Design and Evaluation of a Security-Integrated Anomaly Detection Framework for IoT-Based Vaccine Cold Chain Kits

Sai Harshit B [1] [0009-0002-2376-5914] and Dr Arun Kumar B R [2] [0000-0002-8659-6102]

[1] BMS Institute of Technology and Management, Bengaluru, Karnataka, India
[2] BMS Institute of Technology and Management, Bengaluru, Karnataka, India
sai.harshitbalaji@gmailcom
arunkumarbr@bmsit.in

Abstract. Cold-chain logistics rely on IoT sensors to ensure safe storage and transport of temperature-sensitive goods such as vaccines. These systems are vulnerable to both operational anomalies (e.g., temperature, humidity, geofence breaches) and cyberattacks that manipulate telemetry (e.g., spoofing, replay, suppression). Existing methods like Isolation Forests and risk fusion models detect statistical outliers or weight anomalies using Common Vulnerability Scoring System(CVSS) / Exploit Prediction Scoring System (EPSS) / Device Vulnerability Density (DVD) scores, but fail to capture malicious data manipulations and often raise false positives. This paper proposes a risk-aware lightweight, layered anomaly detection framework that integrates per-feature temporal detectors, a cyberattack detector based on cross-sensor consistency, and a fusion module incorporating device vulnerability scores. Using a simulated IoT dataset, we demonstrate improved anomaly detection, reduced false positives through temporal persistence, and the ability to flag malicious manipulations. The proposed framework advances IoT cold-chain monitoring with a security-aware and explainable approach.

Keywords: Cold-chain logistics, IoT security, Anomaly detection, Cyberattack detection, Risk-aware machine learning, Temporal persistence, Isolation Forest, CVSS, EPSS, DVD.

1 Introduction

Cold-chain logistics ensures that temperature-sensitive goods such as vaccines, pharmaceuticals, and perishable food items are transported and stored within stringent environmental conditions. The increasing adoption of IoT-enabled devices has made it possible to continuously monitor storage conditions, geofencing, and tampering events during transport. However, these devices face two categories of risks: operational anomalies, arising from equipment malfunction or environmental deviation, and cyberattacks, where adversaries manipulate telemetry data to conceal unauthorized actions.

Operational anomalies include violations of temperature thresholds, prolonged humidity excursions, tampering of containers, or deviation from geofenced routes. Conventional anomaly detection models can flag such deviations, but they often generate false positives when transient noise or single-sample spikes occur. On the other hand, cyberattacks such as replaying previously normal telemetry, suppressing sensor outputs, or injecting smoothed values can deceive detection systems, resulting in undetected breaches of security.

Existing approaches, such as Isolation Forests and risk fusion models that incorporate vulnerability scores, fall short in this dual context. They either treat anomaly detection as a blind statistical exercise or rely only on device posture without validating telemetry integrity. This creates a critical gap in cold-chain IoT: the inability to simultaneously detect operational deviations and malicious manipulations in a lightweight, IoT-suitable manner.

This research addresses the gap by introducing a layered anomaly detection framework that combines operational anomaly detectors, a cyberattack detector exploiting cross-sensor inconsistencies, and a fusion module that incorporates device-level risk posture into the final decision. The framework is designed to be explainable, IoT-friendly, and deployable in real-world cold-chain monitoring scenarios.

2 Problem Statement

The rapid adoption of IoT-enabled vaccine cold-chain kits (temperature, humidity, GPS sensors) promises real-time assurance of vaccine safety, yet these devices are vulnerable to firmware tampering, sensor spoofing, and network-based attacks. While anomaly detection models (Isolation Forest, geofence-based route deviation) can flag operational anomalies, they cannot detect **intentional cyber exploits**.

Currently, there is **no unified framework** that combines:

- **Proactive security assessment (VAPT)**,
- **Standardized control compliance (NIST SP 800-53/82)**, and
- **Quantitative risk metrics (CVSS, EPSS, Device Vulnerability Density - DVD)**

to **contextualize anomalies with cyber risk posture** of each IoT node in the vaccine cold-chain system. This gap leaves vaccine safety dependent on devices that may be compromised, silently bypassing anomaly detectors and poisoning blockchain records.

3 Objectives

1. Design a cyber-risk-driven anomaly detection framework for IoT vaccine cold-chain kits, which:
    - Incorporates vulnerability risk scoring inspired by VAPT results, including CVSS (severity), EPSS (exploit likelihood), and Device Vulnerability Density (DVD) metrics tailored for IoT devices.
    - Maps vulnerability impacts and compliance gaps to relevant NIST standards, primarily SP 800-82 for IoT/industrial controls.
1. Develop a risk-weighted anomaly detection pipeline combining:
    - Temperature + Humidity + tamper + geofence route deviation anomaly detection with cross sensor cyber anomaly detection
    - Risk scores as weighting features to modulate alert confidence and urgency
2. Build a dashboard for security posture + anomaly correlation, showing:
    - Device-level vulnerability risk (CVSS/EPSS/DVD)
    - Live anomaly alerts mapped to affected devices
3. Validate the framework on simulated IoT vaccine logistics data, comparing:
    - Baseline anomaly detection vs
    - Security-augmented anomaly detection (after VAPT hardening)

4 Literature Survey

Recent advancements in IoT-enabled vaccine cold-chain management have demonstrated significant potential for improving real-time environmental monitoring and supply chain efficiency. Jiang et al. introduced an IoT framework focusing on leveraging temperature and humidity sensors to detect anomalies in vaccine storage, thus improving logistical efficiency but without explicitly addressing cyberattack resilience. Expanding on this, Harrabi et al. developed a resource-constrained embedded system using deep learning on ESP32 devices to identify temperature anomalies in real time, thus improving detection capabilities on edge devices.

Ansari et al. presented a case study improving vaccine supply chain efficiency during COVID-19, focusing on logistics rather than security. Similarly, Alshadi et al. proposed an IoT-based cold storage and transportation management solution aimed at better monitoring but with limited coverage of cybersecurity concerns. Ahmad et al. examined barriers in IoT implementation for cold chain systems, highlighting technical and security challenges.

