VessiGuard: AI-Driven Anomaly Detection for Maritime Cyber Defence

Ravi Varma Kanumuri, Laavanya Rachakonda, Hosam Alamleh, Bilge Karabacak and Ulku Clark Yaylacicegi
University of North Carolina, Wilmington, USA

rk1260@uncw.edu
rachakondal@uncw.edu
alamlehh@uncw.edu
karabacakb@uncw.edu
clarku@uncw.edu

Abstract: The maritime industry is undergoing a rapid digital transformation, driven by the adoption of technologies such as the Automatic Identification System (AIS), advanced navigation software, and onboard Internet of Things (IoT) sensors. These innovations have significantly improved operational efficiency, safety, and situational awareness. However, this increasing reliance on interconnected digital systems also expands the sector's exposure to cyber threats. Traditional rule-based monitoring and siloed intrusion detection systems often fail to identify coordinated multi-modal attacks, leaving vessels vulnerable to sophisticated, stealthy intrusions. The dual nature of this transformation underscores the urgent need for more sophisticated and adaptive cybersecurity strategies. This study introduces VessiGuard, an AI-driven anomaly detection system designed to detect and mitigate abnormal vessel behaviour as an early indicator of potential cyber intrusions, system failures, or operational anomalies. The approach leverages two complementary artificial intelligence techniques: Long Short-Term Memory (LSTM) neural networks, which model temporal dependencies in vessel movement, and Isolation Forest algorithms, which excel at detecting rare and unusual behaviour patterns. By fusing navigational telemetry with operational technology (OT) sensor readings, specifically engine temperature and fuel consumption, the model creates a unified, cross-domain anomaly score that is robust against single-variable manipulation. A prototype anomaly detection system was implemented and evaluated using controlled simulation and publicly available maritime datasets that reflect real-world operational scenarios. Results demonstrate that VessiGuard effectively detects anomalies, including GPS spoofing, sensor drift, and structured interference. Experimental validation indicates a detection accuracy of approximately 94.2% for trajectory anomalies and 92.8% for sensor deviations. Furthermore, the system demonstrates modality-specific responsiveness, identifying operational sensor faults in under four minutes while accurately accumulating evidence for trajectory deviations within five to eight minutes. This work presents a practical pathway towards adaptive, data-driven cybersecurity solutions by situating anomaly detection within the broader maritime operational ecosystem. The findings highlight how AI-based anomaly detection can complement existing maritime defence mechanisms, support decision-making under dynamic threat conditions, and improve incident response readiness. Furthermore, the results lay the groundwork for future research into autonomous and semi-autonomous detection architectures, ultimately contributing to a more resilient, secure, and intelligence-driven digital maritime domain.

Keywords: Maritime cybersecurity, Anomaly detection, Machine learning, Cyber-Physical systems, Cyber resilience, AIS security

1. Introduction

The maritime industry underpins global trade, with over 90% of goods transported by sea. As shipping operations evolve, vessels increasingly rely on digital technologies, such as the Automatic Identification System (AIS), GPS navigation, onboard Internet of Things (IoT) sensors, and advanced communication networks, to enhance operational efficiency, safety, and situational awareness (Kim and Joe, 2024). However, this digital transformation has also introduced complex cybersecurity risks. Modern ships now function as floating cyber-physical systems, and their growing connectivity exposes them to threats such as GPS spoofing, data manipulation, unauthorised access, and coordinated cyberattacks on operational technology (OT) systems. These vulnerabilities can disrupt navigation, endanger crew safety, and jeopardise global supply chains (Bobrovnikova et al., 2022; School, 2022).

Detecting such threats remains challenging. Unlike mechanical failures, cyberattacks are often stealthy, adaptive, and embedded within legitimate data streams. Conventional rule-based systems struggle to detect these subtle anomalies in real time, often generating excessive false alarms or missing sophisticated attacks. As maritime systems become more interconnected (IMO, 2017a; IMO, 2017b; BIMCO et al., (2020), there is an urgent need for intelligent, automated, and adaptive cybersecurity solutions capable of context-aware, early detection of abnormal behaviours (Maganaris et al., 2024; Shit et al., 2025).

This paper introduces VessiGuard, an AI-driven anomaly detection system designed to enhance maritime resilience through intelligent anomaly detection. It employs advanced artificial intelligence (AI) techniques, including Long Short-Term Memory (LSTM) neural networks and Isolation Forest algorithms, to analyse AIS, GPS, and sensor data for deviations from expected operational patterns. These models detect subtle temporal anomalies and rare behavioural outliers as indicators of cyber intrusions, system malfunctions, or unauthorised activities.

