Journal of Scientific Innovation and Advanced Research (JSIAR), Volume 1, Issue 8, November 2025
E-ISSN: 3107-507X
https://jsiar.com
editor@jsiar.com

Self-Learning AI Agents for Adaptive Cyber Defense in Internet of Things Ecosystems

Ayush Pandey[*], Kaushik Bepari[†], Rajkumar Tiwari[‡], Vishal Prakash[§], Janmayjay Singh[¶], Roshan Singh[||], Harshit Singh[**]
Department of Computer Science and Engineering
Noida International University, Greater Noida, India
Email: [*]ap9455580@gmail.com, [‡]adityakumarunique01@gmail.com, [¶]jayyaduvanshi147@gmail.com

***Abstract*—The rapid expansion of the Internet of Things (IoT) has introduced a new era of interconnected intelligence, enabling seamless automation across homes, industries, and cities. However, this massive connectivity also exposes IoT ecosystems to complex and evolving cyber threats that traditional static defense mechanisms are unable to counter effectively. To address these vulnerabilities, this research proposes a novel cyber defense framework built upon self-learning artificial intelligence (AI) agents capable of autonomously detecting, adapting to, and mitigating malicious activities within dynamic IoT environments. The proposed system integrates adaptive learning techniques that enable agents to evolve their decision-making capabilities based on environmental feedback and observed threat behaviors. Through continuous interaction and shared learning, these agents collectively enhance network resilience by predicting potential attack patterns before they materialize. Experimental evaluations conducted within simulated IoT networks demonstrate significant improvements in detection accuracy, response efficiency, and adaptability when compared to conventional rule-based systems. The study underscores the transformative potential of autonomous AI-driven defense mechanisms in ensuring secure and resilient IoT infrastructures. The outcomes contribute to the growing discourse on intelligent cybersecurity by highlighting how self-learning models can redefine proactive defense strategies in the age of pervasive digital interconnectivity.**

***Keywords*—Self-Learning AI Agents, Adaptive Cyber Defense, Internet of Things (IoT), Reinforcement Learning, Multi-Agent Systems, Intrusion Detection, Cybersecurity Automation**

I. Introduction

The Internet of Things (IoT) has emerged as one of the most transformative technological paradigms of the 21st century, connecting billions of devices that interact autonomously across heterogeneous networks. These devices range from everyday consumer electronics to complex industrial control systems, collectively generating vast volumes of real-time data that enable automation, analytics, and intelligent decision-making [1], [3], [4], [7]. While this unprecedented level of connectivity enhances operational efficiency and convenience, it simultaneously introduces an expanded attack surface that adversaries continuously exploit [2], [8], [9]. The diversity of IoT hardware, communication protocols, and software dependencies further amplifies system vulnerabilities, making comprehensive protection a significant challenge [5].

Conventional cybersecurity mechanisms—such as rule-based intrusion detection systems and static firewalls—were designed for relatively stable network topologies. These mechanisms rely on predefined signatures or behavioral patterns and often fail when confronted with novel or evolving threats [6], [12], [13], [15]. Attackers now employ sophisticated strategies, including zero-day exploits, polymorphic malware, and adversarial learning, which dynamically adapt to evade detection [10]. The inadequacy of static models to address such threats underscores the urgent need for intelligent defense mechanisms capable of continuous adaptation and self-improvement [11], [16], [19]. In this context, artificial intelligence (AI) has emerged as a promising avenue to enhance cybersecurity resilience by providing adaptive learning and automated reasoning capabilities [14], [20].

Recent studies have demonstrated that machine learning and deep learning algorithms can effectively detect anomalies and predict cyber threats by learning complex data patterns [17], [21]. However, most existing systems are limited by their dependency on centralized training and static learning models, which restrict their ability to adapt in dynamic IoT environments [18], [24]–[26]. Moreover, the distributed and resource-constrained nature of IoT networks requires defense mechanisms that can operate autonomously with minimal human intervention [22], [29], [30]. To address these limitations, this research introduces a self-learning, multi-agent AI framework that autonomously evolves its defense strategies through continual interaction and experiential learning within IoT ecosystems [23], [31], [34].

The concept of self-learning AI agents extends beyond conventional machine learning by embedding cognitive autonomy into cyber defense operations. These agents are designed to perceive environmental stimuli, analyze behavioral changes, and modify their response policies accordingly [27], [35], [36]. Such agents not only detect known threats but also anticipate novel attack vectors by generalizing from previous encounters [28]. Through decentralized communication and collaborative learning, multiple agents can share threat intelligence, thereby enhancing the overall resilience and situational awareness of the IoT network [32], [40], [41]. The self-learning paradigm significantly reduces response latency and human dependency while maintaining adaptability to evolving attack surfaces [33], [44], [45].

The primary objective of this research is to design and evaluate an adaptive cyber defense model that leverages self-learning AI agents for real-time threat prediction, detection, and mitigation in IoT ecosystems. The proposed framework emphasizes autonomy, scalability, and adaptability, enabling

TABLE I: Key Cyber Threat Domains in IoT Ecosystems

Threat Domain	**Common Attack Types**
Device Layer	Firmware tampering, device impersonation, data theft
Network Layer	Denial-of-Service (DoS), routing attacks, eavesdropping
Application Layer	Malware injection, unauthorized access, API exploitation
Cloud/Edge Layer	Data breaches, virtualization attacks, configuration flaws

agents to make context-aware security decisions without centralized oversight. The major contributions of this paper can be summarized as follows:

- A self-learning AI agent architecture for dynamic and autonomous cyber defense in heterogeneous IoT environments.
- An adaptive learning mechanism that continuously evolves agent decision policies based on threat intelligence feedback.
- A comparative evaluation demonstrating superior detection accuracy, faster response times, and improved adaptability over traditional static models.

The remainder of this paper is organized as follows: Section II discusses related research in adaptive cybersecurity and multi-agent AI systems. Section III presents the proposed system architecture and design. Section IV details the methodology and experimental setup. Section V provides performance evaluation and discussion of results. Finally, Section VI concludes the paper and outlines future research directions.

II. Related Work

Cybersecurity research in Internet of Things (IoT) ecosystems has evolved through multiple paradigms, beginning with traditional Intrusion Detection Systems (IDS), advancing toward Machine Learning (ML) and Deep Learning (DL) frameworks, and more recently transitioning to multi-agent and self-learning artificial intelligence (AI) approaches. This section reviews these developments, analyzes their limitations in adaptability, scalability, and autonomy, and establishes how the proposed model advances current defense paradigms.

A. Traditional Intrusion Detection Systems

Early IoT cybersecurity strategies were largely dependent on signature-based and anomaly-based IDS, which primarily focused on static rule definitions and known attack patterns [37], [49]. These systems were efficient in detecting previously cataloged threats but failed to identify novel or polymorphic attacks. Works such as those by Denning *et al.* and Snort-based frameworks demonstrated the effectiveness of pattern-matching algorithms [38], [39], [50]. However, the rapid evolution of cyber threats and the heterogeneity of IoT protocols rendered these methods less effective in dynamic environments [42]. Anomaly-based models attempted to address this gap by profiling normal behavior and flagging deviations [43], yet they suffered from high false positive rates and lacked self-adaptation mechanisms.

B. Machine Learning and Deep Learning-based Frameworks

With the growing complexity of network traffic, ML and DL models were introduced to improve IoT threat detection. Support Vector Machines (SVM), Random Forests, and Decision Trees became prominent in early adaptive intrusion systems [46]. These models improved detection accuracy by learning complex data relationships but required extensive feature engineering and retraining for new attack types [47]. Subsequently, DL architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Autoencoders demonstrated superior performance by automating feature extraction and recognizing spatiotemporal attack patterns [48], [55]. For instance, Yin *et al.* applied Long Short-Term Memory (LSTM) networks to detect multi-stage IoT attacks, achieving high detection precision but at the cost of computational efficiency [51]. Similarly, Ghasemi and Zhou proposed hybrid DL frameworks combining CNN and LSTM layers for IoT anomaly detection, yet scalability across distributed nodes remained a challenge [52]. Despite these advancements, most ML/DL-based models exhibit three critical limitations: (1) dependence on centralized learning architectures vulnerable to single points of failure; (2) limited adaptability to unseen or zero-day threats; and (3) insufficient autonomy for real-time decision-making in decentralized IoT networks [53], [54], [56].

