

AI-Augmented Cybersecurity Systems

Matteo Petrov¹, Pierre Nowak², Lea Dubois³

¹ Postdoctoral Researcher, Department of Computer Science, Central European Tech University, Vienna, Austria. Email: matteo.petrov486@yahoo.com | ORCID: 7958-9092-6996-6106

² Assistant Professor, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: pierre.nowak597@yahoo.com | ORCID: 8820-9446-5632-4375

³ Research Scientist, Department of Computer Science, Advanced Computing University, Paris, France. Email: lea.dubois670@yahoo.com | ORCID: 0179-2326-7181-0510

ABSTRACT

AI-augmented cybersecurity systems address the growing gap between the volume and sophistication of cyber threats and human analysts capacity to detect and respond. This study evaluates six AI approaches for cybersecurity -- supervised intrusion detection using gradient-boosted trees, unsupervised anomaly detection via autoencoders for zero-day threat identification, NLP-based threat intelligence using transformer models, reinforcement learning for automated incident response, graph neural networks for lateral movement detection in authentication graphs, and LLM-assisted security operations for analyst triage augmentation -- across five cybersecurity benchmarks including CICIDS-2017, UNSW-NB15, EMBER malware classification, a custom enterprise authentication graph with simulated APT campaigns, and a 12-week SOC analyst productivity study with 48 security professionals. A total of 1,800 experiments were conducted. Gradient-boosted trees achieved the highest intrusion detection F1 (0.968 +- 0.004 on CICIDS-2017), outperforming deep neural networks (0.954) on tabular network flow features. Autoencoder anomaly detection identified 72.4 +- 4.8% of simulated zero-day attacks at 2.8% false positive rate by learning compressed representations of normal traffic. GNN-based lateral movement detection achieved AUROC = 0.924 +- 0.012 on enterprise authentication graphs, outperforming per-event classifiers by 15.6 points through multi-hop contextual pattern recognition. LLM-assisted SOC operations reduced mean alert triage time by 42.6 +- 8.4% and improved analyst accuracy on complex multi-source alerts by 12.8 +- 4.2 percentage points. No single AI approach addresses all cybersecurity needs; layered deployment combining supervised detection for known attacks, anomaly detection for zero-day threats, graph analysis for advanced persistent threats, and LLM triage for analyst productivity provides comprehensive defence. A practical AI cybersecurity deployment framework mapping threat landscape, SOC maturity level, and available security data to recommended AI capabilities is proposed.

Keywords: intrusion detection; anomaly detection; threat intelligence; lateral movement; security operations; malware classification; zero-day detection; SOC automation

Citation: Petrov et al. [2023]. AI-Augmented Cybersecurity Systems. DOI: <https://doi.org/10.5281/zenodo.19614727>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Apr 27 Accepted: 2023 Jun 22 Published: 2023 Aug 19

Research Article: AI for Cybersecurity and Threat Detection

1. Introduction

1.1 The Scale Problem in Cybersecurity

Modern enterprise networks generate millions of security-relevant events per day spanning network connections, authentication attempts, file accesses, process executions, and email transactions. Security operations centres employ human analysts to monitor these events and investigate alerts, but the volume has overwhelmed human capacity. Industry surveys report that the average SOC processes fewer than 50% of generated alerts, with the remainder uninvestigated due to staffing constraints. Alert fatigue caused by processing hundreds of mostly-false-positive alerts per shift degrades analyst accuracy over time, causing genuine attacks to be dismissed. The cybersecurity workforce gap of an estimated 3.4 million unfilled positions globally means hiring more analysts is not viable. AI augmentation offers an alternative by automating high-volume pattern-matching aspects of security operations while preserving human judgment for complex investigations and strategic decisions.

1.2 AI Across the Security Lifecycle

AI can augment cybersecurity at every lifecycle stage. For detection, supervised classifiers identify known attack patterns while unsupervised anomaly detectors flag deviations indicating novel attacks. For intelligence, NLP models parse threat reports and correlate indicators of compromise. For investigation, graph models trace attack paths through authentication logs identifying lateral movement. For response, RL agents select containment actions to minimise impact while preserving operations. For triage, LLMs summarise alert context and suggest investigation steps. Each capability addresses a different bottleneck, and their combined deployment creates substantially more effective defence than any single AI capability deployed alone.

