Ocean Engineering 371 (2027) 129432

Contents lists available at ScienceDirect

Ocean Engineering

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

Research paper

A data-driven framework for spatiotemporal pattern mining of cyber incidents in digital maritime transport systems

Huanhuan Li [a,*], Jincheng Gong [b], Yunfeng Zhao [c], Zaili Yang [c,**]

[a] School of Engineering, University of Southampton, Southampton, UK
[b] School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China
[c] Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK

ARTICLE INFO

Keywords:
Maritime cyber incidents
Spatiotemporal pattern mining
Cyber risk analysis
Maritime transport systems

ABSTRACT

The rapid digitalisation of maritime transport systems has increased the exposure of vessels, ports, and offshore infrastructures to cyber threats, posing growing risks to operational safety, service continuity, and system resilience. Despite growing concern over maritime cybersecurity, limited research has examined the spatial and temporal patterns of cyber incidents at a system level. This study develops a data-driven Maritime Cyberattack Spatiotemporal Mining (MC-STM) framework to explore geographical concentrations, temporal associations, and co-occurrence patterns in reported maritime cyber incidents. An enhanced Maritime Cyber Attack Database (MCAD++) is established, containing 465 reported incidents up to December 2025, with structured Risk Influential Factors (RIFs) representing different aspects of cyber incidents. The framework combines association rule mining to identify recurring RIF combinations, transfer entropy to examine directional associations between annual RIF time series, and HDBSCAN and Moran's I to investigate spatial clustering patterns. The results identify three geographical clusters of reported incidents in North America, Europe and Asia under the selected HDBSCAN configuration. Association rule mining reveals recurring combinations of RIFs involving system availability and information integrity. Transfer entropy suggests several directional associations before multiple-testing correction, but none remain statistically significant after Benjamini–Hochberg false discovery rate correction. Global spatial autocorrelation in recorded consequence severity is weak and sensitive to the neighbourhood specification. Overall, the framework provides a structured analysis of geographical concentrations and recurring risk-factor combinations in reported maritime cyber incidents. The findings offer an empirical basis for maritime cybersecurity monitoring and the assessment of cyber resilience, while highlighting the need to account for reporting bias and uncertainty when interpreting historical incident data.

1. Introduction

The maritime industry is undergoing an unprecedented digital transformation driven by the widespread adoption of automation, connectivity, artificial intelligence, and data-driven technologies (Khan et al., 2026; H. Li et al., 2024). Modern vessels, ports (Belmoukari et al., 2023), and offshore infrastructures increasingly rely on interconnected cyber-physical systems, including integrated bridge systems (H. Chen et al., 2024), Automatic Identification Systems (AIS) (Yang et al., 2024), satellite communications, industrial control systems, port community systems, and digital logistics platforms. These technologies have significantly improved operational efficiency, situational awareness, and decision-making capabilities across maritime transportation systems (Shi et al., 2026). Against this backdrop, the concept of cyber seaworthiness has emerged (Jessen, 2026; Schinas and Metzger, 2023), extending conventional seaworthiness considerations to the cyber domain. In increasingly digital and automated maritime systems, this requires cyber risks to be managed in a manner consistent with the safe and reliable operation of vessels and associated maritime systems. Existing regulatory developments, including the International Maritime Organization (IMO) framework for maritime cyber-risk management and the IACS Unified Requirements, reflect the increasing importance of cyber resilience to operational safety and fitness for service. However, empirical evidence linking observed cyber-incident patterns with these

* Corresponding author.
** Corresponding author.
E-mail addresses: Huanhuan.Li@soton.ac.uk (H. Li), Z.Yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.oceaneng.2026.129432
Received 22 June 2026; Received in revised form 3 November 2026; Accepted 7 November 2026
Available online 12 November 2026
0029-8018/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

broader operational requirements remains limited.

Recent cyber incidents demonstrate that maritime cyberattacks are no longer isolated technical events but can directly undermine system reliability and operational continuity (Abdelmagid et al., 2025; Peng et al., 2025). For example, large-scale interference with Global Navigation Satellite System (GNSS) signals has caused widespread disruption of AIS-based vessel tracking, reducing situational awareness and increasing navigation and collision risks (Duzenli et al., 2025). Similarly, cyberattacks targeting port and customs information infrastructures have resulted in prolonged operational shutdowns, congestion, and cascading economic losses across maritime supply chains (Senarak, 2025). These incidents highlight that cyber threats in maritime transport increasingly manifest as system-level hazards, with the potential to propagate across spatially distributed and temporally interdependent components of global shipping networks.

Existing research on maritime cybersecurity has contributed valuable insights through literature reviews, qualitative analyses, and case-based investigations, clarifying attack typologies, affected assets, and potential consequences (Belmoukari et al., 2023; Soner, 2025). Rule-based and expert-driven modelling approaches, such as combinations of Failure Mode and Effects Analysis (FMEA) and Bayesian networks, have been widely adopted to formalise cyber risk reasoning in maritime contexts (Park et al., 2023). In parallel, event-based datasets, most notably the Maritime Cyber Attack Database (MCAD), have provided an important empirical foundation for analysing reported incidents globally (Pijpker et al., 2024). These efforts have significantly advanced understanding of maritime cyber risk; however, they remain limited in their ability to characterise how cyber threats evolve, interact, and unfold across space and time.

Three key limitations can be identified in the current literature (Clavijo Mesa et al., 2024; Peng et al., 2025; Yu et al., 2023). First, the empirical basis for spatiotemporal risk modelling remains weak, as most studies rely on relatively small, manually curated datasets that restrict reproducibility and large-scale analysis (Li et al., 2023; Li and Yang, 2023). Second, interactions among multiple Risk Influential Factors (RIFs), such as attack type, targeted system, operational consequence, and organisational context, are often examined in isolation, with limited attention to their joint dependency structures and recurring combinations (Zhao et al., 2025). Third, existing analytical frameworks rarely integrate temporal dependency modelling with spatial pattern analysis, limiting their capacity to capture non-random clustering, directional dependencies, and emergent system-level behaviour (Ali et al., 2025; Aydin et al., 2025). These limitations constrain the development of reliability-oriented risk assessment methods capable of supporting early warning, maintenance planning, and safety-informed decision-making in complex maritime systems.

To address these research gaps, this study develops a data-driven Maritime Cyberattack Spatiotemporal Mining (MC-STM) framework using an enhanced Maritime Cyber Attack Database (MCAD++). The framework integrates association rule mining, transfer entropy analysis, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Moran's I statistics to investigate maritime cyber incidents from complementary interaction, temporal, and spatial perspectives. Specifically, it examines recurrent co-occurrence patterns among RIFs, explores directional temporal associations among annual RIF series, and investigates geographical concentrations of reported incidents and spatial associations in incident severity. By applying these analytical methods to a common structured incident dataset, the framework provides an integrated empirical basis for interpreting maritime cyber risk in relation to operational resilience and the broader objective of maintaining cyber seaworthiness in increasingly digitalised maritime transport systems.

The main contributions of this study are fourfold, each corresponding to a specific research question and analytical method.

(1) Co-occurrence patterns: Association rule mining is used to identify recurrent combinations of RIFs and construct an interaction network that characterises their co-occurrence patterns in reported maritime cyber incidents.
(2) Exploratory temporal associations: Transfer entropy analysis is applied to annual RIF time series to examine potential directional associations. The resulting network provides an exploratory view of temporal relationships, although none of the identified associations remains statistically significant after correction for multiple testing.
(3) Spatial concentration: HDBSCAN is employed to identify density-based geographical clusters of reported maritime cyber incidents and to characterise their regional differences in incident concentration and dominant risk factors.
(4) Spatial autocorrelation of severity: Global and Local Moran's I statistics are used to assess the spatial distribution and autocorrelation of incident severity. Local results are evaluated with Benjamini–Hochberg (BH) False Discovery Rate (FDR) correction to account for multiple testing.

The structure of this paper is organised as follows: Section 2 reviews the existing literature on maritime cybersecurity and spatiotemporal data mining, highlighting key advances and remaining research gaps. Section 3 describes the proposed spatiotemporal pattern-mining framework and analytical methods. Section 4 presents the data and experimental results in detail. Section 5 discusses the findings, their theoretical and practical implications, and the limitations of the study. Finally, Section 6 concludes the paper and outlines directions for future research.

2. Literature review

This section reviews existing research on maritime cybersecurity and spatiotemporal data mining, with particular attention to the analytical methods used to investigate cyber-risk factors and incident patterns. It then identifies the limitations of existing approaches and establishes the research gaps addressed by the proposed framework.

2.1. Review of the studies on maritime cybersecurity

Recent studies have considered how conventional notions of seaworthiness should evolve as vessel operations become increasingly dependent on digital, connected, and automated systems. Schinas and Metzger (2023) discuss the need to reconsider seaworthiness in the context of highly automated shipping, while Dogan (2025) and Balci et al. (2024) emphasise the growing importance of cyber resilience and cyber-risk management to vessel fitness for service. The growing emphasis on cyber resilience is also reflected in maritime regulatory requirements. IMO Resolution MSC.428(98) requires cyber risks to be addressed within existing safety management systems under the ISM Code, while subsequent industry requirements have further incorporated cyber resilience into ship design and operation (IMO, 2017). In this study, cyber seaworthiness is therefore used as a broader operational context for interpreting maritime cyber risk, rather than as a directly measured outcome.

Empirical research on maritime cybersecurity has expanded in recent years, with increasing attention to the collection, classification, and analysis of reported cyber incidents (Afenyo and Caesar, 2023; Alcaide and Llave, 2020; Bielawski and Lazarowska, 2022; Miller et al., 2025). Meland et al. (2021) analysed and summarised 46 maritime cybersecurity incidents between 2010 and 2020. Li et al. (2024) conducted an in-depth analysis of threats in critical systems within the maritime industry, examining real-world cyber events impacting the sector. Pijpker et al. (2024) systematically collected and catalogued all known publicly available cybersecurity incidents affecting the GMTS, compiling the data into the MCAD.

As the maritime industry deepens its understanding of cybersecurity, an increasing number of studies focus on assessing cybersecurity risks and identifying risk factors within maritime transport systems. Bayesian networks (BNs) (Li et al., 2014) can effectively handle uncertainty and information gaps, providing quantitative support for risk assessments. They have thus become one of the key methods for maritime cybersecurity risk analysis. For example, Park et al. (2023) proposed a new maritime cyber threat risk assessment framework, combining FMEA and RBN to evaluate the risk levels of identified threats and determine which threats contribute most to overall maritime cybersecurity risks. Mohsendokht et al. (2024) introduced a data-driven BN approach for cybersecurity risk analysis, utilising recorded cyber events over the past two decades to identify the factors that most significantly affect maritime safety. Aydin et al. (2025) proposed a risk assessment method for cybersecurity threats in Maritime Autonomous Surface Ships (MASS) operations, integrating an improved Z-value with BN modelling to address uncertainty and enhance the reliability of expert judgements.

Failure Mode, Effects, and Criticality Analysis (FMECA) (Sun et al., 2023) is used to systematically identify potential failures and their impacts, prioritising them based on severity, occurrence probability, and detectability. As its advantages in risk management have become increasingly recognised, FMECA has also been introduced into the maritime cybersecurity field. For example, Amro and Gkioulos (2023) combined FMECA with the Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework to propose a new method for cybersecurity risk assessment. This approach introduces adversarial attack techniques, compensating for the limitations of traditional FMECA in Cyber-Physical Systems (CPS) risk assessments. Uflaz et al. (2024) further integrated FMECA parameters with BN modelling and utilised the Dempster-Shafer theory to incorporate expert opinions, providing a systematic evaluation of network risks in bridge navigation systems in the maritime sector.

In addition to FMECA and BN models, other methods have also been proposed to assess maritime cybersecurity. Kanwal et al. (2024) evaluated the relationships between six key dimensions affecting cybersecurity performance in the maritime industry using Structural Equation Modelling (SEM). Their findings provide important guidance for strengthening cybersecurity practices in the shipping industry and offer directions for future maritime cybersecurity research. Soner et al. (2024) combined the Shipboard Operation Human Reliability Analysis (SOHRA) method with the National Institute of Standards and Technology (NIST) cybersecurity framework to conduct a quantitative assessment of vessel network resilience and human-related risks.

In recent years, Deep Learning (DL) methods have also been introduced to maritime cybersecurity research. Kumar et al. (2023) developed a DL framework for modelling Cyber Threat Intelligence (CTI) and identifying threat types. This framework effectively detects threats and recognises attack types by automatically extracting potential patterns from the IoT-enabled Maritime Transportation System (MTS) network.

Maritime cybersecurity risk has become an increasing concern for both the global shipping industry and academia. As maritime cyber threats become increasingly complex, there is a growing need for analytical approaches that complement established methods, such as BN, FMECA, SEM, and DL, by capturing interactions among risk factors and their temporal and spatial characteristics. Simultaneously, more effective risk management and mitigation strategies are being proposed, helping shipping companies enhance their cybersecurity defences and ensure the safety and reliability of maritime transport.

2.2. Review of spatiotemporal mining methods in maritime applications

Maritime spatiotemporal data-mining techniques have been widely applied to various scenarios, such as vessel trajectories, piracy incidents, maritime accidents, traffic flow, and environmental pollution. Through spatiotemporal analysis, these techniques can effectively capture vessel interactions, the spatial distribution of piracy activities, risk factors associated with maritime accidents, spatiotemporal patterns of pollution emissions, and traffic flow dynamics, thereby driving innovative applications in maritime safety and environmental protection.

Ship trajectory prediction, combined with maritime spatiotemporal data-mining techniques, can identify potential interactions between vessels during their voyages by analysing their spatiotemporal patterns, thereby enabling more accurate predictions of future positional changes. This not only provides crucial data-driven support for maritime management but also enhances maritime safety. For example, Feng et al. (2024) proposed an association rule-based spatiotemporal co-occurrence pattern-mining algorithm to uncover the potential spatiotemporal co-occurrence behaviour of large groups of vessels during their voyages. Wang et al. (2024) combined spatiotemporal interaction awareness and multimodal trajectory prediction to model temporal, spatial, and cross-domain features, thus achieving accurate trajectory predictions. You et al. (2020) introduced the Spatiotemporal-Sequence-to-Sequence (Seq2Seq) method for real-time short-term ship trajectory prediction.

With ongoing technological advancements, researchers have begun to utilise Graph Neural Networks (GNNs) to capture interaction information between vessels, further improving prediction accuracy. For instance, Zhang et al. (2023) employed the Gated Spatiotemporal Graph Aggregation Network (G-STGAN) to address the issues of GNNs' reliance on predefined graph structures to capture vessel static attributes and the inability of Recurrent Neural Networks (RNNs) to capture long-term temporal dependencies. The Graph Attention Network (GAT), based on GNNs, has proven advantageous as a tool for dynamic interaction modelling, especially in capturing complex interactions between vessels. For example, Wang et al. (2023) combined Long Short-Term Memory (LSTM) with spatiotemporal-aware GAT for trajectory prediction. Gong et al. (2025) used a time attention module and Transformer-based Graph Convolution (TGConv) to capture spatiotemporal vessel dynamics, employing Gaussian heatmaps for probabilistic trajectory modelling and a Vision Transformer (ViT) to capture trajectory uncertainty. Furthermore, Zhang et al. (2024) applied the Dynamic SpatioTemporal Graph Attention Network (DSTGAN) for Multi-Vessel Trajectory Prediction (MVTP), effectively addressing the challenges of dynamic interactions and long-term pattern modelling between vessels.

In the study of piracy incidents, spatiotemporal analysis can identify trends and high-risk scenarios, providing data support for effective prevention strategies. For instance, Tsioufis et al. (2024) applied the Analytic Hierarchy Process (AHP) and spatiotemporal analysis to investigate the spatiotemporal patterns of piracy attacks and uncover key factors influencing pirate behaviour. Liang et al. (2024) analysed trends in piracy events from spatial, temporal, and spatiotemporal perspectives, using visualisation tools to identify clusters, outliers, and hotspots, providing data support for anti-piracy policy decision-making. Meanwhile, Fan et al. (2023) applied Kernel Density Estimation (KDE), k-means clustering, and buffer zone analysis to visually analyse piracy events from 2010 to 2022, improving the understanding of the spatial patterns of piracy. Additionally, Li and Yang (2023) developed a piracy event database and employed Fast Adaptive Dynamic Time Warping (FADTW) to uncover spatiotemporal patterns of piracy and used DBSCAN clustering to identify high-risk areas for pirate activities.

In maritime accident research, spatiotemporal techniques play a crucial role, especially in analysing the root causes of accidents and assessing their risks. Li et al. (2024) proposed a data-driven risk model to study the underlying causes and mitigation measures for global vessel collision accidents. Guo et al. (2025) introduced the Human Systems and spatiotemporal Maritime Accident Causal Chain Datasets (HS-MACs) to analyse the impact of human errors on accidents, revealing the interrelationships between human factors, accident environments, and spatiotemporal elements through correlation and causal analysis.

Spatiotemporal techniques provide effective tools for in-depth analysis of maritime environmental pollution and ecological impacts, particularly in the areas of vessel emissions and marine pollution

monitoring. For example, Huang et al. (2017) applied spatiotemporal analysis to explore the spatial distribution and variation patterns of vessel emissions, uncovering key information such as high-pollution areas, peak emission times, and the most environmentally harmful vessel types. X. Chen et al. (2024) used the Ship Traffic Emission Assessment Model (STEAM2) to estimate pollution emissions from different types of merchant vessels under various operational conditions and visualised the spatiotemporal distribution of emissions using ArcGIS. Additionally, Jayathilaka et al. (2022) assessed the spatiotemporal distribution of plastic particles along the Sri Lankan coastline, particularly focusing on the spread of plastic particles following the MV X-Press Pearl incident, exploring the link between pollution levels and upper ocean dynamics, which was used to track future plastic particle leaks.

In maritime traffic flow research, spatiotemporal techniques play a significant role, particularly in improving the accuracy of traffic flow predictions and understanding complex traffic patterns. Yuan et al. (2025) proposed an end-to-end Marine Traffic Flow (MTF) prediction framework based on AIS data, integrating time and spatial features to enhance the accuracy of maritime traffic flow predictions. Ma et al. (2024) combined a GAT to extract spatial features and used a Temporal Convolutional Network (TCN) and a Bidirectional Gated Recurrent Unit (BiGRU) to capture temporal features, significantly enhancing prediction accuracy and stability.

Existing studies demonstrate the broad application of spatiotemporal mining across maritime domains, including vessel trajectory analysis, piracy, maritime accidents, environmental monitoring, and traffic-flow prediction. As summarised in Table 1, the methods range from association-rule mining and density-based clustering to spatiotemporal neural networks and graph-based models, reflecting the different analytical requirements of these applications.

