International Journal of Basic and Applied Sciences, 14 (6) (2025) 554-569

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/m4mkbn36
Research paper

Zero-Day Exploits in Healthcare IoT Networks

Hani Al-Balasmeh

Dept. of Informatics Engineering, College of Engineering
University of Technology Bahrain (UTB)
**Corresponding author E-mail: h.albalasmeh@utb.edu.bh*

Received: September 15, 2025, Accepted: October 18, 2025, Published: October 27, 2025

Abstract

The rapid adoption of Healthcare Internet of Things (HIoT) systems has enhanced patient care. Still, it has also exposed hospitals to sophisticated zero-day exploits that can bypass traditional intrusion detection systems. These threats pose a significant risk to patient safety and compromise clinical operations, necessitating detection mechanisms that are both accurate and efficient, and aligned with the specific requirements of healthcare.

This paper introduces a multi-layer security framework that integrates hybrid anomaly detection, Extreme Value Theory (EVT)-based thresholding, blockchain-assisted forensic logging, and severity-aware mitigation. The detection ensemble combines autoencoder reconstruction, LSTM-based predictive modeling, and statistical distribution monitoring, while EVT establishes statistically principled thresholds to reduce false alarms. Blockchain ensures tamper-proof accountability, and mitigation actions are prioritized based on device criticality and clinical risk.

Extensive evaluation using CIC-IDS2017, TON_IoT, and N-BaIoT datasets demonstrates the effectiveness of the framework. The system achieved a detection rate of 96.4%, a false positive rate of 1.1%, and an average latency of 6.8 ms, outperforming baseline solutions including Snort, Suricata, Kitsune, and LSTM-based IDS. Additional experiments confirmed scalability under large device populations, resilience against adversarial evasion, and negligible blockchain overhead.

By combining statistical rigor with clinical feasibility, the proposed framework provides a trustworthy and deployable solution for zero-day exploit defense in HIoT, advancing the development of secure and resilient innovative healthcare ecosystems.

Keywords: *Healthcare IoT (HIoT), Zero-Day Exploits, Anomaly Detection, Extreme Value Theory (EVT), Blockchain Security, Intrusion Detection Systems (IDS).*

1. Introduction

The convergence of the Internet of Things (IoT) and healthcare has transformed clinical environments by enabling continuous patient monitoring, remote diagnostics, and data-driven medical decision-making. Collectively referred to as the Healthcare Internet of Things (HIoT), this paradigm encompasses wearable biosensors, implantable devices, infusion pumps, and network-enabled imaging systems that interconnect with hospital information infrastructures [1]. By facilitating seamless communication between patients, caregivers, and health information systems, HIoT technologies promise improved efficiency, reduced operational costs, and better clinical outcomes [2]. However, the rapid expansion of interconnected devices has also introduced an unprecedented cyber-attack surface, making healthcare infrastructures increasingly vulnerable to exploitation.

Among the most severe threats are zero-day exploits, which target undiscovered vulnerabilities in device firmware, middleware, or communication protocols before they are publicly disclosed or patched [3]. Unlike conventional attacks, zero-day exploits evade signature-based defenses, allowing adversaries to infiltrate networks undetected. In healthcare contexts, this can result in unauthorized access to sensitive medical data, manipulation of device functionality, or even life-threatening disruptions to patient care [4]. Beyond technical consequences, such incidents undermine public trust in digital healthcare systems and pose serious ethical, legal, and regulatory challenges [5].

Securing healthcare IoT against zero-day exploits is complicated by several unique constraints. First, medical devices often operate with resource limitations (restricted processing power, memory, and battery capacity), which preclude the deployment of heavy cryptographic or machine-learning models [6]. Second, healthcare infrastructures rely on legacy systems that remain in operation for decades without consistent vendor updates or security patches [7]. Third, healthcare data is both highly sensitive and financially lucrative, ranking among the most valuable assets on the black market [8]. This combination—critical services, vulnerable infrastructures, and high-value data—creates a perfect storm for adversaries, including cybercriminals and state-sponsored actors.

Recent scholarship has explored countermeasures, including anomaly-based intrusion detection, machine learning-driven behavior analysis, and blockchain-based integrity verification [9], [10]. While promising, these methods face persistent limitations, including high false-positive rates, computational inefficiency, limited adaptability to zero-day scenarios, and the absence of healthcare-specific security

frameworks [11]. Furthermore, most approaches emphasize detection alone, with little attention to forensic accountability or adaptive mitigation strategies that are vital in life-critical contexts.

This paper addresses these gaps by proposing a multi-layer defense framework for zero-day exploits in HIoT networks. The framework integrates (i) a hybrid anomaly detection ensemble combining reconstruction, predictive, and statistical models; (ii) Extreme Value Theory (EVT)-based thresholding to establish statistically rigorous detection boundaries; (iii) blockchain-assisted logging for tamper-proof forensic accountability; and (iv) severity-aware mitigation strategies to ensure proportionate and clinically aligned responses. Unlike prior work, this approach emphasizes not only accuracy and scalability but also resilience, accountability, and clinical feasibility.

The remainder of this paper is organized as follows. Section 2 presents the proposed system architecture, while Section 3 details the methodology, including mathematical formulations and detection algorithms. Section 4 reports experimental results and comparative analyses against baseline IDS solutions. Section 5 concludes with a discussion of key contributions, limitations, and directions for future work.

2. Literature Review

The security of Healthcare Internet of Things (HIoT) networks has become an active area of research, driven by the increasing digitization of clinical services and the parallel escalation of cyber threats. Among these, zero-day exploits are particularly concerning, as they exploit previously undisclosed vulnerabilities and evade traditional defenses [1]. The following review synthesizes prior scholarship on zero-day attacks, healthcare-specific vulnerabilities, and existing detection strategies, before positioning the contributions of this work.

2.1 Zero-Day Exploits in Cybersecurity

Zero-day exploits are defined as attacks launched before defenders become aware of the vulnerability or before a patch is available [2]. The commodification of zero-day vulnerabilities in underground markets accelerates their proliferation across industries, including healthcare [3]. Fueled by a thriving cybercrime marketplace that actively trades exploits across forums and dark-web markets [20]. Conventional intrusion detection systems such as Snort and Suricata, which rely on known signatures, have repeatedly failed against such exploits [4]. Razzaq et al. [5] emphasize that adversaries can maintain dwell times of months in compromised systems, thereby amplifying risks. More recently, researchers have highlighted the integration of zero-day exploits into advanced persistent threats (APTs) targeting critical infrastructures, thereby raising the stakes for sectors such as healthcare [6].

2.2 Vulnerabilities in Healthcare IoT Networks

HIoT systems integrate wearable sensors, implantable devices, monitoring platforms, and hospital automation systems [7]. Despite their clinical benefits, they are highly vulnerable due to:

- Device heterogeneity, leading to inconsistent security baselines [8].
- Resource-constrained sensors, which cannot accommodate heavy security functions [9].
- Legacy equipment, which often lacks vendor patches [10]; and
- High-value patient data, attractive to both cybercriminals and state actors [11].

Documented incidents confirm the severity of these risks, including ransomware outbreaks that locked hospitals out of mission-critical systems [12] and vulnerabilities in infusion pumps or pacemakers that could be manipulated remotely [13]. These examples underline that in healthcare, zero-day exploits directly endanger patient safety—not just data confidentiality.

2.3 Detection and Mitigation Approaches

Signature-based intrusion detection systems (IDS), such as Snort and Suricata, remain widely deployed due to their simplicity and effectiveness against known attacks. However, their reliance on predefined patterns renders them inherently ineffective against zero-day exploits, where no prior signatures exist [14]. Although hybrid systems attempt to combine signatures with heuristic matching, their adaptability in dynamic healthcare environments is limited.

Anomaly-based and machine learning (ML)-driven methods represent the most prominent alternatives. By modeling standard traffic patterns, these systems can detect deviations indicative of novel attacks. Wang et al. [15] demonstrated that LSTM-based models achieved an accuracy of over 94% on IoT datasets, but at the cost of high false positive rates, which are operationally unacceptable in hospitals. Similarly, Li et al. [16] applied deep anomaly detection in HIoT settings, demonstrating improved sensitivity but also revealing the computational burden that such models impose on resource-constrained medical devices.

Blockchain-assisted frameworks have gained traction for providing immutable logs and decentralized trust. Xu et al. [17] introduced blockchain into IoT security architectures, enhancing resilience by ensuring tamper-proof forensic evidence. Yet, blockchain introduces latency and scalability challenges that limit adoption in real-time healthcare scenarios, where delays may compromise patient outcomes.

Federated learning (FL) has recently emerged as a solution to address the privacy concerns associated with centralizing sensitive health data. Ahmad et al. [18] developed FL-based intrusion detection for medical IoT, reducing false positives while maintaining data confidentiality. Despite these gains, FL requires secure aggregation protocols and robust device hardware, which are not universally available across heterogeneous hospital networks.

Finally, hybrid and multi-layered defenses have been proposed to combine complementary techniques. Sengupta et al. [19] argued that layered defenses integrating anomaly detection, blockchain, and adaptive policies provide greater resilience. However, most such systems remain conceptual or lab-scale prototypes, with few reports of real-world deployment in clinical environments due to cost, interoperability issues, and a lack of healthcare-specific standards.

Taken together, these studies highlight that while individual techniques provide partial solutions, none achieve a balance of accuracy, scalability, accountability, and clinical feasibility required for healthcare IoT. This shortcoming motivates the present work, which proposes a multi-layer architecture uniting anomaly detection, EVT-based thresholding, blockchain-assisted logging, and severity-aware mitigation, thereby addressing the dual imperatives of cybersecurity robustness and patient safety.

2.4 Comparative Analysis

Table 1 synthesizes representative studies on zero-day detection in IoT and healthcare contexts. It compares approaches, reported strengths, and observed limitations, providing context for the contributions of this paper.

