UDC 004.056:004.85

THE TRANSFORMATIVE IMPACT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING ON CONTEMPORARY NETWORK SECURITY PARADIGMS

ТРАНСФОРМАЦІЙНИЙ ВПЛИВ ШТУЧНОГО ІНТЕЛЕКТУ ТА МАШИННОГО НАВЧАННЯ НА СУЧАСНІ ПАРАДИГМИ МЕРЕЖЕВОЇ БЕЗПЕКИ

Levinskyi Maksym[1], Shapo Vladlen[2], Levinskyi Valeriy[3], Volovshchykov Valeriy[4]

NU «Odesa Maritime Academy»[1]; Naval Institute NU «Odesa Maritime Academy»[2];
Odesa National University of Technology[3], Odesa, Ukraine
National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine[4]
ORCID: https://orcid.org/0000-0002-6544-5110[1], https://orcid.org/0000-0002-3921-4159 [2]
https://orcid.org/0000-0002-3563-528X [3], https://orcid.org/0000-0003-4454-2314 [4]
E-mail: MaxLevinskyi@gmail.com[1], Vladlen.Shapo@gmail.com[2], ValeryLevinskyi@gmail.com[3],
Valeriy.Volovshchykov@khpi.edu.ua[4]

DOI: 10.15673/atbp.v17i4.3322

Abstract *- This work provides a comprehensive overview and analysis of the pivotal role of Artificial Intelligence (AI) and Machine Learning (ML) in revolutionizing network security. As cyber threats escalate in complexity and volume, traditional, primarily reactive security measures demonstrate inherent limitations. AI/ML offers potent alternatives by analyzing vast datasets to identify subtle compromise patterns, detect anomalies deviating from normal baselines, and orchestrate autonomous, timely responses to evolving threats with reduced human intervention. The core objective of this work is to systematize and analyze current knowledge on key AI/ML applications in this domain, including intelligent intrusion detection/prevention systems (IDPS), advanced phishing mitigation, robust malware classification, and behavioral analytics for identifying insider threats and anomalous activities. While highlighting the substantial defensive enhancements AI/ML introduces, the paper critically examines significant adoption challenges: training data quality and representativeness, optimizing the balance between false positives and negatives, the persistent threat of adversarial attacks designed to deceive ML models, and crucial ethical considerations regarding data privacy, algorithmic bias, and the need for transparency. Emerging trends like privacy-preserving learning, the drive towards Explainable AI (XAI) for trustworthy decisions, and increasingly sophisticated automated security operations (SecOps) indicate a future where AI is fundamentally woven into resilient, adaptive network defenses. The key outcome of this analysis is the confirmation of a paradigm shift from static, signature-based protection towards dynamic, continuously learning security frameworks capable of co-evolving with the perpetually shifting cyber threat landscape.*

Анотація *– ця стаття надає комплексний огляд та аналіз ключової ролі штучного інтелекту (ШІ) та машинного навчання (МН) у революціонізуванні мережевої безпеки. Оскільки кіберзагрози зростають за складністю та обсягом, традиційні, переважно реактивні заходи безпеки демонструють властиві обмеження. ШІ/МН пропонують потужні альтернативи, аналізуючи величезні набори даних для виявлення тонких шаблонів, виявлення аномалій, що відхиляються від нормальних базових варіантів подій, та оркеструючи автономні, своєчасні відповіді на еволюціонуючі загрози зі зменшеним втручанням людини. Основна мета цієї роботи – систематизувати та проаналізувати поточні знання щодо ключових застосувань ШІ/МН у цій сфері, включаючи інтелектуальні системи виявлення/запобігання вторгненням (IDPS), передові методи пом'якшення фішингу, надійну класифікацію шкідливого ПЗ та поведінкову аналітику для ідентифікації внутрішніх загроз та аномальної активності. Підкреслюючи значні оборонні покращення, які вносять ШІ/МН, стаття критично аналізує суттєві виклики впровадження: якість та репрезентативність навчальних даних, оптимізацію балансу між хибними спрацьовуваннями та пропусками загроз, постійну загрозу ворожих атак, спрямованих на обман*