Maritime spatiotemporal mining techniques have demonstrated great potential across multiple fields, not only enhancing maritime safety, environmental protection, and management capabilities but also providing a scientific basis for predicting and preventing various maritime risks. As technology continues to develop, maritime spatiotemporal mining will continue to drive innovative applications in related fields, contributing to the sustainable development of global shipping and marine environments.

2.3. Research gaps

Despite the growing attention to maritime cybersecurity, several important gaps remain in understanding how cyber incidents develop across increasingly digitalised maritime transport systems. Previous studies have provided valuable insights into cyber threats, system vulnerabilities, and individual incidents, but the interaction, temporal, and spatial characteristics of maritime cyber risk have received comparatively less attention. In particular, four research gaps can be identified.

Table 1
Spatiotemporal mining methods in maritime.

Fields	References	Methods
Ship trajectory	Feng et al. (2024)	Association-rule co-occurrence mining
	Wang et al. (2024)	Multimodal spatiotemporal prediction
	You et al. (2020)	Spatiotemporal Seq2Seq model
	Zhang et al. (2023)	Gated Spatiotemporal Graph Aggregation
	Wang et al. (2023)	LSTM with spatiotemporal GAT
	Gong et al. (2025)	Temporal attention, TGConv and ViT
	Zhang et al. (2024)	Dynamic Spatiotemporal GAT
Piracy incident	Tsioufis et al. (2024)	AHP-based piracy spatiotemporal analysis
	Liang et al. (2024)	Spatiotemporal piracy visualisation
	Fan et al. (2023)	KDE and k-means clustering
	Li and Yang (2023)	FADTW + DBSCAN piracy analysis
Marine environment	Huang et al. (2017)	Spatiotemporal ship emission mapping
	Chen et al. (2024)	STEAM2 emission assessment
	Jayathilaka et al. (2022)	Plastic pollution distribution
Traffic flow	Yuan et al. (2025)	End-to-end AIS traffic flow
	Ma et al. (2024)	GAT, TCN and BiGRU fusion

(1) Limited availability of structured incident data for system-level cyber-risk analysis

Maritime cyber incidents are reported through heterogeneous and largely unstructured sources, which limits their use for systematic quantitative analysis. Many previous studies have relied on relatively small incident collections. For example, Meland et al. (2021) analysed 46 incidents, while earlier versions of the MCAD contained approximately 160 cases. Pijpker et al. (2024) expanded MCAD to over 290 incidents, although much of the information remained unstructured. More recent studies by Zhao et al. (2026c, 2026b) introduced LLM-based information extraction, with the resulting data primarily used for BN analysis. There remains a need for structured incident data that can support the joint analysis of risk-factor interactions and their temporal and spatial characteristics. To address this need, the present study develops MCAD++, comprising 465 reported incidents up to December 2025 and ten standardised RIFs, providing a consistent empirical basis for interaction and spatiotemporal analysis.

(2) Insufficient understanding of recurring interactions among maritime cyber-risk factors.

Maritime cyber incidents rarely arise from a single factor in isolation. Attack characteristics, affected systems, operational conditions, and incident consequences may occur together in recurring combinations. Existing studies have mainly examined individual factors or their dependencies within predefined risk models. Park et al. (2023), for example, used FMEA and Bayesian Networks but treated factors independently, while Mohsendokht et al. (2024) decoded dependencies using BN but focused on probabilistic contributions rather than co-occurrence patterns. Zhao et al. (2026a) further highlighted the fragmented treatment of digitalisation, cyber risk, and governance in smart-port research. However, empirical evidence on which RIFs repeatedly co-occur across historical maritime cyber incidents remains limited. Identifying such patterns is important for understanding combinations of conditions associated with cyber incidents rather than considering individual risk factors independently.

(3) Limited exploratory analysis of temporal associations among maritime cyber incidents

The temporal dimension of maritime cyber risk remains comparatively underexplored. Incident records are commonly analysed as independent observations, providing limited insight into whether changes in particular risk factors may be associated with subsequent changes in others. Soner (2025), for example, used Fuzzy Cognitive Mapping to examine port cyber-risk scenarios, while other modelling approaches have relied largely on simulated or synthetic scenarios. Empirical investigations of directional temporal associations using historical maritime cyber incident records remain limited. Exploratory analysis of annual RIF time series may therefore help identify potential directional statistical associations and generate hypotheses for further investigation. However, such associations require validation using longer time series and appropriate statistical testing before they can be regarded as established temporal dependencies.

(4) Limited characterisation of spatial concentration and severity patterns

The geographical dimension of maritime cybersecurity also remains insufficiently characterised. Previous studies have reported regional differences in incident frequency, but have paid less attention to whether incident locations form identifiable spatial concentrations and whether incident severity shows significant spatial dependence. For example, Mohsendokht et al. (2024) compared maritime cyber incidents by region without formally examining spatial autocorrelation. This distinction is relevant to maritime transport because cyber exposure and incident consequences may vary across ports, shipping regions, and other geographically concentrated maritime activities. A combination of location-based clustering and spatial autocorrelation analysis is therefore needed to distinguish concentrations of incident occurrence from local patterns in incident severity.

These gaps highlight the need for a data-driven understanding of the interaction, temporal, and spatial patterns of cyber incidents in digital maritime transport systems.

3. Methodology

This section introduces the overall analytical framework and its main components, including structured feature extraction, RIF interaction network construction, exploratory temporal association analysis using transfer entropy, HDBSCAN-based spatial clustering, and spatial autocorrelation analysis using Moran's I. These methods examine RIF co-occurrence patterns, nominal directional temporal associations, and spatial patterns in reported maritime cyber incidents.

3.1. The proposed framework

The framework aims to examine interaction, temporal, and spatial patterns in reported maritime cyber incidents through an integrated data-driven approach. It consists of four steps.

(1) Structured feature extraction and validated dataset construction.

This step establishes the foundation for analysis by transforming unstructured incident reports into a standardised dataset. The structured RIFs are extracted.

(2) RIF co-occurrence network analysis of cyber risk factors.

Association rule mining is applied to the structured dataset to identify recurrent co-occurrence patterns among RIFs. The extracted rules are used to construct a weighted directed interaction network, which is examined using graph-theoretic measures to characterise the structural positions and associations of individual RIFs.

(3) Exploratory temporal association analysis using transfer entropy.

Transfer entropy is applied to annual RIF time series to examine nominal directional statistical associations. An exploratory transfer entropy network is constructed from associations meeting the uncorrected permutation-based significance and stability criteria. The resulting network is used to characterise potential temporal association patterns and generate hypotheses for further investigation.

(4) Spatial clustering and spatial autocorrelation modelling.

HDBSCAN is used to identify geographical concentrations of reported incidents, while Global and Local Moran's I statistics are applied

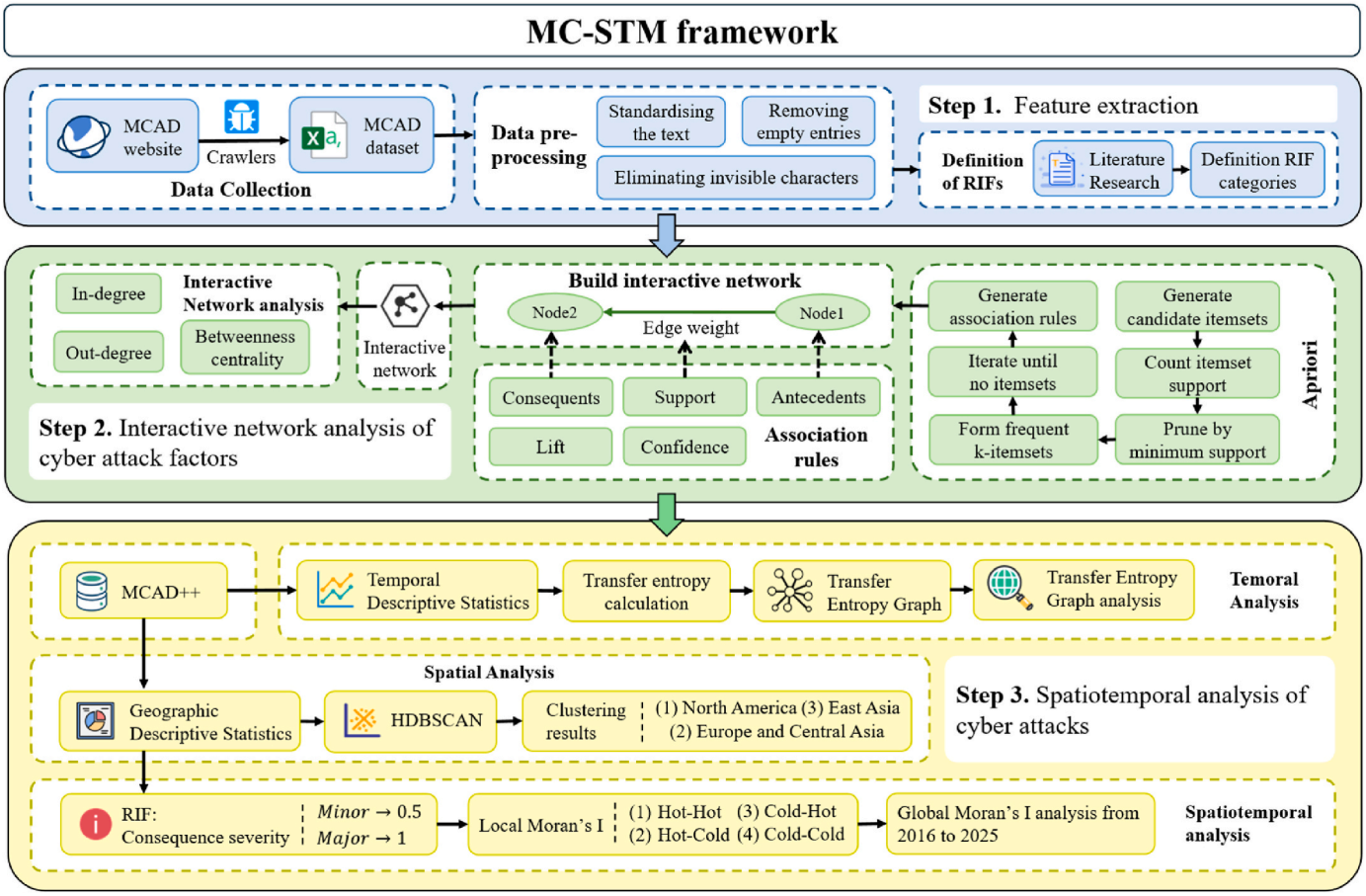


Fig. 1. Overall framework.

to examine spatial autocorrelation in incident severity. These analyses distinguish patterns of incident occurrence from spatial associations in severity within the available records.

As illustrated in Fig. 1, the MC-STM framework integrates structured RIF extraction, interactive network analysis, and spatiotemporal analysis to characterise patterns in reported maritime cyber incidents. The resulting evidence may support cybersecurity monitoring and risk communication in the maritime sector.

3.2. Feature extraction and dataset construction

The structured empirical foundation for this study is built upon the MCAD (Pijpker et al., 2024), published by NHL Stenden's Maritime IT Security research group. As one of the most comprehensive publicly available repositories of maritime cyber incidents, MCAD provides valuable information regarding attack events affecting vessels, ports, offshore infrastructures, and other maritime operational environments.

Several inconsistencies were identified within the original records. For example, phishing incidents were occasionally categorised as data breaches, while distributed denial-of-service (DDoS) attacks were broadly labelled as hacking events. Such inconsistencies may obscure the underlying relationships among cyber threats, affected maritime assets, operational consequences, and geographic locations. To support systematic cyber risk analysis, the original MCAD was therefore further refined and transformed into an enhanced analytical dataset, referred to as MCAD++.

A total of 465 maritime cyber incident reports, collected up to December 2025, were processed through a structured data integration procedure comprising three stages.

(1) Data ingestion and preprocessing. Raw incident reports were collected, standardised, and screened for completeness. Duplicate records, incomplete entries, and formatting inconsistencies were removed to ensure data quality and analytical consistency. This process established a reliable empirical basis for subsequent risk pattern identification.

(2) Definition of RIFs. To represent the multidimensional characteristics of maritime cyber incidents, a unified schema of ten RIFs was developed. These factors capture key dimensions associated with cyber risk formation and propagation within maritime transport systems, including exploited asset, target, cyber threat type, consequence type, attack intent, consequence severity, attack origin, attacking country/region, victim country/region, and incident timing, as illustrated in Fig. 2. Drawing upon established cyber risk and maritime security taxonomies (Curti et al., 2023), each factor was associated with a predefined set of states to ensure semantic consistency across incident records. In addition, "Other" and "Unknown" categories were retained to preserve information completeness and avoid artificial classification bias.

(3) Structured mapping and dataset consolidation. Building upon the preliminary extraction framework proposed by Zhao et al. (2026c, 2026b), each incident report was systematically re-examined and mapped to the newly defined RIF structure using an LLM-based classification pipeline. Classification quality was further assessed using five LLM judges on randomly sampled incident records. The ensemble assessment score was 83.63%, with an incorrect rating rate of 1.96%. The mean pairwise quadratic-weighted Cohen's kappa was 0.727. These results reflect LLM-based assessment rather than validation against human-expert annotations. Detailed results are presented in Tables S1 and S2 of the Appendix.

The resulting MCAD++ dataset provides a standardised representation of maritime cyber incidents that supports the analysis of RIF co-occurrence structures, directed statistical associations, and

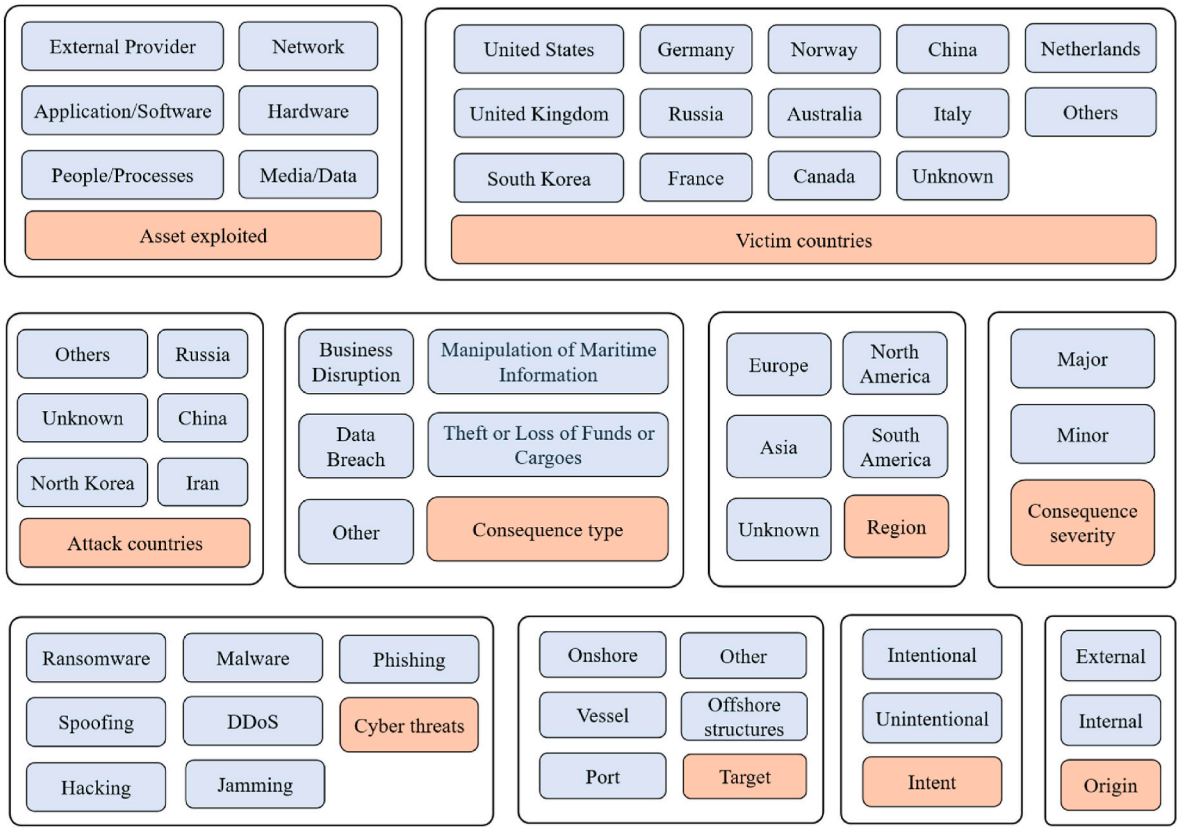


Fig. 2. RIF categories.

spatiotemporal distributions. For the spatial analyses, geographic coordinates refer to the reported location of the affected maritime asset, as recorded in MCAD, rather than to the presumed origin of the attacker. Records with missing or invalid geographic coordinates were excluded from the corresponding spatial analyses.

By transforming heterogeneous incident narratives into a structured and analytically consistent representation, MCAD++ enables systematic investigation of interactions among cyber-risk factors and their temporal and spatial characteristics across maritime transport systems. It therefore provides the common empirical basis for the interaction, temporal, and spatial analyses conducted within the proposed MC-STM framework.

To facilitate reproducibility and further research, the complete analysis code is publicly available at the project GitHub repository https://github.com/jc-gong/maritime-cyberattack-ST-mining. The original MCAD remains accessible at https://maritimecybersecurity.nl.

3.3. RIF interaction network construction

The Apriori algorithm (Chen et al., 2026) can mine frequent patterns from maritime cybersecurity data, thereby identifying RIFs with high co-occurrence. By converting the seven RIFs (e.g., Target, Asset exploited, Cyber threats, Consequence severity, Consequence type, Intent, and Origin) into item sets, frequent combinations are filtered using support and confidence, thereby extracting recurrent patterns. Victim countries and attacking countries/regions were excluded from the Apriori analysis because their high spatial granularity and sparse distributions tend to generate highly localised association patterns. While these variables are valuable for geographic hotspot analysis, their inclusion in association rule mining may reduce the ability to identify recurrent and transferable cyber co-occurrence patterns. The analysis therefore focuses on RIFs that better capture common threat characteristics and system-level patterns across maritime transport systems. The calculation method for association rules is shown in Algorithm 1, with the core formula being:

$$S(X) = \frac{N_X}{N} \tag{1}$$

$$C(X \Rightarrow Y) = \frac{S(X \bigcap Y)}{S(X)} \tag{2}$$

$$L(X \Rightarrow Y) = \frac{C(X \Rightarrow Y)}{S(Y)} \tag{3}$$

where $S(X)$ represents the support of the itemset X, $C(X \Rightarrow Y)$ denotes the confidence of the rule $X \Rightarrow Y$, and $L(X \Rightarrow Y)$ indicates the lift of the rule. These metrics effectively measure the co-occurrence relationships between different RIFs, providing a basis for constructing systematic association patterns. Following existing research methods (Xie et al., 2025), association rules with a lift greater than 1 are retained for further analysis, as they indicate positive associations between the corresponding factors.

