Vol. 9, Issue: 2025

Article

ISSN: 2312-2668

International Journal of Information Sciences Management (IJIMS)

Security Framework in IoT-based Systems: Integrating Blockchain, Deep Learning, Reinforcement Learning, and Game Theory

Maryam Fattahi

Security Framework in IoT-based Systems: Integrating Blockchain, Deep Learning, Reinforcement Learning, and Game Theory

Maryam Fattahi
Amirkabir University of Technology, Tehran Iran
m.fattahi@aut.ac.ir

Abstract:

This study examines the relationship between Social Responsibility, Knowledge Management, and sustainable development in Hayel Saeed Ana'am Group in Yemen. The research aims to explore how social responsibility influences knowledge management processes and contributes to achieving the Sustainable Development Goals. Data were collected using a structured questionnaire administered to 191 participants, and the instrument's reliability and validity were confirmed. The data were analyzed using SPSS 28 and SmartPLS 4.0.5.9. The results reveal that Social Responsibility significantly influences both Knowledge Management and Sustainable Development. Furthermore, Knowledge Management acts as a mediator between Social Responsibility and Sustainable Development. The findings underscore the importance of integrating social responsibility into organizational strategies to enhance knowledge management practices, which in turn drive sustainable development outcomes. The study contributes to the growing body of knowledge on the interconnectedness of these dimensions, providing valuable insights for organizations seeking to improve their sustainability practices through responsible management and knowledge processes.

Keywords: Social Responsibility; Knowledge Management; Sustainable Development.

Received:
July 24, 2025
Review Process:
August 01, 2025
Accepted:
August 28, 2025
Available Online:
September 25, 2025

Introduction

IoT-based Systems (IoTS) have emerged as a fundamental component for essential infrastructure across multiple sectors, such as industrial control systems, healthcare, transportation, and intelligent grid networks. These systems amalgamate physical processes with computational and communication assets, facilitating real-time oversight, regulation, and enhancement (Mohammed et al., 2024). The intrinsic interconnection of IoTs makes them more prone to all sorts of cyber threats, including false data injection attacks, denial-of-service attacks, and jamming attacks, which can compromise their security, reliability, and operational performance (Dai et al., 2024; Balogun et al., 2024).

For instance, medical cyber-physical systems used in the medical domain have high vulnerability to cyber attacks, which could steal sensitive patient information or interrupt critical life-sustaining services (Mohammed et al., 2024; Balogun et al., 2024). Equally, the smart grid infrastructures are exposed to a connected series of cyber attacks that are able to compromise the grid operations and cause mass-scale power failures (Farraj et al., 2024). The critical role CPS play in society underlines the necessity for their security and resilience in view of emerging threats.

Traditional security mechanisms often don't bring out the complexity and dynamic nature of CPS, which needs innovative approaches. In this context, emerging technologies in blockchain, deep learning, and game theory hold very promising solutions for CPS security.

The decentralized and immutable blockchain structure provides a guarantee for data integrity, setting up a robust foundation for secure communications (Dai et al., 2024; Mohammed et al., 2024). Deep learning techniques enable real-time anomaly detection through complex patterns in system behavior (Selim et al., 2024; Costa et al., 2024). Besides, the game theory brings in strategic insights in adversarial interactions to improve the defense against advanced attackers (Ge et al., 2024; Wang et al., 2024).

Several studies have proposed new methods to face the security issues in CPS. Mohammed et al. (2024) pointed out that the integration of deep learning and blockchain may be an effective method to secure healthcare data in industrial CPS. Dai et al. (2024) reported a blockchain framework to improve cyber-resilience against FDI attacks of microgrid distributed control systems. Similarly, Balogun et al. (2024) proposed a deep learning approach embedded with blockchain to enhance the security of next-generation medical IoTs.

The area of CPS security has had wide applications of game theory. Zhang et al. (2024) analyzed a game-theoretic approach for enhancing constraint-following control in CPS under cyber threats. Farraj et al. (2024) analyzed an attack-mitigation strategy of switching attacks in smart grid systems by a game-theoretical model. Moreover, Wang et al. (2024) presented the application of a Stackelberg game-theoretical model for optimal attack strategy for CPS.

Deep learning, especially in the realm of deep reinforcement learning, has surfaced as an effective instrument for adaptive control and the mitigation of cyber threats. Selim et al. (2024) conducted an exploration of deep reinforcement learning aimed at protecting distribution systems from cyber attacks. Costa et al. (2024) employed reinforcement learning to improve control efficacy within adversarial contexts, demonstrating its adaptability and resilience.

Based on the literature review, this study takes a step forward to propose a holistic security framework for IoTs that integrates blockchain, deep learning, and game-theoretic strategies. It contributes in the following ways:

- Blockchain-Based Resilient Architecture
- Attack detection by applying Deep learning & Reinforcement learning approach
- Game-Theoretic & Reinforcement learning technique for Attack Mitigation

1. Proposed Framework

The proposed system presents a strong framework for IoTs defending against cyber-attacks based Figure 1. It fuses blockchain technology, deep learning, reinforcement learning, and game theory into a multi-layer structure.

Each stage's output becomes the input for the next:

1. Layer 1 Output (secure data) → Layer 2 Input (anomaly detection).
2. Layer 2 Output (anomalies) → Layer 3 Input (reinforcement learning for action selection).
3. Layer 3 Output (actions) → Layer 4 Input (game-theoretic allocation for resource optimization).

Final Workflow (Incorporating Logical Flow):

1. Input: Sensor data, system states.
2. Blockchain: Validate and secure data.
3. Anomaly Detection: Identify attack patterns using deep learning.
4. Reinforcement Learning: Optimize defensive responses.
5. Game Theory: Allocate resources efficiently.

This hierarchical flow ensures all components work seamlessly, addressing both operational and security challenges in IoTS environments.

