Zero-Trust Marine Cyberdefense for IoT-Based Communications

Subjects: Engineering, Marine | Computer Science, Artificial Intelligence | Computer Science, Information Systems
Contributor: Ebuka Chinaechetam Nkoro, Judith Nkechinyere Njoku, Cosmas Ifeanyi Nwakanma, Jae-Min Lee, Dong-Seong Kim

Integrating Explainable Artificial Intelligence (XAI) into marine cyberdefense systems can address the lack of trustworthiness and low interpretability inherent in complex black-box Network Intrusion Detection Systems (NIDS) models. XAI has emerged as a pivotal focus in achieving a zero-trust cybersecurity strategy within marine communication networks.

Keywords: Explainable Artificial Intelligence (XAI) ; zero-trust security ; IoT ; Network Intrusion Detection Systems (NIDS) ; marine cybersecurity ; communications ; transportation ; cybersecurity ; machine learning

1. Introduction

Approximately 71% of the Earth's surface is subaqueous, significantly influencing the territorial, geographical, and economic landscapes of nation-states. In the current era of cyberwarfare [1], maritime organizations bear the crucial responsibility of establishing, managing, and securing marine networks to mitigate the risks of breaches and vulnerabilities. The frequency of cyber incidents in the maritime sector has witnessed a notable increase in recent years, exemplified by the 2020 ransomware attack on industry giants such as MAERSK. This attack resulted in substantial financial losses, estimated to be between 200 and 300 million USD. Additionally, intranet breaches targeting the International Maritime Organization (IMO) have raised security and reputational concerns [2]. The genesis of major marine cyberattacks often stems from vulnerabilities in Internet of Things (IoT) and Internet of Underwater Things (IoUT) sensors, which malicious actors exploit to initiate and perpetuate intrusions into maritime systems [3].

To strengthen marine cyberdefense systems, previous research has investigated the use of Artificial Intelligence (AI) frameworks to improve maritime Network Intrusion Detection Systems (NIDS) [4], thus guaranteeing faster and more reliable detection of cyberattacks such as Distributed Denial of Service (DDoS) attacks, ransomware, phishing, and backdoor attacks. Strong learning algorithms, such as deep neural networks, have been used to guarantee highly accurate predictions in marine NIDS, due to their ability to capture the spatial relations of IoT/IoUT network traffic data and detect malicious threats [5].

Major challenges within the introduction of AI in marine NIDS are outlined below:

(i) The prevalence of false alarm rates, fake distress calls, and especially the lack of explanations regarding the black-box AI algorithms used to predict marine cyberattacks [6]. Meanwhile, marine cyberdefense systems now require Explainable AI (XAI) frameworks and human-in-the-loop interactions for security experts to provide reliable and trustworthy predictions of marine cyberthreats.

(ii) Most XAI interpretation methods reported in the current literature focus majorly on visual explanations and still lack quantitative XAI metrics that can aid expert decisions or methods.

To overcome the above-highlighted challenges, visual, quantitative, and human-in-the-loop XAI can be employed to salvage the challenges of reliability and transparency in marine NIDS. A current cyberdefense paradigm, Zero-trust Architecture (ZTA), as proposed by the United States Department of Defence (DoD) in 2022, highlights a holistic approach that embodies real-time network traffic monitoring, strong authentication, and continuous evaluation of the confidence levels of AI-based NIDS models to address transparency and reliability in cybersecurity issues. ZTA adopts the "trust no one, verify everything" principle, thus providing NIDS experts with better understanding, reliability, and authentication of network users and mitigation of security threats just-in-time [7]. Within this strategy, explainable NIDS are layered to understand and prevent the stealthy advances of attackers whose aim is to tamper with the confidentiality, integrity, and availability of marine cyberspace.

For example, a marine cybersecurity expert might wish to investigate the following question: "How certain is this NIDS model's prediction of a normal or DDoS attack, and what training features led to the NIDS prediction?" To address the lack of transparency and model trustworthiness of most NIDS models, the growing area of XAI aims to address the major reasons for model distrust and provide security experts with insight-driven feedback for the improved security posture of their organizations [8]. Although recent works have begun studying XAI, only a few of them have addressed cybersecurity

concerns related to marine cyberdefense. Other works have not provided quantitative and secure methods that help to differentiate malicious alerts and improve expert decisions and model trustworthiness [9].