This work focuses on the implementation and evaluation of its core anomaly detection engine. A prototype system, validated using simulated maritime datasets, demonstrates effective detection of anomalies such as GPS spoofing, sensor manipulation, and unauthorised access attempts.

By combining state-of-the-art AI with a modular system design, VessiGuard represents a significant step towards autonomous, adaptive, and data-driven maritime cybersecurity. Beyond real-time anomaly detection, it lays the groundwork for future integration of resilient response mechanisms, contributing to safer and more secure maritime operations in an increasingly digital era.

2. Related Work

The maritime industry's increasing digitalisation has prompted significant research into cybersecurity, particularly in anomaly detection and secure data handling. Existing studies have explored diverse approaches ranging from trajectory-based monitoring and sensor analysis to intrusion detection systems. This section reviews these developments, highlighting their strengths and limitations, and identifies the gap the proposed VessiGuard system aims to address.

AIS-based Vessel Behaviour Analysis: AIS data forms the backbone of most maritime anomaly detection systems. Traditional approaches rely on rule-based mechanisms such as geofencing and speed thresholds to flag deviations from expected vessel behaviour. While these methods are straightforward, they often suffer from high false-positive rates and limited adaptability (Riveiro et al, 2018). Machine learning methods have enhanced AIS-based monitoring by learning standard movement patterns, using models like Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks to capture temporal dependencies. Despite their effectiveness, these models are often data-hungry and require high-quality labelled datasets, which are scarce in the maritime domain (Davari et al, 2021; Ribeiro et al, 2023). Additionally, most AIS-focused studies operate exclusively on navigational trajectories and do not integrate physical sensor states, leaving them vulnerable to multi-modal cyber-physical spoofing.

*Sensor and Engine Anomaly Detection***:** Beyond navigation data, several studies have examined anomaly detection using sensor and machinery data. Deep learning models such as convolutional neural networks (CNNs) and autoencoders have been used to monitor operational parameters, while Transformer-based models identify complex, non-linear patterns and deviations from normal operating conditions (Thakur et al, 2025). Such methods effectively detect cyber-induced manipulation or faults but face deployment challenges due to sensor variability across vessels and the need for domain-specific feature engineering (Amro et al, 2022). Moreover, many approaches focus solely on detecting anomalies without integrating incident response mechanisms or cross-referencing them with navigation data.

Intrusion Detection Systems in Maritime Networks: The convergence of information technology (IT) and operational technology (OT) aboard vessels introduces new network-based vulnerabilities. Intrusion detection systems (IDS) have been developed to identify malicious traffic, unauthorised access, or abnormal communication patterns (Wang et al, 2022; Brenner et al, 2023). Unsupervised learning techniques such as clustering and one-class classification have been applied to network traffic data to detect outliers without requiring extensive labelled datasets (Talukder et al, 2024; Nguyen et al, 2020). However, most IDS approaches are limited in their ability to correlate network anomalies with physical system states, offering little insight into the physical impact of cyber events.

2.1 Summary and Research Gap

Although substantial progress has been made in maritime cybersecurity, most approaches target isolated aspects, such as AIS-based anomaly detection, sensor monitoring, or network intrusion detection, rather than providing a unified defence solution. These siloed approaches result in fragmented situational awareness, which is insufficient for detecting coordinated or multi-modal cyberattacks. Furthermore, few frameworks combine real-time detection with secure logging and automated incident response; a brief comparison is mentioned in Table 1.

Table 1: Brief comparison with the existing State-of-the-Art Literature

Research Area	Data Type	Methods Used	Strengths	Limitations	Gap Addressed by VessiGuard
Rule-based AIS anomaly detection Riveiro et al. (2018)	AIS trajectories	Geofencing, speed rules	Simple, interpretable	High false positives; no temporal modeling	ML-based sequence learning
Deep learning AIS models (RNN/LSTM) Davari and Aguiar (2021)	AIS/GPS	Sequential neural networks	Captures temporal dependencies	Requires high-quality labeled data; ignores sensors	fuses navigation + sensor behaviors
Sensor anomaly detection Nguyen et al. (2020)	Engine, fuel, compass	Autoencoders, CNNs, Transformers	Detects subtle OT faults	Vessel-specific; noisy logs; no link to navigation	unites OT + navigation streams
Maritime IDS Amro et al. (2022)	Network packets	One-class, clustering, MTL	Detects malicious traffic	No physical context; blind to sensor/trajectory attacks	cyber-physical anomaly detection
AIS deep-generative models Wang et al. (2022)	AIS	Probabilistic trajectory modeling	Strong spatial-temporal modeling	Cannot detect OT manipulation	integrates OT and GPS/compass features

VessiGuard addresses these gaps by integrating AI-based anomaly detection, cross-domain data fusion, and laying the foundation for integrity mechanisms into a cohesive, modular architecture. This approach enhances detection accuracy, ensures secure event traceability is possible, and lays the foundation for adaptive, automated maritime cybersecurity systems.

