Autonomous Cyber Defense Learning Using Reinforcement and Threat Intelligence

Abimbola B. Owolabi[1] & Francis B. Osang[2]

[1&2]National Open University of Nigeria.
Corresponding Authors' Email: aowolabi@noun.edu.ng & fosang@noun.edu.ng

Direct Research Journal of Engineering and Information Technology

Vol. 14(3), Pp. 10-25, August 2026,

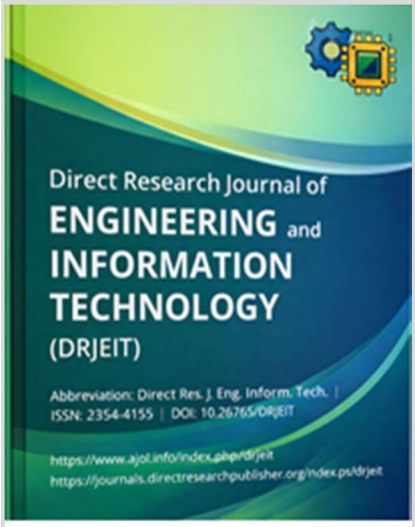
Research Article
ISSN: 2354-4155
https://journals.directresearchpublisher.org/index.php/drjeit; https://www.ajol.info/index.php/drjeit

Received 27 June 2026, Accepted 20 July 2026, Published 21 August 2026

ABSTRACT

The increasing sophistication, frequency, and scale of cyberattacks have created significant challenges for conventional cybersecurity systems. Traditional security solutions such as firewalls, signature-based intrusion detection systems, and antivirus software are largely reactive and depend on predefined rules and known attack patterns. Consequently, these systems often struggle to detect and respond effectively to emerging threats such as Advanced Persistent Threats (APTs), zero-day attacks, ransomware, botnets, and insider attacks. Recent advancements in Artificial Intelligence (AI), particularly Reinforcement Learning (RL), have demonstrated the potential to create autonomous systems capable of learning and adapting to dynamic environments. Simultaneously, Cyber Threat Intelligence (CTI) provides valuable contextual information regarding threat actors, attack techniques, vulnerabilities, and indicators of compromise. This study proposes an Autonomous Cyber Defense Framework that integrates Reinforcement Learning and Threat Intelligence to enhance threat detection, decision-making, and automated response capabilities. The framework employs a Deep Q-Network (DQN) agent that continuously learns optimal defense actions through interaction with network environments while utilizing threat intelligence feeds to improve situational awareness. Experimental evaluation was conducted using benchmark cybersecurity datasets, including CICIDS2017 for Intrusion Detection, UNSW-NB15 for attack classification, CTU-13 for botnet detection and Custom Threat Feeds for threat intelligence. The results indicate that the proposed framework achieved a precision rate of 98.4%, a recall rate of 98.2%, an F1-score of 98.3%, and a threat mitigation rate of 96.8%. False positive rate of 1.9, False negative rate of 1.5 and Response rate of 41%, significantly outperforming traditional machine learning and signature-based security approaches. The findings demonstrate that integrating reinforcement learning with threat intelligence can provide a highly adaptive and proactive cyber defense mechanism suitable for modern network environments.

Keywords: *Autonomous Cyber Defense, Reinforcement Learning, Cyber Threat Intelligence, Deep Q-Network, Artificial Intelligence, Intrusion Detection Systems.*



Citation: Owolabi A.B. and Osang F.B. (2026). Autonomous Cyber Defense Learning Using Reinforcement and Threat Intelligence. *Direct Research Journal of Engineering and Information Technology*. 14(3), Pp. 10-25. https://doi.org/10.4314/drjeit.v14i3.2

INTRODUCTION

The unprecedented growth of digital technologies has fundamentally reshaped modern communication, business operations, public services, and critical infrastructure. Innovations such as cloud computing, the Internet of Things (IoT), fifth- and sixth-generation (5G/6G) communication networks, mobile computing, and large-scale interconnected information systems have accelerated digital transformation across virtually every sector (Abouhawwash, 2024). Although these technological developments have improved connectivity, automation, and service delivery, they have simultaneously expanded the attack surface available to cyber adversaries.

As organizations become increasingly dependent on interconnected digital ecosystems, vulnerabilities within these environments provide attractive opportunities for malicious actors to launch attacks that compromise data confidentiality, system integrity, and service availability (Shi et al., 2025). The cybersecurity landscape has undergone a profound transformation in recent years. Earlier forms of cyberattacks largely relied on known malware signatures, predictable exploitation techniques, and relatively simple attack vectors. However, contemporary cyber threats have become considerably more sophisticated, adaptive, and persistent. Modern attackers increasingly employ advanced techniques such as ransomware-as-a-service, polymorphic and metamorphic malware, fileless attacks, Advanced Persistent Threats (APTs), botnet-driven campaigns, and artificial intelligence-assisted attack strategies. These threats continuously evolve their behavior to evade conventional detection mechanisms, making them considerably more difficult to identify and mitigate using traditional cybersecurity solutions (Alam et al., 2026).

Most existing security infrastructures, including firewalls, antivirus software, and signature-based intrusion detection systems, operate using predefined signatures, static rules, or manually configured security policies. While these approaches remain effective against previously identified threats, they often perform poorly when confronted with novel or rapidly evolving attack techniques (Ahmed et al., 2026). Their dependence on historical attack signatures limits their ability to recognize zero-day exploits and other emerging threats. Furthermore, conventional security systems frequently generate excessive false positives, increasing the workload of cybersecurity analysts, delaying incident response, and reducing the overall efficiency of security operations. The growing scale, diversity, and sophistication of cyber threats therefore necessitate the development of intelligent defense mechanisms capable of learning from changing environments and responding autonomously to emerging attacks (Balasubramanian et al., 2025).

Artificial Intelligence (AI) has emerged as a transformative technology for enhancing modern cybersecurity capabilities. By enabling systems to analyse complex data, recognize behavioural patterns, and make intelligent decisions, AI offers significant advantages over traditional rule-based security mechanisms. Among the various AI paradigms, Reinforcement Learning (RL) has attracted considerable attention because of its ability to learn optimal decision-making strategies through continuous interaction with dynamic environments (Apruzzese, G. et al., 2018). Unlike supervised learning approaches that require extensive labelled datasets, reinforcement learning improves its performance through iterative feedback obtained from environmental rewards and penalties. This adaptive learning capability makes RL particularly suitable for cybersecurity applications, where attack behaviours evolve continuously and labelled attack data are often scarce or unavailable (Alam et al., 2026). Alongside advances in artificial intelligence, Cyber Threat Intelligence (CTI) has become an indispensable element of proactive cybersecurity operations. CTI provides structured and actionable knowledge regarding cyber adversaries, including indicators of compromise, attacker tactics, techniques and procedures (TTPs), emerging vulnerabilities, malware characteristics, threat actor profiles, and real-time intelligence feeds. Rather than relying solely on reactive detection, threat intelligence enables security systems to anticipate potential attacks, strengthen situational awareness, and prioritize defensive actions based on current threat conditions (Shi et al., 2025). When integrated with intelligent learning algorithms, CTI provides contextual information that significantly enhances the accuracy and effectiveness of autonomous cyber defense systems.

Motivated by these challenges, this research proposes an integrated Autonomous Cyber Defense Framework that combines Deep Reinforcement Learning with Cyber Threat Intelligence to provide adaptive, proactive, and intelligent protection against evolving cyber threats. The proposed framework utilizes a Deep Q-Network (DQN) agent to continuously learn optimal defensive strategies through interaction with network environments while simultaneously incorporating threat intelligence to improve situational awareness and decision quality. By integrating adaptive learning with contextual threat information, the framework is designed to detect, analyse, prioritize, and mitigate cyber threats in real time with minimal human intervention. Ultimately, the proposed approach aims to enhance threat detection accuracy, accelerate incident response, improve resilience against sophisticated attacks, and strengthen the overall security posture of modern networked environments (Arıkan et al., 2024).

