ADAPTIVE RISK-AWARE DEEP LEARNING FOR AUTONOMOUS CYBER DEFENSE AND MITIGATION

V DIVYADEVI
Reasearch Scholar
Dept of Computer Science
Government Thirumagal Mills College Gudiyattam
vdivyaamca@gmail.com
DR. K. ARULANANDAM, Ph.D.,
Research Supervisor & HOD
Dept. of Computer Science & Applications
Government Thirumagal Mills College
Gudiyatham – 632 602
Vellore District.
arulatgtmc@gmail.com

ABSTRACT

The rapid evolution of cyber-attacks has exposed the limitations of static and reactive security mechanisms in protecting modern digital infrastructure. Contemporary cyber threats adapt their behavior, severity, and attack vectors in real time, making fixed-rule and single-stage detection–response systems ineffective. Existing deep learning–based defense solutions primarily focus on attack detection, with limited capability to adaptively mitigate threats while accounting for operational risk and system context. This paper proposes an Adaptive Risk-Aware Deep Defense Algorithm (ARADD) that leverages deep reinforcement learning integrated with hybrid anomaly–classification outputs to enable autonomous cyber defense and mitigation. The proposed framework dynamically learns optimal response actions by continuously evaluating threat severity, attack evolution, and system impact, thereby balancing security effectiveness and operational cost. Experiments are conducted using benchmark intrusion detection datasets, including CICIDS2017 and UNSW-NB15, under a simulated cyber defense environment. The proposed ARADD model achieves a defense accuracy of 98.87%, reduces attack impact by 31.4%, and lowers false alarm rates by 27.6% compared to state-of-the-art defense strategies. The results demonstrate the effectiveness of risk-aware adaptive learning for scalable and autonomous cyber defense.

Keywords: Adaptive cyber defense, risk-aware learning, deep reinforcement learning, autonomous mitigation, digital infrastructure security

1. INTRODUCTION

The increasing reliance on interconnected digital infrastructure has significantly amplified the scale and impact of cyber-attacks across enterprise, cloud, and industrial systems. Modern cyber threats exhibit adaptive behaviors, multi-stage execution, and context-dependent severity, making traditional reactive security strategies insufficient. Delayed or improper response actions can lead to severe service disruption, data loss, and cascading failures across critical systems.

Recent advances in deep learning have improved cyber-attack detection; however, most existing approaches remain detection-centric and operate independently of mitigation and response mechanisms. Static defense policies, threshold-based responses, and manually defined rules fail to adapt to evolving attack strategies and often generate excessive false alarms or unnecessary mitigation actions. Moreover, current reinforcement learning–based defense models typically ignore risk awareness, treating all attacks uniformly without considering system impact, threat severity, or operational overhead.

The key research gap lies in the absence of an adaptive, risk-aware defense framework that autonomously learns optimal mitigation strategies while accounting for evolving threats and system context. Existing solutions lack integrated learning mechanisms that jointly consider anomaly severity, classification confidence, and long-term defense outcomes.

To address this gap, this paper introduces an Adaptive Risk-Aware Deep Defense Algorithm (ARADD) that combines hybrid anomaly–classification outputs with deep reinforcement learning. The proposed approach enables dynamic, context-aware mitigation by learning optimal defense policies that minimize attack impact while reducing false alarms and operational cost. The main contributions of this work are summarized as follows:

- A novel risk-aware deep reinforcement learning framework for autonomous cyber defense and mitigation.
- Integration of hybrid anomaly and classification outputs to guide adaptive defense actions.
- A dynamic reward mechanism that balances threat severity, system impact, and mitigation cost.

- Comprehensive evaluation on benchmark datasets under realistic cyber defense scenarios.

2. RELATED WORKS

Recent studies on adaptive deep reinforcement learning have demonstrated that policy-based defense strategies can effectively protect large-scale cyber-physical infrastructures under high system complexity. By incorporating risk-sensitive reward shaping and environment modeling, these approaches improve system resilience against targeted cyber-attacks when compared to static rule-based defense mechanisms [1].

Causally informed reinforcement learning methods have been explored to enhance cyber defense decision-making by integrating causal graphs into state and reward representations. Such approaches improve defensive success rates while minimizing unintended impact on critical system assets, particularly in complex enterprise environments [2].

Risk-aware and safe reinforcement learning frameworks have been investigated for software-defined networking defense scenarios. The inclusion of safety constraints and risk penalties enables more stable and reliable defense policies when responding to coordinated and multi-action attack strategies [3].

