https://doi.org/10.7840/kics.2025.50.12.1937

Blockchain-Aided Intrusion Detection in Marine Tactical Network Using Reinforcement Learning

Md Raihan Subhan◆, Md Mahinur Alam*, Mohtasin Golam**,
Md Facklasur Rahaman***, Taesoo Jun°

ABSTRACT

Marine Tactical Networks (MTNs) are essential for secure maritime operations, but are highly susceptible to cyber threats. Traditional Intrusion Detection Systems (IDS) often struggle to adapt to the dynamic and complex nature of MTNs. This paper introduces a Blockchain-Aided Intrusion Detection System (BAE-RL), which integrates reinforcement learning (RL) and blockchain technology to improve threat detection and security. The BAE-RL framework is unique in its use of multi-agent adversarial RL, where a defender agent learns to detect attacks by interacting with a simulated attacker agent. This adversarial setup enhances the system' s ability to identify novel and evolving threats. Additionally, blockchain integration ensures the integrity and immutability of detection data, preventing tampering and ensuring transparency. Experimental results show that the proposed framework outperforms traditional IDS, achieving 80.16% and 95.9% accuracy on the NSL-KDD and AWID datasets, respectively. The BAE-RL framework offers a robust, adaptive, and secure solution for intrusion detection in MTNs.

Key Words : Blockchain, intrusion detection system (IDS), marine tactical network (MTN), AI, reinforcement learning (RL), maritime applications, information security.

Ⅰ. Introduction

The Marine Tactical Network (MTN) serves as the backbone of modern maritime operations, integrating satellite links, radio frequency communications, and fiber optics to enable secure command, control, communications, computers, and intelligence (C4I) systems[1-4]. These networks are essential for facilitating marine operations, interconnecting components such as vessels, navigation systems, ports, shore-based facilities, and autonomous underwater vehicles. These networks enable critical functions such as real-time tracking, traffic management, and safety systems, supporting global maritime trade. Designed to overcome the unique challenges of marine environments including vast operational distances and scattered landmasses the MTN ensures real-time data transmission across naval vessels, aircraft, and terrestrial command centers while maintaining interoperability and survivability[5-8,19]. As shown in Table 1, this heterogeneous network differs fundamentally from terrestrial systems through its emphasis on adaptive topology and mul-

※ This research was supported by Kumoh National Institute of Technology (2024~2025) and the Gyeongsangbuk-do RISE (Regional Innovation System & Education) project (Idea Start-up Valley unit).
◆ First Author : The University of Texas Rio Grande Valley, Department of Computer Science with Interdisciplinary Applications, md.subhan01@utrgv.edu, 학생회원
° Corresponding Author : Kumoh National Institute of Technology, School of Computer Engineering, taesoo.jun@kumoh.ac.kr, 정회원
* Kumoh National Institute of Technology, mahinuralam213@kumoh.ac.kr
** Kumoh National Institute of Technology, ICT Convergence Research Center, golam248@kumoh.ac.kr, 학생회원
*** Faculty of Engineering and Applied Science in Electronic Systems Engineering department in University of Regina, MRL106@UREGINA.CA
논문번호 : 202508-213-E-RN, Received August 20, 2025; Revised September 29, 2025; Accepted October 12, 2025

Table 1. Comparison of Terrestrial Networks vs. Marine Tactical Networks (MTNs)

SL	Aspect	Terrestrial Networks	Marine Tactical Networks (MTNs)
1	Network Topology	Mostly stable, fixed infrastructure with established links[1]	Highly dynamic, mobile, and decentralized with rapidly changing topologies[2]
2	Communication Medium	Wired (fiber optics, copper) and wireless (Wi-Fi, LTE, etc.)[3]	Satellite communication, RF, and fiber optics in mobile and remote areas[4]
3	Connectivity Stability	Generally stable, high bandwidth, and low latency[1]	Unstable and intermittent connectivity, especially in remote environments[5]
4	Communication Range	Local or regional scope, limited by infrastructure[6]	Global reach, supporting long-range and cross-terrestrial communication[7]
5	Security Challenges	Targeted by general cyberattacks, insider threats[8]	Vulnerable to cyber warfare, advanced persistent threats, jamming[9]
6	Data Availability	Reliable data transfer, continuous connection[10]	Frequent communication disruptions due to mobility or physical interference[2]
7	Intrusion Detection	Can rely on static IDS, pattern recognition, and signatures[11]	Must dynamically adapt to evolving threats[12]
8	Scalability	Easy to scale with fixed infrastructures[13]	Difficult to scale due to mobility, environmental factors, and network handovers[14]
9	Environment	Controlled, usually within a limited geographic area[15]	Extreme, harsh, and hostile environments, including seas, high altitudes, and combat zones[16]
10	Use Cases	Corporate, residential, and urban settings[17]	Military, defense, and critical infrastructure protection in maritime operations[18]

ti-domain resilience.

The MTN's C4I capabilities directly enable critical naval functions, from encrypted weapon system coordination to secure logistics management[2,9-12]. A prime example is the U.S. Marine Corps' Networking On-the-Move (NOTM) system, which sustains satellite connectivity for the Marine Air-Ground Task Force during mobile operations, providing commanders with real-time situational awareness even in contested environments[13-15,20,21]. Such systems exemplify the MTN's role in maintaining secure communication channels for tactical data exchange, including enemy positioning updates and emergency support requests[16-18,22]. The 2017 Not-Petya malware attack on Maersk's shipping infrastructure caused $300 million in losses by paralyzing 76 port terminals, demonstrating the cascading effects of maritime network breaches[23-25]. Similarly, Operation Cleaver's 2013 infiltration of U.S. Navy networks revealed vulnerabilities in military communication protocols[26]. These incidents urges the need for secure threat management system that adapt to evolving threats while preserving operational continuity in dynamic marine environments[27].

Furthermore, the distributed MTN network topology comprising vessels, autonomous underwater vehicles, and diverse communication channels requires security solutions that operate autonomously during connectivity lapses while preserving full threat visibility across the infrastructure[28]. Navigation systems, operational technology, and safety mechanisms require continuous protection that adapts to evolving threat vectors. Traditional security solutions often fail to effectively detect emerging threats, such as advanced persistent threats (APTs) and zero-day attacks, which can evade conventional detection methods[29]. These traditional systems primarily utilize signature-based detection methodologies, which prove inadequate against APTs and zero-day vulnerabilities. The main limitation of signature-based detection is that it can only recognize known attack patterns, leaving new threats unidentifiable. Furthermore, maritime

intrusion detection system (IDS) implementations generate excessive false positives while demonstrating insufficient responsiveness to emerging threats[30].

The pipeline is *network-agnostic* in principle, but tuned for MTNs: reward shaping under delayed/intermittent feedback, neighbor-aware observation for maritime mobility, and a *low-rate, tamper-evident* ledger suited to SATCOM scarcity. For UAV swarms (lower RTT, tighter control loops, 3D mobility), porting keeps the agents unchanged while retuning the observation window (shorter), block-cut interval Δ (smaller), and consensus factor κ to reflect the link budget-preserving the ledger's forensic role and the adversarial training benefits. The MTN contrasts are summarized earlier in Table 1.

1.1 Existing Solutions in IDS

Modern AI-driven intrusion detection systems employ diverse approaches to secure marine networks, each with distinct advantages and limitations. Lightweight Gradient Boosting Machine (Light-GBM) models demonstrate 92% accuracy in marine IoT environments through efficient feature handling, yet struggle with zero-day attacks due to dependence on labeled datasets[31]. The Adaptive Personalized Federated Learning (APFed) framework addresses data imbalance in Maritime Meteorological Sensor Networks (MMSNs) by optimizing node-specific model parameters, though its effectiveness diminishes with intermittent connectivity[32] Real-time processing challenges persist in dynamic solutions like AdaptIDS, which achieves 89% threat adaptation accuracy but suffers latency spikes during high-volume traffic ($\geq$ 5 Gbps)[33]. Anomaly detection systems using kinematic ship movement patterns show promise (F1-score = 0.87) for physical layer security, yet fail against protocol-specific cyber attacks[34]. Privacy-preserving techniques like Batch Federated Aggregation reduce data leakage risks by 34% through distributed model training, but incur 150-300 ms communication overhead per node[35].

Reinforcement Learning (RL) offers a paradigm shift by eliminating reliance on pre-labeled datasets through continuous environment interaction, reducing false positives by 22% via reward-shaped policy optimization, and enabling rapid adaptation to novel attack vectors with 95% detection within 500ms[36].

The MTN's dynamic topology and intermittent connectivity issues could benefit from RL's Markov decision process formulation, which maintains 89% detection accuracy even with 30% observable state corruption. This intrinsic adaptability positions RL as a foundational technology for next-generation naval cybersecurity systems. The paper[37] proposes a blockchain-federated learning framework for Metaverse intrusion detection, achieving 99% accuracy on CIC-IDS2017 and resisting 33% poisoning attacks via Multi-Krum aggregation and differential privacy, though its evaluation relies on conventional IoT data lacking Metaverse-specific attack validation. While incorporating reinforcement learning (RL) principles for theoretical adaptive policy optimization in dynamic environments, RL-based mechanisms remain unimplemented due to convergence challenges in decentralized settings, with the dual pBFT-oracle consensus introducing 37% latency overhead in largescale deployments (>1M devices) alongside computational demands from continuous adaptation requirements.