моделей МН, та важливі етичні міркування щодо конфіденційності даних, алгоритмічної упередженості та потреби у прозорості. Нові тенденції, такі як навчання для збереження конфіденційності, прагнення до пояснюваного ШІ (XAI) для надійних рішень, та все більш складні автоматизовані операції безпеки (SecOps), вказують на майбутнє, де ШІ буде фундаментально вплетений у стійкі, адаптивні мережеві захисти. Ключовим результатом цього аналізу є підтвердження зміни парадигми від статичного, сигнатурного захисту до динамічних, постійно навчаючихся фреймворків безпеки, здатних спів-еволюціонувати з постійно мінливим кіберландшафтом загроз.
Keywords: artificial intelligence (AI); machine learning (ML); network security; cybersecurity; intrusion detection systems (IDS); intrusion prevention systems (IPS); Explainable AI (XAI)

Ключові слова: штучний інтелект (ШІ); машинне навчання (МН); мережева безпека; кібербезпека; системи виявлення вторгнень (IDS); системи запобігання вторгненням (IPS); пояснюваний ШІ (XAI)

I. INTRODUCTION

Robust network security is foundational in today's hyper-connected digital world, safeguarding sensitive data, ensuring operational continuity of critical infrastructure, and preserving trust in online interactions. As organizations increasingly rely on complex digital infrastructures, the cyber threat landscape has expanded alarmingly in frequency, volume, and sophistication. These challenges extend beyond traditional IT infrastructure to specialized industrial systems, where AI-driven diagnostics—such as those enhancing the reliability of ship electric power installations [1]—demonstrate parallel advancements in preemptive failure detection that complement security frameworks. Threat actors range from individuals to organized crime and state-sponsored entities. Traditional defenses like firewalls, signature-based IDS, and antivirus software struggle against advanced, polymorphic, or zero-day attacks. Their reliance on static rules limits agility against rapidly evolving adversary tactics, techniques, and procedures (TTPs). This capability gap drives demand for more intelligent, predictive, and adaptive security solutions. Artificial Intelligence (AI) and its subfield Machine Learning (ML) offer transformative potential here. They enable systems to learn complex patterns from vast network data (packets, flows, logs), discern subtle anomalies indicative of threats often missed by humans or traditional tools, and initiate timely, often automated responses, significantly reducing detection/response times with less direct human oversight.

II. LITERATURE ANALYSIS

Understanding modern network security requires acknowledging the dynamic threat environment characterized by [2, 3, 4]:

- Increased Sophistication: polymorphic malware, fileless attacks, advanced persistent threats (APTs), and refined social engineering tactics are common;
- Volume and Velocity: automated tools launch attacks (DDoS, brute-force, scans) at scales overwhelming manual defenses;
- Expanded Attack Surface: proliferation of IoT, cloud adoption, mobile workforces, and interconnected supply chains increase potential vulnerabilities;
- Zero-Day Exploits: attacks targeting unknown vulnerabilities bypass signature-based defenses entirely.

Conventional security technologies face inherent limitations:

- Reactive posture: signature-based tools require prior threat knowledge, leaving systems vulnerable to novel attacks.
- Inability to handle polymorphism: malware designed to change signatures evades detection.
- Scalability issues: manual log analysis is often infeasible; rule-based systems become unwieldy.
- High false positive rates: simple anomaly rules generate noise, leading to alert fatigue.
- Context-Blindness: rule-based systems lack contextual understanding to differentiate malicious from unusual-but-legitimate behavior.

III. THE PURPOSE AND TASKS OF THE RESEARCH

The primary objective of this paper is to provide a structured overview and critical analysis of the transformative impact of AI and ML on modern network security paradigms. This work synthesizes current knowledge by exploring the fundamental principles, examining key applications and enabling technologies, dissecting significant implementation challenges, and highlighting emerging trends and future research directions. Given the rapid evolution of both cyber threats and AI capabilities, including significant advancements reported in recent years [5], such a synthesis is crucial for understanding the current state-of-the-art and navigating the complexities of adopting these powerful technologies effectively.