3. Proposed Framework

3.1 System Overview and Design Rationale

VessiGuard is designed as a modular anomaly detection system that fuses navigation, telemetry, and operational sensor data using machine-learning techniques. Rather than serving as a full cybersecurity framework, VessiGuard focuses specifically on detecting anomalous patterns that may indicate spoofing, sensor manipulation, or cyber-physical irregularities. The system employs a hybrid architecture that combines sequential and statistical learning to enhance robustness across various types of anomalies.

The architecture is built around four core layers: data acquisition, anomaly detection, score fusion, and a conceptual human-centered interface. The current implementation primarily focuses on the data acquisition and anomaly detection layers. The design rationale rests on two core principles:

- Cross-modal Integration: A unified system that analyzes AIS/GPS signals alongside OT sensor data improves detection accuracy against cyber incidents that frequently affect multiple data sources simultaneously.
- Complementary ML Specialisation: LSTM autoencoders specialise in capturing temporal vessel behaviour, while Isolation Forest models detect abrupt deviations in sensor data. A fused approach mitigates weaknesses inherent in using either model independently.

3.2 Data Acquisition Layer

The foundation of VessiGuard is a data ingestion pipeline that aggregates multi-modal maritime telemetry from diverse onboard sources. This includes AIS broadcasts, GPS navigation data, compass readings, engine and fuel sensor metrics, and other operational parameters essential to vessel performance. The system operates with simulated datasets and real-time feeds, ensuring compatibility with maritime communication standards, such as NMEA 0183 and NMEA 2000 (Kim, H., & Joe, I., 2024). This layer performs data cleaning, normalisation, resampling, and temporal alignment to ensure consistency before feeding into the ML models. The

preprocessing step is essential for reducing false alarms and improving the reliability of anomaly detection by ensuring that the downstream ML models receive stable, synchronised inputs.

3.3 AI-Driven Anomaly Detection Engine

The core of VessiGuard's current implementation is the Anomaly Detection Engine, which comprises two distinct, complementary models:

1. Long Short-Term Memory (LSTM) Autoencoder: This model learns temporal dependencies in vessel behaviour, capturing patterns in continuous signals like vessel movement or sensor readings. It reconstructs expected trajectories; deviations in the reconstruction error indicate sequence-level anomalies (Hochreiter and Schmidhuber, 1997). This model detects temporal drifts or trajectory inconsistencies.
2. Isolation Forest (IF) Algorithm: This lightweight, unsupervised model excels at detecting rare, non-sequential deviations that differ statistically from the majority of the data (Liu, Ting and Zhou, 2008). IF is particularly effective at identifying subtle irregularities in sensor or telemetry data that may signal unauthorised access or latent system faults. It detects statistically rare events in static or low-dimensional feature spaces.

Together, these models form a hybrid detection pipeline. The LSTM autoencoder learns vessel motion patterns over time, while the Isolation Forest detects instantaneous deviations in sensor readings, providing complementary strengths. This dual-model strategy ensures resilience against both gradual and abrupt irregularities in maritime datasets.

4. Implementation, Methodology, and Experimental Setup

4.1 Research Question and Objectives

This study investigates the following research question: How can machine learning algorithms, specifically LSTM and Isolation Forest models applied to AIS, GPS, and OT data, enhance real-time anomaly detection in maritime cybersecurity? To address this, the primary objective is to develop and evaluate a prototype system capable of identifying operational and navigational anomalies, indicating potential cyber intrusions or system failures, and providing actionable intelligence to vessel operators before incidents escalate.

4.2 Scope of the Study

Four core areas define the scope of this work:

- Prototype Development: Designing a functional system to monitor AIS, GPS, and simulated OT telemetry.
- Model Application and Real-Time Detection: Applying LSTM for sequential behaviour modelling and Isolation Forest for multi-dimensional outlier detection.
- Simulation-Based Evaluation: Validating detection accuracy, false alarm rates, and responsiveness using simulated datasets.
- Integration Considerations: Ensuring compatibility with existing vessel monitoring workflows without significant infrastructure changes.