The benefit of blockchain technology applied for vaccine supply chain traceability is explored by Shiri et al. and Yadav et al. , who pointed out increased transparency and tamper-resistance but also the persisting vulnerability of IoT endpoints. Taj et al.

surveyed IoT-based supply chain management approaches, underscoring the increasing importance of integrating security.

Research published in IJNRD emphasized the role of IoT for efficient vaccine transport and real-time tracking, while a broad review identified the need for more cohesive, secure frameworks in digital vaccine supply chains.

On cybersecurity, Siraparapu et al. reviewed IoT threat landscapes, recommending automated vulnerability assessments. Radanliev et al. and Waqdan et al. advanced dynamic risk scoring approaches using CVSS and EPSS metrics, supporting proactive detection. The Data Security Council of India provided guidelines focusing on threat modeling relevant to such IoT deployments.

Further analysis of vulnerabilities was explained in the article titled "Analysis of Security Vulnerabilities for IoT Devices" from the *Journal of Information Processing Systems*, vol. 20 emphasizing the need for continuous monitoring. Li et al. proposed predictive models for estimating IoT device risks based on vulnerability data, offering pathways for preemptive actions.

Multiple studies (eg. *Computers & Industrial Engineering*, 2025.[17][18]and *Procedia Computer Science*, vol. 227, 2023. [19]) investigated blockchain-enabled vaccine logistics, each underscoring operational resilience but usually lacking integration with real-time security assessment.

Surveys on anomaly detection in IoT cold chains by Corradino et al. catalog methods but noted limited integration with device risk scoring. Works by Wright et al. , Harrou et al. , and Zou et al. developed machine learning and statistical fault detection techniques with application to sensor data in sensitive cold chain logistics. Viswanath et al. introduced trust models combining sensor and risk data, but practical implementation challenges remain. Lastly, Gillespie et al. demonstrated real-world anomaly detection applicable to cold supply logistics, confirming practical feasibility.

This comprehensive review indicates that while component technologies are mature, a unified framework combining operational anomaly detection, cyberattack resilience, continuous adaptive risk scoring (CVSS, EPSS, DVD), and regulatory compliance mapping remains an open challenge. Our work aims to bridge this gap with a novel layered approach tailored for vaccine cold-chain IoT.

5 Research Gap

The gaps identified from the literature survey were as follows:

1. **Operational focus only:** Most existing anomaly detection approaches target environmental deviations such as temperature spikes or route breaches, but

they do not account for malicious telemetry manipulations like spoofing or replay.

2. **Ad-hoc security testing:** Security assessments are usually limited to one-time penetration tests or static analysis of IoT firmware, with no automated or continuous integration into the anomaly detection process.

3. **Compliance gap:** While NIST provides well-defined control guidelines for IoT and industrial systems, many prototypes overlook mapping detection results to these standards, leaving compliance unaddressed.

4. **Risk metrics underutilized:** Established scoring models such as CVSS (severity), EPSS (exploitability), and DVD (vulnerability density) are rarely applied in IoT cold-chain contexts, reducing the ability to prioritize alerts by actual device risk.

5. **Fragmented design:** Security testing and anomaly detection are often developed in isolation, preventing the creation of unified, security-aware anomaly detection frameworks for cold-chain IoT.

6 Novelty of the Proposed Work

Building on the identified research gaps, this work introduces several innovations that distinguish it from prior IoT anomaly detection approaches:

1. **Dual-layer anomaly coverage.** In contrast to methods focused solely on operational anomalies, our framework integrates both operational deviations (temperature, geofence, tamper) and cyber manipulations (timestamp spoofing, replay).
2. **Continuous risk-aware integration.** Security posture metrics (CVSS, EPSS, DVD) are dynamically incorporated into thresholding and updated in response to anomaly outcomes, replacing one-time assessments with continuous risk evolution.
3. **Compliance alignment.** Detection outputs are explicitly mapped to NIST control families, bridging anomaly detection with compliance requirements.
4. **Risk-driven prioritization.** By embedding CVSS, EPSS, and DVD into anomaly scoring, the system can prioritize alerts according to device-level risk, enabling resource-aware response.
5. **Unified risk-aware fusion.** Although operational and cyber detectors are trained separately, their results are integrated through a fusion layer tied to security posture, reducing fragmentation and enabling closed-loop risk management.

These elements collectively establish the novelty of a **lightweight, security-aware, and compliance-ready framework** for IoT cold-chain monitoring.

7 Contributions

The main contributions of this paper are as follows:

1. **A layered, risk-aware anomaly detection framework** that combines temporal anomaly detectors, a supervised cyberattack detector, and a risk-aware fusion layer.
2. **Synthetic IoT telemetry datasets** with both operational and cyber anomalies, designed to emulate cold-chain scenarios and allow reproducible evaluation.
3. **A dynamic posture update mechanism** that integrates CVSS, EPSS, and DVD metrics into detection and continuously adjusts device risk profiles.
4. **A comprehensive experimental evaluation** against two baselines — Isolation Forest and Risk Fusion — demonstrating superior recall and balanced performance on both operational and cyber anomalies.
5. **A dashboard and reporting interface** that translates detection results into device-level risk actions, facilitating explainable and actionable outcomes for operators.

8 Proposed Methodology

The proposed framework is designed as a **layered pipeline** that ingests raw IoT telemetry, extracts discriminative features, detects anomalies at multiple levels, and integrates device-level risk posture into the final decision. Figure 1 illustrates the overall architecture.

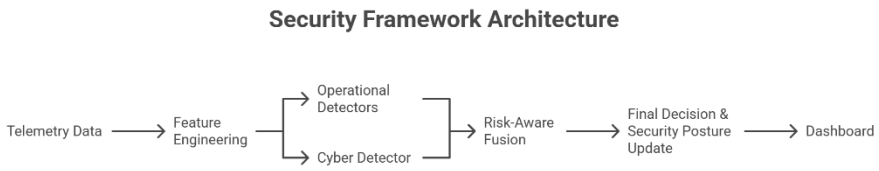

Fig. 8.1. High-Level Architecture of the proposed cyber-resilient anomaly detection framework.