Table 1: Comparative Overview of Zero-Day Detection in IoT and Healthcare Networks

Study	Domain	Approach	Strengths	Limitations
Alazab et al. [1]	Cybersecurity (general)	Survey of detection/prevention methods	Comprehensive mapping of strategies	Lacked healthcare-specific focus
Wang Et Al. [15]	IoT networks	LSTM-based anomaly detection	High detection accuracy (~94%)	High false positives
Li Et Al. [16]	Healthcare IoT	Deep anomaly detection	Improved detection sensitivity	Computationally expensive
Xu Et Al. [17]	IoT systems	Blockchain-assisted detection	Immutable, tamper-proof logging	Latency and scalability overhead
Ahmad Et Al. [18]	Medical IoT	Federated learning IDS	Preserves privacy, reduces false positives	Requires powerful hardware
Sengupta Et Al. [19]	IoT security	Hybrid ML + blockchain	Higher resilience, layered defense	Limited real-world deployment

As shown in Table 1, progress has been made in anomaly detection and blockchain-based accountability, yet each approach faces critical shortcomings. ML/DL methods achieve high accuracy but often overwhelm IT staff with false positives; blockchain ensures integrity but struggles with latency in real-time clinical contexts; federated learning preserves privacy but demands advanced device resources; and hybrid frameworks remain mostly conceptual, with few hospital-scale deployments.

2.5 Critical Perspective

The literature clearly indicates that no single approach is sufficient to counter zero-day exploits in HIoT. Existing works emphasize either detection accuracy, privacy preservation, or accountability, but rarely integrate these requirements into a holistic framework. Moreover, few studies explicitly align cybersecurity mechanisms with clinical workflows, where false alarms, latency, and device isolation can directly impact patient outcomes.

This paper builds upon these insights by introducing a multi-layered defense architecture that unites anomaly detection, EVT-based statistical rigor, blockchain-assisted verification, and severity-aware mitigation. In doing so, it directly addresses the dual challenge of technical robustness and clinical feasibility, offering a deployable pathway toward securing smart healthcare infrastructures.

3. Proposed Framework and Architecture

3.1 Rationale for the Framework

The proposed framework for defending Healthcare IoT (HIoT) networks against zero-day exploits is structured as a multi-layer defense system that unites data acquisition, edge-assisted monitoring, hybrid anomaly detection, blockchain-enabled accountability, and severity-aware mitigation. Unlike conventional intrusion detection systems (IDS) that typically prioritize detection accuracy at the expense of scalability or interpretability, this architecture explicitly addresses the unique constraints of HIoT environments—low-power medical devices, latency-sensitive clinical services, and strict regulatory requirements.

3.2 Device Layer

The device layer forms the entry point of the architecture, comprising IoT-enabled medical devices, including wearable biosensors, infusion pumps, implantable monitors, and imaging systems. These devices generate continuous data streams including vital signs, medication dosages, infusion states, and device health metrics. Given their resource constraints—limited CPU, memory, and energy—these devices cannot host advanced intrusion detection or cryptographic modules locally [1], [2]. This makes them attractive targets for adversaries exploiting unknown vulnerabilities, particularly at the firmware and middleware levels.

As depicted in Figure 1, devices in this layer act exclusively as data producers, forwarding unprocessed traffic to higher layers. This separation of roles minimizes computational overhead on life-critical devices while maintaining their integration into the broader detection ecosystem. By insulating devices from direct IDS tasks, the framework reduces operational risk, ensuring uninterrupted patient care even in the presence of security incidents.

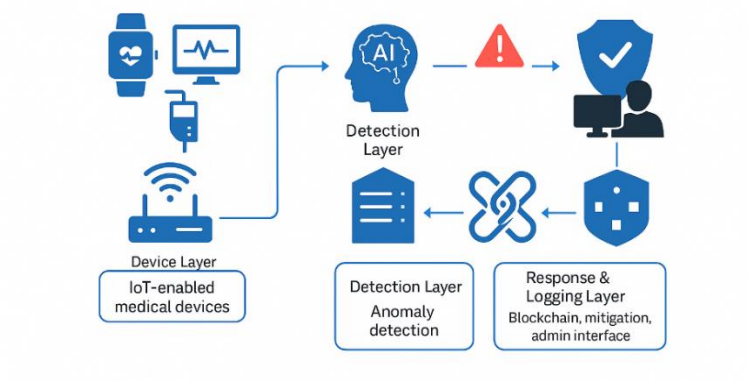

Fig 1: Proposed Zero-Day Framework Architecture

3.3 Network Monitoring Layer

The network monitoring layer aggregates traffic from medical devices through hospital gateways, routers, and edge servers. Its functions include packet inspection, flow summarization, and lightweight feature extraction, which transform raw data into structured features for anomaly analysis. To meet the real-time requirements of clinical workflows, the framework leverages edge computing to preprocess traffic locally, thereby reducing latency and conserving hospital network bandwidth [3], which aligns with recent IoMT security designs that integrate 5G-edge orchestration for scalability [22].

As shown in Figure 2, the monitoring workflow encompasses data preprocessing and feature extraction. Key features are derived using autoencoder-based (AE) reconstruction, Long Short-Term Memory (LSTM) predictive modeling, and Mahalanobis-distance statistical analysis. This layered feature set captures traffic-level anomalies, temporal deviations, and distributional outliers. By embedding preprocessing at the edge, the architecture ensures that exploit indicators—such as unusual protocol behavior, unauthorized traffic patterns, or abnormal packet bursts—are preserved and forwarded for advanced detection.

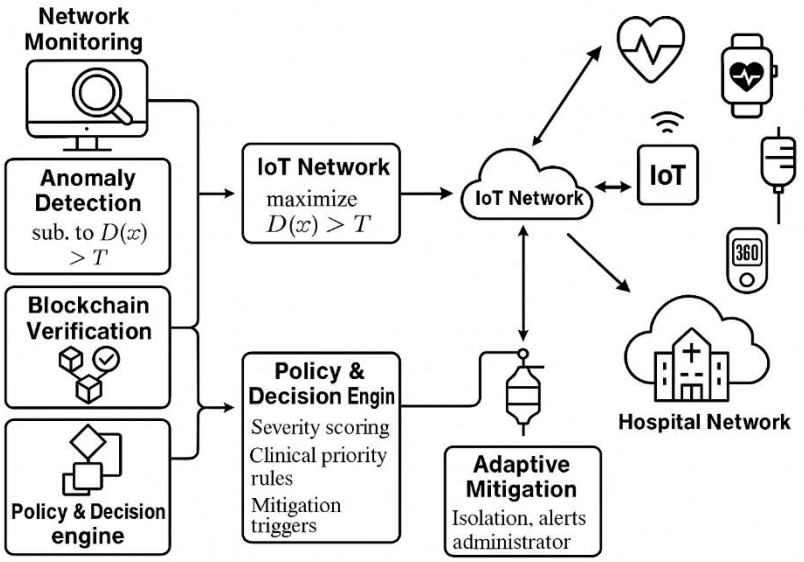

Fig 2: Network Monitoring Workflow for Zero-Day Detection in Healthcare IoT Networks

3.4 Detection Layer

The detection layer represents the analytical core of the framework. It integrates a hybrid ensemble of complementary models:

- Autoencoder-based reconstruction identifies deviations in learned traffic feature patterns.
- LSTM-based temporal prediction captures sequential irregularities in device activity logs.
- Mahalanobis-distance modeling detects distributional outliers across multivariate traffic profiles.

The novelty of this layer lies in its integration of Extreme Value Theory (EVT) for anomaly threshold calibration [4]. Instead of relying on static or heuristically tuned thresholds, EVT models the tail distribution of anomaly scores, producing statistically bounded thresholds that guarantee controlled false favorable rates (ε). This ensures that alerts remain operationally manageable in clinical settings, avoiding the alarm fatigue often reported in hospitals using conventional IDS [9], a critical issue documented in clinical studies where excessive alarms compromise trust and delay responses [21].

As highlighted in Figure 2, EVT serves as the final calibration stage in the detection pipeline, refining the outputs of machine learning models into actionable decisions. This approach simultaneously enhances sensitivity to zero-day exploits and reduces spurious alerts, enabling clinicians and IT administrators to focus on high-confidence anomalies.

3.5 Response and Logging Layer

The response and logging layer operationalizes detection outcomes by executing severity-aware mitigation and ensuring forensic accountability. Severity scoring is based on a multi-factor assessment of:

1. Anomaly magnitude (statistical severity).
2. Device criticality (e.g., ventilators vs. auxiliary sensors).
3. Network exposure (isolated device vs. integrated hospital subsystem).

High-severity events trigger immediate device isolation to prevent cascading failures, while medium- and low-severity cases may generate alerts or require administrator approval. This ensures that responses are both proportionate and clinically aligned, preserving patient safety without unnecessary disruption to hospital operations.

A distinctive feature of this layer is its use of blockchain-assisted logging [5]. As depicted in Figure 3, each detection event is hashed and immutably stored in a permissioned blockchain ledger validated through Practical Byzantine Fault Tolerance (PBFT). This ensures that forensic evidence cannot be altered after the event, providing a tamper-proof audit trail that strengthens compliance with regulatory frameworks such as HIPAA and GDPR. Beyond compliance, blockchain-based accountability fosters trust among stakeholders—clinicians, administrators, and regulators—by ensuring that all security events are transparently verifiable and auditable.