4.3 Project Motivation and Contribution

VessiGuard addresses the gap in deployable maritime security solutions by presenting a practical, lightweight anomaly detection system. Its contributions are threefold: first, it introduces a novel **hybrid ML approach** that combines LSTM-based sequence reconstruction with Isolation Forest outlier detection for multimodal vessel data. Second, the system features a **real-time, edge-ready design** suitable for deployment on vessel-based computing hardware. Finally, it establishes an **operational integration pathway**, demonstrating how advanced anomaly detection can be embedded within broader cybersecurity architectures.

4.4 Dataset, Preprocessing, and Feature Engineering

The dataset comprises Automatic Identification System (AIS) and Global Positioning System (GPS) telemetry (latitude, longitude, Speed Over Ground (SOG), Course Over Ground (COG)) alongside simulated Operational Technology (OT) data, including engine temperature and fuel rate (Varma, 2024). Data were resampled to a 1-minute cadence and filtered to remove invalid MMSI, zero coordinates, or implausible speeds (> 50 kn). Preprocessing involved deduplication, chronological sorting, and interpolating short gaps (≤ 5 min) along great-

circle paths, while longer gaps initiated new segments. GPS coordinates (WGS84) were converted to a local east–north–up (ENU) frame to **stabilise** kinematic computations. Feature engineering involved computing kinematic metrics (displacements Δx/Δy, acceleration, jerk, turn rate, and cross-track error) and geofencing parameters (distance to coastline/centreline and harbour proximity) over a sliding window of T. Behavioural indicators included stop/start flags, loitering scores, and sudden heading flips. OT features comprised z-scored engine temperature and fuel rate residuals against a kinematic load model, alongside window aggregates (means, variances, and stop fractions).

4.5 Methodology

The VessiGuard prototype models vessel behaviour using sliding windows of telemetry data (T x F) derived from AIS, GPS, and simulated OT sensors. The system computes an anomaly score s(X) ∈ [0,1], where higher values indicate greater deviation, targeting real-time inference on edge hardware. The framework integrates two complementary models. First, a **Sequence Model** employing a two-layer LSTM autoencoder (hidden size 128, dropout 0.2) reconstructs sequential inputs, using reconstruction error as the anomaly signal. To support transferability across fleets with limited data, bottleneck adapters are inserted between layers, allowing fine-tuning of only ~5% of parameters. The loss function combines feature-weighted mean squared error with a temporal smoothness penalty:

$$\mathcal{L} = \sum_{t=1}^{T} w \, \|\hat{x}_t - x_t\|_2^2 \; + \; \lambda \sum_{t=2}^{T} \|(\hat{x}_t - \hat{x}_{t-1}) - (x_t - x_{t-1})\|_2^2$$

Window-level errors are **normalised** to produce the score s_{LSTM}, with per-timestep scores providing interpretability.

In parallel, a **Tabular Outlier Model** using an Isolation Forest (n_estimators=300, contamination=0.01–0.10) processes aggregated features to detect abrupt sensor deviations, such as fuel spikes or temperature drifts. Its outputs are scaled to s_{IF} ∈ [0,1]. Finally, these outputs are fused to provide a robust anomaly score.

$$s = \alpha * s\_LSTM + (1 - \alpha) * s\_IF, \alpha \in [0.3,0.7]$$

Alerts are triggered when s≥ τ, with the threshold τ selected to **maximise** the F1 score or maintain a target false alarm rate.

4.6 Simulation and Validation Setup

Validation was conducted entirely in a controlled simulation environment using injected anomalies such as:

- *GPS Spoofing*: Artificial trajectory deviations from the vessel trajectory were injected to simulate navigation signal tampering (Varma, 2024).
- *Sensor Drift*: Gradual manipulation of OT readings through artificial spikes and drops
- *Compass Perturbations*: Sudden heading flips to mimic compass spoofing
- *Structured Noise*: Periodic false signals to evaluate how well the models handled structured noise.

Cross-validation was performed on time-split datasets to avoid leakage of future information into training. In addition, region-holdout experiments were conducted, where one port's data was excluded from training to test generalisation. The Isolation Forest and LSTM-AE models were validated independently and in fused mode to assess improvements from score combination.

5. Results and Discussion

5.1 Simulation Scenarios and Evaluation Metrics

Validation was conducted in a controlled simulation environment using injected anomalies to test system resilience: **GPS Spoofing** (trajectory deviations), **Sensor Drift** (gradual OT manipulation), **Compass Perturbations** (sudden heading flips), and **Structured Noise**. Performance was evaluated using **Detection Accuracy** (proportion of true anomalies detected), **False Alarms per Hour (FA/h)** (reliability metric), and **Precision–Recall AUC** (PR-AUC) for robustness on imbalanced data.

