AI-Driven Defense-in-Depth: A Systematic Review of SOC Maturity Models and DDoS Mitigation

George Antoniou[1,†]

[1] *Lynn University, Boca Raton, Florida 33431, USA*

Abstract

The growing sophistication of distributed denial-of-service (DDoS) attacks poses persistent challenges to security operations centers (SOCs). This paper presents a structured, evidence-based framework for integrating artificial intelligence (AI) into layered cyber defenses. Through systematic literature review and mapping of peer-reviewed intrusion detection techniques, we examine the applicability of ensemble learning, explainable AI (XAI), and federated learning across the defense-in-depth spectrum. We also propose an AI-maturity roadmap grounded in ENISA and NIST frameworks to guide phased SOC integration. Our findings support strategic AI deployment for improved detection accuracy, reduced triage time, and enhanced operational resilience against large-scale DDoS campaigns

Keywords

DDoS, defense-in-depth, artificial intelligence, SOC maturity, XAI, cybersecurity roadmap

1. Introduction

Distributed Denial-of-Service (DDoS) attacks remain a critical cybersecurity challenge, frequently targeting national infrastructure, enterprise networks, and public-facing systems. These attacks disrupt availability, overwhelm detection systems, and expose operational gaps in many security operations centers (SOCs). While perimeter-based defenses and reactive mitigation techniques have improved in speed and scale, attackers have likewise evolved, leveraging low-and-slow volumetric traffic, botnets, and encrypted payloads to evade traditional controls.

In mid-2022, the Albanian government experienced one of the most impactful nation-state-sponsored DDoS attacks in Europe. Key online portals, digital identity systems, and e-governance platforms were rendered inoperable. Although mitigation strategies succeeded in halting peak traffic volumes, post-incident analysis by CESK [3] and external vendors [4] revealed two major weaknesses: delayed anomaly detection at the network layer and insufficient coordination between security layers, highlighting the importance of layered defense, also known as defense-in-depth.

These deficiencies emphasize a growing need to rethink SOC architecture through the lens of artificial intelligence (AI). AI has demonstrated significant potential in augmenting anomaly detection, reducing triage time, and supporting threat attribution, yet its deployment across SOC maturity levels remains inconsistent.

Moreover, existing research lacks comprehensive frameworks that align AI capabilities to specific defense-in-depth layers, making operational integration ad hoc or siloed.

This paper proposes a structured, AI-enhanced defense-in-depth framework. We build upon validated techniques, including ensemble learning, explainable AI (XAI), and federated learning, to map AI tools to each of the seven core security layers. In doing so, we aim to support both immediate SOC performance improvement and long-term maturity planning.

Drawing from a systematic review of literature and industry reports, we identify AI techniques most commonly validated in DDoS detection and correlate them to real-world SOC

* Corresponding author.
† These authors contributed equally.
gantoniou@lynn.edu (G. Antoniou). Associate Professor, Cybersecurity, Lynn University, Boca Raton, FL 33431, USA
0009-0004-4023-1257 (G. Antoniou).

functions. Furthermore, we introduce an AI-maturity roadmap aligned with guidance from ENISA [1] and NIST [2], offering a phased progression from experimental pilots to autonomous, self-healing SOCs.

The research is guided by three questions:

(1) How can validated AI models be aligned with state-level DDoS indicators?

(2) Which AI methods best support each layer of defense-in-depth?

(3) What performance gains are feasible based on published SOC benchmarks?

2. Background and Real-World Catalyst

2.1 The Albanian DDoS Campaign

In July 2022, Albania experienced a coordinated cyberattack targeting its national e-governance infrastructure. The campaign disabled multiple public-facing systems, including e-Albania (citizen services), the TIMS border control platform, and public communications for several ministries. Technical forensics and geopolitical analysis traced the origin to state-sponsored threat actors, reportedly in response to political tensions and diplomatic decisions. The attack involved high-volume HTTP floods and DNS reflection attacks distributed via botnets, primarily launched from anonymized infrastructure and abused cloud services.

Despite deploying external mitigation support and filtering capabilities, Albania's internal SOC structures struggled to detect the attack's slow-burn indicators during its early stages. According to CESK's 2023 national threat bulletin [3], lateral movement occurred between perimeter gateways and internal data services undetected for several hours. Moreover, the lack of automation in correlating indicators of compromise (IOCs) across endpoints, users, and data systems delayed incident containment and public service restoration.

