AIoT-Blockchain Security for Supply Chain Threats in IEC 61850 Substations Using Informer-Powered Reinforcement Learning

Lilia Tightiz, *Member, IEEE,*, L. Minh Dang and Ki-Woong Park, *Member, IEEE*

***Abstract*—IEC 61850 substations enable fast-speed digital communication among intelligent electronic devices (IEDs) for power system automatic control, monitoring, and protection. Their remote configurability and interoperability, however, make them vulnerable to highly advanced cyberattacks, mainly supply chain attacks. While existing methods, such as intrusion detection systems (IDS) and machine learning (ML)-based anomaly detection, provide partial protection, they often lack resilience against evolving attacks and real-time mitigation capabilities. We present an artificial intelligence of things (AIoT) blockchain security framework that uses Informer-augmented proximal policy optimization (PPO) for adaptive cyber defense, along with Hyperledger Fabric, for tamper-proof and automated security enforcement. The novelty of the proposed framework over state-of-the-art research lies in its combination of anomaly detection, dynamic threat mitigation, and auditable policy execution. Our security tests demonstrate robustness against zero-day and synthetic adversarial attacks, while preserving privacy and integrity. Experimental findings demonstrate that Informer-PPO attains 98.4% detection accuracy and 35 ms response time, representing improvements of 3.6%, 5.0%, and 9.1% in accuracy and 32.7%, 51.4%, and 63.2% faster response time compared to Transformer-PPO, long short-term memory (LSTM)-PPO, and convolutional neural network (CNN)-PPO baselines, respectively. Blockchain-enabled policy enforcement is accomplished within 42–50 ms, facilitating scalable real-time protection for IEC 61850 substations.**

***Index Terms*—IEC 61850, Supply Chain Security, Proximal Policy Optimization, Informer, Blockchain, Cybersecurity, Smart Grid.**

I. Introduction

THE digital communication-based automated and remote control processes changed the automation of substations, productivity, and real-time monitoring. The IEC 61850 pioneered such a global extent of open communication between substation devices. Interoperability and high-speed data exchange through IEC 61850 enable real-time tracking of a substation's IEDs [1]. Smooth communication is provided with the key protocols such as generic object-oriented substation event (GOOSE) messages, which enable event-driven data exchange with high rates; sampled value (SV) messages, which allow precise synchronization of measurement data; and manufacturing message specification (MMS) messages, which enable remote monitoring and control in supervisory control and data acquisition (SCADA) systems.

However, the digitalization and interoperability provided by IEC 61850 also introduce crucial security vulnerabilities. These involve message injection, spoofing, tampering, and unauthorized firmware updates that attackers can use to create significant disruptions [2]. The 2015 Ukraine power grid cyberattack is a well-known instance of how hacked substation communication systems can lead to large-scale blackouts and infrastructure collapses [3], [4]. In particular, supply chain vulnerabilities, i.e., malicious firmware updates and unauthorized IED configurations, have emerged as a key threat vector. Although traditional IDS and ML approaches offer pattern-based anomaly detection, they are often ineffective against zero-day attacks and novel threat vectors [5]. Reinforcement learning (RL), a subfield of ML, is particularly beneficial for adaptive cyber-physical systems, as it can learn optimal mitigation actions through interactions with the environment and without requiring any labeled data [23]. These characteristics position RL as a highly promising solution for real-time substation cybersecurity. However, RL approaches are confronted by computational inefficiencies as well as stability problems when applied in high-dimensional, time-critical applications such as IEC 61850 substations [7].

To overcome these limitations, researchers have investigated augmenting technologies like blockchain, which provides data integrity, secure device identity, and tamper-proof logs via decentralized consensus protocols [8]. While blockchain is inherently reactive to integrity enforcement, it is not inherently sensitive to rapidly evolving cyber threats and lacks real-time flexibility. Likewise, while RL is adaptive, it lacks the ability to capture temporal dependencies over long horizons in evolving attack patterns.

To this end, we propose a holistic security framework that integrates synergistically long-range anomaly detection through Informer-based time series modeling, adaptive threat mitigation through PPO-based reinforcement learning, and tamper-proof enforcement through blockchain smart contracts. The combined system fills the essential gap in existing substation cybersecurity solutions by providing real-time detection, learning-based response, and verifiable policy enforcement.

Besides, while this study primarily targets IEC 61850 substations, the AIoT-Blockchain framework developed herein readily applies to larger smart grid networks and other IoT-

L. Tightiz is with the Department of Computer Engineering, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006, Korea e-mail: liliatightiz@sejong.ac.kr.

L. M. Dang is with the Institute of Research and Development, Duy Tan University, Da Nang, 550000, Viet Nam, and Faculty of Information Technology, Duy Tan University, Da Nang, 550000, Viet Nam

K.W. Park is with Department of Computer and Information Security, Sejong University, Seoul 05006, email: woongbak@sejong.ac.kr

Corresponding author: K.W. Park

supported systems. Its modular characteristics of anomaly detection, reinforcement learning-based decision-making, and blockchain-based enforcement can be used to protect distributed energy resources, microgrids, electric vehicle charging systems, and smart city energy management systems, where dynamic cyber-physical threat landscapes and decentralized device authentication are similarly critical.

Based on the aforementioned motivations and design considerations, the key contributions of this paper are summarized as follows:

- **Informer-Powered PPO for Cyber Threat Mitigation**: We propose an Informer-powered PPO reinforcement learning approach that can learn long-range temporal dependencies to identify and prevent sophisticated cyber-attacks on IEC 61850 substations.
- **Blockchain-Enhanced Supply Chain Security**: We leverage blockchain technology for device authentication, firmware integrity, and tamper-proof security logging.
- **Smart Contract-based Autonomous Threat Response**: We utilize smart contracts for autonomous, policy-based enforcement of mitigation actions with little human involvement.
- **Performance Validation and Comparative Evaluation**: We evaluate the framework on Kitsune and ERENO datasets and benchmark it against CNN-PPO, LSTM-PPO, and Transformer-PPO baselines. The system achieves 98.4% and 97.6% detection accuracy with an average response latency of 35 ms, outperforming the baselines and recent state-of-the-art approaches, as detailed in Tables III and V.

The rest of this paper is organized as follows. Section II presents a summary of existing literature on IEC 61850 cybersecurity, focusing on existing ML-based, RL-based, and blockchain-based solutions and their shortcomings. Section III presents the envisioned AIoT-Blockchain security architecture, with a specific focus on informer-based anomaly detection, PPO-based adaptive threat mitigation, and blockchain-based enforcement through smart contracts. Section IV describes the experimental configuration for the implementation setup, the deployed datasets, the experimental evaluation, and the performance metrics. Section V presents the experimental results, scalability evaluations, security analyses, and considerations for real-world deployment. Section VI concludes the paper by highlighting main findings and directions for the future, including adversarial robustness and scalable deployment.

II. Literature Review

Cybersecurity research for IEC 61850 substations has evolved in a number of domains, including IDS, ML, RL, and blockchain-based authentication. While significant progress has been achieved, most existing solutions tackle individual components of the security stack and do not offer an end-to-end solution that ensures long-term anomaly detection, adaptive mitigation, and tamper-proof enforcement. This section reviews the relevant contributions and identifies their strengths and limitations in the context of IEC 61850 substations.

Duman et al. [11], [12] studied supply chain security in IEC 61850 substations using attack graphs and security posture metrics. In their 2019 work [11], the authors simulated attack scenarios to count risks, while their 2024 follow-up [12] proposed cost-effective control actions. Although these research studies are useful for attack surface understanding and defense prioritization, they do not impose adaptive detection or mitigation methods in real-time.

Ustun et al. [13] proposed an ML-based IDS that identifies anomalies in GOOSE messages. Their supervised system performs adequately to identify attacks that are known with labeled data. However, the heavy dependence on labeled datasets makes such systems ineffective against zero-day attacks or emerging intrusions. In a similar direction, Nhung-Nguyen et al. [14] designed a deep neural network-based IDS for GOOSE traffic, demonstrating high detection accuracy under controlled conditions. Yet, like other supervised systems, it lacked mitigation mechanisms and provided no latency evaluation, leaving questions about its applicability in real-time substation operations. Lian et al. [15] further emphasized these shortcomings of conventional IDS architectures, concluding that most will fail to generalize in rapidly evolving substation environments.

Blockchain-based authentication frameworks have also been investigated. Ghosh et al. [16] put forward a blockchain-based decentralized authentication framework for smart grids, highlighting tamper-proof device identity. Gayo et al. [17] expanded on that by integrating blockchain with IEC 61850 communication for secure microgrid hardware. While the two systems ensure data integrity and auditability, they do not employ adaptive learning to react dynamically to threats. In parallel, Park et al. [18] developed a machine learning-based anomaly detection system for SV and GOOSE traffic that further attempted to classify whether anomalies stemmed from malicious activity or benign faults. This partial restoration capability represents a step toward adaptivity, but their study still lacked blockchain integration and did not report latency, limiting its applicability in real-time critical infrastructure. Beyond cybersecurity, blockchain technology and AIoT have also shown promise in other areas of supply chain management. Lakhan et al. [19], for example, proposed a blockchain-enabled AIoT framework for improving sustainable supply chain systems through improved inventory transparency, logistics optimization, reducing operating expenses, and reducing carbon footprints. All these developments reflect the widespread applicability of AIoT and blockchain beyond security. However, specific to AIoT-enabled critical infrastructure, particularly real-time control systems such as IEC 61850 substations, cybersecurity challenges remain comparatively under-explored. This paper directly addresses this gap by proposing an integrated AIoT-Blockchain architecture for substation resiliency.

In the area of RL, Mohamed et al. [22] utilized RL to simulate adversarial strategies on frequency and voltage controls. Although this work focuses on increasing the complexity of attacks, no countermeasures are designed in this work. Said et al. [24] used Q-learning in an attempt to design an IDS with the ability to learn new threats. However, Q-learning and its variants are not efficient with scalability and convergence in the context of large state spaces, as noted by Shateri et al. [26] (DDQN) and Cai et al. [27] (DQN) in their respective

TABLE I: Comparison of Our Informer-PPO Approach with IEC 61850 Security Studies and Related Methods

Approach	Attack Modeling	ML	RL	Blockchain	Adaptive	Key Results	Strengths (+) / Weaknesses (–)
Duman et al. (2019) [11]	Attack Graphs	✗	✗	✗	✗	k-Supply metric for substation supply chain	+ Early quantitative metric – No IDS or mitigation
Duman et al. (2024) [12]	Risk Metrics	✗	✗	✗	✗	HFS improves posture under budget limits	+ Practical hardening guidance – No anomaly detection
Mohamed et al. (2023) [22]	Grid RL	✗	✓	✗	✗	RL-based attacker strategies	+ Captures adversarial complexity – No countermeasures
Ghosh et al. (2024) [16]	Device Auth	✗	✗	✓	✗	Blockchain-based device authentication	+ Tamper-proof authentication – No IDS adaptivity
Ustun et al. (2021) [13]	IDS ML	✓	✗	✗	✗	95.1% detection on GOOSE	+ Good supervised accuracy – No zero-day coverage
Said et al. (2024) [24]	RL-based IDS	✗	✓	✗	✗	Q-learning IDS for DDoS	+ Learns new threats – Limited scalability
Jin et al. (2022) [29]	PPO Security	✗	✓	✗	✗	PPO scheduling in cloud computing	+ RL stability – Not IEC 61850-specific
Nhung-Nguyen et al. (2024) [14]	IDS for GOOSE	✓	✗	✗	✗	≈98% detection accuracy	+ High accuracy – No mitigation, latency unreported
Park et al. (2024) [18]	Attack/Fault Cls.	✓	✗	✗	✓	Attack–fault classification for SV/GOOSE	+ Partial restoration logic – No blockchain, latency unreported
Zaboli & Hong (2025) [30]	GenAI ADS	✓	✗	✗	✗	>98% detection using synthetic GOOSE	+ Zero-day detection – Detection-only, no RL or enforcement
Our Approach (Informer-PPO)	RL+Blockchain	✓	✓	✓	✓	98.4% Kitsune, 97.6% ERENO, 35 ms latency	+ **Unified IDS + RL + Blockchain** + **Real-time mitigation with on-chain verification** – Prototype consensus; production deployments should adopt RAFT

studies. These methods are efficient in small cases but face computational inefficiencies in large substation domains.

