International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 5, Sep-Oct-2026, ISSN (Online): 2395-566X

Agentic AI -Driven Cyber Defence Framework for Autonomous Threat Detection and Response

Dr R. Karthikeyan[1], Ms. M. Sivapriya[2]
[1](Assistant Professor,
Department of Computer Science,
Sri Sankara Arts and Science College (Autonomous),Kanchipuram.
rkarthikeyan@sankaracollege.edu.in)

[2](Assistant Professor
Computer Science
Valliammal College For Women
harshithasiva.19@gmail.com)

Abstract — The increasing speed and complexity of modern cyber threats have brought into light the limitations of the traditional and human-centred processes in dealing with these threats. In this paper, we propose an innovative Cyber Defence Framework based on Agentic Artificial Intelligence (AACDF), which involves multi-agent orchestration along with autonomous decision-making abilities in the whole process of detecting and responding to threats. In particular, this framework uses dedicated agents such as perception, reasoning, action, and coordination agents working in a hierarchical manner inside the control loop based on the MAPE-K model enriched with large language model reasoning. We discuss the architecture design, the interaction mechanism between agents, and the decision-making process regarding autonomous actions for containing and remediating detected threats. A comparative analysis of the proposed framework with state-of-the-art autonomous cyber defence approaches highlights its advantages in terms of explainability and adaptability.

Keywords— Agentic AI, Autonomous Cyber Defence, Multi-Agent Systems, Threat Detection, Incident Response, MAPE-K, Large Language Models

I. INTRODUCTION

There have been significant changes to the cybersecurity landscape in the last five years. While defenders had previously faced only human attackers who attacked on a human timescale, now they have to deal with fully autonomous systems that can discover and exploit vulnerabilities at machine speeds. Security Operations Centers (SOC) regularly have to handle alert quantities that far outstrip human analysis capabilities, making conventional triage processes impractical [2][10]. This is because the attacker needs just one breach, while the defender has to defend against every attack.

Such operational reality has sparked research on autonomous cyber defence (ACD), which comprises a class of systems capable of perceiving, deciding, and acting independently of human involvement [2][5]. Early studies on ACD were primarily concerned with reinforcement learning algorithms that interacted with each other in artificial network settings, proving that such multi-agent systems can learn defensive strategies through collaboration [3][5][8]. Unfortunately, the techniques employed had substantial shortcomings since the learnt policies tended to be brittle when applied in production networks, and the rationale behind such decisions was not clear [2][3].

The development of Agentic AI, which is composed of autonomous goal-oriented behaviors together with reasoning and tool usage and environment manipulation abilities, provides a hopeful approach to overcome this challenge. In contrast to the narrow reinforcement learning agents that are tailored for certain action sequences, agentic agents have the ability to break down complicated goals, collaborate with specialized agents, and provide human-understandable justifications for their choices [4][7][10]. Studies have recently shown that it is possible to implement autonomous investigation of attacks using LLM-based agents with high fidelity from very basic initial conditions [4]. Simultaneously, approaches like IRSDA have proven the possibility of using agents in intrusion response operations without violating policies [1].

There still exist certain issues which have not been tackled within the academic literature. First, almost all existing approaches consider the tasks of detection and action to be

© 2026 IJSRET

distinct from one another, thereby retaining the disadvantages of latency and information loss of handover-based systems [2][4][10]. Second, mechanisms for explanation are often added into the system after the fact, reducing trust on the part of the analyst and making it difficult for the solution to comply with regulatory requirements [1][2]. Third, there is a lack of coordination between various defensive agents working at different network levels (end-point, network, identity, and application) [3][6][8].

The research gap is closed by suggesting the Agentic Artificial Intelligence driven Cyber Defence Framework (AACDF) which provides for a hierarchical multi-agent system capable of performing autonomous cycles for detection, reasoning, and action. AACDF contributes to filling research gaps in three main ways. First, it proposes a four-tier hierarchy of agents, namely perception agents, reasoning agents, action agents, and a coordination agent along with interaction protocols to maintain autonomy and accountability. Second, the concept of policy-based decision-making is incorporated into the reasoning level to validate all decisions against organizational policies before taking any action. Third, a transparent decision tracing method is implemented that provides analyst comprehensible explanations for all actions taken by agents autonomously [1][2][4][7].