2. Zero-Trust Marine Cyberdefense for IoT-Based Communications

2.1. Cyberdefense in Marine Networks

The broad term “marine cyberdefense” represents the security of a broad range of marine sectors, including vessels, offshore and onshore facilities, navigation and transport systems, and cargo systems that rely on networked IoT and IoUT technologies that facilitate day-to-day marine operations [10][11]. Each marine system, depending on the type of application, is enabled with peripherals such as microphones, cameras, sound and image processing units, GPS units, and a collaborative communication mechanism where sensor nodes broadcast their data packets to neighboring nodes until data exchange is achieved. Due to the tremendous amount of data generated from the integration of these multiple marine systems, cybercriminals now leverage the vulnerabilities in IoT and IoUT communications to perpetuate marine cyberattacks [12]. Meanwhile, the surface attacks witnessed in marine organizations may vary specifically from regular cyber scenarios in terms of the underlying network infrastructure compromising of complexities in marine supply chain systems, marine GPS vulnerabilities, and even marine communication jamming attacks, which ordinary businesses do not usually witness [13].

Severe cyberattacks have been reported by shipping industries and the IMO [14] concerning the prevalence of attackers (hacktivists, terrorists, digital pirates, and ransomware groups) who disrupt marine networks in the form of DDoS attacks or steal confidential information for financial gain [15]. As shown in **Figure 1** and **Figure 2**, there have been several cyberattacks detected by leading marine cybersecurity experts in 2023 alone, showing a continual increase in cyber incidents within marine environments (shipping, supply chain, energy, yard, port, defense, marine organizations, and vessel operations) [16]. These proliferated attacks can be linked to the fast-paced stealthiness of modern attackers, vulnerabilities of IoT and IoUT technologies, and most especially, lack of real-time defense mechanisms such as NIDS [2].

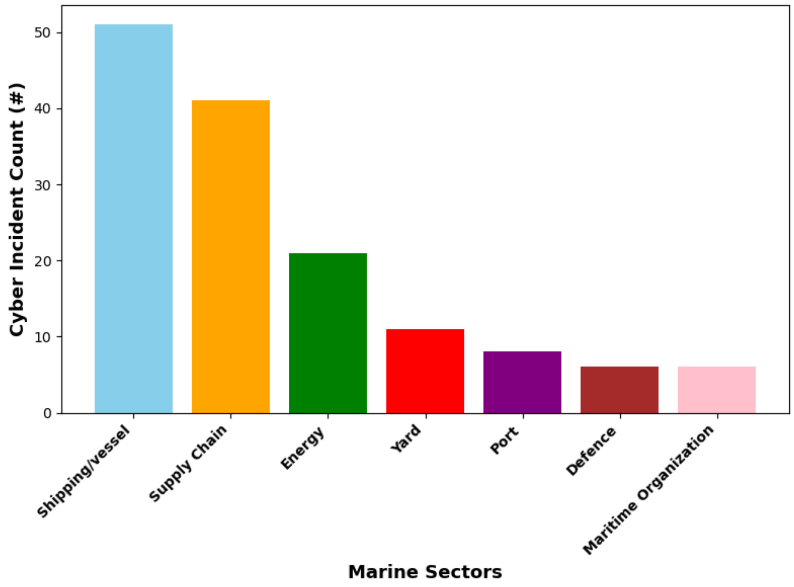

Figure 1. Cyberattack incidents within various marine sectors in 2023 where # signifies the cyber incident count.

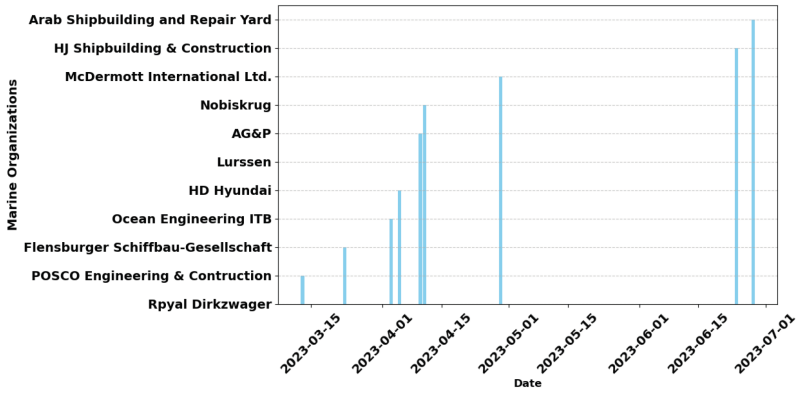

Figure 2. Timeline of cyber incidents in yard marine industry 2023.

As shown in **Figure 3**, marine networks employ automated IoT and IoUT systems that foster ship-to-ship communication, which optimizes marine productivity while reducing operational costs. Marine network communication is characterized by interoperable nodes such as base stations, coastal units, and the Software-Defined Network (SDN) controller. The

network control center transmits and receives several different wireless technologies, such as Long-term Evolution Advanced (LTE-A), Wireless Fidelity (Wi-Fi) networks, satellites, and acoustic communications (buoys) [17][18]. All marine communication nodes are geographically distributed among the different regions, including coastal, offshore, open-sea, or underwater communication endpoints to facilitate continuous communications and the running of the marine industry.