Algorithm 1. Apriori algorithm

Input: Transaction database D; minimum support threshold Sup_{min}; minimum confidence threshold $Conf_{min}$
Output: Set of frequent item sets F; set of association rules R
1. $L_1 \leftarrow \{\{i\} | support(\{i\}) \geq Sup_{min}\}, F \leftarrow L_1, k \leftarrow 2$
2. While $L_{k-1} \neq \varnothing$ **do**
3. $C_k \leftarrow Generate_candidates(L_{k-1}, k)$
4. for each transaction $t \in D$ **do**
5. for each candidate $c \in C_k$ **do**
6. If $c \in t$ **then** $count[c] \leftarrow count[c] + 1$
7. $L_k \leftarrow \{c \in C_k | count[c] \geq Sup_{min}\}, k \leftarrow k + 1, F \leftarrow F \cup L_k$
8. end while
9. $R \leftarrow \varnothing$
10. for each frequent itemset $f \in F$ with $|f| \geq 2$ **do**
11. for each non-empty subset $A \subset f$ **do**

(continued on next column)

(continued)

12. $B \leftarrow f - A$
13. If $count[f]/count[A] \geq Conf_{min}$ **then**
14. $R \leftarrow R \cup \{A \rightarrow B\}$

Based on the retained association rules, a weighted directed RIF interaction network is constructed to characterise the co-occurrence structure among RIFs. Each RIF is represented as a node, and a directed edge is established from the antecedent to the consequent of each rule. Rule confidence is assigned as the edge weight, such that higher weights represent stronger conditional associations between the corresponding RIFs.

Three node-level measures, out-strength, in-strength, and betweenness centrality, are used to characterise the structural positions of RIFs within the network. For a node i, out-strength is defined as

$$S_i^{out} = \sum_j w_{ij}, \tag{4}$$

where w_{ij} denotes the confidence weight of the directed edge from i to j. Out-strength therefore represents the aggregate confidence of association rules in which a RIF appears as an antecedent. Similarly, in-strength is defined as

$$S_i^{in} = \sum_j w_{ji}, \tag{5}$$

and represents the aggregate confidence of rules in which the RIF appears as a consequent.

For the calculation of betweenness centrality, the confidence weight w_{ij} of each retained directed edge $i \rightarrow j$ is converted into a distance cost as $d_{ij} = 1 / w_{ij}$, ensuring that stronger conditional associations correspond to shorter distances. Betweenness centrality is then calculated on the resulting distance-weighted network to quantify the extent to which a RIF lies on shortest paths connecting other nodes. Together, out-strength, in-strength, and betweenness centrality characterise the outgoing association strength, incoming association strength, and bridging role of individual RIFs within the interaction network.

3.4. Spatiotemporal analysis

3.4.1. Exploratory temporal association analysis using transfer entropy

The spatiotemporal analysis examines two complementary dimensions of reported maritime cyber incidents: temporal associations among RIFs and the geographical distribution of incident records. Transfer entropy is applied to annual RIF time series to explore potential directional associations (Bossomaier et al., 2016), while HDBSCAN is used to identify spatial concentrations of reported incidents (McInnes et al., 2017).

For the temporal analysis, annual incident counts were compiled for each RIF category to construct a set of time series. Transfer entropy was then estimated between pairs of these series to explore potential directional associations. As the dataset covers 2010–2025, each series contains only 16 annual observations. A history length of $k = 1$ was therefore used to limit the loss of observations associated with longer lag structures. Transfer entropy measures the additional information provided by the past values of one series about the subsequent values of another, beyond that contained in the target series' own history. For two RIF time series, X and Y, the transfer entropy from X to Y is defined as follows:

$$TE_{X \rightarrow Y} = \sum p\left(y_{t+1}, y_t^k, x_t^k\right) \log \frac{p\left(y_{t+1} | y_t^k, x_t^k\right)}{p\left(y_{t+1} | y_t^k\right)} \tag{6}$$

where y_{t+1} represents the value of sequence Y at time $t+1$, $x_t^k = (x_t, x_{t-1}, \ldots, x_{t-k+1})$, and $y_t^k = \left(y_t, y_{t-1}, \ldots, y_{t-k+1}\right)$ are the historical RIF sequences of length k.

To assess the statistical significance of transfer entropy, a permutation test (Chen et al., 2025) is conducted. The source sequence X is held fixed, while the temporal order of the target sequence Y is randomly permuted to generate a surrogate target Y'. Transfer entropy is then recalculated for each surrogate sequence, yielding $TE_{X\to Y'}$. This procedure is repeated $B = 500$ times using a fixed random seed (seed = 0) to construct the null distribution. The one-sided empirical p-value is calculated with finite-sample correction as:

$$p = \frac{r+1}{B+1} \tag{7}$$

where r is the number of surrogate transfer entropy values greater than or equal to the observed $TE_{X\to Y}$. A directed association was considered to meet the nominal, uncorrected threshold when its permutation p-value was below 0.05.

The robustness of the identified dependencies to the randomisation procedure is further assessed using five random seeds {0, 1, 42, 123, 2025}. For each candidate directed edge, a stability score is defined as the proportion of seeds for which the dependency remains statistically significant. Only edges meeting the nominal threshold across all five seeds were retained in the exploratory network. This criterion assesses sensitivity to the randomisation seed and does not correct for multiple testing. The transfer entropy value obtained under the reference configuration (seed = 0) is assigned as the edge weight. This stability criterion reduces the influence of seed-specific permutation outcomes on the resulting network structure.

To account for multiple testing, the BH procedure was applied to the 870 possible directed relationships among the 30 RIF nodes, with the FDR set to 0.05. With 500 permutations, the minimum attainable empirical p-value is $1/(B+1) \approx 1/501 \approx 0.002$. Under the BH procedure, at least one ordered p-value must satisfy $(p(k) \le (k \times 0.05)/870$ for any hypothesis to be rejected. Given the limited resolution of the permutation-based p-values, a sufficiently large group of small p-values would be required to satisfy the corresponding BH thresholds. In the present analysis, no relationships remained statistically significant after correction. The limited number of annual observations and the large number of candidate tests further constrain statistical power. Accordingly, the transfer entropy network is presented as an exploratory analysis based on uncorrected permutation p-values, rather than as evidence of statistically significant temporal dependencies

The resulting exploratory transfer entropy network represents each RIF as a node and each nominally retained directional association as a directed edge, with the corresponding transfer entropy value assigned as the edge weight. Three node-level measures, out-strength, in-strength and betweenness centrality, are used to characterise the structural positions of RIFs within the network.

For a node i, out-strength is denoted as

$$S_i^{out} = \sum_j TE_{i\to j}, \tag{8}$$

where the summation is taken over all retained outgoing edges. It therefore represents the aggregate transfer entropy associated with the outgoing directional associations of a given RIF.

In-strength is defined as

$$S_i^{in} = \sum_j TE_{j\to i}, \tag{9}$$

where the summation is taken over all nominally retained incoming edges. It represents the aggregate transfer entropy associated with the incoming directional associations of a given RIF.

Betweenness centrality characterises the extent to which a RIF occupies a bridging position along paths connecting other nodes in the exploratory transfer entropy network.

For the calculation of betweenness centrality, the transfer entropy weight of each retained edge is transformed into a distance cost so that larger transfer entropy values correspond to shorter distances. Specifically, for a retained directed edge $i\to j$ with $TE_{i\to j} > 0$, the distance is defined as $d_{i\to j} = 1/TE_{i\to j}$. Edges that do not satisfy the retention criteria are treated as absent. Betweenness centrality is then calculated on the resulting directed distance-weighted network to quantify the extent to which each RIF lies on shortest paths connecting other nodes.

These measures describe the outgoing, incoming, and bridging roles of individual RIFs within the exploratory transfer entropy network. The identified associations should not be interpreted as evidence of causal relationships.

3.4.2. HDBSCAN

The identification of spatial clustering patterns is also a crucial component of maritime cybersecurity research. Traditional density-based clustering methods often face issues with parameter sensitivity and unstable clustering results when dealing with data of non-uniform density. In contrast, HDBSCAN can identify robust spatial clustering structures under varying density conditions by constructing a mutual reachability distance graph and a hierarchical clustering tree (Shu et al., 2026). It evaluates all possible density thresholds and automatically selects the optimal clusters based on stability assessment.

As shown in Fig. 3, the geographic coordinates of reported maritime cyber incidents are used as the spatial input for clustering. Coordinates are represented in the WGS84 geographic coordinate system (EPSG:4326), and pairwise spatial distances are calculated using the Haversine metric to account for the curvature of the Earth. Records with missing or invalid coordinates, defined as longitudes outside [−180, 180] or latitudes outside [−90, 90], are excluded prior to clustering. HDBSCAN is then applied to identify spatial concentrations of reported incidents based on local variations in point density. The algorithm constructs a hierarchy from mutual-reachability distances and evaluates the persistence of candidate clusters across density levels, while observations that do not belong to sufficiently stable clusters are classified as noise. Cluster extraction is performed using the Excess of Mass (EOM) criterion, which favours clusters that remain stable over a broad range of density levels. First, the mutual reachability distance between event points is calculated based on the density relationship between the points, measuring their proximity. Then, a hierarchical clustering tree is built based on the computed distances, progressively merging points into clusters. Next, stability evaluation determines the optimal number and structure of clusters. Finally, points that do not meet clustering criteria are automatically marked as noise or outliers to exclude low-density event points. Through these steps, HDBSCAN identifies density-based geographical clusters of reported incidents while classifying observations outside sufficiently stable clusters as noise. This enables the identification of spatial concentrations without requiring the number of clusters to be specified in advance.

3.4.3. Moran's I spatiotemporal analysis

Moran's I (Tiefelsdorf and Boots, 1997) is used to measure the spatial autocorrelation of data and can be divided into Global Moran's I and Local Moran's I. Global Moran's I reflects the average level of spatial correlation in the overall pattern and is used to analyse the temporal trend of overall spatial autocorrelation in consequence severity across different years. Local Moran's I is used to identify high-value and low-value clustering areas, revealing spatial differences in the distribution of incidents at the local level.

In the analysis of maritime cybersecurity incidents, Local Moran's I can be used to identify spatial clusters and outliers in incident severity, providing information on the spatial structure of reported events. In this study, consequence severity was employed as the attribute variable c in the Local Moran's I analysis, representing the magnitude of operational disruption and system impact associated with maritime cybersecurity incidents. The variable is discretised as follows: Major is assigned a value of 1, and Minor is assigned a value of 0.5. Because Moran's I is invariant to non-zero linear transformations of the variable, the specific numerical

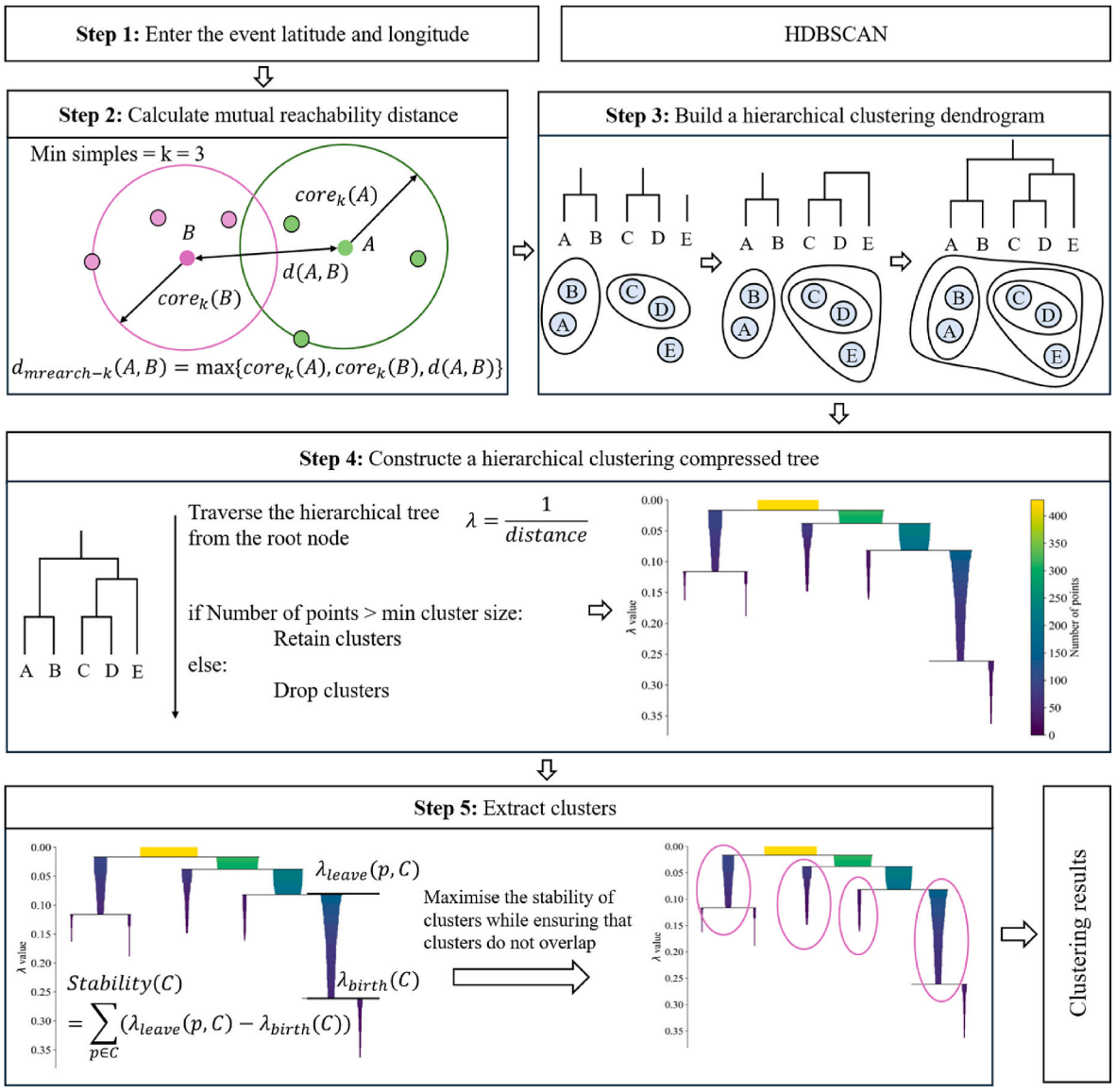


Fig. 3. HDBSCAN algorithm flowchart.

values assigned to the two categories do not affect the results.

The Global Moran's I is calculated as follows:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij}(c_i - \bar{c})(c_j - \bar{c})}{\sum_i (c_i - \bar{c})^2} \tag{10}$$

where N is the total number of spatial units, c_i and c_j are the consequence severity values for spatial units i and j, respectively, and $\bar{c}$ is the mean consequence severity of all events. w_{ij} is the spatial weight matrix, calculated using the K-Nearest Neighbour (KNN) method with row standardisation applied to ensure comparability across spatial units with different numbers of neighbours. The value of Global Moran's I ranges from [-1, 1], where positive values indicate positive spatial autocorrelation and negative values indicate negative spatial autocorrelation.

Local Moran's I is used to analyse the local correlation between an individual spatial unit and its neighbours, and is defined as:

$$I_i = \frac{(c_i - \bar{c})}{\sum_i (c_i - \bar{c})^2} \sum_j w_{ij}(c_j - \bar{c}) \tag{11}$$

where I_i is the Local Moran's I value for the i-th spatial unit. It measures the local autocorrelation of that unit within the overall spatial pattern. By using Local Moran's I, four types of spatial association patterns can be identified:

(1) High-High: A high-severity observation is surrounded by neighbouring observations with similarly high severity, indicating a local cluster of high severity.
(2) Low-Low: A low-severity observation is surrounded by neighbouring observations with similarly low severity, indicating a local cluster of low severity.
(3) High-Low: A high-severity observation is surrounded predominantly by low-severity neighbours, indicating a spatial outlier relative to its local context.
(4) Low-High: A low-severity observation is surrounded predominantly by high-severity neighbours, likewise indicating a spatial outlier relative to its local context.

Statistical significance for Local Moran's I was assessed using conditional permutation tests with 999 permutations and a fixed spatial weight matrix. Two-sided pseudo p-values were calculated for each valid location by comparing the observed Local Moran's I statistic with its

permutation distribution. The BH procedure was then applied to the valid local tests to control the FDR at $q = 0.05$. Only locations remaining significant after FDR correction were considered to exhibit statistically significant local spatial association; all other local patterns were treated as exploratory.

Global Moran's I was calculated for successive years to examine changes in the spatial autocorrelation of consequence severity, while Local Moran's I was used to identify the geographical distribution of local spatial associations. These analyses characterise the spatial structure of recorded incident severity and its variation over time, providing a basis for further investigation of regional patterns.

4. Experimental results and analysis

4.1. Data description

The MCAD++ dataset contains 465 cybersecurity incidents reported between 2000 and December 2025. Fig. 4 presents the temporal distribution of incidents across the study period. Only three incidents were recorded between 2000 and 2009, whereas the number of reported incidents increased substantially from 2010 onwards. This increase broadly coincides with the growing digitalisation and connectivity of maritime transport systems, although changes in incident reporting and data availability may also have contributed to the observed trend. Accordingly, the subsequent analyses focus on the 462 incidents reported between 2010 and 2025, which provide a more consistent empirical basis for examining contemporary maritime cyber-risk patterns.

Reported incidents peaked in 2023, followed by a decline in 2024 and 2025. This decline should be interpreted cautiously, as it may partly reflect variations in database update cycles and reporting practices rather than a genuine reduction in incident occurrence. To maintain temporal consistency and focus on the period most relevant to contemporary maritime digitalisation, subsequent analyses are restricted to incidents reported between 2010 and 2025. This period encompasses substantial growth in the adoption and integration of digital technologies across maritime transport systems and provides a consistent empirical window for examining temporal dependencies and spatial patterns in reported cyber incidents.

Based on the analysis of the RIFs, distributional visualisations were generated for Asset Exploited, Target, Consequence Severity, Consequence Type, Intent, Cyber Threats, Threat Origin, and Attack Countries, Region, and Victim Countries. As shown in Fig. 5, several notable distributional characteristics can be observed.