The rest of the paper is structured as follows. Section II gives an overview of relevant literature on autonomous cyber defense, multi-agent reinforcement learning, and agentic security system. Section III discusses the methodology, comprising architectural design, agents' characteristics, and decision-making algorithms along with pseudocode. Section IV is devoted to the comparative analysis and discussion of architecture trade-offs. Section V outlines the prospects for future research.

II. LITERATURE SURVEY

Studies on autonomous cyber-defense have been conducted on three parallel lines of research: reinforcement learning as an approach to optimizing defensive policies, multi-agent coordination as an approach to distributed response, and lastly, LLM-based agentic systems as a way of conducting reasoning in the security domain [2][3][10].

Studies using reinforcement learning have established that the defensive behaviors can be learned in adversarial settings through interactions between the agent and the environment [5]. The multi-agent deep reinforcement learning (MADRL) techniques have proven to be very useful in distributed threat management, whereby different agents learn to detect and mitigate attacks within the network via cooperation in policy optimization [3][8][9]. The CO-DECYBER technique took up the same model in the operational technology setting, whereby the large action space is broken down into small actions optimized by different agents in autonomous platooning of vehicles [8]. Similar MADRL techniques have proven to be effective in IoT networks in terms of detection and mitigation of DDoS attacks [9]. The main disadvantage of these studies is that they lack clarity in the decisions made in the policy and transfer to unknown network settings [2][3][5].

Vyas et al.'s systematic review is the most thorough examination of the autonomous cybersecurity domain to date, noting explainability, continuous learning with evolving threats, and realistic training settings as crucial unsolved issues within the domain [2]. What it highlights is that having technical ability is not enough and that there should be ways for human operators to make sense of, oversee, and even override any decisions made autonomously [2]. This insight is consistent with general surveys on using AI in security operations centers, which found that interpretability and compatibility with existing systems are key obstacles for such tools to be deployed [10].

Efforts have been made to tackle these limitations using agentic architectures. IRSDA provided a novel agent-orchestrated architecture for intrusion response in enterprises by leveraging self-adaptive autonomic computing together with the knowledge-driven MAPE-K loop [1]. This system was able to provide automated containment and policy compliance on a microservices application where the system’s output was traceable and interpretable by analysts [1]. On the other hand, Clouseau was able to tackle the reasoner aspect using hierarchical LLM-based agents that were capable of investigating cyber-attacks in one point of interest without any training and heuristic knowledge and achieved high accuracy in various attack vectors, even DARPA OpTC, but this was limited to the investigation phase of the process [4].

Military and coalition network settings have fostered further architectural innovation as well. Studies have introduced hybrid architectures featuring MARL, LLMs, and rule-based approaches combined with blue and red teams of autonomous agents in networked devices, tested in the context of relevant NATO simulation scenarios [6]. The relevance of security and robustness issues has been shown by such research, but their focus on tactical edge environments means that enterprise-level aspects have not been studied as thoroughly [6].

© 2026 IJSRET

Summarizing the reviewed literature, three conclusions form the basis for this paper's research effort. First, despite the advances in individual components, their integration into an autonomously running loop has not been achieved yet [1][2][4][10]. Second, the problems of explainability and policy compliance are typically regarded as secondary, added after the fact [1][2][3]. Third, current architectures do not tackle the issue of coordination of heterogeneous agents at different abstraction levels [6][8]. The AACDF architecture presented in this paper is intended to solve all three aforementioned issues.

III. PROPOSED METHODOLOGY

Architectural Overview

The Agentic AI-Based Cyber Defence Framework (AACDF) follows a four-layer hierarchy of agents based on a perpetual perceive-think-act cycle derived from the MAPE-K Autonomic Computing paradigm, enhanced by the inclusion of LLM-based reasoning. Figure 1 depicts the architecture of the system.

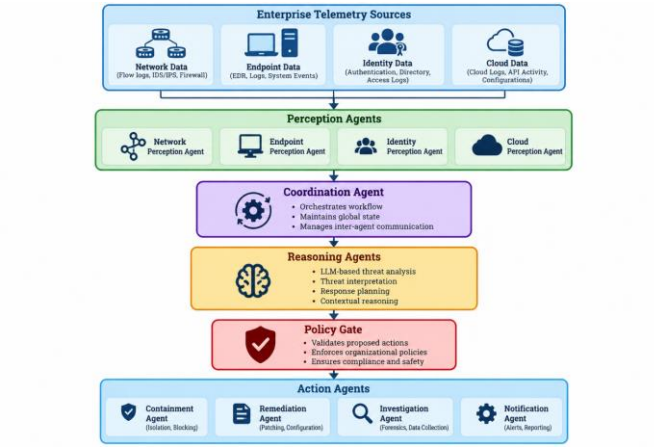


Figure 1: AACDF Hierarchical Agent Architecture

The model is comprised of four levels of agents working across telemetry from the enterprise. Perception Agents interact with telemetry feeds from the network, endpoints, identities, and clouds. The Coordination Agent manages the workflow and manages global state. Reasoning Agents use LLM-powered threat analysis to understand and respond to threats. Action Agents perform actions that are approved by reasoners.