Jin et al. [29] explored the application of PPO to cloud computing security environments. PPO offers improved training stability because of its clipped objective function and is thus more suited than traditional RL methods. However, PPO alone cannot effectively learn long-range dependencies that are inherent in complex cyberattack patterns, and its utility in dynamic infrastructure like substations is therefore limited. More recently, Zaboli & Hong [30] proposed a Generative AI-based anomaly detection system for IEC 61850 substations. Their approach introduced synthetic, protocol-compliant GOOSE datasets to enhance training balance and realism, and achieved strong anomaly detection accuracy. Yet, the system remained detection-only, with no integration of RL-driven adaptivity or blockchain-secured enforcement, which limits its practical deployment in real-time substation defense.

To surpass these limitations, Informer-based time-series models have been shown to be effective tools. Informer, founded on ProbSparse self-attention, is efficient in capturing long-range temporal dependencies at reduced computational costs. Sun et al. [31] demonstrated the application of Informer in energy management, achieving effective memory usage and improved detection performance. Shi et al. [32] developed InforTest, an anomaly detection framework based on Informer for robotic systems. These studies confirm Informer's ability to detect long patterns in complex data. However, none of them combine Informer with RL or use it in a real-time substation application.

Even though this study focuses primarily on substation environments, the applicability of suggested long-term anomaly detection integration with adaptive RL and blockchain-secured enforcement also extends to more generic IoT-fueled cyber-physical systems. Certain applications, such as smart city infrastructure, distributed energy resources, and industrial internet of things (IIoT) applications, might utilize the modular nature of the framework in enhancing resilience in security in a similar fashion.

Table I provides a comparative summary of both IEC 61850-focused security studies and related works in reinforcement learning, machine learning, and blockchain that inform substation defense design. Some recent efforts have advanced supervised anomaly detection for GOOSE traffic, others have introduced fault-aware classification and partial restoration capabilities, and still others have applied generative models for synthetic data-driven anomaly detection. Yet these approaches remain limited to isolated components of the security stack. None of them unify long-term anomaly detection, adaptive reinforcement learning, and decentralized blockchain enforcement within a single framework. Our Informer–PPO system is the first to integrate these dimensions, delivering real-time detection, auditable on-chain policy execution, and scalable mitigation tailored for IEC 61850 substations.

III. Proposed Framework

Increasing complexity in supply chain attacks on IEC 61850 substations demands an adaptive and robust security framework that can dynamically detect, mitigate, and prevent evolving threats. The AIoT-Blockchain security framework herein integrates Informer-based deep learning models for anomaly detection, RL for adaptive threat mitigation, and blockchain for data integrity assurance and device authentication. Unlike

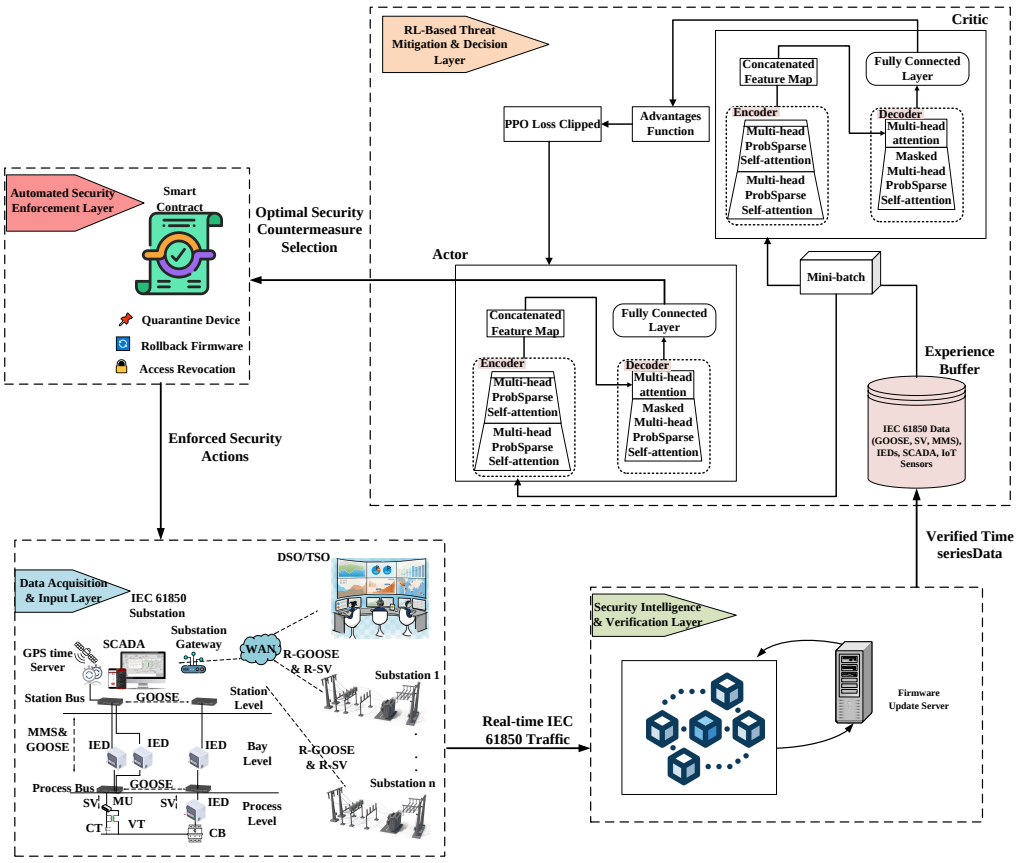


Fig. 1: System overview of the Informer–PPO framework with blockchain-verified enforcement.

traditional IDSs relying on pre-defined rules or static machine learning-based systems, the suggested framework leverages PPO-based RL to learn optimal attack mitigation policies in real-time continuously. The proposed framework consists of four key components. AIoT-based anomaly detection with Informer models for processing IEC 61850 traffic and supply chain-focused anomaly detection before propagation. Supply chain integrity through blockchain guarantees that every device deployed and firmware patches are verified and untampered. RL-based adaptive attack mitigation selects optimal countermeasures against recognized threats in real-time. Automated security response via smart contracts provides real-time security policy enforcement with zero human intervention. Fig. 1 illustrates the envisioned system design by integrating an anomaly detection module, IoT-enabled sensors, RL-based decision-making, blockchain authentication, and an automated security response mechanism. These hardware components combined form a robust cybersecurity system for IEC 61850 substations and effectively deter or counter possible cyberattacks. The following are descriptions of each module and its purpose.

A. AIoT-Enabled Data Acquisition and Anomaly Detection

IEC 61850 process, bay, and station levels (Fig. 1) generate high-frequency, time-critical telemetry. Our sensors and taps collect (i) electrical measurements from current transformers (CTs), voltage transformers (VTs), and circuit breakers (CBs) through merging units (MUs) that transmit sampled value (SV) streams, (ii) protection/control events from intelligent electronic devices (IEDs), and (iii) network traffic such as GOOSE, SV, and MMS. Although routable GOOSE (R-GOOSE) and routable SV (R-SV) support wide-area operation, this work focuses on classical substation links and leaves wide-area threat detection for future investigation.

Given a multivariate sequence $X_{t-n:t}$, we compute an anomaly score

$$A_t = g(X_{t-n}, \ldots, X_t)\,, \tag{1}$$

and raise an alert when $A_t > \theta$. Here, $g(\cdot)$ is learned from historical telemetry to model normal behavior, and θ is a calibrated threshold. The score A_t is fed into the RL state along with device-integrity indicators, enabling mitigation policies that are sensitive to both traffic dynamics and supply chain status.

B. Blockchain for Device Integrity and Supply Chain Authenticity

To resist device impersonation and unauthorized firmware changes, we adopt a permissioned blockchain for provenance and audit. Each asset (IED, CT, VT, MU, CB) carries an immutable identity record and firmware hash on-chain; updates must match signed, approved artifacts. Time-critical protections such as GOOSE blocking remain off-chain to respect sub-4 ms deadlines, whereas non-time-critical actions,

such as quarantine, rollback, and credential revocation, are enforced via smart contracts to ensure non-repudiation and verifiable policy compliance. Implementation details of consensus, leader election, validation, and block addition are given in Section IV and empirically linked to enforcement latency in Section V-D. For experimental verification, we used the SOLO consensus algorithm for its deterministic ordering and low overhead; for deployment, a RAFT ordering service is recommended, as it provides crash-fault tolerance through leader election, heartbeat synchronization, and log replication, thereby ensuring robust block ordering prior to peer validation and commit (see Section IV).

Each device D_i is anchored by a cryptographic digest:

$$H(D_i) = \text{Hash}(\text{Device ID}, \text{Manufacturer}, \text{Firmware Version}), \tag{2}$$

and every firmware update F_t is validated against this record before commit. The blockchain enforces:

1) *Device provenance:* immutable registration of IEDs, CTs, VTs, MUs, and CBs;
2) *Firmware integrity:* rollback or rejection of updates not matching ledger hashes;
3) *Tamper-proof event logging:* auditable records of configuration and access changes.

Smart contracts operationalize these checks by isolating compromised devices, revoking unauthorized access, or logging security violations. Where latency permits, GOOSE and MMS commands are cross-verified against blockchain records, while critical protection flows remain off-chain to meet sub-4 ms deadlines. The end-to-end enforcement path, including endorsement policies, membership service provider (MSP) identity checks, access control list (ACL) enforcement, multi-version concurrency control (MVCC) validation, and final block commit, is described in Section IV, with performance results in Section V-D.

C. RL-Based Adaptive Threat Mitigation (PPO with Informer)

To effectively address cyberattacks in IEC 61850 substations in practice, our solution in this paper combines the decision-making capability of PPO with the long-sequence modeling capability of Informer. The complex structure and integration of this Informer-PPO framework are elaborated in the following three parts.

1) Reinforcement Learning Setup for Cyber Mitigation: The increasing sophistication and evolving nature of cyberattacks on IEC 61850 substations necessitate a mitigation process capable of continuous learning and adaptation. In the direction of such a framework, the proposed methodology integrates PPO-based RL and an added improvement from Informer in both the Actor and Critic networks. This approach enables the system to continuously update decision-making policies to counter cyberattacks with the help of long-range temporal dependencies in time-series data.