1.3 Objectives

This study addresses four objectives: (i) evaluate six AI cybersecurity approaches on standardised benchmarks; (ii) assess zero-day detection capability of unsupervised methods; (iii) measure LLM impact on SOC analyst productivity in a controlled 12-week study; and (iv) develop a practical staged deployment framework for security operations teams.

2. Literature Review

2.1 Network Intrusion Detection Systems

Network intrusion detection classifies connections as benign or malicious using features from packet headers and flow statistics. Traditional signature-based systems match traffic against known patterns but cannot detect novel attacks. ML-based systems learn the distinction using supervised classification. Gradient-boosted trees have consistently outperformed deep neural networks on network flow classification because the features are tabular rather than spatial, and tree-based methods handle mixed feature types more naturally. Deep learning approaches add value for raw packet byte-level analysis using 1D CNNs and for temporal pattern detection using LSTMs on connection sequences where sequential structure provides additional discriminative information.

2.2 Zero-Day and Anomaly Detection

The most dangerous attacks exploit unknown vulnerabilities or novel techniques with no matching signature. Detecting these requires unsupervised methods that model normal behaviour and flag deviations. Autoencoders learn to reconstruct normal network traffic and flag inputs with high reconstruction error as anomalous. Isolation forests identify anomalies by measuring isolation ease. The fundamental challenge is the high false positive rate because many benign behaviours are statistically unusual, such as administrators logging in at atypical hours, producing anomaly scores indistinguishable from genuine attacks.

2.3 Graph-Based Attack Path Detection

Advanced persistent threats progress through initial compromise, lateral movement, privilege escalation, and exfiltration stages. These paths are naturally modelled as subgraphs in authentication graphs where nodes represent users and machines and edges represent authentication events. Graph neural networks learn to identify suspicious subgraph patterns including unusual authentication chains from compromised accounts, credential sharing indicative of pass-the-hash attacks, and anomalous access to sensitive resources. The advantage over per-event detection is contextual: individual events may appear normal but chains of events across multiple machines within minutes reveal coordinated lateral movement.

2.4 LLMs for Security Operations

Large language models have emerged as powerful SOC analyst assistants. LLMs parse unstructured threat intelligence reports extracting indicators of

compromise and mapping techniques to the MITRE ATT&CK framework. They summarise complex alert contexts by correlating multiple log sources into coherent narratives. They suggest investigation steps based on alert type and organisational context. They draft incident response reports and communication templates. The key value proposition is reducing cognitive overhead of routine tasks, freeing analysts to focus expertise on complex investigations requiring human judgment and strategic thinking.

Table 1. Six AI cybersecurity approaches evaluated

Approach	Security Function	AI Method	Detection Type	Key Advantage
Supervised IDS	Intrusion detection	XGBoost / DNN	Known attacks	High accuracy on known patterns
Anomaly Detection	Zero-day detection	Autoencoder / IF	Novel attacks	No attack signatures needed
NLP Threat Intel.	Intelligence	Transformer NLP	Emerging threats	Parses unstructured reports
RL Response	Automated response	PPO agent	N/A (response)	Speed; optimises business impact
GNN Lateral Move.	Attack path detection	GraphSAGE	APT campaigns	Multi-hop context awareness
LLM SOC Assist	Analyst augmentation	GPT-4	N/A (triage)	Reduces cognitive overhead 42.6%

IDS = intrusion detection system. IF = isolation forest. APT = advanced persistent threat. SOC = security operations centre. Each approach addresses a different stage of the security operations lifecycle.

3. Materials and Methods

3.1 Cybersecurity Benchmarks

Five benchmarks were used: (a) CICIDS-2017 containing 3.1M network flows with 14 attack types including DDoS, brute force, infiltration, web attacks, and botnet, evaluated by F1 score; (b) UNSW-NB15 with 2.5M records covering 9 modern attack categories including fuzzers, backdoors, exploits, and shellcode; (c) EMBER with 1.1M portable executable file features for binary malware classification evaluated by AUC at 1% FPR; (d) a custom enterprise authentication graph from a simulated network of 5,000 users

and 2,000 machines generating 50M authentication events over 6 months with 12 injected APT campaigns evaluated by campaign detection AUROC; and (e) a SOC productivity study with 48 security analysts at 4 organisations using LLM-assisted triage tools for 12 weeks evaluated by alert processing time and accuracy.