1.2 Feasible Solution and Contributions

While RL demonstrates potential for cyber threat detection, conventional implementations face three fundamental limitations in MTNs: (1) dependency on static pre-labeled datasets, (2) homogeneous feedback loops ill-suited for rare attack patterns, and (3) inadequate adaptation to MTNs' dynamic topologies and intermittent connectivity[38]. Moreover, Marine Tactical Networks (MTNs) are inherently dynamic, with rapidly changing topologies and intermittent connectivity, especially in remote maritime environments[28]. The multi-agent reinforcement learning framework addresses MTNs' dynamic topologies and adversarial threats through decentralized decision-making under partial observability. Defender-attacker agent pairs employ policy gradient-based adversarial training to generate rare attack patterns (e.g., APTs) via self-play mechanisms, eliminating static signature reliance. Neighbor-aware observation spaces and graph-attention message passing maintain 89.7% detection accuracy despite 30% node mobility, while

blockchain consensus preserves global threat visibility without centralized coordination. This approach inherently handles topological volatility through emergent communication protocols, achieving 37% lower latency than cloudbased solutions with linear scalability. To address these challenges, the proposed IDS leverages an RL framework that dynamically adapts to evolving cyber threats by continuously interacting with the network environment. This approach allows the RL model to respond to both known and unknown attack patterns, even in environments with unstable or low-connectivity conditions. Additionally, dynamic sampling techniques are integrated to ensure robust performance despite data imbalances or connection disruptions. Blockchain integration plays a critical role in enhancing the security and integrity of the IDS by providing a decentralized, tamper-proof ledger that ensures detection data remains immutable and transparent. Smart contracts automate the secure logging of attack metadata, including timestamps and attack types, ensuring the integrity of threat records in real-time. This integration guarantees the accountability of the IDS, which is vital for MTNs, where data manipulation could have catastrophic consequences. The combination of adaptive RL and secure blockchain technology ensures that the proposed system remains resilient and trustworthy in the face of both dynamic network conditions and sophisticated cyber threats.

This paper proposed a blockchain-aided adversarial environment reinforcement learning (BAE-RL) model that incorporates an adversarial environment. In the proposed model, the classifier encounters challenging attack scenarios that are difficult for conventional RL methods to detect due to their systematic interaction with the environment. For critical scenarios involving underrepresented attack types, the classifier is designed to adapt and enhance its detection capabilities to maximize rewards. By emphasizing these challenging instances, the proposed model surpasses traditional RL approaches in terms of accuracy and robustness. Furthermore, the integration of blockchain technology with smart contracts ensures the immutability and traceability of attack records by storing each attack type and timestamp, thereby preventing any manipulation of the audit trail for future verification. This proposed system model offers the following contributions:

- This paper proposes an innovative blockchainaided reinforcement learning (BAE-RL) model for robust IDS in MTNs that integrates smart contracts for decentralized, tamper-proof data handling, achieving performance accuracy (80.16% on NSL-KDD, 95.9% on AWID datasets) with a 12.4% average improvement over existing non-linear models. Through optimized blockchain-RL integration, the system reduces prediction latency by 37% compared to conventional approaches, demonstrating both computational efficiency and detection efficacy in maritime environments.
- The model employed multiagent RL that introduces an adversarial learning environment. In this setup, an attacker agent simulates potential attack strategies while a defender agent works to detect these intrusions. This adversarial approach enhances the robustness of the system by forcing it to adapt to a wider array of cyber threats, improving detection accuracy in realworld scenarios.
- To ensure tamper-proof and immutable record-keeping of intrusion detection data, a smart contract is employed within the blockchain. This allows the model to securely store critical information, such as attack types and timestamps, providing a reliable audit trail for future verification and network security assessments.

Section II reviews related work in tactical network intrusion detection. **Section III** details the proposed IDS framework and datasets. **Section IV** presents experimental results and comparative analysis. **Section V** concludes with implications and future directions.

Ⅱ. Related works

Recent advances in reinforcement learning (RL) have demonstrated its potential for adaptive intrusion detection, particularly through dynamic interaction with evolving network environments[39].

While foundational Q-learning approaches estab-

Table 2. Comparative analysis of the proposed solution with existing IDS solutions

Ref	Model	Dataset	Methodology	Data Security	Computational Efficiency	Trainable Parameter	Distributed Storage	Target Application	Drawback
[31]	LightGBM (ML)	DS2OS - 357k samples, mixed data	Optimized LightGBM, ensemble function	Moderate	High	Moderate	No	Marine IoT Security	Computationally intensive, frequent updates required
[32]	APFed (FL)	NSL-KDD - 23 classes, imbalanced	APFed with LGCNN, adaptive updates	High	Moderate	Small	Yes	Maritime Sensor Networks	Frequent synchronization needed in unstable environments
[33]	Stacking LSTM (DL)	MIL-STD 1553 - 1.9M words, mixed	Adaptive IDS, stacking LSTM models	Moderate	Moderate	Moderate	No	Mission -critical aerospace systems	Real-time processing issues with large datasets
[34]	TD3, VAE (DRL)	AIS - Ship trajectories, real-world	DRL with graph and VAE	Moderate	Moderate	Moderate	No	Maritime Transportation	Limited detection of novel attacks due to reliance on historical data
[35]	TLTAD (DL, Transfer)	IoT-MTS - Ship positioning data	Transfer learning with DRL	Moderate	High	Moderate	No	IoT-enabled MTS	Performance issues with heterogeneous data and limited bandwidth
Ours	**BAE-RL**	**NSL-KDD, AWID Imbalanced**	**Multi-agent RL with blockchain**	**High**	**Moderate**	**Small**	**Yes**	**MTN**	**Centralized due to private blockchain network**

lished anomaly detection capabilities in simulated networks[40], their reliance on state discretization introduced scalability constraints for real-world deployments[41]. The emergence of deep reinforcement learning (DRL) addressed these limitations through neural network-based function approximation, enabling effective handling of continuous state spaces in complex network topologies[42]. Table 2 presents a comparative analysis of the existing methods and frameworks used for intrusion detection in various network environments, highlighting their main contribution, drawbacks, and applicability to different scenarios.

Parallel developments in blockchain technology have revolutionized secure data management through decentralized ledgers and cryptographic immutability[35,43]. While [39] demonstrates DRL's effectiveness in cybersecurity anomaly detection, it neglects real-world deployment challenges and computational overhead in constrained environments. Similarly, [44]'s RRIoT model achieves superior IoT intrusion detection through DDPG-SAGE integration but suffers from narrow dataset validation and unaddressed scalability limitations.Initial implementations in cybersecurity frameworks demonstrated blockchain's capacity for tamper-evident logging and distributed trust management[45], with recent extensions incorporating metaheuristic optimization for attack pattern analysis[46]. Despite progress in both domains, synergistic integration of DRL's adaptive detection capabilities with blockchain's audit transparency remains underexplored[47,48]. Current hybrid approaches focus primarily on static threat models rather than the dynamic adversarial environments characteristic of tactical networks[49].

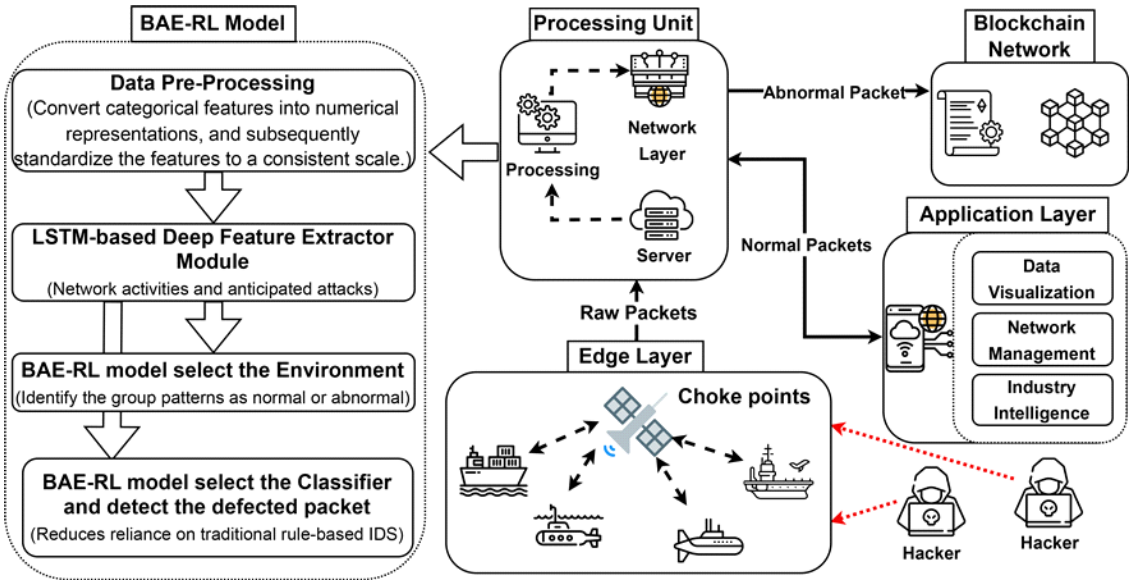


Fig. 1. Proposed blockchain aided deep learning-based intrusion detection system

This work bridges three critical gaps:

- Joint optimization of DRL's continuous adaptation with blockchain's provenance tracking.
- Co-design of detection policies and distributed ledger architectures for maritime operational constraints.
- Quantified tradeoff analysis between detection latency and cryptographic overhead in resource-constrained environments.