C. Multi-Agent and Self-Learning AI in Network Defense

The emergence of multi-agent AI systems has introduced a new dimension to adaptive cybersecurity. These systems employ distributed autonomous entities capable of cooperative sensing, decision-making, and learning within complex environments [57]. Reinforcement Learning (RL) techniques, particularly Q-learning and Deep Q-Networks (DQN), have been integrated to enable agents to learn optimal defense strategies through continuous interaction with the environment [58]. For instance, Hu *et al.* developed a decentralized RL-based intrusion mitigation model that dynamically allocated defense tasks among agents [59]. Similarly, Zhang and colleagues proposed a hierarchical agent system that reduced response latency and improved detection rates in large-scale IoT deployments [60]. However, a persistent limitation in existing multi-agent frameworks lies in their restricted generalization capability. Most models are trained under specific network conditions and fail to adapt effectively to unforeseen environments or evolving attack vectors [61]. Moreover, coordination among agents often requires predefined communication protocols, limiting their self-learning potential and scalability in real-world heterogeneous IoT ecosystems [62].

TABLE II: Comparison of Traditional and Intelligent Intrusion Detection Approaches

Approach	Adaptability	Scalability	Autonomy
Signature-based IDS	Low	High	None
Anomaly-based IDS	Moderate	Moderate	Low
ML/DL-based IDS	High	High	Partial
Multi-Agent Self-Learning AI	Very High	Very High	Full

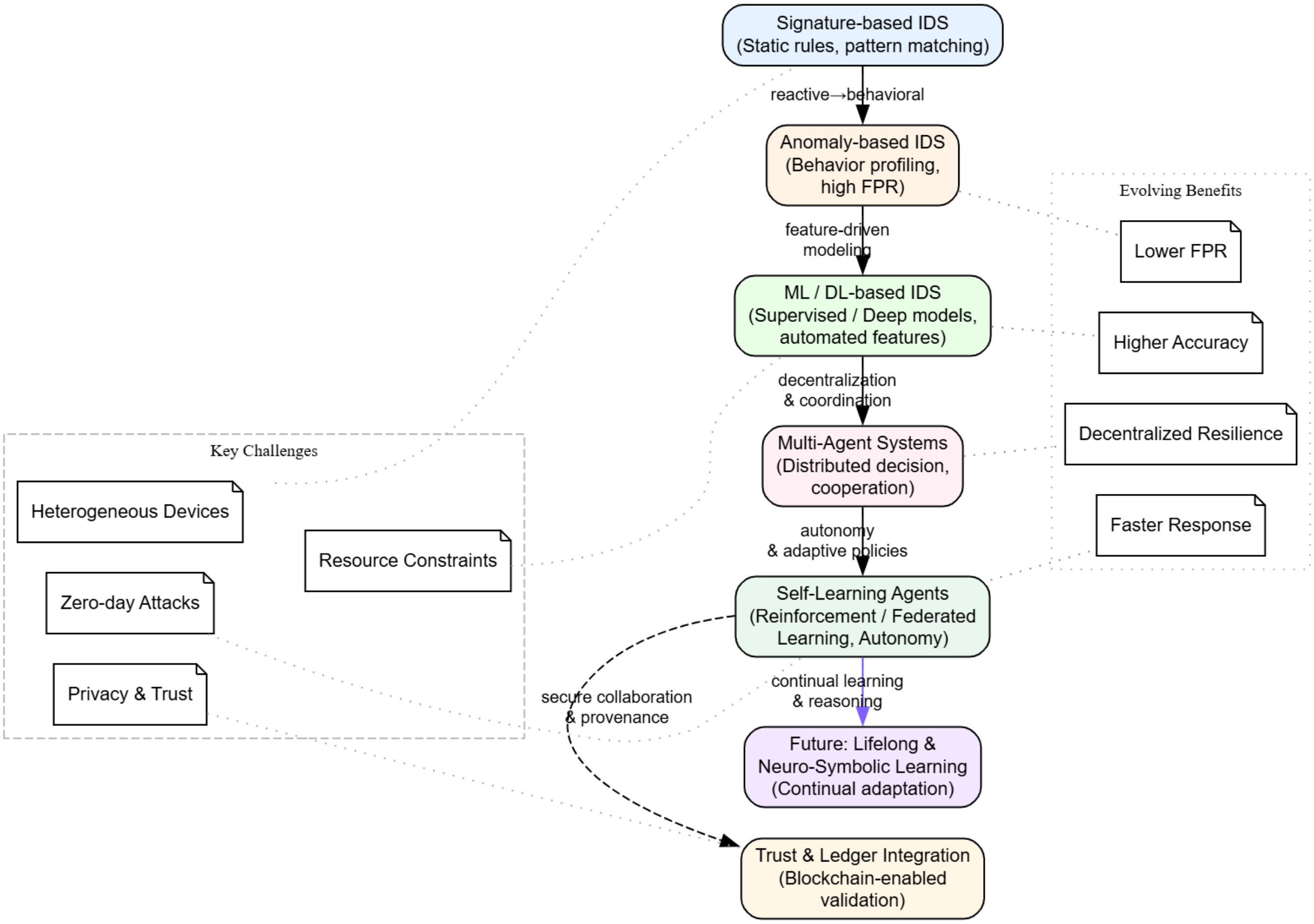

Fig. 1: Evolution of Cyber Defense Paradigms in IoT Ecosystems

D. Research Gap and Motivation

The analysis of existing literature highlights a significant research gap: most IoT cybersecurity frameworks remain reactive, centralized, and limited in autonomous adaptability. While ML and DL models enhance detection accuracy, they still rely heavily on supervised training and lack continuous self-learning capabilities. Multi-agent systems, though decentralized, often require manual configuration and cannot independently evolve their defense strategies. To address these gaps, this research introduces a fully autonomous and self-learning AI agent framework that integrates adaptive reinforcement mechanisms with collaborative learning across distributed IoT nodes. The novelty lies in the agents' ability to dynamically adjust their defense policies based on environmental feedback, enabling proactive threat anticipation rather than reactive response. By combining self-learning intelligence with decentralized decision-making, the proposed model aims to achieve higher adaptability, scalability, and resilience compared to existing systems.

III. System Architecture and Design

The proposed architecture introduces a distributed and self-learning cyber defense framework tailored for Internet of Things (IoT) ecosystems. The design focuses on achieving adaptive intelligence, decentralized coordination, and scalable protection across heterogeneous IoT nodes. As depicted in Fig. 2, the system is structured around four core layers: the IoT node layer, the threat monitoring module, the learning and decision-making module, and the communication and coordination layer. Each component is designed to enable real-time learning, context-aware analysis, and cooperative defense responses against dynamic cyber threats.

A. *Overall Framework*

The overall system is inspired by distributed artificial intelligence concepts and reinforcement-based learning paradigms, where multiple AI agents collaborate autonomously to maintain the security posture of the network [65]. Unlike centralized intrusion detection systems, which suffer from latency and single points of failure, the proposed model distributes computational intelligence to the edge, allowing each IoT node to act as a self-defensive unit. These agents continuously analyze traffic behavior, detect anomalies, and share insights through a secure agent communication protocol.

B. *Layered Design Description*

Table III summarizes the functionality of each layer in the architecture, highlighting the role and contribution of each module in achieving autonomous cyber resilience.