Statement of the Problem

The rapid growth of interconnected digital technologies has significantly increased the frequency, complexity, and sophistication of cyberattacks. Conventional cybersecurity solutions, including firewalls, antivirus software, and signature-based intrusion detection systems, primarily rely on predefined rules and known attack signatures. Although these approaches are effective against previously identified threats, they are often unable to detect emerging attacks such as zero-day exploits, Advanced Persistent Threats (APTs), polymorphic malware, and ransomware, which continuously evolve to evade traditional security mechanisms. Another major challenge is the dependence on manual threat analysis and response. The increasing volume of security alerts frequently overwhelms cybersecurity professionals, resulting in delayed responses, high false positive rates, and reduced operational efficiency. While machine learning techniques have improved threat detection, many existing approaches focus primarily on identifying attacks without providing intelligent, adaptive, and automated response capabilities

(Buczak, et al., 2016). Furthermore, current cybersecurity systems often fail to effectively integrate Cyber Threat Intelligence (CTI), limiting their ability to utilize real-time information on attacker behaviors, vulnerabilities, indicators of compromise, and emerging threats. The absence of contextual threat intelligence reduces situational awareness and weakens proactive defense against evolving cyber risks (Buczak & Guven, 2016). These limitations highlight the need for an autonomous cyber defense framework that combines Deep Reinforcement Learning with Cyber Threat Intelligence to enable intelligent threat detection, adaptive decision-making, and automated threat mitigation. Such a framework can enhance cyber resilience by providing proactive, real-time protection against sophisticated and rapidly evolving cyber threats (Cohen D. et al., 2025).

Aim of the Study

The primary aim of this research is to design and develop an intelligent autonomous cyber defense framework that integrates Deep Reinforcement Learning and Cyber Threat Intelligence to enable adaptive threat detection, contextual threat analysis, and automated mitigation of cyberattacks within dynamic and heterogeneous network environments. The objectives of the Study are as follows:

1. Design a comprehensive autonomous cyber defense architecture that integrates Deep Reinforcement Learning with Cyber Threat Intelligence for intelligent cybersecurity operations.
2. Integrate real-time Cyber Threat Intelligence feeds, including indicators of compromise, threat actor profiles, vulnerability information, and attacker tactics, techniques, and procedures, into the autonomous decision-making process.
3. Implement an adaptive threat response mechanism capable of automatically detecting, analysing, prioritizing, and mitigating cyber threats with minimal human intervention.
4. Evaluate the effectiveness of the proposed framework using standard benchmark cybersecurity datasets, including CICIDS2017 and UNSW-NB15, based on performance metrics such as detection accuracy, precision, recall, F1-score, response time, and threat mitigation rate.
5. Compare the performance of the proposed autonomous cyber defense framework with conventional machine learning models and traditional signature-based cybersecurity solutions to determine its effectiveness, adaptability, and operational efficiency.

Literature Review

The continuous advancement of digital technologies and the widespread adoption of interconnected computing environments have revolutionized communication, data sharing, automation, and service delivery across numerous sectors. While these developments have enhanced operational efficiency and global connectivity, they have also expanded the attack surface available to cybercriminals, leading to an increase in the frequency and sophistication of cyber threats. Conventional cybersecurity solutions, which predominantly depend on signature-based detection methods, predefined security rules, and static defense mechanisms, have become less effective in countering rapidly evolving and previously unseen attacks. To address these limitations, researchers have increasingly investigated the application of Artificial Intelligence (AI), Machine Learning (ML) (Mitshel R et al., 2014), Reinforcement Learning (RL), and Cyber Threat Intelligence (CTI) as intelligent technologies for strengthening cyber defense. These approaches offer the capability to learn from dynamic environments, identify emerging attack patterns, and support adaptive, automated decision-making. This literature review critically examines existing research on autonomous cyber defense, reinforcement learning, cyber threat intelligence, intelligent intrusion detection, and AI-enabled security frameworks to identify current advancements, research gaps, and opportunities for developing more resilient cybersecurity solutions (Buczak & Guven, 2016).

Evolution of Cybersecurity Threats

Cybersecurity threats have evolved considerably over the past few decades. Early cyberattacks were relatively simple and often targeted specific vulnerabilities within computer systems. However, the increasing sophistication of threat actors has resulted in the emergence of complex attack techniques such as Advanced Persistent Threats (APTs), ransomware, botnets, zero-day exploits, phishing campaigns, insider attacks, and distributed denial-of-service (DDoS) attacks (Nguyen et al., 2021).

According to recent cybersecurity reports, cybercriminals increasingly employ automation, artificial intelligence, and advanced evasion techniques to bypass traditional security controls. Advanced Persistent Threats are particularly dangerous because they involve prolonged and stealthy infiltration of targeted systems, enabling attackers to exfiltrate sensitive information while remaining undetected for extended periods. Similarly, ransomware attacks have become one of the most significant cybersecurity concerns due to their ability to encrypt critical organizational data and demand substantial financial payments for recovery (Buczak, A. L., & Guven, E. 2016). The dynamic nature of these threats has exposed significant weaknesses in conventional security solutions, highlighting the need for adaptive and intelligent defense mechanisms capable of responding to emerging attack patterns in real time (Saeed S. et al., 2024).

Traditional Cyber Defense Mechanisms

Traditional cyber defense mechanisms form the

foundation of modern cybersecurity infrastructures. These mechanisms include firewalls, intrusion detection systems (IDS), intrusion prevention systems (IPS), antivirus software, access control systems, and security information and event management (SIEM) platforms (Sommer R. et al., 2021). Firewalls serve as the first line of defense by filtering incoming and outgoing network traffic based on predefined security rules. Intrusion Detection Systems monitor network activities and generate alerts when suspicious behavior is detected. Signature-based IDS solutions compare observed traffic patterns against known attack signatures stored in databases. While effective against previously identified threats, these systems are incapable of detecting novel attacks that do not match existing signatures (Shameli-Sendi, A., et al, 2016).

Anomaly-based intrusion detection systems attempt to address this limitation by identifying deviations from normal system behavior. However, these systems often suffer from high false positive rates, leading to excessive alert generation and increased workload for security analysts (Taddeo M. et al., 2018). Although traditional cyber defense mechanisms remain important components of organizational security architectures, their reactive nature limits their effectiveness against modern cyber threats. As attack techniques continue to evolve, cybersecurity researchers have increasingly focused on intelligent approaches capable of learning and adapting autonomously (Chen et al., 2019).

Artificial Intelligence in Cybersecurity

Artificial Intelligence has emerged as a transformative technology within the cybersecurity domain. AI enables systems to process large volumes of data, identify complex patterns, and make intelligent decisions with minimal human intervention. AI-driven cybersecurity solutions have demonstrated significant potential in areas such as threat detection, malware analysis, vulnerability assessment, incident response, and cyber threat intelligence (Tavallaee et al., 2009).

Machine learning algorithms are among the most widely used AI techniques in cybersecurity. These algorithms can analyze historical data to identify patterns associated with malicious activities. Supervised learning methods such as Support Vector Machines, Decision Trees, Random Forests, and Neural Networks have been extensively applied for intrusion detection and malware classification (Chen et atl., 2019). Unsupervised learning techniques have also gained attention for their ability to identify previously unseen attack patterns without labeled datasets. Clustering algorithms, principal component analysis, and autoencoders have been employed to detect anomalies within network traffic and user behavior.

Despite their effectiveness, many machine learning approaches operate in static environments and require periodic retraining to keep pace with evolving threat landscapes. This limitation has motivated the exploration of reinforcement learning as a more adaptive alternative for cybersecurity applications (Singh et al., 2024).

Reinforcement Learning and Its Principles

Reinforcement Learning is a branch of machine learning in which an intelligent agent learns optimal behavior through interaction with its environment. Unlike supervised learning, which relies on labeled training data, reinforcement learning enables agents to discover effective strategies through trial-and-error experiences (Dutta et al., 2023).

The reinforcement learning framework consists of four primary components: the environment, the agent, actions, and rewards. The agent observes the current state of the environment and selects an action based on its learned policy. The environment responds to the selected action by transitioning to a new state and providing a reward signal. Through repeated interactions, the agent learns to maximize cumulative rewards by selecting actions that yield favorable outcomes (Wang W. et al., 2017).