RL learns within the framework of a Markov decision process (MDP), where an agent is paired with an environment to learn an optimal policy that maximizes cumulative rewards. The state space S in this case is the current security posture of the substation, consisting of real-time IEC 61850 network data, anomaly scores, device integrity status, and historical attack trend. The action space A is possible mitigation actions, i.e., quarantining an infected device, rolling back firmware to a verified state, or blocking suspicious network traffic. The system transitions to states based on policy $\pi(a|s)$, which takes security states to the best responses to threats.

At each timestep t, the RL agent receives a state s_t, selects an action a_t, and transitions to the next state s_{t+1}, where the reward r_t captures the effectiveness of the employed security action. We developed a reward function that promotes optimal threat response by rewarding effective mitigation while penalizing missed attacks (false negatives) more heavily than false alarms (false positives). This ensures a security-first approach while minimizing unnecessary operational disruptions. It is defined as:

$$\begin{aligned} r_t = \alpha_1 R_{\text{mitigation}} - \alpha_2 \mathbb{1}_{\text{FN}} - \beta_1 R_{\text{operational disruption}} \\ - \beta_2 \mathbb{1}_{\text{FP}} - \gamma \mathbb{1}_{\text{inaction}} \end{aligned} \tag{3}$$

where:

- $\mathbb{1}_{\text{FN}}$ is an indicator function that equals 1 if a false negative (missed attack) occurs, and 0 otherwise.
- $\mathbb{1}_{\text{FP}}$ is 1 if a false positive security action is triggered, and 0 otherwise.
- $\mathbb{1}_{\text{inaction}}$ is 1 if no action is taken when an attack occurs, and 0 otherwise.
- $R_{\text{mitigation}}$ measures the effectiveness of security actions ($0 \leq R \leq 1$).
- $R_{\text{operational disruption}}$ measures the negative impact of a security action on system operations.
- $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma$ are tunable weights.

This arrangement rewards proper attack detection and response, punishes missed attacks more than false alarms, and discourages pointless security measures that may cause interruptions in normal activities.

To optimize learning of the policy, PPO is employed, which improves the earlier policy gradient techniques by offering a clipped surrogate objective to stabilize training. The PPO objective function is defined as:

$$\begin{aligned} J(\theta) = \mathbb{E}\Bigg[\min \Bigg(& \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} A_t, \\ & \text{clip}\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, 1-\epsilon, 1+\epsilon \right) A_t \Bigg) \Bigg] \end{aligned} \tag{4}$$

where:

- θ is the set of trainable parameters in the current policy network (Actor),
- π_θ and $\pi_{\theta_{\text{old}}}$ are the current and past policy probabilities of taking action a_t in state s_t, respectively,
- A_t is the advantage estimate quantifying the relative value of action a_t in state s_t,
- ϵ is the clipping parameter (typically between 0.1 and 0.2) employed to clip the policy update ratio, preventing destabilizing updates.

PPO, using this clipped equation, provides strong convergence by constraining large deviations between successive policy improvements.

2) Informer-Augmented Actor-Critic Networks: While PPO provides stable and efficient policy optimization through its clipped objective, the effectiveness of cyber-mitigation in IEC 61850 substations depends critically on the representational ability of the neural network architecture. Advanced attacks, such as staged firmware injections or distributed GOOSE spoofing, evolve across long horizons; capturing these requires models that can extract long-term temporal dependencies rather than short-range correlations.

To this end, we employ the Informer architecture [33] as the function approximator for both Actor and Critic networks. Informer is a transformer variant designed for long-sequence forecasting, using ProbSparse self-attention to prioritize the most informative keys. This reduces complexity from $O(L^2)$ to $O(L \log L)$, enabling efficient modeling of thousands of timesteps while preserving accuracy for rare or stealthy patterns in substation traffic.

Formally, the Actor network parameterized by θ^A maps an observed state s_t to a probability distribution over mitigation actions:

$$\pi_{\theta^A}(a_t|s_t) = \text{softmax}\left(\text{Informer}_{\theta^A}(s_t)\right), \tag{5}$$

where $\text{Informer}_{\theta^A}(s_t)$ is the encoded representation of s_t and θ^A are the Actor's parameters. The Critic network parameterized by θ^C estimates the state-value function:

$$V_{\theta^C}(s_t) = \text{Informer}_{\theta^C}(s_t), \tag{6}$$

where θ^C denotes the Critic parameters.

This joint use of Informer ensures temporal consistency: the Actor leverages long-range dependencies for adaptive action selection, while the Critic evaluates rewards with extended context, stabilizing PPO training. By comparison, CNN-based IDSs capture only short local patterns, LSTMs suffer from vanishing gradients and short-term bias, and standard Transformers incur quadratic complexity that is prohibitive for long telemetry streams. Informer's sparsity-aware mechanism overcomes these limitations, dynamically attending to delayed or rare behaviors such as stealthy device compromises.

Therefore, the Informer-augmented PPO framework provides a principled solution to the scalability and responsiveness challenges of IEC 61850 cybersecurity. It delivers real-time decision-making under bursty and asynchronous traffic, offering resilience against multi-stage intrusions while remaining computationally feasible for deployment.

3) Fusion Architecture and Operational Workflow: Building on the Actor–Critic formulation, the Informer encoder-decoder backbone jointly supports policy learning and value estimation, enabling long-range temporal dependencies in GOOSE, SV, and IED telemetry to be captured for adaptive mitigation.

During training, the PPO agent interacts with a simulated adversarial environment in which cyberattacks are dynamically generated from historical incidents. Across episodes, the agent iteratively reduces the gap between predicted and observed outcomes, converging to policies that yield stable and effective countermeasures.

At runtime, mitigation actions are prioritized by operational urgency. Time-critical responses, such as breaker tripping or blocking compromised GOOSE traffic, are executed directly through SCADA/IED interfaces to comply with IEC 61850's strict sub-4 ms deadlines. Non-time-critical actions, including device quarantine, firmware rollback, or credential revocation, are encapsulated in digitally signed blockchain transactions. These are validated asynchronously and immutably recorded, ensuring tamper-proof auditability without introducing latency to urgent protective functions.

This dual-path execution mechanism balances the immediacy of direct enforcement with the decentralized trust of blockchain-backed actions. By decoupling urgent protections from consensus delays, the framework ensures both operational safety and verifiable compliance.

D. Automated Security Enforcement via Smart Contracts

Preventing cyberattacks on IEC 61850 substations requires adaptive responses that are both timely and tamper-proof. Conventional centralized enforcement is prone to delays and manipulation. In our framework, blockchain smart contracts autonomously enforce security policies, ensuring that RL-selected countermeasures are applied without human intervention. Once the PPO agent determines the optimal action, it is encapsulated in a digitally signed transaction and validated on the ledger, rendering responses such as device quarantine, firmware rollback, or credential revocation immutable and auditable.

The sequence of enforcement proceeds as follows:

1) **Threat detection and policy selection:** The RL agent chooses an optimal countermeasure based on AIoT telemetry and historical attack context.
2) **Contract invocation:** The chosen action is submitted to a designated smart contract.
3) **Validation and execution:** The contract checks policy rules and context, verifies authorization, and enforces the action.
4) **Tamper-proof logging:** The decision and its outcome are irreversibly recorded on the blockchain for auditability.

1) Mathematical Model for Security Enforcement: A smart contract SC is a self-executing function that maps an action request a_t to an execution decision $\mathcal{E}_t$, considering system constraints:

$$\mathcal{E}_t = SC(a_t, P, C) \tag{7}$$

where, a_t is the security action determined by PPO at time t. P is the set of pre-defined policy rules (e.g., firmware rollback conditions, access control policies). C is the current security context, comprising device authentication status and substation operating state. Smart contract logic is designed to ensure that:

$$\mathcal{E}_t = \begin{cases} \text{execute } a_t, & \text{if } (a_t \in P) \text{ and } (C \text{ is compliant}) \\ \text{reject } a_t, & \text{otherwise} \end{cases} \tag{8}$$

This blocks illegitimate security actions only from being enforced and prevents unauthorized interference.

Algorithm 1 Adaptive Cyber Defense via Informer-PPO with Smart Contract Enforcement

Input: IEC 61850 telemetry streams (GOOSE, SV, MMS); initialized PPO parameters (θ_A, θ_C) and hyperparameters $(\gamma, \lambda, \epsilon, \eta)$; blockchain connection and deployed smart contracts (`InvokeContract()`, `VerifyAndEnforce()`).
Output: Trained policy $\pi^{\star}_{\theta_A}$; per-timestep mitigation decisions a_t; commit receipts (`MitigationCommitted`); audit log entries $\{\mathcal{L}_t\}$.
Start
Initialize replay buffer $\mathcal{B} \leftarrow \varnothing$
for each training episode **do**
  Collect telemetry s_t from IEC 61850 traffic
  Compute anomaly score and device integrity status
  Construct augmented state $s'_t \leftarrow [s_t, \text{anomaly score}, \text{device status}, \text{attack history}]$
  Select action $a_t \sim \pi_{\theta_A}(s'_t)$ using policy in Eq. 5
  Apply a_t in simulator; observe reward r_t and next state s'_{t+1}
  Store transition $(s'_t, a_t, r_t, s'_{t+1})$ in buffer $\mathcal{B}$
  Estimate advantage A_t with PPO objective (Eq. 4)
  Update Actor θ_A and Critic θ_C via clipped loss
  if a_t requires blockchain enforcement **then**
    Construct signed payload $\mathcal{T} \leftarrow \texttt{Sign}(a_t, \text{timestamp}, \text{deviceID})$
    Submit payload: $\texttt{InvokeContract}(\mathcal{T})$
    $\mathcal{E}_t \leftarrow \texttt{VerifyAndEnforce}(\mathcal{T})$ {returns commit event}
    Log enforcement $\mathcal{L}_t \leftarrow (\mathcal{T}, \mathcal{E}_t)$
  end if
end for
End

2) Types of Security Actions Enforced by Smart Contracts: The smart contract platform supports various security enforcement actions that are activated according to the anomaly severity:

- Device Quarantine: When an IED, CT, VT, MUs, or CB is detected to act suspiciously, the smart contract isolates it from the network so that the attack is not propagated further.
- Firmware Rollback: If a device's firmware has been tampered with or changed in any fashion, the contract examines its original authenticated version on the blockchain and causes a rollback.
- Revocation of Access: All unauthorized access attempts are tracked, and if detected, the smart contract revokes access or rejects malicious traffic.

3) Blockchain-Based Auditability and Compliance: Every enforced security measure is cryptographically signed and immutable on the blockchain. This ensures:

- Non-repudiation: Every security event can be traced back to its source.
- Tamper resistance: Once a measure has been enforced, it cannot be changed.
- Regulatory compliance: Enforcement based on smart contracts is cybersecurity policy compliant for critical infrastructure.

4) Interaction Between Smart Contracts and RL-Based Mitigation: Upon selecting an action by the PPO agent, the intelligent contract carries it out automatically, making unauthorized or late modifications impossible. The seamless integration facilitates real-time mitigation of cyber threats without losing the integrity and robustness of IEC 61850 substations. With the combination of RL-adaptive decision-making and blockchain-enforced automated compliance, this approach promises a tamper-evident, autonomous, and secure cybersecurity solution for modern power substations.