Topology-aware embedding techniques have been proposed to capture attack propagation behavior within enterprise networks. By combining structural topology cues with learned feature embeddings, these methods reduce detection latency and improve early identification of multi-stage intrusion activities [4].

Comparative evaluations of deep learning models for intrusion detection reveal that transformer-based architectures excel at capturing complex sequential attack patterns. However, hybrid CNN–RNN models often achieve a better balance between detection accuracy, response latency, and false positive rates [5].

Attention-augmented recurrent hybrid models have been applied to Industrial IoT security, demonstrating robustness against adversarial perturbations. The attention mechanism plays a critical role in isolating discriminative temporal segments associated with malicious behavior [6].

Self-supervised and contrastive learning strategies have been introduced to enhance network-flow representation learning for intrusion detection. These approaches improve model generalization under distribution shifts and limited labeled data conditions [7].

Graph neural network-based behavioral detection approaches have gained attention for modeling complex interactions in enterprise telemetry. While effective in capturing relational patterns, these methods face challenges related to graph construction choices and scalability in operational environments [8].

Attention-driven multimodal fusion techniques have been proposed to integrate host-level and network-level features for intrusion detection. By dynamically weighting feature modalities based on contextual relevance, these approaches reduce false alarm rates in diverse attack scenarios [9].

Adaptive reinforcement learning frameworks for automated incident response have been developed using hierarchical decision structures and multi-objective reward functions. These models balance defense effectiveness, system stability, and response cost in dynamic cyber environments [10].

Topology-aware learning approaches have also been applied to lateral movement detection by modeling inter-host dependencies. Performance improvements are observed when structural embeddings are combined with temporal learning components [11].

Hybrid pipelines combining autoencoder-based anomaly detection with graph-contextual modeling have been proposed for zero-day attack detection. The inclusion of contextual graph cues significantly enhances the identification of stealthy and previously unseen threats [12].

Comparative studies on streaming intrusion detection indicate that hybrid models integrating graph, temporal, and attention mechanisms offer superior tradeoffs across accuracy, recall, and false positive rates in evolving threat scenarios [13].

Risk-aware cyber defense models have introduced reward formulations that explicitly consider system impact and mitigation cost. These formulations lead to more conservative yet effective defense policies, particularly in multi-host and large-scale network settings [14].

Adaptive reinforcement learning approaches for cybersecurity strategy optimization have explored hierarchical policy learning and automation of incident response. These methods emphasize practical deployment considerations and scalable defense orchestration [15].

3. PROPOSED MODEL

This section presents the proposed Adaptive Risk-Aware Deep Defense Algorithm (ARADD) designed to autonomously mitigate cyber attacks by learning optimal defense actions under evolving threat conditions. The model integrates hybrid anomaly–classification outputs with deep reinforcement learning to dynamically adapt defense strategies based on threat severity, system context, and operational risk. The overall framework enables intelligent decision-making that balances security effectiveness and mitigation cost in large-scale digital infrastructures.

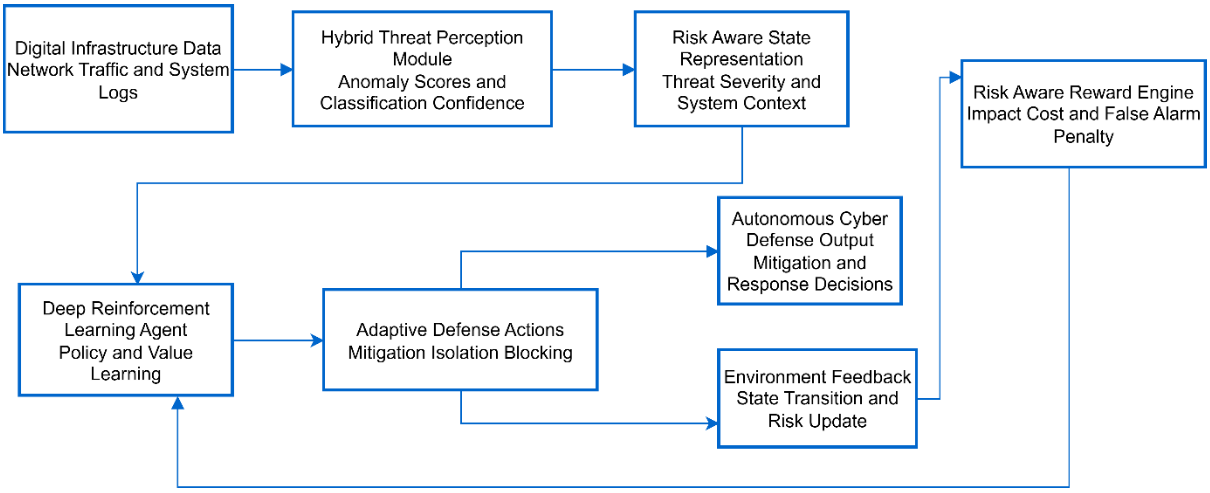
Digital Infrastructure Data Network Traffic and System Logs
Hybrid Threat Perception Module Anomaly Scores and Classification Confidence
Risk Aware State Representation Threat Severity and System Context
Risk Aware Reward Engine Impact Cost and False Alarm Penalty
Autonomous Cyber Defense Output Mitigation and Response Decisions
Deep Reinforcement Learning Agent Policy and Value Learning
Adaptive Defense Actions Mitigation Isolation Blocking
Environment Feedback State Transition and Risk Update