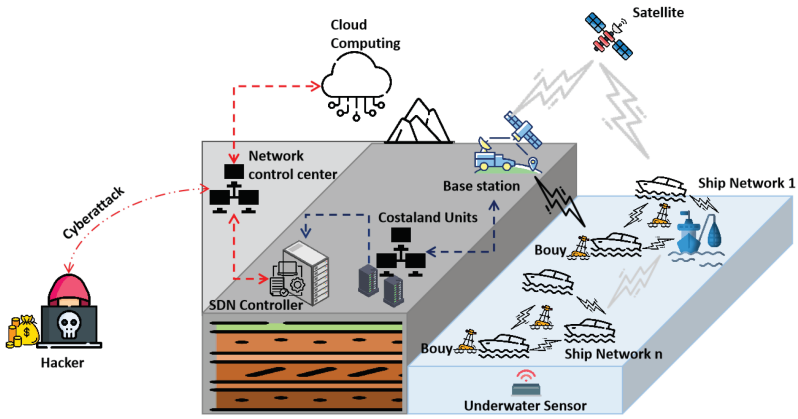

Figure 3. Illustration of marine network communications and cyberthreats.

However, stealthy hacker groups can leverage network system vulnerabilities to perpetuate cyber incidents such as vessel engine failures, manipulation of control systems, and jamming attacks [19]. Recovering from cyberattacks has become very expensive. In January 2023 a hacker group 'PLAY' published the private data of four European Union marine IT companies on the darkweb after December 2022's ransomware infection. Within the same year, another recognized advisor for the maritime industry, Det Norske Veritas (a leading classification society and a recognized advisor for the maritime industry), was reported as a fresh target, where hackers compromised the data of ship management companies, which account for 21% of the total share of the marine industry [20].

2.2. NIDS for Cyberdefense in Marine Networks

To overcome marine cyber incidents in a very responsive manner, the use of network intrusion detection systems has been employed by security organizations and in the domain of marine security to detect and automate potential attacks, thus preserving marine security postures and organizational reputation. NIDS methods, as explored by previous works [21], can be categorized as follows: signature-based, anomaly-based, flow-based, and machine-learning-enabled NIDS aimed at detecting anomalous network traffic while minimizing the number of false-positive predictions.

Traditional machine-learning algorithms have been explored in the field of IoT-enabled NIDS for secure marine network operations. In a study on IoT botnet attack detection, Alqahtani, Mathkour, and Ismail (2020) proposed a method based on the optimized extreme Light Gradient Boosting (LGB) algorithm for detecting and protecting IoT devices from dangerous large-scale botnet attacks. Researchers' previous work [22] also utilized the LGB due to its fast computation, which is vital for fast and cost-efficient computation in IoT networks. The results of the multi-class results yielded a 95% accuracy while predicting 12 diverse cyberattack types using the 2023 EdgeIIoT dataset.

Unlike traditional machine-learning classifiers, which are limited in their ability to extract features from massive data, considering the extensive cyber traffic in real life, the use of deep-learning algorithms for network traffic classification has been preferred in modern research [5].

Earlier work addressed the gap in explainable NIDS models within the domain of marine cyberdefense. Furthermore, this study was supplemented with an additional dataset, well-investigated feature selection methods, visual XAI, and a significant quantitative XAI interpretation of the proposed "black-box" neural network model. In comparison with the tree-based algorithm used for the classification task in researchers' prior work [22], neural network NIDS models are inherently not easily interpretable [8].

Meanwhile, Hou et al. in [23], proposed an intrusion detection framework for hydrographic station network anomalies. The proposed approach utilized a hybrid CNN and BiLSTM method using the NSL KDD dataset while obtaining an F1-score of 87.35%. Although their approach was effective in identifying deep features, the low accuracy of their results cannot be ignored, taking into consideration efficiency requirements and the need for the low false-positive rates required for a zero-trust model in marine networks.

Xin et al. in [24], proposed a Generative Adversarial Network (GAN) approach to process the imbalanced NSL-KDD dataset [25] for IDS in marine networks. Within their method, a data generation module was initiated to improve minority class samples, using the OPTICS denoising algorithm. The classification accuracy of the authors' proposed data augmentation method yielded a micro-average accuracy of 95% with five classes of network traffic. A decentralized training method using a federated learning approach for marine IDS was investigated by authors in [4]. Their federated learning technique was designed to save computing and storage overhead, with an accuracy of 87%, 500 rounds of

training, and the use of the old NSL-KDD dataset. Dataset dimensionality in the domain of NIDS availability, suitability, and dimensionality in the domain of NIDS has become a bottleneck for the efficient and effective correlation of network traffic for improved model accuracy. Therefore, the use of obsolete datasets such as NSL-KDD may not fit the current demands of modern networks.