5.2 Results and Analysis

GPS Trajectory Anomalies (LSTM Autoencoder): The LSTM autoencoder effectively captured temporal dependencies in vessel trajectories, as shown in Figure 1. The simulated vessel trajectory, together with the GPS anomalies identified by the LSTM autoencoder, is visualised. The underlying track (shown in light blue)

represents the vessel's nominal movement pattern, while the red markers highlight locations where the model flagged deviations from expected behaviour. These anomalies occur in spatial clusters, indicating periods where the vessel's motion diverged from its learned trajectory dynamics, consistent with spoofing, drift, or injected positional perturbations. Because the LSTM autoencoder reconstructs the expected temporal evolution of latitude–longitude sequences, abrupt or sustained inconsistencies result in elevated reconstruction errors that allow the model to isolate abnormal segments of the track.

The model achieved a detection accuracy exceeding 94.2% with a PR-AUC of 0.91, typically identifying anomalous behaviour within 5-8 minutes of its onset. This delay reflects the need for gradual deviation patterns to accumulate sufficient temporal discrepancy before being distinguished from legitimate maneuvering. The visualization clearly demonstrates how the system localizes and clusters GPS anomalies along the vessel's route, supporting real-time situational awareness for maritime monitoring.

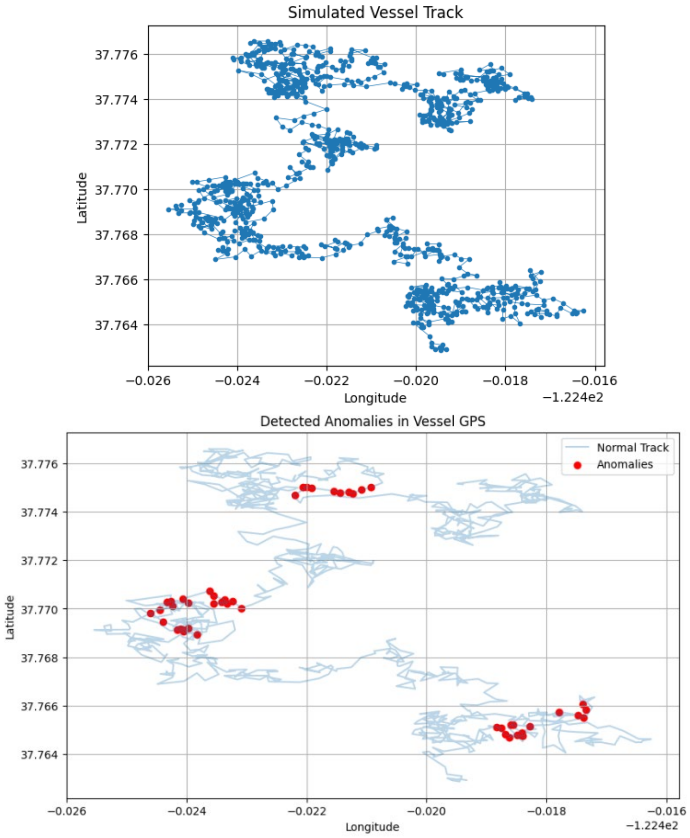

Figure 1: Detected GPS Trajectory Anomalies

Fuel level and engine temperature results, as shown in Figure 2, illustrate the behaviour of both fuel-level measurements and engine-temperature readings over time, with anomalous points highlighted in red. The fuel-level plot shows several abrupt drops to zero, indicating sensor tampering, data corruption, or sudden, unrealistic depletion events. Similarly, the engine-temperature plot reveals sharp downward spikes to zero and occasional extreme upward bursts, all of which deviate significantly from normal engine operating ranges.

The Isolation Forest model successfully identified each of these deviations as outliers, demonstrating its effectiveness in capturing both instantaneous spikes and sustained abnormal values in OT sensor streams. Detection accuracy reached 92.8%, with a very low false-alarm rate (FA/h < 0.12). Furthermore, the model exhibited a short detection latency (typically under 4 minutes) due to the sharp and non-temporal nature of these anomalies, which the Isolation Forest is well-suited to identify quickly. These results confirm the model's ability to reliably detect fuel and engine abnormal behaviours, as summarised in Table 2.

