

Toward Responsible AI-Augmented Cyber Defense: Pattern Recognition, Defense-in-Depth, and the Case for Human–AI Collaboration

Mustafa S. Aljumaily¹, Hayder Kareem Abed², and Nawar S. Alseelawi³

¹*Research and Development (R&D) Department, Daw Alfada Company, Baghdad, Iraq, mustafa.s@daw-alfada.com*

²*CEO, Daw Alfada Company, Baghdad, Iraq, haider.k@daw-alfada.com*

³*University of Misan, Maysan, Iraq, nawar.alseelawi@uomisan.edu.iq*

Abstract

Cybersecurity literature has extensively documented the operational benefits of artificial intelligence (AI) for threat detection, incident response, and prevention, while raising qualitative concerns about over-automation, algorithmic bias, and analyst-skill erosion. What remains largely absent is a formal, falsifiable model connecting three constructs that recur across this literature: Defense-in-Depth Theory, the Artificial Intelligence Theory of Pattern Recognition, and human–AI collaboration in security operations. This paper develops such a model. We formalize layered defense as a Bernoulli detection cascade in which AI augmentation enters multiplicatively across layers; we formalize each layer’s pattern-recognition behavior as a Neyman–Pearson/Bayesian detector with a derived closed-form optimal threshold; and we formalize human–AI triage as a capacity-constrained cascade with an explicit, quantifiable trade-off between detection probability and false-alarm (“alert fatigue”) rate. A Monte Carlo/analytical simulation evaluated at illustrative but realistic operating points shows that (i) AI augmentation compounds across defense layers, delivering its largest marginal gains exactly where traditional layering saturates, and (ii) full human review of AI-flagged alerts is not optimal: increasing analyst capacity toward 100% coverage cuts false alarms by roughly 20-fold but simultaneously lowers system-level detection probability, because imperfect analyst accuracy is then applied to every alert rather than a filtered subset. These results give the widely repeated qualitative recommendation of “balanced human–AI collaboration” a precise, testable form and suggest an interior-optimum capacity ratio as a concrete design target for security operations centers (SOCs), including those securing IT/OT-converged critical infrastructure.

Keywords: Artificial Intelligence; Cybersecurity; Defense-in-Depth; Pattern Recognition; Human–AI Collaboration; Bayesian Detection; Security Operations Center; Alert Fatigue.

1 Introduction

The cybersecurity threat landscape has grown in scale and technical sophistication faster than traditional, rule-based defenses can adapt, a pattern documented across recent reviews of AI-driven security [1, 2]. Two theoretical constructs recur throughout this literature as organizing frameworks: Defense-in-Depth Theory, which holds that layering independent security controls narrows the attack surface more than any single control could; and the Artificial Intelligence Theory

of Pattern Recognition, which holds that machine learning systems can detect deviations from learned normal behavior that static, signature-based systems cannot [2]. Both theories are used descriptively—to organize a narrative account of where AI helps—but neither has, to our knowledge, been operationalized as a quantitative model that predicts how much a given layering strategy, AI augmentation level, or human-review policy will change system-level security outcomes.

This gap matters because the same literature identifies a genuine, unresolved tension: AI improves detection speed and coverage, but over-reliance on automation is repeatedly flagged as a risk to analyst skill retention and to the diversity of judgment a security operations center (SOC) can bring to novel threats [1, 2]. Recommendations to “balance automation with human oversight” appear throughout, but without a model that specifies what balance means quantitatively, such advice is difficult to act on or to falsify.

This paper addresses that gap with three contributions:

1. A formal Defense-in-Depth cascade model in which AI-driven pattern recognition augments each layer’s detection probability multiplicatively across the miss-probability product, giving a precise account of why AI’s marginal value is largest where layering alone saturates (Sections 4.2–4.4).
2. A Neyman–Pearson/Bayesian formalization of the pattern-recognition layer itself, including a closed-form cost-weighted optimal threshold, connecting the qualitative “AI detects anomalies” claim to standard detection theory (Section 4.3).
3. A capacity-constrained human–AI triage model that produces a specific, counterintuitive, and testable prediction—that full human review of AI-flagged alerts is not detection-optimal—replacing the general “human oversight is good” recommendation with an interior-optimum design target (Section 4.5, Section 6.3).