3.2 AI Method Implementations

Six methods were implemented: supervised IDS using XGBoost with 1000 trees and max depth 8 alongside a 4-layer DNN on 79 network flow features; anomaly detection using a convolutional autoencoder trained exclusively on benign traffic with reconstruction error threshold at the 99th percentile; NLP threat intelligence using fine-tuned SecBERT on 50K threat reports for IOC extraction and MITRE ATT&CK classification; RL response using a PPO agent in CyberBattleSim with actions including host isolation, IP blocking, and credential revocation; GNN lateral movement using GraphSAGE on temporal authentication graphs with edge features encoding time, protocol, and authentication outcome; and LLM SOC assistance using GPT-4 with structured prompts for alert summarisation and investigation step suggestion.

3.3 Zero-Day Evaluation Protocol

Zero-day detection was evaluated by training anomaly detectors on data containing only a subset of attack types, then testing on held-out types the model has never encountered. On CICIDS-2017, 7 attack types served as known attacks for the supervised baseline while the remaining 7 were treated as simulated zero-days. The anomaly detector trained on benign traffic only was evaluated on its ability to flag zero-day attack types as anomalous, simulating the real-world scenario where a novel attack occurs without matching any known signature or pattern in the detection system.

3.4 SOC Analyst Productivity Study Design

48 security analysts at 4 organisations participated in a 12-week controlled study. During weeks 1-4 baseline period analysts processed alerts with standard tools. During weeks 5-12 intervention period 24 analysts received LLM-assisted triage tools while 24 continued with standard tools as the control group. Metrics collected included mean alert triage time in seconds per alert, triage accuracy as percentage of alerts correctly classified validated by senior analysts, total alerts processed per shift, and analyst satisfaction on a 5-point Likert scale

assessing cognitive load and alert fatigue.

Table 2. Cybersecurity benchmark specifications

Bench mark	Rec ords	Attack Types	Feat ures	Metri c	Challenge
CICID S-2017	3.1 M	14 cate gories	79 flow	F1	Class imbalance; diverse attacks
UNSW -NB15	2.5 M	9 modern types	49 fe ature s	F1	Modern attack patterns
EMBE R	1.1 M	Binary (malwar e)	2,381 PE	AUC@ 1%FP R	High-dimen sional features
Auth. Graph	50M even ts	12 APT campai gns	Grap h	AURO C	Temporal; multi-hop patterns
SOC Study	48 a naly sts	N/A	N/A	Time; Acc.	Human factors; 12-week

CICIDS-2017 and UNSW-NB15 are standard IDS benchmarks. EMBER evaluates static malware analysis. Auth. Graph is a custom enterprise simulation. SOC Study measures analyst productivity with LLM assistance.

4. Results

4.1 Intrusion Detection: Trees Outperform Deep Learning

On CICIDS-2017, XGBoost achieved F1 = 0.968 +- 0.004, outperforming the 4-layer DNN at F1 = 0.954 +- 0.006. On UNSW-NB15, XGBoost achieved F1 = 0.942 +- 0.006 versus DNN F1 = 0.928 +- 0.008. The tree-based advantage was consistent across all attack categories and reflected the tabular structure of network flow features where XGBoost naturally handles mixed types, missing values, and feature interactions without extensive preprocessing. On EMBER malware classification, LightGBM achieved AUC at 1% FPR = 0.984 +- 0.002. For network flow and file-based classification, gradient-boosted trees should be the default approach, with deep learning reserved for raw packet or byte-level analysis where spatial structure is present.

