The Eastasouth Journal of Information System and Computer Science
Vol. 4, No. 02, December 2026, pp. 36-48
ISSN: 3025-566X, DOI: 10.58812/esiscs.v4i02

Capacity Aware Cyber Threat Detection and Safe Response for Critical Infrastructure

Celeste Yvonne Ramirez
BBW University of Applied Sciences, Berlin, Germany

Article Info

Article history:

Received Nov, 2026
Revised Nov, 2026
Accepted Nov, 2026

Keywords:

Alert Triage;
Critical Infrastructure;
Digital Twins;
Federated Learning;
Intrusion Detection;
Operational Technology

ABSTRACT

A cyber detector can classify industrial traffic accurately and still overwhelm the team responsible for investigating its alerts. An equally important failure occurs when a plausible detection triggers a response that interrupts a physical process. We develop an operational technology architecture that separates detection, incident review, and authorization to intervene. Sites score events locally, collaborate on model updates when privacy and availability requirements permit, and consult a read-only digital twin before proposing a control-sensitive response. An analytical stress test quantifies the interaction between incident prevalence, detector sensitivity, false-positive rate, and review capacity. At an assumed prevalence of 0.1%, sensitivity of 90%, and false-positive rate of 0.1%, a stream of 100,000 review units yields 190 expected alerts, including 90 true positives. A review budget of 100 units requires a false-positive rate near 0.010% to retain that sensitivity and accommodate all expected alerts. A second analysis shows that the usefulness of twin corroboration depends on its behavior specifically among detector-positive cases. The results are conditional planning estimates. A chronological, cross-site evaluation protocol specifies how to test whether the proposed operating points are attainable under incident clustering, telemetry loss, and changing plant conditions.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Name: Celeste Yvonne Ramirez
Institution: BBW University of Applied Sciences, Berlin, Germany
Email: celesteramirez.analytics@gmail.com

1. INTRODUCTION

During a shift change, maintenance laptop traffic can resemble reconnaissance. The same traffic on an unattended control segment may require immediate attention. The detector sees evidence for both interpretations, but the analyst must determine which one fits the operating state and whether a proposed intervention could interrupt service. This distinction matters in water treatment, electricity distribution, transportation, and building automation, where a compromised engineering workstation and an unnecessary loss of controller connectivity can both carry physical consequences.

Operational technology security guidance makes reliability, performance, and safety central to control selection. A response architecture therefore needs explicit authority boundaries between the analytics service and the systems responsible for operating the process [1].

A U.S. critical infrastructure study uses real incident examples, including Colonial Pipeline, to discuss how AI can support detection and response while leaving important governance and implementation questions open. Its incident discussion motivates a design that keeps operator authority explicit [2].

Intrusion detection research has long distinguished performance on an experimental traffic distribution from the demands of deployment. Among the reasons are an uncertain definition of abnormal traffic and the difficulty of collecting representative examples of attacks and legitimate behavior [3].

The research questions are whether a detector can meet an explicit analyst budget at realistic low prevalence, which conditions make independent process corroboration useful, and how to test both claims without granting the analytics stack control privileges. The paper contributes a gated architecture, a closed-form capacity analysis, and a prospective replay protocol that treats detection and intervention as separate decisions.

2. RESEARCH DESIGN AND EVIDENCE

The manuscript combines purposive evidence synthesis and an analytical experiment. The supplied bibliography provides the initial corpus across threat detection, distributed security, and digital twins. Supplemental sources anchor operational technology safety, cross-silo training, secure aggregation, and deployment-oriented intrusion detection. The synthesis uses the supplied citation records to identify design themes and uses accessible publisher text for specific findings. It is a targeted design synthesis with incomplete full-text coverage. Consequently, the paper neither pools benchmark performance across datasets nor infers study outcomes from titles.

A published Edge-IIoT experiment combined attention over flow features, an extreme learning machine, optimization of hyperparameters, and class-weighted focal loss. It reported binary F1 of 0.9942 and a multiclass macro-F1 of 0.9859. Those results motivate class-aware baselines and rare-attack reporting. Their scope is the dataset and evaluation used in that study [4].

A comparative study on textual threat classification reported stronger Transformer results than LSTM results under its shared preprocessing setup. An operational comparison must distinguish this text-classification task from chronological scoring of network flows. We therefore reserve text models for incident notes and intelligence enrichment unless a separate flow-based evaluation supports their use [5].

The experiment evaluates expected counts under stated assumptions. It uses a fixed exposure volume and specified event rates, so the outputs can be reproduced directly from the equations. The analysis measures workload implications and feasibility boundaries. Estimating field performance, uncertainty across sites, and intervention safety requires the independent replay and prospective studies described later.

3. ARCHITECTURE AND THREAT MODEL

The system has three trust domains. The operational site controls data collection and response authorization. A collaboration service coordinates model releases without receiving raw event streams in the federated configuration. A read-only twin service estimates process state and can reject an action proposal when its state estimate is stale or incomplete. A site can continue local monitoring when the collaboration or twin service fails.

The attacker can inject malformed or misleading telemetry, compromise a participating training client, delay messages, replay old updates, or manipulate the information used by the twin. The threat model assumes that safety interlocks and the independent operator authorization channel remain outside the analytics trust domain.

Simultaneous compromise of those controls is beyond the modeled protection boundary and must be addressed in the facility safety case.