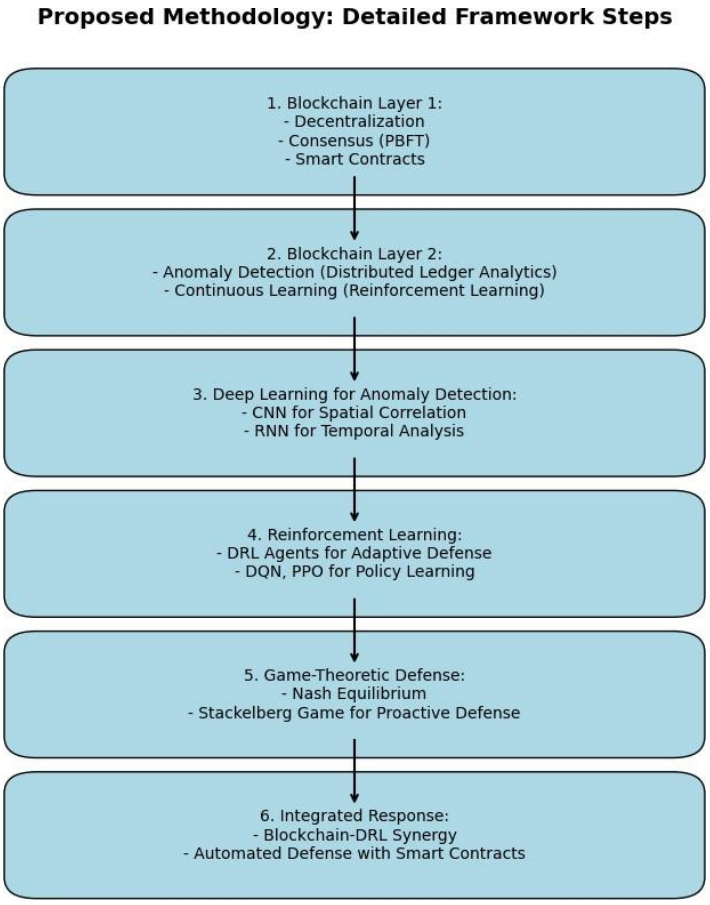


Figure 1. *Proposed Methodology: Framework Steps*

2. Blockchain-Based Attack-Tolerant Architecture

The security and resilience of the IoTs can surely help provide seamless operational capability in significant domains such as health care, smart grids, and industrial automation. Given the opportunity, we introduce a robust two-layer blockchain architecture for the protection of the IoTS networks against sophisticated cyber attacks. A fault-tolerant communication framework forms Layer 1, while advanced defense mechanisms comprise Layer 2.

Blockchain Layer 1: Fault-Tolerant Communication Architecture

This layer aims at enabling secure, decentralized, and tamper-proof communication among IoTS networks. By leveraging the inherent features of blockchain, Layer 1 ensures integrity, authenticity, and non-repudiation of the inter-device communications.

1. Decentralization and Distributed Ledger:

Blockchain provides security based on a distributed ledger. The collapse of one or more nodes cannot shut down the whole system due to the consensus process in which every node of the IoTs network takes part; hence, it can provide the required operation and reliability. As shown by Mohammed et al. (2024) and Dai et al. (2024), this decentralization enhances fault tolerance, hence network reliability under attack conditions.

2. Immutability and Data Integrity:

Each transaction or message placed onto the blockchain is encrypted into the blockchain

through a process called cryptographic hashing; meaning, once it is written, it cannot be changed.

Input: System data and network status.

Objective: Securely and reliably transmit data across IoTS nodes.

A transaction T_i is hashed using a secure hash function (Control Law):

$$H(T_i) = SHA256(T_i)$$

where $H(T_i)$ is the hashed value. Any tampering attempts would result in a hash mismatch, thereby making such type of attack easily detectable.

3. Consensus Mechanisms:

Agreement among nodes is achieved through consensus algorithm, like PBFT. PBFT offers low latency and high throughput and is, hence, ideal for real-time IoTS applications. This consensus ensures that invalid transactions are rejected and any valid transaction is added to the blockchain. Consensus Condition: For a transaction to be validated:

$$\sum_{i=1}^{N} 1(V_i = True) \geq Q$$

Where:

- N = number of nodes.
- V_i = verification of i^{th} node.
- Q = quorum threshold for consensus.

Output: Secure, tamper-proof communication for anomaly detection.

4. Smart Contracts:

Smart contracts automate the enforcement of security policies. Such a contract can specify rules over many aspects like access control, authentication, and automated threat responses. A simple rule may take the form:

$$if\ (Anomaly\ Detected) \Rightarrow Trigger\ Response\ Action$$

5. Authentication and Access Control:

Blockchain's public-key cryptography secures the authentication of devices and users. Each device has a unique identification via a key-pair (K_{public}, $K_{private}$) where:

$$Message\ Signature = Encrypt(Message, K_{private})$$

Verification via the public key K_{public} ensures that only messages via an authorized device are passed to and from it.

Layer 1 Implementation Steps

1. Blockchain Network Setup: Design a private or consortium blockchain that best suits the IoTS performance needs, balancing transaction speed, security, and scalability.

2. Node Configuration: Blockchain nodes will be deployed on sensors, actuators, and controllers, which are IoTS devices, considering computational compatibility.

3. Consensus Algorithm Selection: The PBFT protocol provides a low-latency, high-throughput environment; otherwise, other consensus protocols might be used depending on the application scenario.

4. Smart Contract Development: Smart contracts that define access policies and interaction protocols shall be developed and deployed to.

5. Infrastructure Integration: Integrating the blockchain layer with IoTS infrastructure to ensure smooth working without much overhead.

Blockchain Layer 2: Advanced Defense Mechanism

On top of the secure communication provided by Layer 1, the second layer incorporates mechanisms that detect and respond to sophisticated attacks. It leverages distributed ledger analytics, machine learning, and reinforcement learning to ensure IoTS stability in dynamic threat environments.

Input: Secure transaction data from Layer 1.

Objective: Detect anomalies and trigger real-time responses.

1. Distributed Ledger Analytics for Anomaly

Detection:

Transaction patterns and data recorded in the blockchain are always observed for any anomaly, which might indicate an attack. As an example, this might be an alert if there is a sudden surge of traffic from a particular node. The metrics used by the anomaly detection model include:

$$Ai = \frac{Observed\ Pattern - Expected\ Pattern}{Expected\ Pattern}$$

Where $A_i > Threshold$, showing anomaly detection.

2. Forensic Analysis and Auditing:

The immutable nature of the blockchain records enables forensic investigation through tracing the sources of an attack, analyzing compromised nodes. This feature is beneficial to know the lifecycle of attacks and hence understand how to mitigate them in the future.