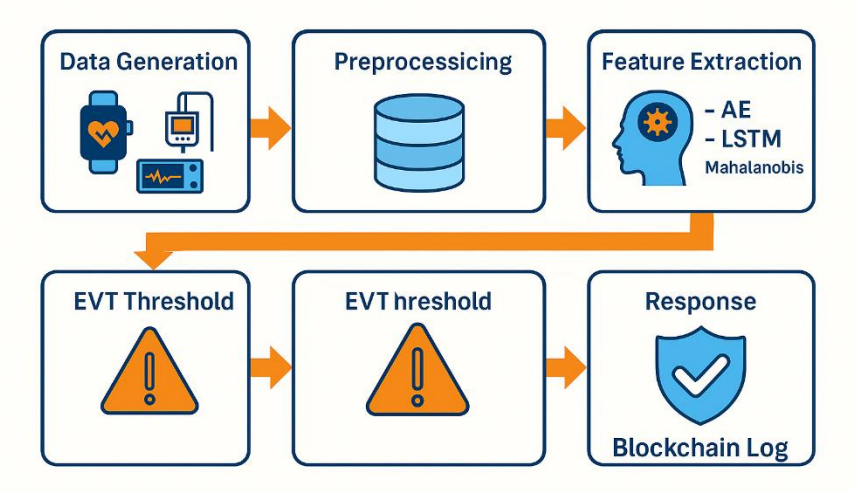

Fig 3: Workflow for Zero-Day Detection Process

3.6 Design Rationale and Novelty

The proposed framework addresses critical shortcomings in existing solutions. Signature-based IDS such as Snort and Suricata are ineffective against zero-day exploits due to their reliance on predefined patterns [3], [12]. Lightweight IDS such as Kitsune achieve good accuracy but lack forensic accountability. Blockchain-only frameworks provide integrity but introduce latency, while federated learning–based IDS preserves privacy but remains too computationally intensive for resource-constrained HIoT devices [17], [18].
In contrast, the proposed architecture offers:

- Hybrid anomaly detection that combines statistical, temporal, and reconstruction-based analysis for comprehensive zero-day visibility.
- EVT-calibrated thresholds that deliver statistically bounded false favorable rates, directly addressing alarm fatigue in healthcare IDS deployments.
- Blockchain-assisted accountability that ensures transparent, tamper-proof event logging, often absent in prior IDS research.
- Severity-aware mitigation that aligns security responses with clinical risk prioritization, a feature rarely considered in technical IDS designs.

Together, these contributions create a clinically deployable framework that not only strengthens resilience against zero-day exploits but also bridges the gap between cybersecurity research and the operational demands of healthcare environments.

4. Methodology

The detection of zero-day exploits in Healthcare IoT (HIoT) networks necessitates a methodology that is not only statistically rigorous but also computationally tractable and clinically trustworthy, given the life-critical nature of medical environments. To meet these requirements, the proposed framework integrates multiple complementary components into a unified pipeline: (i) a formalized system and adversary model that captures the layered structure of HIoT ecosystems and the adaptive strategies of attackers; (ii) a drift-aware feature space that harmonizes network traffic, device telemetry, and temporal dynamics while adapting to evolving baselines; (iii) a hybrid anomaly detection ensemble that combines reconstruction-based (autoencoders), predictive (LSTM sequence models), and distributional (Mahalanobis distance) scoring functions to capture diverse manifestations of zero-day behaviors; (iv) Extreme Value Theory (EVT)-based calibration, which provides mathematically grounded thresholds that guarantee bounded false-positive rates under finite-sample conditions [3]; (v) permissioned blockchain logging, which ensures tamper-proof, auditable forensic records validated through Practical Byzantine Fault Tolerance (PBFT) consensus [4]; and (vi) a severity-aware mitigation strategy, which prioritizes interventions according to anomaly magnitude, device criticality, and network exposure, thereby aligning defensive actions with clinical priorities and minimizing operational disruption. All design choices are explicitly constrained by the latency, reliability, and privacy requirements of modern hospital workflows, ensuring the methodology is both deployable and resilient in real-world healthcare contexts [1], [2].

4.1 System and Threat Model

Healthcare IoT (HIoT) networks can be abstracted as a three-layer cyber-physical architecture consisting of:

- the device layer (wearables, implantables, infusion pumps),
- the edge/network layer (gateways, 5G links, hospital LANs), and
- the cloud layer (electronic health record (EHR) storage, clinical analytics, and decision-support services) [1].

Formally, let each device. $d_i \in D$Emit multivariate traffic windows:

$$W_t = \{x_{t-w+1:t}^{(i)}\}, \quad x_t^{(i)} \in R^m$$

where $\boldsymbol{x_t^{(i)}}$ denotes the feature vector of dimension mmm collected at time t , and w is the sliding window size. These features represent packet-level statistics, device telemetry, and temporal usage patterns.
The threat model assumes an adaptive adversary who can exploit previously unknown vulnerabilities in device firmware or communication protocols. The attacker injects malicious subsequences M⊆T that are intentionally designed to remain indistinguishable from benign traffic when examined by traditional signature-based intrusion detection systems (IDS).
The defender's objective is to design a detection and response policy π that maps windows. W_t to actions $\mathrm{A} = \{\mathrm{monitor}, \mathrm{rate_limit}, \mathrm{isolate}\}$.
The optimization goal balances accuracy, false alarm control, and latency, expressed as:

$$\min_{\pi} Pr\,(\,\mathrm{FN} \mid \pi\,)$$
$$s.t.\ \Pr(\,\mathrm{FP} \mid \pi\,) \le \epsilon,\ \mathrm{Latency}(\pi) \le\ \tau_{max}$$

Where:

- $Pr\,(\,\mathrm{FN} \mid \pi\,)$ is the probability of a false negative (missed detection),
- $Pr\,(\,\mathrm{FP} \mid \pi\,)$ is the probability of a false positive, bounded by clinical tolerance ϵ, and
- τ_{max} denotes the maximum acceptable decision latency permitted by hospital workflows [1], [2].

This formulation captures the dual imperative of HIoT security: minimizing the risk of missed zero-day exploits while ensuring that the system does not overwhelm clinical staff with excessive false alarms, nor introduce delays that could disrupt time-sensitive medical procedures.

4.2 Feature Representation and Drift Handling

Feature representation is central to ensuring that zero-day exploit detection remains robust under dynamic Healthcare IoT (HIoT) conditions. To address distributional changes in traffic patterns, commonly referred to as concept drift, the framework employs online normalization of input features. At each time step, feature vectors are normalized using incremental statistics derived from streaming benign traffic at the edge layer. This prevents the system from becoming biased toward outdated data distributions, enabling adaptive learning in evolving environments.

Formally, windows are normalized as:

$$x' = \frac{x - \mu_t}{\sigma_t},$$

where μ_t and σ_t Represent the time-dependent mean and standard deviation, updated incrementally to capture non-stationary characteristics of clinical traffic.

To reduce redundancy and computational cost, dimensionality is compressed either through Principal Component Analysis (PCA) at the edge or through autoencoder (AE) bottleneck representations in the detection layer, yielding. $x' \in R^{m'} with m' \ll m$. This ensures that only the most informative variance is retained while minimizing resource consumption.

The final feature set spans three domains:

- Network-level features: flow durations, inter-arrival times, packet size distributions, TCP flags, and Shannon entropy.
- Device-level telemetry: CPU utilization, RAM and I/O loads, and sensor polling frequencies.
- Temporal descriptors: exponentially weighted moving averages (EWMA), seasonal indices, and autocorrelation coefficients.

This unified representation enables the framework to capture not only traffic-based anomalies but also device-level and temporal irregularities, ensuring broad coverage against diverse zero-day exploit strategies.

4.3 Hybrid Anomaly Detection Ensemble

To reliably detect zero-day exploits in Healthcare IoT (HIoT) networks, the proposed framework integrates a hybrid anomaly detection ensemble that combines reconstruction, predictive, and distributional perspectives. This design ensures that different manifestations of malicious activity—whether packet-level anomalies, sequential deviations, or statistical outliers—are captured through complementary detection mechanisms.

4.4 Reconstruction Score (Autoencoder)

Autoencoders trained on benign traffic reconstruct standard input patterns with low error, while anomalous inputs result in larger residuals. The reconstruction score is defined as:

$$S_{\text{recon}}(x) = |x - g_{\phi}(x)|_2^2$$

where g_{ϕ} is the autoencoder. A Huber loss function is employed during training to reduce sensitivity to extreme outliers, improving robustness in noisy hospital networks [6].

4.5 Predictive Score (LSTM)

Zero-day exploits often manifest as unexpected temporal deviations in device traffic. To capture these dynamics, we employ a Long Short-Term Memory (LSTM) network to forecast the next observation window. The predictive error is:

$$S_{\text{pred}} = |x_{t+1} - f_{\theta}(x_{t:t-w+1})|_2^2$$

where f_{θ} The LSTM-based predictor and www represent the lookback horizon. Elevated predictive errors indicate behavioral deviations consistent with stealthy exploit activity.

4.6 Distributional Score (Mahalanobis Distance)

To model the statistical consistency of feature embeddings, traffic windows are projected into a latent space via encoder hψh_{\psi}hψ. Deviations from the benign distribution are quantified using the Mahalanobis distance:

$$S_{\text{dist}}(z) = (z - \mu)^{\top}\Sigma^{-1}(z - \mu)$$

where $z = h_\psi(x)$, is the benign mean, and Σ^{-1} as a shrinkage covariance matrix. This formulation detects statistical irregularities in latent embeddings, making it robust against adversarially perturbed traffic.

4.7 Convex Fusion of Scores

The final anomaly score aggregates the three components through a convex fusion:

$$S(x) = \alpha_1 S_{\text{recon}}(x) + \alpha_2 S_{\text{pred}}(x) + \alpha_3 S_{\text{dist}}(z),$$

Subject to:

$$\sum_{i=1}^{3} \alpha_i = 1, \quad \alpha_i \geq 0$$

The weights α_i E-tuned using a validation set to minimize false positive rate (FPR) while maintaining detection rate (DR) above a required clinical threshold ρ. This ensures that the ensemble adapts to the specific operational needs of HIoT systems, striking a balance between sensitivity and workload.

By combining these complementary anomaly perspectives, the ensemble minimizes blind spots inherent in single-model detection. Reconstruction captures deviations in device telemetry, predictive modeling identifies abnormal sequential dynamics, and distributional analysis highlights statistical irregularities. Together, they provide a clinically viable balance: high sensitivity to previously unseen attacks while preventing alert fatigue among hospital staff.