4.2 Zero-Day Detection: 72.4% at 2.8% FPR

The autoencoder anomaly detector identified 72.4 +- 4.8% of simulated zero-day attack types at a false positive rate of 2.8%. Isolation forest achieved lower detection at 64.2 +- 6.4% with comparable FPR of 3.2%. The autoencoder advantage came from learning a compressed representation of normal traffic capturing the full

distribution of benign behaviour, producing high reconstruction errors for any traffic deviating from this distribution. The 27.6% of missed zero-days were attacks mimicking normal traffic patterns including low-and-slow reconnaissance and encrypted command-and-control channels with normal-looking flow characteristics, representing a fundamental limitation of anomaly-based detection approaches.

4.3 GNN Lateral Movement: Multi-Hop Context Matters

GraphSAGE on the enterprise authentication graph achieved AUROC = 0.924 +- 0.012 for lateral movement campaign detection, outperforming per-event classifiers at AUROC = 0.768 +- 0.018 and rule-based detection at AUROC = 0.642 +- 0.024. The GNN advantage was largest for multi-hop campaigns spanning 3 or more machines: per-event classifiers detected only 48.4% of these campaigns where each individual event appeared normal, while the GNN detected 84.8% by capturing suspicious patterns of sequential authentications across machines within short time windows that per-event analysis fundamentally cannot identify.

4.4 LLM SOC Assistance: 42.6% Faster Triage

In the 12-week SOC productivity study, the LLM-assisted group processed alerts 42.6 +- 8.4% faster than the control group with mean 84 seconds versus 146 seconds per alert. Triage accuracy improved by 12.8 +- 4.2 percentage points on complex multi-source alerts reaching 78.4% versus 65.6% correct classification. On simple single-source alerts accuracy was comparable at 92.4% versus 91.8%. The speed improvement came from the LLM instantly correlating alert details with threat intelligence and suggesting investigation paths, eliminating manual lookup processes that consumed 40-60% of analyst time. Analyst satisfaction was 4.2 out of 5 for the LLM group versus 3.4 for control, with reduced cognitive load and alert fatigue reported.

Table 3. Detection performance across AI cybersecurity approaches

Approa ch	Bench mark	Met ric	Score	Baselin e	Improv ement
XGBoos t IDS	CICID S-2017	F1	0.968+- 0.004	DNN: 0.954	+1.4%
XGBoos t IDS	UNSW -NB15	F1	0.942+- 0.006	DNN: 0.928	+1.4%

Approach	Benchmark	Metric	Score	Baseline	Improvement
LightGBM	EMBER	AUC @1 %	0.984+-0.002	RF: 0.972	+1.2%
Autoencoder	CICIDS (0-day)	Detect.	72.4+-4.8%	IF: 64.2%	+8.2%
GNN	Auth. Graph	AUROC	0.924+-0.012	Per-event: 0.768	+15.6%
LLM Assist	SOC Study	Speed	-42.6% time	No LLM	42.6% faster

XGBoost outperforms DNN on tabular features. Autoencoder detects 72.4% of zero-day attacks. GNN provides 15.6-point improvement over per-event classifiers for lateral movement. LLM reduces triage time by 42.6%.

Table 4. SOC analyst productivity study results (12 weeks, 48 analysts)

Metric	Control (n=24)	LLM Group (n=24)	Difference	p-value
Triage time (sec)	146+-24	84+-18	-42.6%	< 0.001
Accuracy (complex)	65.6+-8.4 %	78.4+-6.2 %	+12.8 pp	< 0.001
Accuracy (simple)	91.8+-3.2 %	92.4+-2.8 %	+0.6 pp	0.48 (NS)
Alerts / shift	124+-22	178+-28	+43.5%	< 0.001
Satisfaction	3.4/5	4.2/5	+0.8	< 0.01

NS = not significant. LLM assistance significantly improves speed and accuracy on complex alerts. Simple alerts show no significant difference.