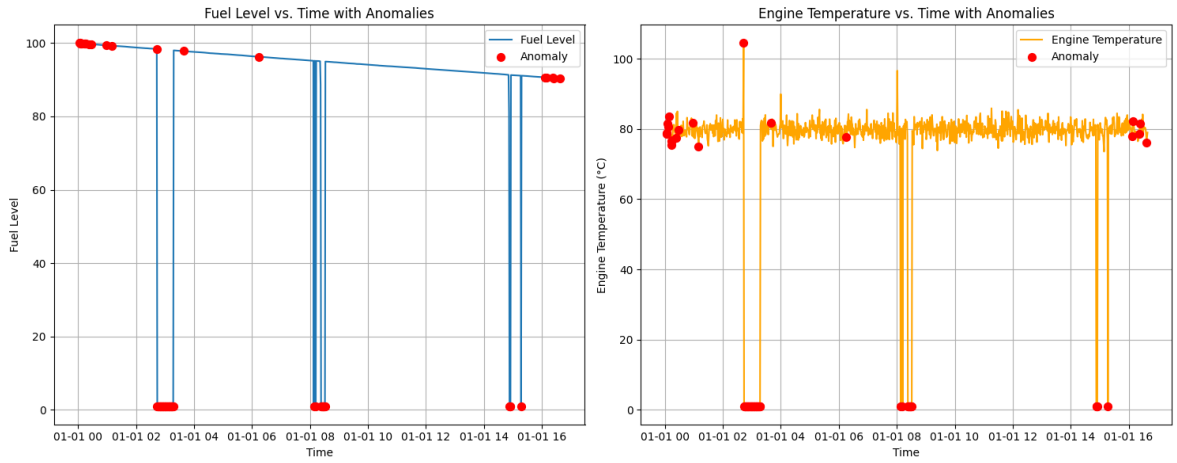

Figure 2: Detected Fuel and Engine Sensor Anomalies

The data, when visualised as a time-series plot, displays the temporal distribution of detected anomalies grouped into four distinct clusters based on their occurrence times, as shown in Figure 3. Each cluster represents a period in which multiple anomalies occurred in close succession, separated from other clusters by longer inactive intervals. Cluster 0 corresponds to the earliest anomaly window, while Cluster 1 captures a sequence of mid-range anomalies. Cluster 2 highlights a later group of anomalies, and Cluster 3 captures the final anomalies.

The clustering pattern illustrates how anomalies tend to emerge in grouped time segments rather than as isolated events, indicating periods of concentrated unusual system behaviour. This analysis supports the system's ability to not only detect individual anomalous readings but also characterise broader anomaly episodes, contributing to more contextual and situationally aware maritime monitoring.

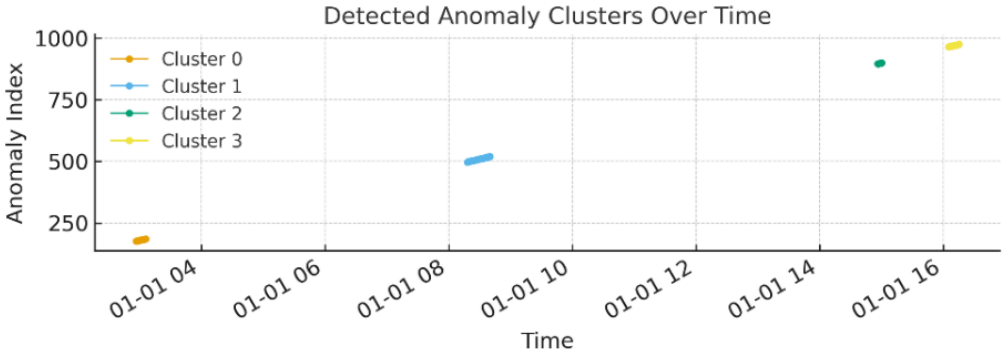

Figure 3: Detected Anomaly Clusters Over Time

5.3 Model Performance Comparison

Table 2: Performance summary of anomaly detection models across modalities.

Model	Data Type	Detection Accuracy	PR-AUC	FA/h	Detection Latency
LSTM Autoencoder	AIS / GPS (temporal)	94.2 %	0.91	0.14	5–8 min
Isolation Forest	Fuel & Temp Sensors	92.8 %	0.88	0.12	< 4 min
Isolation Forest	Compass Signals	90.3 %	0.85	0.16	6–10 min

5.4 Insights and Operational Implications

The results demonstrate that modality-specific anomaly detection substantially enhances precision and interpretability across the vessel's navigation and operational data streams. The LSTM autoencoder effectively identifies trajectory-based inconsistencies (Figure 1), capturing subtle spatial deviations that emerge only through temporal sequence modelling. In contrast, the Isolation Forest excels at detecting abrupt fuel-level and engine-temperature irregularities (Figure 2), flagging sudden drops, spikes, and implausible readings with minimal delay. The temporal grouping of anomalies (Figure 3) further reveals that abnormal events tend to occur

in concentrated clusters rather than as isolated incidents, providing valuable contextual insight for operational monitoring.