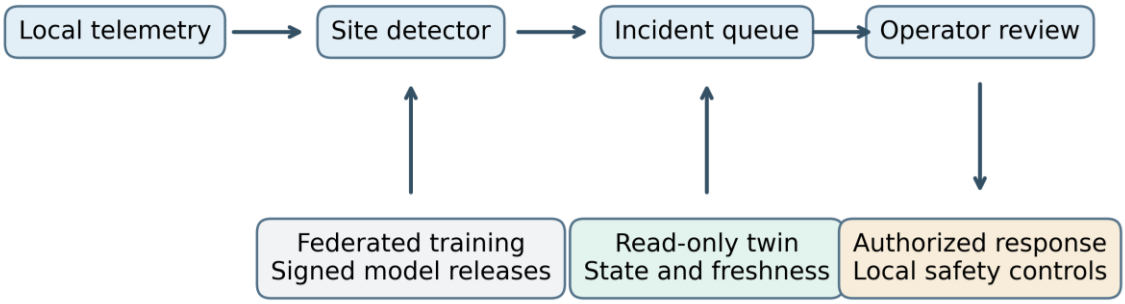


Figure 1. Trust Boundaries for Monitoring and Response
Note. The twin informs triage. The local operator and safety controls govern authorization.

3.1 Detection and Incident Construction

Each site collects passive network flows, authentication events, asset inventory snapshots, and process observations when those observations are available through approved interfaces. All records carry event time and arrival time. The scoring service builds feature from records that have arrived by the decision timestamp. It attaches missingness flags and rejects features derived from subsequent investigation or incident closure. It preserves asset identifiers locally and publishes aggregate evaluation counts only when disclosure rules permit.

The detector should begin with a signature baseline, regularized logistic regression, and a tree ensemble. A temporal model is evaluated only after these baselines establish the benefit of lagged observations. The incident builder groups candidate windows by site, asset group, and a prespecified time interval. It stores every constituent score and timestamp and counts the grouped incident as one review unit. Capacity planning uses the distribution after grouping. Evaluating window-level false-positive rates against an incident-level budget can produce a misleading estimate when benign activities generate bursts of correlated windows.

3.2 Collaboration and Response Authority

Federated averaging provides a reference training method that aggregates locally computed updates across participants. For this architecture, it is a baseline against which to measure the value of collaboration across sites with unequal traffic profiles and uneven attack exposure [6]. Research on real-time cloud-native processing for federated learning motivates measuring the time from local update submission to release of the aggregated model under variable network and compute loads [7].

Secure aggregation can reveal an aggregate of client vectors while shielding individual updates under the protocol assumptions. The protocol is relevant to training confidentiality. Authorization, model poisoning, inference about contributors from repeated aggregates, and release rollback require additional controls [8].

The coordinator signs a release manifest with the training round, participant policy, feature schema, and model hash. Sites check those fields and run a local acceptance replay before promotion. A new round needs a minimum cohort size, an approved model version, and a deadline. Sites continue serving the last accepted model during dropouts. Clipping and robust aggregation must be designed together with confidentiality. A server that sees

only an aggregate cannot directly verify the norm of each hidden update. A deployable design must specify whether it relies on attested clients, a different cryptographic protocol, or a different visibility model.

The twin supplies a time-stamped estimate of process state and a separate measure of its freshness. A twin reading derived from the same compromised sensor path as the detector carries shared failure risk. When the provenance check fails, the interface marks corroboration unavailable and routes high-consequence incidents for review. Permitted recommendations range from requesting another observation to isolating an enterprise endpoint through an already authorized playbook. Any control-affecting action remains subject to an independent operational check, a valid approval, and a rollback plan.

Table 1. Decision Rights at Each Stage

Stage	Required evidence	Decision right
Candidate scoring	Schema version and event availability	Site detection service
Queue admission	Grouped incident, score and provenance	Site triage policy
Corroboration	Fresh process estimates with traceable inputs	Read-only twin service
Control-sensitive action	Interlock checks and documented approval	Authorized site operator
Model promotion	Signed manifest and local replay acceptance	Site release owner

4. ANALYTICAL METHODS

4.1 Capacity Constraints

Let N denote the number of grouped review units per period, π the attack prevalence per unit, r the true-positive rate, and f the false-positive rate. Expected true alerts equal $N\pi r$, expected false alerts equal $N(1 - \pi)f$, and the total is the sum. Precision in this mean-rate calculation is πr divided by $[\pi r + (1 - \pi)f]$ when expected alerts are positive. The benchmark assumes constant rates within the analysis period. The expectation holds for correlated units with common marginal rates, but correlation affects day-to-day queue variation.

The base-rate problem in intrusion detection explains why benign traffic can account for most alerts when attacks are rare [9]. Here, the same relationship is expressed as a workload constraint by specifying how many grouped incidents analysts can review in a period.

For capacity B, the largest feasible false-positive rate at a fixed sensitivity is $f_{max} = (B/N - \pi r)/(1 - \pi)$, assuming $0 \le \pi < 1$ and a nonnegative numerator. If expected true alerts alone exceed capacity, the specified sensitivity and review-all rule are infeasible. The feasible bound is a design requirement for a score threshold, not a claim that the receiver operating characteristic contains that point. In practice the threshold is selected on a calibration period, then evaluated on a locked test period.

The base scenario uses N = 100,000, $\pi = 0.001$, r = 0.90, and B = 100 per period. False-positive rates range from 0.00005 to 0.01. To assess the sensitivity of staffing requirements to changes in incident frequency, the stress test varies π from 0.0001 to 0.002. Prevalence represents the fraction of grouped units containing an adjudicated attack. A deployment should estimate it from a sample of both alerted and non-alerted units because routine incident tickets omit attacks that the system never escalates.

4.2 Corroboration and Action Loss

Define $a = P(C = 1 \mid D = 1, Y = 1)$ and $b = P(C = 1 \mid D = 1, Y = 0)$, where D indicates detector alert, C indicates usable positive corroboration, and Y

denotes the attack label. A corroboration requirement for an accelerated review lane retains overall sensitivity ra and false-positive rate fb. Estimating a and b conditional on detector alerts avoids an independence assumption between the two channels. Low values of b favor the accelerated lane, while low values of a can delay serious incidents. All detector alerts remain eligible for ordinary review, with availability and urgency flags.