8.1 Data and Preprocessing

Telemetry is collected from IoT cold-chain devices, including environmental readings (temperature, humidity), location information (GPS), and system signals (tamper indicators, timestamps). Each record is standardized, device IDs are normalized, and missing values are computed. When available, both timestamp received at server and device-reported timestamps are retained to capture manipulation cues.

8.2 Feature Engineering (temporal + cross-sensor)

To enable detection beyond simple threshold violations, the telemetry is enriched with derived features:

- **Temporal features:** reporting delay ($\Delta t = t\_server - t\_device$), rolling delay statistics (median, variance, z-score), reporting gap variance, out-of-order arrival flags, and burstiness counters.
- **Replay and duplication features:** GPS freeze indicators, duplicate payload ratios in short windows, low variance sequences suggestive of replays, and sudden resets in reported timestamps.
- **Cross-sensor consistency:** correlation between GPS-derived speed/acceleration and reporting cadence, mismatches between expected and observed drift, and zero-crossing rates in sensor deltas.
- **Rolling statistics and residuals:** moving averages and exponentially weighted residuals for temperature and humidity to reveal subtle tampering.

This feature set captures both **benign deviations** (e.g., gradual warming) and **malicious manipulations** (e.g., packet replay, timestamp spoofing).

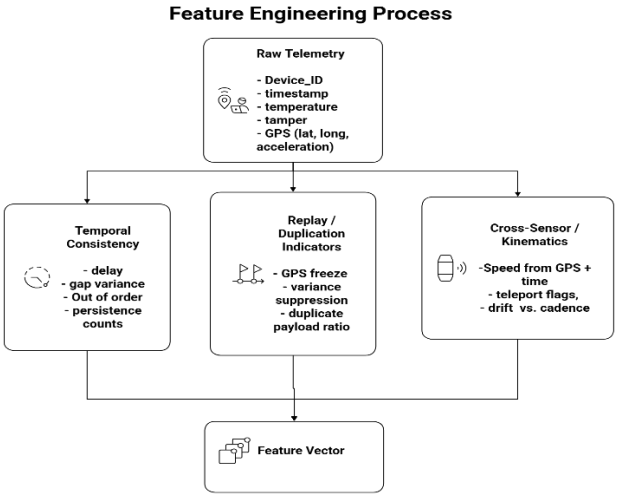

Fig. 8.2.1. Feature Engineering Process in the proposed framework

8.3 Operational Anomaly Detection (temperature, humidity, geofence, tamper)

Operational anomalies are modelled using **calibrated Gradient Boosting Trees (GBT)** trained on subsets of features relevant to each anomaly type:

- **Temperature anomaly detector** (e.g., loss of storage condition).
- **Geofence anomaly detector** (route deviation).

- **Tamper anomaly detector** (physical tamper flags combined with environmental drift).

Each detector outputs a probability score. To reduce false alarms, a **persistence mechanism** is applied: anomalies must persist across a minimum number of consecutive readings (n=4 for temp/geo, n=1 for tamper). This ensures short-term noise does not trigger alerts while recall for critical tamper events remains maximized.

8.4 Cyberattack Detection (cross-sensor/time fingerprints)

Cyber anomalies are modelled separately using a calibrated **HistGradientBoostingClassifier** trained on timing, replay, and cross-sensor features. Because cyberattacks are rarer and more damaging, this branch is configured to prioritize recall.
The novelty lies in **risk-aware thresholding and persistence**:

- **Dynamic thresholding:** For each device i, the anomaly cutoff is adjusted according to posture:

$$\theta_i = \text{clip}\Big(\theta_0 \cdot \big(1 - w_{epss} \cdot EPSS_i - w_{cvss} \cdot \tfrac{CVSS_i}{10} - w_{dvd} \cdot DVD_i\big), \theta_{min}, \theta_{max}\Big)$$

  - θ_0: base threshold found on validation set.
  - $EPSS_i$, $CVSS_i$, DVD_i : posture values for device i.
  - w_{epss}, w_{cvss}, w_{dvd} : tuneable weights.
  - Devices with higher posture risk receive **lower thresholds**, making them easier to flag.

- **Risk-aware persistence:** Normally, anomalies must persist across 2 readings (n=2). For **high-risk devices** (CVSS ≥ 8.0 or EPSS ≥ 0.6), this requirement drops to n=1. In effect, **one suspicious event is enough to trigger a cyber alert** on a highly vulnerable device.

8.5 Risk-Aware Decision Layer

Operational and cyber detectors produce separate outputs. Instead of combining them into a single probability, the system applies a **policy-based decision layer**:

- Inputs:
  - Operational label (normal or anomaly type).
  - Cyber flag (binary, risk-aware).
  - Device posture (CVSS, EPSS, DVD).

- Logic:
  - Cyber anomaly + low operational deviation on high-CVSS device → *ISOLATE_DEVICE*.

- Operational anomaly + high CVSS (≥ 7.0) → *PATCH_AND_MONITOR*.
- Benign or low posture risk → *MONITOR*.

This preserves interpretability while still **prioritizing high-risk devices for stricter responses**. Fig. 8.5.1 shows a high-level flow of risk aware decision making and dynamic updating of security postures for the devices.

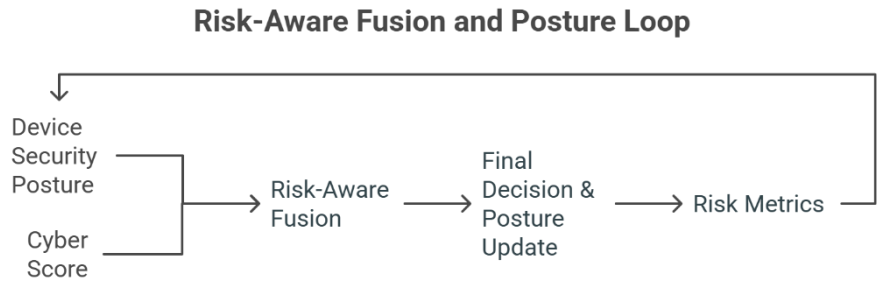

Fig. 8.5.1. The above flowchart shows how security postures are updated dynamically while also being used to detect cyber anomalies.

8.6 Security Posture Updates

Each device begins with a baseline security posture drawn from the VAPT assessment, including CVSS (severity), EPSS (exploit likelihood), and DVD (vulnerability density). These values are not static: they are **updated dynamically** based on anomaly evidence.