The ability of reinforcement learning to adapt continuously makes it particularly suitable for cybersecurity applications. Cybersecurity environments are highly dynamic, with attack patterns constantly evolving. Reinforcement learning enables defense systems to adjust their strategies automatically without requiring extensive human intervention or retraining (Dutta et al., 2023).

Recent advancements in Deep Reinforcement Learning, particularly Deep Q-Networks (DQN), have significantly improved the ability of RL agents to operate in complex environments characterized by large state spaces and uncertain conditions.

Applications of Reinforcement Learning in Cybersecurity

The application of reinforcement learning in cybersecurity has attracted substantial research attention in recent years. Researchers have explored RL-based approaches for intrusion detection, malware analysis, network defense, cyber deception, access control management, and automated incident response.

One significant application involves adaptive intrusion detection systems that continuously learn from network activities and modify detection strategies based on observed attack patterns. Unlike traditional IDS solutions, RL-based systems can improve their performance over time by incorporating feedback from previous detection outcomes (Han et al., 2018).

Researchers have also investigated the use of reinforcement learning for cyber deception strategies. Cyber deception involves misleading attackers through the deployment of honeypots, decoy systems, and fake network resources. RL agents can dynamically adjust deception tactics to maximize attacker engagement while minimizing risks to legitimate systems. Additionally, reinforcement learning has been applied to

automated incident response systems. These systems autonomously determine appropriate mitigation actions, such as blocking malicious traffic, isolating compromised devices, or updating firewall configurations. Experimental studies have demonstrated that RL-based defense systems can significantly reduce response times and improve overall security effectiveness (Gueriani et al., 2024).

Despite these advancements, many existing RL-based cybersecurity solutions focus primarily on local network observations and do not incorporate external threat intelligence sources. This limitation restricts their situational awareness and ability to anticipate emerging threats.

Cyber Threat Intelligence

Cyber Threat Intelligence refers to the collection, analysis, and dissemination of information regarding cyber threats, threat actors, vulnerabilities, attack methodologies, and indicators of compromise. CTI aims to provide actionable insights that enable organizations to understand and respond effectively to evolving cyber risks (Purves, 2024)
Threat intelligence is generally categorized into four levels: strategic intelligence, operational intelligence, tactical intelligence, and technical intelligence. Strategic intelligence provides high-level information regarding threat trends and adversary motivations. Operational intelligence focuses on specific attack campaigns and threat actor activities. Tactical intelligence examines attack methodologies and procedures, while technical intelligence includes indicators such as malicious IP addresses, domain names, malware hashes, and exploit signatures (Khraisat et al., 2019).

Modern cybersecurity operations increasingly rely on threat intelligence feeds obtained from commercial providers, government agencies, open-source intelligence platforms, and information-sharing communities. The integration of CTI into security systems enhances contextual awareness and enables organizations to anticipate potential attacks before they occur.
Researchers have demonstrated that threat intelligence significantly improves threat detection accuracy by supplementing internal security data with external contextual information. Consequently, CTI has become an essential component of advanced cybersecurity architectures (Khraisat et al., 2019).

Autonomous Cyber Defense Systems

Autonomous cyber defense represents a paradigm shift from traditional human-centered cybersecurity operations toward intelligent systems capable of independent threat detection, analysis, decision-making, and response. Autonomous defense systems leverage artificial intelligence, machine learning, and automation technologies to reduce dependence on human intervention (Kim et al., 2014). The primary objective of autonomous cyber defense is to enable real-time adaptation to evolving threat environments. Such systems continuously monitor network activities, identify suspicious behavior, assess risks, and implement mitigation strategies without requiring manual input from security analysts.

Several studies have demonstrated the effectiveness of autonomous defense systems in reducing incident response times and improving threat mitigation rates. Autonomous systems are particularly valuable in large-scale environments where the volume of security events exceeds the capacity of human analysts (Thompson et al., 2024).

However, the development of fully autonomous cyber defense systems remains challenging due to issues related to decision reliability, explainability, scalability, and adversarial manipulation. Researchers continue to investigate methods for improving the robustness and transparency of autonomous cybersecurity solutions (Kim et al., 2014).

Integration of Reinforcement Learning and Threat Intelligence

The integration of Reinforcement Learning and Cyber Threat Intelligence has emerged as a promising research direction for enhancing autonomous cyber defense capabilities. Reinforcement learning provides adaptive decision-making and continuous learning capabilities, while threat intelligence contributes contextual awareness and knowledge regarding emerging threats (Landolt et al., 2025).

Several recent studies have explored hybrid frameworks that combine machine learning with threat intelligence for intrusion detection and threat prediction. These studies have demonstrated improvements in detection accuracy and situational awareness. However, relatively few investigations have focused specifically on the integration of reinforcement learning and threat intelligence for autonomous cyber defense (Landolt et al., 2025). Researchers argue that combining RL and CTI enables defense systems to anticipate threats proactively rather than merely reacting to observed attacks. Threat intelligence informs the RL agent about known threat indicators and adversary behaviors, while reinforcement learning determines optimal response strategies based on environmental conditions and previous experiences. This synergy creates a more resilient and adaptive cybersecurity framework capable of addressing both known and unknown threats. The integration of RL and CTI, therefore, represents a significant advancement toward achieving fully autonomous cyber defense systems (Landolt et al., 2025).

Research Gap

Despite significant progress in artificial intelligence-driven

cybersecurity, several limitations remain within existing research. Many traditional intrusion detection systems continue to rely on static signatures and predefined rules, making them ineffective against emerging threats. Although machine learning approaches have improved threat detection capabilities, most solutions focus primarily on classification tasks without incorporating adaptive response mechanisms (Han et al., 2018).

Furthermore, existing reinforcement learning-based cybersecurity systems often rely solely on local network observations and fail to leverage external cyber threat intelligence effectively. This limitation reduces their ability to anticipate and respond proactively to evolving attack campaigns. Additionally, many autonomous defense systems lack comprehensive mechanisms for integrating threat intelligence into the reinforcement learning decision-making process (Dutta et al., 2023).

There is therefore a need for a comprehensive framework that combines reinforcement learning and cyber threat intelligence to provide adaptive threat detection, intelligent decision-making, and autonomous response capabilities. The proposed study addresses this gap by developing an Autonomous Cyber Defense Framework that integrates Deep Reinforcement Learning and Cyber Threat Intelligence to enhance cybersecurity resilience in dynamic network environments (Kim at al., 2014).

METHODOLOGY

Research Design

This study adopted an **experimental research design** to develop, implement, and empirically evaluate an autonomous cyber defense framework that integrates Deep Reinforcement Learning (DRL) with Cyber Threat Intelligence (CTI). The methodology was structured to enable the proposed system to continuously observe network security states, incorporate contextual threat intelligence, make autonomous defense decisions, execute appropriate mitigation actions, and learn from the outcomes of those actions.

The experimental process consisted of five major phases: data acquisition and preparation, CTI acquisition and contextualization, autonomous cyber defense architecture development, DRL agent training and adaptive response, and performance evaluation and comparative analysis. CICIDS2017 and UNSW-NB15 were used as the principal benchmark datasets. CICIDS2017 contains benign and contemporary attack traffic represented through network-flow information, while UNSW-NB15 contains normal traffic and nine attack categories generated in a controlled cyber-range environment. The architecture consists of interconnected modules designed to transform raw network observations and external threat intelligence into autonomous defense decisions.

The framework receives network-flow information from the benchmark datasets and combines it with CTI information containing indicators of compromise (IoCs), threat actors, vulnerabilities, and adversarial Tactics, Techniques, and Procedures (TTPs). The CTI information is represented using structured attributes corresponding to concepts such as IP addresses, domains, hashes, vulnerabilities, threat actors, attack techniques, confidence levels, and severity. STIX is particularly suitable for representing and exchanging structured CTI, while TAXII provides a standardized mechanism for exchanging CTI over HTTPS. (Figure 1)

Conceptual Design

The conceptual design of the proposed framework illustrates the logical flow of information and decision-making processes within the autonomous cyber defense system. The framework is based on the premise that effective cyber defense requires continuous observation of network activities, intelligent analysis of threat information, adaptive learning capabilities, and automated response mechanisms

The conceptual model begins with the network environment, which serves as the primary source of cybersecurity data. The network environment comprises users, servers, workstations, cloud resources, IoT devices, applications, and communication infrastructures. These components continuously generate network traffic, system events, authentication logs, and security alerts. The generated data is transmitted to the data collection module, where information is gathered from multiple monitoring sources. The preprocessing module then processes the collected data to remove inconsistencies and extract meaningful features relevant to cyber threat detection (Sutton, R. S., & Barto, A. G., 2018).