If intervention has a false-action loss Lfalse and waiting has a missed-action loss Lmiss, a simplified decision compares (1 − q)Lfalse with qLmiss for calibrated attack probability q. The resulting cutoff is Lfalse/(Lfalse + Lmiss) under zero correct-decision loss and fixed costs. This threshold illustrates the effect of asymmetric losses. Site interlocks, operator approval, and state validity remain binding regardless of the estimated probability. Cost estimates should distinguish detection delay from the harm that an intervention can itself cause.

A retail banking study provides a cross-domain example of separating probability calibration from cost-sensitive threshold selection using a leakage-safe pipeline [10]. For the proposed operational technology evaluation, fit the calibration map on a chronological calibration period and assess both calibration and action thresholds on the locked test period.

5. RESULTS

Table 2. Expected Alert Load at 0.1 Percent Attack Prevalence

False-positive rate	True alerts	False alerts	Total alerts	Precision
1.000%	90.00	999.00	1089.00	8.3%
0.100%	90.00	99.90	189.90	47.4%
0.010%	90.00	9.99	99.99	90.0%
0.005%	90.00	5.00	95.00	94.7%

At a false-positive rate of 0.1%, the expected queue contains 189.9 alerts, or about 1.90 times the review budget. Expected precision is 47.4%. At 0.01%, the same sensitivity produces 99.99 expected alerts with approximately 90.0% precision. The difference is driven by the number of benign units exposed to even a small false-positive rate. Expected counts are fractional because the analysis describes repeated periods with the same rates.

At the base operating point, the capacity bound is approximately 0.01001%. Raising prevalence to 0.2% at the same sensitivity yields 180 expected true alerts, which exceeds the budget even at zero false positives. Under that surge, a team must distinguish queue admission from the staffing assumption. Preserving incident visibility while deferring low-urgency review is a separate operational policy that needs measured delay and escalation limits.

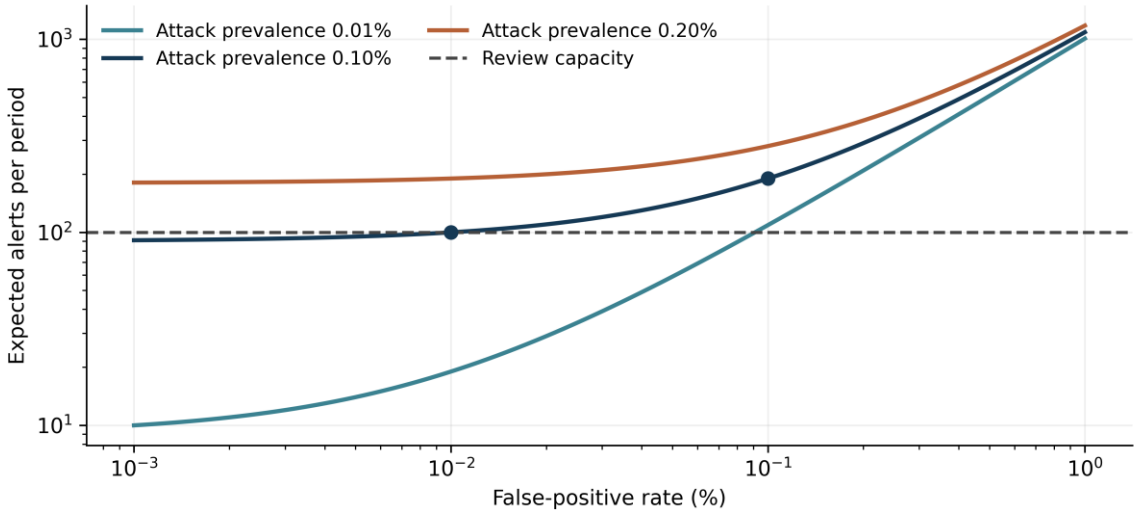


Figure 2. Expected Queue Load Under Three Prevalence Scenarios
Note. Each curve assumes 100,000 review units and 90% sensitivity. Values are analytical scenarios.

Table 3. Corroboration Scenarios at the Base Detector Operating Point

Conditional a	Conditional b	True in fast lane	False in fast lane	Fast-lane precision
90%	10%	81.0	10.0	89.0%
90%	50%	81.0	50.0	61.9%
50%	10%	45.0	10.0	81.8%
50%	50%	45.0	50.0	47.4%

At a = 0.90 and b = 0.10, the accelerated lane contains 81 expected true cases and about 10 expected false cases, with precision near 89.0%. At the same a and b = 0.50, precision falls to 61.9% and the expected accelerated lane grows to about 131 cases. A low a makes the lane look manageable by retaining fewer true cases. Under a = 0.50 and b = 0.10, 45 of the 90 detector-positive attacks stay outside the accelerated lane. These cases expose the trade-off between queue size and delay even when accelerated review looks efficient.

6. INTERPRETATION AND DEPLOYMENT REQUIREMENTS

The capacity result changes what counts as a useful comparison. A model that improves sensitivity at a false-positive rate above the feasible bound may still provide value for retrospective hunting. At that same operating point it cannot support review of every alert within the staffing budget. The evaluation therefore compares methods at a matched review burden and reports the operating points each model can actually achieve.

The twin should be judged against a simple process rule using the same inputs and timestamps. A complex twin can help where state depends on delayed observations or interactions between assets. Its complexity creates additional failure modes, including stale estimates, schema drift, and an apparently plausible state reconstructed from corrupt inputs. When corroboration is unavailable, a separate unknown state avoids silently equating a missing measurement with evidence of normal operation.

6.1 Governance and Intelligence

Work on cyber-economic resilience through AI and management information systems motivates tracking service disruption alongside detection metrics. Hours of unavailable service and deferred maintenance can reveal burdens hidden by alert counts [11].

Research on U.S. supply-chain exposure motivates including supplier

remote-access paths and downstream dependencies in the impact register [12].

Work on national critical-infrastructure threat response informs the choice to distinguish internal escalation from notification to an external partner [13].

The governance framework for information-system cyber risk motivates recording the owner, approval, and expiry of every accepted detection exception [14].

Work on big-data analytics within management information systems informs the use of a common incident record that links technical observations with triage decisions [15].