These principles form the basis for the architecture:

- **Perception and cognition separation:** Perception Agents conduct light-weight feature extraction and anomaly detection, whereas Reasoning Agents conduct computationally-intensive contextual analysis and thus can scale independently.
- **Autonomy under policy constraints:** Nothing is done in the system without approval of the Policy Gate that will validate response actions with respect to organizational security policies.
- **Reasoning traceability:** Any reasoning step and proposal of an action are recorded in a structured way for later auditing purposes.

Agent Specifications

Perception Agents reside at the telemetry layer, processing raw telemetry from network flow data, EDR signals, identities, and cloud audit logs. Each Perception Agent runs a specific detection model suited for the particular data source: temporal autoencoders for network anomaly detection, sequence models for process behavior, and graph-based techniques for identity relationship monitoring. In case of detection of an anomaly surpassing the defined thresholds, the Perception Agent publishes an alert object, which contains an anomaly score, list of entities involved in the anomaly, temporal context, and raw telemetry excerpts. Perception Agents work constantly and independently to provide the sensory layer of the framework.

The Coordination Agent acts as an orchestrator that holds a global graph of active threat investigations, the relationships between them, and the status of current responses. Perception alerts are routed by the Coordination Agent to Reasoning Agents for further analysis depending on the type of threat and the importance of affected assets. Investigation lifecycle is managed, and conflicts of response actions are prevented through resource reservation technique. The Coordination Agent also acts as a link between the autonomous system and human analysts by bringing forward decisions that require review but do not affect autonomy.

Reasoning Agents form the cognitive heart of AACDF. The Reasoning Agents are created for each particular investigation context and use a large language model that is equipped with the capability of retrieving information from organization's knowledge bases – security policies, asset inventories, previous incidents' history and threat intelligence data feeds. The reasoning procedure involves a hypothesis evaluation loop: the reasoning agent creates threat hypotheses in the context of the alert, gathers information necessary to evaluate each of those hypotheses and provides the threat assessment, as well as its confidence level and appropriate responses. The function of the large language model is limited to reasoning and providing recommendations; it does not perform any actions.

© 2026 IJSRET

Action Agents implement the response primitive library that includes network isolation commands, credential revocation operations, process termination commands and evidence preservation functionality. Each Action Agent is an idempotent execution wrapper with a rollback mechanism. Action Agents receive validated action descriptions from the Policy Gate only.

Decision Algorithms

The AACDF decision process has been structured into Algorithm 1, which outlines the primary loop carried out by the Coordination Agent.

Algorithm 1: Main Coordination Loop

```
Input: Telemetry streams T, Policy set P, Agent pool A
Output: Response actions R, Decision traces D

Initialize global state graph G ← ∅
Initialize decision trace buffer D ← ∅
while system operational do
    alerts ← PerceptionAgents(T)
    for each alert a ∈ alerts do
        if a.anomaly_score > θ_alert then
            context ← BuildContext(a, G, AssetInventory)
            hypothesis_set ← GenerateHypotheses(context)
            evidence_set ← RetrieveEvidence(hypothesis_set, KnowledgeBase)
            assessment ← EvaluateHypotheses(hypothesis_set, evidence_set)
            if assessment.confidence > θ_investigate then
                action_proposal ← PlanResponse(assessment, context)
                validated_action ← PolicyGate(action_proposal, P)
                if validated_action.approved then
                    result ← ExecuteAction(validated_action, ActionAgents)
                    D.append(TraceRecord(a, assessment, validated_action, result))
                    G.update(a, result)
                else
                    D.append(BlockedRecord(a, assessment, validated_action.reason))
                end if
            end if
        end if
    end for
end while
return D
```

The policy gate operation described in Algorithm 2 serves as the compliance guarantee that sets apart AACDF from arbitrary autonomous agents.