To officially initiate blockchain enforcement, all action a_t created by RL is converted into a digitally signed transaction payload named $\mathcal{T}$. This payload includes action, timestamp, and device ID, and is uploaded to the blockchain smart contract for verification and enforcement. After receiving $\mathcal{T}$, the smart contract processes $\mathcal{T}$, verifies against policy and context, and sends an execution result $\mathcal{E}_t$ as to whether or not the action was approved and enforced. These transactional variables are reflected in the control flow described by Algorithm 1. This algorithm summarizes the end-to-end operational logic of our proposed RL-blockchain framework, outlining how the system collects real-time substation data, determines the optimal mitigation actions using the Informer-PPO agent, and enforces them through smart contracts on the blockchain.

IV. Implementation & Experimental Setup

A. Simulation Environment and Blockchain-Based Enforcement

A simulation model of the IEC 61850 digital substation system was developed using **OMNeT++**, a network simulator based on discrete events, and the **NS3** network communication simulation framework between IEDs, CTs, VTs, MUs, CBs, SCADA, and HMI devices. The testbed simulates real-time GOOSE, SV, and MMS traffic under benign and adversarial settings, including supply chain-driven attacks, to validate the resilience of the framework to coordinated substation attacks.

TABLE II: Summary of Datasets Used: Simulated Attacks, Features, and Preprocessing Pipeline

Dataset	Attack Scenarios	Protocol(s)	Main Features	Preprocessing and Use in Paper
Biswas et al. [29]	9 cyberattacks + 3 benign	GOOSE	Breaker status, current values, `stNum`, `sqNum`, timestamps	Feature extraction from `.pcapng` and `.csv`, scenario-level labeling, substation configuration language (SCL)-based time-series alignment. Used for supervised training and GOOSE anomaly detection.
Kitsune [30]	9 real-world attack scenarios	transmission control protocol (TCP)/internet protocol (IP), address resolution protocol (ARP), secure sockets layer (SSL)	115-dimensional statistical features (packet flow, timing, entropy)	Informer pretrained on real traffic anomalies. Validates generalization under real network noise. Used for cross-domain transfer learning.
ERENO [31]	7 attack scenarios	GOOSE, SV	Protocol timing, flow statistics, replay and injection markers	Used for benchmarking detection robustness on industrial control systems (ICS)-specific attacks. Attacks are annotated, and data is high-fidelity from substation simulation.
Custom OMNeT++–NS3	5 synthetic attacks	GOOSE, SV, MMS (simulated)	Device identity, control message content, firmware hashes	Simulated MMS spoofing, unauthorized IEDs, and firmware injection. Used to test blockchain enforcement (rollback/quarantine) and RL-triggered response.

Blockchain-Based Enforcement System: A multi-peer Hyperledger Fabric network was implemented to establish a tamper-proof security enforcement system for IEC 61850 substations. The three key constituents of the blockchain network are:

- Ordering Service – Ensures transaction consistency and blocks are committed in sequence.
- Peer Nodes – Every peer node consists of a distributed ledger, verifying firmware updates, authentication logs, and security actions.
- Certificate Authority (CA) – Provides cryptographic identity management for IEDs so that the network is joined only by enrolled and authenticated devices.

Transaction and block addition path: (1) The RL agent submits a proposal to endorsing peers that satisfy a 2-of-3 organizations policy. (2) Endorsers simulate chaincode and return signed read/write sets (with MVCC versions). (3) The client assembles endorsements and sends the transaction to the orderer. (4) The ordering service batches transactions using `BatchTimeout`/`MaxMessageCount` and cuts a block. (5) Each peer validates the block by checking (i) endorsement policy signatures, (ii) MSP/identity, and (iii) MVCC conflicts. (6) *Valid* transactions are committed to the ledger and applied to world-state; *invalid* ones are flagged but not applied. On commit, an application event (`MitigationCommitted`) is emitted and used to mark a mitigation as verified.

Smart Contract Logic: Go-written smart contracts were used as Chaincode on peer nodes to enforce security policy automatically. The contracts compare firmware update requests with cryptographic hashes of blockchain-stored, signed firmware versions. In the case of a discrepancy, the contract triggers automatic security enforcement actions such as firmware rollback or device quarantine.

Integration with RL-Based Security System: The blockchain interacts with the RL-based security system through a RESTful API, which offers:

- Threat response enforcement – Receiving attack mitigation directives from the PPO RL-based Informer augmentation.
- Tamper-proof security logging – Saving security actions (quarantine, rollback, access revocation) immutably on the blockchain ledger.
- Real-time firmware integrity validation – Providing ongoing attestation to SCADA and IEDs.

Each transaction is hashed for auditability and non-repudiation. Dynamic simulation of transaction load was employed to examine performance involving enforcement latency and scalability under adversarial stress.

B. Blockchain Configuration and System Parameters

The Hyperledger Fabric network was initialized with three peer organizations and a single ordering service. During prototyping and experiments in a controlled environment, the SOLO consensus algorithm was utilized. With SOLO, transactions are deterministically ordered with low latency and overhead, making it suitable for early experimental verification of blockchain-based enforcement in IEC 61850 substations.

For deployment, however, we recommend a crash-fault-tolerant RAFT ordering service. RAFT elects a leader among orderers using randomized timeouts and heartbeat messages; the leader serializes transactions into a replicated log, and once a majority of orderers confirm replication, a block is cut and

broadcast to peers. Client SDKs automatically redirect requests to the active leader after re-election, making leader changes transparent to applications.

Each of the smart contract transactions, as implemented in our Hyperledger Fabric testbed deployment, included a hashed firmware signature, device ID, and enforcement action, such as firmware rollback or device quarantine. The transaction lifecycle proceeded through the standard Fabric endorsement–ordering–commit path: (1) the RL agent submits a proposal to endorsing peers; (2) endorsers simulate the chaincode and return signed read/write sets with version metadata; (3) the client assembles endorsements and submits to the orderer; (4) the ordering service batches transactions and cuts a block; (5) each peer validates the block by checking endorsement policy signatures, MSP identities, and MVCC version conflicts; and (6) valid transactions are committed to the ledger and applied to world state, while invalid ones are flagged but preserved for audit. Upon commit, an application event (`MitigationCommitted`) is generated, and a mitigation is considered verified only after this event is received.

All performance tests were conducted on a 12-core Intel CPU and 32 GB RAM workstation. To address IEC 61850 substation operational requirements, Hyperledger Fabric was selected as the blockchain platform over others such as Ethereum and Corda. Fabric has native support for permissioned networks, low-latency transaction endorsement, and modular chaincode architecture, all of which are fundamental to security-critical industrial infrastructures. Unlike Ethereum, Fabric eliminates gas fees, enables private data sharing among selected participants, and ensures deterministic transaction finality in controlled membership networks. While Corda has some privacy advantages and financial optimization features, it is less flexible in its modular smart contract design or as versatile overall as Fabric for substation cybersecurity enforcement.

Blockchain performance testing results with transaction throughput and latency details are given in Sections V.D and V.E.

C. Dataset Description and Trace Composition

This research hires four datasets to pretrain, train, and experiment with the proposed RL-based anomaly enforcement and detection framework. The datasets consist of two publicly available IEC 61850-specific datasets (Biswas et al. [34] and ERENO) [35], a general-purpose real-world network intrusion dataset (Kitsune) [36], and one synthesized custom dataset created within our OMNeT++–NS3 substation testbed. Table II summarizes each dataset's attack scenarios, protocol coverage, extracted features, and preprocessing. The IEC 61850 Security Dataset by Biswas et al. [34] includes 12 scenarios, 9 cyberattacks, and 3 benign, exclusively concerning GOOSE protocol communications. They include attacks such as GOOSE flooding, spoofed state/sequence number manipulation, false current injection, and control message tampering. Each scenario comprises '.pcapng' traces with timestamped GOOSE packets among 18 IEDs and corresponding '.csv' logs of present magnitude and breaker status. Preprocessing consisted of extracting timestamps, GOOSE control fields (`stNum`, `sqNum`), and aligning them with SCL configuration files for time-series structural modeling. For estimating generalization on real data, we used the Kitsune Network Attack Dataset [36], with nine labeled attack scenarios recorded from an operational IoT surveillance network. Although it does not cover IEC 61850 protocols, it covers ARP spoofing, replay injection, and denial-of-service (DoS) attacks impersonating real traffic anomalies. We pre-trained the Informer model on Kitsune to learn generalized adversarial patterns and then fine-tuned it on IEC 61850-specific datasets. The ERENO IEC 61850 IDS Dataset [35] was utilized for the realistic evaluation of substation-centric attacks. It contains seven scenarios of GOOSE and SV traffic under replay, flooding, and masquerading conditions. Each data instance includes network-level flow features and labels indicating whether an attack occurred. This data set was used to test the resilience of our model in high-fidelity substation environments. To test attack vectors not present in any public dataset, we have developed a custom dataset within the OMNeT++–NS3 environment. This simulated data set includes GOOSE, SV, and MMS traffic. Attack scenarios include unapproved IED registration, firmware injection, and MMS spoofing. We used this dataset to test our blockchain-enforcement logic, e.g., automatic rollback and quarantine responses triggered by RL-based detection outputs. These aggregated datasets enable end-to-end training and validation in both realistic and diverse adversarial environments. The public datasets offer the promise of benchmarking relevance, whereas our synthetic traces enable testing of state-of-the-art scenarios involving protocol-level and control-level substation attack threats.

D. Synthetic Attack Generation in OMNeT++–NS3

To complement public data sets and study complex policy enforcement behaviors, new adversarial test cases were created on our OMNeT++–NS3-based IEC 61850 substation testbed. Such test cases include attack vectors beyond the bounds of existing data sets, for example, spoofed MMS control messages, malicious IED registration attempts, and firmware injection on device integrity. The anomaly detection module and the blockchain response system were tested across these synthetic threats.

The subsequent types of attacks were specifically designed to initiate blockchain-enforced actions, such as firmware rollback and device quarantine, in a bid to combat unauthorized behavior and maintain substation stability.

Scenario Categories: The public and synthetic traces combined dataset is organized as follows:

1) Normal Operations:
   - Authentic exchanges on GOOSE, SV, and simulated MMS protocols.
   - Secure firmware update logs on the blockchain ledger.
   - Verified exchanges among SCADA, IEDs, MUs, CTs, and VTs.
2) Cyberattack Situations:
   - Firmware Injection Attacks: Illegal firmware updates that compromise device functionality.

- IEC 61850 Message Spoofing: Spoofed GOOSE messages duplicating valid protection and control signals.
- Unauthorized Device Alterations: Attempts to register insurgent IEDs or alter installed substation equipment.
- Data Manipulation Attacks: Present and voltage measurement tampering affecting the operation of the relay.
- DoS Attacks: Disrupting substation network communications by flooding with excessive traffic.