(1) Asset Exploited: Application/Software represents the largest category (40.69%), followed by Hardware (22.3%), Network (12.55%), and People/Processes (3.68%), indicating the predominance of software-related assets among the recorded incidents.
(2) Target: Shipping companies (23.59%), vessels (22.3%), and ports (20.13%) constitute the three largest identified target categories, highlighting the exposure of both maritime operational assets and associated organisations to cyber threats.
(3) Cyber Threat Type: Ransomware is the most frequently recorded cyber threat type (53.25%), followed by Spoofing (16.45%) and Hacking (9.52%), demonstrating the prominence of ransomware within the collected maritime cyber incident records.
(4) Consequence Type: Business Disruption (51.08%) and Data Breach (25.97%) are the two most frequently recorded consequence types, highlighting operational continuity and information security as major areas affected by the recorded incidents.
(5) Intent: Intentional incidents account for 99.13% of the dataset, whereas only 0.87% are classified as unintentional, indicating that the overwhelming majority of recorded incidents involve deliberate cyber activities.
(6) Consequence Severity: Incidents classified as Minor account for 55.41%, compared with 44.59% classified as Major, indicating that Minor consequences are more frequently recorded than Major consequences in the dataset.
(7) Threat Origin: External threats account for 98.92% of the recorded incidents, compared with 1.08% attributed to internal origins.
(8) Attack-Origin Country: The attack origin is unknown for 64.5% of the recorded incidents. Among the identified country categories, Russia represents the largest share (18.83% of all incidents), followed by China (6.06%), Iran (3.25%), and North Korea (2.16%). These distributions should be interpreted cautiously because attribution information is unavailable for a substantial proportion of incidents and cyberattack attribution itself is subject to uncertainty.
(9) Region: Europe (37.88%), Asia (25.54%), and North America (20.78%) account for the largest shares of recorded incidents, while other regions collectively account for 15.8%. These proportions characterise the geographical distribution of incidents captured in the database and should not be interpreted directly as regional cyber-risk levels because incident reporting and data availability may vary geographically.
(10) Victim Countries: Given the large number of affected countries, the four most frequently recorded identified victim countries are

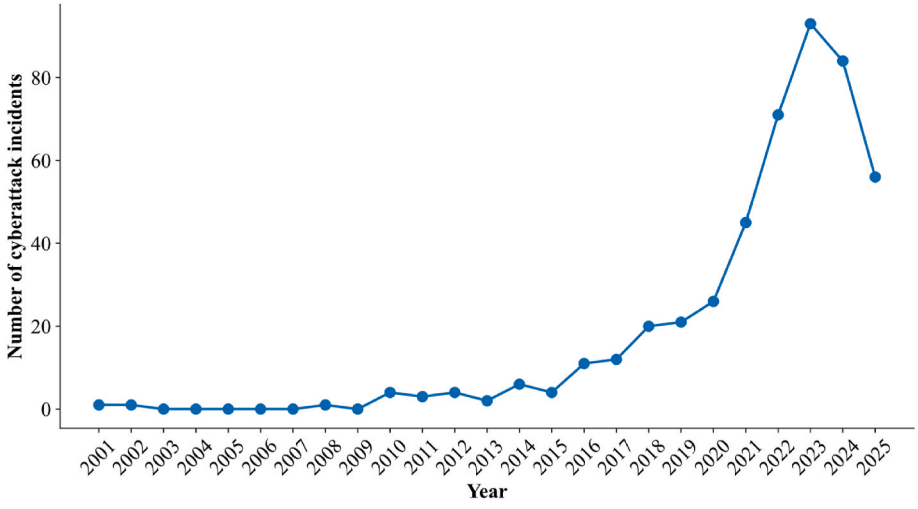


Fig. 4. Timeline of cyberattack events.

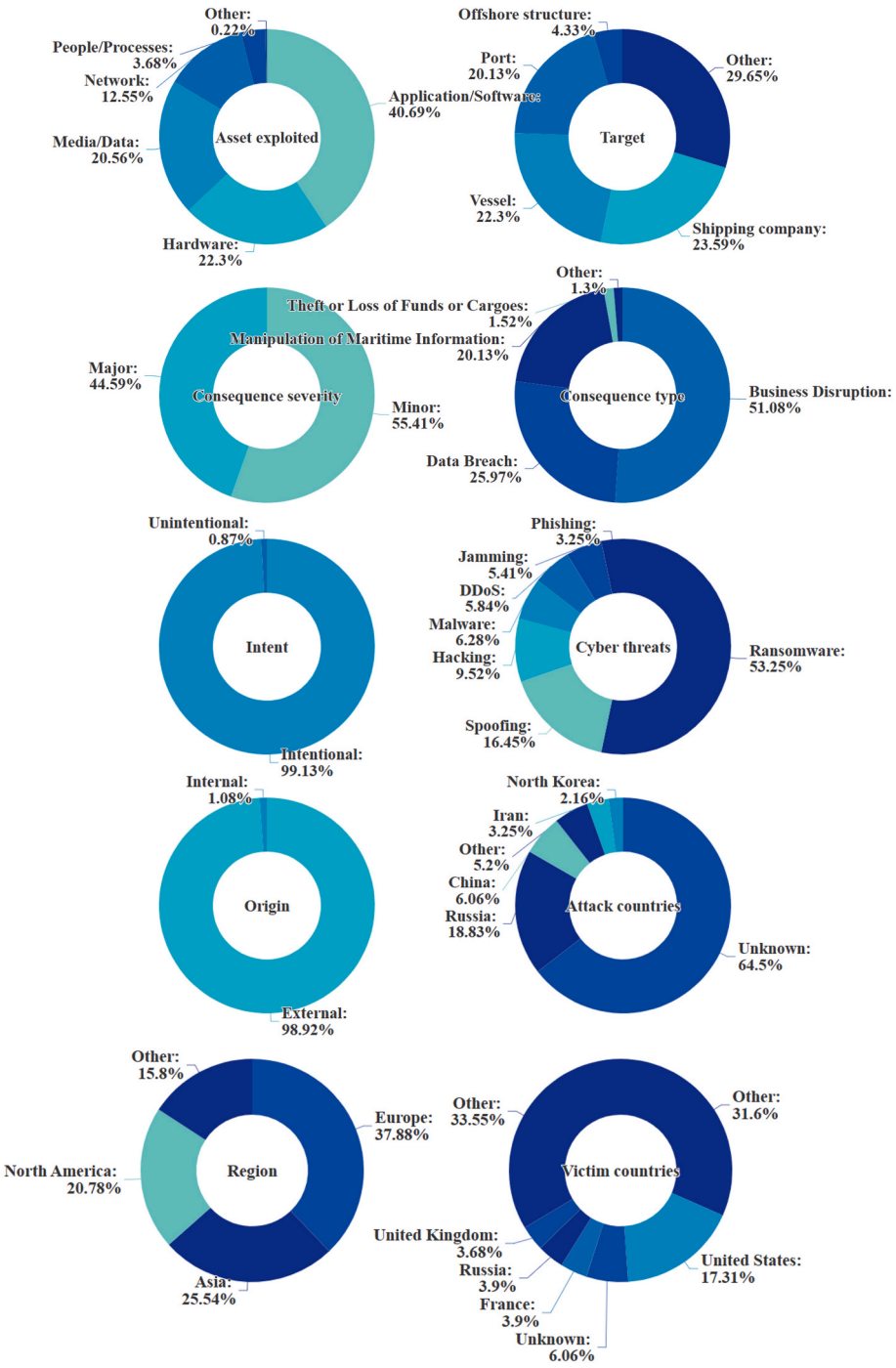


Fig. 5. RIF distribution of maritime cyberattacks.

presented separately, with the remaining identified countries grouped as "Others" and cases without country information classified as "Unknown". The United States represents the largest individual victim-country category (17.31%), followed by France (3.9%), Russia (3.9%), and the United Kingdom (3.68%). However, these distributions should be interpreted as characteristics of the reported incident dataset rather than direct measures of national cyber risk.

4.2. Apriori analysis

4.2.1. Parameter selection and method justification

The Apriori algorithm requires three threshold parameters: minimum support, minimum confidence, and minimum lift. To select an appropriate configuration, a grid search was performed across candidate support and confidence thresholds while holding the minimum lift at 1.0. Parameter selection considered both the number of retained rules and their mean lift, with the aim of preserving a sufficiently rich set of associations without retaining a large number of weak rules.

Specifically, minimum support was varied over {0.05, 0.10, 0.15, 0.20} and minimum confidence over {0.30, 0.40, 0.50, 0.60, 0.70}, with minimum lift fixed at 1.0. For each parameter combination, the number of retained rules and their mean lift were recorded (Table 2). Based on this sensitivity analysis, a minimum support of 0.15 and a minimum confidence of 0.60 were selected for the subsequent analysis. This configuration retained 52 rules with a mean lift of 1.759, providing a

Table 2
Sensitivity analysis of Apriori parameters.

Support	Confidence	No. of rules	Mean lift	Support	Confidence	No. of rules	Mean Lift
0.05	0.3	131	1.436	0.1	0.3	105	1.484
	0.4	117	1.465		0.4	97	1.502
	0.5	93	1.564		0.5	81	1.590
	0.6	71	1.677		0.6	60	1.728
	0.7	59	1.672		0.7	49	1.723
0.15	0.3	84	1.502	0.2	0.3	52	1.071
	0.4	79	1.515		0.4	52	1.071
	0.5	67	1.602		0.5	42	1.079
	0.6	**52**	**1.759**		0.6	30	1.095
	0.7	44	1.722		0.7	27	1.057

practical balance between the number of identified associations and their average strength. Lower confidence thresholds substantially increased the number of retained rules but generally yielded lower mean lift, whereas increasing the support threshold reduced the number of rules without a corresponding improvement in mean lift.

The selected thresholds were chosen to retain recurrent and sufficiently strong co-occurrence patterns. A minimum support of 0.15 requires each retained itemset to occur in at least 15% of the incident records, thereby excluding associations based on relatively infrequent co-occurrences. A minimum confidence of 0.60 requires the consequent to occur in at least 60% of records containing the antecedent. In addition, only rules with a lift greater than 1.0 are retained, indicating that the antecedent and consequent co-occur more frequently than would be expected under statistical independence.

To assess the suitability of Apriori for the present analysis, its results were compared with those obtained from chi-square tests of independence applied to the same dataset. At a significance threshold of $p < 0.05$, the chi-square analysis identified 14 statistically significant RIF pairs. In comparison, the algorithm initially generated 52 association rules under selected thresholds, of which 41 had lift greater than 1 and were retained for subsequent interactionnetwork analysis. The two approaches provide complementary information: chi-square tests assess whether pairs of categorical variables exhibit statistically significant dependence, whereas Apriori characterises recurrent co-occurrence patterns through support, confidence, and lift. Apriori also represents associations in rule form, allowing antecedent and consequent roles to be distinguished when constructing the RIF interaction network. This representation is particularly suited to the objective of identifying recurrent multi-factor patterns and characterising their network structure. Apriori was therefore adopted as the primary method for the RIF interaction analysis, with the chi-square results serving as a complementary comparison.

4.2.2. Association rule discovery and network analysis

By applying Apriori association rule mining (Wang et al., 2026) to the RIFs, Table 3 presents the top ten rules ranked by lift, revealing several key co-occurrence patterns in cybersecurity incidents. One of the strongest associations is observed between spoofing as a cyber threat and the manipulation of maritime information (Lift = 4.706). The high lift indicates that these two factors co-occur substantially more frequently than would be expected under statistical independence, while the high confidence for this rule shows that maritime information manipulation is frequently observed among incidents involving spoofing. This pattern highlights a strong empirical association between spoofing and information integrity risks in the reported incident data.

A similarly strong association is observed between hardware exploitation and maritime information manipulation (lift = 4.293). Their frequent co-occurrence points to the close connection between physical equipment and information functions in maritime systems, particularly where sensors, communication devices, and navigation equipment support the collection and exchange of operational information.

Table 3
Top ten association rules with the highest lift.

Antecedents	Consequents	Support	Confidence	Lift
Cyber threats-Spoofing	Consequence type-Manipulation of Maritime Information	0.156	0.947	4.706
Consequence type-Manipulation of Maritime Information	Cyber threats-Spoofing	0.156	0.774	4.706
Asset exploited-Hardware	Consequence type-Manipulation of Maritime Information	0.193	0.864	4.293
Consequence type-Manipulation of Maritime Information	Asset exploited-Hardware	0.193	0.957	4.293
Cyber threats-Spoofing	Asset exploited-Hardware	0.156	0.947	4.249
Asset exploited-Hardware	Cyber threats-Spoofing	0.156	0.699	4.249
Consequence type-Manipulation of Maritime Information	Target-Vessel	0.162	0.806	3.617
Target-Vessel	Consequence type-Manipulation of Maritime Information	0.162	0.728	3.617
Asset exploited-Hardware	Target-Vessel	0.173	0.777	3.484
Target-Vessel	Asset exploited-Hardware	0.173	0.777	3.484

Vessel targets are also strongly associated with maritime information manipulation (lift = 3.617). This pattern is consistent with the growing dependence of vessels on integrated navigation, communication, and monitoring systems, and indicates that information-related risks are closely associated with vessel operations in the reported incidents.

Overall, the strongest association rules are concentrated around spoofing, hardware exploitation, vessel targets, and maritime information manipulation (Table 3). These RIFs frequently occur together and, in several rules, appear in both antecedent and consequent positions. Rather than representing isolated relationships, they form a closely connected co-occurrence structure within the incident data. The direction of an association rule, however, distinguishes the antecedent from the consequent and should not be interpreted as evidence of temporal or causal dependence.

The prominence of these associations is consistent with the increasing integration of physical equipment and digital information systems in maritime operations. Navigation, situational awareness, and operational decision-making depend heavily on the integrity and availability of information generated and exchanged across these systems. Although spoofing, hardware vulnerabilities, and information manipulation have been examined in previous maritime cybersecurity studies (Rodriguez-Diaz et al., 2026; Timothy et al., 2026), the present

analysis quantifies how strongly these factors co-occur in reported incidents and shows how they are connected within the broader RIF interaction network. This provides an empirical view of recurring cyber-risk patterns in maritime transport without implying causal relationships among the identified factors.

After selecting association rules with Lift > 1, a cyberattack event RIF interaction network (Fig. 6) is constructed. In this network, antecedents and consequents are represented as nodes, while rule confidence is assigned as the weight of directed edges. The resulting weighted directed graph provides a representation of recurrent RIF associations, allowing the connectivity and structural positions of individual factors to be examined using network measures.

As illustrated in Fig. 7, out-degree, in-degree, and betweenness centrality characterise the structural positions of individual RIFs within the association-rule network. Weighted out-degree represents the aggregate strength of associations in which a RIF appears as an antecedent, whereas weighted in-degree represents the corresponding strength when it appears as a consequent. Betweenness centrality captures the extent to which a RIF lies on paths connecting other nodes. These measures distinguish RIFs according to their outgoing, incoming, and bridging roles within the network.

Media/Data (#13) has a high weighted out-degree (3.674), indicating strong aggregate associations with other RIFs when it appears as an antecedent. This prominent network position is consistent with the central role of data and information assets in maritime operations, where communication, navigation, and operational functions increasingly rely on digitally generated and exchanged information. Application/Software (#11) shows a similarly high weighted out-degree (3.596), reflecting its strong connectivity with other RIFs and the widespread integration of software-based functions into contemporary maritime systems.

Business Disruption (#7) also exhibits high weighted outgoing connectivity (3.369) and has a betweenness centrality of 0.024. The latter indicates that it occupies a bridging position between other RIFs, suggesting that Business Disruption contributes to the connectivity of otherwise less directly connected parts of the network.

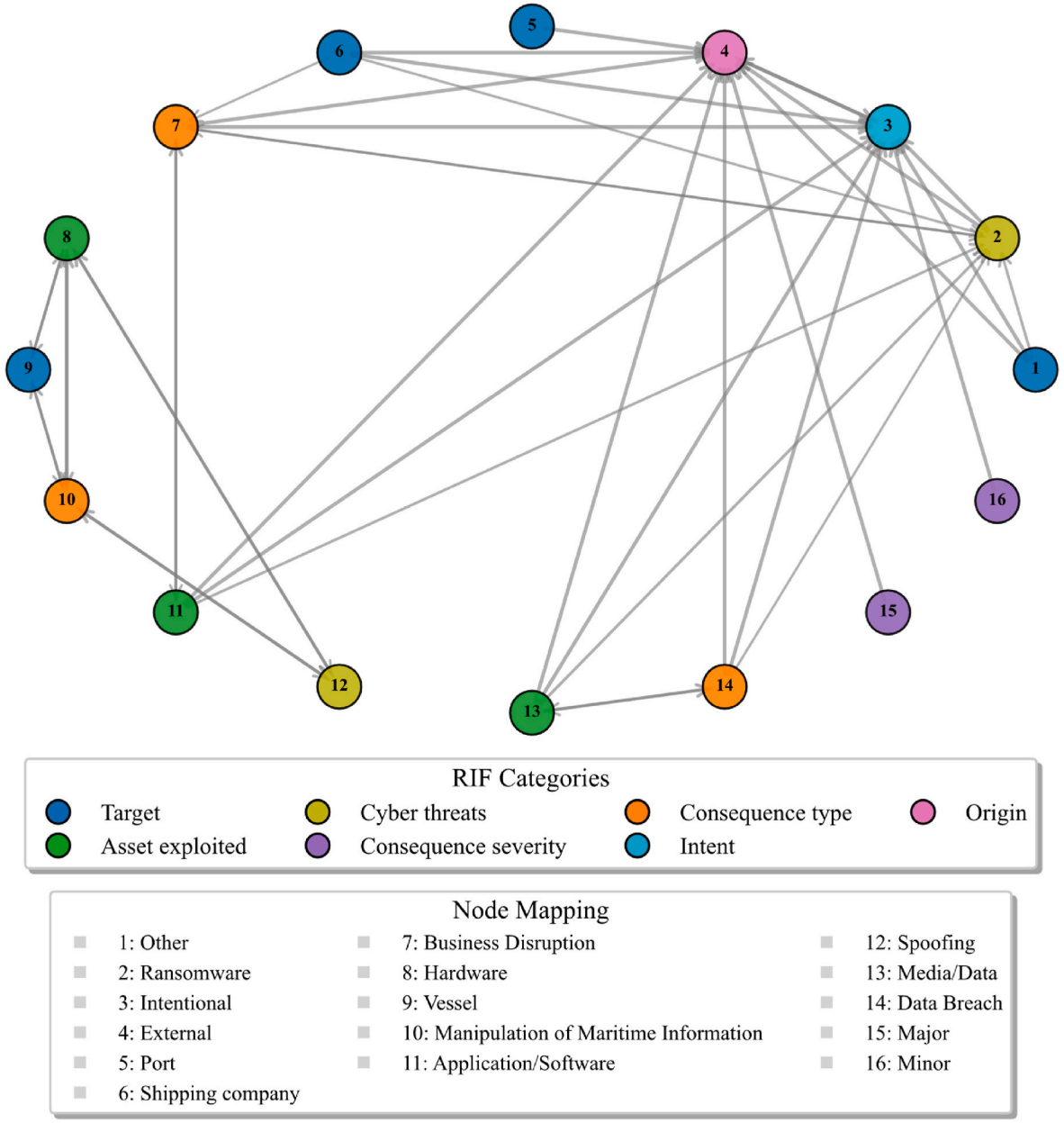


Fig. 6. Cyberattack event RIF interaction network.