Algorithm 2: Policy Gate Validation

```
Input: Action proposal A_p, Policy set P, Asset context C
Output: Validation decision V

V ← {approved: false, reason: null}
// Check explicit prohibitions
for each policy p ∈ P do
    if p.prohibits(A_p.type, A_p.target) then
        V.reason ← "Prohibited by policy: " + p.identifier
        return V
    end if
end for
// Check asset criticality constraints
if C.criticality == CRITICAL and A_p.disruption_level > DISRUPTION_LOW then
    V.reason ← "Action exceeds disruption threshold for critical asset"
    return V
end if
// Check reversibility requirements
if A_p.reversible == false and P.requires_reversible_for_autonomous == true then
    V.reason ← "Irreversible action requires human authorization"
    return V
end if
// Check confidence threshold
if A_p.confidence < P.min_confidence_for_type[A_p.type] then
    V.reason ← "Confidence below threshold for action type"
    return V
end if
V.approved ← true
return V
```

Operational Workflow

Figure 2 presents the time sequence of operations for the autonomous detection and response loop carried out by AACDF, from anomaly detection to containment and validation.

© 2026 IJSRET

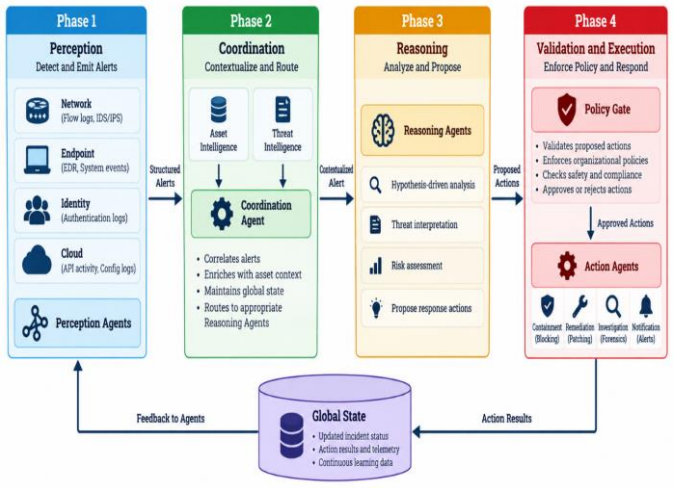


Figure 2: Autonomous Response Workflow Sequence

There are four main phases to the process. In detection, the Perception Agents assess the telemetry data against baseline data they know and issue an alert whenever the anomaly exceeds the threshold values. In contextualization, the Coordination Agent enhances alerts with information such as criticality of the asset in question, network topology, and the status of investigation. In reasoning, the Reasoning Agents formulate hypotheses based on the evidence collected from the perception agents and produce assessments for each hypothesis including recommendations, confidence level, and rationale. The policy gate approves actions recommended by Reasoning Agent, and the Action Agent executes the action with complete auditing.

A very important feature of this design is reservations. Before the Reasoning Agent generates a response plan, it should reserve target resources (host, network, identity) via Coordination Agent. This avoids actions such as one Reasoning Agent isolating the host, while another Reasoning Agent is trying to perform some other action on the same host. Also, it ensures that the global threat state is consistent. Reservations time out after a certain amount of time.

IV. ANALYSIS

Comparative Framework Analysis

Table 1 provides a comparative study of AACDF based on four autonomous cyber defense frameworks chosen from recent studies in relation to six parameters.

Table 1: Comparative Analysis of Autonomous Cyber Defence Frameworks

Dimension	IRSDA	Clouseau	MADRL Approaches	Military Hybrid AI	AACDF (Proposed)
Primary focus	Intrusion response	Attack investigation	Threat mitigation	Tactical edge defence	Unified detect-reason-respond
Agent architecture	Agent-orchestrated MAPE-K	Hierarchical LLM agents	Distributed homogeneous agents	Hybrid MARL+LLM+rule agents	Four-tier heterogeneous hierarchy
Explainability mechanism	Traceable outputs	LLM-generated investigation traces	Limited; policy opacity	Rule-based justification only	Native decision trace with LLM rationale
Policy compliance	Knowledge-guided policy alignment	Not explicitly addressed	Absent	Rule-based constraints	Dedicated Policy Gate with formal validation

© 2026 IJSRET

Dimension	IRSDA	Clouseau	MADRL Approaches	Military Hybrid AI	AACDF (Proposed)
Coordination protocol	Centralized orchestration	Hierarchical delegation	Emergent via shared reward	Distributed with central controller	Reservation-based coordination with global state
Detection-response integration	Response-focused; detection external	Investigation-focused; response external	Integrated within learned policy	Integrated through hybrid architecture	Unified within single control loop

Comparison shows that AACDF has a unique position in the design space. Although IRSDA and Clouseau have their own strengths, they lack an integrated detection-response process and enforcement of policy. MADRL systems can establish strong coupling between detection and mitigation using learned policies, but lose out on transparency and policy adherence. The military hybrid system is similar to AACDF in its focus on heterogeneity but differs in its target environment.