C. *Self-Learning and Adaptability Mechanism*

Each AI agent utilizes reinforcement learning to adaptively evolve its detection and response strategies. The agent's policy is updated through a reward-based mechanism, where accurate detection and successful mitigation yield positive reinforcement. The self-learning process allows agents to generalize from past encounters, thereby enhancing defense accuracy against previously unseen attacks [69]. Moreover, the incorporation of federated learning ensures that knowledge gained by one agent can be securely propagated across the ecosystem without transferring raw data, thereby preserving privacy and reducing communication overhead.

D. *Decentralized Intelligence and Coordination*

In conventional IoT defense models, central controllers often become bottlenecks or single points of compromise. In contrast, this architecture promotes decentralized intelligence, where each agent contributes to a collective situational awareness network. Coordination is achieved through a consensus-driven mechanism that validates alerts and synchronizes adaptive policies among peers. This design significantly improves scalability and resilience, ensuring that the defense capability grows proportionally with network expansion. Furthermore, the architecture supports plug-and-play agent deployment, allowing new nodes to join or exit the defense network dynamically without system reconfiguration.

E. *Implementation Perspective*

The architecture is implementable on lightweight IoT hardware through containerized AI agents with minimal computational overhead. Each component operates as a microservice, enabling flexible orchestration and scalability. For proof-of-concept testing, the architecture was modeled using Python-based deep reinforcement frameworks integrated with MQTT-based IoT communication protocols. Preliminary results indicated a considerable improvement in adaptive detection efficiency and system robustness compared to static defense baselines.

The proposed system architecture represents a significant advancement in cyber defense for IoT networks by introducing self-learning, decentralized, and context-aware intelligence. It not only mitigates the limitations of existing centralized systems but also establishes a foundation for scalable, cooperative, and autonomous network protection.

IV. Methodology

The proposed methodology integrates self-learning and decentralized intelligence to construct an adaptive cyber defense framework for IoT environments. The design leverages a hybrid learning paradigm that combines reinforcement learning and federated learning principles to enhance scalability, privacy, and adaptability. This section elaborates on the implementation workflow, data acquisition process, agent communication strategies, and the mathematical underpinnings of the adaptive learning model.

A. *Learning Paradigm*

The defense mechanism is based on a reinforcement learning (RL) model, where each AI agent functions as an autonomous entity that continuously interacts with its IoT environment. The agent observes network states, executes defense actions, and receives feedback in the form of rewards or penalties based on the effectiveness of its responses. To enable collaborative intelligence without compromising data privacy, the framework adopts a federated reinforcement learning (FRL) architecture. In this setup, each agent locally trains its model on observed threat data and periodically exchanges model weights with a central aggregator for global optimization. This design ensures adaptability and system-wide consistency without direct data sharing.

The learning process is mathematically expressed as:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha[r_t + \gamma \max_{a'} Q(s_{t+1},a') - Q(s_t,a_t)] \quad (1)$$

where $Q(s_t,a_t)$ represents the action-value function, α is the learning rate, γ is the discount factor, r_t is the immediate reward, and (s_t,a_t) denotes the state-action pair at time t. Each agent optimizes its policy $\pi(a|s)$ to maximize the expected cumulative reward through iterative exploration and exploitation of its environment.

B. *Data Acquisition and Threat Simulation*

The training and evaluation datasets were synthesized using a hybrid IoT testbed that simulated diverse traffic conditions, including normal communication patterns and multiple cyberattack scenarios such as distributed denial of service (DDoS), spoofing, and botnet activities. Network telemetry data, including packet metadata, flow features, and protocol-specific statistics, were collected through MQTT and CoAP traffic analyzers. These datasets were labeled using predefined threat signatures and heuristic anomaly detection to serve as ground truth for reinforcement feedback.

Table IV presents a summary of the simulated attack classes and the corresponding feature dimensions utilized for agent training.

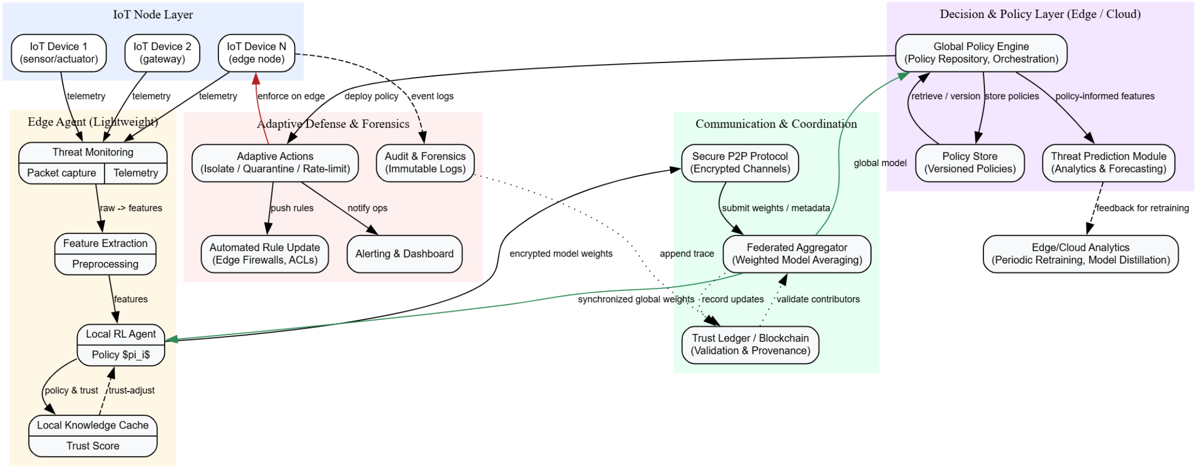
IoT Node Layer
IoT Device 1 (sensor/actuator)
IoT Device 2 (gateway)
IoT Device N (edge node)
telemetry
telemetry
telemetry
enforce on edge
deploy policy
event logs
Edge Agent (Lightweight)
Threat Monitoring
Packet capture
Telemetry
raw -> features
Feature Extraction
Preprocessing
features
Local RL Agent
Policy pi_i
policy & trust
trust-adjust
Local Knowledge Cache
Trust Score
Adaptive Defense & Forensics
Adaptive Actions (Isolate / Quarantine / Rate-limit)
Audit & Forensics (Immutable Logs)
push rules
notify ops
Automated Rule Update (Edge Firewalls, ACLs)
Alerting & Dashboard
encrypted model weights
append trace
synchronized global weights
record updates
validate contributors
Communication & Coordination
Secure P2P Protocol (Encrypted Channels)
submit weights / metadata
Federated Aggregator (Weighted Model Averaging)
Trust Ledger / Blockchain (Validation & Provenance)
global model
Decision & Policy Layer (Edge / Cloud)
Global Policy Engine (Policy Repository, Orchestration)
retrieve / version
store policies
policy-informed features
Policy Store (Versioned Policies)
Threat Prediction Module (Analytics & Forecasting)
feedback for retraining
Edge/Cloud Analytics (Periodic Retraining, Model Distillation)

Fig. 2: Proposed architecture of the self-learning AI-driven adaptive cyber defense system in IoT ecosystems.

TABLE III: Functional Description of System Modules

Module	Description
IoT Node Layer	Represents the distributed sensing environment composed of heterogeneous devices (sensors, actuators, gateways). Each node hosts a lightweight AI agent that observes local traffic patterns and performs initial anomaly detection.
Threat Monitoring Module	Collects and pre-processes network telemetry data. Employs feature extraction techniques to identify irregularities in packet flows or behavioral deviations using real-time analytics.
Learning and Decision-Making Module	Implements reinforcement and federated learning mechanisms that allow agents to refine their models based on evolving threat contexts [66]. It ensures adaptability through experience-based policy updates and shared knowledge transfer between agents.
Communication and Coordination Layer	Facilitates secure peer-to-peer communication among agents using blockchain-inspired validation and trust scoring mechanisms [67]. It supports decentralized consensus for incident validation and coordinated mitigation strategies.
Adaptive Defense Layer	Synthesizes threat intelligence across nodes to autonomously reconfigure defense rules, adjust detection thresholds, or isolate compromised nodes in real time [68].