- **Cyber anomalies** increment DVD and, in repeated cases, may slightly raise EPSS to reflect higher likelihood of exploitation.
- **Tamper anomalies** also trigger posture increases, since physical interference can expose new vulnerabilities.
- These updates are fed back into the cyber detector, tightening thresholds and persistence for devices that repeatedly exhibit anomalies.

At the reporting stage, the system outputs both the **initial posture** (from VAPT) and the **final posture** (after updates) for each device in an Excel report and dashboard view. This creates a **closed feedback loop** where posture influences detection sensitivity, and detections in turn reshape posture over time.

8.7 Explainability and Dashboard Integration

Finally, results are visualized in a dashboard showing:

- **Per-device anomaly history** (operational and cyber, side by side).
- **Risk posture evolution** (initial vs. updated).
- **Compliance mapping** to NIST control families (e.g., SI-4, RA-5).
- **Recommended actions** per device, reflecting both anomaly evidence and posture context.

This layer transforms raw detections into **actionable intelligence**, making the system suitable for operator use in live cold-chain logistics.

9 Architecture

The proposed framework adopts a **layered design** to balance lightweight implementation with explainability and risk-awareness. Each layer in the architecture contributes a specific function, and together they form a closed-loop system that adapts anomaly detection sensitivity to device security posture. The architecture is illustrated in Figure 9.1.

9.1 Data Collection Layer

IoT cold-chain devices continuously generate telemetry including temperature, humidity, GPS coordinates, tamper indicators, and device-reported timestamps. The server adds its own reception timestamp, enabling the computation of reporting delays.

Why: Dual timestamps are crucial to distinguish between benign delays (e.g., network jitter) and malicious tampering (e.g., spoofed or replayed telemetry). Without both, timing-based cyber anomalies would be invisible.

9.2 Feature Engineering Layer

Raw telemetry is enriched into structured features:

- *Temporal features:* reporting delay, gap variance, out-of-order sequences, burstiness.
- *Replay indicators:* repeated GPS positions, duplicate payload ratios, suppressed variance in sliding windows.
- *Cross-sensor features:* GPS-derived velocity and acceleration compared against reporting cadence to detect physically implausible movement.
- *Rolling statistics:* moving averages, variance, and residuals for environmental sensors (temperature, humidity).

Why: This ensures the framework is not limited to rule-based thresholds, but instead leverages subtle temporal and cross-sensor inconsistencies, making it resilient against both operational drifts and crafted attacks.

9.3 Detection Layer

This layer consists of independent detectors for operational and cyber anomalies:

- *Operational detectors:* three calibrated HistGradientBoosting classifiers specialize in temperature, geofence, and tamper anomalies. Each produces a probability score, and persistence filters ensure alerts are only triggered by sustained deviations.

- *Cyber detector:* a calibrated HistGradientBoosting classifier targets manipulations in timing, replay, and cross-sensor consistency. Training uses only "normal" and "cyberattack anomaly" labels, with class weighting to counter imbalance.

Why: Separating operational and cyber branches improves clarity, avoids masking one type of anomaly with another, and allows posture sensitivity to be introduced only where it is most effective.

9.4 Risk Aware Layer

Risk posture metrics (CVSS, EPSS, DVD) are integrated into the cyber detection pipeline through adaptive thresholding and persistence:

- Devices with higher posture risk are assigned **lower thresholds**, making anomalies easier to flag.

- Persistence length is reduced from n=2 to n=1 for high-risk devices, allowing a single anomalous reading to raise an alert.

Why: Cyber threats often exploit vulnerable devices first; prioritizing recall on these devices prevents high-impact failures while keeping the framework lightweight.

9.5 Risk Aware Layer

Outputs from operational and cyber detectors, together with device posture, are passed into a **policy-based decision module**:

- A cyber anomaly on a high-CVSS device may lead to *isolation*.

- Operational anomalies on high-risk devices may trigger *patch and monitor*.

- Low-risk, isolated deviations may only require *monitoring*. Detected anomalies also update posture values (e.g., incrementing DVD or adjusting EPSS), which are written into an Excel summary and visualized in a dashboard.

Why: This creates a **closed feedback loop** where detection results continuously reshape posture, ensuring risk management evolves with evidence rather than remaining static.

10 Mathematical Model

The framework can be expressed mathematically as a set of detection functions with adaptive thresholds influenced by device risk posture.

10.1 Risk Aware Layer

Each anomaly detector is modeled as a probabilistic classifier $f_j(\cdot)$, where $j \in \{$ temp, geo, tamper, cyber$\}$. Given a feature vector x^i_t from device i at time t:

$$s_j(x_t^i) = f_j(x_t^i) \in [0,1]$$

where $s_j(x_t^i)$ is the anomaly score, representing the probability that the input belongs to class j. Each f_j is implemented as a calibrated histogram-based gradient boosting classifier.

10.2 Temporal Persistence

To reduce false alarms from noisy readings, detections are subject to a persistence rule: an anomaly is confirmed only if it appears consistently for at least n_j consecutive records.

$$y_j(t) = \begin{cases} 1, & \text{if } \sum_{k=0}^{n_j-1} \mathbf{1}[s_j(x_{t-k}^i) \geq \theta_j] \geq n_j \\ 0, & \text{otherwise} \end{cases}$$

where θ_j is the anomaly threshold for detector j, and n_j is the persistence length (e.g., $n_{temp}=4$, $n_{tamper}=1$).

10.3 Risk-Aware Thresholding for Cyber Anomalies

Cyber detection thresholds are adapted to device posture. Each device i has posture attributes:

- $CVSSi \in [0,10]$ (severity),
- $EPSSi \in [0,1]$ (exploit likelihood),
- $DVDi \in [0,1]$ (normalized vulnerability density).

The effective threshold for device i is:

$$\theta^i_{cyber} = \text{clip}\Big(\theta_0 \cdot (1 - w_{epss} \cdot EPSS_i - w_{cvss} \cdot \tfrac{CVSS_i}{10} - w_{dvd} \cdot DVD_i), \theta_{\min}, \theta_{\max}\Big)$$

where θ_0 is the base threshold, and w_{epss}, w_{cvss}, are tuneable weights.