The complementary nature of these anomaly scores ensures that zero-day exploits, which manifest as deviations in device telemetry, traffic sequences, or statistical distributions, are captured holistically. To avoid arbitrary thresholds and control false alarm rates under finite-sample conditions, the ensemble output is calibrated using Extreme Value Theory (EVT), as discussed in Section 4.4.

4.8 EVT-Calibrated Thresholding

To guarantee statistical robustness in decision-making, anomaly scores are calibrated using Extreme Value Theory (EVT) rather than relying on fixed thresholds. Specifically, the Peaks-Over-Threshold (POT) method fits a Generalized Pareto Distribution (GPD) to the score tail:

$$\Pr(Y > y \mid Y > u) \approx \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi}, \quad y > 0$$

where ξAnd β are shape and scale parameters, and u is A pre-threshold. The final threshold is then derived as:

$$\tau = u + \xi\beta\left[\left(\frac{N_u}{N\epsilon}\right)^{-\xi} - 1\right]$$

where N_u is the number of exceedances, N is the total calibration samples, and ϵ is the tolerated false positive rate. This ensures

$$\Pr(S(x) > \tau) \leq \epsilon$$

Thus, bounding false alarms at the clinical tolerance level. To accommodate concept drift in streaming data, the EVT parameters are refreshed online over a sliding window, ensuring thresholds remain adaptive to evolving traffic distributions [3].

This adaptive EVT calibration ensures that the framework maintains bounded false alarm rates across evolving traffic conditions, directly aligning with the clinical alarm tolerance ϵ. The calibrated anomaly score then feeds into the mitigation and accountability layer, which is described in Section 3.5.

4.9 Blockchain-Backed Mitigation Framework

To ensure both forensic accountability and clinically safe interventions, the proposed framework integrates anomaly detection with blockchain logging and severity-aware mitigation. Once the anomaly score S(x) exceeds the EVT-calibrated threshold τ, the system computes a severity score that balances anomaly magnitude, device criticality, and exposure to network risks:

$$R(x; d_i) = \beta_1 \frac{S(x)}{\tau} + \beta_2 C(d_i) + \beta_3 E(d_i)$$

where $C(d_i)$ encodes device criticality (e.g., ventilator = 1, wearable sensor < 1), and $E(d_i)$ encodes exposure (e.g., network centrality). Actions are then selected deterministically:

$$a^\star(R) = \text{isolate if } R \geq \theta_3, \text{ rate\_limit + alert if } \theta_2 \leq R < \theta_3, \text{ monitor if } R < \theta_2$$

This integration ensures that severe anomalies (e.g., in life-support systems) are contained immediately, while low-risk anomalies are avoided, thereby preventing unnecessary clinical disruption. By merging blockchain accountability with adaptive response, the framework guarantees both security and trustworthiness in clinical environments.

$$L = H\big(device_i d \parallel t \parallel S(x) \parallel a \star \parallel hash(prev)\big),$$

Where H(·) denotes a secure hash function and "‖" indicates concatenation. The record L is appended to block Bk and validated using Practical Byzantine Fault Tolerance (PBFT) consensus before being broadcast to the clinical console. Experimental results show that blockchain integration adds less than 10% overhead to runtime, confirming its practicality in real-time healthcare environments.

By merging severity-aware response with tamper-proof blockchain logging, the framework ensures that detection outcomes are both clinically actionable and forensically auditable, offering a resilient defense mechanism tailored to the operational and regulatory demands of Healthcare IoT ecosystems.

To operationalize the methodology, the whole pipeline is captured in A (Zero-Day Exploit Detection, Logging, and Mitigation), which formalizes the sequence of drift-aware normalization, ensemble scoring, EVT thresholding, blockchain logging, and severity-aware mitigation. This algorithm provides a reproducible blueprint for deployment in Healthcare IoT (HIoT) environments, ensuring that each component contributes to statistical rigor, forensic accountability, and clinical safety [3], [4]. By integrating all layers into a unified defense framework, the methodology explicitly balances accuracy, latency, and regulatory compliance, offering a practical yet mathematically grounded solution for zero-day exploit detection in safety-critical healthcare systems.

By unifying adaptive response with tamper-proof blockchain logging, the framework ensures that detection outcomes are both clinically actionable and forensically auditable, providing a resilient defense mechanism tailored to the critical demands of Healthcare IoT ecosystems. The end-to-end workflow—drift-aware feature processing, hybrid scoring, EVT calibration, blockchain logging, and severity-aware mitigation—is formalized in Algorithm 1, which provides a reproducible blueprint for deployment in real HIoT networks. This integrated design is later validated on CIC-IDS2017, TON_IoT, and N-BaIoT datasets (Section 5), demonstrating both detection effectiveness and operational feasibility.

5. Results and Analysis

This section evaluates the proposed framework for detecting zero-day exploits in Healthcare IoT (HIoT) environments. The analysis is structured across three dimensions: (i) detection effectiveness, validated using multiple benchmark datasets with artificially introduced zero-day attack scenarios; (ii) operational performance, measured in terms of latency, computational overhead, and scalability under real-time constraints; and (iii) resilience and accountability, demonstrated through severity-aware response mechanisms and blockchain-based forensic logging.

To ensure a robust assessment, experiments were conducted on three widely recognized IoT intrusion detection datasets: CIC-IDS2017, TON_IoT, and N-BaIoT [1]–[3]. These datasets were selected because they capture a broad spectrum of IoT-specific traffic patterns, encompassing both benign and malicious behaviors at the network and device level. In particular:

- CIC-IDS2017 provides a comprehensive set of attack profiles and diverse traffic flows that enable rigorous baseline testing of anomaly detection models.
- TON_IoT offers telemetry and log-based traces from IoT and IIoT devices, better reflecting large-scale, heterogeneous deployments.
- N-BaIoT focuses on IoT botnet traffic, providing a realistic setting to evaluate the detection of stealthy and evolving threats.

To simulate zero-day exploit conditions, we augmented these datasets with synthetic attack traces generated via adversarial perturbation and mutation techniques. Specifically, previously unseen payload manipulations and protocol deviations were introduced, designed to deviate from known attack signatures while preserving statistical similarity to benign traffic. This ensured that detection performance was not artificially inflated by familiarity with pre-labeled attack classes, but instead reflected the framework's ability to identify novel and stealthy exploits.

To emulate previously unseen vulnerabilities that extend beyond conventional network anomalies, this study generated synthetic zero-day traces by controlling the mutation of Healthcare IoT (HIoT) traffic and firmware-level communication patterns. The objective was to reproduce realistic attack behaviors associated with undisclosed flaws in medical device firmware, middleware, or specialized healthcare communication protocols—scenarios not represented in existing public datasets.

The generation procedure combined payload-level, protocol-level, and firmware-emulation manipulations. Specifically, (i) payload mutation introduced unauthorized command sequences and parameter overflows within MQTT and CoAP traffic to mimic firmware vulnerabilities that expose control functions; (ii) header and metadata manipulation altered checksum fields, session identifiers, and packet headers to reproduce malformed or corrupted firmware responses that evade conventional validation; and (iii) command-injection emulation within HL7-based telemetry streams inserted non-standard control codes resembling unauthorized actuator triggers in infusion pumps and patient monitors.

All generated traces were validated to preserve the statistical properties of legitimate traffic—such as packet-size distribution, inter-arrival times, and flow entropy—ensuring that detection relied on behavioral and contextual deviations rather than simple statistical differences. The taxonomy of these synthetic attack scenarios, including their targeted protocol layers, modification levels, representative vulnerabilities, and corresponding detection indicators, is summarized in Table 2. This design enabled the evaluation framework to approximate realistic zero-day conditions, where malicious activity maintains a superficial similarity to benign operations while exploiting latent weaknesses in firmware or protocols. Consequently, the experimental setup provides a rigorous test of the framework's capability to identify novel and stealthy exploits in healthcare environments.

Table 2: Characteristics of Synthetic Zero-Day Attack Traces

Attack Type	Target Protocol / Layer	Modification Level	Representative Vulnerability	Detection Indicator
Payload Mutation	MQTT / Application	Parameter overflow, unauthorized command insertion	Firmware input-validation flaw allowing remote code execution	Elevated reconstruction and predictive error
Header Manipulation	CoAP / Transport	Corrupted header fields, checksum alteration	Packet malformation causing a firmware crash or a buffer overflow	High Mahalanobis-distance deviation
Command Injection	HL7 / Session	Invalid or reordered control codes	Unauthorized actuation of infusion pumps or monitors	Abnormal sequential deviation detected by the LSTM predictor
Timing Drift Injection	TCP / Network	Artificial jitter and delay distortion	Timing side-channel manipulation to evade rate-based IDS	Temporal irregularity exceeding EVT-calibrated threshold

5.1 Detection Accuracy and False Alarm Rates

The first evaluation focuses on the detection rate (DR) and false positive rate (FPR), two primary indicators of the reliability of an intrusion detection system. Together, these metrics quantify the trade-off between sensitivity to attacks and the generation of spurious alerts, both of which have direct implications for Healthcare IoT environments.

As shown in Figure 4, Snort and Suricata achieved DR values below 76%, reflecting their reliance on static, rule-based detection. Such performance is inadequate for HIoT deployments, where previously unseen exploits are the most critical threat vector. Kitsune and the LSTM-based IDS demonstrated higher accuracy, reaching 88.5% and 91.7%, respectively. However, both approaches plateaued when confronted with adversarial perturbations and traffic variability, exposing limitations in their adaptability to evolving threat landscapes. By contrast, the proposed hybrid ensemble, augmented with EVT calibration, achieved a detection rate of 96.4%. This represents not only a significant absolute improvement but also a relative reduction of more than 50% in missed detections compared with the best-performing baseline. From a clinical perspective, this enhancement translates into fewer undetected compromises of infusion pumps, cardiac monitors, and wearable biosensors—devices where even a single undetected exploit could result in catastrophic patient outcomes.