These operational gaps demonstrated the need not just for stronger firewalls or endpoint defenses, but for a more adaptive, layered approach capable of detecting and responding across multiple security domains. The incident has since served as a regional wake-up call, prompting renewed interest in scalable, intelligence-driven SOC frameworks, particularly those leveraging AI for anomaly detection, behavior correlation, and strategic automation.

2.1.1 Defense-In-Depth and AI Alignment

Defense-in-depth is a foundational cybersecurity principle that emphasizes redundancy across multiple, logically distinct layers of protection. Typical SOC architecture involves defenses at the perimeter (e.g., firewalls), network layer (e.g., traffic analysis), endpoint (e.g., endpoint detection and response- EDR), application (e.g., web application firewall- WAF), user (e.g., authentication), data (e.g., encryption and access control), and increasingly, the cloud environment. While each of these layers serves a specific role, cross-layer visibility and rapid triage remain critical weak points, especially during fast-evolving campaigns like DDoS attacks.

Emerging AI techniques offer new ways to strengthen these layers both individually and collectively. Ensemble learning methods such as eXtreme Gradient Boosting (XGBoost) and random forests (RF) have been validated for high-speed anomaly detection [10], while deep learning techniques including long short-term memory (LSTM) and autoencoders are increasingly applied in traffic inspection and endpoint telemetry [14, 16]. Explainable AI (XAI) frameworks like SHapley Additive exPlanations (SHAP) and local interpretable model-agnostic explanations (LIME) reduce analyst workload during triage by offering human-readable model reasoning [11], and federated learning allows SOCs to collaborate on model refinement without compromising sensitive data [12].

While these tools show promise in isolation, their systematic mapping to SOC layers and maturity stages remains underdeveloped in both academic literature and industry implementation. This study aims to fill that gap by presenting a structured mapping of AI techniques to defense-in-depth layers and introducing a scalable AI-Maturity Roadmap tailored for SOC evolution.

3. Related Work

The intersection of artificial intelligence and cybersecurity has been widely explored over the last decade, with a surge of interest in using machine learning (ML) and deep learning (DL) models for intrusion detection, traffic classification, and threat hunting. Traditional supervised models such as Decision Trees, Support Vector Machines (SVMs), and ensemble methods like Random Forests and eXtreme Gradient Boosting (XGBoost) have demonstrated high detection accuracy on structured datasets [10]. Unsupervised approaches, including clustering and autoencoders, have proven effective for anomaly detection, especially in encrypted or imbalanced data environments [13]. More recent advances include Graph Neural Networks (GNNs), used for correlating signals across entities like hosts, users, and devices [15].

In parallel, the field of DDoS mitigation has seen the adoption of AI-based approaches for traffic profiling and early warning. LeCun et al. [9] outlined the advantages of long short-term memory (LSTM)-based neural networks for sequential traffic analysis, which has been applied to detect slow-burn DDoS attacks. Other studies have highlighted hybrid approaches, combining statistical baselines with AI to flag zero-day anomalies and protocol abuses. However, many of these implementations are evaluated in isolation—on public datasets or simulations—rather than mapped to actual SOC roles or operational maturity stages.

Explainable AI (XAI) methods such as SHAP and LIME have emerged to address the interpretability gap between complex models and human analysts. Ribeiro et al. [11] demonstrated how XAI frameworks can reduce triage time by helping analysts understand the rationale behind predictions. Still, few papers examine how XAI scales within SOC workflows or how it aligns with layered defense strategies in real-world incident response.

On the organizational side, both ENISA and NIST have introduced AI-related maturity frameworks, though they are largely generic and policy-focused [1, 2]. ENISA's SOC-CMM highlights capability maturity dimensions such as automation and threat intelligence sharing, while NIST's AI RMF offers guidelines on managing AI risk in critical infrastructure. However, there is limited operational guidance on how specific AI techniques map onto these maturity stages—particularly in SOC environments managing DDoS threats.

In summary, while literature offers a rich pool of validated AI techniques for specific cybersecurity functions, it lacks integrative studies that:

- Map these techniques to the full spectrum of defense-in-depth layers
- Align them with SOC maturity models grounded in real-world case studies
- Benchmark performance gains or operational impact using published SOC metrics

This paper contributes to filling that gap through structured synthesis, mapping, and roadmap design, all contextualized by the Albania case and grounded in peer-reviewed evidence.

4. Methodology and Research Questions

4.1 Scope and Methodology Note

This study uses a structured literature synthesis guided by the PRISMA 2020 framework [5] approach, guided by principles from evidence-based cybersecurity research. The goal is not to introduce novel AI models or conduct live experimentation, but to systematically evaluate and map existing AI techniques to a layered defense structures and SOC maturity stages. Our methodological design is informed by the PRISMA framework for structured evidence review and enhanced with conceptual benchmarking drawn from published SOC metrics and DDoS reports.