3. Continuous Learning and Adaptation:

Reinforcement learning models are integrated to dynamically adjust the defense strategies. These methods utilize a reward function, which is maximized once an attack has been successfully mitigated. The reward has been summarized as follows :

$$R = Mitigation\ Success - False\ Positive\ Penalty - Resource\ Utilization\ Cost$$

Over time, the model continues to modify the strategies to ensure excellent system performance.

4. Real-Time Response and Automation:

The smart contract will automatically trigger the defense mechanisms, which may be a node isolation or resource reallocation, depending on the detection outcome. For example:

$$if\ A_i > Threshold\ Trigger\ Isolation\ Protocol$$

Layer 2 Implementation Steps

1. Anomaly Detection Model Training: Train the models on historical and simulated IoTS attack data to learn patterns related to attacks.
2. Integration with Reinforcement Learning: Deploy RL agents that learn with new attack vectors and continuously optimize resource allocations.
3. Real-Time Deployment: Integrate the detection models with blockchain analytics for real-time monitoring and response.
4. Smart Contract-Based Automation of Defense: Build smart contracts that will automatically execute some predefined defensive action upon the detection of any anomaly.

The proposed blockchain-based architecture integrates these two layers to provide a highly resilient, real-time anomaly detection, and efficient response mechanisms to safeguard IoTs networks against complex cyber-attacks.

4. Deep Learning and Reinforcement Learning-Based Attack Detection

The proposed system involves the second layer for the detection and mitigation of attacks in IoTs environments through utilizing advanced deep learning and reinforcement learning techniques. This layer copes with real-time anomaly detection and adaptive mechanisms in order to ensure continued security with operational efficiency for IoTs.

Deep Learning Models

Deep learning is the backbone of this anomaly detection system. It utilizes supervised learning algorithms to build models showcasing, from IoTs data streams, deviations from normal patterns of operation.

Key Components of Deep Learning Models:

1. Data Preprocessing and Feature Extraction:

-Noise Removal: Most sensor data, like temperature, pressure, and voltage, tends to be noisy. Cleaning such datasets involves smoothing filters or noise-reduction algorithms to remove inconsistencies.

- Feature Extraction: The process of extracting important features using Fourier Transform

for frequency analysis or Principal Component Analysis (PCA) for reducing the dimensions and retaining the important information.

The Fourier Transform decomposes the time-series data into frequency components, for example:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j 2\pi f t} dt$$

where *X(f)* represents the frequency domain and *x(t)* is the time-domain signal.

2. Anomaly Detection Models:

- CNNs: CNNs are able to detect spatial correlations within data such as sensor readings in smart grids. The convolution operation extracts feature in the form of spikes or anomalies:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

where f is the input signal and g is the kernel function.

- RNNs: RNNs, in particular Long Short-Term Memory (LSTM) networks record temporal dependencies in sequences. LSTMs handle long-term dependencies by means of gating mechanisms:

$$h_t = o_t, \tanh(c_t)$$

where h_t is the hidden state, o_t is the output gate, and c_t) is the cell state.

Output: Alert and features of anomalies, which serve as input to reinforcement learning.

3. Real-Time Operativity:

Models analyze streams of data in real time; deviations from learned patterns flag potential cyber- attacks. For instance, during smart grid operation, if a sensor reports unusual voltage spikes, then such a signal is flagged for anomaly by the system. The response mechanisms are triggered, including the isolation of compromised components and/or alerting of the administrators.

4. Scalability:

Deep learning models are designed to scale; therefore, IoTS environments can handle increased volumes of data. Parallel processing and deployment on edge devices reduce latency and hence provide real-time detection efficiently as indicated by Costa et al. (2024).

Deep reinforcement learning

Deep reinforcement learning (DRL) enhances the adaptability of the system by learning about optimal defense policies in dynamic, rapidly changing threat landscapes. The DRL agent interacts with IoTs to learn a best policy through the feedback.

Input: Features of detected anomalies and system states.

Objective: Optimize defense strategies dynamically using reinforcement learning.

Key Components of DRL

1. Learning Environment:

- Abstracting IoTs environment into states, actions, and rewards:
  - States (S) : Represent system status - sensor readings or network health.

- Actions: Isolate nodes, reallocate bandwidth, or reconfigure network paths.

- Rewards: Represent the action's impact on preventing ransomware-quantify to describe the successful mitigation or wasting of resources.

In that respect, the agent aims at maximizing the cumulative reward:

$$Gt = \sum_{k=0}^{\infty} R_{t+k+1}$$

where γ is the discount factor that strikes a balance between immediate rewards and long-term rewards.

2. Agent Architecture:

- DQN: Use neural networks to approximate the Q-value function and make decisions to optimize discrete actions:

$$Q(s,a) = Q(s,a) + \alpha [R + \max Q(s', a') - Q(s, a)]$$

where *Q (s, a)* is the expected reward for taking action a in state s.

- Proximal Policy Optimization (PPO): This efficiently balances exploration and exploitation for continuous action spaces to ensure stable learning.

3. Adaptation to Evolving Threats:

DRL models are constantly updated with new policies to keep up with emerging data and attack patterns. This makes the system resistant to various new, complex attacks.

4. Integration with Blockchain:

- DRL agents make use of blockchain for maintaining immutable records of the events occurring on the system. This helps in training on reliable datasets and creating tamper-proof logs for forensic analysis.

The proposed system takes in deep learning's pattern recognition capabilities, combines them with the adaptability of DRL, thereby ensuring attack detection and mitigation in IoTS environments that is both robust and scalable. This integrated approach addresses current challenges while preparing for future, more sophisticated cyber threats.

4. Game-Theoretic Defense Strategies

Game theory provides a mathematical approach toward modeling the interactions between attackers and defenders in IoTs. In the context of the proposed framework, this approach allows. it to further forecast, analyze, and counteract cyber-attacks. The goal of the defense strategies is to optimize resource allocation and response mechanisms w.r.t the IoTS resiliency.

Input: System state, anomaly features, and defensive actions from reinforcement learning.

Objective: Allocate resources optimally for attack mitigation.

Game Theory-Based Defense

Game theory models the conflict between attackers and defenders as strategic games in which both adversaries try to maximize respective utilities, namely the success of an attack or the efficiency of a defense.

Nash Equilibrium

- Concept: Nash equilibrium is the stable state where none of the players, either the attacker or the defender, can improve his utility by changing his strategy unilaterally.