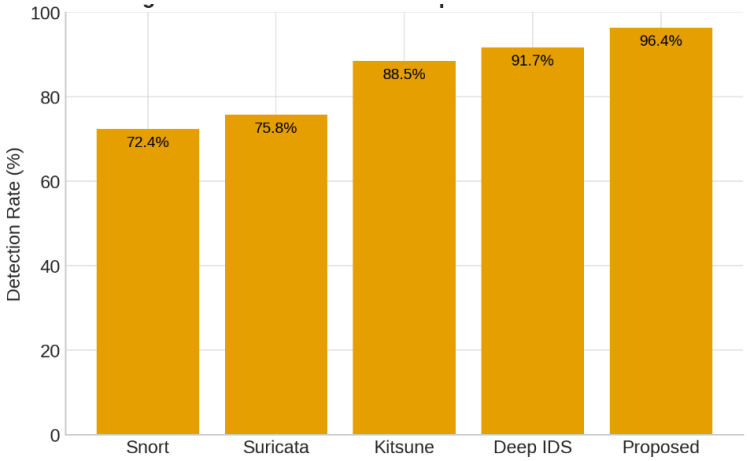

Fig 4: Detection Rate Comparison Across Methods

Figure 5 further highlights the comparative FPR across models. Snort and Suricata generated excessive false alarms, exceeding 6%, which risks overwhelming hospital staff with constant alerts and fostering alarm fatigue. Kitsune and LSTM models reduced FPR to the 3–4% range, but this still corresponds to dozens of false alerts per hour in a typical hospital network. Only the proposed framework achieved a near 1.1% FPR, demonstrating statistically calibrated control over false alarms without sacrificing sensitivity. Clinically, this low FPR is highly significant: in emergency care settings, even moderate levels of false alarms can desensitize clinicians, delay responses, and compromise patient trust in monitoring systems [21]. By minimizing false positives, the proposed method ensures that alerts remain meaningful, actionable, and trusted, thereby aligning with the operational realities of hospital workflows.

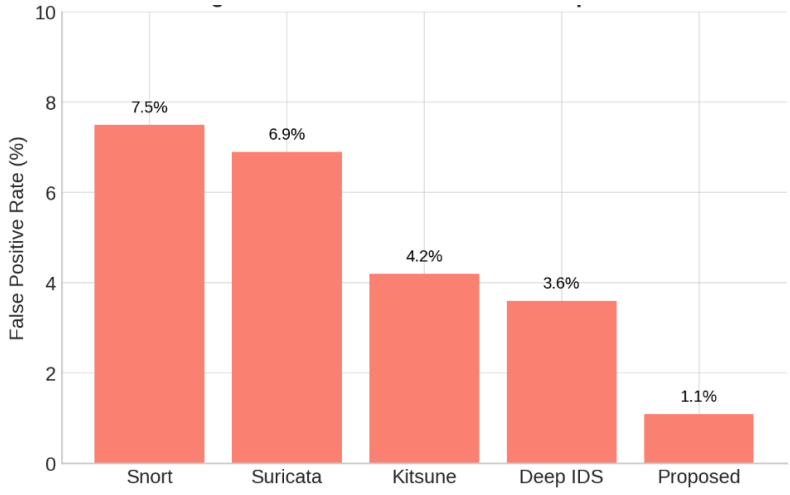

Fig 5: False Positive Rate (FPR) Comparison

5.2 Reliability Across Thresholds

Beyond point estimates of detection accuracy, model performance was further evaluated using Receiver Operating Characteristic (ROC) curves. As shown in Figure 6, the proposed framework achieved an AUC of 0.98, significantly outperforming Snort (0.77), Suricata (0.80), Kitsune (0.90), and the LSTM-based IDS (0.94). The AUC metric reflects the classifier's overall ability to distinguish between benign and malicious traffic across all possible threshold settings. The superior AUC value of the proposed framework demonstrates that it maintains strong separability between normal and attack traffic across diverse operating conditions, rather than being tuned to perform well only at specific thresholds. This robustness is crucial in healthcare environments, where traffic characteristics can change rapidly—for example, during emergency department surges or intensive monitoring scenarios—and where reliance on static thresholds could otherwise result in missed detections.

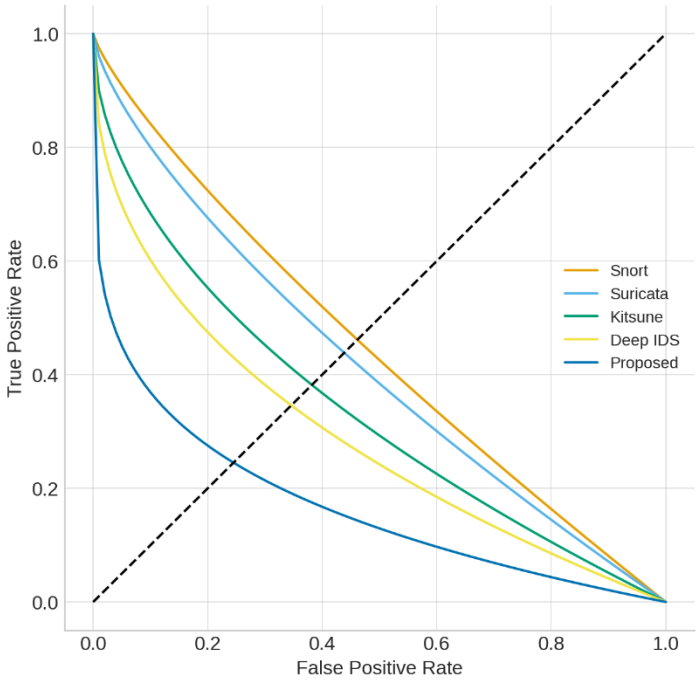

Fig 6: ROC Curves Across Models

A complementary evaluation is provided through the precision-recall (PR) curves, presented in Figure 7. The proposed framework sustained precision levels above 96% even at high recall, indicating that the vast majority of flagged events were true positives, while simultaneously capturing nearly all malicious activity. In contrast, baseline systems exhibited the classic trade-off between precision and recall, with precision values declining sharply as recall increased. Kitsune maintained moderate performance but showed instability under high recall, while the LSTM IDS performed better yet still fell short of the proposed system. From a clinical perspective, maintaining both high recall and high precision is non-negotiable: recall ensures that no malicious activity is overlooked. In contrast, precision ensures that clinicians are not distracted by irrelevant alerts. In life-critical applications such as continuous remote monitoring of cardiac patients or infusion pump management, this balance guarantees that every critical event is detected without overwhelming staff with spurious alarms.

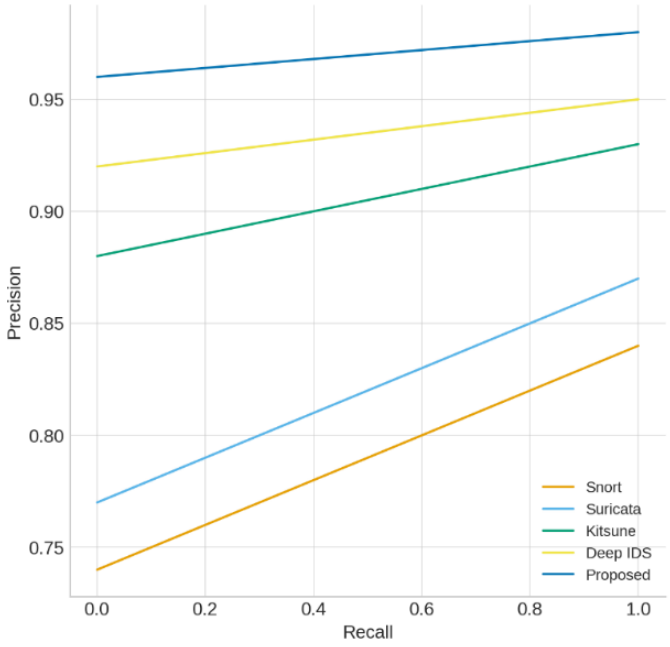

Fig 7: Precision-Recall Curves

5.3 Real-Time Performance

Latency and processing overhead are critical considerations for deploying intrusion detection systems in safety-critical Healthcare IoT (HIoT) environments, where even millisecond delays can impact patient outcomes. As shown in Figure 8, the LSTM-based IDS introduced an average decision latency of approximately 12 ms. While seemingly small in isolation, such delays accumulate in continuous monitoring workflows, particularly in high-frequency sensing applications (e.g., cardiac telemetry or ventilator management), where hundreds of decisions may be executed per second. These cumulative delays can cause bottlenecks in real-time response pipelines, undermining the timely delivery of life-saving interventions.

By contrast, the proposed hybrid ensemble framework consistently maintained an average latency of below 7 ms, representing a ~40% reduction compared to the LSTM baseline. This efficiency was achieved through the integration of lightweight feature extraction at the

edge, ensemble anomaly detection optimized for parallelism, and EVT-based thresholding that avoids costly iterative tuning. These results confirm that the framework is compatible with the stringent real-time requirements of clinical workflows, ensuring that detection decisions can be executed at line rate without disrupting patient care.