2.3. Zero-Trust Cyberdefense in IoT

The zero-trust security architecture, as recently published by the National Institute of Standards and Technology (NIST), is a paradigm shift towards rethinking the network security and protection of organizational assets. The strength of ZTA principles in IoT and marine cyberdefense lies in its skepticism [26], i.e., "assume breach, verify explicitly, privilege access only", and not blind trust, thus supporting multi-level authorization/scrutiny to achieve fine-grained security controls. ZTA embraces five core tenets, as shown in **Figure 4**, namely:

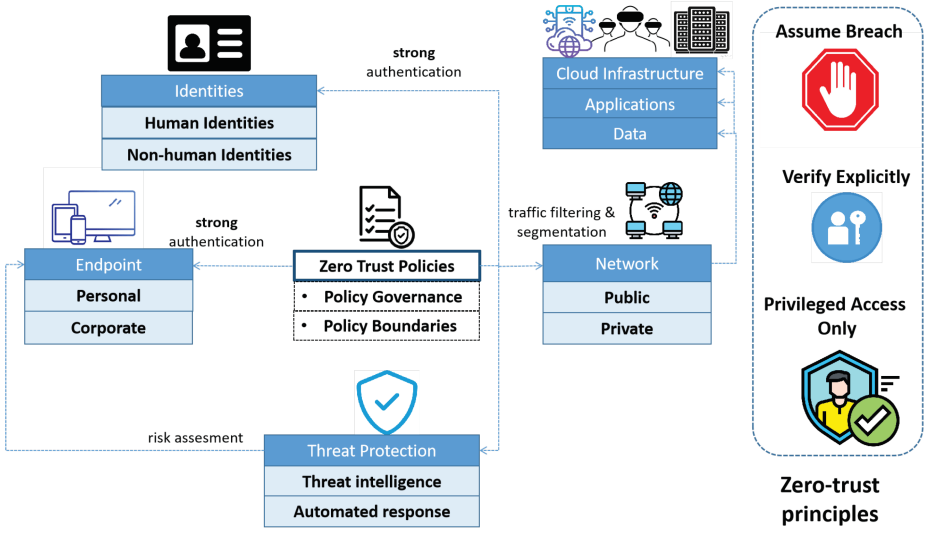

Figure 4. A visual illustration of the zero-trust cybersecurity principles, governed by strong authentication, filtering, threat intelligence, and zero-trust policies.

i Resource segmentation;

ii Ubiquitous authentication;

iii Strong encryption;

iv Principle of least privilege;

v Intelligent real-time threat monitoring.

The US DoD released a ZTA framework for integrated threat intelligence and remediation [27]. Therein, machine-learning analytics, real-time network traffic monitoring, and orchestration capabilities were employed to enforce the DoD's data/enterprise security against cyberthreats. In addition, evaluating the confidence levels of the DoD's ML models, devices, users, and resources is routinely performed to ensure minimal security vulnerabilities. Recent advancements towards the Industry 5.0 paradigm now require zero-trust network-based access using AI for effective cybersecurity and real-time monitoring, controlling, and allocations of production sequences to prevent false rates and maximize productivity [28].

Current academic research has also investigated the employment of the ZTA for strong IoT security. Recent studies in [29] have proposed that the ZTA model will address most of the security concerns in 5G networks, where security models can dynamically detect/identify the malicious activities of users, devices, or applications. Proof-of-concept experiments using blockchain have also been employed to satisfy security requirements in edge computing networks [30]. The simulation results showed that ZTA in edge networks can satisfy successful edge node authentication with good time constraints. A recent trend towards adopting ZTA NIDS using deep learning has also yielded the increased security of network devices by calculating the security scores/awareness of imminent security threats [31], thus satisfying the ZTA demand for real-time monitoring and mitigation against threats. Syed et al. [32] presented a comprehensive survey, highlighting the relevance of AI-based NIDS for full ZTA realization in IoT-based networks. The survey, like previous ZTA policies in [27], outline the need to evenly distribute the zero-trust principle within machine–machine communications, IoT devices, security protocols, and even AI models. The authors in [33] have also emphasized how AI-based NIDS methods can be employed for resilient zero-trust IoT defense. Here, AI models can establish a probabilistic relation between a Cyber–Physical System (CPS) hypothesis (i.e., likelihood of attacks) and the supporting evidence (i.e., signs of attack activities); thus, even the slightest malicious activities can still be detected in real-time, and with high confidence, as long as enough

evidence is accumulated. However, the development of AI algorithms requires scrutiny and interpretability to ensure that they make predictions as required.