E. Evaluation Metrics

To rigorously evaluate our Informer-PPO–Blockchain security framework, we use a complete set of measures that quantify detection performance, mitigation effectiveness, and enforcement delay, as follows:

1) Anomaly Detection Measures: To contrast the performance of different IDS architectures (rule-based, CNN-PPO, LSTM-PPO, Transformer-PPO, and Informer-PPO), we utilized the following typical classification measures:
   - Precision: The number of predicted anomalies that were actual attacks.
   $$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$
   - Recall (Sensitivity): The number of actual attacks detected correctly.
   $$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$
   - False Positive Rate (FPR): The proportion of normal events incorrectly classified as attacks.
   $$\text{FPR} = \frac{FP}{FP + TN} \tag{11}$$
   - False Negative Rate (FNR): The proportion of actual attacks missed by the model.
   $$\text{FNR} = \frac{FN}{FN + TP} \tag{12}$$
   - F1-Score: Harmonic mean between recall and precision.
   $$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$
   - Matthews Correlation Coefficient (MCC): Measures the quality of binary classifications even in the presence of class imbalance.
   $$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{14}$$

   where:
   - TP: True Positives (properly detected attacks)
   - TN: True Negatives (properly ignored normal traffic)
   - FP: False Positives (normal traffic incorrectly labeled as attacks)
   - FN: False Negatives (detection failures, failed-to-detect attacks)

   These metrics were calculated on test subsets of every dataset—Biswas [34], ERENO [35], and Kitsune [36]—to provide insight into real-world, protocol-specific, and cross-domain generalizability.
2) Attack Mitigation Metric: To quantify the effectiveness of the RL agent in selecting effective responses under adversarial conditions, we use mitigation success rate (MSR), calculated as follows:
   $$MSR = \frac{\text{Number of Correctly Mitigated Attacks}}{\text{Total Number of Attack Attempts}} \times 100\% \tag{15}$$
   The mitigation of a threat is considered accomplished when the chosen RL action is both optimal for the threat type and system policy and is successfully enforced and confirmed by smart contract enforcement logs on the blockchain. This metric reflects the end-to-end performance of the detection + decision-making + enforcement loop.
3) System Responsiveness Metrics
   - Detection + Mitigation Latency: Average time (in ms) from when an anomaly is detected to when the RL agent makes an action decision.
   - Blockchain Enforcement Latency: Time to validate, authorize, and execute a security action on-chain.

   These measures were recorded with timestamp logs in the OMNeT++–NS3 testbed and Hyperledger Fabric chaincode responses.

V. Results & Discussion

This section gives a comprehensive evaluation of the proposed Informer-PPO–Blockchain system on a range of performance indicators, including accuracy of anomaly detection, mitigation efficiency, model convergence, computational overhead, and enforcement of blockchain. All experiments were conducted on a custom IEC 61850 substation simulation setup in OMNeT++–NS3, with adversarial scenarios taken from three datasets: a synthetic IEC 61850 testbed, the Biswas et al. [34] dataset, and two additional benchmark datasets included, Kitsune [36] and ERENO [35], for generalization and robustness across heterogeneous cyber-physical infrastructures. Performance metrics are defined in Section IV.E and used throughout this section for consistency and reproducibility.

A. Anomaly Detection Performance

Table III illustrates the performance of four comparison methods—rule-based IDS, CNN-PPO, LSTM-PPO, Transformer-PPO, and proposed Informer-PPO, in anomaly detection as evaluated on four different datasets: our own in-house custom-developed synthetic IEC 61850 testbed, the Biswas et al. [34], Kitsune [36], and ERENO [35]. Important metrics include accuracy, precision, recall, F1-score, FPR, FNR, and MCC, all of which are discussed in Section IV.E.

TABLE III: Anomaly Detection Performance Across Datasets and Models

Model	Dataset	Accuracy (%)	Precision	Recall	F1-Score	FPR (%)	FNR (%)	MCC
Rule-Based IDS	Synthetic Testbed	85.4	0.78	0.82	0.80	7.9	13.6	0.65
CNN-PPO	Synthetic Testbed	90.9	0.86	0.87	0.87	4.6	9.2	0.76
LSTM-PPO	Synthetic Testbed	93.3	0.89	0.91	0.90	3.0	6.7	0.82
Transformer-PPO	Synthetic Testbed	94.7	0.91	0.92	0.91	2.6	5.4	0.85
Informer-PPO	**Synthetic Testbed**	**95.9**	**0.93**	**0.94**	**0.94**	**2.1**	**4.3**	**0.88**
Rule-Based IDS	Biswas dataset	87.3	0.80	0.84	0.82	7.1	12.3	0.69
CNN-PPO	Biswas dataset	91.2	0.87	0.88	0.88	4.3	8.6	0.78
LSTM-PPO	Biswas dataset	93.6	0.90	0.92	0.91	3.1	6.2	0.84
Transformer-PPO	Biswas dataset	94.8	0.92	0.92	0.92	2.5	5.2	0.86
Informer-PPO	**Biswas dataset**	**96.1**	**0.94**	**0.94**	**0.94**	**2.0**	**4.0**	**0.89**
CNN-PPO	Kitsune dataset	90.1	0.86	0.85	0.86	5.7	8.9	0.75
LSTM-PPO	Kitsune dataset	94.0	0.91	0.92	0.91	3.2	5.8	0.85
Transformer-PPO	Kitsune dataset	95.2	0.93	0.94	0.93	2.4	4.6	0.88
Informer-PPO	**Kitsune dataset**	**98.4**	**0.97**	**0.98**	**0.98**	**1.1**	**2.3**	**0.94**
CNN-PPO	ERENO dataset	89.5	0.85	0.83	0.84	6.0	9.1	0.74
LSTM-PPO	ERENO dataset	92.4	0.88	0.91	0.89	3.8	6.0	0.81
Transformer-PPO	ERENO dataset	94.2	0.90	0.92	0.91	2.7	5.8	0.84
Informer-PPO	**ERENO dataset**	**97.6**	**0.95**	**0.96**	**0.95**	**1.5**	**3.1**	**0.91**

On the synthetic IEC 61850 testbed, emulating time-critical GOOSE/SV communication patterns and baseline substation cyber-traffic, the Informer-PPO achieves 95.9% accuracy, significantly outperforming Transformer-PPO (94.7%), LSTM-PPO (93.3%), CNN-PPO (90.9%), and rule-based IDS (85.4%). The model also demonstrated an improved balance between F1-score (0.94) and MCC (0.88), indicating good quality of prediction even under conditions of class imbalance. Synthetic testbeds, however, fail to capture real-world noise and operating diversity, yet they represent important controlled settings for establishing baseline detection performance and validating convergence stability before deployment on dynamic infrastructures.

On the Biswas dataset, which records structured IEC 61850 substation attack flows, Informer-PPO achieved 96.1% accuracy and a high MCC of 0.89, outperforming all Transformer-, CNN-, and LSTM-based RL counterparts. The detection capability of the model for stealthy firmware injection and spoofing behavior further enhances its usability for protocol-layer anomaly detection.

Inference on the Kitsune dataset, with realistic anomalies and heterogeneous IoT traffic, showed that the Informer-PPO generalizes well. It was 98.4% accurate with nearly perfect precision (0.97) and recall (0.98). This result shows excellent resistance to noise and cross-domain traffic patterns that were not observed at pretraining. Compared to Transformer-PPO (95.2%), CNN-PPO (90.1%), and LSTM-PPO (94.0%), the Informer-based structure provided greater temporal coverage and reduced FNR by a large margin.

For the ERENO dataset, simulating real-time hardware-based IEC 61850 operational environments with real-time GOOSE/SV payloads, Informer-PPO again led the pack with 97.6% accuracy, 0.95 F1-score, and 0.91 MCC. This points towards the ability of the framework to react to timing-critical event sequences as well as communication delays, both of which have been victimized by sequence models earlier.

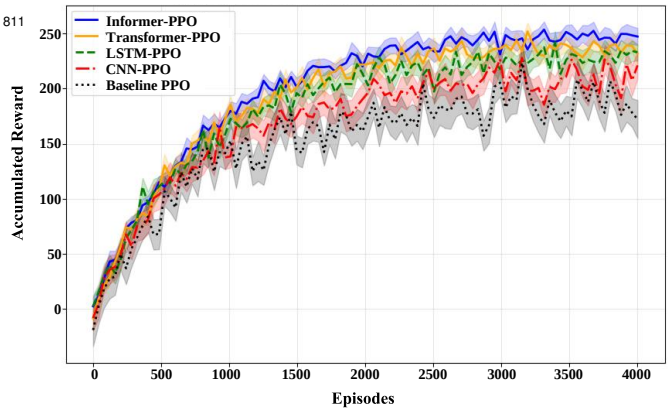


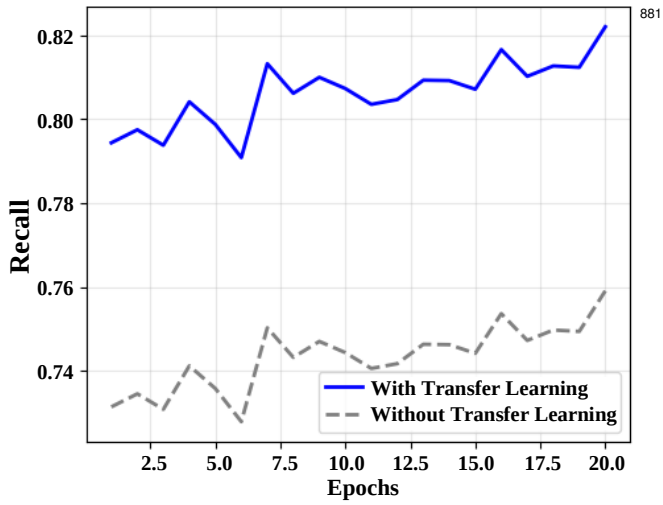
Fig. 2: Training Process of RL Agent: Cumulative Reward Over Episodes

These results indicate that Informer-PPO not only performs significantly better on synthetic and benchmark datasets but also lays a solid foundation for subsequent real-time mitigation and blockchain-secured policy enforcement.

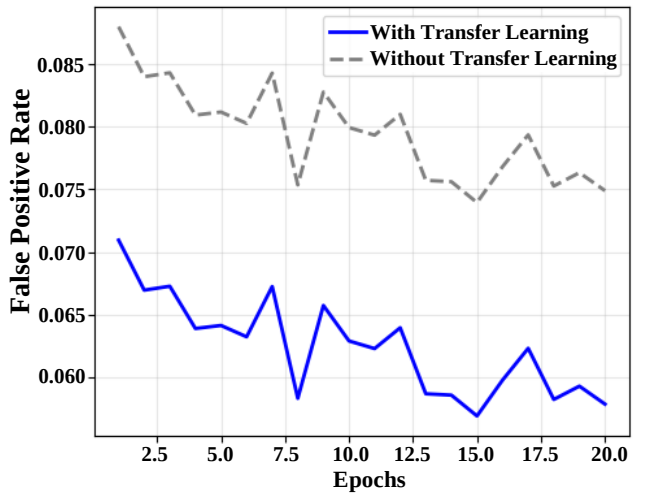
B. RL Agent Training and Convergence

We monitored the average episodic reward over training for all PPO-based models to compare the RL agents' learning dynamics and convergence behavior. The cumulative reward curves over 4,000 episodes on the synthetic IEC 61850 testbed are presented in Fig. 2, where all models were trained from scratch to enable fair baseline comparisons.