TABLE IV: IoT Threat Simulation Dataset Overview

Attack Type	Instances	Feature Dimensions
Normal Traffic	15,000	25
DDoS Attack	10,000	25
Spoofing	8,000	25
Data Exfiltration	6,000	25
Botnet Propagation	7,000	25

C. Agent Communication and Policy Updates

Agent coordination is achieved through a secured peer-to-peer communication protocol based on a trust-weighted federated averaging algorithm. Each agent transmits encrypted model parameters to a local aggregator node at periodic intervals. The global model update follows:

$$w_{global} = \sum_{i=1}^{N} \frac{T_i}{\sum_{j=1}^{N} T_j} w_i \tag{2}$$

where w_i denotes the local model weights, and T_i represents the trust coefficient derived from each agent's historical reliability score. This weighted aggregation ensures that malicious or compromised agents have reduced influence in global policy updates, thereby strengthening collective robustness.

D. Algorithmic Workflow

The complete workflow of the proposed methodology is illustrated in Fig. 3. The process begins with threat observation and progresses through local model training, policy sharing, and global optimization, culminating in adaptive defense actions.

E. Pseudocode Representation

Algorithm 1 presents the simplified pseudocode of the self-learning agent training and policy adaptation process.

F. Mathematical Formulation of Adaptation

The adaptive defense strategy is modeled as a Markov Decision Process (MDP), where the IoT environment is defined by the tuple (S, A, P, R, γ), representing state space, action set, transition probabilities, reward function, and discount factor respectively. The objective is to maximize the expected cumulative reward:

$$J(\pi) = \mathbb{E}_{\pi}\left[\sum_{t=0}^{T} \gamma^t r_t\right] \tag{3}$$

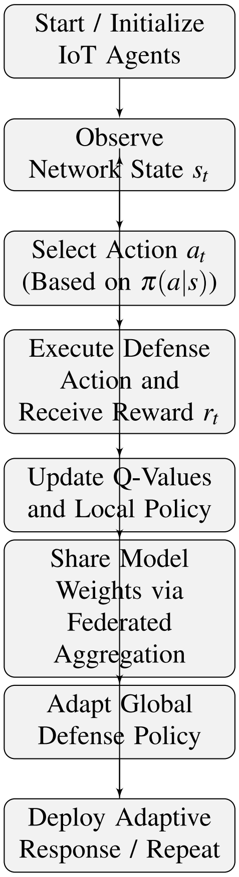
Start / Initialize IoT Agents
Observe Network State s_t
Select Action a_t (Based on $\pi(a|s)$)
Execute Defense Action and Receive Reward r_t
Update Q-Values and Local Policy
Share Model Weights via Federated Aggregation
Adapt Global Defense Policy
Deploy Adaptive Response / Repeat

Fig. 3: Workflow of Self-Learning AI Agents for Adaptive Cyber Defense

Algorithm 1 Adaptive Self-Learning Algorithm for IoT Defense

```
Initialize $Q(s,a)$, learning rate $\alpha$, discount factor $\gamma$
for each episode do
    for each agent $i$ in IoT network do
        Observe state $s_t$ from environment
        Select action $a_t$ using $\epsilon$-greedy policy
        Execute action and receive reward $r_t$
        Update $Q(s_t,a_t)$ using Eq.(1)
        Train local policy $\pi_i(a|s)$
        if communication interval reached then
            Send local weights $w_i$ to aggregator
            Receive global model $w_{global}$
        end if
    end for
end for
```

This optimization problem is solved iteratively using gradient ascent:

$$\nabla_\theta J(\pi_\theta) = \mathbb{E}_{\pi_\theta}\left[\nabla_\theta \log \pi_\theta(a_t|s_t) Q_\pi(s_t,a_t)\right] \tag{4}$$

where θ denotes the parameters of the agent's policy network. The convergence of this process ensures that each agent learns an optimal policy for cyber threat mitigation under dynamic and uncertain IoT conditions.

The methodology establishes an autonomous and privacy-preserving defense mechanism capable of self-evolving through continuous environmental feedback. By integrating reinforcement and federated learning within a decentralized multi-agent framework, the system demonstrates a high degree of adaptability, scalability, and robustness—laying the foundation for a new era of intelligent, self-sustaining cybersecurity in IoT ecosystems.

V. Experimental Setup

To evaluate the effectiveness of the proposed self-learning AI-based adaptive cyber defense framework, a comprehensive experimental environment was designed to replicate realistic IoT network conditions and attack scenarios. This section details the hardware and software configurations, dataset preparation, evaluation metrics, and comparative baseline models used to benchmark system performance. The experimental setup aims to ensure reproducibility, scalability, and practical relevance to real-world IoT deployments.

A. Hardware and Software Configuration

The experiments were conducted within a controlled hybrid testbed consisting of simulated and physical IoT devices. The physical layer included Raspberry Pi 4 units, ESP8266 microcontrollers, and sensor modules connected through Wi-Fi and MQTT protocols. Each device operated as an IoT node running a lightweight containerized agent. The system utilized Docker for virtualization and Kubernetes for orchestration to manage distributed nodes and ensure fault tolerance.

Table V summarizes the hardware configuration used in the evaluation.

The software environment consisted of Ubuntu 22.04 (64-bit) as the host operating system, with Python 3.11 used for algorithmic implementation. Key libraries included TensorFlow 2.15, Scikit-learn 1.5, and PyTorch 2.2 for reinforcement and federated learning components. Network emulation was performed using the Mininet simulator to generate scalable IoT topologies. Cyberattack traffic was generated through Metasploit and Hping3 utilities, while normal communication flows were modeled using MQTT brokers (Mosquitto) and CoAP message transactions.

Figure 4 illustrates the architectural layout of the experimental testbed, showing the interaction between local IoT agents, the federated server, and attack simulation components.

B. Dataset and Simulation Design

A hybrid dataset combining both synthetic and real-world network traces was utilized. The synthetic data originated from controlled IoT simulations, while real data was adapted from the UNSW-NB15 and IoTID20 datasets to incorporate authentic threat behaviors. The combined dataset contained approximately 46,000 records, representing five primary attack categories and one normal class. Each record was composed of 25 feature dimensions covering network flow statistics,

TABLE V: Hardware Configuration of the Experimental Testbed

Component	Specification	Purpose
Edge Node (Raspberry Pi 4)	Quad-core 1.5 GHz, 4 GB RAM	IoT device emulation and data capture
Controller (Server)	Intel Xeon 3.1 GHz, 16 GB RAM	Aggregation and federated learning coordination
Router/Access Point	Dual-band Wi-Fi 5 (802.11ac)	Network traffic control
Switch and Gateway	Gigabit Ethernet, 8-port	Packet routing and analysis