2.4. XAI for Cyberdefense

The development of trustworthy, transparent, and reliable algorithms has gained tremendous momentum in modern AI development. Recently, in October 2023, President Biden issued an executive order on safe, secure, and trustworthy AI, which requires that developers provide interpretable and trustworthy models during training to satisfy the safety/security of AI-enabled CPSs, software, and networks [34]. To address the potential risks (bais, transparency, privacy) and challenges associated with the widespread adoption of AI in IoT, there has been a huge interest in the field of XAI models by cybersecurity experts and researchers in IoT-enabled CPSs. In the domain of NIDS, for example, diverse questions such as “Why should we implicitly trust the predictions of NIDS?” [9] and “How certain is this model’s prediction of a cyberattack?”. These are the basic answers that an explainable NIDS seeks to address.

Within the domain of XAI, two categories of XAI periods (time of model explanation) post-hoc and ad-hoc, have been adopted for interpretable NIDS [8]. As illustrated in **Figure 5**, the current XAI taxonomy can be classified within time, complexity, and scope requirements. Concerning timely requirements, the ad-hoc explainability method provides model explanations during its decision process, while the post-hoc methods offer explainability information after model prediction in terms of intrinsic explanations, such as feature contributions to the model output. Commonly used post-hoc explainability methods in NIDS are SHAP and LIME. The SHAP and LIME explainers have become a favorite choice for explainable NIDS due to their model-agnostic features, while providing explanations to deep-learning-based NIDS models, which are fairly opaque in their decision-making process. Meanwhile, the complexity of interpretation methods can be classified as either extrinsic (model-agnostic) or intrinsic (model-specific). For example, tree-based (rule-based) NIDS models are inherently interpretable depending on the dataset or complex nature of training. As regards the scope of interpretations, XAI techniques can be categorized as either global or local [8].

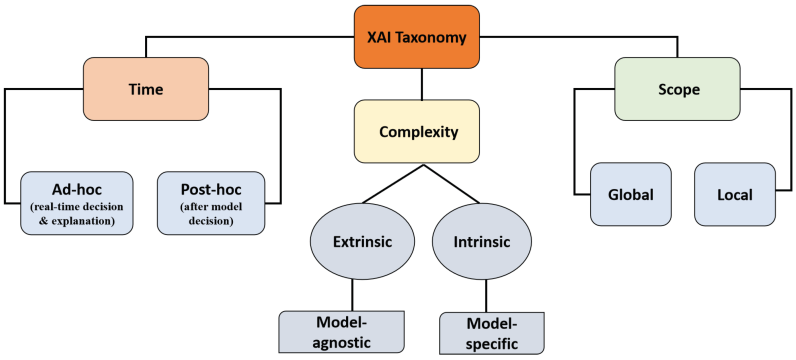

Figure 5. A summarized visual taxonomy of XAI methods.

Within current XAI frameworks for NIDS, the SHAP and LIME XAI methods possess specific suitability and efficiency with network traffic data, and they are an alternative to computer vision-based XAI methods such as Generalization of the Class Activation mapping (Grad-CAM), Guided Grad-CAM, and axiom-based Grad-CAM, which may be computationally expensive while converting network traffic (text) to images, given the real-time demands for security, efficiency, and interpretability. The SHAP explainer provides the marginal value of contributions made by a feature or subset of features within a model’s prediction. Similarly, the LIME explainer can generate local surrogate models to approximate the decision-making process of a complex model, providing interpretable explanations for individual predictions by highlighting important model features [35].

Current research into various cyberattacks, such as phishing attacks, botnets, and fraud, is gaining better insights, proper visualizations, and deeper forensics into the nature of these attacks. Additionally, significant features for model training can be identified to perform effective/trustworthy cyberdefense [8]. As depicted in **Figure 6**, NIDS approaches in IoT CPSs can model explainability and trustworthiness to evaluate the credibility of the predicted cyberattacks. Other layers of the zero-trust model cover areas such as physical barriers or mechanisms, which can help marine organizations to prevent, monitor, or detect unauthorized access to their assets through the use of locks, rails, CCTVs, badges, PC locks, turnstiles, and alarms. Important perimeter defenses includes identity access, perimeter security, compute, application security, and data integrity.