The utility functions for attackers and defenders are:

$U_A(S_A, S_D) = Success\ Rate\ of\ Attack - Cost\ of\ Attack$

$U_D(S_A, S_D) = Effectiveness\ of\ Defense - Cost\ of\ Defense$

where S_A and S_D are the strategies of attackers and defenders.

- Application: - It models the scenarios in which the attackers choose targets, such as nodes or sensors, and defenders allocate resources like computation or bandwidth to protect the IoTS.

- The equilibrium helps find an optimum between the attack efforts and defense measures for the most efficient use of resources (Zhang et al., 2024).

- Example: In a smart grid, a defender defends some vital nodes, while the attacker tries to disturb the distribution of energy. Under Nash equilibrium, the defender's resource allocation would minimize the success of the attack by striking a balance between system cost and efficiency.

Stackelberg Game

- Definition: A Stackelberg game is a hierarchical game-theoretic model where one player's (leader's) strategy commitment is made first, followed by the other player's strategy to act accordingly.

The defender's optimization problem is:

$$\max_{S_D} \min_{S_A} U_D(S_A, S_D)$$

where the defender anticipates the attacker's best response. Output: Optimized resource allocation strategy.

- Application: - The defender acts as the leader, anticipating potential attack strategies and committing resources preemptively.
  - The attacker, as the follower, adapts its strategy based on the defender's commitment.
  - This proactive approach enables the defender to mitigate threats before they escalate (Farraj et al., 2024).
- Example: During the event of a attack, the defender predicts that an attacker will commit resources to flood network channels. Pre/delaying bandwidth for operationally critical system use would mean that the system is continued without interruption.

Optimization Goals:

- Minimizing the impact of successful attacks
- Maximizing the value of the defensive resources assigned
- Minimizing costs involved in deploying protective countermeasures

Resilient Nash Equilibrium Seeking

The framework uses resilient Nash equilibrium models in order to enhance IoTS defenses under attack conditions, taking adversarial behavior and environmental uncertainties into account.

Advanced Game-Theoretic Models

- These models extend the traditional Nash equilibrium by incorporating aspects such as:
  - Uncertainty Modeling: Incomplete information about the attacker's strategy or unexpected system failures.
  - Dynamic Adjustments: Continuous updates of defensive strategies whenever new threats emerge or changes occur in system conditions.
- Risk Management: Strike a balance among security, efficiency in operation, and limitation of resources (Cai et al., 2024).

False Data Injection (FDI) Defense

- Scenario:
  - In attacking agents inject false data into sensors to mislead decision-making processes.
  - Resilient Nash equilibrium models help defenders allocate resources and monitor and secure vulnerable sensors.
- It is via game-theoretic analysis that the defender identifies the critical sensors based on the impact each sensor has on system stability and deploys her defense resources in that form.
- Adaptive mechanisms redeploy resources dynamically when attackers change their focus on other parts of the system.

Multi-Agent Systems

- Application:
- Agents in multi-agent IoTS are both defenders and functionals of the system, for instance, interconnected smart grids.
  - Resilient Nash equilibrium ensures collaboration between agents on resource sharing to defend the whole system.

The approach combines game-theoretic models, such as Nash equilibrium and Stackelberg games, with proactive dynamic methods for defending against sophisticated cyberattacks on IoTS. The embedding of resilient Nash equilibrium makes the system dependable in unforeseen situations, therefore highly adaptive and secure.

5. Enhanced Security by combination of methods

The framework goes further to strengthen the security in IoTS with the integration of Blockchain Technology and Deep Reinforcement Learning (DRL). In this hybrid approach, the immutability and decentralization from blockchain merge with the predictive and adaptive capabilities of DRL, thereby delivering a scalable, robust, and adaptive security solution in dynamic and high-risk environments.

Integration of Blockchain into Security

Blockchain provides IoTS with a secure platform for storing and managing data, ensuring that IoTS data, including all system

activities, sensor readings, and control commands, are recorded in nature with an immutable record due to its decentralized and tamper-proof features. This, in turn, enhances the reliability of the data used for training DRL models and hence ensures integrity in decision-making processes when an attack scenario is involved.

- Immutability: sensor data and system logs are kept within the blockchain ledger. Attackers will not be able to tamper with any historical data, thus allowing the proper training of DRL models and the effective execution of the defensive strategy. Dai et al. (2024).

- Decentralization: Blockchain retains data across multiple nodes. By doing this, single points of failure do not exist anymore. This makes critical information accessible during attacks.

- Transparent and Verifiable Data: All data in the blockchain are audited and may become verifiable by rightful parties, which will instill more confidence in inbuilt IoTS defense mechanisms.

Deep Reinforcement Learning for Adaptive Defense

The DRL models interface with the blockchain layer to offer adaptive and predictive defense mechanisms. The DRL agents interact with the environment of the IoTs to learn an optimal policy for providing maximum resilience for the system and reducing the cyberattack impact.

Agents Training

- Modelling Environment: IoTs comprising sensors, actuators, and communication networks is modelled to present an environment where DRL agents can interact. States S, actions A, and rewards R define the learning environment.

- Reward System:

- Positive rewards given for the mitigation of attacks, system stability, and resource utilization.

- Negative rewards introduced for delaying response, resource wastage, or other such actions that result in the failure to prevent the disruption of a system.

$$R_t = \begin{cases} +1, & \text{if attack mitigated} \\ -1, & \text{if attack successful} \end{cases}$$

Algorithms and Architectures

- Deep Q-Networks: Discrete action spaces are used for node isolation and communication channel switching.

- Proximal Policy Optimization: Continuous action spaces like dynamic adjustment of control parameters during the attack flow are handled by PPO.

- Actor-Critic Methods: Integrates value-based methods with policy-based approaches for faster convergence and more robust decisions.

Blockchain-DRL Synergy

This would be the leverage of the unique strengths of both blockchain and DRL in enhancement of the general security and adaptability of the system.

- Data Integrity for Training:

- Blockchain creates one source of verified historical and real-time data on which to train DRL models.

- Ensuring that DRL agents are trained based on sound data would lead to much more effective and secure defense.

- Decentralized Decision-Making:

-Agents deployed at various nodes use DRL and, therefore, can act independently, while their coordination through blockchain scales up the defenses uniformly within the network.

- Smart Contracts for Automated Defense

- Blockchain smart contracts apply automated defensive measures suggested by the DRL agents. Once a DRL agent detects an attack, for instance, a smart contract may automatically change the setting of the network or notify the operators.