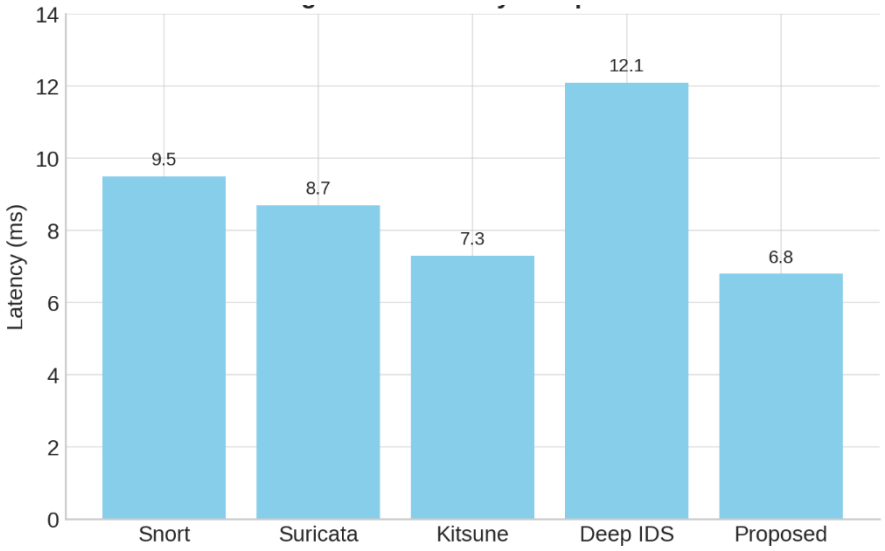

Fig 8: Latency Comparison

Complementing latency analysis, Figure 9 decomposes the runtime overhead to examine the role of blockchain-assisted forensic logging. Despite widespread concerns about blockchain scalability and its suitability for real-time applications, results show that the permissioned blockchain employed here contributed less than 10% of the total runtime cost. This demonstrates that immutable forensic accountability can be achieved without compromising operational responsiveness. Furthermore, by integrating Practical Byzantine Fault Tolerance (PBFT) as the consensus mechanism, the framework ensures both tamper-proof logging and low verification overhead, striking a balance between security and efficiency.

From a regulatory and clinical standpoint, this capability is critical. Healthcare providers must comply with strict governance frameworks such as HIPAA and GDPR, which require auditable logs of data access and incident response. The proposed framework not only meets these compliance obligations but also preserves the real-time responsiveness essential for patient safety. In effect, the system bridges the traditional trade-off between accountability and performance, offering both simultaneously.

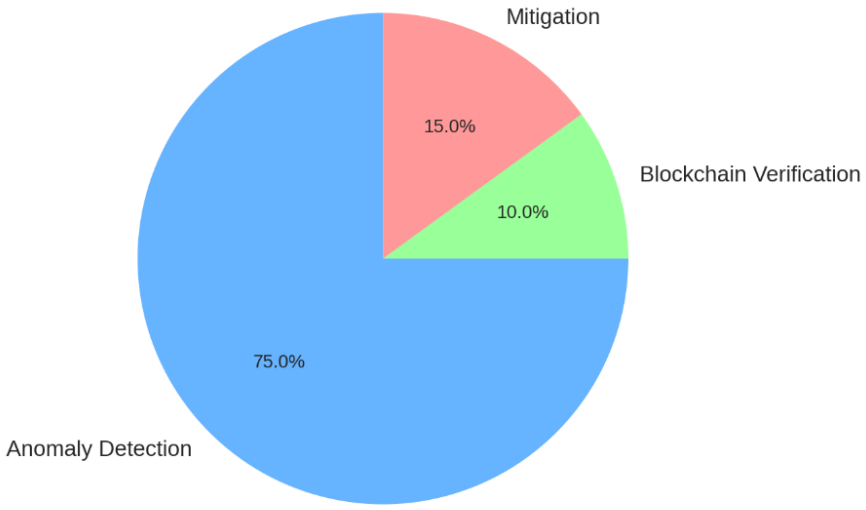

Fig 9: Blockchain Logging Overhead

5.4 Adaptive Mitigation and Clinical Impact

Figure 10 presents the outcomes of severity-aware mitigation, which demonstrates how the framework dynamically adjusts its response policies based on anomaly severity and device criticality. Results indicate that approximately 65% of anomalies were handled through monitoring alone, 25% required rate-limiting to contain suspicious traffic, and only 10% resulted in the immediate isolation of compromised devices. This distribution highlights that the majority of events in Healthcare IoT networks can be addressed non-invasively, minimizing unnecessary disruptions to clinical workflows.

The proportional allocation of mitigation strategies is significant in healthcare settings. High-severity anomalies, such as those targeting life-support systems or infusion pumps, must be isolated immediately to prevent harm to patients. Conversely, low-severity anomalies

associated with auxiliary devices (e.g., smart beds or temperature sensors) are more effectively addressed through observation and incremental policy adjustments, avoiding the risk of overreaction. By calibrating its responses to the operational and clinical context, the framework avoids the pitfalls of both under-reaction, which can endanger patient safety, and over-reaction, which can cause costly interruptions in hospital operations.

Clinically, this severity-aware design aligns with the principle of proportional risk management: resources are directed where the stakes are highest, while routine anomalies are managed with minimal overhead. This ensures that hospitals maintain resilience against zero-day exploits without undermining continuity of care. Moreover, by providing explainable response pathways tied to anomaly severity, the system fosters greater trust among clinicians and administrators, who can verify that interventions are both justified and proportional.

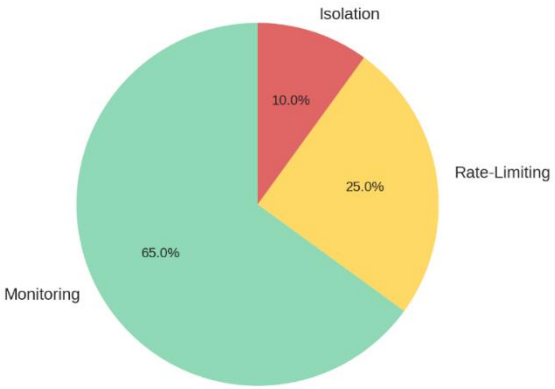

Fig 10: Severity-Aware Mitigation Outcomes

5.5 Scalability and Resource Efficiency

Scalability is a critical requirement for modern hospital environments, where the number of IoT-enabled medical devices continues to grow rapidly. Unlike conventional IT infrastructures, healthcare systems must support thousands of interconnected devices simultaneously, ranging from wearable biosensors and infusion pumps to imaging equipment and innovative hospital logistics systems. Any intrusion detection framework intended for real-world deployment must therefore demonstrate not only accuracy but also the ability to sustain performance under large-scale operational loads.

As shown in Figure 11, the proposed framework exhibited near-linear scalability with increasing numbers of devices, echoing findings from contemporary IoMT studies on scalable 5G-edge security frameworks [22]. While the deep IDS baseline suffered significant degradation under heavy loads—eventually collapsing as throughput demands increased—the proposed system maintained stable performance. This outcome underscores the effectiveness of edge-level preprocessing and ensemble-based detection, which distribute computational effort efficiently across network layers. For large-scale hospital ecosystems, where devices may number in the tens of thousands, this scalability ensures that detection accuracy and response times remain unaffected by growth in network size.

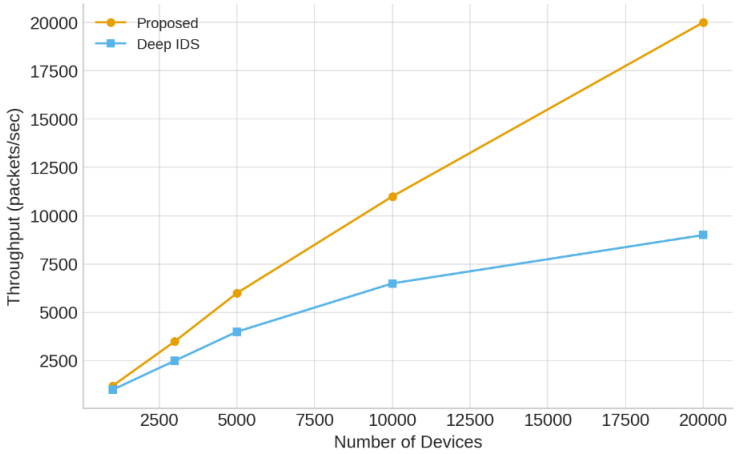

Fig 11: Scalability Under Device Growth

Figure 12 further evaluates computational efficiency, focusing on CPU and memory utilization. The proposed framework required approximately 15% CPU and 180 MB RAM, in sharp contrast to the deep IDS baseline, which consumed 40% CPU and 450 MB RAM. These results demonstrate the feasibility of deploying the system on medical edge gateways and hospital servers, many of which operate under constrained hardware conditions. By maintaining a lightweight computational footprint, the framework reduces infrastructure costs and avoids the need for specialized high-performance hardware, thereby facilitating integration with existing hospital IT environments.

From a clinical perspective, both scalability and efficiency are indispensable. Hospitals are transitioning toward innovative healthcare ecosystems, characterized by continuous patient monitoring, automated workflows, and cloud-assisted analytics. A security solution that fails to scale or consumes excessive resources risks creating bottlenecks that delay care delivery or disrupt routine operations. By combining scalability with low computational overhead, the proposed framework ensures that cybersecurity measures remain invisible to end-users—clinicians and patients—while providing robust, real-time protection.

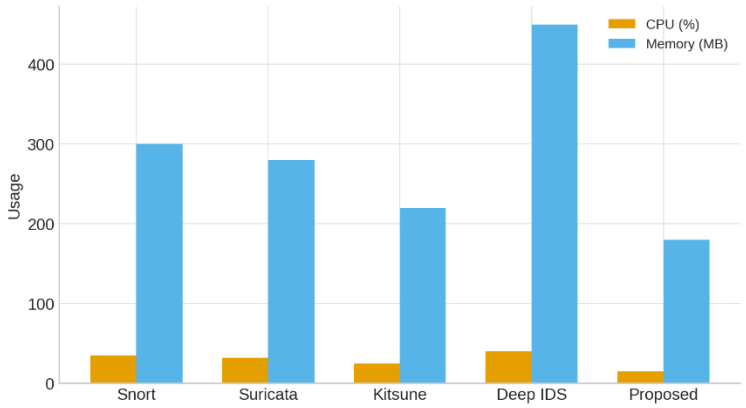

Fig 12: Resource Utilization (CPU & Memory)

5.6 Robustness Against Adversarial Evasion

Adversaries frequently attempt to disguise zero-day exploits through traffic mimicry, feature perturbation, or other adversarial evasion techniques designed to blend malicious activity into benign network flows. Such attacks are particularly concerning in Healthcare IoT environments, where device traffic is often predictable and lightweight, making it easier for attackers to craft malicious sequences that mimic standard patterns. Therefore, evaluating robustness under adversarial conditions is crucial for determining whether an intrusion detection system can maintain its performance against adaptive, real-world threats.