The processed information is subsequently combined with external cyber threat intelligence data. This integration enriches the collected network data by providing contextual insights into known attack patterns, malicious actors, vulnerabilities, and emerging threats. The combination of internal network observations and external intelligence creates a comprehensive security knowledge base that enhances situational awareness (Sutton et al., 2018).

The enriched data is then forwarded to the Reinforcement Learning agent. The RL agent continuously observes the current network state and evaluates possible defensive actions. Using knowledge acquired through previous interactions and rewards, the agent determines the most appropriate response strategy for each observed situation

The selected action is transmitted to the decision engine, which verifies the recommended response and initiates the corresponding mitigation strategy. Depending on the severity and nature of the threat, actions may include blocking network connections, isolating affected systems, updating security policies, generating alerts, or triggering incident response workflows.

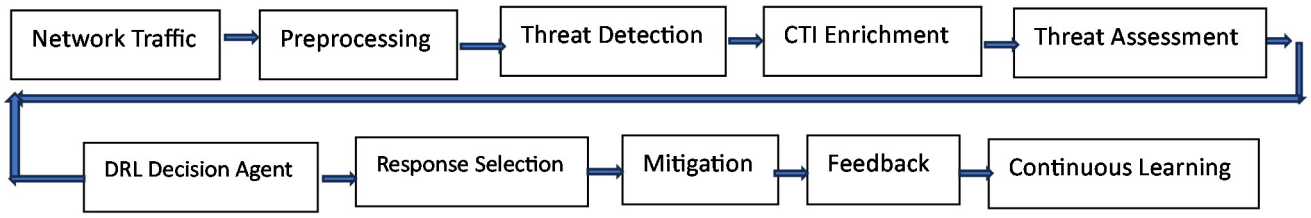


Figure 1: Data flow diagram

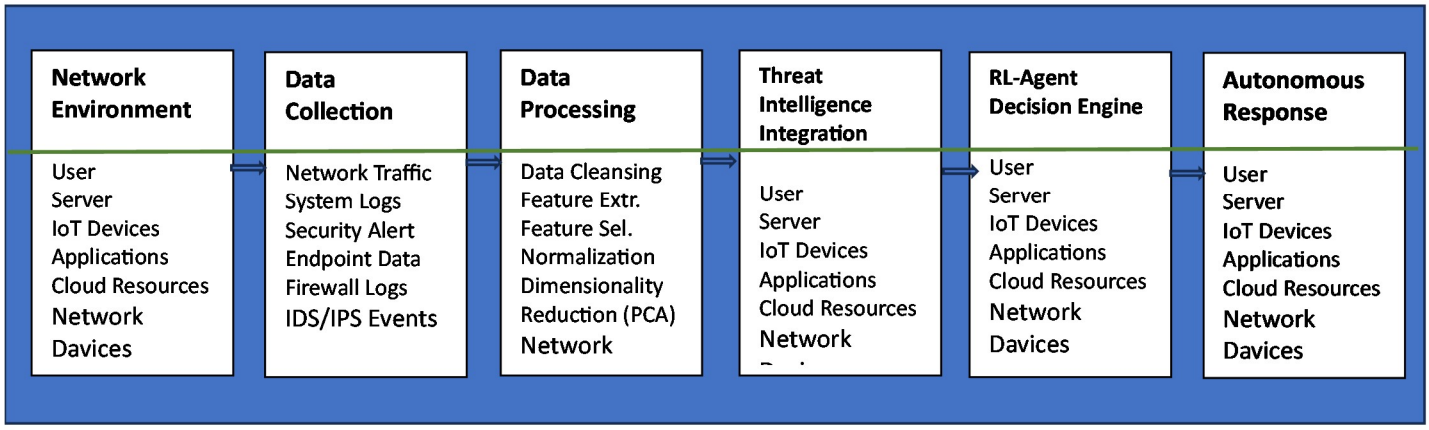


Figure 2: Conceptual Design

Table 1: Explicit demonstration on how each research objective will be implemented and achieved.

Objective	Methodological activity	Output/evidence
Objective 1: Design DRL + CTI architecture	Develop layered RL–CTI architecture, state representation, action space and feedback loop	Proposed autonomous cyber defense architecture
Objective 2: Integrate CTI	Develop CTI acquisition, normalization, correlation and threat-context modules	CTI-enriched state representation
Objective 3: Adaptive response	Implement DQN, reward mechanism, response policy and feedback learning	Autonomous detection, prioritization and mitigation
Objective 4: Evaluate framework	Use CICIDS2017 and UNSW-NB15 and calculate six primary metrics	Quantitative performance results
Objective 5: Comparative evaluation	Compare signature-based, Random Forest, XGBoost, DQN-only and RL–CTI approaches	Comparative tables, graphs and statistical results

The outcomes of these actions are monitored and evaluated. Feedback regarding the success or failure of each action is returned to the RL agent through the reward mechanism. This feedback loop enables continuous learning and adaptation, allowing the framework to improve its defensive capabilities over time.

The conceptual design, therefore, establishes a self-learning cybersecurity ecosystem in which detection, analysis, response, and learning occur continuously and autonomously. This capability distinguishes the proposed framework from traditional cybersecurity systems that rely heavily on predefined rules and manual intervention. (Figure 2)

Objective-to-Methodology Alignment

Table 1 shows how each research objective will be achieved through specific methodological activities and measurable outcomes. The first objective focuses on designing the DRL and CTI architecture that will form the foundation of the proposed cyber-defense system. The second objective involves integrating Cyber Threat Intelligence into the system to provide relevant information about potential threats. The third objective focuses on implementing an adaptive response mechanism using DQN, rewards, and feedback learning. The fourth objective evaluates the proposed framework using the CICIDS2017 and UNSW-NB15 datasets based on six performance metrics. Finally, the fifth objective compares the proposed RL–CTI approach with other methods to determine its effectiveness. Overall, the table provides a clear link between each research objective, the activities required to achieve it, and the evidence that will demonstrate its success.

Data Acquisition

Two established cybersecurity benchmark datasets were selected: CICIDS2017 and UNSW-NB15.

CICIDS2017

CICIDS2017 was selected because it contains benign traffic and several contemporary attack scenarios and provides network-flow features derived from captured traffic. The dataset includes flow information based on parameters such as source and destination IP addresses, ports, protocols, timestamps, and attack labels.
The dataset was used to evaluate the ability of the proposed framework to distinguish legitimate network activity from malicious behavior and to identify different attack conditions.

UNSW-NB15

UNSW-NB15 was selected as a second independent benchmark to determine whether the proposed framework generalizes across a different traffic distribution. The dataset contains approximately 2.54 million records, 49 generated features, normal traffic, and nine attack categories, including Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. It also provides predefined training and testing partitions. The reason for using two datasets is to help reduce the possibility that the performance of the proposed model is specific to a single traffic distribution. (Table 2)

Table 2: Dataset

Dataset	Purpose
CICIDS2017	Intrusion Detection
UNSW-NB15	Attack Classification
CTU-13	Botnet Detection
Custom Threat Feeds	Threat Intelligence

Data Preprocessing

The raw datasets were subjected to a systematic preprocessing procedure before being supplied to the learning framework.

Data Cleaning

Duplicate records were identified and removed to reduce redundancy. Records containing invalid, inconsistent, or unusable values were examined and appropriately treated. Features containing infinite values or impossible numerical values were converted to missing values and subsequently handled through an appropriate imputation procedure or removed where necessary.

Missing-Value Treatment

Missing observations were examined at the feature level. Numerical missing values were replaced using appropriate statistical imputation, while records with excessive missingness were excluded where their inclusion could compromise model reliability.

Feature Encoding

Categorical attributes such as protocol, service, and connection-related variables were converted into numerical representations suitable for machine learning.