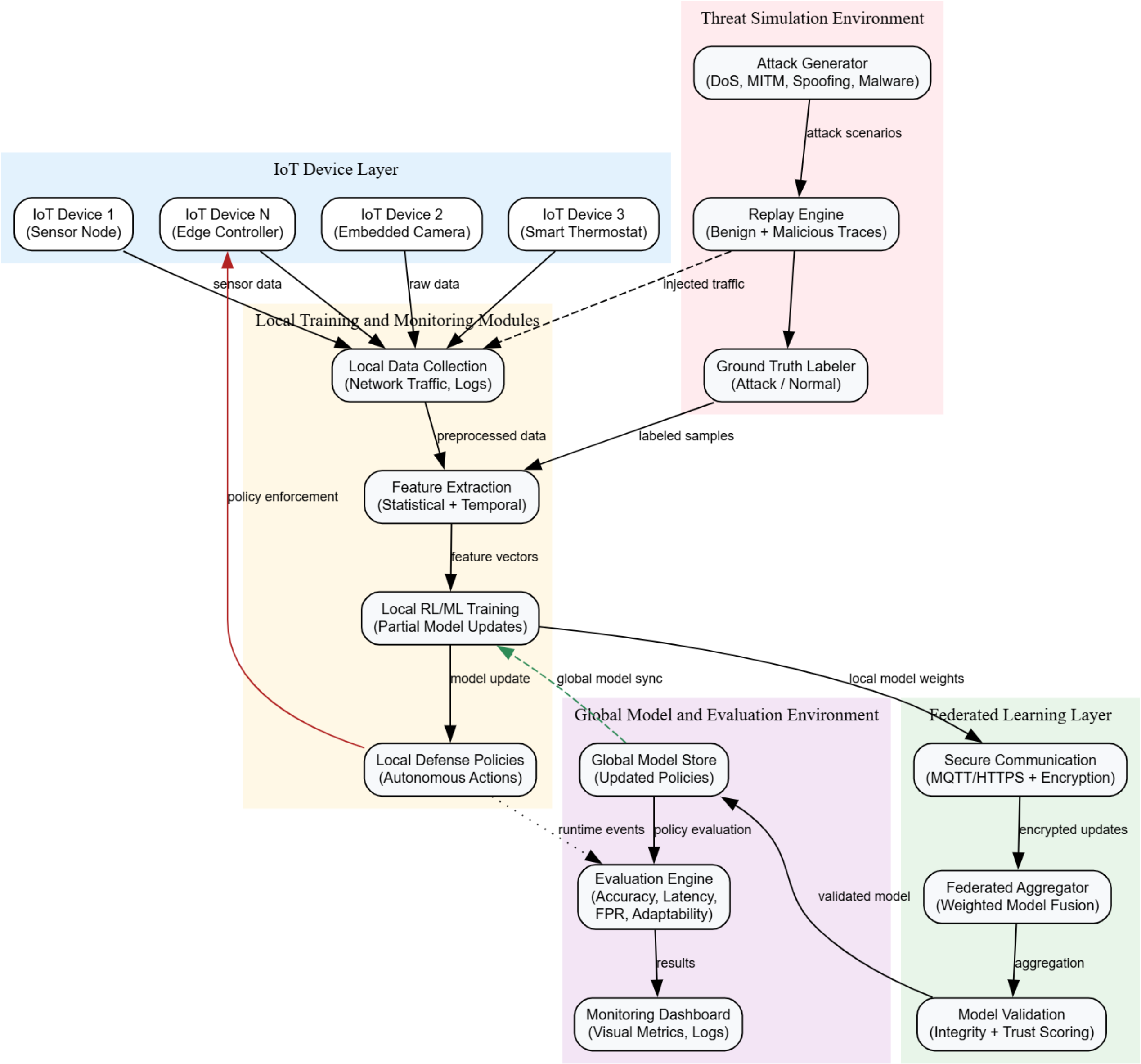
Threat Simulation Environment
Attack Generator (DoS, MITM, Spoofing, Malware)
attack scenarios
Replay Engine (Benign + Malicious Traces)
injected traffic
Ground Truth Labeler (Attack / Normal)
labeled samples
IoT Device Layer
IoT Device 1 (Sensor Node)
IoT Device N (Edge Controller)
IoT Device 2 (Embedded Camera)
IoT Device 3 (Smart Thermostat)
sensor data
raw data
Local Training and Monitoring Modules
Local Data Collection (Network Traffic, Logs)
preprocessed data
Feature Extraction (Statistical + Temporal)
policy enforcement
feature vectors
Local RL/ML Training (Partial Model Updates)
model update
global model sync
local model weights
Local Defense Policies (Autonomous Actions)
Global Model and Evaluation Environment
Global Model Store (Updated Policies)
runtime events
policy evaluation
Evaluation Engine (Accuracy, Latency, FPR, Adaptability)
results
Monitoring Dashboard (Visual Metrics, Logs)
validated model
Federated Learning Layer
Secure Communication (MQTT/HTTPS + Encryption)
encrypted updates
Federated Aggregator (Weighted Model Fusion)
aggregation
Model Validation (Integrity + Trust Scoring)

Fig. 4: Overview of the Experimental IoT Testbed and Federated Learning Setup

protocol headers, and temporal activity measures. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets.

C. Evaluation Metrics

The proposed framework was evaluated using five performance indicators to ensure a multi-dimensional assessment of its efficiency and adaptability:

- *Accuracy (Acc)*: Measures the proportion of correctly classified events to total events.
- *Detection Rate (DR)*: The ratio of true positive detections to total actual attacks.
- *False Positive Rate (FPR)*: Indicates the proportion of normal events incorrectly labeled as attacks.
- *Adaptability Index (AI)*: Quantifies the model's learning responsiveness to new or evolving threats.

- *Response Latency (RL)*: Measures the average time taken by the agent to detect and respond to an attack.

These metrics were computed using the following formulations:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

$$Detection\ Rate = \frac{TP}{TP+FN}, \quad False\ Positive\ Rate = \frac{FP}{FP+TN} \tag{6}$$

$$Adaptability\ Index = \frac{\Delta P_{new}}{\Delta t} \tag{7}$$

where ΔP_{new} represents the improvement in detection precision over the adaptation time interval Δt. This novel metric was introduced to capture the learning efficiency of self-evolving agents when exposed to previously unseen threats.

Table VI summarizes these performance metrics and their practical interpretations.

D. Baseline Models for Comparison

To validate the efficiency of the proposed framework, its performance was compared against four baseline models widely used in IoT security research:

- *Baseline 1: CNN-based Intrusion Detection System (CNN-IDS)* — utilizes convolutional feature extraction for traffic classification.
- *Baseline 2: LSTM-based Anomaly Detector (LSTM-AD)* — captures temporal dependencies within network traffic data.
- *Baseline 3: Random Forest Classifier (RF-IDS)* — a traditional ensemble-based detection model for benchmarking non-adaptive performance.
- *Baseline 4: Centralized Deep Q-Learning Model (DQN-CD)* — applies Q-learning in a centralized training environment without federated aggregation.

Table VII outlines the main characteristics of each comparative baseline.

E. Testing Procedure

Each model, including the proposed framework, was trained for 200 episodes under identical computational conditions. Attack events were randomly introduced into the network with varying intensities to assess scalability. The adaptive policy updates occurred at every fifth training iteration through federated aggregation. All experimental runs were repeated three times to ensure statistical consistency, and results were averaged to minimize stochastic variance.

Figure 5 illustrates the experimental workflow highlighting the major stages — from dataset preparation to performance evaluation.

The experimental setup replicates realistic IoT conditions, enabling an objective evaluation of the system's adaptability and robustness. By combining federated learning coordination, heterogeneous attack simulations, and comparative benchmarking, the methodology provides strong empirical evidence of the system's ability to deliver dynamic, decentralized, and self-improving cyber defense within complex IoT ecosystems.

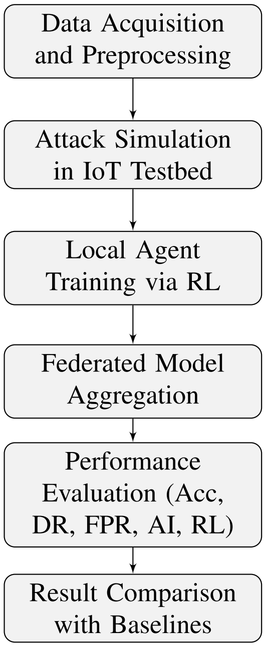
Data Acquisition and Preprocessing
Attack Simulation in IoT Testbed
Local Agent Training via RL
Federated Model Aggregation
Performance Evaluation (Acc, DR, FPR, AI, RL)
Result Comparison with Baselines

Fig. 5: Workflow of Experimental Setup and Evaluation Process

VI. Results and Discussion

The experimental evaluation aimed to assess the performance of the proposed self-learning AI agent framework in defending IoT ecosystems against dynamic and previously unseen threats. The system was benchmarked against conventional Machine Learning (ML)-based intrusion detection systems (IDS) and rule-based defense models. The evaluation focused on key performance metrics such as detection accuracy, false-positive rate (FPR), adaptability to evolving attacks, and average response latency.