The model is validated with an analytical/Monte Carlo simulation using illustrative Gaussian-separated detectors representative of a network intrusion detection system (IDS), an endpoint detection and response (EDR) layer, and an OT/SCADA anomaly monitor—a layer combination directly relevant to IT/OT-converged critical infrastructure such as oil and gas SCADA environments.

2 Related Work

Recent integrative and systematic reviews converge on three application areas for AI in cybersecurity [1, 2, 3]. First, AI-driven threat detection extends signature-based intrusion detection systems (IDS) with supervised and unsupervised machine learning—random forests, support vector machines, and neural networks for classification; clustering and principal component analysis for outlier detection—and deep learning architectures (CNNs, RNNs/LSTMs) for both network-traffic and malware analysis [1, 4, 5]. Second, AI-powered response systems automate parts of the incident-response lifecycle through Security Orchestration, Automation, and Response (SOAR) platforms and AI-driven threat hunting, reducing the time between detection and containment [1]. Third, AI-enhanced prevention techniques—predictive threat intelligence, AI-assisted vulnerability management, and behavioral biometrics—shift security posture from reactive to proactive [1, 2].

Running alongside these operational gains, both reviews raise a consistent set of concerns: adversarial manipulation of AI models, data-quality and representativeness problems that produce

biased detection outcomes, the opacity (“black-box” nature) of AI decision processes, and the risk that operational reliance on AI erodes human analytical skill over time [1, 2]. Ejjami [2] frames these concerns within two theories—Cybersecurity Defense-in-Depth Theory and the Artificial Intelligence Theory of Pattern Recognition—and argues that responsible AI deployment requires combining AI’s predictive capability with human oversight, transparency (explainable AI), and equitable access, particularly for smaller organizations with constrained AI budgets.

These reviews are, by design, qualitative syntheses: they identify recurring themes and theoretical constructs across a large literature rather than deriving quantitative predictions from them. Neither offers a mechanism for computing how much detection improvement a given layering-and-AI-augmentation strategy should be expected to deliver, nor a decision rule for setting a human-review capacity that is provably better than full automation or full human review. This paper takes the two theoretical constructs those reviews foreground and gives each a formal, testable mathematical structure, which we then validate computationally.

3 Theoretical Framework

We adopt the same two guiding theories used in the reviewed literature and make each one operational:

3.1 Cybersecurity Defense-in-Depth Theory

Defense-in-Depth Theory holds that layering independent, heterogeneous security controls (network, endpoint, application, and—in industrial environments—OT/SCADA-specific controls) reduces overall risk more effectively than optimizing any single control, because an attacker must evade every layer to succeed [2, 6]. We formalize this as a probabilistic miss-cascade in Section 4.2.

3.2 Artificial Intelligence Theory of Pattern Recognition

This theory holds that AI systems can learn the statistical structure of “normal” behavior from data and flag deviations from it, enabling detection of threats that do not match any known signature [2]. We formalize this as a binary hypothesis test on a learned anomaly score, following standard Neyman–Pearson/Bayesian detection theory, in Section 4.3.

3.3 Synthesis: Human–AI Collaboration as a Triage Cascade

Neither theory, individually, addresses where human judgment enters the pipeline or how its capacity should be allocated. We treat human–AI collaboration as a second, downstream cascade stage—an AI-prioritized triage queue with a capacity-constrained human reviewer—and derive its performance characteristics in Section 4.5. This synthesis is the paper’s central theoretical contribution: it connects Defense-in-Depth and Pattern Recognition Theory to a third, previously under-formalized construct using the same probabilistic language.