The Informer-PPO model described achieves faster and more convergent convergence than its Transformer-, CNN-, and LSTM-based counterparts. Specifically, Informer-PPO stabilizes its cumulant reward after ∼1,100 episodes, while Transformer-PPO converges slightly later, around 1,400



(a) Early Recall Comparison



(b) Early False Positive Rate Comparison

Fig. 3: Effect of Transfer Learning on Fine-Tuning Performance of Informer-PPO pretrained on Kitsune.

episodes. LSTM-PPO requires approximately 1,800 episodes to converge, and CNN-PPO experiences noisier and slower learning beyond 2,500 episodes. Baseline PPO has the slowest learning trajectory with instability in reward accumulation during the training process. Informer's ProbSparse attention enables it to learn long-term temporal dependencies without memory bottlenecks or computational inefficiencies, which are typical issues with recurrent or dense-attention models.

For cross-domain adaptability evaluation, Informer-PPO was first pre-trained on the Kitsune dataset and then fine-tuned on the Biswas and ERENO datasets. Transfer learning achieved an early-stage recall gain of 6.3% and FPR reduction of 1.7% over scratch training. Results, as represented in Figures 3a and 3b, confirm the ability of knowledge transfer to facilitate quicker convergence and enhancement of generalizability across other cyber-physical environments.

Furthermore, the Informer-PPO consistently exhibited lower reward variance over episodes (Fig. 2), reflecting enhanced training stability. This combination of stable training, fast convergence, and effective domain transfer is an indication of the dominance of Informer-PPO for real-time smart grid security applications.

C. Attack Mitigation Success Rate

In RL–based cybersecurity systems for critical infrastructure, effectively mitigating detected threats is as essential as accurately detecting them. The proposed architecture integrates a policy execution layer backed by a smart contract deployed on a Hyperledger Fabric blockchain network, enabling countermeasures such as firmware rollback, device isolation, and access revocation for the reinforcement learning agent.

MSR, which is defined in Section IV.E, calculates the percentage of successful mitigation of attacks out of total attempts. A mitigation action qualifies as a success only when the selected policy is enforced on-chain and marked as verified by the blockchain network.

To demonstrate this dimension numerically, Fig. 4 shows the MSR of four RL-based agents, CNN-PPO, LSTM-PPO, Transformer-PPO, and Informer-PPO, on four sample IEC 61850 attack types: firmware injection, GOOSE spoofing, unauthorized configuration, and data manipulation. The Informer-PPO agent shown here consistently has the highest success rate, over 92% in all categories, and up to 93.1% for unauthorized configuration events. Transformer-PPO performs better than LSTM-PPO, achieving MSR scores around 90–91% across different attack scenarios. LSTM-PPO performs fairly (approximately 88–90%), while CNN-PPO demonstrated comparatively lower mitigation success rates (approximately 84–86%), indicating challenges in modeling sequential adversarial patterns.

This relative performance gap is primarily attributed to the temporal modeling capacity of the Informer architecture. Compared to CNNs, which rely on localized kernels, and LSTMs, which are plagued by memory decay and vanishing gradients when dealing with long sequences, and Transformers, which struggle with computational scalability for longer sequences, the Informer employs a ProbSparse attention mechanism. It can thus maintain global temporal context and dynamically re-prioritize necessary time steps, enabling it to forecast multi-stage attack development and select context-aware mitigation policies.

These results build upon the Informer-PPO's advantage in temporal reasoning, response consistency, and compliance with on-chain policies, making it perfect for real-time protection in smart substations.

D. Latency and Computational Overhead

This subsection examines the end-to-end anomaly prevention pipeline's response latency and computational overhead to ascertain the proposed framework's real-time viability in IEC 61850 substations, including the RL agent inference and blockchain-based implementation.

1) Response Time Analysis: As evident from Table V, the Informer-PPO model presented here has an average

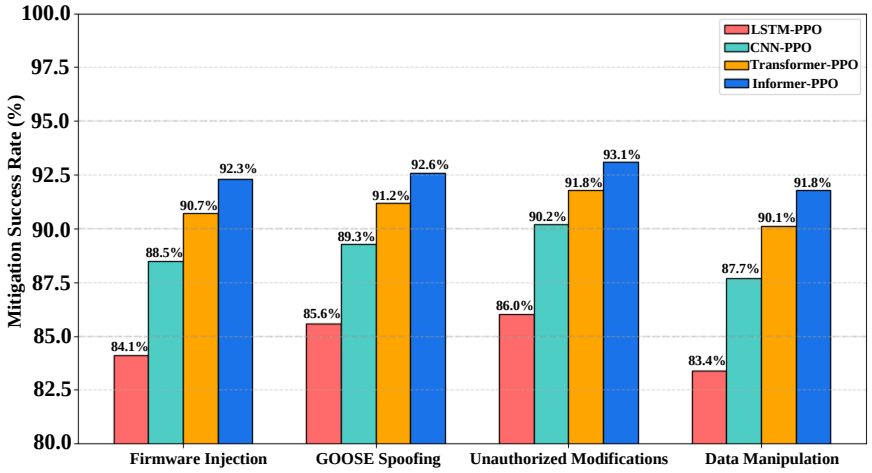


Fig. 4: Comparative Mitigation Success Rate by Model and Threat Type.

end-to-end response time of 35 ms for anomaly detection and mitigation, outperforming Transformer-PPO (52 ms), LSTM-PPO (72 ms), CNN-PPO (95 ms), and the baseline IDS (110 ms). This latency includes both the inference time of the RL agent and actuation delays initiated through smart contracts. These values make the use of Informer-PPO in actual field deployment feasible in substation settings that demand rapid detection-to-mitigation switching.

2) Blockchain Enforcement Delay: Blockchain adoption introduces additional processing delays to enforcement and verification, as shown in Fig. 5. Delays of more than three critical security measures, firmware rollback, access revocation, and device quarantine were all executed with standard transaction loads with Hyperledger Fabric. The following delays were incurred:

- Firmware Rollback: 42 ms
- Access Revocation: 50 ms
- Device Quarantine: 37 ms

Although higher than IEC 61850's protection messaging deadline (e.g., 4 ms for GOOSE), these latencies are acceptable for non-time-critical control flows such as firmware integrity checking and policy-based segmentation. Thus, blockchain enforcement is not employed in time-critical layers but supports auditability and tamper-proof logging for events initiated by the RL agent.

3) Blockchain Transaction Lifecycle and Validation: The observed enforcement delays are directly tied to the Fabric transaction processing pipeline. Each mitigation request is first endorsed by peers according to a 2-of-3 organizations policy, where endorsers simulate the chaincode and return signed read/write sets with version metadata. The client then assembles these endorsements and submits the transaction to the ordering service. In our benchmarks, a SOLO orderer was used for determinism and low overhead; in deployment, a RAFT cluster would be used to provide crash-fault tolerance through leader election, heartbeat messages, and replicated logs. The orderer batches transactions using `BatchTimeout` and `MaxMessageCount`, cuts a block, and broadcasts it to peers. Peers verify endorsement signatures, MSP identities, and MVCC read-set consistency before committing valid transactions to the ledger and discarding invalid ones. Upon commit, an application event (`MitigationCommitted`) is emitted, which the RL agent uses to confirm that a mitigation has been enforced and audit-logged. This path explains why enforcement actions such as firmware rollback (42 ms) and device quarantine (37 ms) remain within practical non-time-critical thresholds while still achieving tamper-proof auditability. The ordering and validation parameters that shaped these latency measurements are summarized in Table IV.

TABLE IV: Ordering and validation parameters of the Fabric testbed used in latency and throughput experiments.

Parameter	Value
Endorsement policy	2 of 3 organizations
BatchTimeout	100 ms (bench); 50–100 ms (deploy)
MaxMessageCount	20 (bench); 10–50 (deploy)
World state	LevelDB; MVCC enabled
Event	`MitigationCommitted` on valid commit

4) Computational Load and Hardware Feasibility: Informer-PPO was executed on a testbed with a 12-core Intel CPU and 32 GB of memory. While the Informer architecture improves long-range dependency handling, it requires slightly more computational power than Transformer-PPO, and notably more than CNN or LSTM equivalents. Inference load profiling shows that any policy choice, i.e., attention scoring and output action selection, completes under 12 ms on a single-threaded CPU core with PyTorch optimizations. GPU acceleration is recommended for high-frequency deployments, especially in multi-substation situations where agents handle diverse traffic in parallel. Future iterations may benefit from integrating quantized inference or model distillation to reduce model complexity

TABLE V: Average detection and mitigation response latency for different models.

Model	Avg. Response Time (ms)
Baseline IDS	110
CNN-PPO	95
LSTM-PPO	72
Transformer-PPO	52
Informer-PPO	**35**

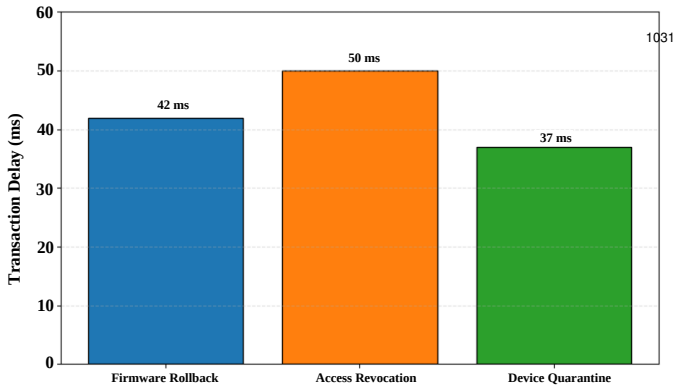


Fig. 5: Blockchain Enforcement Delay by Action Type. Average transaction time for firmware rollback, access revocation, and device quarantine executed via Hyperledger Fabric.

at the cost of policy fidelity.

5) Resource and Bandwidth Requirements: In addition to computational considerations, network bandwidth requirements for the proposed system remain practical for real-world deployments. Typical IEC 61850 telemetry traffic, including GOOSE, SV, and MMS messages, generates less than 1–2 KB/sec per device under normal substation operating conditions. AIoT-driven edge-level preprocessing ensures only anomalous or suspicious behavior is forwarded to the upper-level decision planes. It keeps total network traffic low even as it scales up to thousands of devices. Furthermore, the blockchain enforcement transactions emulated in the framework, being light in weight (approximately 220–300 bytes per transaction), impose minimal extra bandwidth demands. These aspects combined make the feasibility of real-time anomaly countermeasures and policy enforcement possible without overloading the computation and communication infrastructure.

E. Blockchain Enforcement and Scalability

The effectiveness of blockchain integration in substation environments depends not only on enforcement capability but also on scalability under changing rates of transactions. This subsection discusses the system's performance at higher policy submission rates, with attention to throughput, latency, and consensus stability.

Under stress testing, the Hyperledger Fabric network witnessed a steady 20–25 transactions per second (TPS) with policy enforcement latency ranging between 1.7 and 3.2 seconds based on the transaction density. The transaction payload was between 220 and 300 bytes, having no impact on the stability of execution. Transaction payloads varied between 220 and 300 bytes with no perceivable impact on execution stability. Such values guarantee the blockchain component's sufficiency for enforcing non-time-critical security responses, such as configuration rollback, access revocation, and device quarantine.

TABLE VI: Blockchain performance metrics under policy load.