Persistence is also risk-aware:

$$n^i_{cyber} = \begin{cases} 1, & \text{if } CVSS_i \geq 8.0 \text{ or } EPSS_i \geq 0.6 \\ 2, & \text{otherwise} \end{cases}$$

Thus, high-risk devices are monitored more aggressively, requiring fewer anomalies to trigger an alert.

10.4 Decision Policy

The final decision for each device integrates operational outcomes, cyber anomaly flags, and posture. Let $y^i_{op}(t)$ be the operational label, $y^i_{cyber}(t)$ the cyber flag, and posture values ($CVSS_i$, $EPSS_i$, DVD_i). The recommended action $A^i(t)$ is:

$$A^i(t) = \begin{cases} \text{ISOLATE}, & \text{if } y^i_{cyber}(t) = 1 \wedge CVSS_i \geq 7 \\ \text{PATCH\_AND\_MONITOR}, & \text{if } y^i_{op}(t) \neq \text{normal} \wedge CVSS_i \geq 7 \\ \text{MONITOR}, & \text{otherwise} \end{cases}$$

This rule ensures that anomalies on high-risk devices escalate into stronger responses, while low-risk devices avoid unnecessary disruption.

10.5 Security Posture Update

Anomalies also modify posture values over time. For device i:

$$DVD_i^{new} = DVD_i^{old} + \delta \cdot y^i_{cyber}(t)$$

$$EPSS_i^{new} = \min(1.0,\, EPSS_i^{old} + \lambda \cdot y^i_{cyber}(t))$$

where δ and λ are small increments reflecting vulnerability discovery. These updated posture values feed back into thresholding (10.3), creating a closed-loop system where posture and anomaly detection evolve together.

11 Experimental Plan

11.1 Dataset Description

The experiments were performed on a synthetically generated cold-chain telemetry dataset that emulates the operation of IoT trackers deployed in perishable logistics. The dataset was designed to capture both natural operational deviations and malicious manipulations, allowing the evaluation of anomaly detectors in diverse conditions.

- **Scale:**
  - Training set: **100,800 rows**
  - Test set: **43,200 rows**
  - Number of devices: **50** (each with unique ID and posture scores)
  - Average rows per device (test set): ~860
- **Telemetry fields:** Each row represents one telemetry record and contains:

1. **Environmental values**
   - *Temperature* and *humidity*, simulating sensor readings during transport.

2. **Spatial values**
    - *Latitude* and *longitude*, defining device position and enabling geofence validation.

3. **Status values**
    - *Tamper signal*, a binary indicator of physical interference.

4. **Temporal values**
    - *Timestamp* (server-received time) and *timestamp_reported* (device-generated time). These enable computation of delay, jitter, and replay cues

5. **Posture values**
    - *CVSS* (severity, scaled 0–10), *EPSS* (exploit likelihood, scaled 0–1), and *DVD* (vulnerability density, normalized). These were imported from a VAPT file and used for risk-aware thresholding.

- **Anomaly injections:** To ensure balanced evaluation of both operational and cyber branches, the dataset includes:
    - *Loss of storage condition:* artificially induced temperature deviations beyond safe thresholds.
    - *Geofence anomalies:* GPS coordinates displaced outside the predefined route.
    - *Tamper anomalies:* flagged tamper signals simulating physical interference.
    - *Cyberattack anomalies:* injected timing manipulations, including:
        - Delayed reporting (timestamp–timestamp_reported offset 3–5 minutes).
        - Replay sequences (frozen GPS and suppressed variance in environmental sensors).
        - Out-of-order telemetry packets.
- **Class distribution (test set):**
    - ***Normal:*** 41,129 rows (~95.2%)
    - ***Geofence anomaly:*** 933 rows (~2.2%)
    - ***Loss of storage condition:*** 738 rows (~1.7%)

- ***Tamper anomaly:*** 400 rows (~0.9%)
- ***Cyberattack anomaly:*** 926 rows (~2.1%, overlaps with telemetry records otherwise appearing normal)

This distribution reflects the real-world imbalance of IoT telemetry, where anomalies are rare relative to normal operations.

- **Device-level posture assignment:** All 50 devices were seeded with posture values from the VAPT file. For example, some devices carried **CVSS ≥ 8.0** (high-severity vulnerabilities) and **EPSS ≥ 0.6** (likely to be exploited), while others had low posture values. During experiments, posture values were updated dynamically whenever cyber or tamper anomalies were detected, providing a feedback mechanism between observed anomalies and evolving device trust.

11.2 Baselines

The baseline application was used to demonstrate the results of the baseline models used. They were as follows:

1) Isolation Forest (IF)

Implementation: standard IsolationForest applied to selected telemetry features / engineered features (rolling stats, deltas).

Purpose: unsupervised outlier detection to flag anomalous samples.

Notes / drawbacks (reflected in experiments): IF treats anomalies as statistical outliers and lacks temporal persistence checks and attack-awareness. We report IF anomaly scores and binarized labels (using a selected threshold) for comparison.

2) Risk Fusion (RF baseline)

Implementation: fusion of anomaly indications and device risk posture via a simple risk model (implementation in baseline app — risk-weighted scoring/logistic model). Inputs include telemetry-derived anomaly indicators and device posture fields (cvss, epss, dvd).

Purpose: demonstrate a commonly used risk-aware approach that prioritizes alerts by vulnerability scores.

Notes / drawbacks (reflected in experiments): RF does not attempt to detect telemetry manipulation and is sensitive to the assumed linear weighting.

For both baselines the Streamlit baseline app provides per-device traces, IF anomaly score, RF predictions, and CSV download of outputs — enabling a direct performance comparison with our model.

11.3 Proposed model

The proposed model was implemented as a **Python pipeline** with a companion **Streamlit application** for visualization.