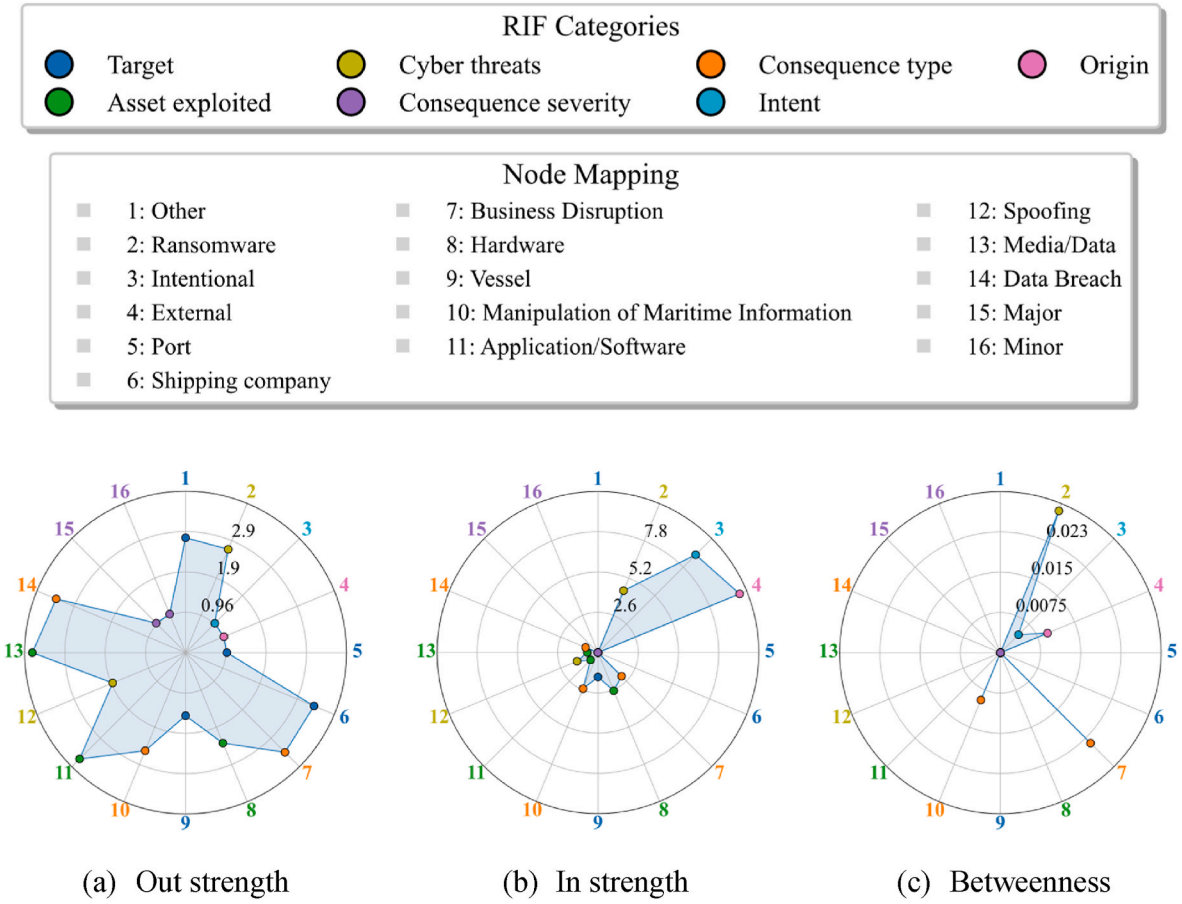


Fig. 7. RIF interaction network node out-degree, in-degree, and betweenness centrality.

In contrast, Origin-External (#4) and Intent-Intentional (#3) have the highest weighted in-degrees (9.964 and 8.976, respectively), indicating strong aggregate associations with RIFs appearing as antecedents in the retained rules. Ransomware (#2) differs in exhibiting substantial connectivity in both directions, with a weighted in-degree of 4.351 and a weighted out-degree of 2.671. Its presence in both antecedent and consequent positions indicates a broadly connected role within the association-rule network.

Overall, the centrality analysis reveals distinct structural roles among the RIFs. Media/Data, Application/Software, and Business Disruption are characterised by strong outgoing connectivity, whereas Origin-External and Intent-Intentional show particularly strong incoming connectivity. Ransomware occupies a more broadly connected position across the rule structure, while betweenness centrality identifies factors that bridge different parts of the network. These network-level patterns complement the individual association rules by showing how recurrent RIF associations are organised within the broader interaction structure.

4.3. Temporal analysis

Throughout this section, temporal changes are examined in terms of the frequency and distribution of maritime cyber incidents documented in MCAD++. These trends reflect changes in recorded incidents rather than necessarily indicating changes in the actual occurrence of maritime cyberattacks, as reporting practices, incident detection, public disclosure, media coverage and database completeness may vary over time.

Fig. 8 illustrates the temporal evolution of recorded maritime cyber incidents across the key RIFs. Regarding target types, both the recorded number and the relative representation of vessel-related incidents increased over the study period, indicating their growing prominence within MCAD++. However, this observed pattern should be interpreted with caution, as it may partly reflect changes in incident reporting and database coverage rather than an increase in the underlying frequency of attacks targeting vessels.

For ‘Asset exploited’, ‘Application/Software’ was the most frequently recorded exploited asset, with recorded incidents reaching their highest level in 2022. From 2023 onwards, the frequency of incidents involving Application/Software remained relatively stable, with a slight decline observed in 2024 and continuing through 2025.

The recorded number of ransomware incidents increased markedly over the study period, making ransomware the most frequently reported cyber threat category in MCAD++. Incidents classified as Minor and Major also increased. In 2025, Major incidents exceeded Minor incidents in the database for the first time. These trends reflect changes in reported incidents and may be affected by variations in reporting, disclosure, detection and database coverage. They do not necessarily indicate an increase in the overall frequency or severity of maritime cyber incidents.

Business disruption was the most frequently recorded consequence type, while incidents classified as intentional and external in origin accounted for the majority of records. These distributions describe the incidents captured by MCAD++ and should not be assumed to represent all maritime cyber incidents. Nevertheless, they provide a useful basis for identifying commonly reported threats and affected assets.

Fig. 9 presents the transfer entropy estimates between the annual RIF time series. The histories of Target–Shipping Company (#2), Target–Port (#4), All (#30), Intent-Intentional (#26) and Origin–External (#28) yielded relatively high TE estimates towards Cyber threats–Jamming (#17). These relationships met the uncorrected

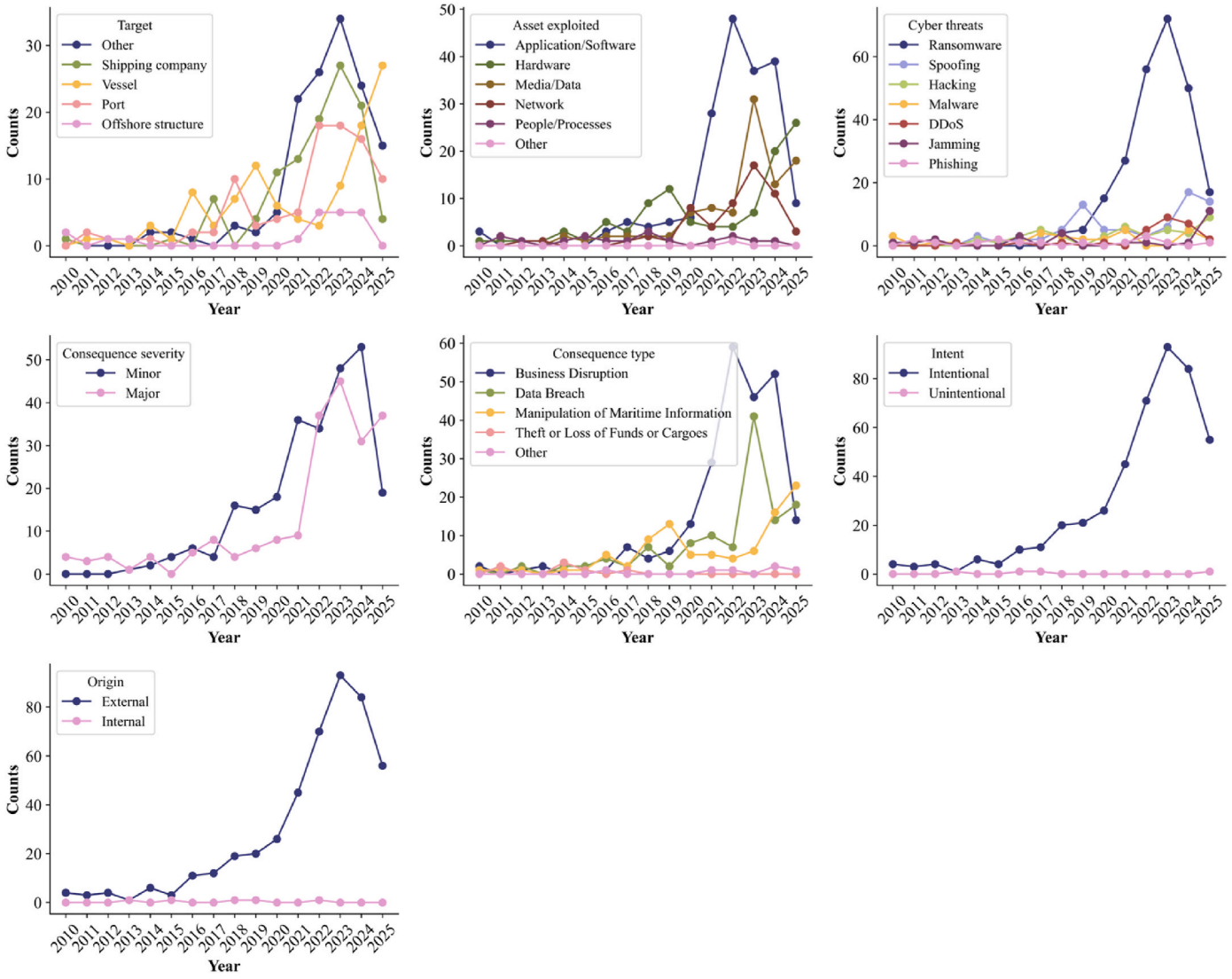


Fig. 8. Time evolution of cyberattack event features.

permutation-test threshold and were retained for exploratory analysis. However, none of the 870 directed relationships remained statistically significant after BH FDR correction. The results are therefore treated as exploratory rather than statistically confirmed directional dependencies.

The TE network was constructed by representing each RIF category as a node. A directed edge was included when the corresponding permutation test gave an uncorrected p-value below 0.05, and the result was retained across all five random seeds. Edge weights were defined by the estimated TE values. The resulting network provides a visual summary of directional associations identified by the exploratory analysis, rather than evidence of established information transfer or causal influence.

This analysis differs from the Apriori association analysis. Apriori identifies frequent combinations and co-occurrence relationships within incident records, whereas TE takes into account the ordering of observations in the annual time series. The two approaches therefore describe different aspects of the recorded data, although the TE findings remain subject to the limitations of the short annual series and multiple testing.

As shown in Fig. 10, the retained TE estimates were generally small. Given that none of the directed relationships survived FDR correction, the network should be regarded as a descriptive representation of the uncorrected results. It does not establish temporal dependencies or predictive relationships between RIF categories.

Fig. 11 summarises the out-strength, in-strength and betweenness centrality of nodes in the exploratory TE network. These measures describe the structure of the retained network and should be interpreted in relation to its uncorrected edges.

Asset exploited–Media/Data had the highest out-strength (2.171), representing the largest sum of outgoing edge weights. Target–Shipping Company, Target–Port, Intent–Intentional, Origin–External and All each had an out-strength of 1.500. Target–Vessel, Cyber threats–Spoofing and Consequence severity–Major had lower values of 1.367. These differences describe the distribution of outgoing edge weights in the exploratory network.

Cyber threats–Jamming had the highest in-strength (14.020), substantially exceeding that of the other nodes. This indicates that several retained incoming edges were directed towards Jamming and contributed to its aggregate edge weight.

Betweenness centrality was low across the network. Intent–Unintentional had the highest value (0.009), followed by Consequence type–Theft or Loss of Funds or Cargoes (0.006), Consequence type–Other (0.005) and Target–Vessel (0.004). These nodes lay on a relatively larger share of shortest paths within the retained

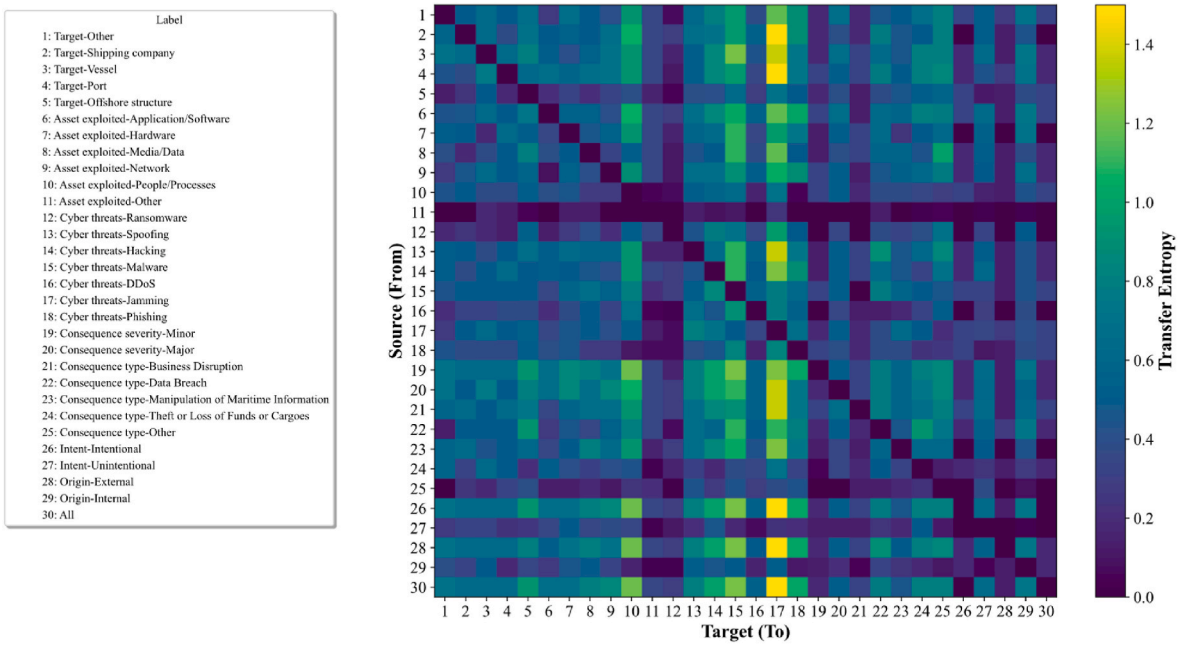


Fig. 9. Transfer entropy matrix.

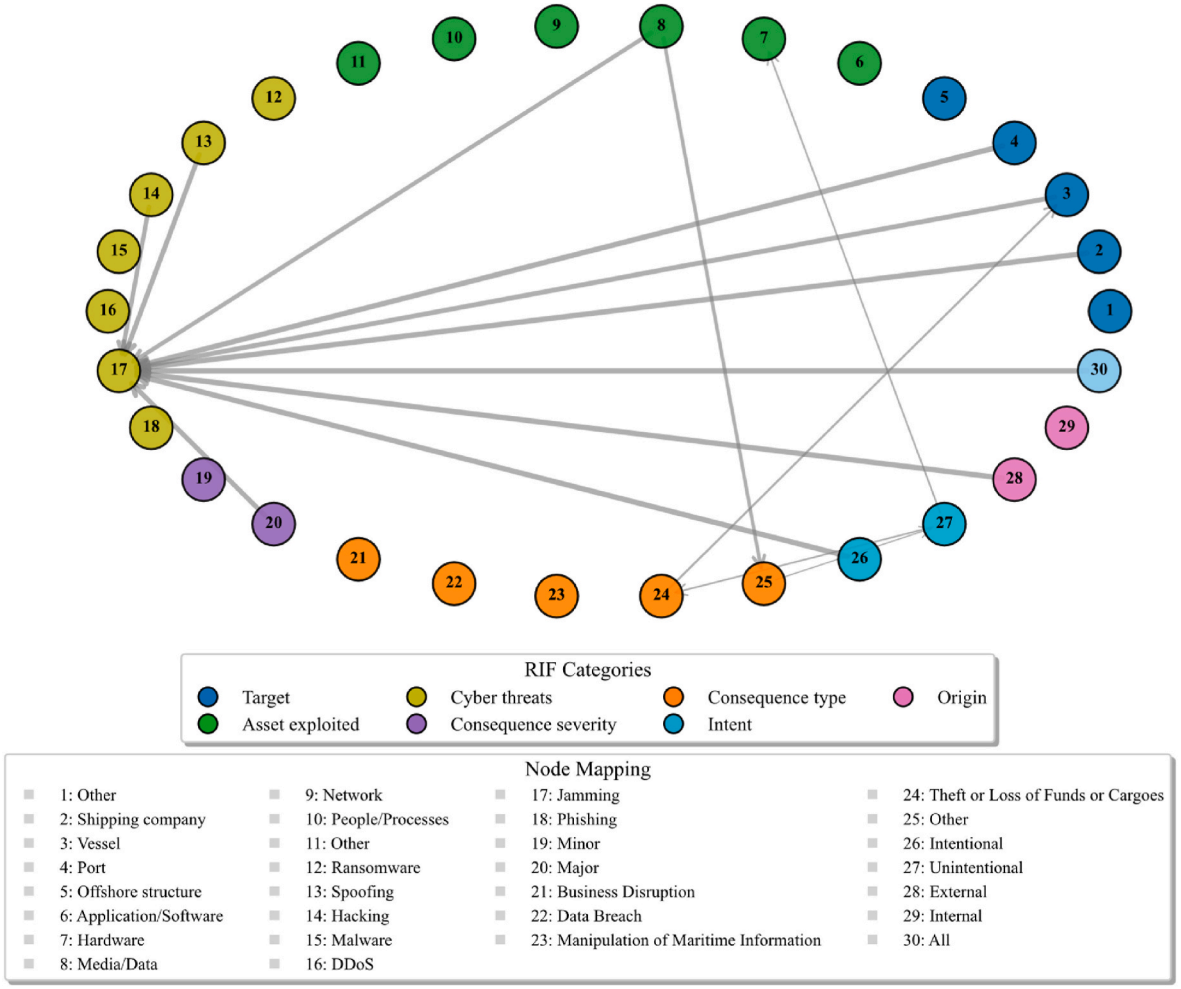


Fig. 10. Transfer entropy graph.