4 Mathematical Model

4.1 Notation

Let L be the number of independent defense layers indexed $i \in \{1, \dots, L\}$ (e.g., network IDS, endpoint EDR, OT/SCADA anomaly monitor). Let H_0 and H_1 denote the benign and attack hypotheses, with prior $\pi_1 = P(H_1)$. Each layer produces an anomaly score s_i and applies a decision threshold τ_i . $d_i(\tau_i) = P(s_i > \tau_i \mid H_1)$ is the layer’s true-positive (detection) rate and $f_i(\tau_i) = P(s_i > \tau_i \mid H_0)$ is its false-positive rate. $\alpha_i \in [0, 1]$ is an AI augmentation factor at layer i . C is the human analyst’s review capacity (alerts per unit time), and d_H, f_H are the analyst’s own detection/false-positive rates on reviewed alerts.

4.2 Defense-in-Depth as a Layered Bernoulli Cascade

Treating the system as detecting an attack whenever any layer flags it, and assuming layer-conditional independence given H_1 :

$$P(\text{system misses attack}) = \prod_{i=1}^L (1 - d_i(\tau_i)) \quad (1)$$

$$P_D(L) = 1 - \prod_{i=1}^L (1 - d_i(\tau_i)) \quad (2)$$

$$P_F(L) = 1 - \prod_{i=1}^L (1 - f_i(\tau_i)) \quad (3)$$

$P_D(L)$ increases monotonically in L with diminishing marginal returns, since each new layer only needs to catch what prior layers missed. $P_F(L)$, however, also increases with L —the formal counterpart of the “alert fatigue” problem noted qualitatively in the reviewed literature [1, 2] and documented empirically in operational SOC settings [7]—motivating the triage stage introduced in Section 4.5. Structurally, Equations (1)–(3) instantiate a classical distributed-detection architecture with an OR fusion rule [8]; the contribution here lies in coupling that structure to AI augmentation (Section 4.4) and to capacity-constrained human triage (Section 4.5).

4.3 Pattern Recognition as a Neyman–Pearson / Bayesian Detector

Each layer’s anomaly score is modeled as $s_i \mid H_0 \sim \mathcal{N}(\mu_{0i}, \sigma_{0i}^2)$ and $s_i \mid H_1 \sim \mathcal{N}(\mu_{1i}, \sigma_{1i}^2)$, with $\mu_{1i} > \mu_{0i}$. For a threshold rule (flag if $s_i > \tau_i$):

$$d_i(\tau_i) = 1 - \Phi\left(\frac{\tau_i - \mu_{1i}}{\sigma_{1i}}\right), \quad f_i(\tau_i) = 1 - \Phi\left(\frac{\tau_i - \mu_{0i}}{\sigma_{0i}}\right) \quad (4)$$

where Φ is the standard normal CDF. Sweeping τ_i traces the layer’s ROC curve; its area (AUC) summarizes discriminative quality [9]. By the Neyman–Pearson lemma [10], the cost-optimal decision rule is a likelihood-ratio test; for equal-variance Gaussians ($\sigma_{0i} = \sigma_{1i} = \sigma_i$) this reduces to a closed-form threshold:

$$\tau_i^* = \frac{\mu_{0i} + \mu_{1i}}{2} + \frac{\sigma_i^2}{\mu_{1i} - \mu_{0i}} \ln \eta, \quad \eta = \frac{c_{\text{FP}} \pi_0}{c_{\text{FN}} \pi_1} \quad (5)$$

where c_{FP} and c_{FN} are the relative costs of a false alarm versus a missed attack. Equation (5) gives a principled, cost-aware way to set each layer’s operating point, directly addressing the reviewed literature’s concern that skewed or arbitrary thresholds produce disproportionate false positives or negatives [2].

