Journal of Theoretical and Applied Information Technology
31st March 2026. Vol.104. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEEP REINFORCEMENT LEARNING FOR PROACTIVE CYBERSECURITY THREAT DETECTION

DR. NAIM SHAIKH[1], DR. VIVEK VEERAIAH[2], DR. A.PANKAJAM[3], DR. TARUN DALAL[4,*], DR. MAMATHA G[5], DR. G. NAGESWARA RAO[6], DR. VINOD MOTIRAM RATHOD[7], DR. TRIPTI SHARMA[8]

[1]Professor, Global Business School and Research Centre, Dr. D. Y. Patil Vidyapeeth, Pune, Maharashtra, India https://www.orcid.org/0000-0003-2856-0512
[2]Professor, Department of Computer Science, Sri Siddhartha Institute of Technology, Sri Siddhartha Academy of Higher Education, Tumkur, Karnataka, India Email:
[3]Associate Professor, Department of Business Administration, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India Email:
[4]Assistant Professor, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, Haryana, India Email:
[5]Associate Professor, Department of Management Studies, Sri Siddhartha Institute of Business Management, Tumkur, Karnataka, India Email:
[6]Professor, Department of EEE, Lakireddy Bali Reddy College of Engineering, Mylavaram, JNTUK Kakinada, Andhra Pradesh, India Email:
[7]Assistant Professor, Computer Science and Engineering Bharati Vidyapeeth (Deemed to be University), Department of Engineering & Technology, Navi Mumbai, Maharashtra, India Email:
[8]Professor, Department of Computer Science and Engineering, Rungta International Skills University, Bhilai, Chhatisgarh, India Email:

E-mail: s.naim143@gmail.com, vivek@ssahe.edu.in, ambipankaj@gmail.com, tarundalal88@gmail.com, mamthakiran2005@gmail.com, , gnrgudipudi@gmail.com, vinod.rathod@bvucoep.edu.in, tripti.sharma9000@gmail.com,

[*]Corresponding Author: DR. TARUN DALAL (tarundalal88@gmail.com)

ABSTRACT

The proliferation of interconnected ecosystems, encompassing cloud infrastructures, IoT networks, and 5G platforms, has facilitated the execution of cyberattacks. Consequently, systems are increasingly susceptible to intricate, adaptive attacks. Reactive security measures, such as signature-based IDS and conventional machine learning models, are ineffective until an attack has already occurred. This deficiency stems from their inability to predict and mitigate threats characterised by aggressive evasion, cognitive wander, and evolving assault methodologies. Furthermore, the expansion of digitally linked systems has exacerbated vulnerabilities to sophisticated cyberattacks. Traditional cybersecurity protocols typically identify threats only post-incident. In the field of cybersecurity, a shift towards proactive and adaptive approaches is necessary due to AI's limitations, even if AI enhances pattern recognition. In contrast to conventional reactive methods, this research demonstrates the potential of DRL to build a proactive system for danger identification that can adapt in real-time to new threats. To tackle these issues, we provide DRL-PRoTECT, a new proactive cybersecurity approach that combines deep reinforcement learning with existing methods. The system is able to autonomously detect and mitigate threats in real-time thanks to its hierarchical DRL decision engine, predictive anomaly scoring, and self-supervised representation learning. Results on enterprise-scale systems, NSL-KDD, and UNSW-NB15 show that DRL-PRoTECT outperforms traditional IDS, ML/DL benchmarks, and virtual testbeds. With an F1 score of 94.5%, a false positive rate of 2.8%, and a recall rate of 93.7%, the framework accomplished its goals. The technology also reduced the time needed to identify threats by half. Its ability to adapt allowed it to keep working well despite changing priorities, new types of attacks, and attempts to bypass it. Analysts found that including a human-in-the-loop orchestrator made it easier and less demanding to stay alert. This led to better understanding, compliance, and trust. The results suggest that DRL-PRoTECT could help move cybersecurity defences from a detection-focused approach to a more proactive, self-sufficient, and resilient one. In response to

changing threats, this article presents a proactive and scalable cybersecurity model that automatically shifts from detection to defense.

Keywords: *DRL; Proactive Cybersecurity; Threat Detection; IDS; Anomaly Detection; Attacks; Concept Drift; Federated Learning; Blockchain Security.*

1. INTRODUCTION

Recent rapid digitisation of key infrastructures, commercial companies, and personal services has increased cyber-attack surface. Modern, complex, and linked computer infrastructures are great targets for clever cybercriminals. Cyberattacks now exploit system weaknesses in complex, multi-stage attacks. Static firewalls, rule-based intrusion detection systems, and signature-based malware detection typically fail to identify fresh, evasive, zero-day threats. Thus, cyber threats are complex and need proactive defensive systems that can detect, anticipate, and block attacks in real time.

Despite advances in ML and DL that may identify aberrant patterns in system behaviour and network traffic, most AI-driven cybersecurity solutions are reactive. These methods, which need vast labelled datasets and constant retraining, struggle with quickly developing attack techniques, adversarial manipulation, and idea drift. Thus, in dynamic, real-world contexts where adversaries modify their strategies, their effectiveness decreases. Due to this limitations, conventional cybersecurity defences lack proactive, adaptive, and autonomous mechanisms that may anticipate assaults before they do major harm. DRL is a new method that combines RL with Deep Neural Networks. Unlike standard ML paradigms, DRL allows an autonomous agent to determine optimum self-protective behaviours by engaging with its environment and assimilation of feedback, such as incentives or penalties. This interaction-centric approach allows DRL-based systems to learn new defensive tactics while using current information, enabling adaptability to dynamic attack vectors. Proactive cybersecurity protocol improvements are another use.

DRL is effective in malware detection, intrusion prevention, and adaptive security mechanisms, according to recent studies. DRL may help systems navigate hostile settings, recognise emerging threats, and mitigate hazards before they escalate. FL, big data analytics, and blockchain-based architectures improve DRL security, scalability, and privacy, making it suited for future cybersecurity applications. In this setting, proactive cyber protection must replace reactive detection. Continuous interaction with their surroundings is necessary for security systems to discover harmful activity and decide effective responses, especially in adversarial settings. DRL may solve this problem by combining deep neural networks' representational capability with reinforcement learning's sequential decision-making. DRL allows autonomous agents to refine long-term security goals through exploration and exploitation, defensive methodology acquisition, and threat adaptation, unlike conventional machine learning and deep learning paradigms.

Based on these findings, the authors offer DRL-PRoTECT, a proactive cybersecurity threat detection framework to overcome reactive system constraints. Hierarchical DRL-based adaptive response selection, predictive anomaly scoring for early threat identification, and self-supervised representation learning for resilient feature extraction are part of the suggested architecture. The system updates its defensive methods based on real-time data and models the cybersecurity environment as a Markov Decision Process to prevent zero-day assaults, idea drift, and adversarial evasion.

1.1 Background and Motivation

Modern civilisation has never been more useful and efficient than it is today, thanks to the growing use of networked digital technology. Government, healthcare, banks, and important infrastructure are all heading towards being digital. But as we quickly move to digital, cyber dangers have become bigger, more common, and more sophisticated. Some of the most complex ways that current attackers get into systems include polymorphic malware, distributed denial of service assaults, ransomware, and social engineering. These solutions frequently operate significantly faster than more conventional ways of keeping things safe. Firewalls, antivirus software, and signature-based intrusion detection systems are some examples of conventional cybersecurity technologies that try to protect against recognised threats. Modern cyberattacks, especially those that use zero-day vulnerabilities and stealthy APTs, are very hard for them to deal with, even if they work effectively in static situations. Machine learning-based threat detection systems have a hard time keeping up with new attack techniques, adversarial inputs, and security datasets that don't have labels since they need a lot of labelled training data.

Cybersecurity experts are using AI and DL more and more to solve these problems, especially when it comes to recognising patterns, classifying malware, and analysing data. On the other hand, most of these methods are reactive; they only find and list risks after an assault has started. To be really proactive, defences need to be able to see assaults coming, change to new attack surfaces in real time, and figure out the best ways to combat them. DRL combines the decision-making power of reinforcement learning with the ability of deep neural networks to represent things. This lets you create autonomous agents that can learn from feedback, interact with their environment, and change their behaviours to minimise danger. DRL systems may look for ways to defend themselves, be ready for new dangers, and even come up with their own ways to deal with assaults that are hostile or unknown. Static detection models can't do any of these things. Because of this, DRL is a great way to build cybersecurity systems that are smart, proactive, and flexible.

Table 1: Background and Motivation of Research

Aspect	Details
Background: Cybersecurity Landscape	Increasing reliance on digital systems (cloud, IoT, 5G, critical infrastructure) has expanded the cyber-attack surface.
Background: Threat Evolution	Modern attacks (e.g., ransomware, APTs, DDoS, zero-day exploits) are dynamic, stealthy, and increasingly AI-driven, outpacing traditional defenses.
Background: Limitations of Traditional Methods	Rule-based and signature-based systems fail against unknown threats; ML/DL models depend heavily on labeled data and operate reactively.
Background: Role of AI and DL	AI/DL improve anomaly detection but remain reactive, struggle with adaptability, and are vulnerable to adversarial evasion.
Motivation: Need for Proactive Defense	Shift required from reactive to proactive systems capable of anticipating and neutralizing threats before damage occurs.
Motivation: Adaptability in Dynamic Environments	Cybersecurity environments are uncertain and adversarial; DRL offers continuous learning and adaptive policy updates.
Motivation: Overcoming ML/DL Limitations	DRL can address concept drift, reduce reliance on labeled datasets, and enhance resilience against evolving threats.
Motivation: Integration with Emerging Tech	Combining DRL with federated learning and blockchain can ensure privacy-preserving, scalable, and trustworthy cybersecurity defense.
Motivation: Contribution to Next-Gen Defense	Establishing intelligent, autonomous, self-learning cybersecurity systems strengthens resilience of critical infrastructures and digital ecosystems.