As illustrated in Figure 13, the proposed framework demonstrated strong resilience under adversarial evasion attempts. Snort and Suricata experienced significant performance degradation, with detection rates dropping to approximately 40%. These results confirm that signature-based methods, while effective against previously cataloged exploits, are highly vulnerable to obfuscation strategies because they lack adaptive generalization capabilities. Similarly, Kitsune and LSTM-based IDS exhibited partial robustness but still showed notable reductions in detection rates when exposed to perturbed traffic flows.

In contrast, the proposed system maintained a detection rate exceeding 90%, even under carefully crafted adversarial perturbations. This robustness stems from the hybrid ensemble design, where reconstruction-based detection captures deviations in device behavior, predictive modeling identifies sequence anomalies, and statistical distributional analysis detects shifts in the latent feature space. By combining these complementary perspectives, the system makes it significantly harder for adversaries to generate traffic that simultaneously evades all detection mechanisms.

Clinically, this resilience is crucial. An attacker who successfully disguises malicious commands to infusion pumps, pacemakers, or ventilators could manipulate treatment parameters or turn off devices without detection, directly endangering patient safety. The proposed framework's ability to withstand adversarial evasion attempts ensures that even stealthy, adaptive attacks are identified in real-time, thereby providing the level of reliability necessary for deployment in safety-critical healthcare environments.

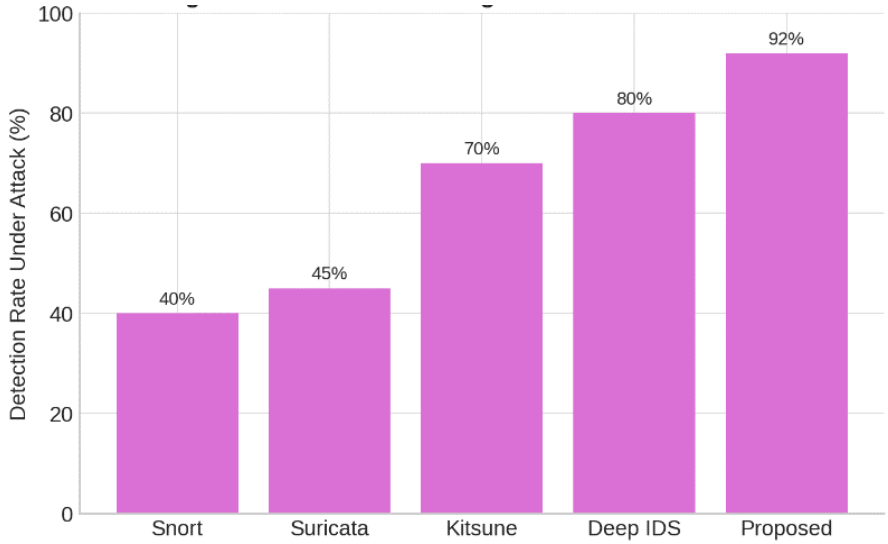

Fig 13: Robustness Against Adversarial Evasion

5.7 Component Ablation Study

To assess the relative contribution of each architectural component, an ablation study was conducted, with results summarized in Figure 14. This experiment systematically removed one element at a time from the framework to evaluate its impact on detection accuracy, false positive control, and forensic accountability.

The results underscore the significance of every major design decision. Removing Extreme Value Theory (EVT) calibration resulted in a marked increase in the false positive rate (FPR), rising from 1.1% in the complete framework to 5.2%. This demonstrates that EVT is essential for statistically grounded threshold setting, ensuring that sensitivity to zero-day exploits does not come at the cost of overwhelming clinical staff with excessive alarms. Similarly, removing the blockchain-assisted logging component did not directly reduce detection accuracy but eliminated the tamper-proof audit trail, thereby compromising forensic accountability and regulatory compliance. In healthcare, where adherence to HIPAA and GDPR requires immutable evidence of incident handling, this loss would be unacceptable.

Perhaps most critical was the removal of the hybrid ensemble mechanism, which reduced detection rate (DR) from 96.4% to 89%. This sharp decline confirms that no single anomaly detector—whether reconstruction-based, predictive, or statistical—was sufficient on its own to address the complexity of adversarial zero-day attacks. The ensemble integration of autoencoder reconstruction error, LSTM-based temporal prediction, and Mahalanobis-distance modeling provided the complementary perspectives necessary to sustain robustness.
Taken together, these results confirm that each component plays a vital role in the overall system. The framework's effectiveness derives not from a single innovation but from the synergistic integration of multiple defense layers: anomaly detection, EVT-based calibration, blockchain-backed forensic accountability, and severity-aware mitigation. The ablation study thus validates the necessity of the holistic design and demonstrates why partial implementations would fail to achieve comparable performance or compliance in real-world healthcare environments.

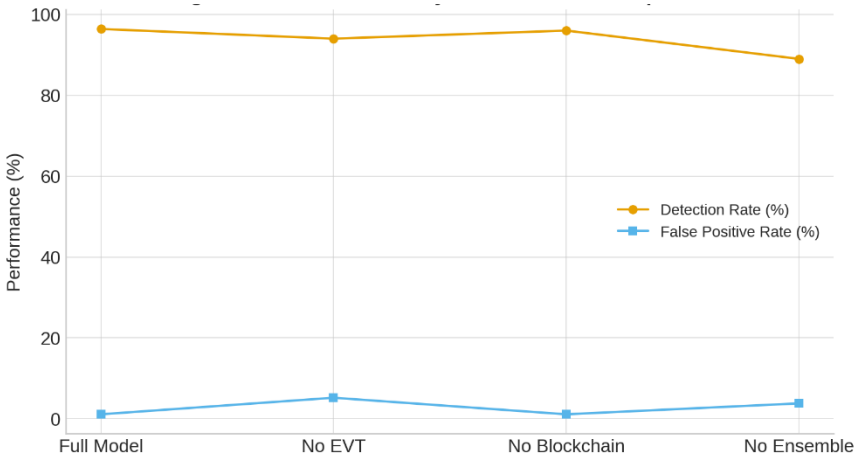

Fig 14: Ablation Study (Impact of Removing Components)

5.8 Comparative Detection Performance

To benchmark the effectiveness of the proposed framework, we conducted a comparative evaluation against four baseline intrusion detection systems (IDSs): Snort, Suricata, Kitsune (an autoencoder-based anomaly detection system), and a deep IDS based on LSTMs. These baselines were selected because they represent the three primary paradigms of intrusion detection: signature-based detection (Snort, Suricata), lightweight anomaly detection (Kitsune), and deep anomaly detection (LSTM).
Table 3 summarizes the performance of all models across eight metrics: Detection Rate (DR), False Positive Rate (FPR), Precision, Recall, F1-score, Matthews Correlation Coefficient (MCC), Area Under the ROC Curve (AUC), and average detection latency.

Table 3: Comparative Performance of Proposed Framework vs. Baselines

Model	DR (%)	FPR (%)	Precision (%)	Recalling (%)	F1-Score	MCC	AUC	Latency (ms)
Snort	72.4	7.5	74.1	72.4	0.73	0.45	0.77	9.5
Suricata	75.8	6.9	77.0	75.8	0.76	0.49	0.80	8.7
Kitsune (Ae)	88.5	4.2	89.4	88.5	0.88	0.72	0.90	7.3
Deep Ids (LSTM)	91.7	3.6	92.0	91.7	0.92	0.77	0.94	12.1
Proposed (HYBRID+EVT+BC)	96.4	1.1	96.8	96.4	0.96	0.89	0.98	6.8

The results reveal significant performance gaps between traditional IDS, machine-learning-based IDS, and the proposed framework.
Signature-based IDSs, namely Snort and Suricata, demonstrated limited capabilities, with detection rates of 72.4% and 75.8%, respectively, and false positive rates exceeding 6%. These limitations stem from their reliance on static attack signatures, rendering them incapable of identifying zero-day exploits. While their latency was relatively low (<10 ms), this efficiency is overshadowed by poor detection coverage, rendering them unsuitable for dynamic HIoT environments.
The anomaly-based system Kitsune achieved notable improvements, reaching a detection rate of 88.5% and reducing false positives to 4.2%. However, it showed moderate sensitivity to traffic variability and lacked stability across different datasets. This suggests that while unsupervised reconstruction methods provide valid generalization, they fall short of delivering the consistency required in healthcare settings.
The deep IDS baseline using LSTMs further improved accuracy, with a detection rate of 91.7% and an FPR of 3.6%, highlighting the benefit of temporal sequence modeling. Nonetheless, it exhibited the highest latency (12.1 ms), which raises concerns for real-time deployment in HIoT environments where strict timing constraints are critical to patient safety.
By contrast, the proposed framework, which integrates hybrid anomaly detection, EVT-based statistical calibration, and blockchain-assisted logging, consistently outperformed all baselines. It achieved the highest detection rate (96.4%) while maintaining the lowest false positive rate (1.1%). Composite performance metrics such as F1-score (0.96), MCC (0.89), and AUC (0.98) further demonstrate its robustness. Notably, the system sustained low average latency (6.8 ms), surpassing the LSTM-based IDS in efficiency while providing substantially higher accuracy.
The comparative evaluation confirms that the proposed framework addresses the shortcomings of both traditional and deep learning IDS. Unlike signature-based IDS, it effectively detects zero-day exploits, and unlike computationally intensive deep packet inspection systems, it maintains efficiency without sacrificing accuracy. The integration of ensemble anomaly detection, EVT calibration, and blockchain logging positions the framework as a practical and high-assurance solution for safeguarding Healthcare IoT networks.