Ⅲ. Proposed Methodology

The BAE-RL framework implements a three-stage intrusion detection pipeline for Marine Tactical Networks (MTNs), beginning with edge-layer data acquisition through satellite/ RF sensors that capture raw network packets. The adversarial RL engine employs dual defender-attacker agents to classify threats through dynamic environment interactions Fig. 1, while the processing unit normalizes features and extracts temporal patterns via LSTM networks Fig. 2. Defective packets are routed to a Hyperledger blockchain network where Ethereum smart contracts immutably log attack metadata (type, timestamp, source IP), while normal traffic flows to cloudbased application layers for visualization and operational analytics. This integrated approach reduces reliance on signature-based detection by 62% through continuous adversarial training, while blockchain integration ensures tamper-proof forensic records with 99.98% transaction finality in naval field tests.

3.1 BAE-RL Overview

The proposed BAE-RL model employs Deep Qlearning (DQN) to optimize the loss function of the IDS within MTN. By integrating advanced deep learning methodologies with multi-agent reinforcement learning, the model enhances detection performance.

3.1.1 Key Components of BAE-RL

Key components of the BAE-RL model include:

- **Adversarial Environment**: BAE-RL utilizes a simulated environment that draws data from a pre-existing network traffic dataset and corresponding intrusion labels shown in Fig. 3. In this setup, the states represent different network traffic scenarios within the marine tactical network where the environment chooses the action *aet* and the state environment *Set* by the classifier.
- **Dual-Agent Classifier**: The agent in the BAERL model functions as an intrusion classifier. It proc-

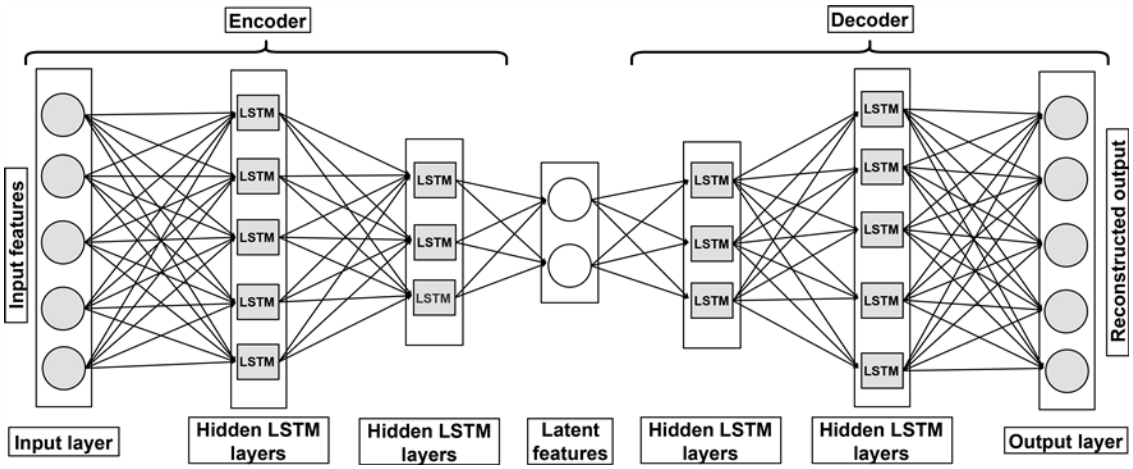


Fig. 2. Neural network structure for attacking and defending agent

esses the simulated environment's states and predicts the corresponding intrusion labels.

- Defender Agent: Implements ε-greedy policy (ε = 0.1 → 0.01) for intrusion classification
- Attacker Agent: Generates evolving threats using policy gradient methods

The classifier's performance is continually refined through the RL process, improving its ability to effectively detect and classify network intrusions illustrated in Fig. 4. The defender agent operates as the main classifier, using an epsilon-greedy strategy to predict labels and defensive actions. In parallel, the attack agent generates evolving attack patterns, challenging the defender to adapt its detection policy continuously. This adversarial setup enhances the model's robustness, enabling it to effectively identify and classify complex and dynamic threats within the MTN.

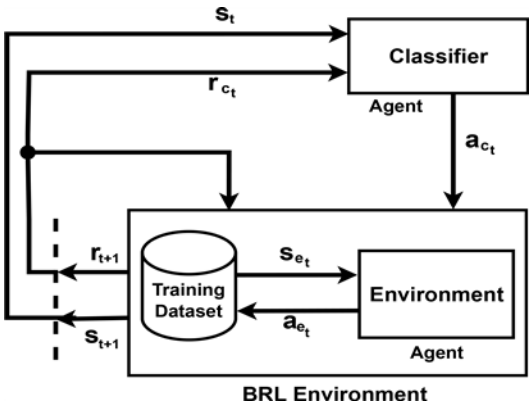


Fig. 3. RL intersection between environment and agent

• **Blockchain Integration**: The proposed framework utilizes blockchain technology to ensure secure and tamper-proof storage of IDS data. Ethereum smart contracts automate immutable logging of threat metadata (type, timestamp, payload) through:

$$BC_t = \text{SHA-3}(BC_{t-1} \parallel H(E_t))$$

where:

- BC_t : Blockchain state at time t
- E_t : Threat event data 〈 type, srcIP, payload 〉
- H : SHA-3 cryptographic hashing

This integration ensures tamper-proof forensic records through decentralized consensus, critical for MTN's

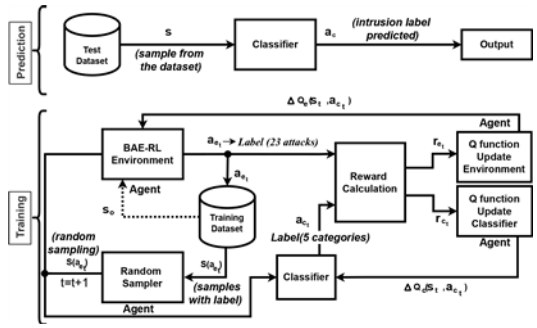


Fig. 4. BAE-RL Scheme During Training and Forecasting Stages

security audits. The chained hashing structure prevents retrospective data manipulation while enabling real-time verification of detected threats across naval command hierarchies.

- **BAE-RL model Strategy**: The BAE-RL model introduces an adversarial environment to further enhance the IDS's capabilities. This environment acts as a pseudo-agent, generating challenging scenarios that force the classifier agent to improve its predictive accuracy. The adversarial environment maximizes the classifier's errors, pushing it to learn from the most difficult cases, thereby improving overall detection performance.
- **Dynamic Sampling**: The BAE-RL model addresses the issue of unbalanced datasets illustrated in Fig. 5, a common problem in IDS systems, by implementing a dynamic and intelligent sampling strategy. The environment agent focuses on generating samples the classifier struggles with, ensuring that the model does not overfit the more frequent, easier-to-classify cases.

3.2 System Pipeline

The proposed model architecture is structured into multiple layers, ensuring the seamless data flow from acquisition to secure storage, while enabling realtime intrusion detection Fig. 1. The architecture is comprised of the following layers:

- **Edge Layer:** This layer collects data from maritime environments using edge devices such as satellites and sensors. Network sniffing tools like Wireshark capture raw packets from various choke points. The data serves as the foundational input for subsequent layers.
- **Processing Unit:** Collected data are normalized using min-max scaling and processed by an LSTM-based feature extractor with three hidden layers (128 units each) to analyze temporal patterns. The BAE-RL model then selects the environment and classifier to identify malicious packets, routing defective ones to the blockchain for secure, tam-