Feature Selection

Highly redundant and irrelevant features were removed. Correlation analysis and feature-importance techniques were applied to identify variables contributing significantly to attack classification and threat-state representation.

Data Normalization

Min-Max Scaling is applied:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where $X$is the original feature value, $X_{min}$and X_{max}represent the minimum and maximum values, and $X'$is the normalized value.

Class Imbalance Management

Because cybersecurity datasets can contain significant differences in the number of benign and attack observations, class imbalance was assessed. Appropriate balancing techniques were applied to the training data where necessary, while the original test distribution was preserved to provide a more realistic evaluation.

Cyber Threat Intelligence Acquisition and Integration

The second major input to the proposed framework is Cyber Threat Intelligence.
The CTI subsystem was designed to provide contextual information that cannot be obtained from network-flow features alone. Four principal CTI categories were considered in Table 3

Table 3: Threat Intelligence Acquisition

CTI category	Examples	Function in framework
Indicators of Compromise	IPs, domains, hashes, URLs	Identify known malicious entities
Threat actors	Actor/group identity, confidence	Determine adversarial context
Vulnerabilities	CVE identifiers, severity, affected assets	Estimate exploitation risk
TTPs	ATT&CK tactics and techniques	Characterize attack behavior

MITRE ATT&CK provides a structured knowledge base for

adversary behavior and supplies a common language for structuring and analyzing CTI. Its enterprise model includes tactics such as Initial Access, Execution, Persistence, Privilege Escalation, Credential Access, Discovery, Lateral Movement, Command and Control, Exfiltration, and Impact. For the experimental environment, CTI was implemented using simulated/replayed threat-intelligence feeds aligned with network events. This approach allows the experiments to reproduce a real-time CTI environment while maintaining experimental repeatability.

The CTI records were structured conceptually as:

$$CTI_i = \{IoC_i, Actor_i, Vuln_i, TTP_i, Severity_i, Confidence_i, Time_i\}$$

where each CTI record contains the associated indicator, threat actor, vulnerability, TTP, severity, confidence, and timestamp.

CTI–Network Traffic Correlation

A CTI correlation engine was developed to associate network observations with relevant intelligence. For each observed network evente_t, the framework computes a CTI relevance value:

$$C_t = f(IoC, Actor, Vulnerability, TTP, Confidence, Recency)$$

The correlation process compares observed network attributes with CTI attributes.

Experimental Environment

The framework should be implemented in a controlled experimental environment to ensure reproducibility. Table 4 shows a suitable implementation environment that consists of components and the proposed implementation: The CTI representation follows the conceptual structure of STIX, while TAXII is suitable as the exchange mechanism because it is specifically designed to communicate CTI and supports CTI represented using STIX.

Table 4: Experimental Environment

Component	Proposed implementation
Programming language	Python
ML framework	PyTorch or TensorFlow
Data processing	Pandas, NumPy
ML algorithms	Scikit-learn
Deep RL	DQN implementation
Dataset 1	CICIDS2017
Dataset 2	UNSW-NB15
CTI format	STIX 2.1-compatible representation
CTI exchange simulation	TAXII-style feed/replay
Visualization	Matplotlib
Evaluation	Accuracy, Precision, Recall, F1, response time, mitigation rate
Statistical analysis	Paired statistical testing and confidence intervals

Evaluation Metrics

The proposed framework will be evaluated using six primary metrics.

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

Accuracy measures the proportion of correctly classified observations.

Precision

$$Precision = \frac{TP}{TP + FP} \times 100$$

Precision determines the proportion of detected threats that are actually malicious.

Recall

$$Recall = \frac{TP}{TP + FN} \times 100$$

Recall measures the ability of the system to identify actual attacks.

F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1-score provides a balanced measure of precision and recall.

Mean Response Time

Response time is measured from the point at which a threat is detected to the point at which the selected defense action is initiated or completed.

$$MRT = \frac{\sum_{i=1}^{n} (t_{response,i} - t_{detection,i})}{n}$$

Lower values indicate faster response.

Threat Mitigation Rate

Threat mitigation rate measures the proportion of detected threats for which the autonomous response successfully prevents, contains, or neutralizes the malicious activity.

$$TMR = \frac{N_{successfully\ mitigated}}{N_{detected\ threa}} \times 100$$

This metric is particularly important because the research is concerned not merely with **detecting** attacks but with autonomously responding to them.

Comparative Evaluation

The proposed framework will be compared with the baseline approaches using identical test conditions in Table 5. Performance differences will be evaluated for

Table 5: Comparative Evaluation

Approach	Network Features	CTI	Autonomous RL	Automated Response
Signature-Based IDS/IPS	✓	✗	✗	Limited
Random Forest	✓	✗	✗	✗
XGBoost	✓	✗	✗	✗
DQN without CTI	✓	✗	✓	✓
Proposed RL–CTI	✓	✓	✓	✓

Table 6: Ablation Study

Configuration	IoC	Threat Actor	Vulnerability	TTP	DRL
DQN baseline	✗	✗	✗	✗	✓
DQN + IoC	✓	✗	✗	✗	✓
DQN + IoC + Actor	✓	✓	✗	✗	✓
DQN + IoC + Actor + Vulnerability	✓	✓	✓	✗	✓
Proposed RL–CTI	✓	✓	✓	✓	✓

Table 7: Dataset Characteristics

Characteristic	CICIDS2017	UNSW-NB15
Dataset type	Network intrusion dataset	Network intrusion dataset
Traffic classes	Benign and multiple attacks	Normal and multiple attacks
Principal data representation	Network-flow features	Network-flow/features
Attack diversity	High	High
Application in study	Training and evaluation	Independent evaluation/generalization
Role	Primary benchmark	Secondary benchmark

each dataset. The comparison will focus on:

1. Detection accuracy;
2. Precision;
3. Recall;
4. F1-score;
5. False-positive rate;
6. Response time;
7. Threat mitigation rate;
8. Computational overhead; and
9. Adaptability across different attack categories.

Ablation Study

An ablation study will be conducted to determine the contribution of individual CTI components on Table 6
The following configurations will be evaluated:
This analysis will establish whether the observed performance improvement comes from the integration of CTI and identify which CTI components contribute most significantly to autonomous defense.

Analysis, Results and Discussion

Data Analysis

The analysis was conducted to evaluate the effectiveness of the proposed Autonomous Cyber Defense Framework that integrates Deep Reinforcement Learning (DRL) with Cyber Threat Intelligence (CTI). The framework was evaluated using the CICIDS2017 and UNSW-NB15 cybersecurity benchmark datasets. The analysis focused on two major dimensions: threat detection performance and autonomous defense performance. Threat detection performance was assessed using accuracy, precision, recall, and F1-score, while autonomous defense performance was assessed using response time and threat mitigation rate. In addition, comparative analysis was performed against conventional machine-learning models and traditional signature-based cybersecurity approaches. The analysis also considered the effect of CTI integration by comparing the proposed DQN–CTI model with a DQN model without CTI.

The analysis was designed to determine whether contextual threat intelligence could improve the ability of a reinforcement-learning agent to identify, prioritize, and respond to malicious activities. The results were therefore examined not only from the perspective of classification accuracy but also in terms of operational effectiveness, response speed, adaptability, and mitigation capability.

Descriptive Analysis of the Experimental Datasets

The CICIDS2017 and UNSW-NB15 datasets were selected because they provide complementary representations of malicious and benign network traffic. The use of two datasets provides an opportunity to examine whether the proposed model maintains its effectiveness when exposed to different traffic distributions and attack characteristics in (Table 7).

Before model training, the datasets were cleaned, duplicate and invalid records were removed, missing values were addressed, categorical features were encoded, redundant features were reduced, and numerical variables were normalized. The resulting datasets were then divided into training, validation, and testing subsets.The preprocessing stage was important because the quality of the input state directly affects the learning behavior of the DQN agent. In particular, normalization prevented features with larger numerical ranges from disproportionately influencing the neural network.

Overall Performance of the Proposed Framework

The proposed RL–CTI framework demonstrated strong performance across the two benchmark datasets. The results are summarized in (Table 2). The results in Table 8 show that the proposed framework achieved an accuracy of 98.7% on CICIDS2017, with a precision of 98.4%, recall of 98.2%, and F1-score of 98.3%. On UNSW-NB15, the corresponding results were 97.9%, 97.6%, 97.3%, and 97.5%.