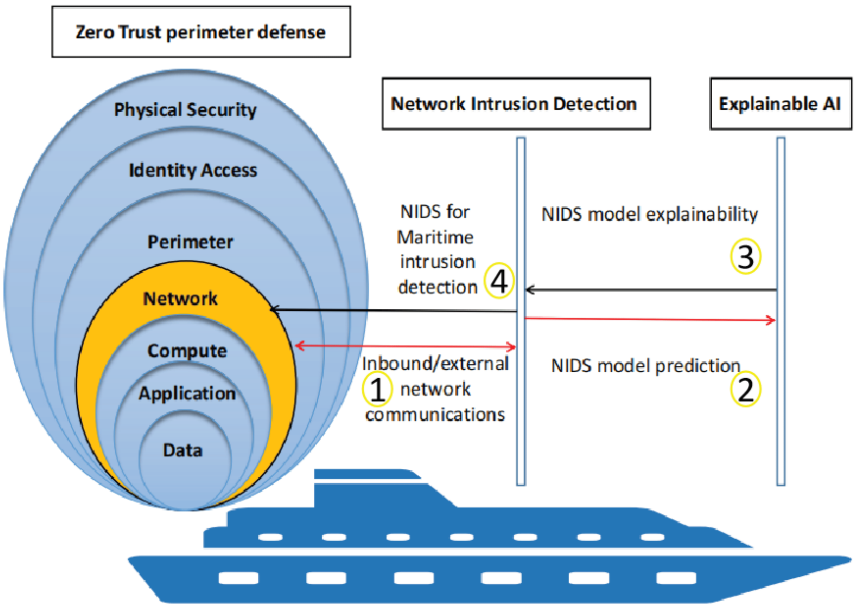

Figure 6. Sequence diagram of the zero-trust perimeter defense strategy for marine networks, with insight-driven feedback using an Explainable AI (XAI) Network Intrusion Detection (NIDS) model.

A malware detection method in [36] investigated an explainable malware detection method to understand the "outstanding" performance of their proposed model in real-world environments. One drawback in their approach was a lack of visual interpretations or XAI methods, for better interpretability. A lack of visual explainability of the XAI frameworks hinders the understanding, debugging, and decision-making within the proposed system. To address the lack of XAI in NIDS, [37] presented an explainable ANN DL model using the CICIDS 2017 dataset [38]. The authors utilized the oracle module, which also showed limited explainability of the model results.

To provide additional insights and forensics into NIDS models, Shruti et al. [39] only explored the visual LIME explainability using naturally transparent machine-learning algorithms, such as decision trees, random forests, and SVM. Another approach [40] using the SHAP explainer investigated a deep-learning approach named the trustworthy explainable artificial intelligence and enhanced krill herd optimisation intrusion detection system to detect breaches in IoT-enabled CPSs. Using the NSL KDD dataset [41] and the CICIDS 2018 dataset [38], their explanations using SHAP yielded insights towards the significant impact of training features in terms of the proposed NIDS model classification accuracy. Mohammed [42] proposed packet-based efficient and explainable IoT botnet detection using machine learning. The SHAP discussions using the Shapley additive explanation also provided transparency to the classifier's prediction process.

Zakaria et al. [9] designed a deep neural network XAI-based framework using the SHAP, LIME, and RuleFit XAI methods to explain their proposed NIDS framework, which was aimed at detecting IoT-related intrusions. However, the proposed system included a non-informative and redundant network traffic feature—source IP address ('srcip')—which is only meaningful within the dataset explored (NSL KDD). The 'srcip' would, by default, gain a high weight, thus dominating the SHAP plot and the model's predictions. This dominance led to a false conclusion that the source IP address is the most critical feature for the proposed NIDS, which may not be the case in a general scenario. The proposed approach does not generalize to the robust defense, interpretation, and transparency of NIDS models. A spoofing detection method in [43] provided both the LIME and SHAP explainability results of cyberattacks in IoT networks. Their work still lacks state-of-the-art evaluation metrics and a better discussion of the explainability results in terms of quantitative decisions and confidence in model predictions. It is therefore expedient, as required by modern defenses in the domain of NIDS, to provide XAI interpretations to bolster the trust and reliability of NIDS prediction.

References

1. Serpanos, D.; Komninos, T. The Cyberwarfare in Ukraine. Computer 2022, 55, 88–91.
2. Park, C.; Kontovas, C.; Yang, Z.; Chang, C.H. A BN driven FMEA approach to assess maritime cybersecurity risks. Ocean Coast. Manag. 2023, 235, 106480.
3. Mohsan, S.A.H.; Li, Y.; Sadiq, M.; Liang, J.; Khan, M.A. Recent Advances, Future Trends, Applications and Challenges of Internet of Underwater Things (IoUT): A Comprehensive Review. J. Mar. Sci. Eng. 2023, 11, 124.
4. Liu, W.; Xu, X.; Wu, L.; Qi, L.; Jolfaei, A.; Ding, W.; Khosravi, M.R. Intrusion Detection for Maritime Transportation Systems With Batch Federated Aggregation. IEEE Trans. Intell. Transp. Syst. 2023, 24, 2503–2514.
5. Dong, B.; Wang, X. Comparison deep-learning method to traditional methods using for network intrusion detection. In Proceedings of the 2016 8th IEEE International Conference on Communication Software and Networks (ICCSN), Beijing, China, 4–6 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 581–585.