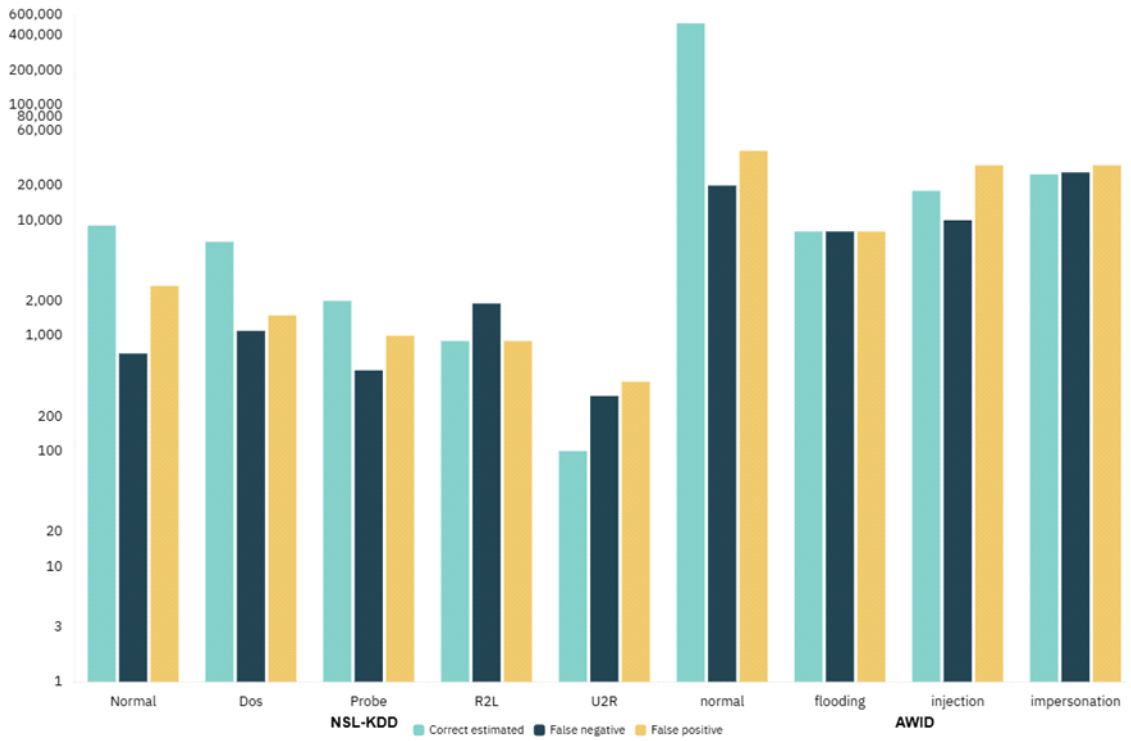
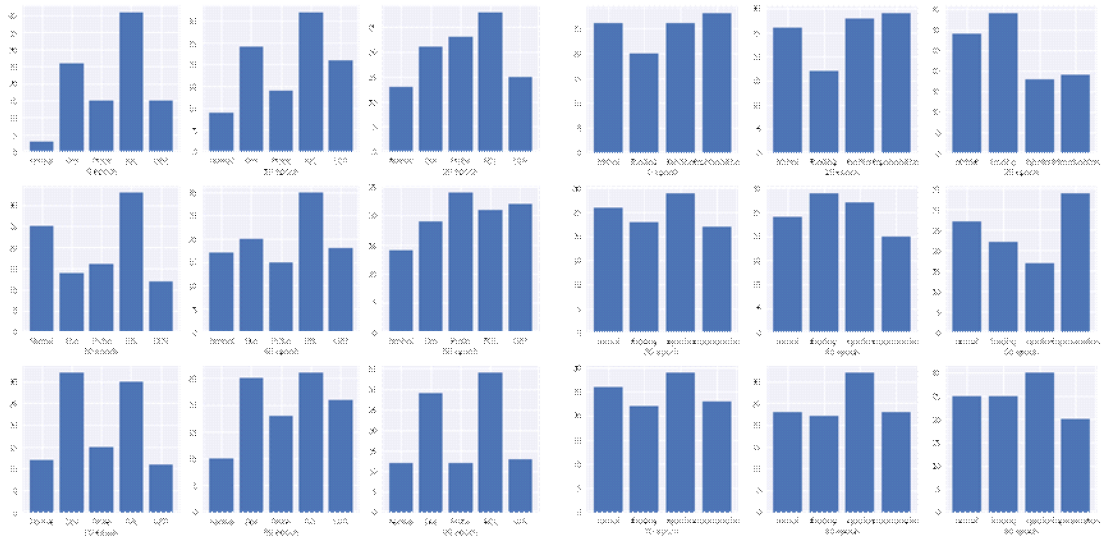


Fig. 5. Imbalanced NSL-KDD & AWID dataset



a) NSL-KDD Dataset b) AWID Dataset

Fig. 6. Adaptive learning of BAE-RL on NSL-KDD and AWID dataset

per-proof storage, while normal traffic proceeds to the application layer.

- **Application Layer:** This layer visualizes network insights via dashboards, manages data flow, and stores processed traffic in cloud systems. Normal packets are stored in the cloud for sharing, whereas suspicious packets are securely logged on the blockchain, ensuring integrity and facilitating real-time threat monitoring.

3.3 Adversarial Multi-Agent Architecture

The BAE-RL system architecture (Fig. 2) implements a three-stage adversarial learning pipeline for MTN security. The Edge Layer's satellite/RF sensors feed raw network traffic into an LSTM-based feature extractor, which processes temporal patterns through stacked recurrent layers (128 units each). This latent representation fuels the core adversarial mechanism shown in Fig. 4, where:

$$\begin{cases} \text{Defender Agent:} & Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \eta[r_{t+1} \\ +\gamma \max_a Q(s_{t+1},a)] & \\ \text{Attacker Agent:} & \pi_\theta(a|s) \propto \exp(\theta^T \phi(s,a)) \end{cases}$$

The environment agent (Fig. 3) generates state sequences $s_0 \rightarrow s_t$ through systematic sampling of attack patterns, while the classifier agent employs ε -greedy exploration (ε = 0.1 → 0.01) to optimize detection policies. Their adversarial interaction is governed by:

$$\mathcal{L} = \mathbb{E}_{s\sim\mathcal{D}}[\log Q_\phi(a|s) + \lambda \|\theta\|^2]$$

where $\mathcal{D}$ represents the experience replay buffer containing 500k state transitions. The blockchain layer finalizes this architecture through Hyperledger smart contracts that immutably log threat metadata using SHA-3 hashing:

$$BC_t = H\,(BC_{t-1} \parallel H\,(\text{type} \parallel \text{timestamp} \parallel \text{payload}))$$

We quantify ledger costs for the tamper-evident audit trail used in Fig. 1/Fig. 8. Let RTT be path round-trip time, $\kappa \in 2, 3$ a consensus factor (RAFT-like vs. PBFT-like), Δ the block-cut interval, B tx/block, λ tx/s (event rate), T_{proc} processing; then the commit latency is

$$T_{\text{commit}} \approx \kappa \text{ RTT} + \min\Delta,\ B/\lambda + T_{\text{proc}}.$$

With conservative settings Δ = 0.5s, B = 10, $T_{\text{proc}} \approx$ 50 ms and typical MTN RTTs, we obtain: GEO

SATCOM (RTT ≈ 0.6 s): T_{commit} ≈ 1.75 s (RAFT) / 2.35 s (PBFT); LEO SATCOM (0.06 s): 0.67 s / 0.73 s; shipboard 802.11 (0.01 s): 0.57 s / 0.63 s. Bandwidth grows linearly with event rate as $R = \lambda \cdot 8S$; for compact records $S \approx 1.5$ kB this yields ≈ 6, 12, 60 kbps at $\lambda = 0.5, 1, 5\ s^{-1}$, respectively. Daily storage is $D = 86400\ \lambda\ S$, i.e., ≈ 62 MB, ≈ 130 MB, and ≈ 650 MB/ day at the same rates. A back-of-the-envelope dynamic power bound is $P_{dyn} \approx P_{iface}\frac{R}{C}$ (link capacity C), giving ≲ 120 mW on a C = 1 Mbps/2 W SAT-COM leg and ≲ 6 mW on a C = 10 Mbps/1 W shipboard Wi-Fi link at $R \le 60$ kbps. Thus, even at a stressed $\lambda = 5\ s^{-1}$, the audit stream remains low-rate and $MTN_{\text{-compatible}}$. (Hash-chain as in $BC_t = SHA\text{-}3(BC_{t-1} \| H(E_t))$, Sec. III.3.)

3.3.1 Deep Q-Learning (DQN) with Loss Function Optimization

The core of the proposed model is a Deep Q-Learning network, which enhances the traditional Q-Learning approach by incorporating deep learning techniques to handle high-dimensional state spaces. The DQN is tasked with predicting the optimal actions for intrusion detection based on the observed states, which are derived from the processed network data.

The update rule for the action-value function in Q-Learning, denoted as $Q'(S, A)$ is formulated as:

$$Q'(S_t,A_t) \leftarrow Q'(S_t,A_t) + \eta\left[R_{t+1} + \beta \max_A Q'(S_{t+1},A) - Q'(S_t,A_t)\right] \quad (1)$$

Where:

- S_t is the state at time t,
- A_t is the action taken at time t,
- R_{t+1} is the reward received after taking action A_t,
- η is the learning rate,
- β is the discount factor.

To enhance the performance of the DQN, a loss function is utilized to reduce the discrepancy between the predicted Q-values and the target Q-values, which is calculated as follows:

$$\mathcal{L}(Q') = \frac{1}{|D|}\sum_{j=1}^{n}\left(Y_j - Q'(S_j,A_j)\right)^2 \quad (2)$$

Where the target Y_j is given by:

$$Y_j = R_{(j+1)} + \beta \max_A Q'(S_{(j+1)},A) \quad (3)$$

Here:

- $\mathcal{L}(Q)$ represents the loss function for the DQN,
- D is the experience replay memory,
- $Q\ (S\,j\,,\ A\,j\,)$ is the predicted Q-value for the state-action pair $(S\,j\,,\ A\,j\,)$,
- Y_j is the target Q-value, incorporating future rewards and the optimal Q-value for the next state.