- Adaptive Policies:

- DRL updates its policies continuously from the new attack data recorded in the blockchain to keep the system updated against emerging threats.

Workflow of the Implementation

1. Data Collection and Storage:
   - IoTs stream data to the blockchain ledger for secure, tamper-proof storage.
2. Model Training:

• DRL models learn from historical and real-time data of the blockchain to develop optimal solutions with regard to attack detection and response.

3. Deployment:

• Trained DRL agents are deployed across the IoTs network. The correctness of final decisions is checked and executed through smart contracts on the blockchain.

4. Continual Learning:

- DRL agents' policies are updated with the newest information on the blockchain as new attack scenarios are encountered to keep it adaptive.

Advantages of Blockchain-Based DRL

- Scalability: The decentralized nature of blockchain and distributed operation of the DRL agents ensure that the system can be easily scaled up for large IoTs environments.
- Robustness: Immutability of blockchain enhances the reliability of DRL model training and decision-making.
- Real-Time Defense: The DRL agent provides an adaptive and fast response against agile threats with reduced attack impact.
- Automation: Defensive actions are automatically executed by smart contracts that reduce human interaction.
- Coordination: Blockchain enables coordination between multiple DRL agents in a distributed IoTS network.

By integrating blockchain with DRL, IoTs will have a robust and adaptive defense mechanism. The blockchain layer assures data integrity and decentralization, while the DRL agents provide intelligent and scalable responses against cyber-attacks that are continuously evolving. This will ensure that IoTs remains secure, reliable, and operational against sophisticated attacks.

Results and Evaluation

The experiments have been carried out with the following settings:

Dataset: IoT Network Intrusion Dataset was preprocessed and normalized in order to meet the requirements for training and testing phases.

- Proposed Framework: The architecture incorporates blockchain, deep learning (CNN and LSTM model), reinforcement learning, and game-theoretic methods to ensure strong security.
- Attack Scenarios:

1. Normal operation-no attack.
2. FDI attack-system output.

To detect attacks and distinguish between normal and abnormal data, a threshold value (Threshold) of 1.5 was considered. This value was chosen in such a way that it could identify data that has undergone abnormal changes due to attacks. The identification criterion is that any data value whose absolute value is greater than the threshold is identified as abnormal data (attack). This threshold value was determined based on an initial examination of the data and the operating conditions of the system.

A blockchain network simulation is performed in IoTs. In this simulation, 10 nodes are defined, each representing a device. The communication between devices is modeled as secure transactions. To identify and distinguish normal data from abnormal data, the transaction data is divided into two categories: normal data with the value Secure and abnormal data representing attacks with the value Compromised.

Table 1 shows the structure of transactions. In this table, the Source and Destination columns represent the source and destination devices, Transaction_Data represents the transaction

status, and Attack_Flag represents whether the transaction is normal or abnormal. For example, the transaction between devices Device_2 and Device_3 is identified as an attack and the Attack_Flag value is 1.

Table 1

Simulated transaction data on the blockchain network

Source	Destination	Transaction_Data	Attack_Flag
Device_0	Device_1	Secure	0
Device_1	Device_2	Secure	0
Device_2	Device_3	Secure	0
Device_3	Device_4	Secure	0
Device_2	Device_3	Compromised	1

In this environment, the actions of the model DRLs are divided into two categories:

1. Action 0: No Action: No defensive action is taken.
2. Action 1: Defensive Action: Defensive action is taken.

In the simulation process, control actions are applied based on the reinforcement learning policy and the system state is updated randomly from the training data. This initial state allows the model to adopt an optimal policy to reduce risk based on the dynamics of real data and the detection of attacks.

To combine blockchain with DRL,a simulated model was used. In this model, the network transaction data is defined in the form of a blockchain, and each transaction has the properties Source, Destination, Transaction Data, and Attack Flag. The DRL model dynamically analyzes transactions and applies a defensive action with a DRL_Action value of 1 if an attack is detected (Attack_Flag value of 1). Otherwise, the DRL_Action value remains at 0.The following table2 shows an example of the updated blockchain data after applying the DRL model.

Table 2

Example of the updated blockchain data after applying the DRL

Source	Destination	Transaction Data	Attack Flag	DRL_Action
Device_0	Device_1	Secure	0	0
Device_1	Device_2	Secure	0	0
Device_2	Device_3	Secure	0	0
Device_3	Device_4	Secure	0	0
Device_2	Device_3	Compromised	1	1

In the above example, the transaction between Device_2 and Device_3 is detected as an attack (Attack_Flag=1) and the DRL model has triggered the associated defensive action.

(DRL_Action=1). This approach demonstrates the ability of the DRL model to detect and respond to cyberattacks in blockchain systems.

Evaluation Scenarios and Results

System Response under Normal and Attack Conditions

Figure 2 depicts the system behavior at three conditions: normal operation and FDI attack. From that the following can be obtained:

Stable and repetitive feature values when the system is under normal conditions.

Severe abnormalities during FDI attacks, which in turn is depicted by variation in the pattern of features.

The system is capable of detecting such variations and isolating them so that operations remain safe.