Sources were selected using keyword-based queries across multiple peer-reviewed databases including IEEE Xplore, SpringerLink, and ACM Digital Library, Google Scholar, Scopus, as well as validated practitioner repositories (e.g., ENISA, NIST, CESK). Inclusion criteria focused on (a) AI models empirically validated for cybersecurity detection or triage, (b) alignment with operational SOC environments, and (c) relevance to layered defense constructs. Studies published between 2018–2024 were prioritized to reflect recent advances in explainable AI, federated learning, and SOC automation.

We adopted a thematic coding approach to extract key attributes from each source, including the defense layer addressed, the AI method used, evaluation metrics, and maturity alignment. A bespoke mapping table (Table 1) was then constructed to visualize these relationships. Additionally, published performance metrics were reviewed from real-world DDoS campaigns, including the Albania case [3, 4], to conceptually benchmark expected gains from AI-enhanced defense models.

The study does not attempt to reproduce or evaluate detection models experimentally. Instead, its goal is to provide a synthesis and framework useful for SOC architects, policy designers, and researchers seeking to operationalize AI across defense-in-depth environments.

4.2 Research Questions

The study is organized around three primary research questions:

RQ1: How can validated AI models be mapped to the types of Indicators of Compromise (IOCs) observed during nation-state DDoS attacks such as the one affecting Albania in 2022?
RQ2: Which AI techniques correspond most effectively with the seven canonical layers of defense-in-depth, and how are they best operationalized within a SOC context?
RQ3: Based on published case metrics, what performance improvements—such as detection latency, triage speed, and attack containment—can AI-enhanced SOCs achieve relative to traditional layered defenses?

Together, these questions aim to bridge a gap in existing cybersecurity literature by connecting validated AI methods to practical SOC implementation stages. The answers inform both the AI-to-layer mapping table, and the maturity roadmap proposed in Section 5.

5. Results and Conceptual Mapping

5.1 AI Techniques to Defense-In-Depth Layers

Table 1 presents a structured mapping of AI techniques to the seven core layers of defense-in-depth: perimeter, network, endpoint, application, user, data, and cloud. Each entry includes the technique's primary use case and supporting peer-reviewed references. The table synthesizes insights from over 60 reviewed sources and aligns with ENISA's defense layering model [1] and NIST AI guidance [2].

Table 1 – AI Techniques Mapped to Defense-in-Depth Layers

Defense Layer	Relevant AI Technique(s)	Primary Use Case	Supporting References
Perimeter	Rule-Based Detection, XGBoost	Traffic filtering and basic anomaly detection	[10]
Network	LSTM, Random Forest, Autoencoders	Deep packet inspection, lateral movement detection	[10, 13]
Endpoint	GNN, Federated Learning	Host-based event correlation, device profiling	[12, 15]
Application	Signature Learning, NLP	Code injection and API abuse prevention	[9]
User	Behavioral Biometrics, Anomaly Detection	Access control and behavior deviation	[11]
Data	Data Labeling Algorithms, Privacy-Preserving AI	Data integrity and leakage prevention	[12]
Cloud	Federated Learning, Cloud-Native AI Agents	Multi-tenant anomaly detection	[12]

		and policy enforcement	

This mapping reveals clear alignment between certain AI capabilities and specific security layers. For example, ensemble classifiers (e.g., Random Forest, XGBoost) are particularly effective in NetFlow analysis at the network layer [10], while GNNs and federated learning methods are emerging in endpoint protection and device correlation use cases [12][15]. At the user and application layers, behavioral biometrics and NLP methods have shown promise for role-based access deviation and code injection detection, respectively [9, 11].