Research on AI and data-system security motivates evaluating the security of the feature pipeline alongside the attacks that the resulting model is expected to detect [16].

Work on AI and cloud integration for cyber defense motivates separating the central coordination service from local execution [17].

A strategic treatment of advanced threats motivates periodic revision of threat scenarios as adversary behavior and plant connectivity change [18].

Work on AI in cyber risk management for IT projects informs the requirement to assign an owner to each mitigation and record its closure [19].

Threat intelligence framed as a management information system informs the routing of incident evidence to both security operations and risk oversight [20].

Research on employee cybersecurity training motivates tabletop evaluation of analyst and operator handoffs [21].

National-scale threat intelligence research informs the use of a shared incident vocabulary for cross-site reporting [22].

Work linking supply chains, energy resilience, and infrastructure protection motivates documenting consequences beyond the initially affected asset [23].

6.2 Detector Extensions

Temporal threat prediction motivates testing whether previous windows improve lead time after features are aligned to their actual arrival timestamps [24].

AI-enhanced detection in management information systems motivates comparing the learned detector with existing alert rules under the same incident grouping policy [25].

Research on self-healing through reinforcement learning motivates a future comparison in an isolated replay or simulation where the effect of unsafe action proposals can be measured [26].

Automated vulnerability assessment motivates asset-aware triage. A known exposure can raise investigation priority without supplying the attack label [27].

Continual learning for insider threats motivates an update schedule with delayed-label replay, version control, and an explicit rollback test [28].

Adversarial learning research motivates stress tests with evasion and poisoning rather than restricting evaluation to clean examples [29].

Work on adversarial biometric verification motivates a separate test for identity-signal manipulation wherever those signals enter the incident score [30].

Binary-image malware classification motivates treating file evidence as its own signal with separately measured latency and provenance [31].

Research on smart-building sensor security motivates including building automation as a transfer setting with distinct maintenance and occupancy cycles [32].

Real-time cyber analytics motivates measuring end-to-end latency from the first available observation through incident admission [33].

Security analytics for live threat identification motivates distinguishing

an early provisional warning from an incident ready for review [34].

6.3 Confidential Collaboration and The Twin

Cryptographic designs for resilient information systems motivate explicit key ownership and documented recovery after key loss [35].

Cloud data-protection research motivates testing cross-tenant access controls separately from detector quality [36]. Work on policy-bound encryption and verifiable access provenance motivates testing whether the storage layer blocks unauthorized cross-tenant access and retains a verifiable record of authorized access to shared artifacts [37].

Federated intrusion detection with secure multiparty computation motivates a baseline that tests aggregation confidentiality and dropout recovery at the same client counts [38].

Homomorphic federated modeling motivates measuring the cost of protected arithmetic as part of the training cycle [39].

Work on blockchain-based access auditing motivates comparing a shared ledger with signed append-only records when multiple organizations require independent verification [40].

Decentralized integrity mechanisms for management information systems motivate linking audit records to artifact hashes and authorization decisions [41].

Privacy-aware cloud data protection motivates documenting how much metadata, including message timing and client availability, remains visible [42].

Research on secure and energy-efficient data centers motivates reporting compute and network cost per accepted training round [43].

Homomorphic cloud analytics research informs the evaluation of ciphertext expansion, numerical precision, and latency when protected computation is included [44].

Cloud-integrated information-system security motivates testing identity boundaries between tenant administrators and the shared coordinator [45].

Auditable federated threat detection motivates requiring a release manifest that an independent reviewer can reconcile with site acceptance records [46].

Self-healing hybrid-cloud analytics motivates rollback and recovery drills when a faulty model is accepted by a subset of clients [47].

Work integrating federated learning, blockchain, and zero trust motivates evaluating each control against its own threat assumption and recording the resulting operational overhead [48].

Hybrid-cloud microservice threat prediction motivates including transient service identities and workload redeployments in the benign traffic replay [49].

Research on distributed-system resilience motivates separating the failure of model coordination from the failure of local scoring [50].

Federated threat intelligence in global cloud information systems motivates testing client populations with different attack classes, maintenance cycles, and label delays [51].

Secure semantic communication research motivates retaining provenance and timing when compact event representations cross an edge link [52].

Trust and optimization research in digital twins motivates treating state confidence and input integrity as separate fields [53].

Zero-touch communications research motivates evaluating automated network adjustments in an isolated environment before exposing action proposals to site operators [54].

Network survivability and traffic-flow analysis motivate checking whether an isolation proposal would

remove the last viable path for a critical service [55].

Research on standardized secure edge intelligence for digital twins motivates a versioned interface for state estimates, timestamps, and failure codes [56].

Work on twin-enabled energy optimization motivates including energy and service effects in the consequence analysis when a response changes the operating regime [57].

7. PROSPECTIVE EVALUATION PROTOCOL

7.1 Cohort Construction and Leakage Controls

Begin with a public industrial traffic benchmark for engineering validation and a second replay stream that combines process-state measurements with incident annotations. Use each public corpus within its own documented scope. To estimate site transfer, collect at least two independent facilities or testbeds with comparable process semantics and distinct capture environments. Preserve all raw captures and time conversion rules under access controls. Publish de-identified event counts and splitting code where agreements permit.

The Secure Water Treatment (SWaT) testbed provides a candidate for the initial replay because it records both network traffic and physical process measurements under normal and attack conditions [58]. Its scripted attacks and single testbed environment limit inference about incident prevalence and cross-site transfer.

Partition each site chronologically into training, calibration, and locked test periods, using a gap at least as long as the largest lookback window. Keep traffic from the same campaign, capture session, or incident in one split. Remove artifact fields such as capturing filenames and explicit attack names. At feature extraction, enforce both the time of the underlying event and the time at which the data entered the monitoring system. Repeat the evaluation with one site held out entirely. Report whether any unlabeled adaptation was allowed and reserve a later period for testing after adaptation.