Detection latency varies significantly by data type: operational sensor anomalies are detected fastest (<4 minutes), whereas trajectory deviations require 5-10 minutes to accumulate sufficient evidence for reliable classification. These findings underscore the value of a hybrid, multi-model architecture in which each modality is paired with the model best suited to its statistical structure. By deploying complementary models within a modular system, VessiGuard maximises anomaly-detection coverage while maintaining low false-alarm rates. The resulting system offers low latency, interpretable outputs, and clear visual diagnostics suitable for real-time maritime monitoring, enabling operators to respond rapidly and prevent escalation during navigation or incident handling.

6. Limitations and Future Work

6.1 Dataset Constraints and Adversarial Robustness

A primary limitation of this study is its reliance on simulated and publicly available datasets. Although sufficient for proof-of-concept validation, these datasets do not fully capture the complexity and noise characteristics of real-world maritime environments. Furthermore, the current evaluation did not include adversarial robustness testing. Consequently, the system's resilience against intentionally crafted and adaptive cyberattacks, specifically designed to mislead machine learning models, remains to be validated. This limits the certainty of performance under sophisticated, targeted threat scenarios.

6.2 Operational Integration and Crew Impact

The current implementation was evaluated offline and not integrated into a real-time bridge system. Therefore, critical operational factors such as **integration challenges with existing navigation equipment**, onboard computational latency, and the **impact of false alarms on crew workload** were not empirically tested. Future deployment must also consider a **cost-benefit analysis** for operators, weighing the expense of retrofitting vessels with edge-computing hardware against the potential risk reduction. Addressing these barriers is essential to ensure the system supports, rather than distracts, vessel crews.

6.3 Future Directions

Future work will transform VessiGuard from a validated prototype into a deployable solution:

- Real-World Deployment and Sensing Expansion: Edge deployment on operational vessels to evaluate performance under bandwidth constraints and incorporate radar, weather, and vessel-to-infrastructure data for enhanced situational awareness.
- Adversarial Robustness Testing: Conducting rigorous stress-testing using adversarial samples to quantify and improve the model's resilience against adaptive evasion attacks.
- Human-Centred XAI Dashboard: Developing a regulation-compliant operator interface with Explainable AI (XAI). This will mitigate **alarm fatigue** by providing context-aware visualisations that help crews quickly distinguish between sensor faults and genuine cyber threats.
- System Integration and Advanced Learning: Implementing secure logging for immutable incident audit trails and exploring hybrid architectures (combining supervised and reinforcement learning) to improve adaptability across diverse vessel types.

7. Conclusion

This research presented VessiGuard, an intelligent anomaly detection system tailored to the evolving cybersecurity needs of the maritime sector. By combining Long Short-Term Memory (LSTM) models for sequential navigational **behaviour** and Isolation Forest algorithms for operational sensor anomalies, the system effectively identifies a wide range of threats, including GPS spoofing, sensor manipulation, and irregular vessel **behaviour**. The results validate the feasibility of modality-specific detection strategies, demonstrating strong detection performance and low false alarm rates. However, this work should be understood as a proof-of-concept, validated using simulation-based scenarios. Beyond its technical contributions, VessiGuard offers a foundational pathway toward adaptive maritime cybersecurity. While currently a prototype, its modular architecture and potential for future integration with existing vessel systems provide a blueprint for industry adoption. As the maritime domain continues its digital transformation, solutions like VessiGuard will be essential for safeguarding shipping infrastructure, ensuring vessels operate securely and resiliently in an increasingly complex threat landscape.

AI Declaration: Artificial intelligence and language support tools were used to prepare this manuscript. OpenAI's ChatGPT was utilised to assist with language refinement, structure enhancement, and technical phrasing, while Grammarly was employed for maintaining style consistency. The authors conducted all conceptualisation, system design, methodology, data analysis, and results interpretation. All AI-generated outputs were critically reviewed and verified for accuracy before inclusion.

Ethics and AI Declaration: This study did not involve human participants, sensitive personal data, or interventions requiring ethical approval. All data used were publicly available or synthetically generated within a controlled simulation; consequently, no institutional ethics clearance was required.

