranking list: degree_list=[], Betweenness_list=[], Clustering coefficient_list=[], Closeness centrality_list=[], Composite indicators_list=[], Create an empty matrix Composite indicators((4,node))

for each node n in the *data* **do**

calculated value of a degree:

$$K_i = \sum_{i \in N} a_{ij}$$

calculated value of a Betweenness:

$$L = \frac{2}{N(N-1)} \sum_{i=1} \sum_{j=i+1} D_{ij}$$

calculated value of a Clustering coefficient:

$$C(i) = \frac{N(i)}{k_i(k_i - 1)}$$

calculated value of a Closeness centrality:

$$E(i) = \sum_{s,t \in N} \frac{\delta_{s,t}(i)}{\delta_{s,t}}$$

Add Ki in degree_list, Add L in Betweenness_list, Add Ci in Clustering coefficient_list, Add Ei in Closeness centrality_list
end for
Combine lists into matrices
for each node n in the $data$ **do**
    Calculate the normalized value

$$H_{(ij)} = \frac{b_{ij} - b_{ij,\min}}{b_{ij,\max} - b_{ij,\min}}$$

    Calculate the value of geometric averaging

$$C_i = \sqrt[4]{\prod_{i=1}^{4} H_i}$$

Take the value composition of the matrix and add it to the list Composite indicators_list
Return degree_list, Betweenness_list, Clustering coefficient_list, Closeness centrality_list, Composite indicators_list

Table A2 Pseudo-codes of Select the node from the network.

Algorithm 2 Select the node from the network

Require: Pilot E&D Incident Chain Data $data$,
Stage1: random selection
for node n Length less than $data$ length **do**
    for i <=1000 **do**
    Initialize feature ranking list: Rankings = []
        While Length less than N length do
            Randomly select a node in data
            If node not in Rankings do
            Append the node to Rankings.
        end while Rankings of randomly selected nodes of length n
    end for a thousand times on average.
end for random selection of nodes from the network
return Rankings
Stage2: intentional selection
Initialize Sorting results list: Sorted = [].
for each node n in $data$ **do**
    Select nodes with high values n
    Append the node n in the Sorted
end for get the sorting result
return Sorted

Table A3 Pseudo-codes of Divide the community.

Algorithm 3 Divide the Module

Require: Pilot E&D Incident Chain Data $data$,
for each node n in the $data$ **do**
    Assign a unique community identifier to each node
    Initializing Modularity Gain f = True
    for Modularity Gain f = True **do**
        Modularity Gain = False
        Move nodes to neighboring communities and compute modularity gain
        If Modularity Gain > 0 **do**
            Modularity Gain f = True
    end for A Module Sequence
end for Module assignment results for nodes
return Module results

Table A4 Pseudo-codes of Module analysis.

Algorithm 4 Module analysis

Required: Pilot E&D Incident Chain Data $data$
Divide the community, initialize an ranking list: Pi= [], Zi= [].
for each node n in the $data$ **do**

3: Calculate the P_i value of the node:

4: $$Z_i = \frac{K_i - \overline{K_{Si}}}{\delta_{K_{Si}}}$$

5: Calculate the Z_i value of the node:

6: $$P_i = 1 - \sum_{s=1}^{N_m} \left(\frac{K_{is}}{K_i} \right)$$

7: Add the value of P_i at Pi, Add the value of Z_i at Zi

8: **end for:**

9: Get the list Pi and the list Zi

Appendix B. The equations appearing in the algorithms of this study

(1) Degree of nodes

The degree of a node in the network refers to the number of edges connected to the node. A higher degree value signifies greater importance of the node within the network.

$$K_i = \sum_{i \in N} a_{ij} \tag{9}$$

where K_i represents the degree value, N is the number of nodes in the network, a_{ij} is the number of connected edges between nodes i and j.

(2) Network diameter and average path length

The network diameter, as expressed in Equation (2), represents the maximum distance between any two points in the network.

$$D = \max_{ij \in N} \left(D_{ij} \right) \tag{10}$$

where D_{ij} denotes the distance between node i and node j.