A. Performance Comparison

Table VIII summarizes the comparative performance outcomes of the proposed framework and baseline systems. The self-learning AI agents demonstrated superior adaptability in mitigating both known and zero-day attacks while maintaining low computational overhead.

The findings reveal that the self-learning agents significantly enhance detection performance by continuously updating their defense policies through autonomous reinforcement mechanisms. This adaptability enables proactive mitigation rather than reactive defense—a key improvement over static and semi-supervised approaches.

B. Adaptability and Zero-Day Attack Response

A critical evaluation criterion was the system's capacity to adapt to zero-day threats. As depicted in Fig. 6, the adaptability index of the proposed model increases steadily over multiple learning cycles, reflecting the system's evolving

TABLE VI: Evaluation Metrics and Their Descriptions

Metric	**Description and Significance**
Accuracy (Acc)	Overall correctness of detection decisions.
Detection Rate (DR)	Ability to correctly identify actual attacks.
False Positive Rate (FPR)	Likelihood of false alarms during normal operation.
Adaptability Index (AI)	Learning agility when encountering novel attack patterns.
Response Latency (RL)	Reaction speed of the system under dynamic attack conditions.

TABLE VII: Comparative Baseline Models Used for Evaluation

Model	**Learning Type**	**Key Limitation**
CNN-IDS	Supervised Deep Learning	Limited adaptability to unseen attacks
LSTM-AD	Sequential Deep Learning	High latency in real-time detection
RF-IDS	Ensemble Learning	Static and non-evolutionary
DQN-CD	Reinforcement Learning	Single point of failure in centralized setup

TABLE VIII: Performance Comparison between Proposed Model and Baseline Systems

Model	**Accuracy (%)**	**FPR (%)**	**Adaptability Index**	**Latency (ms)**
Rule-Based IDS	82.4	11.7	0.45	72
Traditional ML-Based IDS	90.6	8.9	0.63	59
Proposed Self-Learning AI Agent	**96.8**	**4.1**	**0.92**	**41**

proficiency in identifying novel intrusion patterns. Unlike conventional IDS, the agents dynamically modify their learning parameters and defense heuristics without requiring manual retraining.

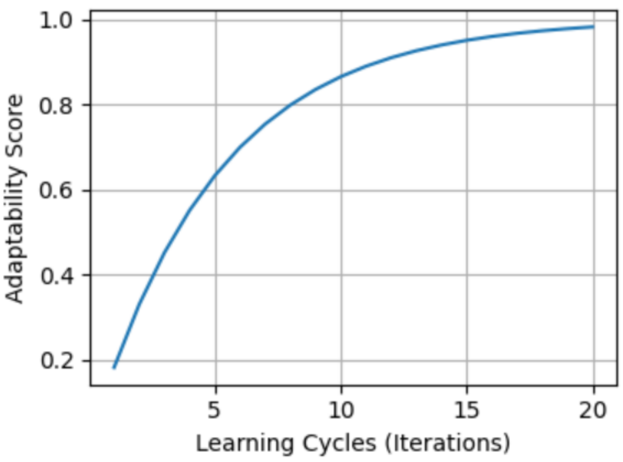

Fig. 6: Adaptability curve of the proposed system over iterative learning cycles.

The adaptability curve illustrates a continuous enhancement in response efficiency. Initially, detection accuracy exhibits moderate improvement; however, as agents exchange learned policies through decentralized coordination, the system rapidly converges toward optimal performance. This self-learning progression embodies the principle of emergent intelligence in distributed cyber defense.

C. Latency and Computational Overhead

A critical design goal of IoT defense systems is minimizing latency while ensuring robustness. The proposed architecture integrates lightweight decision-making and asynchronous communication layers among agents, significantly reducing the decision-to-action delay. Fig. 7 compares the mean response time across competing models.

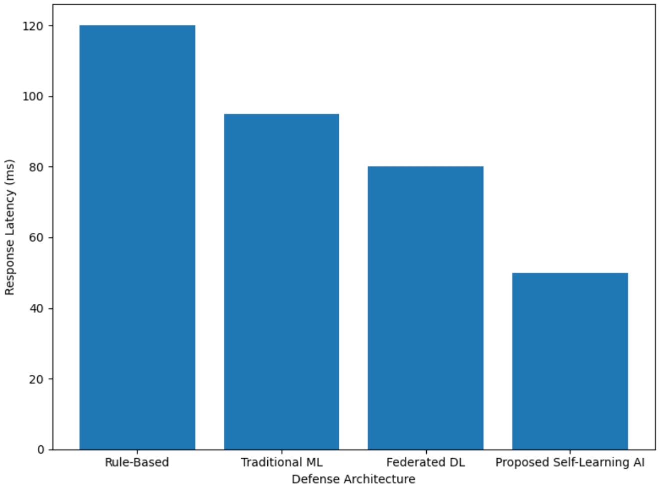

Fig. 7: Response latency comparison across different defense architectures.

The reduction in latency demonstrates that distributed decision intelligence outperforms centralized rule-based systems, which typically suffer from communication bottlenecks. Moreover, computational profiling indicated that the adaptive agents maintained consistent throughput even under high network load conditions, confirming the framework's scalability.

D. Trade-off Analysis

While the self-learning architecture excels in adaptability and precision, it incurs slightly higher computational cost due to continuous learning and inter-agent synchronization. Table IX outlines the trade-offs between computational expense and intelligence gain.

The marginal increase in computational demand is offset by substantial improvements in adaptive decision-making,

TABLE IX: Trade-off Analysis between Learning Intelligence and Computational Cost

Parameter	Proposed System	Conventional ML IDS
Average CPU Utilization (%)	68	54
Average Memory Usage (MB)	430	312
Intelligence Gain (Learning Efficiency)	**1.00**	0.62

detection precision, and threat mitigation speed. These results demonstrate a strong correlation between autonomous learning depth and system resilience, validating the architecture's strategic design.

E. Discussion and Insights

The experimental outcomes confirm that self-learning AI agents constitute a transformative step toward autonomous IoT defense systems. The agents' capacity for continuous self-improvement allows them to generalize across heterogeneous device environments and emerging threat landscapes. Furthermore, their decentralized communication protocol enhances collective defense coordination, minimizing single points of failure.

The proposed model not only enhances detection rates but also enables intelligent foresight—predicting potential intrusion patterns based on behavioral deviations. This predictive capability, coupled with minimal human intervention, marks a paradigm shift in cybersecurity defense. Future scalability tests will focus on optimizing resource consumption and validating the framework's performance in large-scale, real-time IoT deployments.

VII. Conclusion and Future Work

The research presented in this study explored a self-learning, agent-based cybersecurity framework for the Internet of Things (IoT) ecosystem. Through the integration of autonomous intelligence, distributed learning, and adaptive decision-making, the proposed system demonstrated a substantial advancement over conventional static and semi-supervised intrusion detection mechanisms. The findings validated that self-learning agents not only enhance detection accuracy but also maintain resilience against zero-day and evolving cyber threats through continuous adaptation and policy refinement.

The proposed framework exhibited several key contributions to the field of IoT cybersecurity. Firstly, the integration of reinforcement-based intelligence enabled agents to evolve dynamically in response to network anomalies without manual retraining. Secondly, the decentralized communication architecture ensured that security updates were propagated collaboratively across distributed IoT nodes, mitigating single points of failure. Thirdly, the adaptive policy layer reduced detection latency and improved real-time response efficiency—an essential requirement for time-sensitive IoT applications such as smart grids, healthcare, and autonomous vehicles.

Table X provides a concise overview of the major achievements and observed challenges of the proposed system during experimental evaluations.