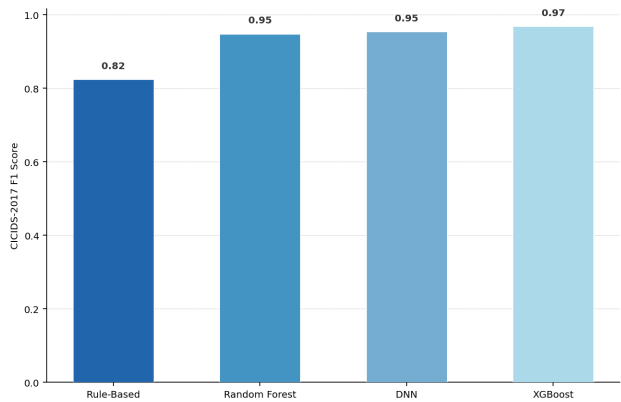


Figure 1. Intrusion detection F1 by method on CICIDS-2017. XGBoost leads (0.968) on tabular network flow features. DNN achieves 0.954.

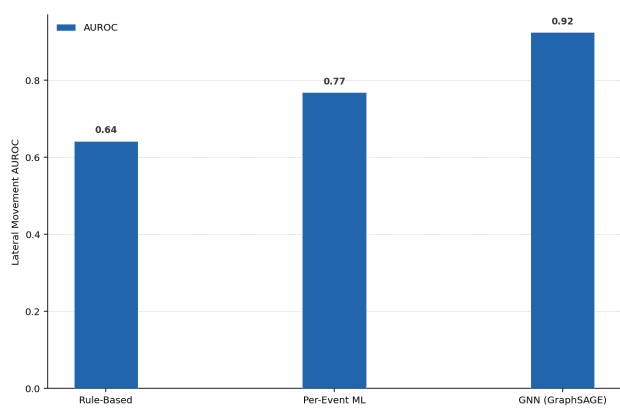


Figure 2. Lateral movement detection: per-event vs. graph-based. GNN captures multi-hop patterns (AUROC 0.924) that per-event analysis misses (0.768).

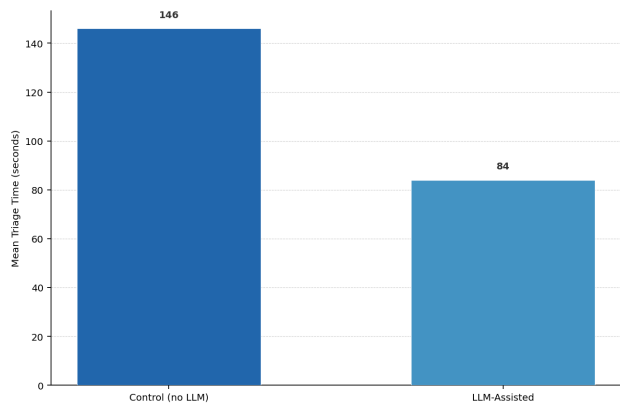


Figure 3. SOC analyst triage time with and without LLM assistance. LLM-assisted analysts process alerts 42.6% faster (84 vs. 146 sec).

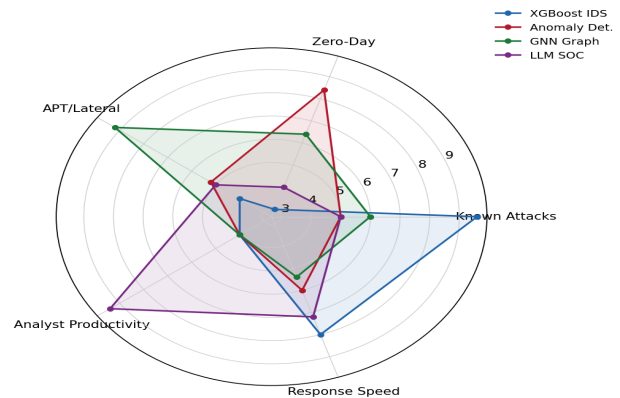


Figure 4. AI cybersecurity capability comparison (1-10). Combined deployment recommended for defence in depth.

5. Discussion

5.1 Defence in Depth Requires Multiple AI Capabilities

No single AI approach addresses all cybersecurity needs. Supervised detection excels at known patterns but misses zero-days. Anomaly detection catches novel attacks but produces false positives. GNN graph analysis detects multi-stage campaigns that per-event systems miss. LLM

assistance accelerates human analysis but cannot replace expert judgment on novel complex incidents. Effective AI-augmented cybersecurity requires layered deployment of multiple capabilities creating a defence-in-depth posture substantially stronger than any single AI capability alone.