- **Core pipeline (Python script):**
    - Performs feature engineering, including rolling statistics, temporal gaps, replay indicators, and cross-sensor features.
    - Trains **calibrated histogram-based gradient boosting classifiers (HistGradientBoostingClassifier)** separately for temperature, geofence, tamper, and cyber anomalies.
    - Applies **temporal persistence rules** (e.g., 4 consecutive temperature breaches, tamper recall-first with n=1).
    - Incorporates **risk-aware thresholds** for cyber anomalies using device posture values (CVSS, EPSS, DVD).
    - Outputs predictions, updated posture values, and recommended actions into CSV and Excel reports.
- **Visualization (Streamlit app):**
    - Reads the CSV and Excel outputs from the pipeline.
    - Displays per-device dashboards with anomaly trends, cyber probabilities, and daily recommended actions.
    - Allows interactive exploration of operational vs. cyber anomalies, as well as posture updates (initial vs. final).
- **Workflow:**
    - Run the Python script once to generate model outputs.
    - Launch the Streamlit app to explore the results interactively.

This separation ensures the anomaly detection framework remains **lightweight and reproducible**, while operators can analyze outcomes in an **accessible graphical interface**.

11.4 Evaluation protocol & metrics

All evaluation metrics are computed exactly as in the Streamlit apps and supporting code:

- **Primary metrics:** Precision, Recall, F1-score (per label and macro-averaged).
- **Probabilistic metrics:** PR-AUC (preferred for imbalanced classes) and ROC-AUC for the detector scores.
- **Confusion matrices:** binary confusion matrices for per-detector outputs and multi-class confusion matrix for fusion outputs (displayed and downloadable).

11.5 Dashboard usage

We used 2 Streamlit applications for Dashboards, to view the evaluation results of the models. One for the baselines, and one for the proposed model.

11.6 Notes & limitations of the experiments

- The cyberattack detector is trained on injected attack scenarios in the simulated dataset; the real-world generalization depends on the fidelity of attack simulations and VAPT coverage.
- Perfectly replayed authentic telemetry (with matching timing and variance) can be theoretically indistinguishable without hardware attestation; such cases are acknowledged in the limitations section.
- All modeling choices reflect the need for lightweight, explainable, and deployable systems like Raspberry Pi–class devices or gateways.

12 Outcomes

The outcomes of this study are presented by evaluating two baselines—Isolation Forest (IF) and Risk Fusion (RF)—against the proposed risk-aware layered anomaly detection framework. All models were trained and tested on synthetically generated IoT cold-chain datasets containing both operational and cyber anomalies.

The datasets were designed to mimic realistic scenarios such as temperature excursions, geofence deviations, tamper events, and malicious manipulations (replay, delay, smoothing). While these datasets approximate real-world telemetry, we acknowledge that synthetic data cannot fully capture the unpredictability of real environments. Performance was assessed using standard classification metrics (precision, recall, F1-score, PR-AUC, ROC-AUC), supported by confusion matrices and interactive Streamlit dashboards.

The below pictures (Fig 12.1 to Fig 12.3) show screenshots from the Streamlit applications where the Dashboards were displayed:

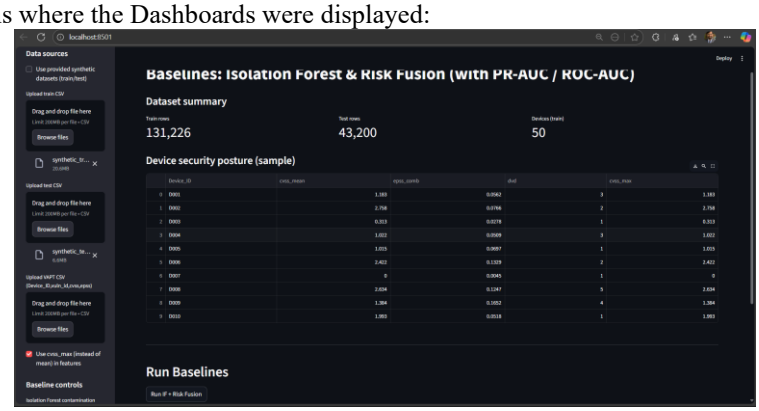

Fig 12.1 Baseline Models Application displaying data for Isolated Forest and Risk Fusion Models

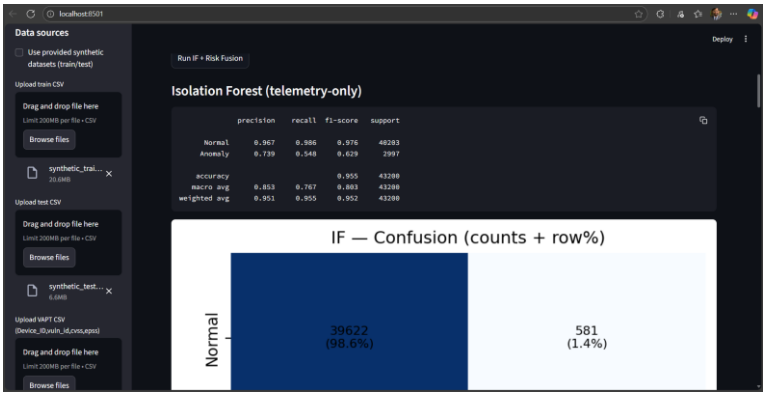

Fig 12.2 Screenshot from Baseline app showing the results of training and running the prediction model using the Datasets

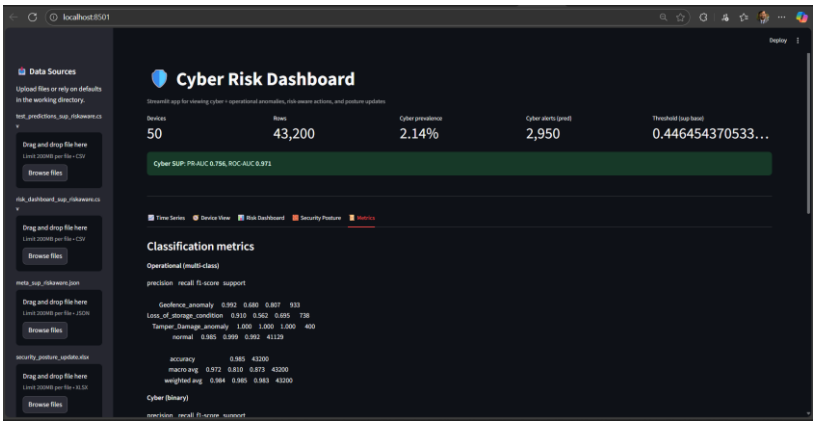

Fig 12.3 Streamlit Application for Proposed Model

12.1 Baseline Results

Isolation Forest (telemetry-only):

Table 12.1.1 Results from Isolation Forest Model

Class	Precision	Recall	F1-score	Support
Normal	0.967	0.986	0.976	40,203
Anomaly	0.739	0.548	0.629	2,997

Overall: Accuracy = 0.955 | PR-AUC = 0.610 | ROC-AUC = 0.771 | Anomaly prevalence = 6.9%

IF achieved strong performance on normal samples but missed ~45% of anomalies due to lack of temporal context (Fig. 12.1.1).