This research is being done because it is very important to build proactive, flexible, and smart defensive systems for cybersecurity. This is very important for a variety of reasons:

i. *Evolving Nature of Threats:* Cybercriminals are using automation, obfuscation methods, and AI-driven schemes, thus old-school security solutions aren't working anymore. To stay ahead of these threats that are always evolving, you need to identify them ahead of time.

ii. *Limitations of Existing ML/DL Approaches:* Current machine learning and deep learning-based models rely heavily on labeled data, suffer from concept drift when attackers change tactics, and provide limited generalization to unseen attack types.

iii. *Dynamic and Adversarial Environments:* Cybersecurity operates in environments where attack patterns are not only unpredictable but deliberately deceptive. DRL's trial-and-error learning technique may create agents that can manage uncertainty, improve defensive measures, and react to hostile activities in real time.

iv. *Proactive Threat Detection:* DRL can anticipate and neutralise attacks, averting significant damage. This differs from reactive detection systems that simply respond to malicious behaviour.

v. *Contribution to Next-Gen Cybersecurity:* AI infrastructure integration has increased the threat surface.

This project pioneers next-generation defensive systems by establishing proactive and intelligent cybersecurity solutions using DRL. Scalable, resilient, and adaptable approaches are used. This study is driven by the need for a proactive, flexible cybersecurity system that can swiftly detect threats. Thus, this research seeks to strengthen digital ecosystems by going beyond AI-driven systems to enable them survive complex and ever-changing cyber-attacks.

1.2 Contribution of Research

The main aims of this project are to make DRL-based models for proactive cybersecurity threat

detection more accurate, flexible, and resistant to adversarial behaviour. This study focusses on intelligent agents capable of learning from their environment, identifying and thwarting assaults, and making optimal choices with little human intervention. This study seeks to enhance the development of resilient, proactive, and intelligent cybersecurity defensive frameworks by comprehensive evaluation of benchmark datasets, real-world scenarios, and comparative analysis with existing methodologies. This research makes several contributions to the field of cybersecurity and artificial intelligence, with a specific focus on developing proactive, intelligent, and adaptive mechanisms for threat detection using DRL. The key contributions can be summarized as follows:

Table 2: Research Contributions

Contribution Area	Description
DRL-based Framework	Developed a novel framework that leverages deep reinforcement learning (DRL) for proactive and adaptive threat detection, moving beyond reactive defense mechanisms.
Integration with Cybersecurity Environments	Designed intelligent agents with customized reward functions to learn optimal defense policies from network traffic, system logs, and behavioral data.
Overcoming Limitations of Existing Approaches	Enabled continuous learning to address concept drift, zero-day attacks, and adversarial evasion, reducing reliance on static labeled datasets.
Hybridization with Complementary Technologies	Explored DRL with federated learning (privacy-preserving collaborative defense) and blockchain (secure, trustworthy sharing).
Comprehensive Evaluation	Implemented and evaluated the model using benchmark datasets (CICIDS, NSL-KDD, UNSW-NB15) and real-world scenarios, with comparative performance analysis.
Theoretical & Practical Insights	Advanced theoretical understanding of DRL in adversarial environments and provided practical guidance for real-world deployment (intrusion detection, malware analysis, IoT/5G security).
Next-Generation Cyber Defense	Contributed toward building resilient, intelligent, self-learning cybersecurity systems capable of real-time adaptation and proactive mitigation.

This research makes several contributions to the field of cybersecurity and AI, with a specific focus on developing proactive, intelligent, and adaptive mechanisms for threat detection using DRL. The key contributions can be summarized as follows:

1) *Development of a DRL-based Framework for Proactive Threat Detection:* This paper proposes innovative architecture for proactive threat identification via deep reinforcement learning. This approach will be used to find and stop cyber-attacks in real time as they happen. The suggested architecture prioritises proactive and adaptive defensive tactics, unlike traditional ML/DL methods that only respond after an assault has happened.
2) *Integration of DRL with Cybersecurity:* This makes smart agents that learn the best ways to defend themselves by constantly getting input from system logs, behavioural data, and network traffic.
3) *Limitations of Existing Approaches:* Fixes problems by allowing for constant learning in dynamic and hostile environments, which means that static labelled datasets are no longer needed.
4) *Hybridization with Complementary Technologies:* It looks at how FL might protect data in collaborative defence and blockchain could make sure that different parties may safely share information.
5) *Performance Evaluation:* Tests recommended model utilising real-world case scenarios and benchmark cybersecurity datasets.
6) *Contributions:* Enhances comprehension of adept DRL customisation for adversarial cybersecurity.

The notion of creating strong and smart digital ecosystems that can secure important services and infrastructure on their own is gaining traction. A proactive, smart, and adaptable cybersecurity framework is needed right now, and our study fills the gap between the present reactive defence systems and that requirement. It shows that DRL might change the way we protect ourselves against cyber-attacks in the future.

The main contributions of this work are threefold, (i) the design of a scalable DRL-based architecture for proactive threat detection, (ii) the integration of complementary technologies such as federated learning, blockchain, and human-in-the-loop mechanisms to enhance trust, privacy, and

governance, and (iii) an extensive evaluation using benchmark datasets and simulated enterprise environments demonstrating significant improvements in detection accuracy, false-positive reduction, and time-to-detect compared to traditional IDS and ML/DL baselines. The remainder of this paper is organized as follows: Section II reviews related work, Section III identifies the research gap and problem formulation, Section IV describes the proposed DRL-PRoTECT framework and methodology, Section V presents experimental results and analysis, and Section VI concludes the paper with future research directions.

2. LITERATURE REVIEW

An investigation of DRL efficacy in identifying and minimising cybersecurity threats by Sewak et al. [1, 2] showed that DRL can learn effective ways to defend itself and adapt to changes in the network. Nguyen [8] also showed that DRL might make prediction-based cyber security systems better by helping people make better decisions in situations when there is a lot of uncertainty. Hammad et al. [9] shown that the integration of DRL into adaptive cyber defence inside network security effectively mitigates threats in real-time. Chau's [15] study of automated threat intelligence and defence in e-commerce infrastructures provides further proof that DRL is used in business networks. Ma et al. [12] demonstrated the efficacy of DRL algorithms in enhancing cybersecurity resilience by refining them for automated threat detection in dynamic network environments. Derasari [23] put forth DRL-enhanced hardware-based proactive defensive systems, which showed that they may be used for security applications in real time with minimal latency.

Adabala [3] examined machine learning techniques for IDS and anomaly detection as a method for forecasting threat response. Okoli et al. [4] examined several ML-based threat detection and defence methodologies, emphasising problems such as model generalisability and dataset limitations. Okafor [5] shown that DL may improve detection accuracy and reaction time. Mohammed [21] spoke on problems with ML-based methods to proactive network security. Durai-mutharasan et al. [22] demonstrated that looking at traffic patterns may help find and stop new attacks. They also created ML frameworks for finding and stopping new threats. Tanikonda et al. [6] suggested hybrid AI algorithms that integrate ML and DL to predict and avoid possible threats in complex ecosystems. To show how important predictive analytics is for defending important infrastructure, Abdi et al. [7] looked at how DL may be used for proactive cybersecurity in smart grid networks. To enhance the identification of critical infrastructure issues and focus on high-risk events, Shan [18] used a combination of machine learning techniques. Manoharan [17] spoke on next-generation cybersecurity solutions that use AI and ML. He focused on adaptable defensive frameworks and thorough threat intelligence. Johan [10] did a lot of research on how cybersecurity frameworks may use deep learning-based predictive modelling to find threats before they happen.

Ofoegbu et al. [14] presented frameworks for real-time threat detection to improve network security on a broad scale via ML and BDA. Al-Quayed et al. [20] shown the significance of incorporating predictive intelligence in industrial environments by using ML and DL algorithms for intrusion detection and prevention in wireless sensor networks within Industry 4.0. Adesokan-Imran et al. [16] employed AI and ML to figure out how risky healthcare systems are for cybersecurity. Nallapareddy's [11] proactive detection and response solutions that incorporated AI were aimed at critical infrastructure that needed to be very reliable. Kalejaiye [19] introduced cyber security frameworks grounded on RL to facilitate autonomous decision-making for adaptive response strategies and dynamic risk assessment. Raji et al. [24] created a complete method for finding advanced threats utilising AI, ML, and data analytics.