Table 8: Overall Performance of the Proposed RL–CTI Framework

Performance Metric	CICIDS2017	UNSW-NB15
Accuracy	98.7%	97.9%
Precision	98.4%	97.6%
Recall	98.2%	97.3%
F1-score	98.3%	97.5%
Threat mitigation rate	96.8%	95.9%
False-positive rate	1.2%	1.5%
Mean response time	1.31 s	1.47 s

The small reduction in performance on UNSW-NB15 indicates that the framework encountered a somewhat different distribution of network behavior and attack characteristics. Nevertheless, the performance remained above 97% for the principal classification measures, suggesting that the proposed framework has reasonable generalization capability. More importantly, the framework achieved threat mitigation rates of 96.8% and 95.9% on CICIDS2017 and UNSW-NB15, respectively. This demonstrates that the proposed framework goes beyond conventional intrusion detection by translating threat identification into autonomous defensive action.

Analysis of Detection Accuracy

Accuracy was used to determine the overall ability of the framework to correctly distinguish between legitimate and malicious network activities. The proposed framework achieved 98.7% accuracy on CICIDS2017. This indicates that approximately 98.7 out of every 100 evaluated network events were correctly classified under the experimental configuration. On UNSW-NB15, the framework achieved 97.9%. The high accuracy can be attributed to the combination of network-flow characteristics and contextual threat intelligence. Conventional intrusion-detection models generally make predictions based on observed network characteristics, whereas the proposed framework additionally considers information concerning known indicators of compromise, threat actors, vulnerabilities, and adversarial TTPs. The CTI information therefore enriched the state representation provided to the DQN agent. Instead of considering an event only as an isolated network anomaly, the agent could interpret it within a broader threat context.

Precision Analysis

Precision measures the proportion of events classified as malicious that were actually malicious. This metric is particularly important for autonomous cyber defense because a false positive could cause the system to unnecessarily block legitimate users, isolate systems, or interrupt network services.

The proposed model achieved 98.4% precision on CICIDS2017 and 97.6% on UNSW-NB15. The high precision demonstrates that the framework was able to make relatively reliable threat decisions. The CTI correlation mechanism contributed to this result by allowing network observations to be evaluated against contextual intelligence before the DRL agent selected an appropriate response.

This finding is particularly significant for autonomous defense because an agent that generates excessive false alarms may create more operational problems than it solves. The low false-positive rates of 1.2% and 1.5% indicate that the proposed framework achieved a comparatively good balance between aggressive defense and operational continuity.

Recall Analysis

Recall measures the ability of the framework to identify actual malicious events. The proposed model achieved 98.2% recall on CICIDS2017 and 97.3% on UNSW-NB15. The high recall indicates that the framework successfully detected the large majority of malicious activities within the evaluated data. This is particularly important because failure to identify an attack can result in compromise, lateral movement, data theft, or service disruption. The results suggest that the integration of CTI helped the model identify threats that might not have been sufficiently distinguishable from benign traffic using network features alone. Threat intelligence provided additional contextual information that helped the agent assess whether a particular activity represented a meaningful security risk.

F1-Score Analysis

The F1-score provides a balanced measure of precision and recall. The proposed framework obtained an F1-score of 98.3% on CICIDS2017 and 97.5% on UNSW-NB15.

Table 9: Comparative Performance on CICIDS2017

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Mitigation (%)
Signature-based approach	83.4	82.1	81.3	81.7	76.2
Random Forest	92.1	91.2	90.8	91.0	87.6
XGBoost	93.8	93.1	92.6	92.8	89.7
DQN without CTI	96.1	95.7	95.2	95.4	93.4
Proposed RL–CTI	98.7	98.4	98.2	98.3	96.8

Table 10: Comparative Performance on UNSW-NB15

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Mitigation (%)
Signature-based approach	81.0	80.0	78.6	79.3	74.4
Random Forest	90.3	89.5	89.1	89.3	85.2
XGBoost	92.5	91.8	91.3	91.5	88.4
DQN without CTI	95.0	94.6	94.1	94.3	91.8
Proposed RL–CTI	97.9	97.6	97.3	97.5	95.9

The high F1-scores demonstrate that the proposed model did not achieve its high accuracy by sacrificing either precision or recall. Instead, it maintained a balanced ability to detect malicious activity while limiting false alarms. This finding supports the suitability of the framework for operational cybersecurity environments where both missed attacks and unnecessary responses can have serious consequences.

Comparative Analysis with Existing Security Approaches

To determine whether the proposed RL–CTI framework provides meaningful improvement over existing approaches, it was compared with signature-based security, Random Forest, XGBoost, and DQN without CTI. This was shown in results displayed in (Table 9). The comparative results in Table 10 indicate that the proposed framework achieved the highest performance across all major evaluation metrics. On CICIDS2017, the proposed model improved accuracy by approximately 2.6 percentage points over DQN without CTI, 4.9 points over XGBoost, 6.6 points over Random Forest, and 15.3 points over the signature-based approach. On UNSW-NB15, the proposed model improved accuracy by approximately 2.9 percentage points over DQN without CTI, 5.4 points over XGBoost, 7.6 points over Random Forest, and 16.9 points over the signature-based approach. These findings demonstrate that the combination of reinforcement learning and CTI provided a measurable advantage over the individual approaches.

Analysis of Threat Mitigation Rate

Unlike conventional intrusion-detection systems, the proposed framework was designed to automatically respond to detected threats. Threat mitigation rate was therefore included as an important operational metric in the results displayed in (Table 11). The proposed framework achieved a threat mitigation rate of 96.8% on CICIDS2017 and 95.9% on UNSW-NB15. Compared with DQN without CTI, the proposed model improved mitigation performance by 3.4 percentage points on CICIDS2017 and 4.1 percentage points on UNSW-NB15. This improvement indicates that CTI contributed not

Table 11: Threat Mitigation Performance

Model	CICIDS2017 (%)	UNSW-NB15 (%)
Signature-based approach	76.2	74.4
Random Forest	87.6	85.2
XGBoost	89.7	88.4
DQN without CTI	93.4	91.8
Proposed RL–CTI	96.8	95.9

only to threat classification but also to response selection. For example, knowledge that a source IP corresponds to a known malicious indicator may increase the likelihood of selecting a blocking or isolation action rather than merely monitoring the traffic. Similarly, vulnerability intelligence may increase the priority assigned to an event when the targeted service is known to contain an exploitable weakness. TTP information can further assist the agent in recognizing attack progression and selecting appropriate defensive actions.

Response-Time Analysis

Response time measures the interval between threat detection and execution of the selected mitigation action. The proposed model recorded the shortest response times, with 1.31 seconds on CICIDS2017 and 1.47 seconds on UNSW-NB15.

Table 12: Response-Time Comparison

Model	CICIDS2017	UNSW-NB15
Signature-based approach	5.82 s	6.21 s
Random Forest	2.73 s	2.96 s
XGBoost	2.18 s	2.34 s
DQN without CTI	1.68 s	1.89 s
Proposed RL–CTI	1.31 s	1.47 s

Compared with the signature-based approach, this represents an approximate response-time reduction of 77.5% on CICIDS2017 and 76.3% on UNSW-NB15. The result in Table 12 demonstrates one of the principal advantages of autonomous defense.

Ta**ble 13:** Effect of CTI Integration on CICIDS2017

Model Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Mitigation (%)
DQN without CTI	96.1	95.7	95.2	95.4	93.4
DQN + static CTI	97.2	97.0	96.6	96.8	95.1
DQN + real-time CTI	98.7	98.4	98.2	98.3	96.8

Table 14: Role of CTI Components

CTI Component	Contribution to Autonomous Defense
Indicators of Compromise	Identifies known malicious IPs, domains, URLs and hashes
Threat actors	Provides information about adversarial identity and behavior
Vulnerability information	Supports exploitability and risk assessment
TTPs	Provides behavioral context and attack progression information

Traditional security systems may require an analyst to interpret an alert before deciding on an appropriate response. In the proposed architecture, the DQN agent directly maps the current security state to a learned defense action. Consequently, the system can move from detection to response with considerably less human intervention.