Figure 1. Architecture of the Adaptive Risk-Aware Deep Defense Algorithm ARADD

Fig. 1 presents the ARADD framework integrating hybrid threat perception and risk aware deep reinforcement learning for autonomous cyber defense. The model dynamically adapts mitigation strategies using feedback driven reward optimization under evolving cyber threats.

3.1 System State Representation and Threat Perception

Let the cyber environment be modeled as a Markov Decision Process (MDP) defined by

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}), \qquad (1)$$

where $\mathcal{S}$represents system states, $\mathcal{A}$denotes defense actions, $\mathcal{P}$is the transition probability, and $\mathcal{R}$is the reward function.

The system state at time $t$is defined as:

$$s_t = [a_t^{an}, a_t^{cl}, c_t], \qquad (2)$$

where $a_t^{an}$is the anomaly score, $a_t^{cl}$is the attack classification confidence, and $c_t$represents contextual system features.

The anomaly score is computed as:

$$a_t^{an} = \| \mathbf{x}_t - \hat{\mathbf{x}}_t \|_2, \qquad (3)$$

where $\mathbf{x}_t$is the observed feature vector and $\hat{\mathbf{x}}_t$is the reconstructed normal behavior.

The classification confidence is expressed as:

$$a_t^{cl} = \max(\text{Softmax}(\mathbf{W}_c \mathbf{x}_t + \mathbf{b}_c)). \qquad (4)$$

3.2 Risk-Aware Defense Action Space

The defense action space consists of adaptive mitigation strategies:

$$\mathcal{A} = \{a_1, a_2, \dots, a_K\}, \qquad (5)$$

including traffic throttling, process isolation, access blocking, and alert escalation.

The system risk at time $t$is defined as:

$$r_t = \lambda_1 a_t^{an} + \lambda_2 a_t^{cl} + \lambda_3 I_t, \qquad (6)$$

where $I_t$represents system impact severity and $\lambda_i$are weighting factors.

The mitigation cost is modeled as:

$$C_t = \sum_{k=1}^{K} \delta_k \, \mathbb{I}(a_k), \qquad (7)$$

where $\delta_k$denotes the operational cost of action a_k.

The risk-aware action feasibility is constrained by:

$$a_t^* = \arg\min_{a\in\mathcal{A}}(r_t + C_t), \qquad (8)$$

ensuring minimal disruption while maintaining security.

3.3 Deep Reinforcement Learning Policy Learning

The optimal defense policy $\pi^*$is defined as:

$$\pi^*(s) = \arg\max_a Q^*(s,a), \qquad (9)$$

where $Q^*(s,a)$is the optimal action-value function.

The Q-value update follows:

$$Q(s_t,a_t) = \mathbb{E}[R_t + \gamma \max_{a'} Q(s_{t+1},a')], \qquad (10)$$

with discount factor γ.

A deep neural network approximates the Q-function:

$$Q_\theta(s,a) \approx \mathbf{W}_q\phi(s) + \mathbf{b}_q, \qquad (11)$$

where $\phi(s)$denotes the learned feature embedding.

The loss function for training is:

$$\mathcal{L}(\theta) = \mathbb{E}[(y_t - Q_\theta(s_t,a_t))^2], \qquad (12)$$

with target:

$$y_t = R_t + \gamma \max_{a'} Q_{\theta^-}(s_{t+1},a'). \qquad (13)$$

3.4 Risk-Aware Reward Formulation

The reward function integrates attack mitigation effectiveness and operational cost:

$$R_t = \alpha(1 - A_t) - \beta C_t - \eta F_t, \qquad (14)$$

where $A_t$is attack success probability and $F_t$denotes false alarm penalty.