References

Amro, A., Oruc, A., Gkioulos, V. and Katsikas, S. (2022). 'Navigation data anomaly analysis and detection', *Information*, 13(3). Available at: https://www.mdpi.com/2078-2489/13/3/104

BIMCO, CLIA, ICS, INTERCARGO and INTERTANKO (2020). *The Guidelines on Cyber Security Onboard Ships (Version 5).* BIMCO, Bagsværd, Denmark.

Bobrovnikova, K., Lysenko, S., Hurman, I. and Kwiecien, A. (2022). 'Machine learning based techniques for cyberattack detection in the Internet of Things infrastructure', in *IntelITSIS*, pp. 411–420.

Bolbot, V., Kulkarni, K., Brunou, P., Banda, O.V. and Musharraf, M. (2022). 'Developments and research directions in maritime cybersecurity: A systematic literature review and bibliometric analysis', *International Journal of Critical Infrastructure Protection*, 39, 100571.

Brenner, B., Hollerer, S., Bhosale, P., Sauter, T., Kastner, W., Fabini, J. and Zseby, T. (2023). 'Better safe than sorry: Risk management based on a safety-augmented network intrusion detection system', *IEEE Open Journal of the Industrial Electronics Society*, 4, pp. 287–303.

Davari, N. and Aguiar, A.P. (2021). 'Real-time outlier detection applied to a Doppler velocity log sensor based on hybrid autoencoder and recurrent neural network', *IEEE Journal of Oceanic Engineering*, 46(4), pp. 1288–1301.

Golosova, J. and Romanovs, A. (2018). 'The advantages and disadvantages of the blockchain technology', in *IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pp. 1–6.

Hochreiter, S. and Schmidhuber, J. (1997). 'Long short-term memory', *Neural Computation*, 9(8), pp. 1735–1780.

Hu, Q., Han, W. and Zhang, H. (2021). 'Ship identity authentication security model based on blockchain', in *4th International Conference on Data Science and Information Technology*, pp. 135–142.

IMO (2017a). *Guidelines on maritime cyber risk management (MSC-FAL.1/Circ.3).* International Maritime Organization, London.

IMO (2017b). *Maritime cyber risk management in safety management systems (MSC.428(98)).* International Maritime Organization, London.

Kim, H. and Joe, I. (2024). 'Enhancing anomaly detection in maritime operational IoT time series data with synthetic outliers', *Electronics*, 13(19). Available at: https://www.mdpi.com/2079-9292/13/19/3912

Liu, F.T., Ting, K.M. and Zhou, Z.-H. (2008). 'Isolation forest', *IEEE International Conference on Data Mining*, pp. 413–422.

Maganaris, C., Protopapadakis, E. and Doulamis, N. (2024). 'Outlier detection in maritime environments using AIS data and deep recurrent architectures', *CoRR*, abs/2406.09966. Available at: https://doi.org/10.48550/arXiv.2406.09966

Nguyen, D., Simonin, M., Hajduch, G., Vadaine, R., Tedeschi, C. and Fablet, R. (2020). 'Detection of abnormal vessel behaviours from AIS data using GeoTrackNet: From the laboratory to the ocean', in *2020 21st IEEE International Conference on Mobile Data Management (MDM)*, pp. 264–268.

Ribeiro, C.V., Paes, A. and de Oliveira, D. (2023). 'AIS-based maritime anomaly traffic detection: A review', *Expert Systems with Applications*, 231, p. 120561.

Riveiro, M., Pallotta, G. and Vespe, M. (2018). 'Maritime anomaly detection: A review', *WIREs Data Mining and Knowledge Discovery*, 8(5), e1266.

School of International and Public Affairs, Columbia (2022). *NotPetya: Lessons learned from the first global cyberattack on critical infrastructure.* Available at: https://www.sipa.columbia.edu/sites/default/files/2022-11/NotPetya%20Final.pdf

Talukder, M.A., Islam, M.M., Uddin, M.A., Hasan, K.F., Sharmin, S., Alyami, S.A. and Moni, M.A. (2024) 'Machine learning-based network intrusion detection for big and imbalanced data', *Journal of Big Data*, 11(1), p. 33.

Varma, R. (2024). *VessiGuard Synthetic Maritime Sensor and Navigation Dataset (Version 1.0).* Unpublished raw data generated by the author.

Wang, W., Bin, J., Zaji, A., Halldearn, R., Guillaume, F., Li, E. and Liu, Z. (2022). 'A multi-task learning-based framework for global maritime trajectory and destination prediction with AIS data', *Maritime Transport Research*, 3, p. 100072.