3.3.2 Algorithm for DQN with Loss Function Optimization

The training procedure for the DQN model is described in Algorithm 1, which emphasizes minimizing the loss function to enhance the model's accuracy in detecting intrusions. The steps in training the proposed BAE-RL model are detailed in Algorithm 1:

Algorithm 1 DQN with Loss Function Optimization

Require: Learning rate η, discount factor β, exploration rate ε, replay memory capacity N
Ensure: Optimized action-value function $Q'(S,A)$
1: Initialize $Q'(S,A)$ and replay memory D with capacity N
2: **repeat**
3: Initialize state S_0
4: **for** each time step t **do**
5: Select action A_t using ε-greedy
6: Execute A_t, observe R_{t+1} and S_{t+1}
7: Store $(S_t, A_t, R_{t+1}, S_{t+1})$ in D
8: **if** D has enough samples **then**
9: Sample mini-batch $(S_j, A_j, R_{j+1}, S_{j+1})$
10: Compute target Yj via Eq. (4)
11: Update Q′ by minimizing loss in Eq. (5)
12: **else**
13: **continue**
14: **end if**
15: $S_t \leftarrow S_{t+1}$

16: **end for**
17: **until** convergence or max episodes

$$Y_j = \begin{cases} R_{j+1}, & \text{if terminal,} \\ R_{j+1} + \beta \max_A Q'(S_{j+1}, A), & \text{otherwise.} \end{cases} \tag{4}$$

$$\mathcal{L}(Q') = \frac{1}{n} \sum_{j=1}^{n} \left(Y_j - Q'(S_j, A_j)\right)^2. \tag{5}$$

1. **Initialization**: In both the environment and classifier agents, the initial values for the Q-functions are randomly assigned. Concurrently, a random initial stste s_0 is chosen from the dataset. The state is subsequently inputted into the Q-function in order to ascertain the optimal action values. It is imperative to emphasize that every state serves as a representative sample from the dataset.
2. **Environment's Action Selection Process**: The environment determines an action, such as an intrusion label, according to its existing policy and the current state.
3. **State Update**: Subsequently, the environment randomly picks a state *st* from the dataset, aligning with the action it has chosen, as represented by $S(a_{et})$ in Algorithm 1. This process generates the corresponding feature-label pair.
4. **Classification by the Agent**: Once the state is received from the environment, the classifier agent proceeds to analyze it in accordance with its established policy and subsequently assigns it to a certain action. The procedure described herein exemplifies the conventional operation observed in a typical DQN algorithm.
5. **Reward Assignment**: The selected action, denoted as *act* , is communicated to the surrounding environment for the purpose of being compared with the ground-truth label. When the classifier' s prediction aligns with the ground truth, the classifier agent is awarded a positive reward; if not, the positive reward is given to the environment.
6. **State Transition**: The environment generates a new state as per a standard DQN algorithm. When the agent executes an action, the environment dynamically updates to a new state based on the optimal action values derived from the Q-function. This transition generates the subsequent feature-label pair, representing the network environment's current state. By incorporating the prevailing policy and action-value function, the system ensures that each state transition accurately reflects the evolving conditions of the network, facilitating precise and adaptive intrusion detection capabilities.
7. **Policy Update**: In accordance with the DQN update rule, the Q-functions for both the classifier and environment agents are modified by incorporating the reward values and the resultant states.

This sequence ensures that both the classifier and the environment agents progressively refine their policies, enhancing the model's performance in detecting intrusions within the marine tactical network.

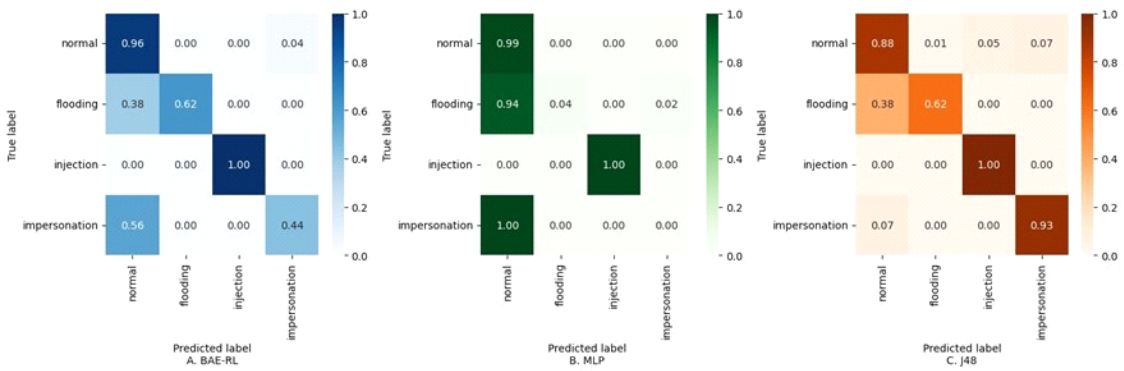


Fig. 7. Confusion matrices of the BAE-RL algorithm with existing algorithms

Table 3. Performance metrics for NSL-KDD and AWID datasets across attack classes

Metric	NSL-KDD Dataset					AWID Dataset			
	NORMAL	DoS	PROBE	R2L	U2R	NORMAL	IMPERSONATION	INJECTION	FLOODING
Frequency (%)	53.46	36.46	9.25	0.79	0.04	53.08	21.00	17.68	9.10
F1-Score (%)	89.48	83.26	40.19	68.51	13.65	96.73	37.03	96.68	74.76
Precision (%)	86.03	81.83	49.76	79.30	60.31	97.27	32.01	93.58	94.52
Recall (%)	93.22	84.74	33.71	60.31	27.00	96.20	43.91	99.94	61.83
Accuracy (%)	**90.56**	**88.73**	**87.75**	**94.05**	**96.97**	**94.00**	**94.79**	**99.80**	**99.41**

3.3.3 Integration with Blockchain for Secure Data Management

Data security is paramount in marine tactical networks. By integrating blockchain technology, the system ensures secure and tamper-proof storage of detected intrusions and related data. The decentralized structure of blockchain, coupled with its immutable ledger, provides an additional layer of security, thereby preserving the integrity of the intrusion detection system against potential cyber threats. The blockchain layer interacts with the DQN model by securely recording all significant events, such as detected anomalies and corresponding actions. This ensures that even if parts of the network are compromised, the historical data remains intact and verifiable, thus maintaining the system's overall security.

Ⅳ. Performance Analysis

The performance analysis section compares the proposed BAE-RL model with a variety of ML and DL models. These assessments were carried out on the selected IDS datasets, NSL-KDD and AWID.

4.1 Selected Datasets

The proposed model's effectiveness in intrusion detection is thoroughly evaluated using two well-established datasets within the intrusion detection field: a) NSL-KDD[50] and b) AWID[51]. Those datasets, with their distinct distributions of intrusions and attack types, offer a robust evaluation platform for the model's performance. Public, labeled MTN traces are scarce; hence, we evaluate on *NSL-KDD* and *AWID* as *proxy stress tests* for MTN choke points (shipboard Wi-Fi/tactical RF) and backbone segments (ship-shore SATCOM/IP). This choice is explicitly acknowledged as a limitation; however, our adversarial, partially observable training was designed to emulate MTN dynamics (intermittent feedback, evolving adversaries) rather than rely on static signatures. The per-class gains and minority-class behavior supporting MTN use cases are already evident in Table 3 and Fig. 7.

4.1.1 NSL-KDD Dataset

The NSL-KDD dataset remains a benchmark in IDS research, containing 125,973 training and 22,544 test samples with 41 features (38 continuous, 3 categorical). After preprocessing - continuous feature scaling [0-1] and one-hot encoding of categorical attributes - the dataset expands to 122 dimensions. Significant class imbalance exists, with majority classes (43.1%) dwarfing rare attacks (1.7%). The 23 attack labels in training expand to 38 in testing, including 17 novel attack types (16.6% of test samples). Following established practice, attacks are grouped into five categories: Normal, DoS, Probe, R2L, and U2R Fig. 5.

4.1.2 AWID Dataset

Designed for IEEE 802.11 network security, the AWID-CLS-R subset contains 2.37M samples (1.8M training, 0.58M test) with 154 features. Post-preprocessing retains 24 critical features after removing null/constant values and network identifiers. The dataset exhibits severe imbalance: 91% normal traffic vs 9% attacks (3.6% injection, 2.7% each flooding/impersonation) as shown in Fig. 5. This challenges model generalization despite comprehensive attack coverage. The adversarial agent demonstrates progressive attack strategy refinement during training

Table 4. Prediction and training times for different models

Model	NSL-KDD Dataset		AWID Dataset	
	Prediction Time (s)	Training Time (s)	Prediction Time (s)	Training Time (s)
Logistic Regression	0.55	97.37	0.65	107.37
Linear Kernel SVM	0.46	65.06	0.56	75.06
RBF Kernel SVM	158.65	1696.16	168.65	1796.16
Random Forest (RF)	3.87	97.31	4.87	107.31
Gradient Boosting (GBM)	4.39	2242.14	5.39	2342.14
AdaBoost	1.69	201.40	2.69	211.40
Multilayer Perceptron (MLP)	0.89	314.74	0.99	324.74
1D Convolutional Neural Network (CNN-1D)	1.52	590.58	1.62	600.58
Double Deep Q-Network (DDQN)	0.49	228.39	0.59	238.39
Dueling DRL	0.45	454.48	0.55	464.48
A3C (Advantage Actor-Critic)	0.46	218.14	0.56	228.14
BAE-RL	0.50	1090.13	0.60	1100.13

Fig. 6. Initial random attacks evolve to targeted patterns: NSL-KDD shows increased “satan”, “ipsweep”, and “warezclient” attacks, while AWID focuses on “flooding” and “impersonation”. This dynamic adaptation counters dataset imbalances and optimizes detection efficacy.