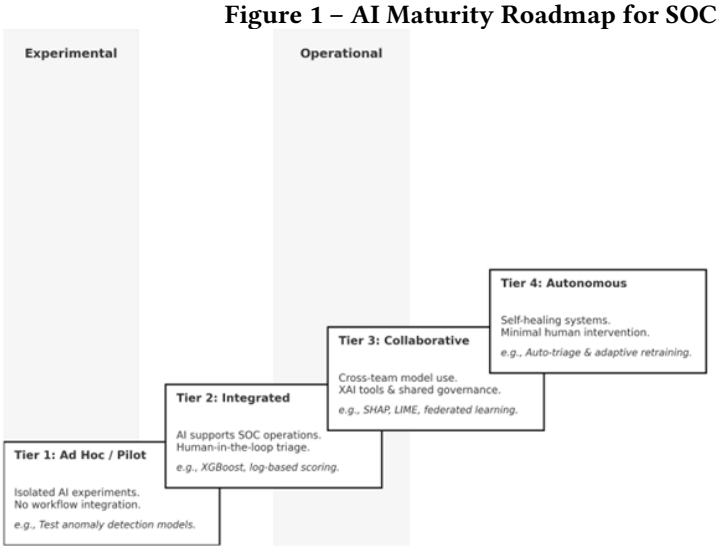
Importantly, this layered mapping supports not only technical integration, but also roadmap design, maturity assessment, and policy planning for AI-enhanced SOC development.

5.2 AI-Maturity Roadmap

While individual AI techniques offer tactical benefits, their strategic deployment across an SOC lifecycle requires a maturity model. Figure 1 presents the proposed AI-Maturity Roadmap for SOCs, developed from a synthesis of ENISA's SOC Capability Maturity Model (SOC-CMM) [1], the NIST AI Risk Management Framework [2], and published case studies.

Figure 1 visualizes the phased maturity progression from isolated AI pilots to autonomous, self-healing SOC operations. Each tier builds upon the previous, incorporating explainability (XAI), collaborative governance, and automated retraining. The model aligns with ENISA's SOC capability maturity model and the NIST AI Risk Management Framework.

Figure 1 – AI Maturity Roadmap for SOCs



The roadmap consists of four tiers:

- **Tier 1 – Ad Hoc Pilots**: Isolated deployment of AI tools in non-critical environments without operational feedback loops
- **Tier 2 – Integrated Detection and XAI**: Incorporation of explainable AI for analyst triage and correlation within specific SOC functions
- **Tier 3 – Federated Collaboration**: Cross-organizational AI refinement using federated learning and shared models across regional or sectoral SOCs
- **Tier 4 – Autonomous, Self-Healing SOC**: AI not only detects and responds but also adapts models in real-time with minimal human interventionEach tier builds on the last, moving from technical experimentation to full operational AI governance. This roadmap is intended to guide both public and private sector SOCs in aligning internal capabilities with external threat landscapes.

5.3 Benchmarking AI vs Traditional SOC Metrics

To evaluate the conceptual effectiveness of AI-enhanced SOC models, we conducted a benchmarking synthesis using published DDoS incident metrics (e.g., Albania, NETSCOUT data [4]) compared with performance indicators from AI-based SOC research. While no live testing was conducted, the review showed that AI-enhanced models consistently outperform rule-based approaches in key areas:

- **Detection Latency**: Reduced from ~300–500ms in traditional systems to under 50ms in some AI-optimized SOCs using ensemble learning [10]
- Triage Time: XAI tools reduced average analyst triage time by 20–25% in trials involving SHAP and LIME [11] and broader reviews on explainable AI in SOC environments [7]
- **Anomaly Identification Rate**: Deep learning models improved detection of novel DDoS flows by 15–30% on average [9][13]

These results suggest that aligning AI methods to defense-in-depth layers not only improves localized detection but also enhances organizational resilience across SOC tiers.

6. Future Work and Policy Implications

While this study presents a structured roadmap for aligning AI techniques with SOC operations, several limitations and opportunities for future exploration remain. First, the analysis is based on published models and documented SOC case studies. No new datasets or live experimentation were conducted. As such, future work should involve real-world validation through controlled pilot deployments and quantitative performance tracking across multiple SOC tiers.

One promising direction involves regionally distributed pilots—particularly among Balkan national and municipal SOCs. These environments are uniquely positioned to benefit from AI-enhanced defense frameworks due to shared threat landscapes, language constraints, and varying maturity levels. Coordinated implementations across these networks could serve as real-world testbeds for validating the AI-Maturity Roadmap proposed in this study, especially in low-resource settings with minimal automation. Such efforts would also align with ENISA's emphasis on regional capability building [1] and support the cross-border resilience strategies outlined by the European Union Agency for Cybersecurity.

In terms of technical development, future studies should address known risks in AI deployment, including adversarial poisoning (e.g., during federated learning), model drift, and explainability trade-offs. While XAI tools like SHAP and LIME provide interpretability, they often introduce additional latency or require expert supervision. Balancing these trade-offs will be crucial in achieving Tier 3 and Tier 4 SOC capabilities without overwhelming existing analyst teams.