6. Nwakanma, C.I.; Ahakonye, L.A.C.; Njoku, J.N.; Odirichukwu, J.C.; Okolie, S.A.; Uzondu, C.; Ndubuisi Nweke, C.C.; Kim, D.S. Explainable Artificial Intelligence (XAI) for Intrusion Detection and Mitigation in Intelligent Connected Vehicles: A Review. Appl. Sci. 2023, 13, 1252.
7. Shore, M.; Zeadally, S.; Keshariya, A. Zero Trust: The What, How, Why, and When. Computer 2021, 54, 26–35.
8. Capuano, N.; Fenza, G.; Loia, V.; Stanzione, C. Explainable Artificial Intelligence in CyberSecurity: A Survey. IEEE Access 2022, 10, 93575–93600.
9. Houda, Z.A.E.; Brik, B.; Khoukhi, L. “Why Should I Trust Your IDS?”: An Explainable Deep Learning Framework for Intrusion Detection Systems in Internet of Things Networks. IEEE Open J. Commun. Soc. 2022, 3, 1164–1176.
10. Ali, E.S.; Saeed, R.A.; Eltahir, I.K.; Khalifa, O.O. A systematic review on energy efficiency in the internet of underwater things (IoUT): Recent approaches and research gaps. J. Netw. Comput. Appl. 2023, 213, 103594.
11. Khan, Z.U.; Gang, Q.; Muhammad, A.; Muzzammil, M.; Khan, S.U.; Affendi, M.E.; Ali, G.; Ullah, I.; Khan, J. A comprehensive survey of energy-efficient MAC and routing protocols for underwater wireless sensor networks. Electronics 2022, 11, 3015.
12. Heering, D.; Maennel, O.; Venables, A. Shortcomings in cybersecurity education for seafarers. In Maritime Technology and Engineering 5 Volume 1; CRC Press: Boca Raton, FL, USA, 2021; pp. 49–61.
13. Jacq, O.; Boudvin, X.; Brosset, D.; Kermarrec, Y.; Simonin, J. Detecting and hunting cyberthreats in a maritime environment: Specification and experimentation of a maritime cybersecurity operations centre. In Proceedings of the 2018 2nd Cyber Security in Networking Conference (CSNet), Paris, France, 24–26 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–8.
14. Lin, W.C. Maritime Environment Assessment and Management Using through Balanced Scorecard by Using DEMATEL and ANP Technique. Int. J. Environ. Res. Public Health 2022, 19, 2873.
15. Akpan, F.; Bendiab, G.; Shiaeles, S.; Karamperidis, S.; Michaloliakos, M. Cybersecurity Challenges in the Maritime Sector. Network 2022, 2, 123–138.
16. Jo, Y. Cyberattack Incidents in Maritime Sector. Available online: https://https://www.cytur.net/ (accessed on 10 October 2023).
17. Koulouras, I.; Bobotsaris, I.; Margariti, S.V.; Stergiou, E.; Stylios, C. Assessment of SDN Controllers in Wireless Environment Using a Multi-Criteria Technique. Information 2023, 14, 476.
18. Liang, M.; Su, X.; Liu, X.; Zhang, X. Intelligent ocean convergence platform based on iot empowered with edge computing. J. Internet Technol. 2020, 21, 235–244.
19. Chen, H.; Yin, F.; Huang, W.; Liu, M.; Li, D. Ocean Surface Drifting Buoy System Based on UAV-Enabled Wireless Powered Relay Network. Sensors 2020, 20, 2598.
20. Jongwoo, A. KR Maritime Cyber Safety News & Report. Available online: https://www.krs.co.kr/Common/Com_Popup/Com_FileDown.aspx?DATA1=7rF67H0cjeYuxn6YdejCySra1U5wS9J0jjGzbttW1YbZqalp5CIKgYVcAVRi6k!_!_!V&DATA2=W241p64Xg7ER4wTHluR9Dw==&DAT (accessed on 9 September 2023).
21. Rehman, M.H.U.; Dirir, A.M.; Salah, K.; Damiani, E.; Svetinovic, D. TrustFed: A Framework for Fair and Trustworthy Cross-Device Federated Learning in IIoT. IEEE Trans. Ind. Inform. 2021, 17, 8485–8494.
22. Nkoro, E.C.; Njoku, J.N.; Nwakanma, C.I.; Lee, J.M.; Kim, D.S. SHAP-Based Intrusion Detection Framework for Zero-Trust IoT Maritime Security. In Proceedings of the 2023 the 2nd International Conference on Maritime IT Convergence (ICMIC), Jeju Island, Republic of Korea, 23–25 August 2023; pp. 1–8.
23. Hou, T.; Xing, H.; Liang, X.; Su, X.; Wang, Z. A Marine Hydrographic Station Networks Intrusion Detection Method Based on LCVAE and CNN-BiLSTM. J. Mar. Sci. Eng. 2023, 11, 221.
24. Su, X.; Tian, T.; Cai, L.; Ye, B.; Xing, H. A CVAE-GAN-based Approach to Process Imbalanced Datasets for Intrusion Detection in Marine Meteorological Sensor Networks. In Proceedings of the 2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), Melbourne, Australia, 17–19 December 2022; pp. 197–203.
25. Kalluri, R.; Mahendra, L.; Kumar, R.S.; Prasad, G.G. Simulation and Impact Analysis of Denial-of-Service Attacks on Power SCADA. In Proceedings of the 2016 National Power Systems Conference (NPSC), Bhubaneswar, India, 19–21 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–5.
26. Stafford, V. Zero trust architecture. NIST Spec. Publ. 2020, 800, 207.
27. Freter, R. Department of Defence (DoD) Zero Trust Reference Architecture, Version 2.0. In Proceedings of the Defense Information Systems Agency (DISA) and National Security Agency (NSA); July 2022. Available online: https://dodcio.defense.gov/Portals/0/Documents/Library/(U)ZT_RA_v2.0(U)_Sep22.pdf (accessed on 9 September 2023).
28. Abuhasel, K.A. A Zero-Trust Network-Based Access Control Scheme for Sustainable and Resilient Industry 5.0. IEEE Access 2023, 11, 116398–116409.
29. Li, S.; Iqbal, M.; Saxena, N. Future industry internet of things with zero-trust security. Inf. Syst. Front. 2022, 1–14.