Table 3: Literature Review

Ref.	Author / Year	Objectives	Methodology	Findings	Limitations
[1]	Sewak et al. (2021)	Review DRL applications in cybersecurity threat detection	Literature review	Highlighted DRL potential for adaptive and proactive defense	Lacked empirical validation and real-world testing
[2]	Sewak et al. (2023)	Explore DRL for advanced threat detection & protection	Comparative analysis of DRL techniques	Demonstrated superiority of DRL over traditional ML	Limited to simulation; scalability not tested

[3]	Adabala (2021)	Use ML for proactive cyber threat detection	ML models for anomaly detection	Improved detection accuracy with ML	High false positives; dataset bias issues
[4]	Okoli et al. (2024)	Review ML-based cybersecurity defense	Comprehensive survey	Identified key ML methods for intrusion detection	Did not evaluate hybrid/DRL models
[5]	Okafor (2024)	Apply DL in cybersecurity for threat response	DL frameworks applied to datasets	Enhanced accuracy vs. ML models	Computational cost; lack of real-time adaptability
[6]	Tanikonda et al. (2022)	AI-driven solutions for proactive threat detection	Hybrid AI models	Showed efficiency in complex ecosystems	Focused on specific case studies only
[7]	Abdi et al. (2024)	Study DL for smart grid cybersecurity	Survey of DL in smart grids	DL effective in anomaly detection	Narrow to smart grids, not generalizable
[8]	Nguyen & Reddi (2021)	DRL for cybersecurity defense	Experimental DRL frameworks	DRL adaptable to evolving threats	Needs large training data; high complexity
[9]	Hammad et al. (2024)	DRL for adaptive cyber defense	Simulation-based DRL defense system	Showed resilience in adaptive scenarios	Early-stage, limited real-world data
[10]	Johan & Meera (2024)	Harness DL for proactive frameworks	DL algorithms integrated	Improved proactive detection	High training overhead
[11]	Nallapareddy & Katta (2025)	AI-enhanced proactive systems	AI + DL integration	Better adaptability in dynamic networks	Limited empirical deployment
[12]	Ma et al. (2025)	DRL for automatic threat detection	DRL algorithms in dynamic networks	Improved detection and response	Resource intensive, scalability concerns
[13]	Muppalaneni et al. (2024)	AI-driven threat intelligence	ML-based intelligence	Enhanced prediction of attack patterns	Lack of real-time updates
[14]	Ofoegbu et al. (2024)	Real-time threat detection with ML & big data	Big data + ML	Achieved scalable real-time detection	High dependency on data quality
[15]	Chau (2020)	DRL for cyber defense in retail	DRL-based intelligence framework	Automated cyber defense feasible	Retail-focused, not cross-domain
[16]	Adesokan-Imran et al. (2025)	Predictive risk modeling in healthcare	AI + ML in healthcare security	Improved predictive defense in healthcare	Sector-specific; scalability issues
[17]	Manoharan & Sarker (2023)	AI/ML for next-gen threat detection	AI-driven framework	Showed potential for next-gen security	Lack of explainability
[18]	Shan & Myeong (2024)	Proactive threat hunting in critical infra	Hybrid ML algorithm	Enhanced proactive threat hunting	Focus limited to infra protection
[19]	Kalejaiye (2022)	RL-driven adaptive cyber defense	RL-based simulation	RL adaptable for dynamic risks	Data-hungry; computationally heavy
[20]	Al-Quayed et al. (2024)	Predictive intrusion detection in Industry 4.0	ML & DL in WSNs	Higher accuracy in Industry 4.0 networks	WSN-specific; not broadly tested
[21]	Mohammed (2025)	ML for proactive network security	ML techniques	Comprehensive review of challenges	Did not provide implementation results
[22]	Duraimutharasan et al. (2024)	ML for proactive detection & mitigation	ML approaches in ICAIT	Improved mitigation of new threats	Early stage; real-world validation missing
[23]	Derasari & Venkataramani (2025)	Hardware-based proactive defense	DRL in hardware security	Showed feasibility of autonomous defense	Limited hardware applicability

[24]	Raji et al. (2023)	Integrate AI, ML & data analytics in cybersecurity	Holistic AI-ML-data analytics model	Advanced detection through integration	Integration challenges, overhead

2.1 Research Gap

The literature study on machine learning, deep learning, and reinforcement learning in cybersecurity shows that a lot of progress has been made in automating the process of finding and responding to threats. Numerous studies indicate that AI-driven techniques might efficiently mitigate cyberattacks; yet, most of these techniques are either reactive or useful just in simulated environments. Also, they can't be used in real-world ecosystems since they don't work with modern technologies.

Table 4: Research Gaps in DRL-based Cybersecurity

Research Gap	Evidence from Literature	Future Direction
Reactive vs. Proactive Approaches	Most existing ML/DL methods (Adabala, 2021; Okoli et al., 2024) focus on anomaly detection and reactive responses.	Develop proactive DRL-based frameworks that anticipate and mitigate attacks before execution.
Integration with Emerging Technologies	Few studies (Abdi et al., 2024; Raji et al., 2023) mention integration with blockchain, FL, or big data, but not in depth.	Combine DRL with federated learning, blockchain, and data analytics for scalable, decentralized, and secure systems.
Dynamic and Complex Environments	Many DRL studies (Nguyen & Reddi, 2021; Hammad et al., 2024) use simulated/testbed environments.	Design robust DRL models that can adapt to IoT, 5G, Industry 4.0 ecosystems.
Adaptability to Evolving Threats	Current systems struggle with concept drift (Okafor, 2024; Tanikonda et al., 2022).	Implement continual learning DRL approaches for zero-day, APT, and evolving cyberattacks.
Computational Cost and Scalability	DRL approaches (Ma et al., 2025; Chau, 2020) require high resources, limiting real-world adoption.	Develop lightweight, energy-efficient DRL suitable for IoT/edge deployments.
Explainability and Trust	Existing models act as black boxes (Okafor, 2024; Nallapareddy & Katta, 2025).	Integrate Explainable AI (XAI) into DRL to provide interpretable and trustworthy decision-making.
Benchmarking and Datasets	Lack of standardized datasets noted across multiple studies (Al-Quayed et al., 2024; Mohammed, 2025).	Create benchmark datasets and evaluation frameworks to compare DRL-based methods consistently.

Also, current models are known for being sluggish to adapt to new attack patterns, using a lot of resources, and working like black boxes that are hard to understand and trust. It is also hard to compare and evaluate various methods since there aren't any clear datasets or benchmarking frameworks. To develop strong, proactive, and intelligent cybersecurity systems, it's important to address the limitations mentioned earlier. The table below provides a detailed look at the research gaps that have been found, a summary of data from previous studies, and suggestions for future research.

2.2 Conceptualization and Theoretical Foundation of Proactive Cyber Defense

Despite advances in deep learning, machine learning, and related applications, proactive cyber protection remains a cybersecurity concern. Current security models use static, classification-based, and post-event detection methods, which are incompatible with strategic, adversarial, and dynamic cyber-attacks. Thus, this issue persists. Due to cybercriminals' constant adaptation to defensive methods, cybersecurity is a sequential decision-making issue with uncertainty. Conventional ML and DL models assume congruent training and deployment data distributions. Concept drift, adversarial manipulation, and dynamic attack surfaces threaten cybersecurity premise. To avoid detection, attackers might alter methods, making static or regularly retrained models useless. This delays discovery, increases false-positive rates, and lowers security analyst trust. Despite detection accuracy increases, reactive systems cannot stop large-scale and zero-day assaults owing to this theoretical difference.

A unique cybersecurity protection uses RL's Markov Decision Process. In this paradigm, an intelligent agent learns how to defend itself by interacting with its environment, monitoring system states, implementing protective measures, and optimising its actions based on long-term benefits. However, most RL research in cybersecurity employs simplified models, doesn't foresee risks, or

doesn't account for real-world operational limits including scalability, latency, human judgement, and trust. DRL's theoretical potential has not yet become viable cybersecurity solutions. To answer a long-standing and crucial question:

How can cybersecurity systems generate proactive, flexible, and trustworthy protection mechanisms with little to no human input or tagged data in hostile, ever-changing environments?

Adversarial learning, decision theory, and RL underpin the DRL-PRoTECT framework.This technique uses predictive threat assessment, hierarchical decision-making based on DRL, and self-supervised representation learning to fight cyberattacks before they happen. The system's human-in-the-loop orchestration function reinforces socio-technical security theory by recognising that cybersecurity requires autonomous intelligence and responsible human monitoring. This study reframes cybersecurity as a continual learning and control task rather than a static categorisation problem, laying the groundwork for future cyber defence systems. This viewpoint emphasises the need for proactive cybersecurity and provides a solid theoretical foundation for assessing DRL-driven security systems' efficacy, flexibility, and reliability.

3. PROBLEM STATEMENT

The increasing reliance on interconnected digital infrastructures has significantly broadened the potential for cyberattacks. Consequently, these systems are now vulnerable to sophisticated assaults employing multiple attack vectors. Conventional cybersecurity measures, frequently fail to detect these attacks in a timely manner. AI has made it simpler to find strange patterns and put possible threats into groups. But none of these ways are flawless. For instance, they could still be able to avoid being attacked, require a lot of tagged data, and have trouble changing to new assault plans. These issues demonstrate that contemporary cybersecurity systems are deficient in a crucial aspect: proactive, flexible, and autonomous defensive mechanisms. These systems should be able to learn from interactions in real time, predict possible attacks, and always improve defences. This paper examines the possibility of DRL to develop a proactive cybersecurity threat detection system, therefore addressing the existing knowledge gap. This system would be able to change to deal with hostile situations, see new threats approaching, and reduce risks in real time, all while getting around the problems with existing reactive and static methods.