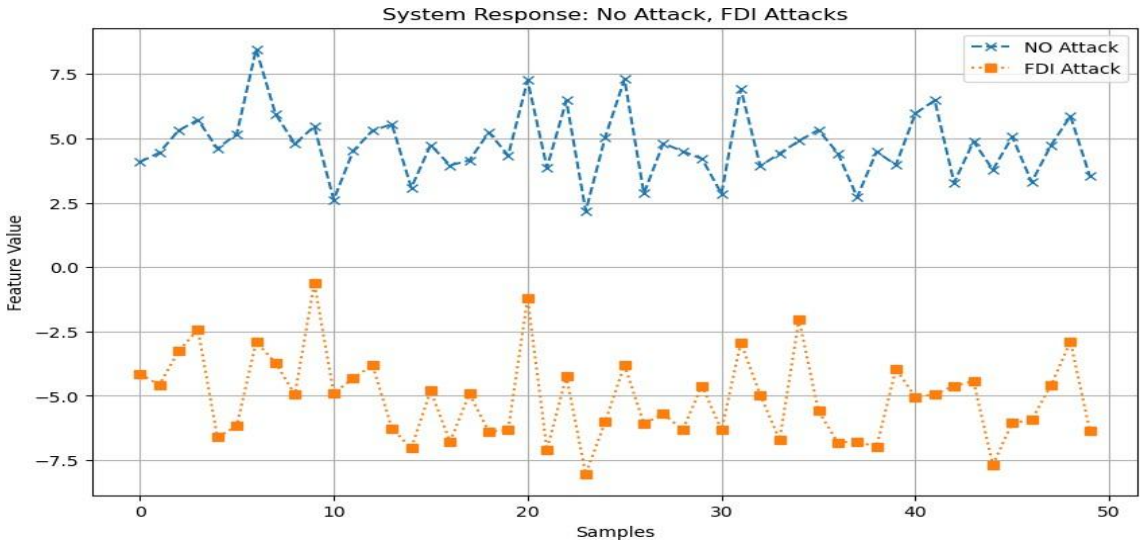


Figure 2. *System Response Under No Attack and Attack Scenarios*

System Behavior Under FDI Attack on Output

Figure 3 shows the system's response to a detected FDI attack. The detected attack produced manipulated output values well outside any normal behavior. The framework operated to clearly detect these anomalies and thus segregated the compromised outputs to ensure appropriate system responses.

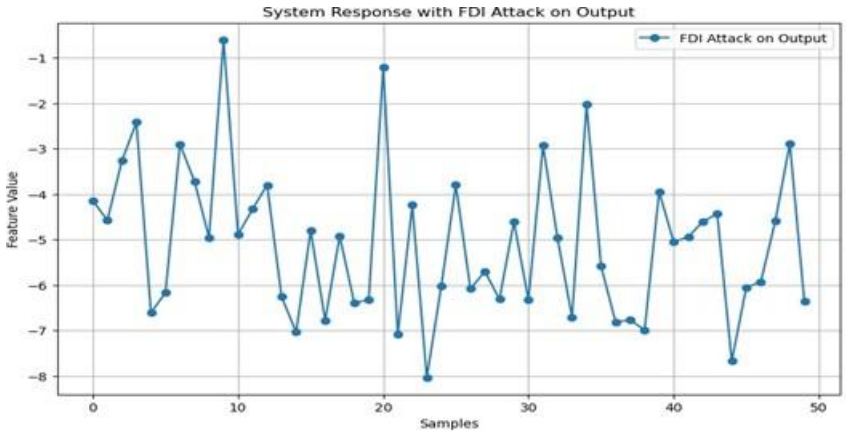


Figure 3. *System with FDI Attack on Output*

Detection of FDI Attacks

Figure 4 is detection of FDI attacks by the proposed detector. The system was able to maintain a high detection rate, flagging malicious modifications to the output data quite well. This quick detection ensured that defensive action was timely and hence prevented the system from further compromise.

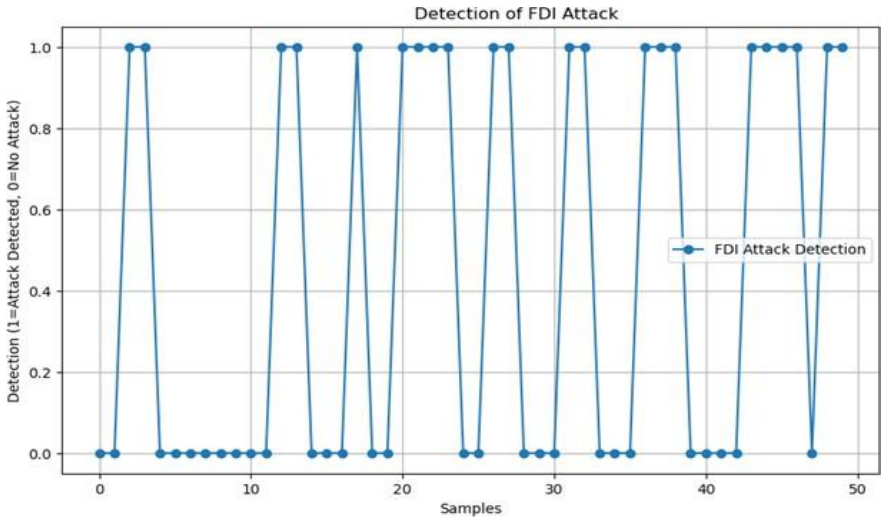
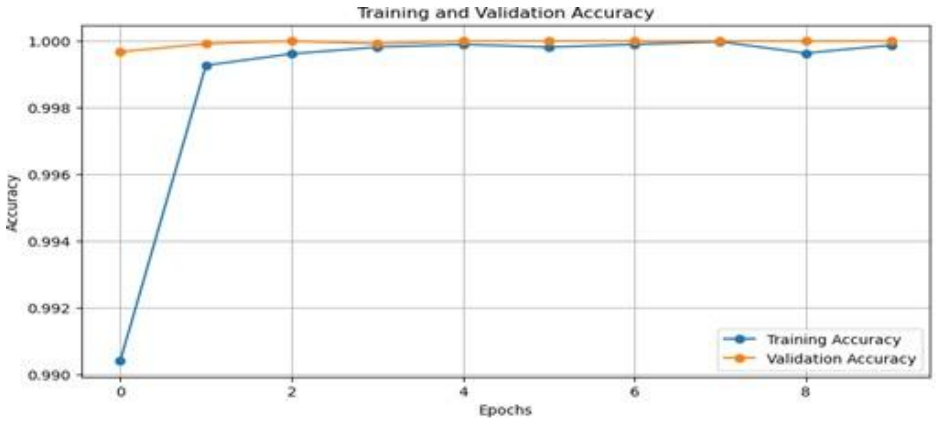