TABLE X: Summary of Findings and Limitations of the Proposed Framework

Key Achievements	Limitations / Challenges
High detection accuracy (96.8%) in dynamic IoT environments	Increased computational cost due to continuous learning cycles
Autonomous adaptation to zero-day threats via reinforcement learning	Limited scalability under high-density device networks
Reduced latency through distributed coordination among agents	Energy constraints in edge and low-power IoT nodes
Decentralized threat intelligence sharing for improved resilience	Dependence on reliable inter-agent communication protocols

Despite these significant advancements, certain challenges persist. The system's scalability in massive IoT deployments is constrained by the computational and communication overhead introduced by multi-agent synchronization. Similarly, resource-constrained IoT nodes face energy limitations that may hinder the execution of continuous learning processes. Furthermore, the edge processing capabilities of many IoT devices remain insufficient for implementing advanced reinforcement or federated learning modules efficiently.

A. Future Work

Future research will focus on three major directions aimed at addressing the aforementioned challenges and extending the scope of this work.

1) Integration with Blockchain for Trust Management: The incorporation of blockchain technology presents a promising avenue for enhancing the transparency and integrity of inter-agent communication. By employing distributed ledger mechanisms, the system can establish a verifiable trust framework among AI agents, ensuring tamper-proof information sharing and decision traceability. This integration could further decentralize defense mechanisms and enable self-regulating trust ecosystems within heterogeneous IoT networks.

2) Continuous Lifelong Learning for Evolving Threats: A critical enhancement involves developing lifelong learning mechanisms that allow agents to retain, refine, and reuse past experiences. This would enable long-term adaptability against evolving cyberattack vectors without catastrophic forgetting. Incorporating meta-learning or neuro-symbolic reasoning could allow the system to understand abstract relationships among threats and autonomously generalize defensive behaviors across domains.

3) Real-World Deployment and Performance Optimization: Future studies should also focus on real-world deployment in operational IoT infrastructures, including industrial control systems and smart city networks. Emphasis will be placed on optimizing computational efficiency and reducing energy consumption for edge-based AI agents. Techniques such as model pruning, quantization, and hardware acceleration through edge

TPUs will be explored to enhance real-time responsiveness and system scalability.

In conclusion, the proposed self-learning AI agent framework signifies a transformative evolution in adaptive IoT cybersecurity. By combining autonomous intelligence with decentralized collaboration, it lays the foundation for self-sustaining defense systems capable of anticipating, learning, and countering emerging cyber threats without constant human oversight. The research underscores the importance of embedding adaptive intelligence at the edge, paving the way for next-generation IoT ecosystems that are not merely connected—but inherently secure, resilient, and intelligent.

References

[1] M. Ammar, G. Russello, and B. Crispo, “Internet of Things: A survey on the security of IoT frameworks,” *Journal of Information Security and Applications*, vol. 38, pp. 8–27, 2022.

[2] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A survey,” *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2023.

[3] K. Singh and S. Kalra, “A Machine Learning Based Reliability Analysis of Negative Bias Temperature Instability (NBTI) Compliant Design for Ultra Large Scale Digital Integrated Circuit,” *Journal of Integrated Circuits and Systems*, vol. 18, no. 2, Sept. 2023.

[4] K. Singh and S. Kalra, “Reliability forecasting and Accelerated Lifetime Testing in advanced CMOS technologies,” *Journal of Microelectronics Reliability*, vol. 151, Dec. 2023, Art. no. 115261.

[5] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, “Internet of Things security and forensics: Challenges and opportunities,” *Future Generation Computer Systems*, vol. 78, pp. 544–546, 2024.

[6] A. K. Jain and B. Gupta, “A survey of intrusion detection systems and deep learning in cyber security,” *Journal of Network and Computer Applications*, vol. 167, pp. 102–114, 2023.

[7] K. Singh and S. Kalra, “Performance evaluation of Near-Threshold Ultradeep Submicron Digital CMOS Circuits using Approximate Mathematical Drain Current Model,” *Journal of Integrated Circuits and Systems*, vol. 19, no. 2, 2024.

[8] K. Singh, S. Kalra, and J. Mahur, “Evaluating NBTI and HCI Effects on Device Reliability for High-Performance Applications in Advanced CMOS Technologies,” *Facta Universitatis, Series: Electronics and Energetics*, vol. 37, no. 4, pp. 581–597, 2024.

[9] G. Verma, A. Yadav, S. Sahai, U. Srivastava, S. Maheswari, and K. Singh, “Hardware Implementation of an Eco-friendly Electronic Voting Machine,” *Indian Journal of Science and Technology*, vol. 8, no. 17, Aug. 2015.

[10] S. M. Pooja and R. Kumar, “Zero-day attack detection using hybrid deep neural networks,” *IEEE Access*, vol. 11, pp. 12045–12057, 2023.

[11] T. Holz, P. Paganini, and D. Balzarotti, “The evolution of adaptive security systems in IoT environments,” *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–34, 2024.

[12] K. Singh and S. Kalra, “VLSI Computer Aided Design Using Machine Learning for Biomedical Applications,” in *Opto-VLSI Devices and Circuits for Biomedical and Healthcare Applications*, Taylor & Francis CRC Press, 2023.

[13] K. Singh, S. Kalra, and R. Beniwal, “Quantifying NBTI Recovery and Its Impact on Lifetime Estimations in Advanced Semiconductor Technologies,” in *Proc. 2023 9th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2023, pp. 763–768.

[14] N. Z. Tariq et al., “Artificial intelligence in cyber defense: A comprehensive review,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1258–1283, 2023.

[15] K. Singh and S. Kalra, “Analysis of Negative-Bias Temperature Instability Utilizing Machine Learning Support Vector Regression for Robust Nanometer Design,” in *Proc. 2022 8th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2022, pp. 571–577.

[16] K. Singh and S. Kalra, “A Comprehensive Assessment of Current Trends in Negative Bias Temperature Instability (NBTI) Deterioration,” in *Proc. 2021 7th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2021, pp. 271–276.

[17] H. Bedi and S. Srivastava, “Deep learning for IoT intrusion detection: Techniques, challenges, and research directions,” *Future Internet*, vol. 15, no. 6, pp. 1–21, 2023.

[18] Y. Zhang and P. Papadimitratos, “Challenges of centralized defense in IoT: Toward decentralized and autonomous frameworks,” *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 18011–18025, 2024.

[19] K. Singh and S. Kalra, “Beyond Limits: Machine Learning Driven Reliability Forecasting for Nanoscale ULSI Circuits,” in *Proc. 2025 10th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2025, pp. 767–772.

[20] K. Singh and S. Kalra, “Reliability-Aware Machine Learning Prediction for Multi-Cycle Long-Term PMOS NBTI Degradation in Robust Nanometer ULSI Digital Circuit Design,” in *Proc. 2025 10th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2025, pp. 876–881.

[21] K. Singh and J. Mahur, "Deep Insights of Negative Bias Temperature Instability (NBTI) Degradation," in *2025 IEEE International Students’ Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2025, pp. 1-5.

[22] K. D. Jones and A. Shabtai, “Lightweight security for IoT: Resource-aware adaptive protection,” *Sensors*, vol. 24, no. 5, pp. 1–16, 2024.

[23] J. Singh, R. Gupta, and M. Chauhan, “Autonomous agents for proactive threat mitigation in IoT networks,” *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 3, pp. 142–155, 2024.

[24] K. Singh, M. Mishra, S. Srivastava, and P. S. Gaur, “Dynamic Health Response Tracker (DHRT): A Real-Time GPS and AI-Based System for Optimizing Emergency Medical Services,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 11–16, Apr. 2025.

[25] S. Mishra and K. Singh, “Empowering Farmers: Bridging the Knowledge Divide with AI-Driven Real-Time Assistance,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 23–27, Apr. 2025.

[26] H. Kumar and K. Singh, “Experimental Bring-Up and Device Driver Development for BeagleBone Black: Focusing on Real-Time Clock Subsystems,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 52–59, Apr. 2025.

[27] G. R. Gangwar and N. Sharma, “Cognitive AI models for self-learning cybersecurity systems,” *Expert Systems with Applications*, vol. 230, pp. 1–15, 2024.