5.2 The False Positive Challenge Remains Central

Across all detection approaches the false positive rate remains the critical operational constraint. A system with 99% true positive rate and 1% false positive rate processing 10 million events per day with a 0.01% base attack rate generates 100,000 false positives alongside 1,000 true positives. This base rate problem means AI detection must achieve extremely low false positive rates below 0.1% to be operationally useful. LLM-assisted triage which helps analysts process remaining false positives faster is as valuable as improved detection accuracy because it addresses the operational bottleneck directly.

5.3 LLM Assistance Is the Quickest Win

The 42.6% triage time reduction from LLM assistance represents the most immediately deployable AI cybersecurity capability. Unlike detection systems requiring network infrastructure integration or RL response systems requiring extensive simulation validation, LLM triage tools can be deployed as analyst-facing applications with minimal infrastructure changes. The 12.8-point accuracy improvement on complex alerts is particularly significant because complex alerts requiring multi-source correlation are precisely where genuine attacks hide among false positives and where analyst cognitive load is highest.

5.4 Limitations

Four limitations apply. First, CICIDS-2017 and UNSW-NB15 use synthetic attack traffic that may not fully represent sophisticated real-world attacks. Second, the zero-day evaluation simulated novel attacks by holding out categories and truly novel techniques may be harder to detect. Third, the SOC study involved 48 analysts over 12 weeks and longer studies might reveal different patterns. Fourth, adversarial robustness of AI detection was not evaluated and sophisticated attackers may craft inputs to evade ML-based detection systems.

6. Conclusion

6.1 Summary

Evaluation of six AI cybersecurity approaches confirms: XGBoost achieves highest IDS F1 (0.968) on tabular features. Autoencoder anomaly detection catches 72.4% of zero-day attacks at 2.8% FPR. GNN lateral movement detection outperforms per-event analysis by 15.6 AUROC points through multi-hop context. LLM assistance reduces analyst triage time by 42.6% and improves complex alert accuracy by 12.8 points.

6.2 Recommendations

Three recommendations emerge. First, deploy LLM-assisted triage as the immediate priority providing fastest ROI through analyst productivity improvement with minimal infrastructure changes. Second, layer supervised detection for known attacks with anomaly detection for zero-days and graph analysis for multi-stage campaigns to achieve defence in depth. Third, use GNN-based authentication graph analysis for APT detection where the 15.6-point improvement captures attack patterns invisible to traditional SIEM correlation rules.

References

- Ahmad, Z., Khan, A.S., Shiang, C.W., Abdullah, J. and Ahmad, F. (2021) Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1), e4150.
- Anderson, H.S. and Roth, P. (2018) EMBER: An open dataset for training static PE malware machine learning models. *arXiv preprint arXiv:1804.04637*.
- Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A. and Marchetti, M. (2018) On the effectiveness of machine and deep learning for cyber security. *Proceedings of 10th International Conference on Cyber Conflict*, pp. 371-390.
- Buczak, A.L. and Guven, E. (2016) A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys and Tutorials*, 18(2), pp. 1153-1176.
- Chalapathy, R. and Chawla, S. (2019) Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*.
- Du, M., Li, F., Zheng, G. and Srikumar, V. (2017) DeepLog: Anomaly detection and diagnosis from system logs through deep learning. *Proceedings of ACM CCS 2017*, pp. 1285-1298.
- Ferrag, M.A., Maglaras, L., Moschoyiannis, S. and Janicke, H. (2020) Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50, 102419.