Adjudicators should use process and security records and document uncertainty. Sample detector-negative units for review and preserve their inclusion probabilities so site-level false-negative estimates can be weighted. Record the time at which evidence became sufficient to classify an incident. Some apparent false positives may be unresolved. Report sensitivity to treating unresolved cases as benign or unknown and retain the primary analysis definition across all models.

7.2 Comparisons and Uncertainty

Compare a signature system, regularized logistic regression, a tree ensemble, and a temporal model using identical information cutoffs. Compare local-only training with federated averaging and an approved centralized benchmark if data-sharing agreements permit. To isolate the effect of the twin, compare model-only triage, a simple corroborating process rule, and the same model with twin-based triage. Maintain identical review capacity and count each grouped incident once. Record the fraction of the queue that remains unreviewed at the deadline, even when all cases were correctly scored. Research on security screening in cloud workload scheduling motivates measuring deadline misses by tenant at the shared coordinator to see whether screening delays fall disproportionately on particular tenants [59].

The primary endpoint is incident recall within the review budget and a prespecified time horizon. Secondary endpoints include precision at capacity, false alerts per site-day, median and 95th-percentile review delay, calibration error, expected workload, availability of corroboration, and proposals blocked by

safety rule. Report these metrics by attack class, site, process mode, and telemetry availability. Estimate uncertainty with paired block resampling at the site-day or campaign level. With few sites, report site-specific results and describe the limitations of any pooled interval.

Report the precision–recall curve and average precision alongside the fixed-budget endpoint. Precision–recall analysis helps interpret performance on imbalanced data because it makes the proportion of true alerts among positive predictions visible across thresholds [60]. Report attack prevalence for each test cohort so that differences in precision can be interpreted in context.

To investigate failure modes, delay telemetry, replay stale updates, introduce a poisoned training participant, and disconnect the twin during high-load periods. Test each condition separately before combining failures. Record how the system signals degraded confidence, which version remains active, and how long it takes to restore normal operation. Record both event time and completion time so the test captures waiting for data as well as model inference. Any automated action should first run in shadow mode. A site operator can then compare suggested decisions with actions actually taken and document why they differ.

Table 4. Evaluation Endpoints and Analysis Units

Question	Endpoint	Unit and comparison
Can the queue fit the shift?	Recall at fixed review capacity	Grouped incidents per site-shift
Does corroboration improve triage?	Fast-lane precision and attack delay	Same detector, paired replay
Does collaboration improve transfer?	Recall and calibration on held-out sites	Local versus federated training
How does the system degrade?	Unavailable corroboration and recovery time	Paired telemetry-loss scenarios
Does the action gate enforce authority?	Unsafe proposals blocked and audit completeness	Shadow-mode response proposals

8. LIMITATIONS

The analytical experiment uses fixed rates and a common review cost. Real shifts vary in attack prevalence, investigator availability, incident severity, and review duration. The capacity bound is a mean requirement. It cannot provide a service-level guarantee for a bursty queue. A prospective workload model should estimate the distribution of queue size and missed deadlines, with incidents clustered by campaign and maintenance activity.

The evidence corpus was supplied for design development. Source access varied, and some records contribute only topical context. The analytical results cannot establish a causal benefit of federated training or twin integration. Its corroboration parameters are scenarios, and the results would change if corroboration quality differed across classes or sites. Production deployment requires field measurements and a site-specific safety assessment.

9. CONCLUSION

At low prevalence, the benign traffic volume can consume most of the review budget even when the detector rarely raises false alarms. The proposed design makes the budget explicit and reserves intervention authority for site operators. A twin can help prioritize detector-positive incidents when it offers timely evidence with low conditional false corroboration. The next empirical question is whether a detector and corroboration policy can meet those

conditions on later periods and unfamiliar sites while preserving operational safety.

REFERENCES

[1] K. A. Stouffer *et al.*, "Guide to Operational Technology (OT) Security," 2023, *Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD*. doi: https://doi.org/10.6028/NIST.SP.800-82r3.

[2] N. Das *et al.*, "AI-driven threat detection and response framework for protecting U.S. critical infrastructure from cyberattacks," *Int. Cybersecurity Law Rev.*, vol. 7, no. 2, pp. 147–163, 2026, doi: https://doi.org/10.1365/s43439-026-00169-5.

[3] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," in *2010 IEEE Symposium on Security and Privacy*, 2010, pp. 305–316. doi: https://doi.org/10.1109/SP.2010.25.

[4] M. Shahiduzzaman *et al.*, "Attention-guided quantum-inspired extreme learning machine with grey wolf optimization for rare-attack detection in IIoT networks," *Discov. Artif. Intell.*, vol. 6, no. 1, p. 1258, 2026, doi: https://doi.org/10.1007/s44163-026-01704-3.

[5] J. Kaur *et al.*, "Comparative analysis of transformer and LSTM architectures for cybersecurity threat detection using machine learning," *EAI Endorsed Trans. AI Robot.*, vol. 4, pp. 1–12, 2025, doi: https://doi.org/10.4108/airo.9759.

[6] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, Pmlr, 2017, pp. 1273–1282. [Online]. Available: https://proceedings.mlr.press/v54/mcmahan17a.html

[7] M. Bashir, K. Syed, S. A. Aetukuri, M. Shukla, V. P. K. Kaluvakuri, and K. S. Mudaliyar, "Real-Time Cloud-Native Processing Techniques for Enhancing Federated Learning in Distributed Applications," in *2026 International Conference on Emerging in Computing Technologies (INCECT)*, 2026, pp. 1–7. doi: https://doi.org/10.1109/INCECT72370.2026.11742885.

[8] K. Bonawitz *et al.*, "Practical secure aggregation for privacy-preserving machine learning," in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191. doi: https://doi.org/10.1145/3133956.3133982.

[9] S. Axelsson, "The base-rate fallacy and the difficulty of intrusion detection," *ACM Trans. Inf. Syst. Secur.*, vol. 3, no. 3, pp. 186–205, 2000, doi: https://doi.org/10.1145/357830.357849.