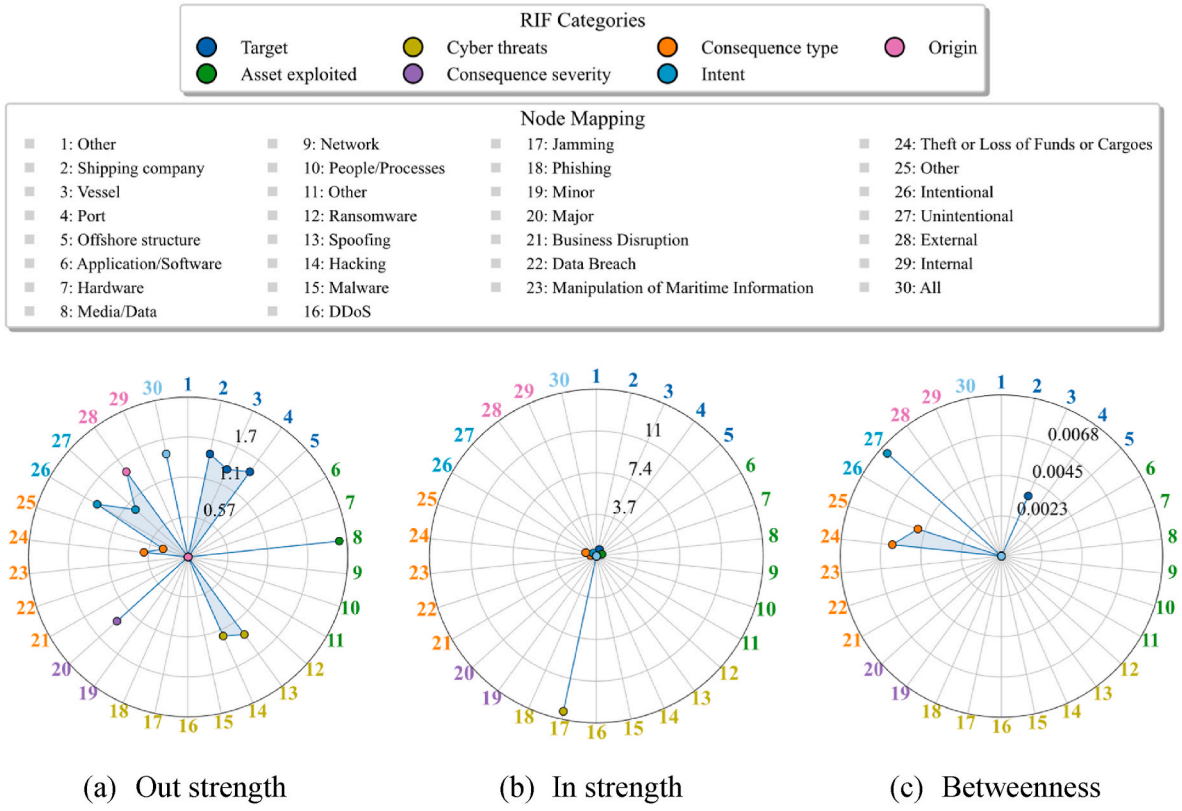


Fig. 11. Transfer entropy node features.

network, although their absolute centrality values were small.

Overall, the exploratory TE analysis identified differences in the distribution and connectivity of uncorrected directional associations among annual RIF time series. The network measures describe these patterns but do not establish statistically significant temporal relationships. As no directed edge remained significant after FDR correction, the TE findings are used only to inform further investigation. They complement the Apriori results by considering temporal ordering, without providing evidence of causal influence or incident propagation.

4.4. Spatial analysis

Using the geographic coordinates recorded in MCAD++, a global distribution map of reported maritime cyber incidents was generated (Fig. 12). These coordinates refer to the reported locations of affected assets, as documented in the original MCAD, rather than to the origins of the attacks. Reported incidents are predominantly concentrated in coastal areas, particularly around regions with major ports and maritime transport activity. This spatial pattern broadly coincides with the geographical concentration of maritime trade and logistics

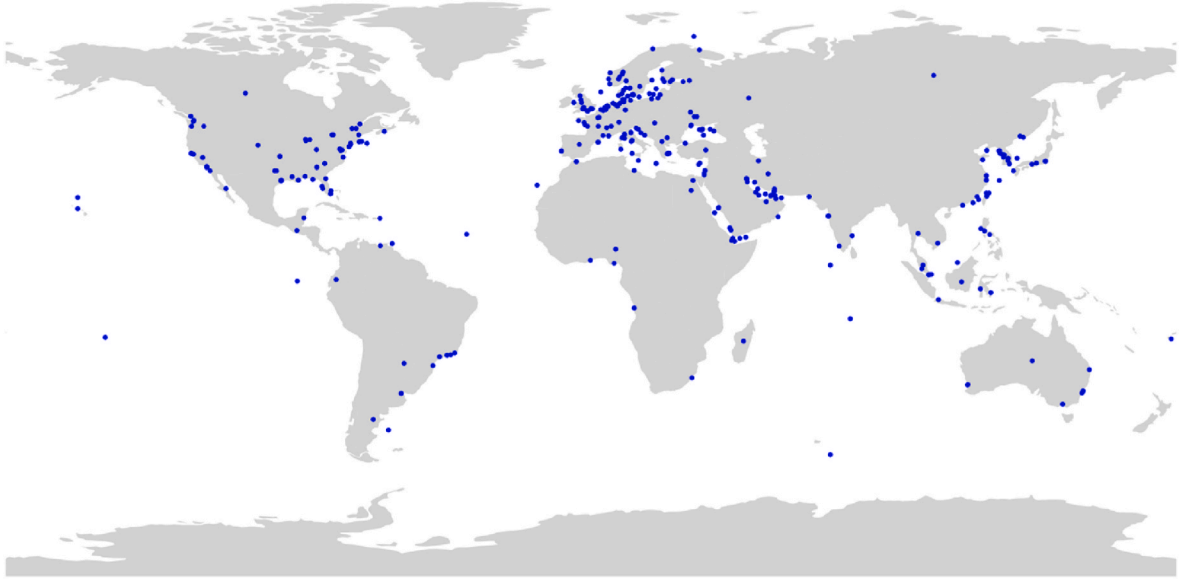


Fig. 12. Global distribution of cyberattacks.

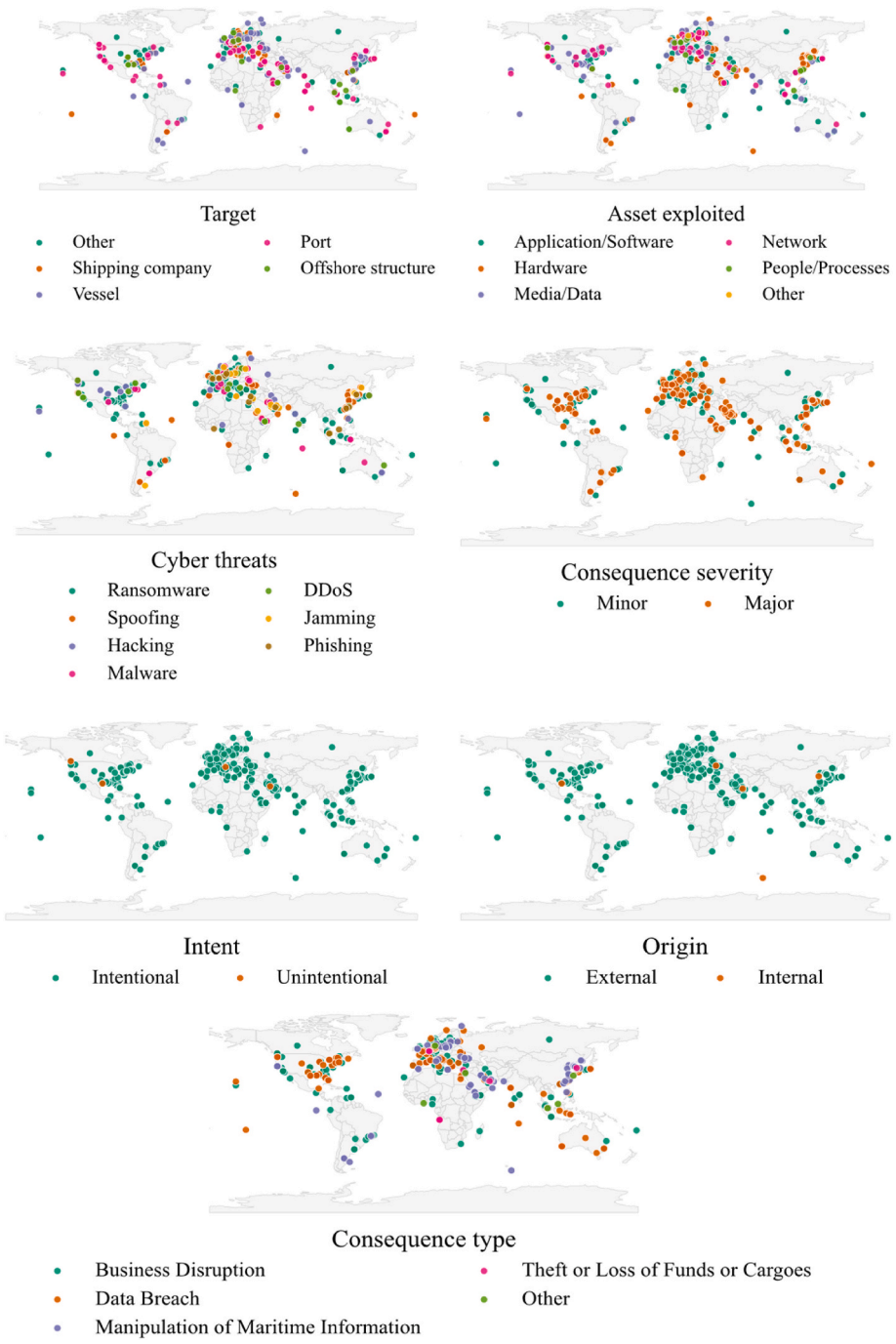


Fig. 13. Global distribution of cyberattacks RIF.

infrastructure, although regional differences in incident reporting and data availability may also contribute to the observed distribution.

Fig. 13 presents the geographical distribution of recorded RIF categories across the incident locations available in MCAD++. The records cover several major maritime regions, although their geographical representation is uneven. These patterns describe the spatial coverage of the dataset and should not be interpreted as direct measures of regional cyber-risk levels, given potential differences in reporting practices and data availability.

HDBSCAN was employed to identify spatial concentrations of maritime cyber incidents within the global maritime domain. Since the choice of min cluster size and min samples can substantially influence cluster stability and noise identification, a sensitivity analysis was conducted. Min cluster size was varied across {10, 20, 30, 40, 50}, while min samples were varied across {5, 10, 20, 30}. For each parameter combination, the number of clusters, noise ratio, and silhouette score were recorded (Table 4).

Silhouette scores were computed on clustered points only, excluding noise points labelled as −1. To select the preferred configuration, the following criteria were applied: (i) noise ratio $\leq$ 0.10, to avoid discarding an excessive proportion of incidents; (ii) at least three clusters, to support geographical interpretation; (iii) highest silhouette score among configurations satisfying criteria (i) and (ii); and (iv) when scores were tied, the smaller min cluster size was preferred to reduce the risk of excluding smaller but meaningful clusters. Under these criteria, the combination of min cluster size = 40 and min samples = 30 was selected, yielding three spatial clusters, a noise ratio of 0.100, and a silhouette score of 0.720. Although min cluster size = 10 and min samples = 10

Table 4
Sensitivity analysis of HDBSCAN parameters.

Min cluster size	Min samples	N clusters	Noise ratio	Silhouette score
10	5	15	0.301	0.628
10	10	14	0.442	0.736
10	20	5	0.169	0.679
10	30	4	0.188	0.693
20	5	6	0.11	0.579
20	10	5	0.08	0.577
20	20	5	0.169	0.679
20	30	4	0.188	0.693
30	5	4	0.093	0.606
30	10	4	0.069	0.568
30	20	4	0.152	0.666
30	30	4	0.188	0.693
40	5	3	0.082	0.705
40	10	3	0.056	0.691
40	20	3	0.087	0.717
40	**30**	**3**	**0.1**	**0.72**
50	5	3	0.082	0.705
50	10	3	0.056	0.691
50	20	3	0.087	0.717
50	30	3	0.1	0.72

produced a higher silhouette score (0.736), its noise ratio was 0.442, which would exclude a large proportion of incidents and is therefore unsuitable for global spatial concentration analysis. Min cluster size = 50 and min samples = 30 produced identical metrics. However, the larger min cluster size may suppress smaller clusters, so min cluster size = 40 was preferred. The selected configuration therefore offers a balanced trade-off between cluster quality, noise exclusion, and geographical interpretability.

To justify the choice of HDBSCAN over traditional density-based clustering, a comparative experiment was conducted against DBSCAN using the same input data and Haversine distance metric. HDBSCAN produces three clusters with a noise ratio of 0.100 and a silhouette score of 0.720, whereas DBSCAN yields nine clusters with a substantially higher noise ratio (0.299) and a lower silhouette score (0.606). These results indicate that HDBSCAN achieves better cluster compactness while assigning a larger proportion of incidents to meaningful spatial clusters, demonstrating its advantage over DBSCAN for this application.

Based on this configuration, three spatial clusters and a set of noise points were identified. These clusters represent the principal concentrations of reported maritime cyber incidents within the study period and provide the basis for subsequent regional characterisation and comparative analysis. As shown in Fig. 14, Cluster 1 (100 incidents) is located in North America, Cluster 2 (241 incidents) in Europe and Central Asia, and Cluster 3 (75 incidents) in Asia, while 46 incidents are classified as noise. The comparative analysis indicates that regions with highly digitalised maritime infrastructures and dense port networks, such as Europe and North America, tend to exhibit higher reported concentrations of cyber incidents. This observation, however, should be interpreted with caution given the potential influence of regional differences in reporting practices.

Fig. 15 presents the temporal distribution of maritime cyber incidents within the three identified spatial clusters. Distinct temporal patterns can be observed across regions. Cluster 1 (North America) shows relatively low incident counts from 2010 to 2020, followed by a sharp increase in 2021 (8 incidents), peaking at 24 incidents in 2023, then declining to 8 incidents in 2025. Cluster 2 (Europe) exhibits a sustained upward trend throughout the study period, with notable increases after 2018 (8 incidents in 2018, 44 in 2024) and reaching its highest level of 47 incidents in 2023 before a slight decline in 2024-2025. Cluster 3 (Asia) displays more modest incident counts overall, with a gradual increase from 2019 onwards, peaking at 10-13 incidents in 2021-2024. The variation in temporal trends across the three clusters may reflect differences in regional operating conditions, reporting practices and database coverage.

The temporal pattern in Europe (Cluster 2) warrants particular attention. Reported incidents increased markedly after 2021, from an average of 3.0 incidents per year in 2010–2018 and 15.7 in 2019–2021 to 44.3 in 2022–2024, with 47 incidents recorded in 2023. The increase was also greater than that observed over the same periods in North America and Asia. This divergence highlights the need to consider region-specific factors when interpreting the increase in reported incidents in Europe, alongside broader changes in digitalisation, maritime activity, and incident reporting. The period from 2022 onwards also coincides with a substantial deterioration in the regional security environment following Russia's invasion of Ukraine. Maritime areas around the Baltic Sea, including the Gulf of Finland and waters in the vicinity of Kaliningrad, have experienced repeated reports of GNSS jamming and spoofing, raising concerns over the reliability of positioning and navigation services for maritime operations (Miętkiewicz, 2025; Westbrook, 2023). Related studies and official reports have also documented heightened cyber and electronic interference affecting transport and critical infrastructure in the wider European security environment (Neumann, 2026; Tombak et al., 2025). These developments provide a relevant geopolitical context for interpreting the increase in reported

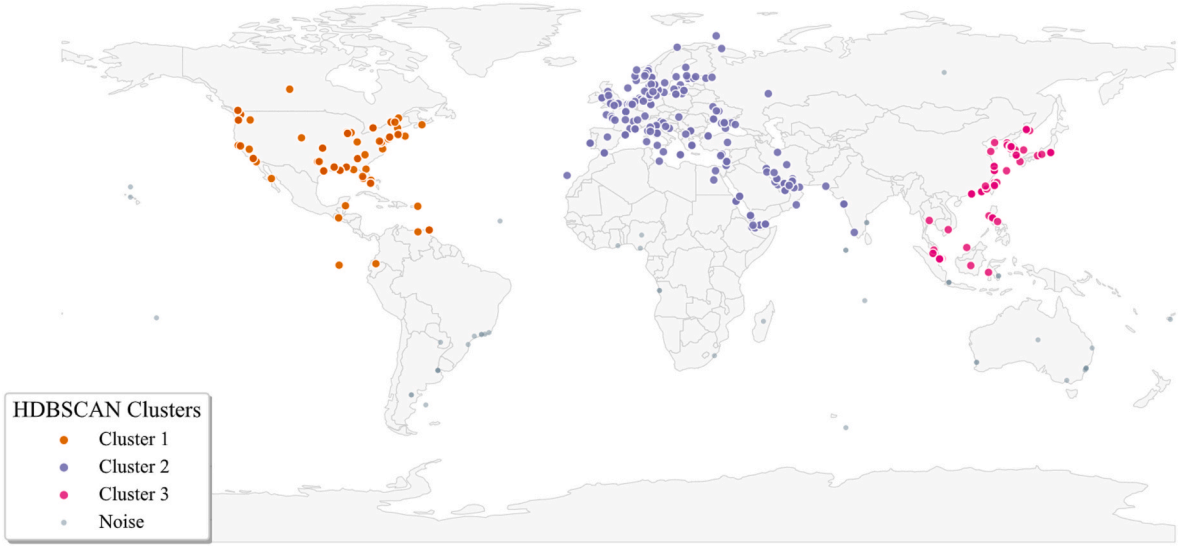


Fig. 14. Global distribution clustering results of cyberattack incidents.

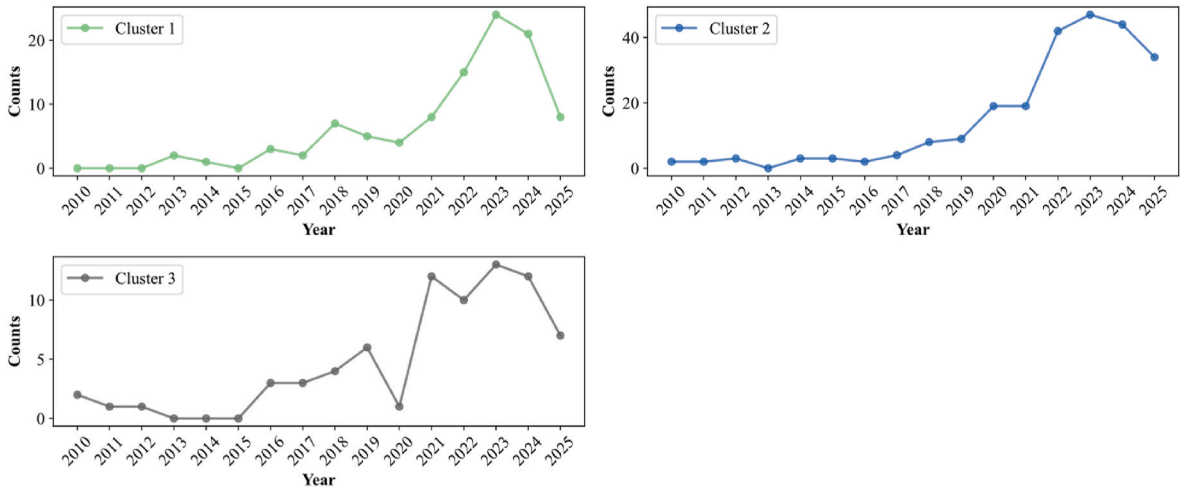


Fig. 15. Yearly trends of spatially clustered cybersecurity incidents.

maritime cyber incidents in Europe. However, the present analysis does not establish a causal relationship between the Russia–Ukraine war, GNSS interference, and the increase observed in MCAD++. Differences in incident detection, media coverage, and disclosure practices across regions may also contribute to the higher number of recorded incidents in Europe.