Metric	Observed Value
Average Throughput (TPS)	20–25
Average Enforcement Latency	1.7–3.2 seconds
Transaction Payload Size	220–300 bytes
Consensus Mode	SOLO (benchmark setting)
Degradation Observed	None up to 25 TPS

Table VI summarizes the blockchain's behavior observed. No performance impact was observed at a rate of up to 25 TPS, reflecting sufficient headroom for deployment in substations with low policy change rates.

The SOLO consensus protocol was used for benchmarking in stress testing due to its low resource requirements and ease of use. However, as SOLO does not deliver distributed crash or Byzantine fault tolerance, it is unsuitable for production-grade deployments. Therefore, future implementations will incorporate formally stronger consensus protocols such as crash fault tolerance using leader election (RAFT) or practical Byzantine fault tolerance (PBFT) in order to enhance resilience, security, and scalability across multi-site substation infrastructures.

To enable large-scale smart grid infrastructures, additional architectural improvements are suggested, such as:

- Transaction batching to reduce per-operation overhead.
- Chaincode optimization to accelerate endorsement and commit processes.
- Sharded blockchain instances to enable simultaneous policy verifications across autonomous substations.

These enhancements, although not achieved in the current work, would enable broader deployment of the proposed framework across multi-site substations that are linked by wide-area networks (WANs). The modular Informer-PPO architecture combined with federated blockchain enforcement provides a promising avenue toward scalable and tamper-evident anomaly mitigation across geographically distributed cyber-physical infrastructures.

Beyond handling blockchain transactions, the broader scalability of the proposed security framework for substations with thousands of devices is achieved through architectural modularity and hierarchical distribution. Specifically, Informer-PPO agents can be instantiated at multiple levels, where local-level agents process anomaly detection and make preliminary mitigation decisions for subsets of IEDs and IoT sensors. Only high-severity or indeterminate cases are forwarded to higher-level decision layers to reduce system-wide computational and communication overhead. Moreover, edge-level AIoT prepro-

TABLE VII: Quantitative Comparison of Detection Accuracy and Mitigation Capability in Related Works

Reference	Detection Accuracy	Mitigation Support	Avg. Response Latency
Duman et al. (2019) [11]	No real detection model	✗	Not Reported
Duman et al. (2024) [12]	No detection model (hardening only)	✗	Not Reported
Ustun et al. (2021) [13]	95.1% (SVM, GOOSE)	✗	Not Reported
Mohamed et al. (2023) [22]	N/A (Attacker Simulation)	✗	N/A
Nhung-Nguyen et al. (2024) [14]	**98%** (GOOSE/DNN)	✗	Not Reported
Park et al. (2024)[18]	ML ADS (SV+GOOSE), attack/fault classification	✓	Not Reported
Zaboli & Hong (2025) [30]	>**98%** (GenAI ADS, synthetic GOOSE datasets)	✗	Not Reported
Proposed Work	**98.4%** (Kitsune), **97.6%** (ERENO)	✓	**35 ms**

cessing enables real-time anomaly filtering, thereby allowing core RL decision engines to remain scalable even under high device densities. Based on observed inference latencies and typical data generation rates, preliminary extrapolation suggests that the framework can realistically support over 5,000–10,000 device deployments with real-time mitigation capacity, given moderate event rates and decentralized control optimization. Future work will validate these estimates in large-scale experimental testbeds.

F. Temporal Modeling Comparison

Modeling temporal dependencies is also critical in cyber-physical threat scenarios where attacks manifest as causally related sequences of events rather than isolated anomalies. To meet this challenge, the Informer-PPO model presented here leverages a self-attention mechanism that prefers temporal relevance over positional locality.

Classical convolutional encoder-based PPO implementations (CNN-PPO) are limited by the lack of ability to recognize cross-timestep dependencies beyond pre-defined receptive fields. Recurrent variants like LSTM-PPO, while sequence-aware, are usually plagued by gradient instability and memory issues in long or irregular event sequences. Transformer-based PPO models fix some of these limitations by applying global attention across sequences, yet they suffer from high computational complexity ($O(L^2)$) and attention diffusion when handling extremely long or noisy sequences.

Conversely, the Informer encoder employs ProbSparse attention to enable the agent to selectively hear critical temporal anchors throughout the observation period—regardless of sequence length. This boosts the agent's ability to coordinate its mitigation maneuvers with delayed system impacts and cross-episode interactions, which are frequently found in coordinated or stealthy cyberattacks.

This architecture enhances the temporal generalization capability of the policy network, allowing it to capture compact yet context-aware representations of dynamic system states. The learned policies provide increased predictive stability and more temporally coherent decision-making, making Informer-PPO extremely well-suited for anomaly-based control systems of smart substations.

G. Comparative Analysis with Related Works

Table VII summarizes recent works on IEC 61850 substation security, anomaly detection, and adaptive defense.

Earlier studies by Duman et al. [11], [12] focused on attack graph analysis and security hardening, but without deployable anomaly detection or mitigation. Ustun et al. [13] applied SVM for GOOSE traffic analysis, achieving moderate accuracy but lacking mitigation support and latency evaluation. Mohamed et al. [22] concentrated on attacker simulations rather than real intrusion detection.

More recent efforts improved detection but remain limited in enforcement. Nhung-Nguyen et al. [14] demonstrated deep neural network-based detection on GOOSE traffic, but without mitigation or latency results. Park et al. [18] combined anomaly detection with attack–fault classification across SV and GOOSE traffic and proposed partial restoration, though latency metrics were not reported. Zaboli & Hong [30] introduced a Generative AI-based anomaly detection system that achieved high accuracy on synthetic GOOSE datasets, but without blockchain integration or measured response times.

In contrast, the proposed Informer–PPO framework achieves high accuracy on both public Kitsune and ERENO datasets, integrates real-time adaptive mitigation, and uniquely provides blockchain-based smart contract validation. The framework maintains an average response latency of 35 ms, positioning our work as the only end-to-end defense framework addressing both technical performance and practical deployability for IEC 61850 substations.

H. Security and Privacy Analysis

The security of the provided framework relies on two collaborating layers: (1) the robustness of the learned RL policies to detect and respond to dynamic attacks and (2) the guarantee of integrity provided by the blockchain-based enforcement mechanism. The security and privacy of the proposed framework are examined across four dimensions: integrity of mitigation execution, resilience to adversarial evasion, robustness against zero-day and synthetic adversarial attacks, and protection of sensitive data during inference and enforcement.

1) Integrity and Tamper Resistance: All mitigation actions are represented as digitally signed transactions and executed through smart contracts on a permissioned blockchain network. This architecture ensures that when a policy action is initiated, it cannot be modified or replayed without validation agreement. Distributed endorsement and immutable block commitment prevent rollback attacks and unauthorized overrides of security

TABLE VIII: Security Properties and Their Implementation in the Proposed Framework

Security Property	How It's Achieved
Integrity	Blockchain immutability via peer endorsement
Non-repudiation	Signed action logs via smart contracts
Resilience to evasion	RL-based dynamic decision policies
Zero-day robustness	Synthetic adversarial testing and cross-domain generalization (Informer-PPO)
Tamper resistance	Verified mitigation through chain consensus
Privacy protection	Anonymized payloads, permissioned chain access

responses, satisfying the minimum requirements for non-repudiation and auditability.

2) Policy Spoofing and Evasion Resilience: As policy actions are learned along state–action trajectories over time and not from static rule-matching, the RL agent is inherently more robust to evasion attacks by attackers. In contrast to conventional threshold-based systems that may be probed by attackers, Informer-PPO's stochastic exploration and long-term reward modeling complicate reverse-engineering of the decision boundary. Further, the blockchain backend guarantees that even if an agent is tricked into taking a wrong step, that step gets recorded, traced, and audited for forensic examination.
3) Privacy Concerns: The system is based upon abstracted attributes that are derived from protocol events, such as GOOSE/SV metadata and action logs and does not require access to user-level or personal identifiers. Transaction payloads sent to the blockchain are small and anonymized at the device layer such that privacy-sensitive information is never revealed or stored in an unsecured state.
4) Robustness to Zero-Day and Synthetic Adversarial Attacks: The proposed framework demonstrates strong resilience against zero-day threats and synthetic adversarial attacks. Through targeted adversarial testing using novel attack scenarios such as spoofed MMS control messages, unauthorized IED registration, and firmware injection, the system consistently achieved high detection and mitigation effectiveness, even without prior exposure to these threat patterns. Moreover, the adaptive learning mechanisms of the Informer-PPO agent enabled successful generalization across heterogeneous environments, significantly reducing FPR and maintaining reliable anomaly detection performance in unfamiliar operational domains. These capabilities confirm the framework's ability to defend against emerging and previously unseen cyber threats in critical infrastructure environments.

A summary of the key security properties assured by the proposed framework and their corresponding implementation mechanisms is provided in Table VIII. All logged events are securely accessible only to permissioned nodes within the consortium network, ensuring that sensitive information remains protected from unauthorized access.

I. Experimental Deployment Considerations

Despite validating the proposed framework through simulation experiments with laboratory-controlled testing, the actual implementation of the AIoT-Blockchain security system in IEC 61850 substations will require specific hardware and networking environments. Industrial-strength edge-computing nodes colocated alongside substations are proposed to implement the Informer-PPO model inference engine as well as a blockchain client. Typical hardware specifications would be at least 16 CPU cores and 64 GB RAM, with the possibility of GPU acceleration for real-time anomaly detection and action generation. At least 1 Gbps LAN Ethernet throughput with internal latency below 5 ms is necessary to support seamless telemetry ingestion and security enforcement. Blockchain transactions for non-time-critical security operations, such as firmware verification and device quarantine, are asynchronously propagated to avoid interference with time-critical protection processes. In large-scale smart grid deployments, secure WAN links and backup blockchain nodes are advised to ensure resilience, high availability, and fault tolerance between substations.

J. Real-World Deployment Considerations

To enable non-disruptive deployment in live IEC 61850 substations, the evolved framework can first be deployed in passive monitoring mode. Using this mode, AIoT sensors and an Informer-PPO-based anomaly detection system would run concurrently with existing operational SCADA systems without interfering with the operational controls. Identified anomalies would be logged for offline analysis to determine the model's reliability. Upon validation, the framework can proceed to an advisory state in which the mitigation steps are reviewed first by human operators prior to enforcement. Finally, upon demonstrated consistency in performance and operational trust, autonomous blockchain-enforced mitigation can be deployed progressively for non-time-critical security operations. This phased integration strategy ensures operational safety, prevents deployment risks, and facilitates the implementation of the system in substation environments without major architectural modifications.

VI. Conclusion

This paper presented an AIoT–Blockchain security framework that integrates Informer-augmented PPO for adaptive cyber defense and blockchain-based authentication within IEC 61850 substations. The proposed framework achieved

a detection accuracy of 98.4% and an average response latency of 35 ms, representing accuracy gains of 3.4-9.2% and response time reductions of 33-63% compared with Transformer-PPO, LSTM-PPO, and CNN-PPO baselines. Blockchain-based enforcement added only 42-50 ms for non-time-critical actions such as firmware rollback and device quarantine, ensuring tamper-proof policy execution without compromising operational feasibility. These findings demonstrate that real-time, auditable, and scalable cybersecurity is attainable within substation constraints. Nonetheless, blockchain latency constrains its application in primary protection, scalability under high-frequency updates remains a challenge, and resilience against adaptive adversarial attacks warrants further investigation. The modular design of the framework allows its extension to broader IIoT ecosystems, including microgrids, distributed energy resources, and smart cities.