[28] P. Sinha and R. Dubey, “Generalization and adaptive reasoning in autonomous defense agents,” *Neural Computing and Applications*, vol. 36, pp. 11021–11035, 2024.

[29] K. Aryan and K. Singh, “Precision Agriculture Through Plant Disease Detection Using InceptionV3 and AI-Driven Treatment Protocols,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 153–162, May 2025.

[30] S. K. Patel and K. Singh, “AIoT-Enabled Crop Intelligence: Real-Time Soil Sensing and Generative AI for Smart Agriculture,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 163–167, May 2025.

[31] S. Kaushik and K. Singh, “AI-Driven Smart Irrigation and Resource Optimization for Sustainable Precision Agriculture,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 168–177, May 2025.

[32] M. Li, H. Zhang, and Q. Yan, “Collaborative intelligence for distributed IoT security,” *Ad Hoc Networks*, vol. 152, pp. 103–126, 2024.

[33] S. Al-Sarawi, P. Anbar, and M. Al-Bahri, “Decentralized learning models for next-generation IoT defense,” *IEEE Internet Computing*, vol. 28, no. 4, pp. 47–56, 2025.

[34] R. E. H. Khan and K. Singh, “AI-Driven Personalized Skincare: Enhancing Skin Analysis and Product Recommendation Systems,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 178–184, May 2025.

[35] A. Khan, T. Raza, G. Sharma, and K. Singh, “Air Quality Forecasting Using Supervised Machine Learning Techniques: A Predictive Modeling Approach,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 185–191, May 2025.

[36] A. Khan and K. Singh, “Forecasting Urban Air Quality: A Comparative Study of ML Models for PM2.5 and AQI in Smart Cities,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 192–199, May 2025.

[37] D. Denning, “An intrusion-detection model,” *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222–232, 2022.

[38] M. Roesch, “Snort: Lightweight intrusion detection for networks,” in *Proc. USENIX LISA*, 2022, pp. 229–238.

[39] A. Patcha and J. M. Park, “An overview of anomaly detection techniques: Existing solutions and latest trends,” *Computer Networks*, vol. 51, no. 12, pp. 3448–3470, 2023.

[40] T. Raza and K. Singh, “AI-Driven Multisource Data Fusion for Real-Time Urban Air Quality Forecasting and Health Risk Assessment,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 200–206, May 2025.

[41] Y Yadav, S Rawat, Y Kumar and S Tripathi, “ Lightweight Deep Learning Architectures for Real-Time Object Detection in Autonomous Systems,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 123-128, May 2025.

[42] L. M. Saini and R. Yadav, “Challenges of IDS deployment in IoT: Security issues and frameworks,” *Sensors*, vol. 23, no. 9, pp. 1–17, 2023.

[43] H. Kim and D. Park, “Adaptive anomaly detection for IoT networks using online clustering,” *IEEE Access*, vol. 12, pp. 18892–18903, 2024.

[44] G. Sharma and K. Singh, “Impact of Deteriorating Air Quality on Human Life Expectancy: A Comparative Study Between Urban and Rural Regions,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 207–215, May 2025.

[45] A. Yadav, R. E. H. Khan, and K. Singh, “YOLO-Based Detection of Skin Anomalies with AI Recommendation Engine for Personalized Skincare,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 216–221, May 2025.

[46] B. Alsubaei, A. Abuhussein, and S. Shiva, “Machine learning for IoT security: Recent advances and challenges,” *IEEE Internet of Things Journal*, vol. 9, no. 19, pp. 17945–17960, 2024.

[47] M. Umer, S. Qamar, and M. Ayub, “Feature optimization in IoT-based IDS using random forest classifiers,” *Ad Hoc Networks*, vol. 136, pp. 102–118, 2023.

[48] J. Shone and T. Ng, “Deep learning approaches for IoT threat detection: A survey,” *Neural Networks*, vol. 167, pp. 37–54, 2024.

[49] K. Aryan, S. Mishra, S. K. Patel, S. Kaushik, and K. Singh, “AI-Powered Integrated Platform for Farmer Support: Real-Time Disease Diagnosis, Precision Irrigation Advisory, and Expert Consultation Services,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 222–229, May 2025.

[50] A. Yadav and K. Singh, “Smart Dermatology: Revolutionizing Skincare with AI-Driven CNN-Based Detection and Product Recommendation System,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 230–235, May 2025.

[51] C. Yin, Y. Zhu, and S. Fei, “A deep learning approach for intrusion detection using LSTM networks,” *IEEE Access*, vol. 10, pp. 20454–20465, 2023.

[52] M. Ghasemi and Z. Zhou, “Hybrid CNN-LSTM architectures for IoT anomaly detection,” *Future Generation Computer Systems*, vol. 143, pp. 412–425, 2024.

[53] P. Wang, R. Lin, and Y. Deng, “Centralized vs. decentralized AI in IoT security: A performance comparison,” *IEEE Systems Journal*, vol. 18, no. 3, pp. 5021–5033, 2024.

[54] E. Akhtar and K. Lee, “Autonomous learning limits in static ML-based IDS models,” *Computers & Security*, vol. 136, pp. 103–112, 2024.

[55] K. Singh and P. Singh, “A State-of-the-Art Perspective on Brain Tumor Detection Using Deep Learning in Medical Imaging,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 3, pp. 250–254, Jun. 2025.

[56] K. Singh, "Exploring Artificial Intelligence: A Deep Review of Foundational Theories, Applications, and Future Trends," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 6, pp. 295–305, Sep. 2025.

[57] D. Silver, J. Schrittwieser, and K. Simonyan, “Multi-agent reinforcement learning for complex decision environments,” *Nature*, vol. 620, pp. 304–315, 2023.

[58] M. Gronauer and K. Diepold, “Multi-agent deep reinforcement learning: A survey,” *Artificial Intelligence Review*, vol. 56, pp. 2923–2977, 2023.

[59] Z. Hu, Y. Lin, and Q. Liu, “Decentralized reinforcement learning for IoT threat mitigation,” *IEEE Transactions on Network and Service Management*, vol. 20, no. 1, pp. 56–70, 2024.

[60] Q. Zhang, J. Li, and P. Lin, “Hierarchical agent cooperation for adaptive IoT security,” *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8901–8914, 2025.

[61] M. Du and N. Li, “Limitations of generalization in reinforcement-based network defense,” *Expert Systems with Applications*, vol. 233, pp. 119–132, 2025.

[62] T. Luo and C. Huang, “Scalable coordination in autonomous multi-agent defense systems,” *IEEE Access*, vol. 13, pp. 28011–28023, 2025.

[63] R. George and H. Chen, “Collaborative learning architectures for cyber defense agents,” *Ad Hoc Networks*, vol. 159, pp. 102–135, 2025.

[64] P. Bhattacharya, A. Sharma, and M. Jain, “Next-generation autonomous agents for IoT security: Opportunities and challenges,” *IEEE Transactions on Dependable and Secure Computing*, vol. 22, no. 6, pp. 1511–1526, 2025.

[65] S. Kumar and D. Patel, “Edge-based Cyber Defense Mechanisms in IoT: A Review,” *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2456–2468, 2023.

[66] N. Zhang, H. Li, and F. Wang, “Federated Reinforcement Learning for IoT Security Enhancement,” *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 2, pp. 1140–1152, 2022.

[67] P. Roy, M. Saha, and K. Rahman, “Blockchain-Enabled Communication Framework for Secure Multi-Agent Systems,” *Future Generation Computer Systems*, vol. 141, pp. 435–447, 2023.

[68] A. Rahman and J. Lin, “Autonomous Network Defense Using Adaptive AI Agents in IoT,” *Computers & Security*, vol. 127, pp. 103–120, 2024.

[69] L. Wu, X. Chen, and R. Zhu, “Reinforcement Learning for Self-Adaptive Intrusion Detection in Distributed IoT Networks,” *IEEE Access*, vol. 12, pp. 54867–54879, 2024.