The average path length, as indicated in Equation (3), corresponds to the average of the sum of distances between all pairs of nodes that have a connected path between them.

$$L = \frac{2}{N\left(N-1\right)} \sum_{i=1} \sum_{j=i+1} D_{ij} \tag{11}$$

where L is the average path length.

(3) Clustering coefficient

The clustering coefficient indicates the degree of linkage between nodes connected to a node and describes the degree to which the nodes are clustered together in a group, as in Equation (4).

$$C(i) = \frac{N(i)}{k_i(k_i - 1)} \tag{12}$$

where C_i is the clustering coefficient; k_i is the number of edges that node i assumes to connect to its neighboring nodes; N_i is the number of edges that actually exists between those nodes connected to node k_i.

(3) Mediated Centrality

Mediated Centrality corresponds to the number of shortest paths passing through a specific node. A higher value indicates stronger centrality, signifying that this node exerts more influence on its neighboring nodes, as described in Equation (5). Mediated Centrality refers to the capability of a node to regulate and serve as a relay between other nodes within the network.

$$E(i)=\sum_{s,t\in N}\frac{\delta_{s,t}(i)}{\delta_{s,t}} \tag{13}$$

where E_i is the mediated centrality; $\delta_{s,t}$ is the number of shortest paths between nodes s and t; $\delta_{s,t(i)}$ is the number of shortest paths between nodes s and t through the node i.

(5) Closeness centrality

Closeness centrality denotes the inverse of the sum of distances from a particular node to all other nodes in the network. It reflects the closeness of the node to other nodes, as described in Equation (6). Indeed, the smaller the closeness centrality of a node, the shorter the distance from this node to other nodes in the network. Therefore, spatially, this node which has a smallest closeness centrality is positioned at the center of the network.

$$F(i)=\frac{1}{\sum_j D(j,i)} \tag{14}$$

where $F_{(i)}$ is the closeness centrality; $\sum_j D(j,i)$ is the sum of distances from node i to all other nodes.

References:

Abreu, D.T.M.P., Maturana, M.C., Droguett, E.L., Martins, M.R., 2022. Human reliability analysis of conventional maritime pilotage operations supported by a prospective model. Reliability Engineering & System Safety 228, 108763. https://doi.org/10.1016/j.ress.2022.108763

Atiyah, A.A., Kontovas, C., Saeed, F., Yang, Z., 2019. Marine Pilot's Reliability Index (MPRI): Evaluation of marine pilot reliability in uncertain environments. Liverpool, ENGLAND, Jul 14-17, 2019.

Aydin, M., Uğurlu, Ö., Boran, M., 2022. Assessment of human error contribution to maritime pilot transfer operation under HFACS-PV and SLIM approach. OCEAN ENGINEERING 266, 112830. https://doi.org/10.1016/j.oceaneng.2022.112830

Aydogdu, Y.V., 2014. A Comparison of Maritime Risk Perception and Accident Statistics in the Istanbul Straight. Journal of Navigation 67 (1), 129-144. https://doi.org/10.1017/s0373463313000593

Boudreau, P., Lafrance, S., Boivin, D.B., 2018. Alertness and psychomotor performance levels of marine pilots on an irregular work roster. Chronobiology International 35 (6), 773-784. https://doi.org/10.1080/07420528.2018.1466796

Butler, G.L., Read, G.J.M., Salmon, P.M., 2022. Understanding the systemic influences on maritime pilot decision-making. Applied Ergonomics 104, 103827. https://doi.org/10.1016/j.apergo.2022.103827

Cao, Y., Wang, X., Wang, Y., Fan, S., Wang, H., Yang, Z., Liu, Z., Wang, J., Shi, R., 2023a. Analysis of factors affecting the severity of marine accidents using a data-driven Bayesian network. OCEAN ENGINEERING 269, 113563. https://doi.org/10.1016/j.oceaneng.2022.113563

Cao, Y., Wang, X., Yang, Z., Wang, J., Wang, H., Liu, Z., 2023b. Research in marine accidents: A bibliometric analysis, systematic review and future directions. OCEAN ENGINEERING 284, 115048. https://doi.org/10.1016/j.oceaneng.2023.115048

China Pilotage Association, 2022. Pilot boarding and disembarkation operations risk control and hidden danger investigation and management manual. Shanghai Pujiang Education Press Co., Ltd. , Shanghai, China.