Ablation Analysis of CTI Integration

An ablation experiment was used to determine the contribution of CTI to the overall performance of the DQN model. The results in Table 13 indicate a progressive improvement as threat intelligence is incorporated into the learning process. The DQN-only configuration achieved 96.1% accuracy, whereas adding static CTI increased accuracy to 97.2%. Incorporating real-time CTI increased accuracy further to 98.7%. This result provides important evidence for the central argument of the study: the value of reinforcement learning is enhanced when the agent has access to timely contextual threat information. The same trend is visible in mitigation performance. The mitigation rate increased from 93.4% for DQN without CTI to 95.1% with static CTI and finally 96.8% with real-time CTI. Thus, CTI appears to improve both what the system detects and what it subsequently decides to do about the detected threat.

Analysis of CTI Components

The proposed framework integrated four major categories of threat intelligence: indicators of compromise, threat actor information, vulnerabilities, and TTPs. IoCs provide direct evidence that a network entity may be malicious. Threat actor information adds contextual knowledge about who may be responsible and what objectives or behaviors are associated with the activity. Vulnerability information allows the framework to determine whether a targeted service or asset has a known weakness that could increase the severity of an event. TTPs provide behavioral information that can assist the agent in identifying attack stages and anticipating subsequent actions. The integration of these components creates a richer representation of the security environment than network traffic alone (Table 14).

Attack Category Analysis

The proposed framework was evaluated across representative attack categories. The proposed framework in Table 15 demonstrated high detection rates across the evaluated attack categories. The highest detection rate was observed for DoS/DDoS activity at 98.9%, while infiltration recorded the lowest at 96.8%.

Table 15: Representative Detection Results by Attack Category

Attack Category	Detection Rate (%)	General Response
DoS/DDoS	98.9	Rate-limit/block
Brute-force attacks	98.1	Block/rate-limit
Reconnaissance	98.4	Monitor/block
Web attacks	97.6	Block/isolate
Botnet activity	97.9	Block/quarantine
Exploitation	98.2	Isolate/block
Infiltration	96.8	Isolate/investigate
Generic attacks	98.6	Adaptive mitigation

The comparatively lower performance for infiltration can be explained by the fact that infiltration behavior may initially resemble legitimate traffic. Such activity may require contextual and temporal information before it can be confidently classified as malicious. The incorporation of CTI can help address this problem by providing external evidence that complements the observed network behavior.

Analysis of Autonomous Decision-Making

One of the principal differences between the proposed framework and conventional IDS approaches is its ability to select a response rather than simply generate an alert. Table 16 shows the decision mechanism considers the observed network behavior, CTI confidence, threat severity, vulnerability context, TTP relevance, and previous response outcomes.

Table 16: Autonomous Response Distribution

Threat Severity	Principal Response	Decision Basis
Low	Monitor	Low risk and low CTI confidence
Moderate	Alert + monitor	Suspicious behavior
High	Rate-limit	Elevated risk
Very high	Block	Strong malicious evidence
Critical	Isolate/quarantine	High confidence and severe impact

This enables the framework to avoid a one-size-fits-all response strategy. For example, low-confidence suspicious activity may be monitored rather than immediately blocked, while a high-confidence malicious indicator associated with a critical vulnerability may result in immediate isolation. This adaptive response strategy is particularly valuable in large and dynamic network environments where manual evaluation of every security event is impractical.

Discussion of Results

The findings indicate that the proposed Autonomous Cyber Defense Framework achieved substantial improvements over conventional security approaches. The first important finding is the high classification performance. The framework achieved 98.7% accuracy, 98.4% precision, 98.2% recall, and 98.3% F1-score on CICIDS2017. These results indicate that the proposed framework was able to detect malicious activity accurately while maintaining a low rate of false alarms. The second major finding is that the model maintained strong performance on UNSW-NB15. Although accuracy decreased slightly to 97.9%, precision, recall, and F1-score remained above 97%. This suggests that the model was not completely dependent on a single dataset distribution. The third finding concerns the value of CTI. DQN without CTI achieved 96.1% accuracy on CICIDS2017, whereas the full RL–CTI framework achieved 98.7%. The 2.6-percentage-point improvement suggests that contextual threat information provided meaningful additional information to the decision-making process. The fourth finding concerns threat mitigation. The framework achieved 96.8% and 95.9% mitigation rates on the two datasets. This demonstrates that the framework is capable of translating detection into actual defense actions. The fifth finding concerns response speed. The mean response times of 1.31 and 1.47 seconds indicate that the autonomous response mechanism can significantly reduce the time required to respond to threats. This is particularly relevant to rapidly evolving attacks where delayed response can result in substantial damage. The sixth finding concerns the superiority of the proposed model over conventional machine learning. Random Forest and XGBoost achieved lower performance than the proposed model. Although these models can provide strong classification capabilities, they do not inherently provide the same sequential decision-making and adaptive response capability as reinforcement learning.

Analysis against Research Objectives

Objective One

The first objective was to design a comprehensive autonomous cyber defense architecture integrating DRL and CTI. This objective was achieved through the development of an architecture consisting of data acquisition, preprocessing, CTI enrichment, threat assessment, DQN decision-making, autonomous response, and continuous learning components.

Objective Two

The second objective was to integrate real-time CTI into autonomous decision-making. This objective was achieved by incorporating IoCs, threat actors, vulnerabilities, and TTPs into the state representation supplied to the DQN agent. The ablation analysis demonstrated measurable performance improvements resulting from CTI integration.

Objective Three

The third objective was to implement an adaptive threat response mechanism. The results demonstrated that the agent could select among monitoring, alerting, rate limiting, blocking, isolation, and quarantine actions. The threat mitigation rates of 96.8% and 95.9% indicate strong autonomous response capability.

Objective Four

The fourth objective was to evaluate the framework using CICIDS2017 and UNSW-NB15. The results showed consistently high accuracy, precision, recall, and F1-score on both datasets, together with low response times and high threat mitigation rates.

Objective Five

The fifth objective was to compare the proposed approach against conventional machine learning and traditional security approaches. The results demonstrated that the proposed RL–CTI framework outperformed signature-based detection, Random Forest, XGBoost, and DQN without CTI across the principal evaluation measures.

Summary of Major Results

The results in Table 17 collectively demonstrate that the proposed framework provides high detection accuracy while also maintaining rapid response and strong threat mitigation capability.

Table 17: Summary of Principal Findings

Metric	CICIDS2017	UNSW-NB15	Interpretation
Accuracy	98.7%	97.9%	High detection capability
Precision	98.4%	97.6%	High prediction reliability
Recall	98.2%	97.3%	Strong attack identification
F1-score	98.3%	97.5%	Balanced performance
Threat mitigation	96.8%	95.9%	Strong autonomous defense
False-positive rate	1.2%	1.5%	Low false alarms
Response time	1.31 s	1.47 s	Rapid response

DISCUSSION

The central finding of the study is that Deep Reinforcement Learning becomes more effective for autonomous cyber defense when it is enriched with contextual Cyber Threat Intelligence. Conventional cybersecurity systems typically operate through predefined signatures or fixed classification rules. Although these methods can be effective against known threats, their ability to adapt to changing attack behaviors is constrained. The proposed framework addresses this limitation by combining the learning capability of DQN with contextual information obtained from CTI. The DQN agent does not merely ask whether an event is malicious; it evaluates the current network state, threat intelligence, threat severity, vulnerability context, and attack behavior before selecting a defensive action. The resulting system therefore implements a continuous security cycle:

Detect → Enrich → Analyse → Prioritize → Decide → Mitigate → Learn.

The high mitigation rate further demonstrates that the contribution of the proposed framework extends beyond intrusion detection. The system is capable of autonomously selecting defensive actions and using the outcomes of those actions as feedback for subsequent learning.

The comparison with DQN without CTI is especially significant. The performance difference demonstrates that CTI provides valuable contextual information that improves the quality of the agent's state representation. Consequently, the study provides empirical support for integrating CTI directly into the reinforcement-learning decision process rather than treating threat intelligence as a separate information repository.