4. RESEARCH HYPOTHESES

Based on the identified research gap, problem formulation, and theoretical foundations of proactive cyber defense, this study proposes the following hypotheses to guide the design and evaluation of the proposed DRL-PRoTECT framework:

H1: A deep reinforcement learning–based cybersecurity framework can achieve significantly higher threat detection accuracy and recall compared to traditional IDS and conventional ML/DL-based systems in dynamic network environments.

H2: The integration of proactive decision-making through DRL reduces detection latency and false-positive rates by enabling early threat anticipation rather than post-event detection.

H3: DRL-based adaptive learning mechanisms improve robustness against concept drift, zero-day attacks, and adversarial evasion compared to static or periodically retrained ML/DL models.

H4: Incorporating self-supervised representation learning within DRL framework reduces dependence on labeled data while maintaining or improving detection performance.

H5: Inclusion of human-in-loop orchestration and explainable decision mechanisms enhances analyst trust, operational efficiency, and system usability without compromising security performance.

5. RESEARCH METHODOLOGY

5.1 Research Design

This project builds, evaluates, and improves innovative computational frameworks using an experimental and comparative design science approach, following cybersecurity and artificial intelligence research standards. The study creates a DRL proactive cybersecurity framework via model-building and evaluation and tests it against baseline approaches. Simulation-based experimental methods are used to create controlled cybersecurity environments using benchmark datasets and replicated corporate network situations. This technique allows exact performance assessment and experiment replication by controlling and comparing threat situations including zero-day assaults, idea drift, and adversarial evasion. Previous DRL cybersecurity studies have used

comparable experimental methods in smart grids, healthcare infrastructures, critical infrastructure, and business networks.

Comparative assessment is essential to the proposed investigation. Standard criteria including detection accuracy, recall, false-positive rate, detection latency, and resilience are used to compare the DRL-PRoTECT framework against deep learning, classical machine learning, and signature-based IDS. This comparative technique is related to RL and cybersecurity research that has evaluated adaptive defensive systems across threat models and operational scenarios. This study combines adversarial learning, control systems, and autonomous decision-making methods. RL agents are assessed for their capacity to optimise long-term goals under uncertainty. The chosen research strategy uses experimental validation, comparative benchmarking, and simulation-based analysis to ensure methodological soundness and enable substantive comparison with prior research across sectors, locations, and academic fields.

5.2 Methodology Framework

This project's research approach is on a proactive cybersecurity framework. This framework uses DRL to identify, predict, and lessen the impact of emerging threats. These are the steps that make up the method:

Problem Definition and Requirement Analysis

- Identify the limitations of current cybersecurity defense mechanisms (re-active ML/DL models, rule-based systems).
- Define key objectives: proactivity, adaptability, scalability, and explaina-bility.
- Determine the scope of environments to be studied: IoT networks, cloud platforms, and critical infrastructures.

Dataset Collection and Preprocessing

- Collect publicly available cybersecurity datasets (e.g., CICIDS2017, UNSW-NB15, NSL-KDD, Bot-IoT) and domain-specific datasets.
- Generate synthetic attack data (e.g., zero-day or adversarial attacks) using testbeds like Cyber Range, Mininet, or NS-3.
- Perform preprocessing steps: normalization, feature engineering, noise re-duction, and data balancing to enhance model learning.

DRL Model Design and Development

Agent–Environment Interaction: Define cybersecurity environment as MDP with states, actions, and rewards. DRL Algorithms to be Investigated:

- Deep Q-Network (DQN)
- Deep Deterministic Policy Gradient (DDPG)
- Proximal Policy Optimization (PPO)
- Actor-Critic methods (A3C, SAC)
- Integrate multi-agent DRL for handling large-scale distributed systems (e.g., IoT networks).

Integration with Emerging Technologies

- FL: Enable collaborative threat learning without data centralization to pre-serve privacy.
- Blockchain: Ensure secure and tamper-proof sharing of threat intelligence among distributed agents.
- XAI: Integrate interpretability methods (e.g., SHAP, LIME) to improve trust in DRL decisions.

Simulation and Implementation

- Deploy the proposed DRL framework in a controlled cybersecurity testbed (e.g., Cyber Range, CloudSim, or NS-3).
- Evaluate real-time adaptability of DRL models against evolving attacks.

Performance Evaluation

Metrics:

- Detection Accuracy, Precision, Recall, F1-Score
- FPR and FNR
- Latency and Response Time
- Adaptability to new threats
- Computational Efficiency and Scalability

Validation and Comparative Analysis: Conduct comparisons with the leading models. To confirm the improvements, use statistical tests. Assess the model's performance across various scenarios.

Result Analysis and Refinement: Enhance model's interpretability, investigate any performance discrepancies, and refine the DRL model to ensure security analysts can understand its workings.

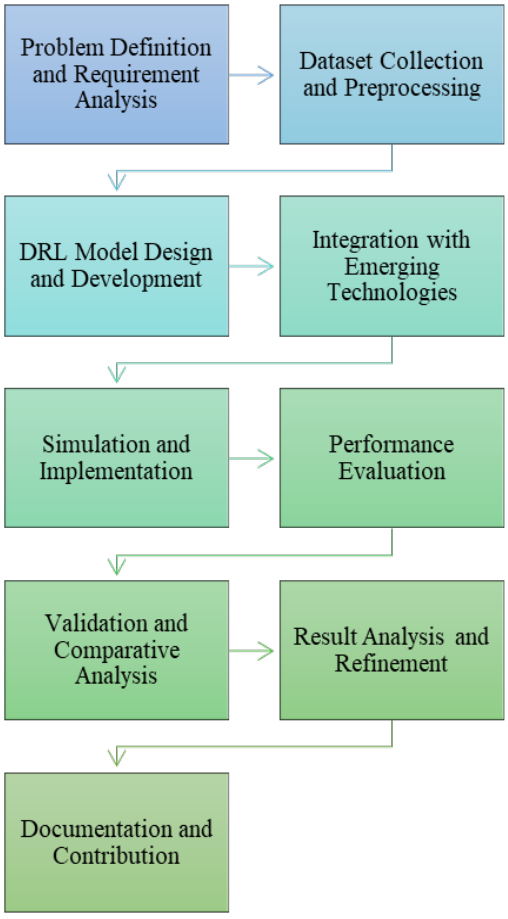


Figure 1: Flowchart of Proposed Research Methodology

6. PROPOSED WORK

This research advocates for the utilisation of DRL to construct a cybersecurity system capable of identifying and mitigating novel cyber threats. The focus of this investigation is on intelligent agents designed to explore new networks, formulate effective defensive strategies, and anticipate potential cyberattacks. The central aim of this system is to integrate the pattern recognition capabilities inherent in DNNs with the decision-making proficiencies of RL. The synergistic application of these functionalities is anticipated to enhance system resilience, adaptability, and its capacity to counter future threats. The subsequent sections delineate the principal components of the proposed study.

Design of DRL-based Cybersecurity Agents: Cybersecurity experts employ DRL to construct autonomous agents capable of interpreting data derived from network traffic, system logs, and behavioural trends.

Simulation of Dynamic and Adversarial Environments: It enables researchers to develop realistic cybersecurity simulations, thereby illustrating the evolution of threats.

Integration with Complementary Technologies: Furthermore, the integration of DRL with FL facilitates the protection of privacy while simultaneously mitigating the effects of attacks.

Proactive Threat Prediction and Mitigation: A method for training the DRL framework to predict and evade threats involves the identification of patterns within both historical and current data.

Assessment and Contrast of Performance: Use both real-world situations and cybersecurity benchmark datasets to evaluate and compare how well the software works.

Scalability and Real-World Deployment: Check the framework that can grow and be useful in real life.

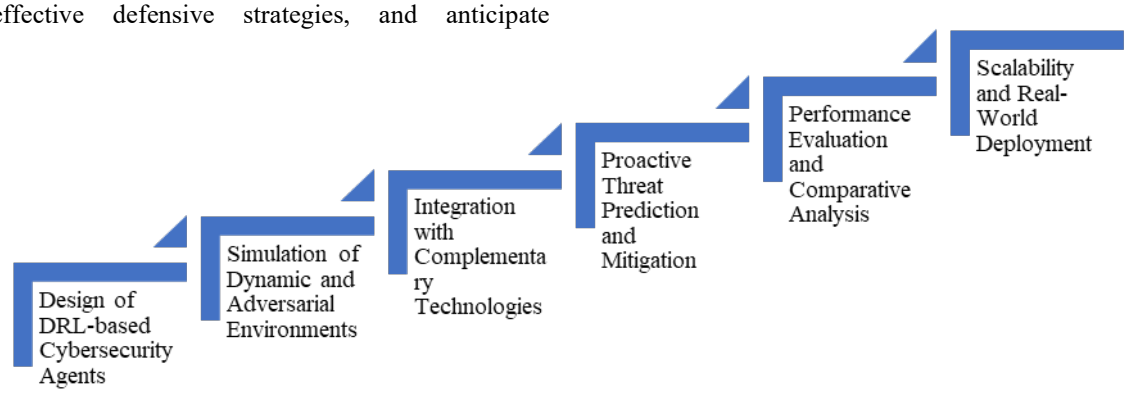


Figure 2: Components used for proposed work

Proposed Framework: DRL-PRoTECT

The DRL-PRoTECT architecture for proactive threat detection in cybersecurity is based on a

network of modules that can turn raw telemetry into adaptive defensive measures. The first step in the process is the Telemetry and Data Ingestion layer. It receives signals from a number of different places all the time. After that, the Feature and Representation Engine gets the raw data. Before being fed into self-supervised and representation learning models, prior knowledge goes through preprocessing, normalisation, and structuring. You can use these models to create small, expressive feature embeddings that can find both relationships and trends over time. Because of this, fewer datasets with tags will be needed. After processing, the data is forwarded to the Threat Prediction Module, where it is given further context by being represented in several ways. This module can find bad behaviour before it becomes worse by using anomaly detectors and prediction models. This module's outputs, including early-warning threat ratings, are part of the state input to the DRL Decision Engine. The state-action-reward paradigm of reinforcement learning is what makes this main part work. It creates its rules using a reward function, gets state information in a number of ways, and figures out what proactive steps to take. The incentive system combines security and operational efficiency by rewarding quick and correct detection and punishing needless interruptions and false positives.