IV. RESEARCH METHODS AND MATERIALS

These limitations create a compelling case for intelligent systems capable of adaptive, behavior-based threat detection.

AI aims to engineer systems with human-like cognitive abilities (learning, reasoning, problem-solving). ML, a core AI subfield, develops algorithms enabling systems to learn patterns from data without explicit programming for every scenario. Key ML paradigms used in security include:

- Supervised Learning: uses labeled data (e.g., 'attack' vs 'normal' traffic) to learn a mapping function for classifying new, unlabeled data. Algorithms include Support Vector Machines (SVMs), Decision Trees/Random Forests, and Artificial Neural Networks (ANNs);

- Unsupervised Learning: uses unlabeled data to discover hidden structures or anomalies. Crucial for detecting deviations from normal behavior without prior attack knowledge. Algorithms include K-Means, DBSCAN, PCA, and Autoencoders;
- Reinforcement Learning (RL): an agent learns optimal actions through trial-and-error interaction with an environment, receiving rewards/penalties. Potential applications include dynamic firewall optimization and automated incident response strategy development.

Applied to voluminous network data (packet captures, flows, logs, telemetry, threat feeds), ML excels at identifying subtle, non-linear patterns indicative of threats, including zero-day exploits, shifting defense towards proactive capabilities. Deep learning variants (CNNs, RNNs/LSTMs, Transformers) further enhance this potential, with recent studies demonstrating improved performance on complex tasks [6].

AI/ML integration significantly enhances threat detection, prevention, and response across various domains. Intelligent Intrusion Detection and Prevention Systems (IDPS): ML models augment traditional IDPS by learning complex patterns of normal and malicious behavior from network traffic. This improves detection of known attacks and enables identification of novel threats via anomaly detection (e.g., unusual port usage, data transfer patterns). Dynamic firewall rules or traffic shaping can be automatically generated [2].

Comparison between traditional and new approaches is shown in Figure 1.

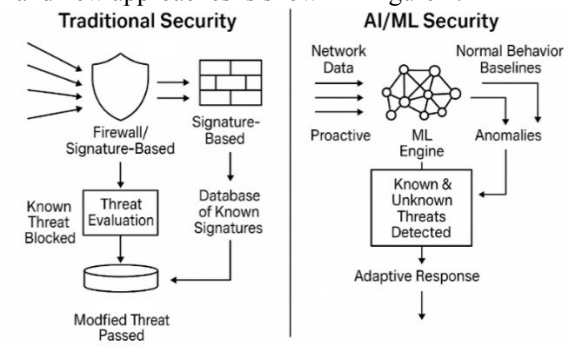

Fig. 1 - Comparison of traditional rule-based security and adaptive AI/ML-driven security paradigms

Advanced Phishing and Social Engineering Mitigation: AI offers multi-layered defense against phishing/BEC by analyzing communication content (Natural Language Processing - NLP for tone, urgency, grammar), sender reputation/behavior, URL/website characteristics (structure, reputation, visual similarity to known sites), and link destinations at click-time. ML models integrate these signals for accurate risk scoring, often leveraging transformer-based language models in newer solutions [7].

An example of AI utilization in Asus router is shown in Figure 2.

Fig. 2 - Router Security Assessment and Threat Prevention using AI

Sophisticated Malware Analysis: ML aids in analyzing the massive volume of malware. Static analysis extracts features from code without execution (byte sequences, API calls, strings) using models like CNNs on binary visualizations [8]. Dynamic analysis monitors behavior in sandboxes (network connections, file/registry changes, system calls). ML learns to differentiate malicious from benign patterns, enabling rapid classification and IoC extraction.