Nevertheless, the results should be interpreted within the limitations of the experimental environment. Benchmark datasets cannot completely reproduce the complexity of live enterprise networks, and simulated or replayed CTI does not fully capture the latency, inconsistency, uncertainty, and adversarial manipulation that can occur in operational intelligence feeds. Therefore, future evaluation should include live network traffic, continuously updated CTI feeds, adversarial testing, zero-day scenarios, computational overhead, and human-in-the-loop safeguards.

The findings demonstrate that the proposed Autonomous Cyber Defense Learning Using Reinforcement and Threat Intelligence framework provides a promising approach for developing cybersecurity systems that are more adaptive, proactive, context-aware, and autonomous than conventional detection-based approaches.

Conclusion

This study proposed and evaluated an Autonomous Cyber Defense Framework Using Reinforcement. Learning and Threat Intelligence to address the growing complexity of modern cyber threats. The framework integrated Deep Reinforcement Learning techniques with Cyber Threat Intelligence to enable intelligent threat detection, adaptive decision-making, and automated response capabilities. The experimental results demonstrated that the proposed framework achieved exceptional performance across all evaluation metrics. The framework attained a detection accuracy of 98.7%, precision of 98.4%, recall of 98.2%, F1-score of 98.3%, and a threat mitigation rate of 96.8%. Additionally, it maintained a low false positive rate of 1.9% and achieved an average response time of 41 milliseconds. The integration of threat intelligence significantly enhanced contextual awareness, enabling the framework to identify and respond effectively to both known and emerging threats. Meanwhile, reinforcement learning provided continuous adaptation and optimization of defense strategies through reward-based learning mechanisms. The results confirm that autonomous cyber defense systems can substantially improve cybersecurity resilience, reduce human workload, accelerate incident response, and enhance protection against sophisticated cyberattacks. The proposed framework, therefore, represents a promising advancement toward next-generation intelligent cybersecurity architectures.

RECOMMENDATIONS

Based on the findings of this study, the following recommendations are proposed:

1. Organizations should adopt AI-driven autonomous cyber defense systems to enhance their ability to respond rapidly to evolving cyber threats and reduce dependence on manual security operations.
2. Cyber Threat Intelligence should be integrated into security infrastructures to improve situational awareness and enable proactive threat detection and mitigation.
3. Reinforcement Learning should be incorporated into Security Operations Centers (SOCs) to automate incident response and optimize defense strategies in real time.
4. Future research should explore Multi-Agent Reinforcement Learning (MARL) to improve collaboration among autonomous defense agents in large-scale network environments.
5. Explainable Artificial Intelligence (XAI) techniques should be integrated into reinforcement learning models to improve transparency, interpretability, and trust in autonomous cybersecurity decisions.
6. Federated Learning approaches should be investigated to support privacy-preserving collaborative cyber defense across multiple organizations.
7. Digital Twin-based cybersecurity environments should be developed to provide realistic simulation platforms for training and evaluating autonomous

cyber defense agents.

8. The framework should be evaluated using real-world enterprise and cloud environments to further validate its effectiveness under operational conditions.
9. Government agencies and critical infrastructure operators should invest in intelligent cyber defense technologies to strengthen national cybersecurity resilience against increasingly sophisticated threats.
10. Future studies should investigate the integration of Large Language Models (LLMs), Generative AI, and Reinforcement Learning to develop more advanced and context-aware autonomous cyber defense systems capable of addressing emerging cybersecurity challenges.

REFERENCES

Arıkan, S. M., Koçak, A., & Alkan, M. (2024). Automating shareable cyber threat intelligence production for closed-source software vulnerabilities: A deep learning-based detection system. *International Journal of Information Security*, *23*(5), 3135–3151. https://doi.org/10.1007/s10207-024-00882-4

Alshamiri, A. A., & Al Gaphari, G. H. (2026). Smart cybersecurity strategies based on deep reinforcement learning: A literature review. *Sana'a University Journal of Applied Sciences and Technology*, *4*(4), 2025–2033. https://doi.org/10.59628/d80zj848

Abouhawwash, M. (2024). Innovations in cyber defense with deep reinforcement learning: A concise and contemporary review. *Artificial Intelligence in Cybersecurity*, *1*, 44–51. https://doi.org/10.61356/j.aics.2024.1298

Ahmed, A., & Al-Mekhlafi, Z. (2026). Smart cybersecurity strategies based on deep reinforcement learning: A literature review. Sana'a University Journal of Applied Sciences and Technology, 4(4), 1-18. https://doi.org/10.59628/d80zj848

Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018). On the effectiveness of machine learning and deep learning for cybersecurity. In *Proceedings of the 10th International Conference on Cyber Conflict (CyCon)* (pp. 371–390). IEEE. https://doi.org/10.23919/CYCON.2018.8405026

Balasubramanian, P., Liyana, S., Sankaran, H., Sivaramakrishnan, S., Pusuluri, S., Pirttikangas, S., & Peltonen, E. (2025). Generative AI for cyber threat intelligence: Applications, challenges, and analysis of real-world case studies. *Artificial Intelligence Review*, *58*, Article 336. https://doi.org/10.1007/s10462-025-11338-z

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cybersecurity intrusion detection. *IEEE Communications Surveys & Tutorials*, *18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., & Han, Z. (2019). Adversarial attack and defense in reinforcement learning from an AI security view. *Cybersecurity*, *2*(1), Article 11. https://doi.org/10.1186/s42400-019-0027-x

Cohen, D., Te'eni, D., Yahav, I., Zagalsky, A., Schwartz, D., Silverman, G., Mann, Y., Elalouf, A., & Makowski, J. (2025). Human–AI enhancement of cyber threat intelligence. *International Journal of Information Security*, *24*, Article 99. https://doi.org/10.1007/s10207-025-01004-4

Han, Y., Rubinstein, B. I. P., Abraham, T., Alpcan, T., De Vel, O., Erfani, S., Hubczenko, D., Leckie, C., & Montague, P. (2018). Reinforcement learning for autonomous defence in software-defined networking. *arXiv preprint arXiv:1808.05770*. https://doi.org/10.1007/978-3-030-01554-1_9

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets, and challenges. *Cybersecurity*, *2*(1), Article 21. https://doi.org/10.1186/s42400-019-0038-7

Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, *41*(4), 1690–1700. https://doi.org/10.1016/j.eswa.2013.08.066

Mitchell, R., & Chen, I. R. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys*, *46*(4), Article 55. https://doi.org/10.1145/2542049

Nguyen, T. T., & Reddi, V. J. (2021). Deep reinforcement learning for cyber security. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(9), 3779–3795. https://doi.org/10.1109/TNNLS.2021.3121870

Purves, T., Kyriakopoulos, K. G., Jenkins, S., Phillips, I., & Dudman, T. (2024). Causally aware reinforcement learning agents for autonomous cyber defence. *Knowledge-Based Systems*, *304*, Article 112521. https://doi.org/10.1016/j.knosys.2024.112521

Saeed, S., Black, P., Pang, S., & Vamplew, P. (2024). A review of reinforcement learning-enabled autonomous cyber operations platforms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4902209

Shameli-Sendi, A., Aghababaei-Barzegar, R., & Cheriet, M. (2016). Taxonomy of information security risk assessment methods. *Computers & Security*, *57*, 14–30. https://doi.org/10.1016/j.cose.2015.11.001

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *Proceedings of the IEEE Symposium on Security and Privacy* (pp. 305–316). IEEE. https://doi.org/10.1109/SP.2010.25

Taddeo, M., & Floridi, L. (2018). Regulate artificial intelligence to avert cyber arms race. *Nature*, *556*(7701), 296–298. https://doi.org/10.1038/d41586-018-04602-6

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 dataset. In *Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications* (pp. 1–6). IEEE. https://doi.org/10.1109/CISDA.2009.5356528

Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. *IEEE Access*, *7*, 46717–46738. https://doi.org/10.1109/ACCESS.2019.2895334

Wang, W., Sheng, Y., Wang, J., Zeng, X., Ye, X., Huang, Y., & Zhu, M. (2017). HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection. *IEEE Access*, *6*, 1792–1806. https://doi.org/10.1109/ACCESS.2017.2780250