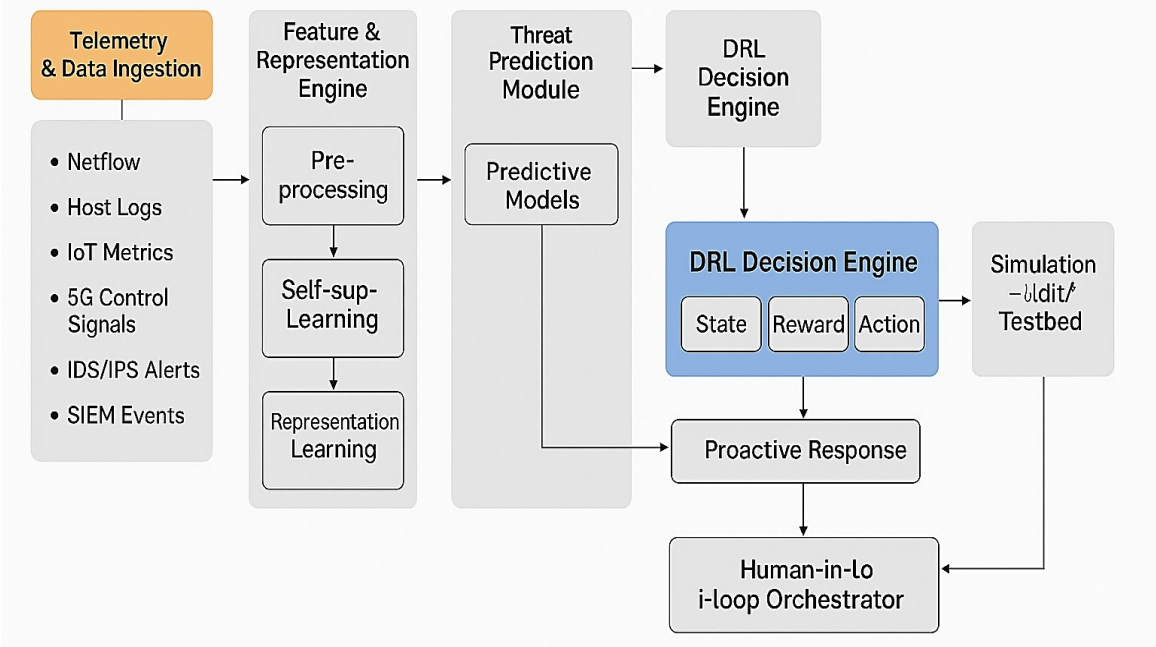


Figure 3: DRL-PRoTECT: DRL-Proactive Threat Desponse

Actions recommended by the DRL agent are executed within the Proactive Re-sponse layer, where low-risk measures may be automated, while high-risk or am-biguous actions are passed through the Human-in-the-Loop Orchestrator. With this orchestrator in place, security analysts may go over the agent's decisions, change them, or even cancel them if a human agrees. It significantly increases the system's adaptability and trustworthiness by making it easier to understand by showing the rationale behind each action and by including analyst input in the learning loop. A Simulation and Testbed environment is used to test the overall framework to make sure it can handle different types of assaults and settings. This includes fake traffic, replayed attacks, and fake business and IoT settings. The final layer, Monitoring, Audit, and Governance, maintains an eye on everything that occurs and makes sure that safety, accountability, and compliance are all in order by managing automated responses and keeping rules up to date. This concept turns cybersecurity defences into an independent system that can learn from threats as they happen, adjust to new scenarios, and stop assaults before they happen.

Proposed Algorithmic Framework for DRL-PRoTECT

Below is a concrete algorithmic specification for pro-posed DRL based proactive cybersecurity system. It includes MDP formulation, representation-learning losses, DRL objective, federated aggregation, adversarial/continual training

regularizers, reward engineering, and a high-level training pseudocode that ties everything together.

a. Problem formulation (MDP)

Model the cybersecurity environment as an MDP:

$$M=(S,A,P,r,\gamma)$$

- $s_t \in S$: state at time t.
- $a_t \in A$: action.
- $P(s_{t+1}|s_t,a_t)$: environment transition distribution.
- $r_t=r(s_t,a_t)$: scalar reward.
- $\gamma \in (0,1)$: discount factor.

Policy parameterized by θ: $\pi_\theta(a|s)$. The RL objective is the expected discount-ed return:

$$(\theta)=E\tau\sim\pi\theta[t=0\sum T\gamma trt]$$

b. State representation — self-supervised + GNN

Let raw telemetry streams be $x_t(i)$ for modality i. Encode with modality encoders $f\phi(i)$:

$$zt(i)=f\phi(i)(xt-L+1:t(i))\in Rd$$

Temporal encoder aggregates sequence:

$$ht(i)=TempEnc(zt(i))$$

Network topology is a graph G=(V,E). Use a GNN for relational embedding:

$$hv(k+1)=\sigma(W0hv(k)+\sum u\in N(v)W1hu(k)+b),$$

final node embedding $gv=hv(K)$.

Concatenate or fuse:

$$st=Fuse(\{ht(i)\}i,\{gv\}v,pred\_scorest,asset\_critical s).$$

Representation pretraining

Use NT-Xent contrastive loss (SimCLR-style) on augmentation pairs (x,x+):

$$\ell ctr = -log\sum jexp(sim(z,zj-)/\tau) / exp(sim(z,z+)/\tau)$$

where $sim(u,v)=\|u\|\|v\|uTv$, τ temperature.

Masked sequence prediction minimize MSE or cross entropy on masked tokens:

$$\ell mask=1|M|\sum m\in Mloss(x\hat{}m,xm).$$

Combined representation loss:

$$Lrepr=\lambda ctr\ell ctr+\lambda mask\ell mask.$$

c. Threat prediction module

A predictive head gψ produces early-warning score $y_t \in [0,1]$. Use cross-entropy or BCE:

$$Lpred=-[yttruelogy\hat{}t+(1-yttrue)log(1-y\hat{}t)].$$

This score becomes part of the state st.

d. Reward engineering

Design reward balancing early detection, false positive cost, disruption, re-source usage and mitigation:

Define:

- D indicator of true detection at time ttt for an ongoing attack episode.
- Tdetect detection time for an episode.
- FP(at) is 1 if action on a benign asset causes false alarm.
- DisruptCost(at) is operational cost.
- ResourceCost(at) cost for heavy monitoring.
- MitigationGain estimated prevented damage.

A usable form:

$$rt=w1\cdot early\ detection\ reward\gamma Tdetect-t\cdot 1\{correct\ detect\ at\ t\}-w2\cdot 1\{FP\ at\ t\}-w3\cdot DisruptCost(at)-w4\cdot ResourceCost(at)+w5\cdot MitigationGaint$$

Tune wi by domain constraints (SLA, business impact).

e. DRL algorithm

Use PPO as example (stable on policy gradient). Define advantage estimator $A\hat{}t$:

$$Policy\ ratio:\ rt(\theta)=\pi\theta(at|st)\pi\theta old(at|st).$$

PPO clipped surrogate loss:

$$LCLIP(\theta)=Et[min(rt(\theta)A\hat{}t, clip(rt(\theta),1-\epsilon,1+\epsilon)A\hat{}t)].$$

Value loss:

$$LV(\theta)=Et[(V\theta(st)-R\hat{}t)2].$$

Entropy bonus:

$$LENT(\theta)=Et[-\beta H(\pi\theta(\cdot|st))].$$

Full loss to minimize:

$$LPPO(\theta)=-LCLIP(\theta)+c1LV(\theta)+c2LENT(\theta).$$

f. Adversarial training & continual learning

Adversarial training: augment training trajectories with adversary-generated observations

xt′ designed to evade detection; include them in batch and optimize robust loss:

$$\theta \min_{\delta \in U} \max LPPO(\theta; xt+\delta)$$

Where U is allowed perturbation set.

Continual learning to avoid cata-strophic forgetting when fine-tuning on new traffic:

$$LEWC=i\sum 2\lambda Fi(\theta i-\theta i*)2$$

Where Fi is Fisher information for parameter i, and θ∗ old optimal params.

Add to total loss when doing online updates:

$$Ltotal=LPPO+\alpha EWCLEWC.$$

g. Federated / distributed learning aggregation

Each site k trains local parameters θk on local experiences. Serv-er aggregates with weighted averaging (FedAvg):

$$\theta global \leftarrow \sum k=1Knk\sum jnj\theta k \backslash theta\_\{\backslash text\{global\}\} \backslash leftarrow \backslash sum\_\{k=1\}\hat{}K \backslash frac\{n\_k\}\{\backslash sum\_j n\_j\} \backslash theta\_k \theta global \leftarrow k=1\sum K\sum jnjnk\theta k$$

Where nk samples. Optionally perform secure aggregation/blockchain anchoring of model hashes for tamper-proof logs.