Figure 4. *Detection of FDI Attack*

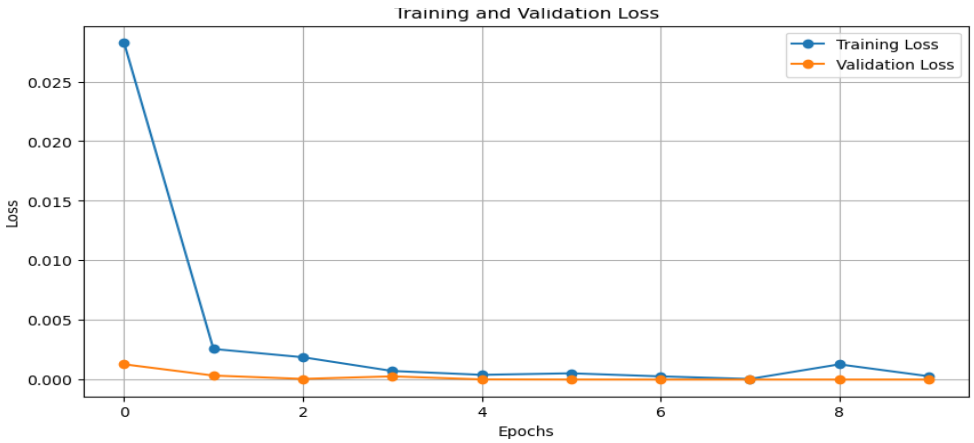
Evaluation Metrics Accuracy and Loss

The accuracy and loss metrics of training and testing are shown in Figures 4 and 5 Key observations include:

- High training and testing accuracy, indicating that the proposed system is very effective in identifying between normal and attack scenarios.
- Low loss values indicate that the proposed deep learning models are robust.



Figures 5. *Accuracy of training and testing*



Figures 6. *Loss metrics of training and testing*

Detection Rate

The proposed integrated framework was effective in detecting FDI attacks, as the detection rate of these attacks was high throughout. Indeed, both types of attacks were correctly detected while minimizing false positives and false negatives based in Figure 7.

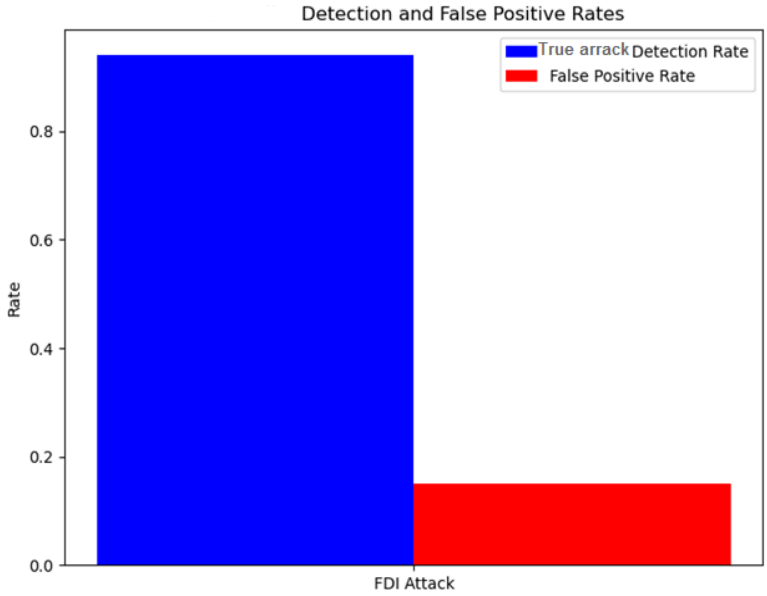


Figure 7. *Detection Rate*

Figure 8 shows the confusion matrix. True Negative (green), the number of instances in which the system correctly detected normal conditions. A relatively high value indicates the system's ability to detect normal conditions.

False Positive (red), the number of instances in which the system incorrectly detected an attack, when there was no attack. A high value indicates a relatively high false positive rate, which may cause false alarms. False Negative (orange), the number of

instances in which the system failed to detect an attack.

A low value indicates a good ability of the system to detect attacks. True Positive (blue), the number of instances in which the system correctly detected an attack. A low value may be due to the complexity of the attacks or the limitations of the detection model.

This graph shows that the system performed well in detecting attacks and normal conditions.

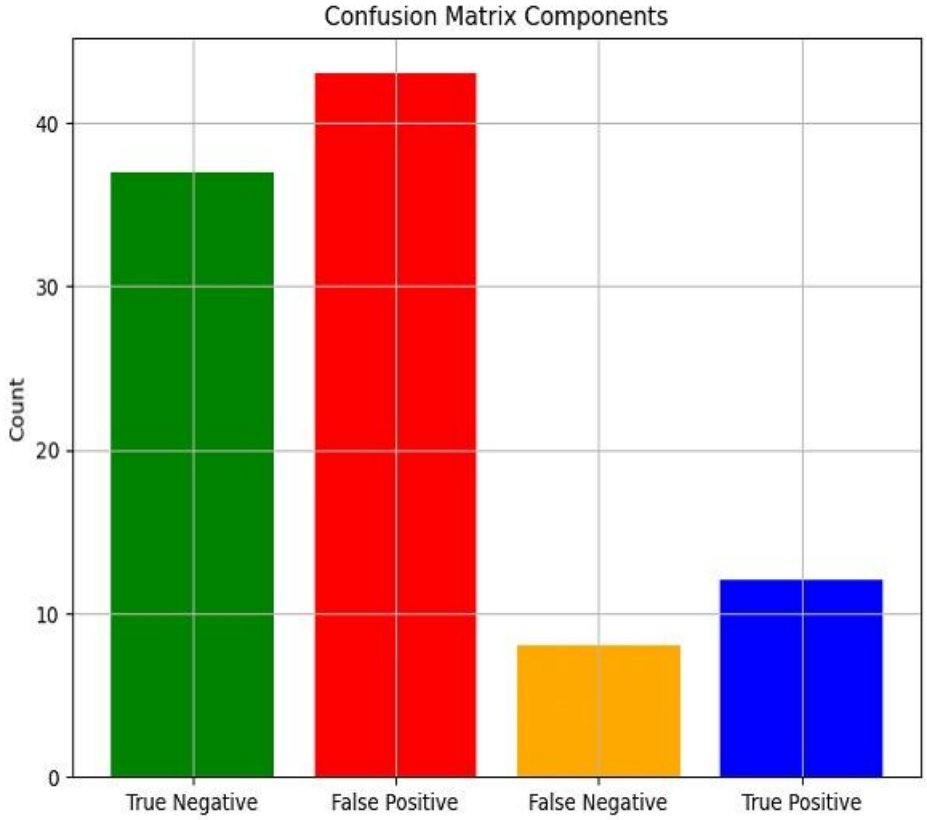


Figure 8. *Confusion matrix*

Conclusion

The developed methodology involves integrating blockchain technology, deep learning models, reinforcement learning, and game theory into one integrated, multilayer security architecture. The system operates in the following layers:

1. Blockchain Layer: This layer enables secure and tamper-proof communication channels and records in the IoTs. Due to the decentralization, the risks concerning single points of failure are eliminated, while using smart contracts automates responses to detected threats, minimizing the response time.
2. Deep Learning Layer: These include real-time anomaly detection using CNNs and LSTM networks. All these models are trained to identify normal system behaviors from those that signal an attack with unparalleled accuracy and very low false alarms.
3. Reinforcement Learning Layer: Adaptive decision-making is performed by DRL. The system can make changes to dynamically adapt to the evolving attack vectors, such as FDI attacks, and turn the defense mechanisms proactive.
4. Game-Theoretic Layer: Attack-defense interactions are modeled using Nash Equilibrium and Stackelberg game theory. It leads to optimal resource allocation and strategic decision making to efficiently mitigate the attacks.