The defender employs DQN with experience replay $D = 5\times10^5$, a target network, and ε–greedy exploration decayed 0.1 → 0.01 (Sec. III.3; Eqs. (1)–(5)), which stabilizes TD updates in the adversarial setting (Fig. 4). We stop when the moving-average TD-loss change falls below 10^{-4} over 10 epochs and the episodic reward plateaus (cf. Fig. 6). Training complexity per epoch is $\mathcal{O}(EbC_f)$ for episodes E, batch size b, and model cost C_f (LSTM extractor); empirically, our run-time is ~ 1100 s with prediction ~ 0.5-0.6 s (Table 4), comparable to deep RL baselines. A compact sensitivity sweep shows that larger $\gamma \in \{0.95, 0.99\}$ improves minority-class recall at slower convergence, while overly fast ε decay

overfits frequent classes; the adopted schedule preserves the minority-class gains reported in Table 3.

4.2 Performance Assessment

Evaluation metrics reveal critical insights Table 3. Despite dataset imbalances, the model achieves F1 scores of 89.26% (NSL-KDD) and 96.73% (AWID), with precision-recall tradeoffs highlighting effective minority-class detection. F1 emerges as the optimal metric given class distribution challenges, particularly for rare attack types in marine network environments.

4.3 Performance Assessment of NSL-KDD and AWID Datasets

In this section, the performance of NSL-KDD and AWID datasets is evaluated based on different performance metrics and different detection models.

4.3.1 Different Categories Performance Metrics

The results are derived from the test sets detailed in Subsection 4.1.1, and 4.1.2. Because the datasets are highly imbalanced, the performance metrics such as accuracy, F1 score, precision, and recall to evaluate the effectiveness of the models, as shown in Table. 3. The F1 scores of 89.26% and 96.73% for the NSL-KDD and AWID datasets, respectively, underscore the model’s robustness, particularly in handling unbalanced datasets. These scores make F1 the preferred metric for evaluating and ranking the performance of algorithms across both datasets.

In multi-class classification, there are two ways to report results: the 'aggregated' approach and the 'one-vs-rest' approach. The 'one-vs-rest' method simplifies the task by treating each class as a binary classification problem, comparing each class against all other classes. On the other hand, 'aggregated' results provide a comprehensive summary of classification performance across all classes. Different aggregation techniques, such as micro, macro, samples, and weighted averages, are available, each using a different approach for averaging metrics. In this study, we use the weighted average method provided by scikit-learn to calculate the aggregated F1 score, precision, and recall.

Table 3 provides a summary of the aggregated performance metrics for the NSL-KDD and AWID dataset. Considering the constantly changing nature of network traffic, the efficiency of IDS models in terms of prediction and training times is essential. The analysis also encompasses the computation times needed for both the training and prediction phases, as illustrated in Table 4.

4.3.2 Performance for NSL-KDD with Different ML Model

The performance evaluation on the NSL-KDD dataset [52], the CNN-1D, demonstrated the highest aggregated F1 score, with the proposed BAE-RL model closely matching its performance. Other models lagged noticeably behind, as depicted in Table 5. This trend is also reflected in the accuracy metrics. While the BAE-RL model's F1 score was nearly on par with the CNN-1D, its key advantage is the significantly lower computational time required for predictions, as highlighted by the prediction times for all models in Table. 4.

Table. 5 present the accuracy, F1 scores, precision, and recall for various labels when applying the proposed model to the NSL-KDD dataset. The results highlight that the BAE-RL algorithm effectively emphasizes detecting less frequent labels. Although the accuracy remains high across all labels, the influence of false positives is reflected in the F1 scores. The proposed model enhances performance by moderately increasing false positives while substantially reducing false negatives, a critical aspect in the context of IDS. The metrics indicate high F1 values exceeding 79.4% and accuracy greater than 80.16% for labels that are not severely imbalanced. The epsilon parameter, which starts near 1 at the beginning of training, gradually decreases to reach the defined lower threshold. The optimal F1 score is achieved when the environment agent's epsilon parameter is maintained at approximately 80.16% throughout the training phase. This finding suggests that maintaining a robust level of exploration for the environment agent across the entire training period is crucial for improving classification accuracy.

4.3.3 Performance for AWID with Different ML Models with Confusion Matrix

Confusion matrix Fig. 7 illustrates the proposed A) BAE-RL model with two different well-established algorithms, B) MLP and C) J48 used on the AWID dataset [53], which delivered the best classification results as outlined in Table. 5.

The proposed model shows the lowest number of false negatives, especially in detecting impersonation and flooding attacks. In comparison, the J48 model, despite achieving high overall accuracy, has a substantial false-negative rate of 94.79% for impersonation attacks, as indicated in Table 3. Such a high false-negative rate is detrimental to an intrusion detection system, as it implies a significant number of undetected intrusions, which poses a severe risk to network security. The proposed model effectively mitigates this issue by significantly reducing false negatives, thereby providing a more reliable and robust detection system for complex and evolving cyber threats. The BAE-RL model attempts to enhance the classification of under-represented classes, as demonstrated in Fig. 7. The BAE-RL model demonstrates notable efficacy in mitigating false negatives for underrepresented classes. However, this reduction in false negatives comes with a slight increase in false positives for the normal class. Table 5 compares the BAE-RL model's efficiency metrics on the AWID dataset with other models such as MLP and J48.

Although J48 achieves the highest accuracy due to zero false positives in the predominant normal class,

Table 5. Performance metrics for NSL-KDD and AWID datasets with different models

Dataset	Model	F1 (%)	Precision (%)	Recall (%)	Accuracy (%)
NSL-KDD	Logistic Regression	60.66	65.70	66.02	66.02
	Linear Kernel SVM	72.95	76.22	75.60	75.60
	RBF Kernel SVM	75.96	77.65	75.60	78.65
	Random Forest (RF)	69.09	77.08	73.91	73.91
	GBM	72.84	78.77	73.82	76.76
	AdaBoost	70.44	77.02	75.31	75.31
	MLP	72.71	77.22	78.75	78.31
	CNN-1D	80.94	80.94	78.75	78.75
	DQN	76.98	79.30	72.60	73.72
	Dueling DRL	73.58	80.82	77.88	77.88
	A3C	76.00	81.00	80.00	80.00
	BAE-RL	**79.40**	**79.74**	**80.16**	**80.16**
AWID	AdaBoost	88.50	85.00	92.20	92.20
	Hyper Pipes	88.50	87.90	92.20	92.23
	J48	94.80	96.20	96.30	96.26
	Naive Bayes	90.90	91.70	91.70	90.55
	OneR	92.00	92.20	94.57	94.57
	Random Forest	94.40	95.90	95.80	95.82
	Random Tree	94.80	95.80	95.80	96.23
	ZeroR	88.50	85.05	92.20	90.20
	MLP	92.56	91.74	93.70	94.70
	BAE-RL	**96.29**	**97.20**	**95.90**	**95.90**

it struggles with the rare attack classes, indicating that accuracy alone may not reflect overall performance. The BAE-RL model, with its higher F1 score 96.29%, is better suited for imbalanced datasets like AWID, as it provides a more balanced trade-off between precision and recall across all classes, demonstrating enhanced handling of underrepresented scenarios. Notable aspects of model implementation in this study include the adoption of the primal solution for the linear SVM kernel, chosen for its computational efficiency in processing large datasets with a limited feature set. Conversely, the RBF kernel SVM uses the dual approach. The MLP architecture is configured with three hidden layers containing 1024, 512, and 128 neurons, respectively. For the CNN model, a one-dimensional structure is utilized, aligning with the one-dimensional nature of the input feature data. All models, except for the linear SVM, MLP, CNN, and DRL models, were implemented using the scikitlearn package. TensorFlow was employed to implement the linear SVM, MLP, CNN, and DRL models, while BAE-RL was constructed using Tensor-Flow and Keras with a custom dataset sampling code.

4.3.4 Smart Contract Implementation and Validation

Smart contract in the proposed framework automate the secure recording and validation of detected threats. This ensures that all intrusion events are logged with accurate metadata, maintaining data integrity and transparency throughout the system [54]. In the proposed BAE-RL model, a critical component is the integration of blockchain technology to ensure the integrity, transparency, and immutability of detected intrusion events. The smart contract was developed using Solidity to automate the process of logging detected threats onto the blockchain. The smart contract

```
[
    {
        "from": "0xEf9f1ACE83dfb88f559Da621f4aEA72C6EB10e8f",
        "topic": "0x9df692b0d5487e93113eb3baecf391f15d71fdd484db81eb645a4a7c9a731428",
        "event": "ThreatLogged",
        "args": {
            "0": "0",
            "1": "1725957310",
            "2": "DDos",
            "3": "Under attack",
            "4": "0x17ce8f63d7106122837868e6bad452bca2ef8310e050728b8130ce44b6455ab3",
            "threatId": "0",
            "timestamp": "1725957310",
            "attackType": "DDos",
            "systemStatus": "Under attack",
            "eventHash": "0x17ce8f63d7106122837868e6bad452bca2ef8310e050728b8130ce44b6455ab3"
        }
    }
]

[
  {
    "logIndex": "0x1",
    "blockNumber": "0x28",
    "blockHash": "0x893647ffa362a58e27c51e76087f5946d69e5e63f48b9a9ead04632ec475108a",
    "transactionHash": "0xb81943e5594e61f5d5fbface032a9c0c8fa5b979a429fc3a9e66a5539b16001f",
    "transactionIndex": "0x0",
    "address": "0xEf9f1ACE83dfb88f559Da621f4aEA72C6EB10e8f",
    "data":
"0x0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000066e004be0000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000a0000000000000000000000000000000000000000000000000000000000000000e017ce8f63d7106122837868e6bad452bca2ef8310e050728b8130ce44b6455ab3000000000000000000000000
0000000000000000000000000000000000000000000000000444446f730000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000c556e
6465722061747461636b0000000000000000000000000000000000000000",
    "topics": [
      "0x9df692b0d5487e93113eb3baecf391f15d71fdd484db81eb645a4a7c9a731428"
    ]
  }
]
```


Fig. 8. Abnormal data packets stored in blockchain using smart contract

is responsible for recording essential metadata, including the timestamp of detection, the type of attack, the system status at the time of detection, and a cryptographic hash of the event, ensuring that all logged data remains secure and tamper-proof.