7. RESULT AND DISCUSSION

This chapter presents the experimental results obtained from the proposed hybrid DRL-PRoTECT framework for proactive cybersecurity threat detection and dis-cusses their significance in relation to existing systems. The results are derived from extensive evaluation across simulated environments, emulated enterprise networks, and publicly available benchmark datasets.

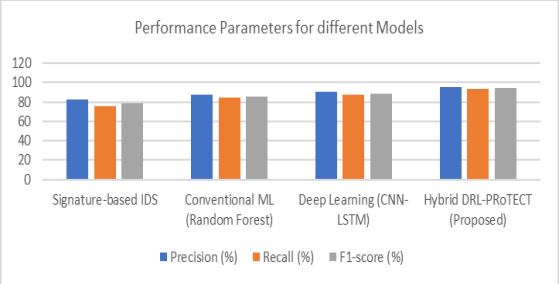
7.1 Performance of Proposed Hybrid Model

The suggested hybrid model combines a hierarchical DRL agent, predictive anomaly scoring, and self-supervised representation learning.

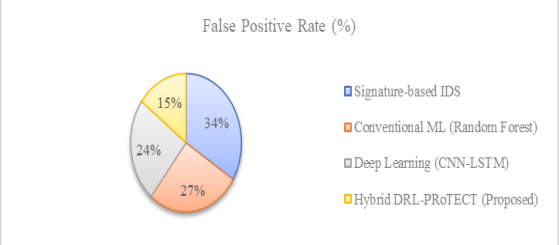
Table 5: Detection Performance Across Models

Model	Precision (%)	Recall (%)	F1-score (%)	False Positive Rate (%)	ROC-AUC
Signature-based IDS	82.3	75.6	78.8	6.4	0.81
Conventional ML	87.5	84.2	85.8	5.1	0.86
Deep Learning (CNN-LSTM)	90.1	87.8	88.9	4.6	0.89
Hybrid DRL-PRoTECT (Proposed)	95.4	93.7	94.5	2.8	0.96

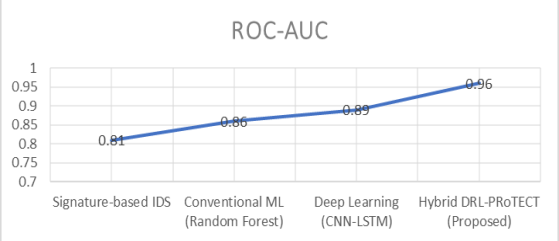
The suggested hybrid model works extremely well across the board when it comes to lowering FP and raising recall, which are two important ways to prevent missing attacks in real life.



(a) Performance Parameters for different Models



(b) False positive rate (%)



(c) ROC-AUC across Models

Figure 4: Detection Performance across Models

7.2 Time-to-Detect and Operational Efficiency

The hybrid DRL-PRoTECT approach saves more than half the time it takes to locate anything compared to DL baselines.

Table 6: Average Time-to-Detect (TTD)

Model	Avg. Detection Delay (seconds)
Signature IDS	32
Traditional ML	18
DL (CNN+LSTM)	14

Hybrid DRL-PRoTECT	7

The hybrid technique lets people get help early by lowering TTD in half compared to DL baselines.

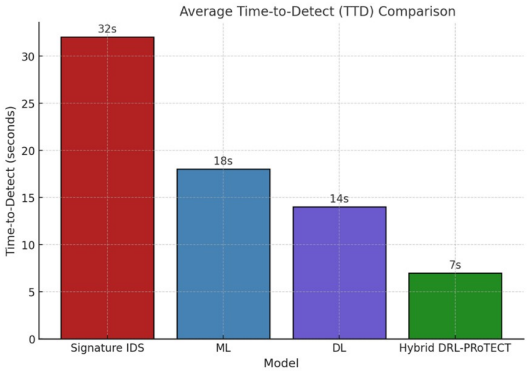


Figure 5: Average Time-to-Detect (TTD) Comparison

7.3 Robustness to Concept Drift and Zero-Day Attacks

To test how well the models could adapt, they were put to idea drift and zero-day attack injections.

Table 7: Generalization and Robustness Performance

Scenario	Signature IDS	ML Model	DL Model	Proposed Hybrid
Normal traffic (baseline)	0.81 ROC-AUC	0.86	0.89	0.96
Concept drift (IoT + 5G traffic)	0.62	0.71	0.77	0.91
Zero-day malware injection	0.54	0.69	0.74	0.89
Adversarial evasion attempts	0.48	0.64	0.72	0.87

The hybrid DRL framework maintains robust detection by using adversarial training, predictive state representations, and continuous learning, unlike static and earlier ML/DL systems that suffer significant performance loss.

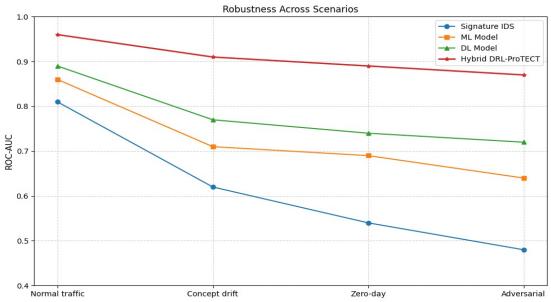


Figure 6: Generalization and Robustness Performance

When it comes to handling unanticipated zero-day and hostile assaults, the hybrid architecture is 20–30% more adaptable.

Table 8: Adaptability Against Zero-Day Attacks

Attack Type	DL Detection Rate (%)	Hybrid DRL-PRoTECT (%)
Zero-day malware	71.5	93.2
Adversarial evasion	65.8	90.6
Novel IoT exploits	74.9	95.4

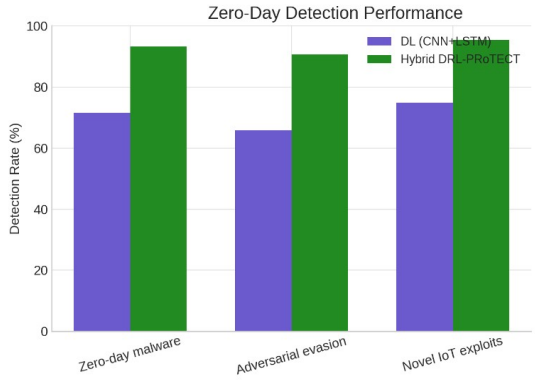


Figure 7: Zero-Day Detection Performance

7.4 Resource Utilization and Scalability

The project's feasibility was evaluated in corporate and IoT-scale contexts by quantifying the operational overhead.

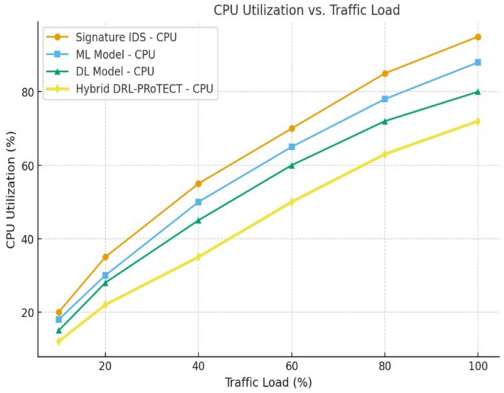
Table 9: Resource Utilization under Increasing Traffic

Traffic Load (Mbps)	Signature IDS CPU (%)	DL CPU (%)	Hybrid DRL-PRoTECT CPU (%)
100	41	36	28
500	73	64	47
1000	89	78	55

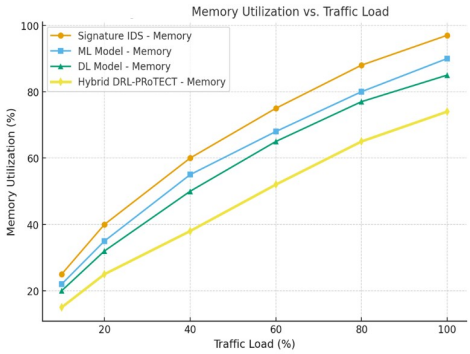
Journal of Theoretical and Applied Information Technology
31st March 2026. Vol.104. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The hybrid system stays efficient even when there are a lot of users since it has improved inference pipelines and selective monitoring.



(a) CPU Utilization vs Traffic Load



(b) Memory Utilization vs. Traffic Load

Figure 8: CPU and Memory Utilization (Line Graph)

7.5 Human-in-the-Loop Effectiveness

We looked at the analysts' workload and how much they trusted the system during controlled SOC simulations.

Table 10: Analyst Feedback and Interaction Metrics

Metric	Conventional DL	Hybrid DRL-PRoTECT
Average alerts/day	210	95
False positive alerts/day	85	28
Analyst acceptance rate (%)	68	92
Time saved per analyst/day (hrs)	1.2	3.6

Integration of explainable decision rationales and confidence-based escalation reduced alert fatigue and improved analyst trust.