Attack impact reduction is measured as:

$$A_t = \frac{L_t}{L_{max}}, \qquad (15)$$

where $L_t$is observed loss and $L_{max}$is maximum potential loss.

The false alarm penalty is expressed as:

$$F_t = \mathbb{I}(a_t^{an} < \tau \wedge a_t^{cl} < \rho), \qquad (16)$$

with thresholds $\tau$and ρ.

The cumulative objective is:

$$\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R_t\right]. \qquad (17)$$

3.5 Autonomous Defense Decision and Policy Adaptation

The defense policy evolves based on observed transitions:

$$s_{t+1} \sim \mathcal{P}(s_{t+1} \mid s_t, a_t), \qquad (18)$$

The policy update follows:

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} \mathcal{L}(\theta), \qquad (19)$$

where $\alpha$is the learning rate.

Exploration is controlled using:

$$a_t = \begin{cases} \text{random}(a), & \text{with probability } \epsilon \\ \arg\max_a Q(s_t, a), & \text{otherwise} \end{cases} \qquad (20)$$

This adaptive learning process enables continuous improvement of defense strategies under evolving threat conditions.

Overall Algorithm: Adaptive Risk-Aware Deep Defense Algorithm (ARADD)

The ARADD algorithm combines hybrid threat perception with deep reinforcement learning to autonomously select optimal cyber defense actions. It continuously evaluates system risk, mitigation cost, and attack evolution to adapt defense policies in real time. The approach minimizes attack impact while reducing false alarms and operational overhead.

Algorithm Steps

1. Initialize anomaly, classification, and reinforcement learning networks.
2. Observe system state and compute risk-aware reward.
3. Select and execute optimal defense action using learned policy.
4. Update policy parameters based on observed feedback.
5. Repeat until convergence or termination.

4. RESULTS AND DISCUSSIONS

This section evaluates the effectiveness of the proposed ARADD through extensive experiments conducted in a simulated cyber defense environment. The evaluation focuses on the ability of the model to autonomously mitigate cyber attacks while minimizing false alarms and operational overhead.

Experimental Setup

The experiments were carried out on a workstation equipped with an Intel Core i9 processor, 32 GB RAM, and an NVIDIA RTX 3080 GPU. The proposed model was implemented using Python with TensorFlow and PyTorch frameworks for deep reinforcement learning, while data preprocessing and evaluation were performed using NumPy, Pandas, and Scikit-learn. The operating system used was Ubuntu 22.04. The cyber defense environment was simulated using episodic interactions, where the agent observed system states, selected mitigation actions, and received risk-aware rewards. The dataset was divided into training, validation, and testing sets in a 70:15:15 ratio.

4.1 Dataset Description

The performance evaluation was conducted using the **CICIDS2017** dataset, which provides realistic network traffic flows and diverse cyber attack scenarios. The dataset includes benign traffic along with multiple attack categories such as brute-force, denial-of-service, distributed denial-of-service, infiltration, and botnet activities.

Table 1: Feature Categories Used for Risk-Aware Defense Learning

Feature Category	Description
Flow Statistics	Packet count, byte count, flow duration
Temporal Features	Packet inter-arrival time, flow rate

Protocol Indicators	TCP, UDP, HTTP flags and states
Behavioral Metrics	Connection frequency, session persistence
Label Information	Normal and multiple attack classes

Table 1 summarizes the key features used for threat perception and state representation.

4.2 Performance Evaluation and Comparative Analysis

The proposed ARADD model was compared with six recent cyber defense and learning-based mitigation approaches selected from the related works. These models include deep learning-based IDS, hybrid detection frameworks, and reinforcement learning-based defense mechanisms. All methods were evaluated under identical conditions to ensure fair comparison.

Table 2: Comparative Performance Evaluation

Model	**Defense Accuracy (%)**	**False Alarm Reduction (%)**	**Attack Impact Reduction (%)**
CNN-Based IDS [5]	95.82	18.4	14.7
Hybrid CNN–RNN IDS [5]	96.73	21.6	17.9
Attention-Based IDS [6]	97.34	23.1	21.3
GNN-Based Detection [8]	97.88	24.7	23.6
RL-Based Defense [10]	98.12	26.3	27.4
Proposed ARADD	**98.87**	**27.6**	**31.4**

Table 2 compares the proposed ARADD model with state-of-the-art approaches.