[10] F. Karim *et al.*, "Calibrated, cost-sensitive, explainable ensemble learning for customer churn prediction in retail banking: a leakage-safe pipeline and single-dataset case study," *Front. Artif. Intell.*, vol. 9, p. 1889639, 2026, doi: https://doi.org/10.3389/frai.2026.1889639.

[11] R. M. Ahsan, B. Uddin, T. Hossen, and S. Das, "Resilient Intelligence: AI and MIS in the Cyber-Economic Era," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 3, no. 02, pp. 151–163, Oct. 2025, doi: https://doi.org/10.58812/esiscs.v3i02.758.

[12] M. A. Goffer *et al.*, "Cybersecurity and supply chain integrity: Evaluating the economic consequences of vulnerabilities in US infrastructure," *J. Manag. World*, vol. 2025, no. 2, pp. 233–243, 2025, doi: https://doi.org/10.53935/jomw.v2024i4.907.

[13] M. A. Goffer *et al.*, "AI-Enhanced Cyber Threat Detection and Response Advancing National Security in Critical Infrastructure," *J. Posthumanism*, vol. 5, no. 3, pp. 1667–1689, Apr. 2025, doi: https://doi.org/10.63332/joph.v5i3.965.

[14] S. N. Hasan, "Embedding cybersecurity risk management into information systems governance: A national-scale framework for organizational resilience, economic stability, and critical infrastructure protection," *Int. J. Appl. Math.*, vol. 39, no. 1s, pp. 237–254, 2026, doi: https://doi.org/10.12732/ijam.v39i1s.1623.

[15] S. N. Hasan *et al.*, "Enhancing cybersecurity threat detection and response through big data analytics in management information systems," *Fuel Cells Bull*, vol. 2023, no. 12, 2023, doi: https://doi.org/10.52710/fcb.137.

[16] S. N. Hasan *et al.*, "The influence of artificial intelligence on data system security," *Int. J. Comput. Exp. Sci. Eng.*, vol. 11, no. 3, 2025, doi: https://doi.org/10.22399/ijcesen.3476.

[17] M. D. Hossain, M. S. Sikder, M. S. Uddin, R. M. Ahsan, B. Uddin, and T. Hossen, "Cognitive Cyber Defense: AI–MIS Integration through Big Data and Cloud Frameworks for Next-Generation Digital Resilience," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 1, no. 02, pp. 140–152, Dec. 2023, doi: https://doi.org/10.58812/esiscs.v1i02.764.

[18] J. Kaur *et al.*, "Advanced cyber threats and cybersecurity innovation—Strategic approaches and emerging solutions," *J. Comput. Sci. Technol. Stud.*, vol. 5, no. 3, pp. 112–121, 2024, doi: https://doi.org/10.32996/jcsts.2023.5.3.9.

[19] F. Mahmud *et al.*, "AI-driven cybersecurity in IT project management: Enhancing threat detection and risk mitigation," *J. Posthumanism*, vol. 5, no. 4, pp. 23–44, 2025, doi: https://doi.org/10.63332/joph.v5i4.974.

[20] S. M. Orthi *et al.*, "AI-Driven Cyber Threat Intelligence as a Management Information System: Integrating Cybersecurity Governance and IT Project Management for Organizational Resilience," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 1, no. 02, pp. 194–213, Dec. 2023, doi: https://doi.org/10.58812/esiscs.v1i02.873.

[21] A. Shan-A-Alahi, M. Mustafizur, K. M. R. Hossan, A. Al Zaiem, and M. M. Rahman, "Cybersecurity training and its influence on employee behavior in business environments," *Comput. Fraud Secur.*, vol. 2024, no. 12, pp. 506–515, 2024, doi: https://doi.org/10.52710/cfs.689.

[22] M. A. Siam *et al.*, "AI-Driven Cyber Threat Intelligence Systems: A National Framework for Proactive Defense Against Evolving Digital Warfare," *Int. J. Comput. Exp. Sci. Eng*, vol. 11, no. 3, 2025, doi: https://doi.org/10.22399/ijcesen.3793.

[23] M. S. Uddin, M. S. Sikder, M. M. Anwar, and F. Hossain, "AI-driven cybersecurity and big data-enabled MIS frameworks: Strengthening supply chain integrity, energy resilience, and critical infrastructure protection," *J. Comput. Sci. Technol. Stud.*, vol. 7, no. 9, pp. 223–232, 2025, doi: https://doi.org/10.32996/jcsts.2025.7.9.26.

[24] P. Chakraborty, "Cyber Threat Prediction Using Temporal Data Analysis Model," in *2026 6th International Conference on Intelligent Technologies (CONIT)*, 2026, pp. 1–10. doi: https://doi.org/10.1109/CONIT69683.2026.11620680.

[25] N. Das *et al.*, "AI-enhanced Cyber Threat Detection: Transforming Security Frameworks in Management Information Systems," in *2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 2025, pp. 1–6. doi: https://doi.org/10.1109/ICECET63943.2025.11472511.

[26] S. N. Hasan *et al.*, "Self-healing cybersecurity systems using RL agents," in *2025 1st International Conference on Advancement in Futuristic Technologies (ICAFT)*, IEEE, 2025, pp. 1–8. doi: https://doi.org/10.1109/ICAFT66710.2025.11452866.

[27] H. Kaur, H. U. Himel, M. H. Rijvi, M. S. Sikder, M. Y. Biplob, and H. Imam, "Automated Vulnerability Assessment Using Machine Intelligence," in *2026 6th International Conference on Intelligent Technologies (CONIT)*, 2026, pp. 1–7. doi: https://doi.org/10.1109/CONIT69683.2026.11620731.

[28] N. Nabi *et al.*, "Continual Learning Pipelines for Detecting Insider Threats Across Corporate Cybersecurity Infrastructures," in *2025 International Conference on Electrical Engineering and Informatics (ICEEI)*, 2025, pp. 1–8. doi: https://doi.org/10.1109/ICEEI68459.2025.11330487.