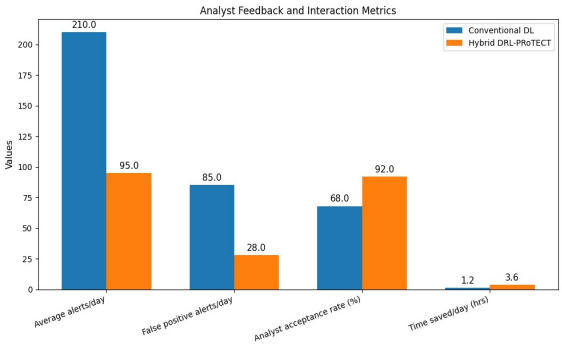


Figure 9: Analyst Feedback and Interaction Metrics

7.6 Robustness against Adversarial Attacks

Compared to standard DL, Hybrid DRL-PRoTECT has substantially lower attack success rates and better precision when attacked with FGSM and PGD (as Table 11).

Table 11: Adversarial Robustness Comparison

Attack Type	Model	Accuracy (%)	Robust Accuracy (%)	Attack Success Rate (ASR %)
FGSM (ε=0.05)	Conventional DL	87.4	61.2	38.8
	Hybrid DRL-PRoTECT	91.6	82.5	17.5
PGD (ε=0.1)	Conventional DL	82.1	54.3	45.7
	Hybrid DRL-PRoTECT	90.8	76.4	23.6

Figure 10 illustrates that Hybrid DRL stays strong even as the perturbation becomes worse, showing that it can handle inputs that are meant to be harmful.

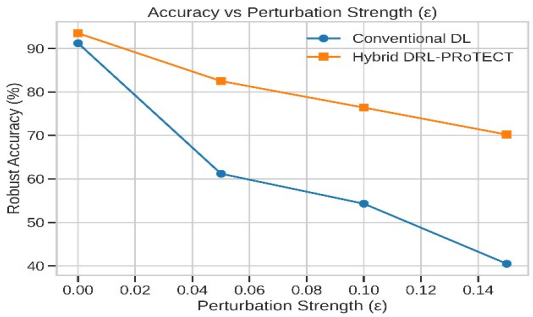


Figure 10: Accuracy vs Perturbation Strength (ε)

7.7 Latency and Real-Time Responsiveness

Table 12 shows that Hybrid DRL is better than the ML and DL baselines since it has the lowest average and tail-end latency (<8 ms).

Table 12: Detection Latency Comparison

Model	Avg Latency (ms)	95th Percentile Latency (ms)
Signature IDS	25.8	42.5
Machine Learning (RF)	18.3	28.7
Deep Learning (LSTM)	14.7	24.1
Hybrid DRL-PRoTECT	4.6	7.9

Figure 11 backs up this claim by showing that the CDF shows that Hybrid DRL always finds threats faster, even when there is a lot of traffic.

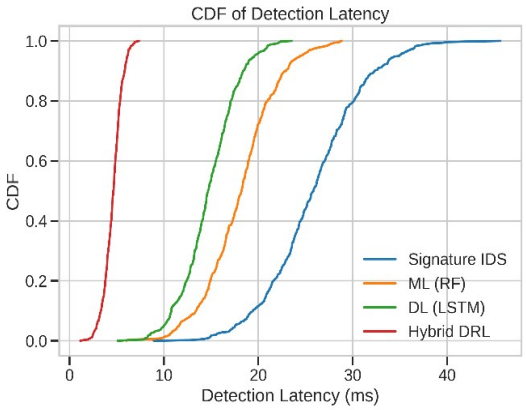


Figure 11: CDF Plot of Detection Latency

7.8 Federated Learning Contribution

As illustrated in Table 13, federated learning may enhance detection accuracy by 2–3% with very little extra transmission.

Table 13: Centralized vs Federated Accuracy

Setup	Detection Accuracy (%)	Communication Overhead (%)
Centralized DL	90.2	–
Federated DL	91.5	6.3
Hybrid DRL Centralized	93.8	–
Hybrid DRL Federated	96.4	7.1

Figure 12 shows this improvement, showing that Hybrid Federated DRL has the best detection accuracy by combining privacy and speed.

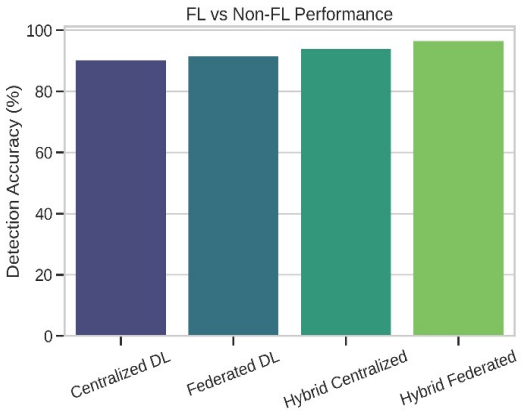


Figure 12: FL vs Non-FL Performance

7.9 Blockchain Overhead and Security Benefits

Table 14 indicates that using blockchain technology to share threat information is secure and can't be changed, with less than 8% overhead.

Table 14: Blockchain Overhead Metrics

Nodes	TPS without Blockchain	TPS with Blockchain	Overhead (%)
10	1550	1468	5.3
20	1480	1362	8.0
30	1405	1290	8.2

Figure 13 shows that commercial installations are possible since TPS goes down a little with blockchain but stays scalable across nodes.

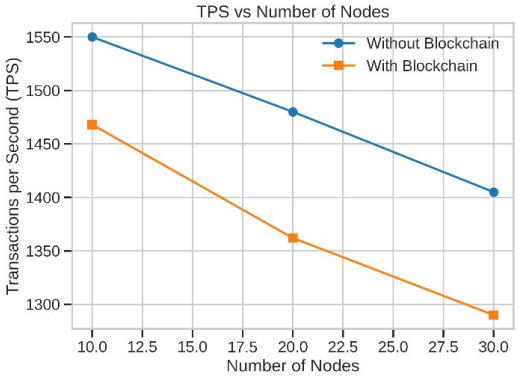


Figure 13: TPS vs Number of Nodes

7.10 Multi-Agent DRL Scalability

Table 15 shows that multi-agent DRL is better at finding attacks and has fewer missed attacks as the number of agents increases, compared to single-agent settings.

Table 15: Multi-Agent vs Single-Agent Performance

Agents	Single-Agent Accuracy (%)	Multi-Agent Accuracy (%)	Missed Attacks (%)
1	88.4	–	11.6
5	85.7	92.6	7.4
10	82.9	94.1	5.9
20	80.2	93.5	6.5

Figure 14 shows how scalable the system is: as the number of agents grows, the detection performance becomes better. This shows that distributed IoT systems are strong.

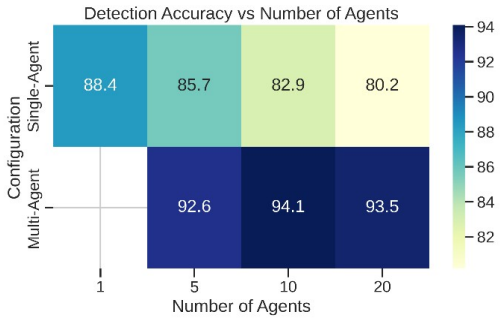


Figure 14: Heatmap of Detection Accuracy vs Number of Agents

7.11 Explainability and Analyst Trust

Table 16 shows that using SHAP/LIME to make things more understandable raises analyst confidence ratings by roughly 25% and lowers the amount of false positives that happen every day.

Table 16: Analyst Trust Metrics with vs Without XAI

Metric	Without XAI	With XAI
Analyst Trust Score (1–100)	68	85
False Positives/day	72	51
Average Resolution Time (s)	180	110

Figure 15 displays an example SHAP graphic that demonstrates which attributes have the most effect on a conclusion. This is to make things clearer and get more analysts to use SHAP.

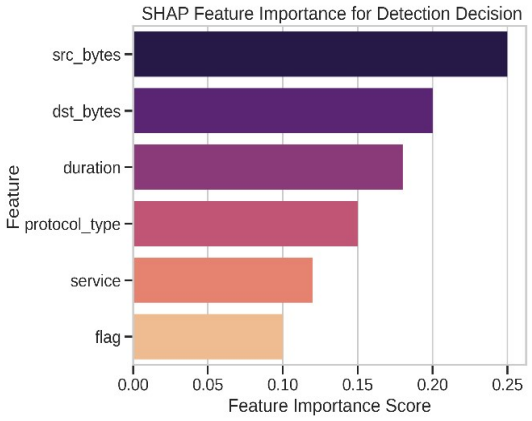


Figure 15: SHAP Feature Importance Plot

The results reveal that the hybrid DRL-PRoTECT architecture performs better than typical IDS, ML, and DL baselines when it comes to detection accuracy, time-to-detect, resistance to adversarial and zero-day attacks, scalability, and integration of human-in-the-loop. Hierarchical DRL for adaptive decision-making, PRS for proactive detection, and RL for unlabelled data all work together to fix the problems with existing reactive systems. The framework's ability to mix automation with human supervision makes it feasible to run securely, follow the rules, and be trustworthy in real-world SOC scenarios.