4.4 AI Augmentation of Layer Detection

Let d_i^0 denote a layer’s baseline (signature/rule-based) detection rate. AI-driven pattern recognition recovers a fraction α_i of the attacks the baseline layer would otherwise miss:

$$d_i^{\text{AI}} = d_i^0 + (1 - d_i^0)\alpha_i \iff 1 - d_i^{\text{AI}} = (1 - d_i^0)(1 - \alpha_i) \quad (6)$$

Substituting into Equation (1):

$$P_{\text{miss}}^{\text{AI}}(L) = \prod_{i=1}^L (1 - d_i^0)(1 - \alpha_i) = P_{\text{miss}}^0(L) \cdot \prod_{i=1}^L (1 - \alpha_i) \quad (7)$$

Equation (7) is the paper’s central structural claim about AI-augmented Defense-in-Depth: AI’s contribution compounds multiplicatively across layers, so even modest per-layer gains ($\alpha_i \approx 0.3$ – 0.4) yield large system-level improvements once $L \geq 2$ – 3 . This gives quantitative form to the frequently repeated but rarely formalized claim that AI “strengthens every layer” of a defense-in-depth architecture [2].

4.5 Human–AI Collaborative Triage Under a Capacity Constraint

The raw alert stream generated by Section 4.2 arrives at rate $R = \pi_1 P_D(L) + (1 - \pi_1)P_F(L)$. An AI triage layer with threshold τ_{AI} prioritizes which alerts reach a human analyst of finite capacity C . Define the review probability $p_{\text{review}} = \min(1, C/R)$. Reviewed alerts pass through the analyst’s own accuracy (d_H, f_H); unreviewed alerts are auto-actioned at AI-only reliability:

$$D_{\text{final}} = P_D(L) [p_{\text{review}} d_H + (1 - p_{\text{review}})] \quad (8)$$

$$F_{\text{final}} = P_F(L) [p_{\text{review}} f_H + (1 - p_{\text{review}})] \quad (9)$$

Two structural properties follow directly. First, F_{final} strictly decreases as capacity increases (since $f_H < 1$), so human review’s clearest value is filtering false alarms—reducing alert fatigue. Second, D_{final} can decrease as capacity increases whenever $d_H < 1$: an imperfect analyst reviewing every alert can be net-negative for detection relative to simply auto-actioning AI-flagged alerts, because the analyst’s own error rate is then applied universally rather than to a filtered subset. This gives a precise, mechanistic reading of the reviewed literature’s qualitative warning that over-reliance on any single layer—including the human layer—can leave a system worse off than a well-designed automated baseline [1].

Choosing τ_{AI} to maximize expected utility subject to the capacity constraint,

$$\max_{\tau_{\text{AI}}} U(\tau_{\text{AI}}) = B \pi_1 d(\tau_{\text{AI}}) - c_{\text{FP}} \pi_0 f(\tau_{\text{AI}}) - c_H R(\tau_{\text{AI}}) \quad \text{s.t.} \quad R(\tau_{\text{AI}}) \leq C \quad (10)$$

gives, via the Lagrangian and the Neyman–Pearson structure of Section 4.3, an implicit equation for the optimal triage threshold τ_{AI}^* solvable by bisection on the Lagrange multiplier until the capacity constraint binds. We solve this numerically in Section 5.

4.6 Composite System Metric

Combining Sections 4.2–4.5 gives the full-pipeline detection probability:

$$D_{\text{system}} = \left[1 - \prod_{i=1}^L (1 - d_i^0)(1 - \alpha_i) \right] \times [p_{\text{review}} d_H + (1 - p_{\text{review}})] \quad (11)$$

with a secondary Alert Fatigue Index, $\text{AFI} = F_{\text{final}}/C$, capturing operational sustainability—a quantity discussed qualitatively but not defined in the reviewed literature [1, 2].