The annual trends in MCAD++ show regional differences in the distribution of reported maritime cyber incidents. North America and Europe account for relatively high concentrations of recorded incidents, highlighting the importance of continued monitoring and cyber resilience planning in these regions. However, these patterns may partly reflect differences in reporting practices, detection capabilities and database coverage, rather than actual differences in exposure to cyber threats. For Southeast Asia and Central Asia, further assessment of cybersecurity preparedness and response capabilities may be warranted, particularly in the context of post-pandemic recovery. These observations provide a basis for regional investigation but do not establish changes in the underlying level of maritime cyber risk.

4.5. Spatiotemporal analysis

Global and Local Moran's I were used to examine spatial autocorrelation in recorded incident severity. Spatial relationships were defined using a KNN weights matrix, with $K = 12$ specified prior to the analysis. This choice was intended to provide a sufficiently broad local neighbourhood while limiting spatial smoothing, rather than to maximise statistical significance. A sensitivity analysis was conducted for $K = 4$–16 to assess the robustness of the results (Table 5).

At $K = 12$, Global Moran's I was 0.027 ($p = 0.050$), indicating very weak positive spatial autocorrelation. The result was sensitive to the neighbourhood definition: the p-value decreased to 0.036 at $K = 16$, whereas values of $K \leq 9$ yielded $p > 0.05$. These findings suggest that global spatial autocorrelation in recorded incident severity is weak and not robust to changes in the spatial weights specification.

Fig. 16 shows Local Moran's I results before and after FDR correction.

Table 5
Sensitivity analysis of K for the spatial weight matrix.

K	Moran_I	p_value	Z-score
4	0.024	0.194	0.897
6	0.026	0.13	1.171
8	0.024	0.111	1.268
9	0.03	0.057	1.628
12	**0.027**	**0.05**	**1.714**
16	0.028	0.036	2.026

In the uncorrected panel, High–High clusters appear at some Middle Eastern ports, while Low–Low clusters appear along the west coast of North America and the Mediterranean coast. After BH-FDR correction, most of these local categories become non-significant. The High–High cluster in the Middle East, the Low–Low cluster along the west coast of North America, and the High–Low outlier in Hawaii remain significant, whereas other local categories are no longer significant. The FDR-corrected results are therefore interpreted as limited local spatial association rather than as strong evidence of widespread local clustering.

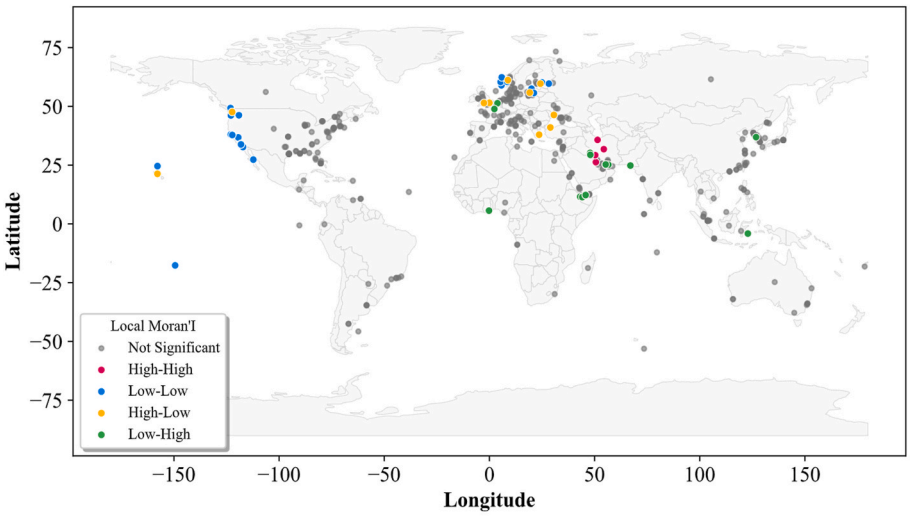
Annual estimation of Global Moran's I requires the number of incident records to exceed the neighbourhood size; therefore, years with 12 or fewer valid records were excluded. Based on the MCAD++ analysis sample, 2018 was the earliest eligible year, and the results for 2018-2025 are presented in Fig. 17. Through this period, Moran's I remained close to zero, with minor fluctuations between positive and negative values and no consistent pattern lasting more than two consecutive years. Only 2019 exhibited statistically significant positive spatial autocorrelation. These changes describe variation in the estimated spatial autocorrelation of recorded consequence severity. Without consistent statistical support across years, they should not be interpreted as evidence of persistent spatial clustering or changes in underlying cyber-risk levels.

5. Discussion and implications

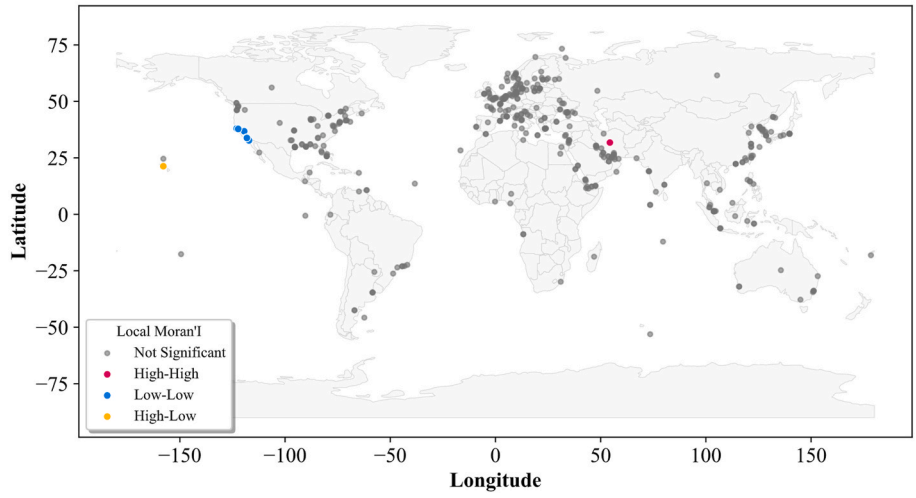
The MC-STM framework provides an integrated, data-driven approach to examining interaction, temporal, and spatial patterns in reported maritime cyber incidents. By combining association rule mining, transfer entropy, HDBSCAN clustering, and spatial autocorrelation analysis, the framework characterises recurrent RIF co-occurrences, explores nominal directional temporal associations, and examines geographical concentrations of reported incidents and spatial associations in incident severity. These complementary analyses extend incident-level description by providing a structured view of how cyber-risk factors are associated across interaction, temporal, and spatial dimensions.

5.1. Interpreting interaction and spatiotemporal patterns in maritime cyber risk

The results indicate that reported maritime cyber incidents exhibit structured relationships among technical, operational, and contextual factors. The association-rule analysis identifies recurrent combinations of threat types, affected assets, and operational consequences, indicating that these RIFs do not appear independently in the incident records. The



(a) Uncorrected Local Moran's I



(b) Local Moran's I after BH-FDR correction

Fig. 16. Spatial clustering of event consequence severity based on Local Moran's I.

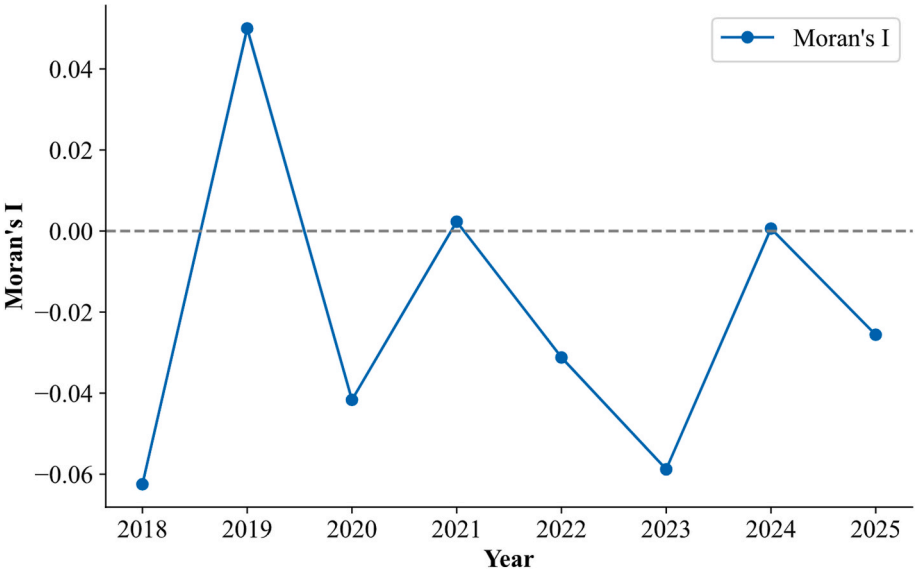


Fig. 17. Temporal analysis of the global Moran's I of event consequence severity.

transfer entropy analysis provides an exploratory view of directional informational associations among annual RIF time series. However, because no edges remained after FDR correction, these findings are presented strictly as exploratory patterns. Taken together, these analyses identify recurrent co-occurrence patterns among RIFs, while the temporal analysis provides preliminary associations that require further investigation using longer time series.

The network analysis further reveals differences in the structural roles of individual RIFs. Some factors exhibit strong outgoing or incoming association strength, whereas others occupy bridging positions within the co-occurrence or exploratory transfer entropy network. These network positions should not be interpreted as evidence that the corresponding factors mediate cyberattack processes. Rather, they indicate how particular RIFs are situated within the statistical structure of the reported incidents. The observed connectivity is consistent with the increasing integration of navigation, communication, information, and operational technologies across maritime transport systems.

The spatial analyses reveal a further dimension of heterogeneity. HDBSCAN identifies geographical concentrations of reported incidents, whereas the spatial autocorrelation analysis captures local patterns in incident severity. The distinction is important: concentrations of incident occurrence do not necessarily correspond to concentrations of high-severity incidents. The results reveal recurring statistical associations among technical, operational, and contextual factors documented in maritime cyber incident reports. These patterns should nevertheless be interpreted with caution, as regional differences in incident reporting and data availability may contribute to the observed spatial structure.

From a broader systems perspective, the interaction, temporal, and spatial results suggest that maritime cyber risk cannot be fully characterised through incident frequency or individual risk factors alone. Relationships among RIFs, exploratory temporal associations, and geographical patterns provide complementary information on the structure of reported cyber incidents. This perspective is particularly relevant to increasingly digitalised maritime transport systems, in which vessels, ports, shipping companies, logistics platforms, and supporting digital infrastructure operate within closely connected operational environments.

5.2. Implications for system-level risk governance

The findings have several implications for maritime cyber risk management. For ports and shipping companies, the recurrent associations among technical assets, cyber threats, and operational consequences suggest that cybersecurity planning should consider relationships across multiple components of maritime operations rather than addressing individual vulnerabilities in isolation. Integrating cybersecurity with business continuity, asset management, and operational resilience planning may therefore provide a more comprehensive basis for managing risks that span interconnected digital and operational systems.

For governments and regulatory authorities, the observed geographical heterogeneity highlights the importance of considering regional context when assessing maritime cyber risk. Spatial evidence derived from incident data may complement existing information used to identify areas requiring closer attention and to inform risk-based resource allocation. Similarly, the nominal directed temporal associations identified among annual RIF series may help generate hypotheses about factors that warrant closer examination.

For classification societies and insurers, the network representation provides information that is not captured by incident-frequency statistics alone. Recurrent co-occurrence patterns, nominal uncorrected directed temporal associations, and the structural positions of individual RIFs may complement existing approaches to cyber-risk assessment by providing additional evidence on how technical, operational, and contextual factors are related within reported incidents. Such information may be particularly relevant when evaluating cyber-related operational exposure across vessels, ports, and associated maritime infrastructure.

At the international level, the observed geographical variation reinforces the importance of accounting for regional context in maritime cybersecurity governance. It also highlights the value of more consistent incident reporting and information sharing across jurisdictions. Greater consistency in the collection and classification of maritime cyber incident data would strengthen the empirical basis for cross-regional comparisons and facilitate more reliable assessment of emerging cyber-risk patterns across interconnected maritime transport networks.

The findings also have implications for technology developers and service providers. Recurrent associations involving information manipulation, technical assets, and cyber threats reinforce the importance of information integrity and the security of interconnected maritime digital systems. Secure software design, vulnerability management, integrity verification, and anomaly detection therefore remain important considerations in the development and operation of critical maritime technologies. The patterns identified in this study can help identify combinations of RIFs that warrant closer technical attention, although they do not demonstrate the effectiveness of any specific cybersecurity control.

Overall, the practical value of MC-STM lies in combining complementary forms of evidence rather than treating any single analytical result as a direct indicator of cyber risk. Association rules characterise recurrent co-occurrence patterns, transfer entropy explores nominal uncorrected directional statistical associations, and spatial analyses reveal geographical concentrations and local severity structures. Their integration provides a broader empirical basis for interpreting reported maritime cyber incidents and for informing subsequent risk assessment, monitoring, and resilience planning.

5.3. Methodological limitations and data constraints

Several limitations should be considered when interpreting the results.

(1) Data coverage and spatial representation. MCAD++ is compiled primarily from publicly available reports and disclosures and therefore does not represent the full population of maritime cyber incidents. Reporting practices vary across regions, organisations, and time, reflecting differences in disclosure requirements, monitoring capabilities, media coverage, and data availability. These differences may affect the temporal and geographical patterns observed in the database. Attack attribution is particularly incomplete: the source country or region is unknown for 64.5% of the incidents, while recorded attributions are based on reported information rather than independently verified attack origins. The coordinates used in the spatial analyses therefore refer to the reported locations of affected maritime assets or operations, not to attacker origins.

The spatial analysis does not account for differences in exposure. Incident counts are not normalised by maritime traffic, port throughput, the number of connected assets, or other measures of maritime activity. Areas with more reported incidents may also have greater operational activity, digital exposure, or reporting coverage. The identified spatial concentrations should therefore be interpreted as patterns within the available incident records rather than as direct measures of geographical cyber-risk intensity.

(2) RIF classification. The transformation of unstructured incident descriptions into structured RIFs introduces classification uncertainty. Source reports vary in completeness and specificity, and ambiguous or context-dependent information may not always be assigned correctly by the LLM-based classification procedure. The resulting RIFs are consequently dependent on both the

information available in the original reports and the classification scheme used in MCAD++, and some degree of classification error is unavoidable.

(3) Statistical constraints. The analyses are subject to the assumptions and parameter settings of the individual methods. HDBSCAN depends on the selected density parameters and distance specification, although alternative parameter settings were examined to assess the stability of the main spatial patterns. Transfer entropy is more strongly constrained by data availability. Annual aggregation over 2010–2025 yields only 16 observations for each RIF, limiting the amount of temporal information available for estimating directional dependencies. Permutation testing and repeated analyses across five random seeds were used to assess the robustness of retained relationships, but they do not remove the uncertainty associated with such a short time series. Crucially, after applying BH-FDR correction to all 870 tests, no edges remained significant. The resulting transfer entropy relationships are therefore reported as exploratory directional statistical associations rather than causal effects or statistically confirmed incident propagation. Moran's I is similarly dependent on the representation of incident severity and the specification of the spatial-weight matrix. Although alternative neighbourhood sizes were examined, the Global Moran's I result remains marginal and sensitive to K. For Local Moran's I, only limited local associations remain significant after BH-FDR correction. These results identify limited local spatial association in recorded severity and do not establish risk diffusion or spillover between locations.

(4) External validation and generalisability. The sensitivity analyses and alternative model specifications provide evidence of internal robustness, but the identified patterns have not been tested using an independent incident dataset. It is therefore unclear to what extent the interaction, temporal, and spatial structures observed in MCAD++ generalise to other incident sources or reporting environments. Future validation using independent incident databases, operational and exposure data, and longer or higher-frequency time series would help determine whether these patterns persist under different data conditions.

6. Conclusions

The study developed the MC-STM framework to investigate structured patterns in reported maritime cyber incidents using the enhanced MCAD++ dataset. The enhanced dataset provides a common empirical basis for the analysis by organising heterogeneous incident information into a consistent set of RIFs. Rather than treating incidents as independent observations, the framework examines relationships among risk-influencing factors and their variation across temporal and geographical dimensions.

The analysis identifies recurring associations among RIFs and geographical differences in reported incident concentrations. Transfer entropy produces exploratory directional associations in annual RIF occurrence patterns, but none remain significant after BH FDR correction; therefore, the temporal association network is presented as an exploratory result based on uncorrected nominal associations. The network results further show that individual RIFs occupy different outgoing, incoming, and bridging positions within the association-rule network and the exploratory temporal network. These results suggest that reported maritime cyber incidents contain structured relationships that complement those derived from incident frequencies or isolated factor analysis alone. These relationships should nevertheless be interpreted as empirical patterns within the available incident records rather than as evidence of causal mechanisms or direct measures of geographical cyber-risk intensity.

MC-STM provides a common analytical basis for examining interaction, temporal, and spatial characteristics of maritime cyber incidents and may complement existing approaches to cyber-risk assessment and monitoring. Further validation using independent incident sources, higher-frequency temporal data, operational exposure measures, and information on maritime system architectures would help establish the extent to which the identified patterns generalise across different maritime environments and reporting conditions.

CRediT authorship contribution statement

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

Declaration of competing interest

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

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 864724).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oceaneng.2026.129432.

References

Abdelmagid, A.M., Javadnejad, F., Mcshane, M., Diaz, R., Pinto, C.A., 2025. A new cyber risk identification and assessment approach of the maritime cyber risks. Enterp. Inf. Syst. 19, 2524848. https://doi.org/10.1080/17517575.2025.2524848.

Afenyo, M., Caesar, L.D., 2023. Maritime cybersecurity threats: gaps and directions for future research. Ocean Coast Manag. 236, 106493.