[29] M. Nazmussakib, M. S. Sikder, M. A. K. P. Hemal, A. Al Zaiem, H. U. Himel, and N. Das, "Adversarial Machine Learning for Robust Cyber Defense Systems," in *2026 International Conference on Computing Theory and Wireless Communications (ICCTWC)*, 2026, pp. 1–9. doi: https://doi.org/10.1109/ICCTWC68241.2026.11583629.

[30] M. Raihan *et al.*, "Adversarial robustness mechanism for safeguarding biometric verification across mobile financial applications," in *2025 International Conference on Electrical Engineering and Informatics (ICEEI)*, IEEE, 2025, pp. 1–7. doi: https://doi.org/10.1109/ICEEI68459.2025.11330502.

[31] U. S. Saha, H. Imam, M. Mustafizur, A. Roy, A. Al Zaiem, and M. M. T. G. Manik, "Deep learning based malware image classification using binary image representation for secure management information systems," in *2026 Technology and Innovation Conference (TIC)*, IEEE, 2026, pp. 1–6. doi: https://doi.org/10.1109/TIC68483.2026.11702958.

[32] M. A. Siam *et al.*, "Cybersecure intelligent sensor framework for smart buildings: AI-based intrusion detection and resilience against IoT attacks," *Sensors*, vol. 25, no. 24, p. 7680, 2025, doi: https://doi.org/10.3390/s25247680.

[33] S. Sultana *et al.*, "AI-Augmented Big Data Analytics for Real-Time Cyber Attack Detection and Proactive Threat Mitigation," *Int. J. Comput. Exp. Sci. Eng.*, vol. 11, no. 3, Jul. 2025, doi: https://doi.org/10.22399/ijcesen.3564.

[34] M. K. Tuhin, H. Kaur, H. Imam, A. Hossain, A. Shan-A-Alahi, and P. Chakraborty, "AI-augmented security analytics for real-time threat identification," in *2026 6th International Conference on Intelligent Technologies (CONIT)*, IEEE, 2026, pp. 1–9. doi: https://doi.org/10.1109/CONIT69683.2026.11621644.

[35] M. T. Bin Ansar, M. S. Sikder, M. K. Tuhin, N. Das, P. Chakraborty, and H. Rahman, "Resilient MIS Design Using Cryptographic and Cyber-Security Techniques," in *Innovative Computing and Communications*, A. E. Hassanien, S. Anand, A. Jaiswal, and P. Kumar, Eds., Cham: Springer Nature Switzerland, 2026, pp. 317–331. doi: https://doi.org/10.1007/978-3-032-30008-9_24.

[36] P. Chakraborty *et al.*, "An Intelligent Cybersecurity Framework for Data Protection in Cloud Computing Environments," *Spectr. Eng. Sci.*, vol. 2, no. 3, pp. 657–677, Oct. 2024, doi: https://doi.org/10.5281/zenodo.21538159.

[37] K. S. Mudaliyar and S. Satish Kumar, "Cloud Computing Data Security Through Policy Bound Encryption and Verifiable Access Provenance in Multi-Tenant Storage," *Eur. J. Appl. Sci. Eng. Technol.*, vol. 3, no. 6, pp. 284–296, Dec. 2025, doi: https://doi.org/10.59324/ejaset.2025.3(6).21.

[38] N. Das *et al.*, "A Privacy-Preserving Federated Intrusion Detection System Leveraging Secure Multi-Party Computation Across Collaborative Cloud Tenants," in *Proceedings of Fifth International Conference on Computing and Communication Networks*, G.-N. Nguyen, A. Swaroop, and P. Shukla, Eds., Cham: Springer Nature Switzerland, 2026, pp. 488–504. doi: https://doi.org/10.1007/978-3-032-14197-2_42.

[39] N. Das *et al.*, "AI-enhanced privacy preservation using homomorphic federated models," in *2025 1st International Conference on Advancement in Futuristic Technologies (ICAFT)*, IEEE, 2025, pp. 1–8. doi: https://doi.org/10.1109/ICAFT66710.2025.11453096.

[40] U. Haldar *et al.*, "Blockchain-Driven Access Control and Compliance Auditing Framework for Federated Cloud Service Providers: Architecture, Prototype and Evaluation," in *Proceedings of Fifth International Conference on Computing and Communication Networks*, G.-N. Nguyen, A. Swaroop, and P. Shukla, Eds., Cham: Springer Nature Switzerland, 2026, pp. 464–487. doi: https://doi.org/10.1007/978-3-032-14197-2_41.

[41] J. Hassan *et al.*, "Blockchain Integration in Management Information Systems a Decentralized Approach to Strengthening Cybersecurity and Data Integrity," in *2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 2025, pp. 1–7. doi: https://doi.org/10.1109/ICECET63943.2025.11472020.

[42] H. U. Himel, M. Mustafizur, T. Akter, R. Anzum, Z. A. Tanni, and A. Hossain, "Privacy-Aware Data Protection Models for Cloud Environments," in *2026 6th International Conference on Intelligent Technologies (CONIT)*, 2026, pp. 1–10. doi: https://doi.org/10.1109/CONIT69683.2026.11621371.

[43] M. D. Hossain, M. S. Uddin, M. S. Sikder, T. Hossen, B. Uddin, and R. M. Ahsan, "Green and secure data centers: Balancing energy efficiency with advanced cybersecurity measures," *J. Comput. Sci. Technol. Stud.*, vol. 6, no. 5, pp. 300–315, 2024, doi: https://doi.org/10.32996/jcsts.2024.6.5.24.