5 Simulation Methodology

The model of Section 4 was implemented in Python (NumPy/SciPy) and evaluated through three experiments, using three illustrative Gaussian-separated layers representative of a network IDS, an endpoint EDR system, and an OT/SCADA anomaly monitor (parameters chosen for realistic but illustrative class separation; the simulation is a proof-of-concept rather than a calibration to a specific deployed system).

Experiment 1 — Defense-in-Depth $\times$ AI augmentation. $P_D(L)$ computed from Equations (2), (6), and (7) for $L = 1, 2, 3$ layers, at AI augmentation levels $\alpha \in \{0, 0.3, 0.6\}$, with each layer’s threshold set to a fixed 5% false-positive rate via Equation (4).

Experiment 2 — Pattern-recognition ROC. Per-layer ROC curves and AUC computed by sweeping τ_i across each layer’s Gaussian pair (Equation (4)).

Experiment 3 — Human–AI cascade. For the fixed 3-layer system at $\alpha = 0.3$ ($\pi_1 = 0.02$ base rate, $d_H = 0.90$, $f_H = 0.05$), D_{final} and F_{final} (Equations (8)–(9)) were computed across analyst capacity ratios $C/R \in [0.12, 1.00]$.

All code and exact parameter values are provided in the accompanying simulation script for reproducibility.

6 Results

6.1 Defense-in-Depth $\times$ AI Augmentation

Table 1 reports the values plotted in Figure 1. Without AI ($\alpha = 0$), a single layer detects 61.6% of attacks; three layers in series raise this to 95.2%. AI augmentation compounds this effect: at $\alpha = 0.3$, three layers reach 98.4%; at $\alpha = 0.6$, 99.7%. Consistent with Equation (7), AI’s marginal contribution is largest precisely where layering alone begins to saturate (the $L = 2 \rightarrow 3$ transition), rather than diminishing as layers alone do.

6.2 Pattern-Recognition ROC

Each layer’s standalone ROC curve shows meaningful but incomplete discriminative power (AUC between 0.86 and 0.94 for the parameters used), underscoring why system-level detection probability

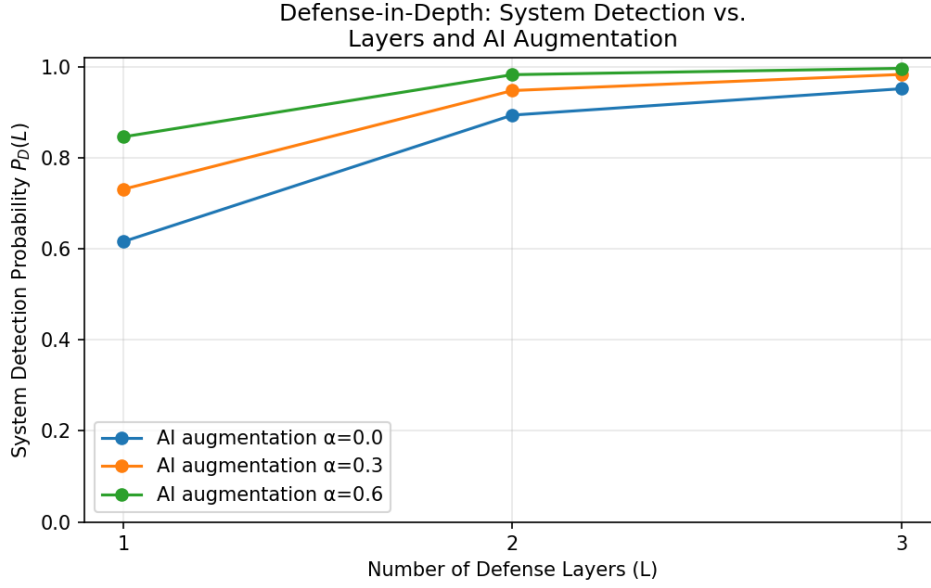


Figure 1: System detection probability $P_D(L)$ versus number of defense layers, at three AI augmentation levels.

Table 1: System detection probability $P_D(L)$ by layer count and AI augmentation level.