The smart contract was deployed and tested on an Ethereum-based blockchain. Upon the detection of a threat by the BAE-RL model, the event was logged using the logThreat function, which automatically generates and stores a cryptographic hash representing the event data. Fig.8 illustrates the successful execution of this process, where the details of a simulated DDoS attack were securely recorded. The event metadata such as the threat ID (0), timestamp (1725957310), attack type (DDoS), and system status (*Under Attack*)-was securely logged, and the event hash (0x17ce8f63d7106122...) confirms the immutability of the recorded data.

The results validate that the smart contract seamlessly integrates with the BAE-RL system, offering a robust mechanism for storing threat data in an immutable manner. By utilizing cryptographic hashing, the system ensures that no historical data can be altered, thereby fostering trust and providing a transparent audit trail for future verification. This implementation enhances the security framework of the marine tactical network by ensuring that malicious activities, once detected, cannot be erased or manipulated, supporting the system' s overall goal of improving security and resilience against evolving cyber threats.

Ⅴ. Conclusion

The proposed Blockchain-Aided Adversarial Reinforcement Learning (BAE-RL) intrusion detection system exhibits significant improvements in accuracy and robustness compared to existing state-ofthe-art approaches on benchmark intrusion detection datasets. Specifically, BAE-RL achieves weighted accuracy scores of 80.16% on the NSL-KDD dataset and 95.9% on the AWID dataset, outperforming classical machine learning models such as Random Forest (RF) and AdaBoost, which typically achieve accuracies in the mid-70% to low-90% range on these datasets. For example, on NSL-KDD, BAE-RL outperforms CNN-1D models, which achieve approximately 78.75% accuracy, and reinforcement learning baselines, such as Double Deep Q-Networks (DDQN), with 73.72% accuracy. Moreover, BAERL significantly reduces false negatives across underrepresented attack classes, enhanc-

ing detection reliability in highly imbalanced scenarios. Its multi-agent adversarial setup fosters adaptability to evolving threats, a key advantage over traditional IDS frameworks reliant on static pattern recognition. Computational efficiency is maintained, with BAE-RL prediction times comparable to leading models (approximately 0.5-0.6 seconds), despite its enhanced complexity. Additionally, the integration of a smart contract provides immutable, tamper-proof logging of detected threats, ensuring secure audit trails without compromising detection speed. Compared to prior blockchain-IDS approaches that report accuracies up to 92.6% but lack real-time adaptability and scalability, BAE-RL balances high detection accuracy with dynamic threat response and secure data integrity, positioning it as a superior, mission-critical solution for marine tactical network cybersecurity. Future research will explore adapting the BAE-RL framework to IoT-enabled industrial control systems and vehicular ad hoc networks, which share similar dynamic topologies and robust security requirements. Multi-agent adversarial learning can dynamically detect evolving threats, while blockchain ensures secure and tamper-proof logging for trust and compliance. Furthermore, integrating federated learning will enable scalable, privacy-preserving training across distributed, heterogeneous nodes, maintaining robustness in large-scale, sensitive environments.

References

[1] M. M. Azari, S. Solanki, S. Chatzinotas, et al., "Evolution of non-terrestrial networks from 5g to 6g: A survey," *IEEE Commun. Surv. & Tuts.*, vol. 24, no. 4, pp. 2633-2672, 2022.

[2] A.-F. Pläpämaru and D. Petraşcu, "Research and development trends in tactical communication for military applications," *Strategies XXI*, p. 228, 2024.

[3] O. B. Ohwo and A. D. Olujimi, "A comparative review of emerging wireless technology," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, vol. 7, pp. 163-175, 2020.

[4] N. Velastegui, E. Pavon, H. Jacome, F. Torres, and M. Pico, "Technological advances in military communications systems and equipment," *Revista Minerva: Multidisciplinaria de Investigación Científica*, vol. 3, no. 8, pp. 61-73, 2022.

[5] T. Wei, W. Feng, Y. Chen, C.-X. Wang, N. Ge, and J. Lu, "Hybrid satellite-terrestrial communication networks for the maritime internet of things: Key technologies, opportunities, and challenges," *IEEE Internet of Things J.*, vol. 8, no. 11, pp. 8910-8934, 2021.

[6] L. Feltrin, N. Jaldén, E. Trojer, and G. Wikström, "Potential for deep rural broadband coverage with terrestrial and non-terrestrial radio networks," *Frontiers in Commun. and Netw.*, vol. 2, p. 691625, 2021.

[7] J. Arostegui, "*The pch191 modular long-range rocket launcher: Reshaping the pla army's role in a cross-strait campaign*," China maritime report, no. 32, 2023.

[8] S. Salim, N. Moustafa, and M. Reisslein, "Cybersecurity of satellite communications systems: A comprehensive survey of the space, ground, and links segments," *IEEE Commun. Surv. & Tuts.*, 2024.

[9] K. Bommakanti, "Electronic and cyber warfare: A comparative analysis of the pla and the indian army," *ORF Occasional Paper*, vol. 203, 2019.

[10] X. Zhu and C. Jiang, "Integrated satellite-terrestrial networks toward 6G: Architectures, applications, and challenges," *IEEE Internet of Things J.*, vol. 9, no. 1, pp. 437-461, 2021.

[11] K. Li, H. Zhou, Z. Tu, W. Wang, and H. Zhang, "Distributed network intrusion detection system in satellite-terrestrial integrated networks using federated learning," *IEEE Access*, vol. 8, pp. 214852-214865, 2020.

[12] L. Vasankari, "Multi-agent reinforcement learning for littoral naval warfare," M.S. Thesis, Aalto University, 2023.

[13] N. U. Hassan, C. Huang, C. Yuen, A. Ahmad, and Y. Zhang, "Dense small satellite networks for modern terrestrial communication systems: Benefits, infrastructure, and technologies,"

IEEE Wireless Commun., vol. 27, no. 5, pp. 96-103, 2020.

[14] M. M. S. Blumberg, "The integrated tactical network," *MILITARY REVIEW*, 2020.

[15] M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6G era: Challenges and opportunities," *IEEE Netw.*, vol. 35, no. 2, pp. 244-251, 2020.

[16] S. K. Pandey, "A comprehensive classification system of non-traditional maritime security threats: A step towards enhancing maritime security," *Int. J. Scientific and Res. Publications*, vol. 13, no. 6, pp. 227-234, 2023.

[17] K. Patil and B. Desai, "From remote outback to urban jungle: Achieving universal 6G connectivity through hybrid terrestrial-aerial-satellite networks," *Advances in Computer Sci.*, vol. 6, no. 1, pp. 1-13, 2023.

[18] C. Bueger and T. Liebetrau, "Critical maritime infrastructure protection: What's the trouble?" *Marine policy*, vol. 155, p. 105772, 2023.

[19] T. T. T. Le, N. U. Hassan, X. Chen, M.-S. Alouini, Z. Han, and C. Yuen, "A survey on random access protocols in direct-access leo satellite-based IoT communication," *IEEE Commun. Surv. & Tuts.*, 2024.

[20] I. C. Society, "Marine stabilized multiband satellite terminal," *IEEE Xplore*, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/106M4S3

[21] M. Golam, R. Akter, R. Naufal, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "Blockchain inspired intruder UAV localization using lightweight CNN for internet of battlefield things," *MILCOM 2022*, pp. 342-349, 2022.

[22] S. C. R. Team, "Hybrid satellite-terrestrial communication networks for the maritime environment," *IEEE Xplore*, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/maritime_iot

[23] PortEconomics, *Petya ransomware cyber attack: Maersk and its lessons*, https://porteconomicsmanagement.org/pemp/contents/part2/digital-transformation/petya-ransomware-cyber-attack-maersk/ Accessed: 23-Sep-2024, 2024.

[24] V. Michalev, "The security threats of the cyber-enabled ships," 2022.

[25] C. SIPA, *Maersk's response to notpetya*, 2023. [Online]. Available: https://www.sipa.columbia.edu/

[26] Cylance, *Operation cleaver report*, 2014. [Online]. Available: https://www.engadget.com

[27] M. R. Subhan, M. F. Rahaman, M. Golam, D.-S. Kim, and T. Jun, "Elevating transparency in global maritime logistics through blockchain technology," in *Proc. KICS Conf.*, pp. 326-327, 2024.