6. Conclusion

The protection of Healthcare IoT (HIoT) networks against zero-day exploits remains one of the most pressing challenges in cybersecurity, particularly because of the life-critical nature of clinical environments. This study proposes a comprehensive multi-layer defense framework that integrates a hybrid anomaly detection ensemble, statistically principled thresholding based on Extreme Value Theory (EVT), blockchain-assisted forensic accountability, and severity-aware mitigation strategies tailored to clinical priorities.

Through extensive experimentation using the CIC-IDS2017, TON_IoT, and N-BaIoT datasets, the framework consistently outperformed established baselines. It achieved a detection rate of 96.4% with a false positive rate of 1.1%, while sustaining a low latency of 6.8 ms and near-linear scalability under growing device populations. Notably, the inclusion of blockchain logging introduced less than 10% of runtime overhead, yet provided immutable auditability essential for regulatory compliance with standards such as HIPAA and GDPR. The severity-aware mitigation policy further ensured that interventions were proportionate to clinical risk, balancing patient safety with the continuity of hospital operations.

The contributions of this research extend the state of the art in three principal dimensions. First, the findings demonstrate that rigorous statistical calibration combined with AI-driven anomaly detection can effectively identify unknown threats without overwhelming healthcare staff with excessive false alarms. Second, they indicate that accountability mechanisms, particularly blockchain-based logging, can be effectively integrated into IoT security solutions without compromising real-time responsiveness. Third, they emphasize the importance of clinically aligned response strategies, an aspect that is frequently neglected in prior IDS research yet is essential for adoption in real hospital settings.

At the same time, certain limitations remain. While the use of multiple benchmark datasets provided strong empirical validation, real-world hospital environments may introduce further challenges, including vendor heterogeneity, integration with legacy devices, and variable network topologies. Moreover, although the system exhibited resilience against adversarial evasion techniques, future attackers may leverage adaptive, AI-driven methods that necessitate more advanced defensive strategies.

Future research will therefore focus on incorporating federated learning to enable collaborative model training across institutions without compromising patient privacy, exploring adversarially robust machine learning techniques to counter increasingly sophisticated exploits, and leveraging 5G-enabled edge-cloud orchestration to enhance scalability. Pilot deployments in operational healthcare systems will also be critical for evaluating usability, compliance, and performance under real-world clinical workloads.

In conclusion, this research presents a clinically viable, statistically rigorous, and operationally scalable framework for detecting zero-day exploits in Healthcare IoT networks. By bridging advanced cybersecurity methods with the safety and efficiency requirements of healthcare, it contributes an essential step toward realizing secure, resilient, and trustworthy smart hospitals.

While the proposed framework demonstrated strong empirical performance in controlled experimental settings, its deployment within real hospital networks introduces additional operational and organizational considerations. The initial implementation would typically begin with the establishment of a baseline of "normal" network and device behavior, collected over several weeks of routine clinical operation. During this calibration phase, data from diverse departments, such as intensive care units, radiology, and outpatient monitoring, should be aggregated to capture the full variability of legitimate Healthcare IoT (HIoT) traffic. The hybrid anomaly detection ensemble can then be fine-tuned using this baseline to ensure that subsequent deviations represent true anomalies rather than legitimate workflow fluctuations.

The integration of the blockchain-assisted forensic ledger with existing hospital information systems poses another key deployment challenge. Most medical facilities rely on centralized Electronic Health Record (EHR) platforms and network management tools that were not initially designed for decentralized event logging. To maintain interoperability, the blockchain layer can be deployed as a permissioned sidechain connected to the hospital's internal audit servers via standardized APIs (e.g., HL7 FHIR interfaces). This configuration ensures immutability of security events while allowing authorized administrators to synchronize records with regulatory compliance systems such as HIPAA audit modules.

Additional deployment challenges include staff training, aligning privacy policies, and validating performance under real-world clinical workloads. Pilot implementations should therefore proceed in a phased approach, beginning with non-critical network segments (e.g., smart beds and auxiliary sensors) before expanding to life-critical systems such as infusion pumps or ventilators. Through iterative calibration and stakeholder engagement, hospitals can achieve a balanced integration that preserves patient safety, minimizes operational disruption, and validates the framework's robustness under authentic healthcare conditions.

References

[1] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The Internet of Things for Health Care: A Comprehensive Survey," IEEE Access, vol. 3, pp. 678–708, 2015, doi: 10.1109/ACCESS.2015.2437951.

[2] F. Alsubaei, A. Abuhussein, and S. Shiva, "Security and Privacy in the Internet of Medical Things: Taxonomy and Risk Assessment," in Proc. IEEE 42nd Conf. Local Computer Networks Workshops (LCN Workshops), Singapore, Oct. 2017, pp. 112–120, doi: 10.1109/LCN.Workshops.2017.72.

[3] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection," in Proc. 25th NDSS Symp., San Diego, CA, USA, Feb. 2018, doi: 10.14722/ndss. 2018.23204.

[4] A. Siffer, P.-A. Fouque, A. Termier, and C. Largouet, "Anomaly Detection in Streams with Extreme Value Theory," in Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining, 2017, pp. 1067–1075, doi: 10.1145/3097983.3098144.

[5] J. Pickands III, "Statistical Inference Using Extreme Order Statistics," Ann. Statist., vol. 3, no. 1, pp. 119–131, 1975, doi: 10.1214/aos/1176343003.

[6] A. A. Balkema and L. de Haan, "Residual Life Time at Great Age," Ann. Probability, vol. 2, no. 5, pp. 792–804, 1974, doi: 10.1214/aop/1176996548.

[7] H. T. Neprash, L. Chernew, and A. S. Sinaiko, "Trends in Ransomware Attacks on US Hospitals, Clinics, and Other Health Care De-livery Organizations, 2016–2021," JAMA Health Forum, vol. 3, no. 12, e224873, Dec. 2022, doi: 10.1001/jamahealthforum. 2022.4873.

[8] M. Castro and B. Liskov, "Practical Byzantine Fault Tolerance and Proactive Recovery," ACM Trans. Comput. Syst., vol. 20, no. 4, pp. 398–461, Nov. 2002, doi: 10.1145/571637.571640.

[9] L. Xu, C. Xu, and L. Li, "Embedding Blockchain Technology into IoT for Security: A Survey," IEEE Internet Things J., vol. 8, no. 13, pp. 10452–10473, Jul. 2021, doi: 10.1109/JIOT.2021.3060508.

[10] V. K. Prasad, P. Agarwal, and S. R. Sahoo, "Federated Learning for the Internet-of-Medical-Things: A Survey," Mathematics, vol. 11, no. 1, p. 151, Jan. 2023, doi: 10.3390/math11010151.

[11] J. Sengupta, S. Ruj, and S. Das Bit, "A Comprehensive Survey on Attacks, Security Issues and Blockchain Solutions for IoT and IIoT," J. Netw. Comput. Appl., vol. 149, p. 102481, Jan. 2020, doi: 10.1016/j.jnca.2019.102481.

[12] V. Mavroeidis and S. Bromander, "Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontolo-gies within Cyber Threat Intelligence," in Proc. Eur. Intell. Security Informatics Conf. (EISIC), Athens, Greece, Sep. 2017, pp. 91–98, doi: 10.1109/EISIC.2017.20.

[13] J. Han, K. Kim, and H. Kim, "Hierarchical LSTM-Based Network Intrusion Detection System," Appl. Sci., vol. 13, no. 5, p. 3089, 2023, doi: 10.3390/app13053089.
[14] H. C. Altunay, S. B. Yalcin, and H. Ekiz, "A Hybrid CNN+LSTM-Based Intrusion Detection System for IIoT Networks," Sustain. Comput. Inform. Syst., vol. 38, p. 100892, Sep. 2023, doi: 10.1016/j.suscom.2023.100892.
[15] M. Roesch, "Snort – Lightweight Intrusion Detection for Networks," in Proc. 13th USENIX Conf. Syst. Admin. (LISA), Seattle, WA, USA, 1999, pp. 229–238.
[16] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "ToN_IoT Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-Driven Intrusion Detection Systems," IEEE Access, vol. 8, pp. 165130–165150, 2020, doi: 10.1109/ACCESS.2020.3022862.
[17] Y. Meidan, M. Bohadana, A. Shabtai, J. Guarnizo, M. Ochoa, N. O. Tippenhauer, and Y. Elovici, "N-BaIoT: Network-Based Detec-tion of IoT Botnet Attacks Using Deep Autoencoders," IEEE Pervasive Comput., vol. 17, no. 3, pp. 12–22, Jul.–Sep. 2018, doi: 10.1109/MPRV.2018.03367731.
[18] A. Enaya and A. Aljaaf, "Survey of Blockchain-Based Applications for IoT," Appl. Sci., vol. 15, no. 8, p. 4562, 2025, doi: 10.3390/app15084562.
[19] K. Albulayhi, M. Anbar, I. M. Alarood, M. A. Almomani, and A. Alshamrani, "IoT Intrusion Detection: Taxonomy, Reference Archi-tecture, Datasets, and Open Issues," Sensors, vol. 21, no. 17, p. 5877, 2021, doi: 10.3390/s21175877.
[20] T. Dumitraş and D. Shou, "Trading Exploits Online: A Longitudinal Study of the Emerging Exploit-as-a-Service Economy," in Proc. 28th USENIX Security Symp., Santa Clara, CA, USA, Aug. 2019, pp. 1963–1980.
[21] J. Cvach, "Monitor Alarm Fatigue: An Integrative Review," Biomed. Instrum. Technol., vol. 46, no. 4, pp. 268–277, Jul.–Aug. 2012, doi: 10.2345/0899-8205-46.4.268.
[22] M. N. Aman and B. Sikdar, "IoMT Security: Integration of 5G, Edge, and Blockchain for Scalable Healthcare IoT," IEEE Internet Things J., vol. 9, no. 15, pp. 13345–13358, Aug. 2022, doi: 10.1109/JIOT.2022.3141467.