Layers (L)	$\alpha = 0$ (no AI)	$\alpha = 0.3$	$\alpha = 0.6$
1	0.6164	0.7315	0.8465
2	0.8942	0.9482	0.9831
3	0.9521	0.9836	0.9969

$P_D(L)$ —not any single detector—is the appropriate unit of analysis for defense-in-depth claims.

6.3 Human–AI Collaborative Triage

Table 2 summarizes the key operating points. Raw (pre-triage) system detection is 0.984 and raw false-alarm probability is 0.143. As analyst capacity increases from 12% to 100% of the raw alert volume, the false-alarm rate reaching action falls roughly 20-fold ($0.126 \rightarrow 0.007$), but detection probability falls from 0.971 to 0.885—a decline directly attributable to the analyst’s own imperfect accuracy ($d_H = 0.90$) now being applied to every alert rather than a filtered subset.

Table 2: Final detection and false-alarm probabilities at selected analyst capacity ratios.

C/R (analyst capacity ratio)	Final detection D_{final}	Final false-alarm F_{final}
0.12 (raw, minimal review)	0.971	0.126
0.25	0.959	0.109
1.00 (full review)	0.885	0.007

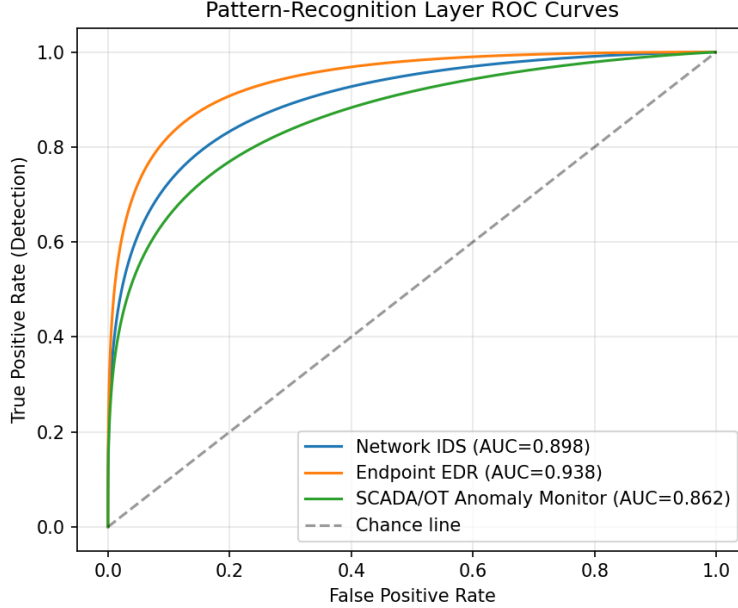


Figure 2: ROC curves and AUC for each of the three illustrative defense layers.

7 Discussion

These results reframe two claims that recur across the reviewed literature in qualitative form, giving each a quantitative, testable structure.

7.1 Defense-in-Depth Is Where AI Pays Off Most

Equation (7) shows that AI’s contribution to a layered architecture is multiplicative rather than additive: AI does not simply add a fixed increment of detection capability, it compounds the miss-probability reduction of every layer beneath it. Practically, this suggests that organizations gain more from deploying moderate AI augmentation ($\alpha \approx 0.3$) across several heterogeneous layers than from concentrating AI investment in a single, highly optimized detector—a specific, resource-allocation-relevant reading of the general claim that “AI strengthens each layer of cybersecurity” [2].

7.2 An Interior Optimum for Human–AI Collaboration

The central and more novel finding is that full human review is not detection-optimal (Section 6.3). This gives the frequently repeated recommendation to “balance automation with human oversight” [1, 2] a specific, falsifiable content: the optimal analyst capacity ratio C/R is an interior point determined by the analyst’s own accuracy d_H relative to 1, not an endpoint of “more human review is always safer.” Where d_H is high (well-trained, unfatigued analysts), the model favors more review; where d_H degrades—plausibly through the very automation-induced skill erosion the reviewed literature warns about [1]—the model favors routing a larger share of alerts to direct AI action, creating a feedback loop worth investigating empirically in future SOC deployments.