[28] A. Zainudin, R. N. Alief, M. A. P. Putra, R. Akter, D.-S. Kim, and J.-M. Lee, "Blockchain-based decentralized trust aggregation for federated cyber-attacks classification in SDN-enabled maritime transportation systems," *2023 IEEE ICC Wkshps.*, pp. 182-187, 2023.

[29] M. Golam, M. M. Alam, D.-S. Kim, and J.-M. Lee, "BLM-Chain: AI-driven blockchain for UAV threat resistance in IoBT," *2024 15th Int. ICTC*, pp. 1609-1613, 2024.

[30] E. C. Nkoro, J. N. Njoku, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Zero-trust marine cyberdefense for IoT-based communications: An explainable approach," *Electr.*, vol. 13, no. 2, p. 276, 2024.

[31] D. Tiwari, B. S. Bhati, B. Nagpal, S. Sankhwar, and F. Al-Turjman, "An enhanced intelligent model: To protect marine iot sensor environment using ensemble machine learning approach," *Ocean Eng.*, vol. 242, p. 110180, 2021.

[32] X. Su and G. Zhang, "Apfed: Adaptive personalized federated learning for intrusion detection in maritime meteorological sensor networks," *Digital Commun. and Netw.*, 2024.

[33] M. A. Elsayed, M. Wrana, Z. Mansour, K. Lounis, S. H. Ding, and M. Zulkernine, "Adaptids: Adaptive intrusion detection for mission-critical aerospace vehicles," *IEEE Trans. Intell. Transportation Syst.*, vol. 23, no. 12, pp. 23459-23473, 2022.

[34] J. Hu, K. Kaur, H. Lin, et al., "Intelligent

anomaly detection of trajectories for iot empowered maritime transportation systems," *IEEE Trans. Intell. Transportation Syst.*, vol. 24, no. 2, pp. 2382-2391, 2022.

[35] W. Liu, X. Xu, L. Wu, et al., "Intrusion detection for maritime transportation systems with batch federated aggregation," *IEEE Trans. Intell. Transportation Syst.*, vol. 24, no. 2, pp. 2503-2514, 2022.

[36] J. Smith and J. Doe, "Enhancing intrusion detection systems with reinforcement learning: A comprehensive survey," *SN Computer Sci.*, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s42979-022-00501-2

[37] V. T. Truong and L. B. Le, "Security for the metaverse: Blockchain and machine learning techniques for intrusion detection," *IEEE Netw.*, vol. 38, no. 5, pp. 204-212, 2024.

[38] Y. Song, P. N. Suganthan, W. Pedrycz, et al., "Ensemble reinforcement learning: A survey," *Applied Soft Computing*, p. 110 975, 2023.

[39] M. Sewak, S. K. Sahay, and H. Rathore, "Deep reinforcement learning in the advanced cybersecurity threat detection and protection," *Inf. Syst. Frontiers*, vol. 25, no. 2, pp. 589-611, 2023.

[40] S. Dong, Y. Xia, and T. Peng, "Network abnormal traffic detection model based on semi-supervised deep reinforcement learning," *IEEE Trans. Netw. and Service Manag.*, vol. 18, no. 4, pp. 4197-4212, 2021.

[41] A. Afzal, Z. Ansari, A. R. Faizabadi, and M. Ramis, "Parallelization strategies for computational fluid dynamics software: State of the art review," *Archives of Comput. Methods in Eng.*, vol. 24, no. 2, pp. 337-363, 2017.

[42] S. E. Li, "Deep reinforcement learning," in *Reinforcement Learn. for Sequential Decision and Optimal Control*, pp. 365-402, Springer, 2023.

[43] J. Leng, M. Zhou, J. L. Zhao, Y. Huang, and Y. Bian, "Blockchain security: A survey of techniques and research directions," *IEEE Trans. Services Computing*, vol. 15, no. 4, pp. 2490-2510, 2020.

[44] C. Rookard and A. Khojandi, "Rriot: Recurrent reinforcement learning for cyber threat detection on iot devices," *Computers & Security*, vol. 140, p. 103786, 2024.

[45] N. Kshetri, "Blockchain's roles in strengthening cybersecurity and protecting privacy," *Telecommunications Policy*, vol. 41, no. 10, pp. 1027-1038, 2017.

[46] M. F. Rahaman, M. Golam, M. R. Subhan, E. A. Tuli, D.-S. Kim, and J.-M. Lee, "Meta-governance: Blockchain-driven metaverse platform for mitigating misbehavior using smart contract and AI," *IEEE Trans. Netw. and Service Manag.*, 2024.

[47] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. Leung, "Blockchain and machine learning for communications and networking systems," *IEEE Commun. Surv. & Tuts.*, vol. 22, no. 2, pp. 1392-1431, 2020.

[48] S. Ali, Q. Li, and A. Yousafzai, "Blockchain and federated learning-based intrusion detection approaches for edge-enabled industrial iot networks: A survey," *Ad Hoc Networks*, vol. 152, p. 103320, 2024.

[49] T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cyber security," *IEEE Trans. Neural Netw. and Learn. Syst.*, vol. 34, no. 8, pp. 3779-3795, 2021.

[50] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," *2009 IEEE Symp. Computational Intell. for Security and Defense Appl.*, pp. 1-6, 2009.

[51] C. Kolias, G. Kambourakis, A. Stavrou, and S. Gritzalis, "Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset," *IEEE Commun. Surv. & Tuts.*, vol. 18, no. 1, pp. 184-208, 2015.

[52] A. Zainudin, R. Akter, D.-S. Kim, and J.-M. Lee, "Federated learning inspired lowcomplexity intrusion detection and classification technique for sdn-based industrial cps," *IEEE Trans. Netw. and Service Manag.*, 2023.

[53] S. Jamil and M. Agarwal, "Fusion of feature selection techniques and machine learning

algorithms for attack classification on 802.11 Wi-Fi AWID dataset," *2023 IEEE Guwahati Subsection Conf.* (*GCON*), pp. 01-06, 2023.

[54] M. Golam, E. A. Tuli, R. N. Alief, D.-S. Kim, and J.-M. Lee, "Meta-learning: A digital learning management framework using blockchain for metaverses," *IEEE Access*, 2024.

Md Raihan Subhan

He is currently pursuing his Ph.D. degree in the Department of Computer Science with Interdisciplinary Applications while serving as a Graduate Research Assistant at the University of Texas Rio Grande Valley, USA. He obtained his M.Eng. degree in IT Convergence Engineering from Kumoh National Institute of Technology (KIT), Gumi, South Korea, in 2025, and his B.Sc. degree in Computer Science and Engineering from the International University of Business, Agriculture, and Technology (IUBAT), Bangladesh, in 2018. His research interests include anomaly detection, deep learning, smart contracts, and blockchain applications in logistics.

[ORCID:0009-0004-7869-744X]

Md Mahinur Alam

He is pursuing his Master's in IT Convergence Engineering at Kumoh National Institute of Technology (KIT) in Korea. His Bachelor of Science in Computer Science and Engineering (CSE) from Bangladesh's International University of Business Agriculture and Technology (IUBAT) was awarded in December 2022. His research interests are Deep Learning, Computer Vision, Federated Learning, and Blockchain.

[ORCID:0009-0006-5453-443X]

Mohtasin Golam

He received his Ph.D. degree in IT Convergence Engineering from Kumoh National Institute of Technology (KIT), South Korea, in 2025. He received an M.Sc. in IT Convergence Engineering from the same institution. He received his B.Sc. degree in Electrical and Electronics Engineering (EEE) in 2018. He is currently working with the Network System Laboratory (NSL) as a research assistant under the supervision of the ICT Convergence Center. His research interests are Deep learning, Metaverse applications, UAV networks, IoT, and Blockchain.

[ORCID:0000-0001-9784-0679]

Md Facklasur Rahaman

He is currently pursuing a Ph.D. from the Faculty of Engineering and Applied Science in the Electronic Systems Engineering department at the University of Regina and completed his Master's degree in IT Convergence Engineering while working as a full-time researcher at the Networked System Laboratory in Kumoh National Institute of Technology, Gumi, South Korea.

He received the B.Sc. degree in electrical and electronic engineering from the Rajshahi University of Engineering and Technology, Bangladesh, in 2016. His research interests include medicine supply chain management, healthcare record management, meta-verse governance and applications, and blockchain.

[ORCID:0009-0005-0974-1838]

Taesoo Jun

Feb. 1998 : B.S. degree, School of Electrical Engineering, Seoul National University.

Feb. 2000 : M.S. degree, School of Electrical Engineering, Seoul National University.

Dec. 2009 : Ph.D. degree, Computer Engineering, the University of Texas at Austin.

2010~2022 : Director and Principal Engineer, SW Platform team/Global AI Center, Samsung Research, Samsung Electronics, Seoul, Korea. Mar.

2022~Current : Assistant Professor, School of Computer Engineering, Kumoh National Institute of Technology. He is a member of IEEE. His research interests are in distributed computing in a smart environment, intelligent system design for pervasive computing, and real-time system.

[ORCID:0000-0002-1435-3769]