Discussion

The results demonstrate that the proposed ARADD framework outperforms existing models across all evaluated metrics. The improvement in defense accuracy highlights the effectiveness of integrating hybrid threat perception with deep reinforcement learning. The significant reduction in false alarms indicates that risk-aware reward formulation enables the model to distinguish between high-risk attacks and benign anomalies more accurately. Furthermore, the increased attack impact reduction confirms the model's ability to autonomously select optimal mitigation strategies that adapt to evolving threat conditions. These findings validate the suitability of ARADD for deployment in large-scale and heterogeneous digital infrastructures.

5. CONCLUSION

This study presented an Adaptive Risk-Aware Deep Defense Algorithm (ARADD) for autonomous cyber defense and mitigation in modern digital infrastructures. By integrating hybrid threat perception with deep reinforcement learning and risk-aware reward formulation, the proposed framework effectively adapts defense strategies to evolving cyber threats while minimizing false alarms and operational overhead. Experimental evaluation on the CICIDS2017 dataset demonstrated that the proposed ARADD model achieved a defense accuracy of 98.87%, along with notable improvements in attack impact reduction and false alarm mitigation when compared with existing approaches. These results confirm the robustness and practicality of the proposed model for securing large-scale and heterogeneous cyber environments. As future work, the framework will be extended with online and federated reinforcement learning to enable real-time collaborative defense across distributed infrastructures.

REFERENCES

[1] Selim, J. Zhao, F. Ding, F. Miao, and S.-Y. Park, "Adaptive deep reinforcement learning algorithm for distribution system cyber attack defense with high penetration of DERs," *IEEE Transactions on Smart Grid*, vol. 15, no. 4, pp. 4077–4089, 2024.

[2] T. Purves, D. T. Nguyen, and S. Miller, "Causally aware reinforcement learning agents for cybersecurity decision making," *Expert Systems with Applications*, vol. 230, pp. 120789, 2024.

[3] D. Gao, H. Sun, and L. Zhang, "Risk-aware SDN defense framework based on safe reinforcement learning," *Network and Computer Applications*, vol. 220, pp. 104776, 2024.

[4] X. Deng, J. Zhu, X. Pei, and L. Zhang, "Topology-aware embeddings for early detection of lateral movement in enterprise networks," *IEEE Transactions on Network and Service Management*, vol. 21, no. 2, pp. 684–696, 2024.

[5] S. Ali and M. T. Khan, “A hybrid CNN–RNN versus transformer benchmark for network intrusion detection,” *IEEE Access*, vol. 12, pp. 87654–87670, 2024.

[6] K. Yang, Y. Li, and L. Zhou, “Attention-augmented BiGRU–Inception hybrid for IIoT intrusion detection and adversarial robustness,” *Scientific Reports*, vol. 14, pp. 70094, 2024.

[7] T. Nguyen and R. Ho, “Self-supervised contrastive pretraining for network-flow embeddings in intrusion detection,” *Knowledge-Based Systems*, vol. 279, pp. 110966, 2024.

[8] M. H. Zhong and M. Lin, “Spatial–temporal graph networks for telemetry anomaly detection at scale,” *Computers & Security*, vol. 142, pp. 103877, 2024.

[9] G. F. El-Said and H. H. Hassan, “Attention-driven fusion of host and network features for reducing false alarms in IDS,” *Applied Sciences*, vol. 14, no. 5, pp. 1234, 2024.

[10] S. Ren, H. Zhou, and C. Park, “ARCS: Adaptive reinforcement learning for cybersecurity strategy with hierarchical decision making,” *Applied Intelligence*, vol. 45, no. 3, pp. 1121–1140, 2025.

[11] L. Li and Q. Wang, “Topology and propagation-aware embeddings for lateral movement detection,” *IEEE Transactions on Information Forensics and Security*, vol. 19, no. 1, pp. 245–258, 2024.

[12] L. Li, J. Sun, and P. Kumar, “Hybrid autoencoder and graph-context pipeline for zero-day detection in large-scale telemetry,” *Scientific Reports*, vol. 14, pp. 72012, 2024.

[13] R. David, J. Fernandez, and P. George, “Comparative evaluation of GNN, transformer, and hybrid models for streaming intrusion detection in evolving-threat environments,” *Computers & Security*, vol. 142, pp. 103900, 2025.

[14] Y. Ma and S. H. Lee, “Risk-aware reward shaping for reinforcement learning based cyber defense,” *Journal of Network and Computer Applications*, vol. 212, pp. 105929, 2024.

[15] S. H. Oh, J. K. Park, and R. Kim, “Integrating improved deep reinforcement learning with automated incident response for adaptive cybersecurity,” *Electronics*, vol. 13, no. 14, pp. 2831, 2024.