[44] S. C. Mohonta *et al.*, "Efficient Homomorphic Encryption Techniques for Confidential Big-Data Analytics over Public Cloud Infrastructures: Algorithms, Performance, and Trade-Offs," in *Proceedings of Fifth International Conference on*

Computing and Communication Networks, G.-N. Nguyen, A. Swaroop, and P. Shukla, Eds., Cham: Springer Nature Switzerland, 2026, pp. 507–528. doi: https://doi.org/10.1007/978-3-032-21499-7_41.

[45] K. S. Mudaliyar and S. S. Kumar, "Intelligent Cybersecurity Frameworks for Data Protection in Cloud-Integrated Management Information Systems," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 2, no. 02, pp. 305–323, Dec. 2024, doi: https://doi.org/10.58812/esiscs.v2i02.1165.

[46] K. S. Mudaliyar and S. S. Kumar, "AEGIS-FL: Auditable Federated Threat Detection for Multi-Tenant Cloud-Edge Systems," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 4, no. 01, pp. 101–121, Aug. 2026, doi: https://doi.org/10.58812/esiscs.v4i01.1217.

[47] K. S. Mudaliyar and S. Satish Kumar, "TrustCloud-FL: A Self-Healing and Privacy-Preserving Architecture for Secure Analytics Across Multi-Tenant Hybrid Clouds," *Eur. J. Appl. Sci. Eng. Technol.*, vol. 4, no. 5, pp. 151–171, Sep. 2026, doi: https://doi.org/10.59324/ejaset.2026.4(5).10.

[48] A. Roy, E. Hossain, H. U. Himel, H. Imam, R. A. Md Rashed, and H. Kaur, "Privacy-Preserving MIS Using Blockchain, Federated Learning and Zero-Trust Models," in *Innovative Computing and Communications*, A. E. Hassanien, S. Anand, A. Jaiswal, and P. Kumar, Eds., Cham: Springer Nature Switzerland, 2026, pp. 210–223. doi: https://doi.org/10.1007/978-3-032-30008-9_16.

[49] A. Shan-A-Alahi *et al.*, "Deep Learning-Based Threat Prediction and Autonomous Response Mechanisms for Containerized Microservices in Hybrid Cloud Deployments," in *Proceedings of Fifth International Conference on Computing and Communication Networks*, G.-N. Nguyen, A. Swaroop, and P. Shukla, Eds., Cham: Springer Nature Switzerland, 2026, pp. 529–554. doi: https://doi.org/10.1007/978-3-032-21499-7_42.

[50] A. Shan-A-Alahi, M. S. Sikder, H. U. Himel, M. T. Bin Ansar, M. K. Tuhin, and H. Kaur, "Resilient Cybersecurity Architectures for Large-Scale Distributed Systems," in *2026 IEEE International Conference for Convergence in Computing Technology (I3CTCON)*, IEEE, 2026, pp. 1–8. doi: https://doi.org/10.1109/I3CTCON68242.2026.11507768.

[51] M. K. Tuhin *et al.*, "Federated Machine Learning for Privacy-Preserving Cyber Threat Intelligence in Global Cloud-Integrated MIS Platforms," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 2, no. 02, pp. 284–293, Dec. 2024, doi: https://doi.org/10.58812/esiscs.v2i02.1123.

[52] U. K. R. Gangula, M. A. Miah, K. Mula, M. Dhakan, Q. T. Sadat, and S. Nayak, "A Lightweight and Secure Semantic Communication Architecture for Edge–IoT–6G Systems," *IEEE Commun. Stand. Mag.*, pp. 1–6, 2026, doi: https://doi.org/10.1109/MCOMSTD.2026.3677038.

[53] E. Hossain *et al.*, "AI-Driven Trust and Bio-Inspired Optimization Integration in Standardized Digital Twin Frameworks for Next-Generation," *IEEE Commun. Stand. Mag.*, pp. 1–6, 2026, doi: https://doi.org/10.1109/MCOMSTD.2026.3688125.

[54] M. R. H. Mahin *et al.*, "Secured and standardized intelligent zero-touch 6G framework for edge-AI applications," *IEEE Commun. Stand. Mag.*, 2026, doi: https://doi.org/10.1109/MCOMSTD.2026.3660159.

[55] K. B. Siddiqa *et al.*, "Assessment of Survivability and Importance Analysis for Networks Managing Intricate Traffic Flows," *IEEE Commun. Stand. Mag.*, vol. 10, no. 3, pp. 101–106, 2026, doi: https://doi.org/10.1109/MCOMSTD.2025.3638981.

[56] S. R. Varanasi, S. S. S. Valiveti, M. Adnan, M. I. Faruk, M. J. Hossain, and M. M. T. G. Manik, "Cross-Domain standardization and secure edge intelligence for Real-Time digital twin deployments in Next-Generation communication systems," *IEEE Commun. Stand. Mag.*, vol. 10, no. 3, pp. 107–112, 2026, doi: https://doi.org/10.1109/MCOMSTD.2026.3662187.

[57] M. I. Faruk *et al.*, "Sustainable Computing and Communication Technologies for Digital Twin–Based Energy Optimization in Smart Infrastructure Systems," *Internet Technol. Lett.*, vol. 9, no. 5, p. e70351, Sep. 2026, doi: https://doi.org/10.1002/itl2.70351.

[58] J. Goh, S. Adepu, K. N. Junejo, and A. Mathur, "A Dataset to Support Research in the Design of Secure Water Treatment Systems," in *Critical Information Infrastructures Security*, G. Havarneanu, R. Setola, H. Nassopoulos, and S. Wolthusen, Eds., Cham: Springer International Publishing, 2017, pp. 88–99. doi: https://doi.org/10.1007/978-3-319-71368-7_8.

[59] K. S. Mudaliyar and S. S. Kumar, "Deadline Performance and Tenant Fairness Under Security Screening in Cloud Workload Scheduling," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 3, no. 01, pp. 198–210, Aug. 2025, doi: https://doi.org/10.58812/esiscs.v3i01.1250.

[60] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS One*, vol. 10, no. 3, p. e0118432, 2015, doi: https://doi.org/10.1371/journal.pone.0118432.