Christensen, M., Georgati, M., Arsanjani, J.J., 2022. A risk-based approach for determining the future potential of commercial shipping in the Arctic. Journal of Marine Engineering & Technology 21 (2), 82-99. https://doi.org/10.1080/20464177.2019.1672419

Dai, H., Yao, E., Lu, N., Bian, K., Zhang, B., 2016. Freeway Network Connective Reliability Analysis Based Complex Network Approach. Beijing, PEOPLES R CHINA, Jul 02-06, 2016.

Demirci, S.M.E., Canimoglu, R., Elcicek, H., 2022. Analysis of causal relations of marine accidents during ship navigation under pilotage: A DEMATEL approach. Proceedings of the Institution of Mechanical Engineers Part M-Journal of Engineering for the Maritime Environment, 237(232):308-321. https://doi.org/10.1177/14750902221127093

Ernstsen, J., Nazir, S., 2018. Human Error in Pilotage Operations. Transnav-International Journal on Marine Navigation and Safety of Sea Transportation 12 (1), 49-56. https://doi.org/10.12716/1001.12.01.05

Feng, Y., Wang, H., Xia, G., Cao, W., Li, T., Wang, X., Liu, Z., 2024a. A machine learning-based data-driven method for risk analysis of marine accidents. Journal of Marine Engineering & Technology, 1-12. https://doi.org/10.1080/20464177.2024.2368914

Feng, Y., Wang, X., Chen, Q., Yang, Z., Wang, J., Li, H., Xia, G., Liu, Z., 2024b. Prediction of the severity of marine accidents using improved machine learning. Transportation Research Part E: Logistics and Transportation Review 188, 103647. https://doi.org/10.1016/j.tre.2024.103647

Feng, Y., Wang, X., Luan, J., Wang, H., Li, H., Li, H., Liu, Z., Yang, Z., 2024c. A novel method for ship carbon emissions prediction under the influence of emergency events. Transportation Research Part C: Emerging Technologies 165, 104749. https://doi.org/10.1016/j.trc.2024.104749

Guimera, R., Nunes Amaral, L.A., 2005. Functional cartography of complex metabolic networks. Nature 433 (7028), 895-900.

Hou, G.Y., Jin, C., Xu, Z.D., Yu, P., Cao, Y.Y., 2019. Exploring evolutionary features of directed weighted hazard network in the subway construction. Chinese Physics B 28 (3), 038901. https://doi.org/10.1088/1674-1056/28/3/038901

Huang, D., Liang, T., Hu, S., Loughney, S., Wang, J., 2023. Characteristics analysis of intercontinental sea accidents using weighted association rule mining: Evidence from the Mediterranean Sea and Black Sea. OCEAN ENGINEERING 287, 115839. https://doi.org/10.1016/j.oceaneng.2023.115839

Lam, C.Y., Tai, K., 2020. Network topological approach to modeling accident causations and characteristics: Analysis of railway incidents in Japan. Reliability Engineering & System Safety 193, 106626. https://doi.org/10.1016/j.ress.2019.106626

Lan, H., Ma, X.X., Qiao, W.L., Deng, W.Y., 2023. Determining the critical risk factors for predicting the severity of ship collision accidents using a data-driven approach. Reliability Engineering & System Safety 230, 108934. https://doi.org/10.1016/j.ress.2022.108934

Lee, J.-S., Yu, Y.-U., 2023. Calculation of categorical route width according to maritime traffic flow data in the Republic of Korea. Journal of Marine Engineering & Technology 22 (5), 222-232. https://doi.org/10.1080/20464177.2023.2223396

Park, Y.A., Yip, T.L., Park, H.G., 2019. An Analysis of Pilotage Marine Accidents in Korea. Asian Journal of Shipping and Logistics 35 (1), 49-54. https://doi.org/10.1016/j.ajsl.2019.03.007