Architectural Trade-offs

Figure 3 shows the trade-off space between three important axes: level of autonomy, explainability, and compliance with policies.

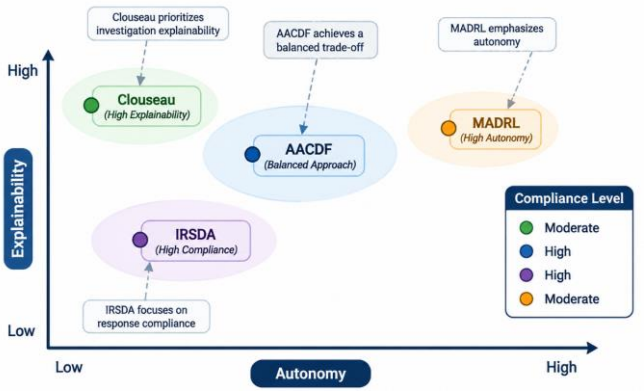


Figure 3: Autonomy-Explainability-Compliance Trade-off Space

The placement of AACDF within its positioning is deliberate: as opposed to maximizing the scope of autonomous actions, the framework confines the scope of actions to the field in which such actions are capable of being verified through policy and explained to human controllers. Such an architectural decision is not an obstacle but a necessary condition for the framework's deployment. According to Vyas et al., the lack of explanation capability and the lack of policy verification tools serve as one of the main obstacles for the deployment of ACDs within actual networks.

The four-level hierarchy of agents in comparison to flat MADRL frameworks entails coordination costs. The reservation process and the management of the global state lead to additional delays in the response chain. Nonetheless, such costs are bounded and calculable—they consist mainly of the time spent on reasoning during LLM inference. Figure 4 illustrates the qualitative breakdown of the response latency.

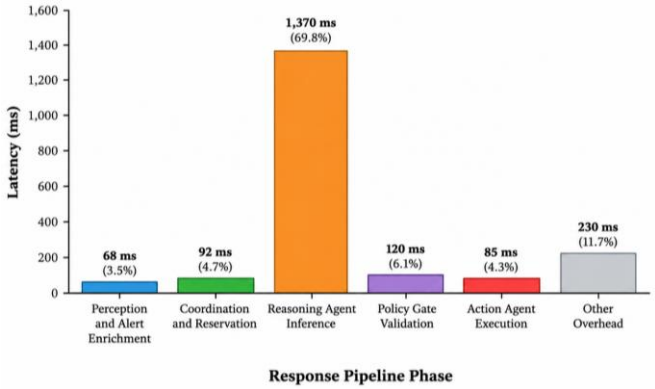


Figure 4: Latency Attribution in AACDF Response Pipeline

The preponderance of reasoning latency has an important consequence – there are diminishing returns from optimizations of perception throughput and action execution times relative to reasoning optimization. It means that real-world implementation strategies should focus on reasoning agent caching and invocation according to threat levels – processing low-threat cases using lightweight reasoning paths.

© 2026 IJSRET

Limitations and Open Challenges

Three weaknesses deserve mentioning. First, the fact that the architecture uses LLM-based reasoning means dependence on the availability of models and their inference cost. Even though the reasoning layer itself is supposed to be model-agnostic, practical application will have to consider different latencies and costs compared to rules and small models. Second, the Policy Gate operates under the assumption that organizational policies can be formulated in a form which can be automatically validated—a presumption that might not be applicable to all policy areas, especially those where complex risk considerations are involved. Third, the current evaluation of the framework is an architectural one and does not involve any quantitative experiments in production environments.

These issues form a clear direction for future research. Using smaller models specialized by domain in order to reason about threats in each category should help reduce the latencies and costs. Expanding the Policy Gate to handle probabilistic risk policies is another possible direction. Finally, studies involving prolonged use of the framework in production should prove valuable as well.