The system's performance was evaluated using the IoT Network Intrusion Dataset simulating normal operations and attack scenarios:

1. Detection Rate: The detection rate for FDI attacks remained consistently high throughout the evaluation, confirming the system's effectiveness in identifying malicious activities. The framework successfully minimized false positives and false negatives, enhancing its reliability in real-world applications.
2. False Positive and Negative Rates: The code extensively measured these metrics, indicating a small rate of false positives and only a few missed detections. This allows the system to ensure that valid activities will not be classified as false, helping maximize the efficiency of operations.
3. Scalability: The modular architecture present in the framework has proved scalable, fully integrative with IoTS architectures, and extendable for system growth.
4. Efficiency: The system could detect and mitigate attacks in runtime with efficiency; thus, it is practically applicable to large-scale IoTS environments.

6. Future Work Direction

While the proposed framework has been effective, there are some points that require further exploration to enhance the capabilities and adapt to future challenges:

1. Advanced Threat Scenarios: Future work can also be directed to the extension of the current framework for complex scenarios including multivector attack, insider threat, and APTs.
2. Hybrid Models: Integration of hybrid models involving deep learning with graph-based approaches would help in better anomaly detection accuracy and scalability.
3. Optimization: Blockchain-based IoTs solutions may bring extra computation and energy overhead. In the future, energy-efficient blockchain consensus mechanisms could be studied, such as PoS or DAGs.

The proposed framework has already demonstrated practical feasibility and scalability through extensive evaluations. Its layered approach comprehensively covers security against all critical cyber attacks, while the integration of blockchain, deep learning, and game theory makes this framework highly robust and adaptive. Addressing the challenges outlined in future work, this

framework can indeed open ways to secure, resilient, and efficient IoTs operations in the ever- changing panorama of cyber threats.

7. Compliance with ethical standards

Author Contribution: All study conception and designs, material preparation, data collection and analysis of this work were performed by Author.

competing interest:

Conflict of interest: The author declares that there is no conflict of interests regarding the publication of this paper.

Human and /or Animals Research: This research doesn't not involve any humans or animals.

Data Availability: There is no additional data availability for this work.

Funding: The author did not receive support from any organization for the submitted work.

References

Mazin Abed Mohammed, Abdullah Lakhan, Dilovan Asaad Zebari, Mohd Khanapi Abd Ghani, Haydar Abdulameer Marhoon, Karrar Hameed Abdulkareem, Jan Nedoma, Radek Martinek. (2024). Securing healthcare data in industrial cyber-physical systems using combining deep learning and blockchain technology.

Jiahong Dai, Jiawei Yang, Yu Wang, Yan Xu. (2024). Blockchain-Enabled Cyber-Resilience Enhancement Framework of Microgrid Distributed Secondary Control Against False Data Injection Attacks.

Bukola Fatimah Balogun, Khushboo Tripathi, Shrikant Tiwari, Shyam Mohan J. S., Amit Kumar Tyagi. (2023). A blockchain-based deep learning approach for cyber security in next-generation medical cyber-physical systems.

Xinrong Zhang, Ye-Hwa Chen, Dongsheng Zhang, Ruiying Zhao, Lei Guo. (2023). A game-theoretic approach of cyberattack resilient constraint-following control for cyber–physical systems.

Bo Zhang, Chunxia Dou, Dong Yue, Ju H. Park, Zhanqiang Zhang. (2023). Attack-Defense Evolutionary Game Strategy for Uploading Channel in Consensus-Based Secondary Control of Islanded Microgrid Considering DoS Attack.

Xin Cai, Feng Xiao, Bo Wei. (2023). Resilient Nash Equilibrium Seeking in Multiagent Games Under False Data Injection Attacks.

Hui Ge, Lei Zhao, Dong Yue, Xiangpeng Xie, Linghai Xie, Sergey Gorbachev, Iakov Korovin, Yuan Ge. (2023). A game theory-based optimal allocation strategy for defense resources of smart grid under cyber-attack.Abdallah Farraj, Eman Hammad, Ashraf Al Daoud, Deepa Kundur. (2023). A Game-Theoretic Analysis of Cyber Switching Attacks and Mitigation in Smart Grid Systems.

Lei Xue, Bei Ma, Jian Liu, Yao Yu. (2023). Jamming attack against remote state estimation over multiple wireless channels: A reinforcement learning-based game theoretical approach.

Zhuping Wang, Haoyu Shen, Hao Zhang, Sheng Gao, Huaicheng Yan. (2023). Optimal DoS attack strategy for cyber- physical systems: A Stackelberg game-theoretical approach.

Shou-Zhou Li, Cheng-Wu Shao, Yan-Fu Li, Zhou Yang. (2023). Game Attack–Defense Graph Approach for Modeling and Analysis of Cyberattacks and Defenses in Local Metering System.

Guoquan Wu, Yujia Wang, Zhe Wu. (2023). Physics-informed machine learning in cyber-attack detection and resilient control of chemical processes.

Chengwei Wu, Wei Pan, Rick Staa, Jianxing Liu, Guanghui Sun, Ligang Wu. (2023). Deep reinforcement learning control approach to mitigating actuator attacks.

Bertinho A. Costa, Francisco L. Parente, João Belfo,

Nicola Somma, Paulo Rosa, José M. Igreja, Joris Belhadj, João

M. Lemos. (2023). A reinforcement learning approach for adaptive tracking control of a reusable rocket model in a landing scenario.

Mahmood Mazare. (2023). Reinforcement learning-based fixed-time resilient control of nonlinear cyber-physical systems under false data injection attacks and mismatch disturbances.