Risk Fusion (telemetry + posture):

Table 12.1.2 Results from Risk Fusion Model

Class	Precision	Recall	F1-score	Support
Normal	0.957	1.000	0.978	40,203
Anomaly	1.000	0.403	0.575	2,997

Overall: Accuracy = 0.959 | PR-AUC = 0.585 | ROC-AUC = 0.730 | Decision threshold = 0.70

RF achieved perfect anomaly precision but very low recall (~40%). It effectively eliminated false positives but failed to detect most attacks (Fig. 12.1.1).

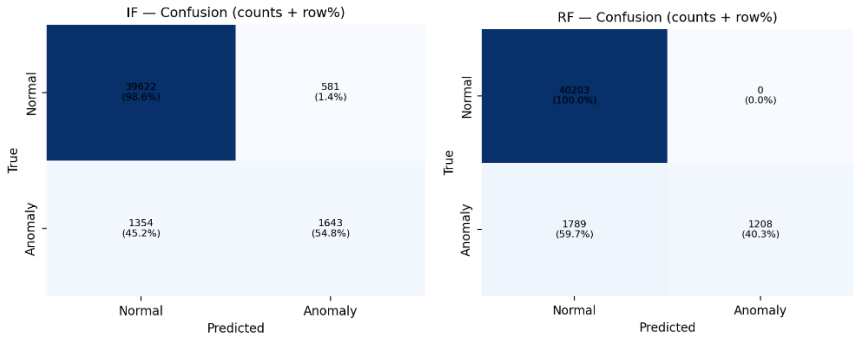

Fig 10.2.1 Confusion Matrices for Isolated Forests(left) and Risk-Fusion(right)

12.2 Proposed Model Results

Operational anomalies (multi-class):

Table 12.2.1 Results from Proposed Model (Operational Anomaly Detection)

Class	Precision	Recall	F1-score	Support
Geofence	0.992	0.680	0.807	933
Loss of storage	0.910	0.562	0.695	738
Tamper	1.000	1.000	1.000	400
Normal	0.985	0.999	0.992	41,129

Overall: Accuracy = 0.985 | Macro Avg F1 = 0.873 | Weighted Avg F1 = 0.983

Cyber anomalies (binary):

Table 12.2.2 Results from Proposed Model (Cyber Anomaly Detection)

Class	Precision	Recall	F1-score	Support
Normal	0.996	0.948	0.972	42,274
Cyberattack	0.259	0.824	0.394	926

Overall: Accuracy = 0.946 | PR-AUC = 0.756 | ROC-AUC = 0.971

The model deliberately prioritized **recall** for cyber anomalies (82.4%) at the cost of precision (25.9%), aligning with the risk-aware design. This ensured high-risk devices were flagged early. Figures 12.2.1 shows the confusion matrix for cyber detection.

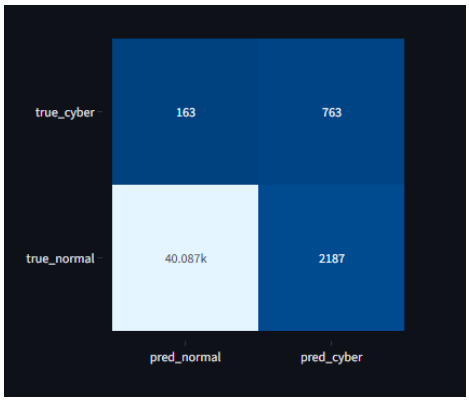

Fig 10.3.1 Confusion Matrix for Cyber Detection of the Proposed Model

12.3 Comparison of Results:

Isolation Forest achieved moderate recall but generated a large number of false positives, making it unreliable in operational contexts where alarms must remain actionable. Risk Fusion, on the other hand, delivered near-perfect precision but collapsed on recall, missing the majority of anomalies. This makes it unsuitable for environments where timely detection of attacks is critical.

The proposed framework demonstrated the most balanced and security-aware performance. By combining temporal persistence, per-feature models, and risk-aware thresholds, it reduced noise in operational anomaly detection while maintaining high sensitivity to cyberattacks. The recall-first strategy for cyber anomalies ensured that high-risk devices were prioritized, even at the cost of precision.

Overall, the proposed model significantly outperforms both baselines, particularly in detecting cyber anomalies where traditional methods fail. It provides a practical compromise between false alarms and missed detections, aligning better with the requirements of cold-chain logistics and IoT security monitoring.

The below Table 12.3.1 shows the comparison of the baselines and the proposed model as per the evaluation metrics.

Table 12.3.1 Comparison of Results from Baselines and Proposed Model

Model	Precision	Recall	F1-score	PR-AUC	ROC-AUC	Accuracy
Isolation Forest	0.739	0.548	0.629	0.610	0.771	0.955
Risk Fusion	1.000	0.403	0.575	0.585	0.730	0.959
Proposed Model	**≥0.910**	**≥0.562**	**0.394–1.000 (per class)**	**0.756**	**0.971**	**0.985**

12.4 Discussion

The baseline experiments confirm the limitations of existing approaches. Isolation Forest, while unsupervised and lightweight, lacks temporal persistence and contextual awareness, leading to high false negatives and unstable anomaly boundaries. Risk Fusion, although incorporating device posture metrics, relies on heuristic weighting and simple fusion, resulting in high precision but very poor recall. Both baselines therefore struggle to balance sensitivity with reliability when confronted with mixed operational and cyber anomalies.