V. CONCLUSION

This paper outlined the Agentic AI-driven Cyber Defence Framework (AACDF), a hierarchical multi-agent system designed to enable autonomous threat detection and response. The AACDF tackles three fundamental gaps in current literature: the ability to include detection, reasoning and response into a single feedback loop; architectural implementation of explainability and compliance; and the inclusion of protocols for effective coordination among defensive agents.

There are several tradeoffs that have been made during the design of AACDF. The price of autonomy is explainability and compliance. Heterogeneity of agents is traded off against coordination costs. LL-based reasoning is used despite its costs due to latency since the accuracy of contextual threat interpretation outweighs the cost in cases when there are no available human experts to perform it. The proposed AACDF framework does not aim to replace security operations but supplement them in order to manage those threats which can be out of the scope of human experts' capabilities.

From the comparison, it can be clearly seen that the proposed framework stands out as unique in comparison to any other existing framework that does not offer an integrated detect-response mechanism combined with native enforcement capabilities and policy decision traceability.

Further research can be carried out on several topics. It is important to carry out empirical validation in production-like settings in order to overcome the architectural evaluation stage. Introducing the framework into current SIEM and SOAR products would facilitate its implementation. Extending the framework to include multi-organizational coalitions where policies would differ between organizations would put the policy abstraction capability of the framework to the test. Finally, introducing some learning capability to the Reasoning Agents based on experience gained by analysts and incidents would tackle the continuous learning challenge.

An approach to practical autonomous cyber defense must involve approaches that do not only work but also are reliable. AACDF is one such attempt towards realizing this goal – an architecture that has been conceived from the bottom up as something that can operate autonomously when needed, that can justify its decisions if called into question, and that will always defer to policy when unable to act. The threat landscape that this effort seeks to address does not have time to wait.

REFERENCES

1. IRSDA: An Agent-Orchestrated Framework for Enterprise Intrusion Response, in Proc. IEEE Int. Conf. on Cyber Security, Macau, China, Dec. 2025, DOI: 10.1109/ICS.2025.11402393.
2. S. Vyas, A. Bolton, V. Mavroudis, and P. Burnap, “Towards the deployment of realistic autonomous cyber network defence: A systematic review,” ACM Computing Surveys, vol. 58, no. 1, pp. 1–36, 2025, DOI: 10.1145/3729213.
3. M. Wang et al., “Multi-agent actor-critics in autonomous cyber defense,” arXiv:2410.09134 [cs.CR], Oct. 2024.
4. Clouseau: A Hierarchical Multi-Agent Approach for Autonomous Attack Investigation, in Proc. IEEE Int. Conf. on Big Data Intelligence and Computing, Honolulu, HI, USA, Dec. 2025, DOI: 10.1109/BDIC.2025.11391875.
5. J. Wiebe, R. Mallah, and L. Li, “Learning cyber defence tactics from scratch with multi-agent reinforcement learning,” arXiv:2310.05939, Oct. 2023.
6. Training Autonomous Cyber Defense Agents: Challenges & Opportunities in Military Networks, in Proc. IEEE Military Communications Conf., Washington, DC, USA, Oct.–Nov. 2024, DOI: 10.1109/MILCOM.2024.10773923.
7. TriShield: Toward an Autonomous and Self-Evolving Cyber Defense via Predictive SOC, Causal Pentest Agents,

© 2026 IJSRET

and Adversarial LLM-CTI, in Proc. IEEE Int. Conf. on Cyber Security, Istanbul, Turkiye, Jun. 2026, DOI: 10.1109/ICS.2026.11648006.

8. M. Cheah, J. Stone, P. Haubrick, S. Bailey, D. Rimmer, D. Till, M. Lacey, J. Kruczynska, and M. Dorn, "CO-DECYBER: Co-operative decision making for cybersecurity using deep multi-agent reinforcement learning," in Computer Security. ESORICS 2023 International Workshops, 2023, DOI: 10.1007/978-3-031-54129-2_37.
9. A Multi-Agent Deep Reinforcement Learning Framework for Detecting and Mitigating DDoS Attacks in IoT Networks, in Proc. IEEE Int. Conf. on Intelligent Systems, Tumkuru, India, Dec. 2024, DOI: 10.1109/ISCON.2024.10872370.
10. AI-Augmented SOC: A Survey of LLMs and Agents for Security Automation, Journal of Cybersecurity and Privacy, vol. 5, no. 4, 2025.

© 2026 IJSRET