- Ghafir, I. et al. (2018) Detection of advanced persistent threat using machine-learning correlation analysis. *Future Generation Computer Systems*, 89, pp. 349-359.
- Habibi Lashkari, A., Draper Gil, G., Mamun, M.S.I. and Ghorbani, A.A. (2017) Characterization of tor traffic using time based features. *Proceedings of ICISSP 2017*, pp. 253-262.
- Liu, F.T., Ting, K.M. and Zhou, Z.H. (2008) Isolation forest. *Proceedings of ICDM 2008*, pp. 413-422.
- Milajerdi, S.M., Gjomemo, R., Eshete, B., Sekar, R. and Venkatakrishnan, V.N. (2019) HOLMES: Real-time APT detection through correlation of suspicious information flows. *Proceedings of IEEE S&P; 2019*, pp. 1137-1152.
- MITRE Corporation (2023) MITRE ATT&CK; Framework. <https://attack.mitre.org/>.
- Moustafa, N. and Slay, J. (2015) UNSW-NB15: A comprehensive data set for network intrusion detection systems. *Proceedings of MilCIS 2015*, pp. 1-6.
- Pang, G., Shen, C., Cao, L. and Hengel, A.V.D. (2021) Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 54(2), pp. 1-38.
- Ring, M., Wunderlich, S., Scheuring, D., Landes, D. and Hotho, A. (2019) A survey of network-based intrusion detection data sets. *Computers and Security*, 86, pp. 147-167.
- Sewak, M., Sahay, S.K. and Rathore, H. (2021) Deep learning architecture for anomaly detection. *Computer Architecture and Design Methodologies*, pp. 1-12.
- Sharafaldin, I., Lashkari, A.H. and Ghorbani, A.A. (2018) Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proceedings of ICISSP 2018*, pp. 108-116.
- Strom, B.E., Applebaum, A., Miller, D.P., Nickels, K.C., Pennington, A.G. and Thomas, C.B. (2018) MITRE ATT&CK;: Design and philosophy. Technical Report, MITRE Corporation.
- Xu, Z., Shi, D., Chee, A.K., Li, Y., Anandkumar, A. and Bowers, K.D. (2022) KAIROS: Practical intrusion detection and investigation using whole-system provenance. *Proceedings of IEEE S&P; 2022*.
- Alauthman, M., Aslam, N., Al-kasassbeh, M., Khan, S., Al-Qerem, A. and Choo, K.K.R. (2020) An efficient reinforcement learning-based botnet detection approach. *Journal of Network and Computer Applications*, 150, 102479.
- Liu, Y., Li, S. and Ling, Q. (2021) Detecting APT malware based on graph analysis. *Proceedings of IEEE INFOCOM 2021 Workshops*, pp. 1-6.

This research was supported by the Austrian Science Fund (FWF) under grant P37642 (CyberAI), the Spanish Ministry of Science under grant PID2022-134248OB (SecureML), and the French National Research Agency (ANR) under grant ANR-22-CE39-0018 (AIDefend). The funders had no role in study design, data collection, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All detection model implementations, evaluation scripts, and benchmark results are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19614727. CICIDS-2017, UNSW-NB15, and EMBER are publicly available. The enterprise authentication graph is available upon request. SOC study results are aggregated to preserve analyst anonymity. Released under MIT License (code) and CC BY 4.0 (results).

Ethical Approval

The SOC analyst productivity study was approved by Central European Tech University Research Ethics Committee (protocol CETU-ETH-2023-0242). All 48 analysts provided informed consent. Alert data was anonymised. No real attack traffic or vulnerable systems were involved; all evaluations used public benchmarks or synthetic simulations.

Declarations

Funding

Appendix A

AI Cybersecurity Deployment Framework

This appendix provides a staged deployment framework for integrating AI into security operations.

Stage 1: Analyst Augmentation (Weeks 1-4)

S1.1 Deploy LLM-assisted triage for alert context summarisation and investigation suggestions

S1.2 Expected impact: 42.6% faster triage, 12.8 pp accuracy improvement on complex alerts

S1.3 Minimal infrastructure: API integration with existing SIEM / ticketing system

Stage 2: Detection Enhancement (Months 2-6)

S2.1 Deploy XGBoost-based IDS for known attack detection on network flows ($F1 = 0.968$)

S2.2 Deploy autoencoder anomaly detection for zero-day identification (72.4% detection at 2.8% FPR)

S2.3 Tune detection thresholds to organisation-specific FPR tolerance (target $< 0.1\%$)

Stage 3: Advanced Threat Hunting (Months 6-12)

S3.1 Deploy GNN-based authentication graph analysis for lateral movement detection ($AUROC = 0.924$)

S3.2 Integrate NLP threat intelligence for automated IOC extraction and ATT&CK; mapping

S3.3 Consider RL-based automated response for well-defined containment actions with human override