Shi, J., Liu, Z., Feng, Y., Wang, X., Zhu, H., Yang, Z., Wang, J., Wang, H., 2024. Evolutionary model and risk analysis of ship collision accidents based on complex networks and DEMATEL. OCEAN ENGINEERING 305, 117965. https://doi.org/10.1016/j.oceaneng.2024.117965

Sui, Z.Y., Huang, Y.M., Wen, Y.Q., Zhou, C.H., Huang, X., 2021. Marine traffic profile for enhancing situational awareness based on complex network theory. OCEAN ENGINEERING 241. https://doi.org/10.1016/j.oceaneng.2021.110049

Sui, Z.Y., Wen, Y.Q., Huang, Y.M., Song, R.X., Piera, M.A., 2023. Maritime accidents in the Yangtze River: A time series analysis for 2011-2020. Accident Analysis and Prevention 180, 106901. https://doi.org/10.1016/j.aap.2022.106901

Sui, Z.Y., Wen, Y.Q., Huang, Y.M., Zhou, C.H., Xiao, C.S., Chen, H.L., 2020. Empirical analysis of complex network for marine traffic situation. OCEAN ENGINEERING 214. https://doi.org/10.1016/j.oceaneng.2020.107848

Tong, Y., Zhen, R., Dong, H., Liu, J., 2023. Identifying influential ships in multi-ship encounter situation complex network based on improved WVoteRank approach. OCEAN ENGINEERING 284, 115192. https://doi.org/10.1016/j.oceaneng.2023.115192

Tuncel, A.L., Akyuz, E., Arslan, O., 2022. Quantitative risk analysis for operational transfer processes of maritime pilots. Maritime Policy & Management, 375-389. https://doi.org/10.1080/03088839.2021.2009133

Turna, İ., 2024. A safety risk assessment for ship boarding parties from fuzzy Bayesian networks perspective. Maritime Policy & Management 51 (1), 1-14. https://doi.org/10.1080/03088839.2022.2112780

Ugurlu, O., Kaptan, M., Kum, S., Yildiz, S., 2017. Pilotage services in Turkey: key issues and ideal pilotage. Journal of Marine Engineering and Technology 16 (2), 51-60. https://doi.org/10.1080/20464177.2016.1262596

Wang, H., Liu, Z., Wang, X., Graham, T., Wang, J., 2021. An analysis of factors affecting the severity of marine accidents. Reliability Engineering System Safety 210, 107513. https://doi.org/10.1016/j.ress.2021.107513

Wang, J., Wang, X., Feng, Y., Cao, Y., Guo, Z., Liu, Z., 2023a. Assessing the Connectivity Reliability of a Maritime Transport Network: A Case of Imported Crude Oil in China. Journal of Marine Science and Engineering 11 (8), 1597. https://doi.org/10.3390/jmse11081597

Wang, J., Zhou, Y., Zhuang, L., Shi, L., Zhang, S., 2023b. A model of maritime accidents prediction based on multi-factor time series analysis. Journal of Marine Engineering & Technology 22 (3), 153-165. https://doi.org/10.1080/20464177.2023.2167269

Wang, X., Xia, G., Zhao, J., Wang, J., Yang, Z., Loughney, S., Fang, S., Zhang, S., Xing, Y., Liu, Z., 2023c. A novel method for the risk assessment of human evacuation from cruise ships in maritime transportation. Reliability Engineering & System Safety 230, 108887. https://doi.org/10.1016/j.ress.2022.108887

Xi, Y., Hu, S., Yang, Z., Fu, S., Qin, T., 2019. The Moderating Effect of Risk Tolerance on the Hazardous Attitudes and Safety Behavior of Maritime Pilots: a Chinese Case. Liverpool, ENGLAND, Jul 14-17, 2019.

Xi, Y., Zhang, Q., Hu, S., Yang, Z., Fu, S., 2021. The effect of social cognition and risk tolerance on marine pilots' safety behaviour. Maritime Policy & Management 48 (1), 1-18. https://doi.org/10.1080/03088839.2020.1847338