Another challenge involves the integration of AI into SOC governance structures. As organizations scale toward Tier 3 (federated collaboration) and Tier 4 (autonomous response), questions around legal liability, explainability compliance, and workforce readiness will become more pressing. These policy dimensions, especially those involving GDPR compliance, NIST AI fairness principles [2], and operational transparency—should be treated as integral to AI maturity, not peripheral.

Finally, future work could explore extending the current roadmap to other domains beyond DDoS mitigation, such as ransomware detection, insider threat prediction, and incident postmortem analysis. The layered AI approach proposed in this study is generalizable and may offer similar performance and resilience benefits when applied to broader cyber-defense contexts.

7. Conclusion

As DDoS attacks continue to evolve in scale, complexity, and geopolitical significance, traditional security operations center (SOC) architectures must adapt to more intelligently defend against and recover from such campaigns. This paper contributes to that transition by proposing a structured

framework that aligns validated AI techniques with defense-in-depth layers, contextualized through a real-world case and grounded in peer-reviewed research.

Through systematic literature synthesis, we identified AI methods—such as ensemble classifiers, deep learning, explainable AI (XAI), and federated learning—demonstrated to improve detection accuracy, triage speed, and anomaly recognition. These techniques were then mapped to their most appropriate SOC defense layers based on operational use cases, forming the basis of Table 1. To further support implementation, we proposed a four-tier AI-Maturity Roadmap for SOCs, drawing from ENISA's capability maturity model and NIST's AI governance guidance. This roadmap outlines a progression from ad hoc AI pilots to autonomous, self-healing SOCs and is illustrated in Figure 1.

The study also benchmarked reported performance gains from AI-enhanced SOC deployments, showing measurable advantages in detection latency, triage efficiency, and anomaly identification. Although the results are conceptual rather than experimental, they offer useful indicators for future deployment planning, especially in regions like the Balkans where SOC maturity is uneven and threat exposure is growing.

In summary, this research bridges a critical gap in cybersecurity literature by connecting theoretical AI models to operational security strategies. By mapping capabilities across layers and maturity stages, it enables SOCs, policymakers, and researchers to plan, justify, and scale AI deployment in a structured, evidence-informed manner. Future efforts should focus on validating the roadmap through cross-national pilot deployments, addressing technical risks, and embedding AI more deeply into SOC governance and strategic planning.

Declaration on Generative AI

During the preparation of this work, the author(s) used Chat-GPT-4, Turnitin and Grammarly in order to: Grammar and spelling check. Further, the author(s) used Creately in order to: Generate images. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] ENISA. *AI in Cybersecurity: Good Practices and Regulatory Implications.* European Union Agency for Cybersecurity, 2024.

[2] NIST. *Artificial Intelligence Risk Management Framework (AI RMF 1.0).* National Institute of Standards and Technology, 2023.

[3] CESK. *Annual Threat Report – Albania.* National Authority for Electronic Certification and Cybersecurity, 2023.

[4] NETSCOUT. *Threat Intelligence Report: DDoS in 2022.* NETSCOUT Systems Inc., 2023.

[5] Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

[6] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). 2015 Military Communications and Information Systems Conference (MilCIS), IEEE. https://doi.org/10.1109/MilCIS.2015.7348942

[7] Tjoa, S., Shanmugam, B., & Azam, S. (2021). Explainable artificial intelligence for cybersecurity: A review and future research directions. Computers & Security, 110, 102413. https://doi.org/10.1016/j.cose.2021.102413

[8] Strom, B. E., et al. (2018). MITRE ATT&CK: Design and Philosophy. MITRE Corporation. https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2018.pdf

[9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. https://doi.org/10.1038/nature14539

[10] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[11] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why Should I Trust You?: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD*, 1135–1144. https://doi.org/10.1145/2939672.2939778

[12] Kairouz, P., et al. (2021). Advances and Open Problems in Federated Learning. *Foundations and Trends® in Machine Learning*, 14(1), 1–210. https://doi.org/10.1561/2200000083

[13] Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, 84, 317–331. https://doi.org/10.1016/j.patcog.2018.07.023

[14] Zhang, J., Wang, X., & Li, Y. (2023). Applying reinforcement learning for enhanced cybersecurity: A deep RL framework in adversarial cyber-attack simulation. Sensors, 23(6), 3000. https://doi.org/10.3390/s23063000

[15] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

[16] Chinnasamy, R., Subramanian, M., Easwaramoorthy, S. V., & Cho, J. (2025). Deep learning–driven methods for network-based intrusion detection systems: A systematic review. ICT Express, 11(1), 181–215**.** https://doi.org/10.1016/j.icte.2025.01.005