ACKNOWLEDGMENT

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) (Project No. RS-2023-00228996, 30%; RS-2024-00438551, 20%, IITP-2025-RS-2021-II211816, 10%), the Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2025 (Project Name: Training Global Talent for Copyright Protection and Management of On-Device AI Models, Project Number: RS-2025-02221620, Contribution Rate: 20%), and the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Project No. RS-2023-00208460, 20%).

REFERENCES

[1] L. Tightiz, H. Yang, "A comprehensive review on IoT protocols' features in smart grid communication," *Energies*, vol. 13, no.11, pp.2762, 2020.

[2] B. Paul, A. Sarker, SH. Abhi, SK. Das, MF. Ali, MM. Islam, MR. Islam, SI. Moyeen, MF. Badal, MH. Ahamed, SK. Sarker, " Potential smart grid vulnerabilities to cyber attacks: Current threats and existing mitigation strategies," *Heliyon*, vol. 10, no. 19, 2024.

[3] M. Beikbabaei, A. Mehrizi-Sani, CC. Liu, "State-of-the-art of cybersecurity in the power system: Simulation, detection, mitigation, and research gaps," *IET Generation, Transmission & Distribution*, vol. 19, no. 1, e70006, 2025.

[4] N. Tatipatri and S. L. Arun, "A Comprehensive Review on Cyber-Attacks in Power Systems: Impact Analysis, Detection, and Cyber Security," *IEEE Access*, vol. 12, pp. 18147–18167, 2024.

[5] R. Chinnasamy, M. Subramanian, SV. Easwaramoorthy, J. Cho , "Deep Learning-driven Methods for Network-based Intrusion Detection Systems: A Systematic Review," *ICT Express*, 2025.

[6] T. T. Nguyen and V. J. Reddi, "Deep Reinforcement Learning for Cyber Security," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 8, pp. 3779–3795, 2023.

[7] N. E. Fard, R. R. Selmic and K. Khorasani, "A Review of Techniques and Policies on Cybersecurity Using Artificial Intelligence and Reinforcement Learning Algorithms," *IEEE Technology and Society Magazine*, vol. 42, no. 3, pp. 57–68, 2023.

[8] K. Park, M. Girdhar, J. Hong, W. Su, A. Herath, CC. Liu, " Machine Learning Based Cyber System Restoration for IEC 61850 Based Digital Substations," *arXiv preprint arXiv:2411.07419*, 2024 Nov 11.

[9] M. Girdhar, J. Hong, W. Su, A. Herath and C. -C. Liu, " SDN-Based Dynamic Cybersecurity Framework of IEC-61850 Communications in Smart Grid," *IEEE Power & Energy Society General Meeting (PESGM)*, Seattle, WA, USA, pp. 1-5, 2024.

[10] S. M. S. Hussain, M. A. Aftab, S. M. Farooq, I. Ali, T. S. Ustun and C. Konstantinou, "An Effective Security Scheme for Attacks on Sample Value Messages in IEC 61850 Automated Substations," *IEEE Open Access Journal of Power and Energy*, vol. 10, pp. 304–315, 2023.

[11] O. Duman, M. Ghafouri, M. Kassouf, R. Atallah, L. Wang and M. Debbabi, " Modeling Supply Chain Attacks in IEC 61850 Substations," *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Beijing, China, 2019, pp. 1–6.

[12] O. Duman, A. Tabiban, L. Wang and M. Debbabi, " Measuring and Improving the Security Posture of IEC 61850 Substations Against Supply Chain Attacks," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–20, 2024.

[13] TS. Ustun, SS. Hussain SS, A. Ulutas, A. Onen, MM. Roomi, D. Mashima, " Machine learning-based intrusion detection for achieving cybersecurity in smart grids using IEC 61850 GOOSE messages," *Symmetry*, vol. 13, no. 5), pp.826, 2021.

[14] H. Nhung-Nguyen, M. Girdhar, YH. Kim, J. Hong, " Machine-learning-based anomaly detection for GOOSE in digital substations," *Energies*, 2024, vol. 17, no. 15, pp. 3745.

[15] Z. Lian, P. Shi, and M. Chen, "A Survey on Cyber-Attacks for Cyber-Physical Systems: Modeling, Defense, and Design," *IEEE Internet of Things Journal*, vol. 12, no. 2, pp. 1471–1483, 2025.

[16] U. Ghosh, L. Njilla, S. Shetty and C. A. Kamhoua, "A Decentralized Smart Grid Communication Framework Using SDN-Enabled Blockchain," *IEEE 21st Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2024, pp. 982–985.

[17] M. Gayo, C. Santos, F. J. R. Sánchez, P. Martin, J. A. Jiménez and M. Tradacete, "Addressing Challenges in Prosumer-Based Microgrids With Blockchain and an IEC 61850-Based Communication Scheme," *IEEE Access*, vol. 8, pp. 201806–201822, 2020.

[18] K. Park, M. Girdhar, J. Hong, W. Su, A. Herath, CC. Liu, "Machine Learning Based Cyber System Restoration for IEC 61850 Based Digital Substations," *arXiv preprint arXiv:2411.07419.* 2024 Nov 11, https://doi.org/10.48550/arXiv.2411.07419

[19] A. Lakhan, Z. A. A. Alyasseri, M. A. Mohammed, B. AL-Attar, J. Nedoma, R. Alubady, S. Memon, and R. Martinek, "Sustainable Secure Blockchain Assisted AIoT and Green Multi-Constraints Supply Chain System," *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2025.3548037.

[20] X. Fu, H. Wang and Z. Wang, "Research on Block-Chain-Based Intelligent Transaction and Collaborative Scheduling Strategies for Large Grid," *IEEE Access*, vol. 8, pp. 151866–151877, 2020.

[21] L. Tightiz L, R. Nasimov, MA. Nasab, " Implementing AI Solutions for Advanced Cyber-Attack Detection in Smart Grid," *International Journal of Energy Research*, vol. 2024, no. 1, pp. 6969383, 2024.

[22] AS. Mohamed, S. Lee, D. Kundur, "Reinforcement Learning for Supply Chain Attacks Against Frequency and Voltage Control," *IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 369–375, 2023.

[23] M. A. Umar and K. Shuaib, "Cyber-Attack Detection in Smart Grids: A Comparative Analysis of Supervised and Semi-Supervised Methods," *2024 6th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, Alkhobar, Saudi Arabia, 2024.

[24] D. Said, M. Bagaa, A. Oukaira, A. Lakhssassi, " Quantum entropy and reinforcement learning for distributed denial of service attack detection in smart grid," *IEEE Access*, vol. 12, pp. 129858-129869, 2024.

[25] L. Tightiz, J. Yoo, " A novel deep reinforcement learning based business model arrangement for Korean net-zero residential micro-grid considering whole stakeholders' interests," *ISA transactions* vol. 137, pp. 471-91, 2023.

[26] M. Shateri, F. Messina, P. Piantanida and F. Labeau, "Privacy-Cost Management in Smart Meters With Mutual-Information-Based Reinforcement Learning," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22389–22398, 2022.

[27] T. Cai et al., "TEMP: Cost-Aware Two-Stage Energy Management for Electrical Vehicles Empowered by Blockchain," *IEEE Internet of Things Journal*, vol. 11, no. 23, pp. 38246–38261, 2024.

[28] SH. Oh,J. Kim , JH. Nah, J. Park, "Employing deep reinforcement learning to Cyber-Attack simulation for enhancing cybersecurity, " *Electronics*, vol .13, no. 3, pp. 555, 2024.

[29] J. Jin and Y. Xu, " Optimal Policy Characterization Enhanced Proximal Policy Optimization for Multitask Scheduling in Cloud Computing," *IEEE Internet of Things Journal*, vol. 9, no. 9, pp. 6418–6433, 2022.

[30] A. Zaboli, J. Hong, " Generative AI for Critical Infrastructure in Smart Grids: A Unified Framework for Synthetic Data Generation and Anomaly Detection," *arXiv preprint arXiv:2508.08593.* 2025 Aug 12, https://doi.org/10.48550/arXiv.2508.08593

[31] Y. Sun, L. Hou, Z. Lv and D. Peng, "Informer-Based Intrusion Detection Method for Network Attack of Integrated Energy System,"

IEEE Journal of Radio Frequency Identification, vol. 6, pp. 748–752, 2022.

[32] Y. Shi, X. Xiao, Q. -L. Han, J. Jin, S. Wen and Y. Xiang, "InforTest: Informer-Based Testing for Applications in the Internet of Robotic Things," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 2, pp. 1499–1507, 2025.

[33] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, W. Zhang, "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol.35, no. 12, pp. 11106–11115, 2021.

[34] P. P. Biswas, H. C. Tan, Q. Zhu, Y. Li, D. Mashima and B. Chen, "A Synthesized Dataset for Cybersecurity Study of IEC 61850 based Substation," *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Beijing, China, 2019, pp. 1–7.

[35] sequincozes. ERENO IEC61850 IDS Dataset. Kaggle, https://www.kaggle.com/datasets/sequincozes/ereno-iec61850-ids, Accessed April 23, 2025.

[36] Yisroel Mirsky, Tomer Doitshman, Yuval Elovici, and Asaf Shabtai. Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection. *The Network and Distributed System Security Symposium (NDSS) 2018*, 2018.

Lilia Tightiz received her Ph.D. in Computer Science and Engineering from Sejong University, Korea, in 2022. With over 15 years of experience in the electric power distribution industry, she has worked as a Power Distribution Engineer specializing in the design, optimization, and maintenance of electricity distribution grids. Her contributions have earned her multiple patents and international awards, including honors from iENA in Nuremberg, the Korean International Women's Invention Exposition, and the Bitgaram International Exposition of Electric Power Technology. She began her academic career as an Assistant Professor in the Department of AI Software at Gachon University and, in 2025, joined the Department of Computer Engineering at Sejong University. Dr. Tightiz has served as a reviewer for IEEE, Springer, and Elsevier journals and as an Associate Editor for e-Prime – Advances in Electrical Engineering, Electronics, and Energy (Elsevier) and Smart Energy Systems (Scilight Press). Her research interests include microgrid energy management, deep reinforcement learning, quantum machine learning in power systems, smart grid communication, IEC 61850, and blockchain technologies.

L. Minh Dang received the BEng degree in information systems from University of Information Technology, VNU HCMC, Vietnam in 2016, and the PhD degree in computer science from Sejong University, Seoul, Republic of Korea in 2021. Starting from 2017, he joined CVPR Lab Sejong University, Republic of Korea and currently a research professor there. His current research interests include computer vision, natural language processing, and artificial intelligence.

Ki-Woong Park received the B.S. degree in computer science from Yonsei University, South Korea, in 2005, the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 2007, and the Ph.D. degree in electrical engineering from KAIST in 2012. He received a 2009–2010 Microsoft Graduate Research Fellowship. He worked as a senior researcher at the National Security Research Institute. He is currently a Professor with the Department of Computer and Information Security at Sejong University, Seoul, South Korea. His current research interests include cloud computing security, zero-trust architecture, AI-driven cyber defense automation, and system-level vulnerability analysis and mitigation in real computing environments.