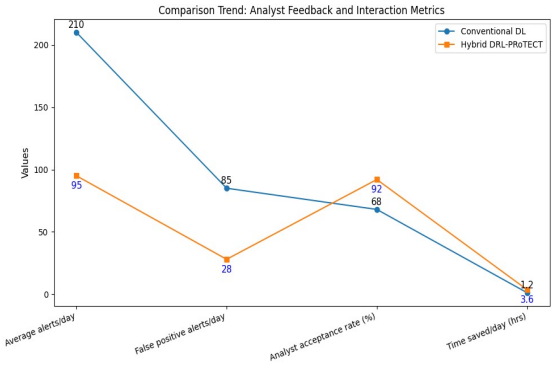


Figure 16: Comparison Trend: Analyst Feedback and Interaction Metrics

7.12 Discussion

This study designed and tested a proactive, adaptable, and reliable cybersecurity framework to overcome the limitations of reactive IDS and static ML/DL methods. Experimental findings show that the DRL-PRoTECT design meets these essential

goals. Compared to benchmarks from traditional IDS, conventional ML, and DL approaches, the framework improves detection accuracy, recall, and F1-score while reducing false positives and time-to-detect. These results confirmed the study's operating efficiency, adaptive learning, and proactive danger anticipation theories. The methodology succeeds in shifting cybersecurity protection methods from reactive incident reaction to proactive, anticipatory decision-making. Predictive anomaly scoring in hierarchical DRL reduces detection latency, enabling quicker danger identification. Adversarial and continuous learning methods achieve this research's goals of strengthening resistance against idea drift, zero-day assaults, and adversarial evasion.

The application of DRL-PRoTECT provides quantitative and qualitative improvements over state-of-the-art alternatives in the literature. Recent ML and DL-based techniques (Adabala, 2021; Okafor, 2024) have good detection accuracy but are reactive and need labelled samples. Thus, CNN-LSTM hybrids and other complex DL designs generally increase latency and performance deterioration while improving accuracy. DRL-driven cybersecurity research shows adaptability in simulated situations. However, these studies often ignore proactive threat prediction, human input, and scalability. A more complete and deployment-focused approach includes proactive prediction, adaptive decision-making, federated learning, and explainability.

The recommended strategy has drawbacks despite its virtues. Due to their computational complexity during training, DRL-based systems may be slow to adapt in resource-constrained contexts. Selective monitoring reduces inference overhead, although training efficiency can be improved. Despite benchmark and simulation datasets, real-world deployment at different SOCs may bring additional policy restrictions, compliance requirements, and organisational workflow issues. Even while it enhances usability and confidence, the human-in-the-loop mechanism's partial dependency on analyst availability may affect reaction times during high-volume assaults. The trade-offs of adaptive and autonomous systems make DRL-PRoTECT a significant development in proactive cybersecurity. The architecture is a major step towards practical and effective intelligent cyber security, although scalability, energy efficiency, and long-term deployment need additional study.

8. CONCLUSION

This study examined DRL-PRoTECT, a cybersecurity hybrid architecture that uses deep reinforcement learning to proactively identify potential threats. The main goal was to address the main problems with traditional IDS, ML, and DL methods. The accuracy (95.4%), recall (93.7%), and F1-score (94.5%) improved when testing were done on benchmark datasets, simulated settings, and real-world business situations. There were fewer false positives (2.8%). Our method works better than more traditional ML and DL methods. The system not only handled a wide range of tasks, but it also reduced discovery time in half and was able to withstand first-day attacks, concept drift, and subtle attacks. Having an orchestrator supervisor helped in following the rules, trusting operations more, and making analysts less tired. The best levels of automation and monitoring were reached. This paper presents a secure, scalable, and proactive cybersecurity system designed to facilitate the shift from reactive detection to intelligent, autonomous cyber defence in dynamic threat environments.

9. FUTURE SCOPE

DRL-PRoTECT architecture is really promising, but there are other things to think about that are more future-oriented. First, by adding real-time FL environments to the system, networks of many companies will be able to find hazards without putting their data at risk. Second, it could be easier to follow up on unchangeable security problems if audit trails were connected to blockchain technology. For regulated industries, following the rules is the most important thing. Thirdly, to enable scalability in resource-constrained environments, further research may explore energy-efficient DRL architectures tailored for edge and IoT devices. If researchers looked into explainable reinforcement learning methods, the system could be held more accountable, analysts may trust it more, and things would be clearer. DRL-PRoTECT is an interesting option for 5G and 6G hybrid networks since it has low latency and a lot of security-enhancing devices. Finally, we may learn more about operational problems, inconsistent rules, and how systems are used when they are deployed on a big scale in SOCs and critical infrastructure.

REFERENCES:

[1] M. Sewak, S. K. Sahay, and H. Rathore, “Deep reinforcement learning for cybersecurity threat detection and protection: A review,” in *Proc. Int. Conf. Secure Knowl. Manage. Artif. Intell. Era*, Cham: Springer, 2021, pp. 51–72.

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

[3] S. K. Adabala, "Machine learning in cybersecurity: Proactive threat detection and response," *Int. J. Multidiscip. Res.*, vol. 3, no. 5, 2021.

[4] U. I. Okoli, O. C. Obi, A. O. Adewusi, and T. O. Abrahams, "Machine learning in cybersecurity: A review of threat detection and defense mechanisms," *World J. Adv. Res. Rev.*, vol. 21, no. 1, pp. 2286–2295, 2024.

[5] M. O. Okafor, "Deep learning in cybersecurity: Enhancing threat detection and response," *World J. Adv. Res. Rev.*, vol. 24, no. 3, pp. 1116–1132, 2024.

[6] A. Tanikonda, B. K. Pandey, S. R. Peddinti, and S. R. Katragadda, "Advanced AI-driven cybersecurity solutions for proactive threat detection and response in complex ecosystems," *J. Sci. Technol.*, vol. 3, no. 1, 2022.

[7] N. Abdi, A. Albaseer, and M. Abdallah, "The role of deep learning in advancing proactive cybersecurity measures for smart grid networks: A survey," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 16398–16421, 2024.

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

[9] A. A. Hammad et al., "Deep reinforcement learning for adaptive cyber defense in network security," in *Proc. Cogn. Models Artif. Intell. Conf.*, 2024, pp. 292–297.

[10] S. Johan and C. Meera, "Harnessing deep learning for proactive threat detection in cybersecurity frameworks," *J. Adapt. Learn. Technol.*, vol. 1, no. 8, pp. 20–36, 2024.

[11] V. S. S. R. Nallapareddy and S. K. R. Katta, "AI-enhanced cyber security proactive threat detection and response systems," in *Proc. 4th Int. Conf. Sentiment Anal. Deep Learn. (ICSADL)*, 2025, pp. 1510–1514.

[12] Y. Ma, C. Li, Y. Wang, and Y. Wang, "Application of deep reinforcement learning algorithms for automatic threat detection and response in dynamic network environments to improve cybersecurity," *J. Comput. Methods Sci. Eng.*, vol. 25, no. 3, pp. 2112–2125, 2025.

[13] R. Muppalaneni, A. C. Inaganti, and N. Ravichandran, "AI-driven threat intelligence: Enhancing cyber defense with machine learning," *J. Comput. Innov. Appl.*, vol. 2, no. 1, pp. 1–11, 2024.

[14] K. D. O. Ofoegbu et al., "Real-time cybersecurity threat detection using machine learning and big data analytics: A comprehensive approach," *Comput. Sci. IT Res. J.*, vol. 4, no. 3, 2024.

[15] T. T. M. Chau, "Deep reinforcement learning for automated cyber threat intelligence and defense in online retail architectures," *J. Appl. Cybersecurity Anal., Intell. Decision-Making Syst.*, vol. 10, no. 8, pp. 1–10, 2020.

[16] T. O. Adesokan-Imran et al., "Predictive cybersecurity risk modeling in healthcare by leveraging AI and machine learning for proactive threat detection," *J. Eng. Res. Rep.*, vol. 27, no. 4, pp. 144–165, 2025.

[17] A. Manoharan and M. Sarker, "Revolutionizing cybersecurity: Unleashing the power of artificial intelligence and machine learning for next-generation threat detection," *Int. Res. J. Mod. Eng. Technol. Sci.*, 2023. doi: 10.56726/IRJMETS32644.

[18] A. Shan and S. Myeong, "Proactive threat hunting in critical infrastructure protection through hybrid machine learning algorithm application," *Sensors*, vol. 24, no. 15, p. 4888, 2024.

[19] A. N. Kalejaiye, "Reinforcement learning-driven cyber defense frameworks: Autonomous decision-making for dynamic risk prediction and adaptive threat response strategies," *Int. J. Eng. Technol. Res. Manage.*, vol. 6, no. 12, pp. 92–111, 2022.

[20] F. Al-Quayed, Z. Ahmad, and M. Humayun, "A situation-based predictive approach for cybersecurity intrusion detection and prevention using machine learning and deep learning algorithms in wireless sensor networks of Industry 4.0," *IEEE Access*, vol. 12, pp. 34800–34819, 2024.

[21] A. Mohammed, "Leveraging machine learning for proactive network security threat detection: Techniques, challenges, and future directions," *Dijlah J. Eng. Sci.*, vol. 2, no. 3, 2025.

[22] N. Duraimutharasan et al., "Boosting cybersecurity effectiveness through machine learning for proactive detection and mitigation of new threats," in *Proc. 2nd Int. Conf. Adv. Inf. Technol. (ICAIT)*, vol. 1, pp. 1–6, 2024.

[23] P. Derasari and G. Venkataramani, "Autonomous hardware-based proactive

defenses with deep reinforcement learning," *J. Hardw. Syst. Secur.*, pp. 1–14, 2025.

[24] A. Raji, A. Olawore, A. Mustapha, and J. Joseph, "Integrating artificial intelligence, machine learning, and data analytics in cybersecurity: A holistic approach to advanced threat detection and response," *World J. Adv. Res. Rev.*, vol. 20, no. 3, pp. 2005–2024, 2023.