Alcaide, J.I., Llave, R.G., 2020. Critical infrastructures cybersecurity and the maritime sector. Transp. Res. Procedia 45, 547–554.

Ali, Z., Hussain, T., Su, C.-L., Khan, I., Jurcut, A.D., Tsao, S.-H., Hu, C.-H., Elsisi, M., 2025. Deep learning-driven cyber attack detection framework in DC shipboard microgrids system for enhancing maritime transportation security. IEEE Trans. Intell. Transp. Syst. 26, 20122–20142. https://doi.org/10.1109/TITS.2025.3589188.

Amro, A., Gkioulos, V., 2023. Evaluation of a cyber risk assessment approach for cyber-physical systems: maritime- and energy-use cases. J. Mar. Sci. Eng. 11, 744. https://doi.org/10.3390/jmse11040744.

Aydin, M., Sezer, S.I., Akyuz, E., Gardoni, P., 2025. Improved Z-number based Bayesian network modelling to predict cyber-attack risk for maritime autonomous surface ship (MASS). Appl. Soft Comput. 180, 113416. https://doi.org/10.1016/j.asoc.2025.113416.

Belmoukari, B., Audy, J.-F., Forget, P., 2023. Smart port: a systematic literature review. Eur. Transp. Res. Rev. 15, 4. https://doi.org/10.1186/s12544-023-00581-6.

Bielawski, A., Lazarowska, A., 2022. Discussing cybersecurity in maritime transportation. Marit. Technol. Res 4, 252151–252151.

Bossomaier, T., Barnett, L., Harré, M., Lizier, J.T., 2016. Transfer entropy. In: An Introduction to Transfer Entropy. Springer International Publishing, Cham, pp. 65–95. https://doi.org/10.1007/978-3-319-43222-9_4.

Chen, H., Wen, Y., Huang, Y., Xiao, C., Dai, H., 2024. From integrated bridge system to marine bridge domain: a computational perspective. Ocean. Eng. 298, 117171. https://doi.org/10.1016/j.oceaneng.2024.117171.

Chen, Liguang, Ning, Xue, Ma, Yongchi, 2026. Investigating human and organizational risk propagation in ship foundering: A data-driven HFACS–Apriori–DEMATEL–FCM framework. Ocean. Eng. 369, 128832. https://doi.org/10.1016/j.oceaneng.2026.128832, 128832. https://linkinghub.elsevier.com/retrieve/pii/S0029801826046664.

Chen, R., Liang, S., Wang, J.-G., Yao, Y., Su, J.-R., Liu, L.-L., 2025. Lag-Specific transfer entropy for root cause diagnosis and delay estimation in industrial sensor networks. Sensors 25, 3980. https://doi.org/10.3390/s25133980.

Chen, X., Dou, S., Song, T., Wu, H., Sun, Y., Xian, J., 2024. Spatial-Temporal ship pollution distribution exploitation and Harbor environmental impact analysis via large-scale AIS data. J. Mar. Sci. Eng. 12, 960. https://doi.org/10.3390/jmse12060960.

Clavijo Mesa, M.V., Patino-Rodriguez, C.E., Guevara Carazas, F.J., 2024. Cybersecurity at Sea: a literature review of cyber-attack impacts and defenses in maritime supply chains. Information 15, 710.

Committee, IMO, 2017. Maritime cyber risk management in safety management systems. Resolution MSC 428 (98). https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/MSCResolutions/MSC.428%2898%29.pdf428,98.

Curti, Filippo, Gerlach, Jeffrey, Kazinnik, Sophia, Lee, Michael, Mihov, Atanas, 2023. Cyber risk definition and classification for financial risk management. Journal of Operational Risk. https://doi.org/10.21314/jop.2022.036. https://www.risk.net/journal-of-operational-risk/7956910/cyber-risk-definition-and-classification-for-financial-risk-management, 2023.

Dogan, F., 2025. Seaworthiness in the Age of Emerging Technologies (Doctoral Dissertation). University of, Reading.

Duzenli, E., Akyuz, E., Kayisoglu, G., Bolat, P., 2025. An integrated risk-based modelling approach for prediction of AIS spoofing attacks on-board ship. Ocean. Eng. 342, 122895. https://doi.org/10.1016/j.oceaneng.2025.122895.

Fan, H., Lyu, J., Chang, Z., He, X., Guo, S., 2023. Spatial patterns and characteristics of global piracy analyzed using a geographic information system. Mar. Pol. 157, 105816.

Feng, C., Xu, J., Zhang, J., Li, H., 2024. A spatiotemporal co-occurrence pattern mining algorithm based on ship trajectory data. Adv. Mech. Eng. 16, 16878132241274449. https://doi.org/10.1177/16878132241274449.

Gong, J., Li, H., Jiao, H., Yang, Z., 2025. Uncertainty-aware ship trajectory prediction via Spatio-temporal Graph Transformer. Transp. Res. Part E Logist. Transp. Rev. 203, 104315. https://doi.org/10.1016/j.tre.2025.104315.

Guo, X., Zheng, Q. qian, Guo, Y., 2025. Maritime accident causation: a spatiotemporal and HFACS-Based approach. Ocean. Eng. 340, 122329. https://doi.org/10.1016/j.oceaneng.2025.122329.

Huang, L., Wen, Y., Geng, X., Zhou, C., Xiao, C., Zhang, F., 2017. Estimation and spatio-temporal analysis of ship exhaust emission in a port area. Ocean. Eng. 140, 401–411. https://doi.org/10.1016/j.oceaneng.2017.06.015.

Jayathilaka, R., Weerakoon, W., Indika, K.W., Arulananthan, K., Kithsiri, H.M.P., 2022. Spatio-temporal variation of plastic pellets dispersion in the coastline of Sri Lanka: an assessment of pellets originated from the X-Press pearl incident during the Southwest monsoon in 2021. Mar. Pollut. Bull. 184, 114145. https://doi.org/10.1016/j.marpolbul.2022.114145.

Jessen, H., 2026. AI-induced modern maritime risks: ensuring “cyber seaworthiness” and legally-compliant data management in international shipping. WMU J. Marit. Aff. 25, 525–546. https://doi.org/10.1007/s13437-026-00426-8.

Kanwal, K., Shi, W., Kontovas, C., Yang, Z., Chang, C.-H., 2024. Maritime cybersecurity: are onboard systems ready? Marit. Policy Manag. 51, 484–502. https://doi.org/10.1080/03088839.2022.2124464.

Kara Balci, Ö., Varan Samut, İ., Ademuni-Odeke, 2024. An evaluation of cyber-worthiness and related factors in maritime safety and maritime security. J. Int. Marit. Saf. Environ. Aff. Shipp. 8, 2411180. https://doi.org/10.1080/25725084.2024.2411180.

Khan, R.U., Yin, J., Afzaal, M., Yang, Z., 2026. Quantitative assessment of language and cultural diversity as human reliability factors in maritime safety. Reliab. Eng. Syst. Saf. 273, 112358. https://doi.org/10.1016/j.ress.2026.112358.

Kumar, P., Gupta, G.P., Tripathi, R., Garg, S., Hassan, M.M., 2023. DLTIF: deep learning-driven cyber threat intelligence modeling and identification framework in IoT-Enabled maritime transportation systems. Ieee Trans. Intell. Transp. Syst. 24, 2472–2481. https://doi.org/10.1109/TITS.2021.3122368.

Li, H., Çelik, C., Bashir, M., Zou, L., Yang, Z., 2024. Incorporation of a global perspective into data-driven analysis of maritime collision accident risk. Reliab. Eng. Syst. Saf. 249, 110187. https://doi.org/10.1016/j.ress.2024.110187.

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

Li, H., Yang, Z., 2023. Towards safe navigation environment: the imminent role of spatio-temporal pattern mining in maritime piracy incidents analysis. Reliab. Eng. Syst. Saf. 238, 109422. https://doi.org/10.1016/j.ress.2023.109422.

Li, K.X., Yin, J., Bang, H.S., Yang, Z., Wang, J., 2014. Bayesian network with quantitative input for maritime risk analysis. Transp. Transp. Sci. 10, 89–118. https://doi.org/10.1080/18128602.2012.675527.

Li, M., Zhou, J., Chattopadhyay, S., Goh, M., 2024. Maritime cybersecurity: a comprehensive review. https://doi.org/10.48550/arXiv.2409.11417.

Liang, M., Li, H., Liu, R.W., Lam, J.S.L., Yang, Z., 2024. PiracyAnalyzer: spatial temporal patterns analysis of global piracy incidents. Reliab. Eng. Syst. Saf. 243, 109877.

Ma, J., Zhou, Y., Chang, Y., Zhu, Z., Liu, G., Chen, Z., 2024. TG-PGAT: an AIS data-driven dynamic spatiotemporal prediction model for ship traffic flow in the port. J. Mar. Sci. Eng. 12, 1875. https://doi.org/10.3390/jmse12101875.

McInnes, L., Healy, J., Astels, S., 2017. Hdbscan: hierarchical density based clustering. J. Open Source Softw. 2, 205.

Meland, Per HÅkon, Bernsmed, Karin, Wille, Egil, Rødseth, Ørnulf Jan, Nesheim, Dag Atle, 2021. A Retrospective Analysis of Maritime Cyber Security Incidents. TransNav, the Int. J. Mar. Navigation Saf. Sea Transport. 15 (3), 519–530. https://doi.org/10.12716/1001.15.03.04. http://www.transnav.eu/Article_A_Retrospective_Analysis_of_Maritime_Cyber_Security_Incidents_Meland,59,1144.html.

Miętkiewicz, R., 2025. Hybrid threats in the baltic sea. The Results of Analysis of Countermeasure Options. Terror. Stud. Anal. Prewencja, pp. 35–70.

Miller, T., Durlik, I., Kostecka, E., Sokolowska, S., Kozlovska, P., Zwolak, R., 2025. Artificial intelligence in maritime cybersecurity: a systematic review of AI-Driven threat detection and risk mitigation strategies. Electronics 14, 1844. https://doi.org/10.3390/electronics14091844.

Mohsendokht, M., Li, H., Kontovas, C., Chang, C.-H., Qu, Z., Yang, Z., 2024. Decoding dependencies among the risk factors influencing maritime cybersecurity: lessons learned from historical incidents in the past two decades. Ocean. Eng. 312, 119078. https://doi.org/10.1016/j.oceaneng.2024.119078.

Neumann, T., 2026. Cybersecurity risk detection in maritime transport using AI-based behavioural modelling. In: Innovations in Maritime Technology and Engineering. CRC Press.

Park, C., Kontovas, C., Yang, Z., Chang, C.-H., 2023. A BN driven FMEA approach to assess maritime cybersecurity risks. Ocean Coast Manag. 235, 106480.

Peng, P., Xie, X., Claramunt, C., Lu, F., Gong, F., Yan, R., 2025. Bibliometric analysis of maritime cybersecurity: research status, focus, and perspectives. Transp. Res. Part E Logist. Transp. Rev. 195, 103971. https://doi.org/10.1016/j.tre.2025.103971.

Pijpker, J., McCombie, S., Johnson, S., Loves, R., Makrakis, G.M., 2024. An Open-Source Database of Cyberattacks on the Maritime Transportation System.

Rodriguez-Diaz, E., Santamaria, M., Prieto, J.M., Alcaide, J.I., 2026. Emerging Maritime threats in global geopolitics: GPS spoofing as a tool of hybrid warfare in strategic Maritime chokepoints. https://doi.org/10.21203/rs.3.rs-8901159/v1.

Schinas, O., Metzger, D., 2023. Cyber-seaworthiness: a critical review of the literature. Mar. Pol. 151, 105592. https://doi.org/10.1016/j.marpol.2023.105592.

Senarak, C., 2025. Toward sustainability and digital resilience: a circular economy cybersecurity framework for seaports. Clean. Logist. Supply Chain 15, 100220. https://doi.org/10.1016/j.clscn.2025.100220.

Shi, Shaoyue, Ma, Bohao, Yan, Ran, Zhang, Mingyang, 2026. Review of intelligent maritime transportation systems facilitated by deep learning: A survey on safe navigation. Commun. Transport. Res. 6 (2), 9640006. https://doi.org/10.26599/commtr.2026.9640006. https://www.sciopen.com/article/10.26599/COMMTR.2026.9640006.

Shu, Y., Dong, A., Liu, C., Gan, L., Song, L., 2026. Anomaly detection of ship behavior based on deep neural networks. Reliab. Eng. Syst. Saf. 266, 111801. https://doi.org/10.1016/j.ress.2025.111801.

Soner, O., 2025. Modeling and analyzing cybersecurity risk propagation in ports using fuzzy cognitive maps: system sensitivity to key threat factors. Ocean Coast Manag. 270, 107857. https://doi.org/10.1016/j.ocecoaman.2025.107857.

Soner, O., Kayisoglu, G., Bolat, P., Tam, K., 2024. Risk sensitivity analysis of AIS cyber security through maritime cyber regulatory frameworks. Appl. Ocean Res. 142, 103855. https://doi.org/10.1016/j.apor.2023.103855.

Sun, J., Wang, H., Wang, M., 2023. Risk assessment of bauxite maritime logistics based on improved FMECA and fuzzy Bayesian network. J. Mar. Sci. Eng. 11, 755.

Tiefelsdorf, M., Boots, B., 1997. A note on the extremities of local moran's I_i s and their impact on global moran's I. Geogr. Anal. 29, 248–257. https://doi.org/10.1111/j.1538-4632.1997.tb00960.x.

Timothy, A., Obiniyi, A., Agbogun, J., 2026. A systematic literature review of cybersecurity risks in autonomous and digitalised maritime systems. Afr. J. Manag. Bus. Res. 22, 226–240. https://doi.org/10.62154/ajmbr.2026.022.01021.

Tombak, M.-L., Zetterman, B.-E., Tapaninen, U., 2025. Cybersecurity risks at port. Transp. Telecommun. J. 26, 276–291. https://doi.org/10.2478/ttj-2025-0021.

Tsioufis, M., Fytopoulos, A., Kalaitzi, D., Alexopoulos, T.A., 2024. Discovering maritime-piracy hotspots: a study based on AHP and spatio-temporal analysis. Ann. Oper. Res. 335, 861–883. https://doi.org/10.1007/s10479-023-05352-z.

Uflaz, E., Sezer, S.I., Tuncel, A.L., Aydin, M., Akyuz, E., Arslan, O., 2024. Quantifying potential cyber-attack risks in maritime transportation under Dempster-Shafer theory FMECA and rule-based Bayesian network modelling. Reliab. Eng. Syst. Saf. 243, 109825. https://doi.org/10.1016/j.ress.2023.109825.

Wang, J., Zhou, H., Zhao, B., Wang, L., Zhuang, L., 2026. Spatiotemporal patterns of global cargo ship accidents: an enhanced association rule mining approach. Reliab. Eng. Syst. Saf. 268, 112045. https://doi.org/10.1016/j.ress.2025.112045.

Wang, M., Wang, Y., Ding, J., Yu, W., 2024. Interaction aware and multi-modal distribution for ship trajectory prediction with spatio-temporal crisscross hybrid network. Reliab. Eng. Syst. Saf. 252, 110463. https://doi.org/10.1016/j.ress.2024.110463.

Wang, S., Li, Y., Xing, H., 2023. A novel method for ship trajectory prediction in complex scenarios based on spatio-temporal features extraction of AIS data. Ocean. Eng. 281, 114846. https://doi.org/10.1016/j.oceaneng.2023.114846.

Westbrook, T., 2023. Radiofrequency interference strategies targeting marine navigation systems: political motives and consequences. J. Balt. Secur. 9, 69–97. https://doi.org/10.57767/jobs_2023_003.

Xie, J., Yang, J., Li, J., He, M., Song, H., 2025. Three-way concept lattice construction and association rule acquisition. Inf. Sci. 701, 121867. https://doi.org/10.1016/j.ins.2024.121867.

Yang, Y., Liu, Yang, Li, G., Zhang, Z., Liu, Yanbin, 2024. Harnessing the power of Machine learning for AIS Data-driven maritime research: a comprehensive review. Transp. Res. Part E Logist. Transp. Rev. 183, 103426.

You, L., Xiao, S., Peng, Q., Claramunt, C., Han, X., Guan, Z., Zhang, J., 2020. St-seq2seq: a spatio-temporal feature-optimized seq2seq model for short-term vessel trajectory prediction. IEEE Access 8, 218565–218574.

Yu, H., Meng, Q., Fang, Z., Liu, J., 2023. Literature review on maritime cybersecurity: state-of-the-art. J. Navig. 76, 453–466. https://doi.org/10.1017/S0373463323000164.

Yuan, Y., Guo, H., Fan, Z., Peng, Y., Zhang, J., Song, X., Shibasaki, R., 2025. A decomposable multi-fusion spatio-temporal marine traffic flow forecasting

algorithm: taking the North Sea and Baltic Sea region as an example. Transp. Res. Part E-Logist. Transp. Rev. 203, 104340. https://doi.org/10.1016/j.tre.2025.104340.

Zhang, X., Liu, J., Chen, K., Gong, P., Liu, Y., Wu, Z., 2024. Learning dynamic interactions and long-term patterns with spatio-temporal graphs for multi-vessel trajectory prediction. IEEE Trans. Intell. Veh.

Zhang, X., Liu, J., Gong, P., Chen, C., Han, B., Wu, Z., 2023. Trajectory prediction of seagoing ships in dynamic traffic scenes via a gated spatio-temporal graph aggregation network. Ocean. Eng. 287, 115886. https://doi.org/10.1016/j.oceaneng.2023.115886.

Zhao, Y., Gong, J., Li, H., Nguyen, T.T., Matthews, C., Yang, Z., 2026a. Cybersecurity in smart port systems: a systematic review and data-driven research agenda. Transp. Policy 187, 104268. https://doi.org/10.1016/j.tranpol.2026.104268.

Zhao, Y., Li, H., Nguyen, T.T., Matthews, C., Yang, Z., 2026b. Integrating prompt-based zero-shot learning into bayesian network for maritime cybersecurity risk analysis. Transp. Res. Part E Logist. Transp. Rev. 216, 105134. https://doi.org/10.1016/j.tre.2026.105134.

Zhao, Y., Li, H., Thanh Nguyen, T., Matthews, C., Yang, Z., 2026c. An automated bayesian network framework for maritime cybersecurity risk assessment powered by large language models. Saf. Sci. 204, 107384. https://doi.org/10.1016/j.ssci.2026.107384.

Zhao, Y., Li, H., Yang, Z., 2025. A new AI solution to maritime cybersecurity risk prediction. In: 35th European Safety and Reliability Conference (ESREL 2025) and the 33rd Society for Risk Analysis Europe Conference (SRA-E 2025). Presented at the 35th European Safety and Reliability Conference (ESREL 2025) and the 33rd Society for Risk Analysis Europe Conference (SRA-E 2025). Research Publishing Services, pp. 2281–2288. https://doi.org/10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P0401-cd. Singapore EXPO, Singapore.