30. Ali, B.; Hijjawi, S.; Campbell, L.H.; Gregory, M.A.; Li, S. A maturity framework for zero-trust security in multiaccess edge computing. Secur. Commun. Netw. 2022, 3178760, 1–14.
31. Lee, B.; Vanickis, R.; Rogelio, F.; Jacob, P. Situational awareness based risk-adapatable access control in enterprise networks. arXiv 2017, arXiv:1710.09696.
32. Syed, N.F.; Shah, S.W.; Shaghaghi, A.; Anwar, A.; Baig, Z.; Doss, R. Zero Trust Architecture (ZTA): A Comprehensive Survey. IEEE Access 2022, 10, 57143–57179.
33. Restuccia, F.; D'Oro, S.; Melodia, T. Securing the Internet of Things in the Age of Machine Learning and Software-Defined Networking. IEEE Internet Things J. 2018, 5, 4829–4842.
34. House, W. FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. Available online: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ (accessed on 1 October 2023).
35. Dieber, J.; Kirrane, S. Why model why? Assessing the strengths and limitations of LIME. arXiv 2020, arXiv:2012.00093. Available online: http://arxiv.org/abs/2012.00093 (accessed on 21 July 2023).
36. Liu, Y.; Tantithamthavorn, C.; Li, L.; Liu, Y. Explainable AI for Android Malware Detection: Towards Understanding Why the Models Perform So Well? In Proceedings of the 2022 IEEE 33rd International Symposium on Software Reliability Engineering (ISSRE), Charlotte, NC, USA, 31 October–3 November 2022; pp. 169–180.
37. Szczepański, M.; Choraś, M.; Pawlicki, M.; Kozik, R. Achieving Explainability of Intrusion Detection System by Hybrid Oracle-Explainer Approach. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8.
38. Sharafaldin, I.; Lashkari, A.H.; Ghorbani, A.A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSp 2018, 1, 108–116.
39. Patil, S.; Varadarajan, V.; Mazhar, S.M.; Sahibzada, A.; Ahmed, N.; Sinha, O.; Kumar, S.; Shaw, K.; Kotecha, K. Explainable Artificial Intelligence for Intrusion Detection System. Electronics 2022, 11, 3079.
40. Sivamohan, S.; Sridhar, S.; Krishnaveni, S. TEA-EKHO-IDS: An intrusion detection system for industrial CPS with trustworthy explainable AI and enhanced krill herd optimization. Peer Peer Netw. Appl. 2023, 16, 1993–2021.
41. Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 8–10 July 2019; IEEE: Piscataway, NJ, USA, 2009; pp. 1–6.
42. Alani, M.M. BotStop: Packet-based efficient and explainable IoT botnet detection using machine learning. Comput. Commun. 2022, 193, 53–62.
43. Alani, M.M.; Awad, A.I.; Barka, E. ARP-PROBE: An ARP spoofing detector for Internet of Things networks using explainable deep learning. Internet Things 2023, 23, 100861.