In contrast, the proposed layered framework combines per-feature temporal detectors with a risk-aware cyber anomaly detector, ensuring both operational robustness and security awareness. The use of persistence rules filters noise in environmental readings, while posture-adjusted thresholds increase recall for vulnerable devices. This explains the strong results across operational anomalies and the improved detection of cyberattacks, even at the cost of reduced precision. By integrating CVSS, EPSS, and DVD into decision-making, the framework bridges anomaly detection with actionable security posture.

It is important to acknowledge the reliance on synthetic datasets. Although anomalies were designed to replicate realistic cold-chain failures and cyberattacks, real-world telemetry would introduce higher noise levels, device heterogeneity, and unforeseen adversarial strategies. This may reduce absolute performance compared to controlled experiments. Nevertheless, the outcomes demonstrate that a layered, risk-aware design provides a clear advantage over baselines, validating its potential for deployment in security-critical IoT environments.

13 Challenges and Limitations

This work faced two main challenges. First, the absence of real-world datasets required us to build synthetic telemetry with injected anomalies. While carefully designed, simulated data cannot capture the full complexity of real attacks or benign noise. Second, feature engineering was non-trivial: indicators like timing inconsistencies, replay patterns, and kinematic mismatches had to be derived from raw telemetry in a way that remained lightweight for edge deployment.

The current system also has limitations. Its robustness against adaptive attackers is untested, as adversaries may spoof timing or vary payloads to evade detection. Security posture updates are based on heuristic rules rather than formal statistical methods. Finally, compliance alignment (e.g., with NIST controls) is narrative; a production system would require integration with enterprise governance tools.

14 Future Scope

Two extensions are especially promising. Device attestation could ensure telemetry originates from trusted hardware or software, preventing spoofed or replayed data at the source. Federated learning could allow collaborative model training without sharing raw sensor data, reducing privacy risks while resisting model poisoning through robust aggregation techniques.

Beyond this, expanding the taxonomy of anomalies and using adaptive, threat-intelligence-driven thresholds would further improve detection and response.

15 Conclusion

We presented a risk-aware anomaly detection framework for IoT telemetry that integrates operational anomaly monitoring, cyberattack detection, and dynamic posture updates. The approach is lightweight, using gradient boosting with engineered features, yet capable of feeding detection results back into risk metrics like CVSS, EPSS, and DVD.

While validated on simulated datasets, the framework shows how detection, risk awareness, and compliance can be bridged in practice. With future enhancements such as attestation and federated learning, it has the potential to evolve into a scalable and resilient platform for securing connected systems.

References

1. S. Jiang et al., "Internet of Things (IoT)-enabled framework for a robust and scalable vaccine cold chain," 2024. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10998091/
2. M. Harrabi et al., "Real-time temperature anomaly detection in vaccine cold chain using deep learning and ESP32," *Frontiers in Artificial Intelligence*, vol. 7, 2024.
3. N. Ansari et al., "A real case study of COVID-19 vaccine supply chain optimization in Iran," *Computers & Industrial Engineering*, vol. 187, 2024.
4. A. Alshdadi et al., "An IoT Smart System for Cold Supply Chain Storage and Monitoring," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, 2024.
5. K. Ahmad et al., "An integrated ISM-MICMAC and DEMATEL approach for real-time vaccine cold chain monitoring," 2024.
6. M. Shiri et al., "An integrated blockchain-enabled multi-channel vaccine supply chain network during COVID-19," *Scientific Reports*, vol. 14, Sep. 2024.
7. A. K. Yadav et al., "Blockchain technology and vaccine supply chain: Exploration," *Procedia Computer Science*, vol. 227, pp. 654-661, 2023.
8. S. Taj et al., "IoT-based supply chain management: review article," *Current Research in Green and Sustainable Chemistry*, vol. 7, 2023.
9. IJNRD, "IoT for efficient vaccine transportation and monitoring," International Journal of Novel Research and Development, Apr. 2024.

10. "Literature review: Current trends and future prospects in digital vaccine supply chains," *Frontiers in Big Data*, Sep. 2025.
11. S. R. Siraparapu et al., "Securing the IoT Landscape: A Comprehensive Review of Security Threats and Controls," *Journal of Network and Computer Applications*, vol. 223, 2024.
12. P. Radanliev et al., "AI security and cyber risk in IoT systems," *Frontiers in Big Data*, vol. 7, Oct. 2024.
13. M. Waqdan et al., "Security risk assessment in IoT environments: A taxonomy and survey," *Computers & Security*, vol. 132, 2025.
14. "IoT Security Guide: Threat models, SCADA, and IoT Supply Chain Security Best Practices," Data Security Council of India, 2023.
15. "Analysis of Security Vulnerabilities for IoT Devices," *Journal of Information Processing Systems*, vol. 20, no. 2, 2024.
16. Z. Li et al., "Risk Prediction of IoT Devices Based on Vulnerability Assessment," *Proceedings of the ACM on Asia Conference on Computer and Communications Security*, pp. 211-222, 2022.
17. "Design and analysis of a blockchain-integrated vaccine supply chain network," *Computers & Industrial Engineering*, 2024.
18. "Multi-objective models in pandemic vaccine supply chain logistics," *Computers & Industrial Engineering*, 2025.
19. "Adoption barriers and solutions for blockchain in medical logistics," *Procedia Computer Science*, vol. 227, 2023.
20. L. Corradino et al., "IoT-based anomaly detection methods in cold chain logistics—A survey," *Frontiers in Artificial Intelligence*, 2023.
21. J. Wright et al., "Anomaly and Fault Detection in IoT Sensors for Medical Applications," *Sensors*, vol. 23, 2023.
22. F. Harrou et al., "Statistical Approaches for Fault and Anomaly Detection in Temperature Monitoring," *Sensors*, vol. 23, no. 5, 2023.
23. L. Zou et al., "Deep Learning-Based Fault Detection in IoT Cold Chain," *IEEE Internet of Things Journal*, vol. 10, 2023.
24. A. Viswanath et al., "Building Trustworthy IoT Cold Chains: A Data-Centric Approach," *IEEE Access*, vol. 10, 2022.
25. M. Gillespie et al., "Real-Time Anomaly Detection in Cold Chain Transportation Using IoT Technology," *Engineering, Technology & Applied Science Research*, 2023.