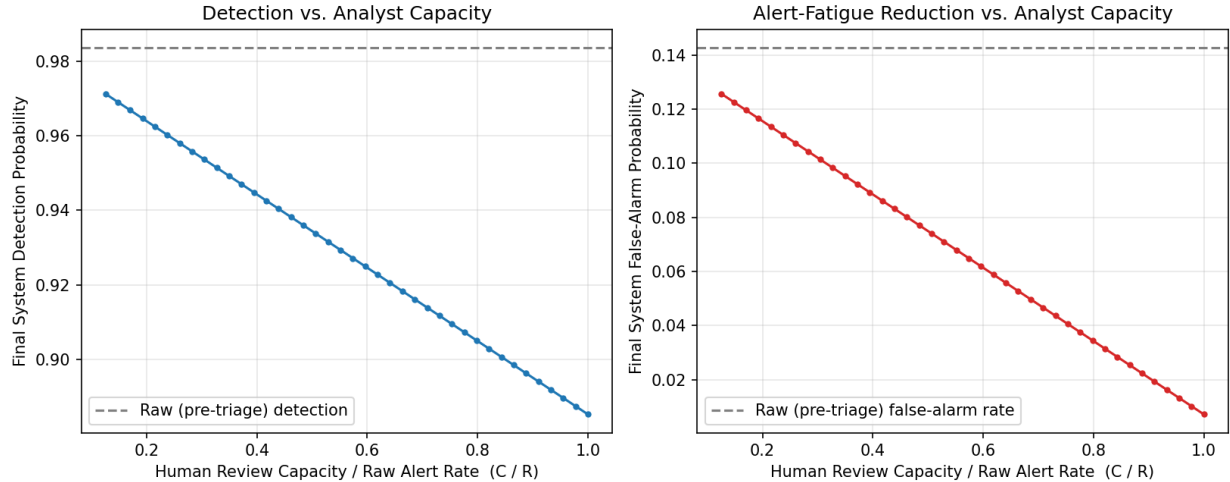


Figure 3: Final system detection probability (left) and false-alarm probability (right) versus human analyst capacity ratio C/R , for the fixed 3-layer, $\alpha = 0.3$ system.

7.3 Implications for IT/OT-Converged Critical Infrastructure

The three-layer configuration used here—network IDS, endpoint EDR, and an OT/SCADA anomaly monitor—mirrors the layer composition of security operations centers protecting IT/OT-converged environments such as oil and gas SCADA networks. In such environments, analyst capacity is often the binding constraint (specialized OT security expertise is scarce), making the capacity-constrained triage model of Section 4.5 directly actionable: Equation (10) offers a principled way to set the AI triage threshold τ_{AI} given an organization’s actual analyst headcount, rather than treating human review as an unlimited resource.

8 Limitations and Future Work

Synthetic, Gaussian-separated detectors. The illustrative parameters are chosen for realistic but unverified class separation. Future work should calibrate μ , σ , and α to empirical detector outputs on public IDS benchmark datasets (e.g., CICIDS2017, NSL-KDD) or to real SOC alert logs.

Layer independence. Equations (1)–(3) assume conditionally independent layers. Real attacks often evade correlated layers together (e.g., an attacker who defeats network-level detection may also be positioned to defeat endpoint detection). Extending the model with a copula or correlated-Bernoulli structure is a natural next step.

Static thresholds. Thresholds τ_i and τ_{AI} are treated as fixed operating points; a reinforcement-learning or adaptive-control extension could let thresholds respond to observed attack-rate drift, addressing the adversarial-evasion concern raised in the reviewed literature [1].

Single triage layer and static analyst accuracy. d_H and f_H are treated as constants; modeling their degradation under alert-volume-driven fatigue would let the model formally test the skill-erosion hypothesis discussed in Section 7.2 rather than only motivating it.

Auto-action reliability. Equations (8)–(9) treat unreviewed AI-flagged alerts as fully actioned detections, which is what drives the decline of D_{final} with capacity. If automated response itself fails, is rate-limited, or is exploited by an adversary, the effective reliability of the unreviewed path falls below one and the interior optimum shifts back toward more human review; modeling that path explicitly is a direct extension.

Uniform review selection. The review probability p_{review} is applied uniformly to attack and benign alert streams, i.e., reviewed alerts are drawn as a random subset. Score-ordered triage—reviewing the highest-priority alerts first—would change the composition of the reviewed subset and generally improve on the uniform baseline reported here, making the present results a conservative bound on well-prioritized triage.

No empirical validation. This paper establishes the model and demonstrates its qualitative and quantitative behavior through simulation; empirical validation against real SOC incident data is left for future work, ideally in collaboration with an operating SOC willing to share anonymized alert and triage-outcome data.

9 Conclusion

This paper formalized two theoretical constructs widely used in the AI-cybersecurity review literature—Defense-in-Depth Theory and the Artificial Intelligence Theory of Pattern Recognition—as an explicit probabilistic cascade, and extended them with a capacity-constrained human–AI triage model. The resulting framework makes two claims precise that the literature otherwise states qualitatively: that AI augmentation compounds multiplicatively across defense layers, and that human oversight of AI-driven security systems has an interior optimum rather than a monotonic “more is better” relationship. Simulation results support both claims under illustrative but realistic parameters. The framework and its optimization structure (Equation (10)) offer SOC designers—including those securing IT/OT-converged critical infrastructure—a concrete, falsifiable basis for allocating AI and human analyst resources, moving the field’s guidance on responsible AI-augmented cyber defense from general principle toward testable design rule.

References

- [1] Chigozie Kingsley Ejeofobiri, Adedoyin Adetumininu Fadare, Olalekan Olorunfemi Fagbo, Valerie O. Ejiofor, and Adetutu Temitope Fabusoro. The role of artificial intelligence in enhancing cybersecurity: A comprehensive review of threat detection, response, and prevention techniques. *International Journal of Science and Research Archive*, 13(2):310–316, 2024. <https://doi.org/10.30574/ijrsra.2024.13.2.2161>.

- [2] Rachid Ejjami. Enhancing cybersecurity through artificial intelligence: Techniques, applications, and future perspectives. *Journal of Next-Generation Research* 5.0, 1(1), 2024. <https://doi.org/10.70792/jngr5.0.v1i1.5>.
- [3] Iqbal H. Sarker, Md Hasan Furhad, and Raza Nowrozy. AI-driven cybersecurity: an overview, security intelligence modeling and research directions. *SN Computer Science*, 2(3):173, 2021.
- [4] Zeeshan Ahmad, Adnan Shahid Khan, Cheah Wai Shiang, Johari Abdullah, and Farhan Ahmad. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1):e4150, 2021.
- [5] Hongyu Liu and Bo Lang. Machine learning and deep learning methods for intrusion detection systems: A survey. *Applied Sciences*, 9(20):4396, 2019.
- [6] National Security Agency. Defense in depth: A practical strategy for achieving information assurance in today’s highly networked environments. Technical report, NSA Information Assurance Solutions Group, n.d.
- [7] Wajih Ul Hassan, Shengjian Guo, Ding Li, Zhengzhang Chen, Kangkook Jee, Zhichun Li, and Adam Bates. NoDoze: Combatting threat alert fatigue with automated provenance triage. In *Network and Distributed System Security Symposium (NDSS)*, 2019.
- [8] Pramod K. Varshney. *Distributed Detection and Data Fusion*. Springer, New York, 1997.
- [9] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- [10] Jerzy Neyman and Egon S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, 231:289–337, 1933.