Behavioral Analytics and Insider Threat Detection: User and Entity Behavior Analytics (UEBA) systems leverage AI/ML to counter insider threats or compromised accounts. By analyzing diverse data sources (logs, network flows, endpoint activity, authentication), they establish dynamic baselines of normal behavior for users and entities. Anomaly

detection flags significant deviations (e.g., unusual logins, access patterns, data movement), enabling early investigation [7].

V. RESEARCH RESULTS

Commercial SIEM, NDR, and XDR platforms increasingly incorporate these AI/ML capabilities (e.g., Splunk UEBA, Darktrace, Vectra AI, Palo Alto Networks Cortex XDR, CrowdStrike Falcon).

AI/ML model efficacy hinges on quality training data. Benchmark datasets facilitate development and comparison:

- KDD Cup 99 / NSL-KDD: foundational but dated, containing simulated labeled connection records.
- CICIDS2017 / CSE-CIC-IDS2018: more contemporary and realistic, featuring labeled network flow data from diverse modern attack scenarios and profiled benign traffic [9].
- UNSW-NB15: blends real normal traffic with synthetically generated attacks, providing rich features.
- CTU-13: focuses specifically on labeled captures of real-world botnet traffic.

While valuable, these benchmarks have limitations (bias, environment specifics). Researchers rely on ML libraries and tools:

- General ML: Scikit-learn (classical ML), TensorFlow/Keras & PyTorch (deep learning).
- Security Specific: Integrated ML in SIEM/SOAR, network analysis tools (Zeek), packet manipulation libraries (scapy), data science platforms (Jupyter).

Effective use requires both cybersecurity domain knowledge and data science skills.

Widespread AI/ML adoption in security faces significant hurdles. Data Challenges: high-quality, labeled attack data is scarce and costly. Class imbalance (rare attacks vs. abundant normal traffic) biases models. Data representativeness across different environments is problematic. Privacy regulations (GDPR, etc.) add complexity.

False Positive/Negative Trade-off: balancing detection accuracy (minimizing missed threats - false negatives) against operational burden (minimizing false alarms - false positives) is crucial. High false positives cause alert fatigue; false negatives can be catastrophic. Optimal tuning is context-dependent and difficult.

Adversarial Machine Learning: models are vulnerable to attacks: evasion (inputs modified to be misclassified as benign), poisoning (training data corrupted to degrade or backdoor the model), model extraction (stealing the model or inferring training data). Developing robust defenses is critical and an active area of research [3, 5].

Explainability and Trust (XAI): many models ("black boxes") lack transparency, hindering trust, verification, and regulatory compliance. Explainable AI (XAI) techniques (e.g., LIME [4], SHAP [5]) aim to provide insights into model decisions, fostering human-AI collaboration, though scaling and validating these explanations remain challenging [6].

Ethical Concerns & Bias: extensive monitoring raises privacy issues. Biases in training data can lead to unfair or inequitable performance across different groups or scenarios. Ensuring fairness, accountability, and transparency is an ethical imperative.

Addressing these requires interdisciplinary effort.

AI/ML in security continues to evolve rapidly:

- Federated Learning: enables collaborative model training across decentralized data sources without sharing raw data, enhancing privacy and allowing broader dataset utilization [3, 4].
- Maturation of Explainable AI (XAI): practical XAI methods are becoming crucial for operator trust, human-AI teaming, debugging, and compliance in security operations [6].
- Hyperautomation in Security Operations (SOAR 2.0): AI drives SOAR platforms towards more intelligent automation, including AI-driven alert triage/correlation, incident analysis (e.g., mapping to MITRE ATT&CK), and adaptive, automated response actions [12].
- AI-Empowered Proactive Threat Hunting: AI assists human hunters by generating high-quality hypotheses from subtle anomalies and automating data exploration, allowing analysts to focus expertise on elusive threats.
- Convergence with Other Technologies: Integration with blockchain (secure threat intel sharing) and quantum computing (opportunities like quantum ML, challenges like post-quantum cryptography) will shape future applications.

VI. DISCUSSION OF THE RESULTS

AI/ML are becoming foundational to modern SOCs, enabling more adaptive, anticipatory, and resilient defenses.

This review demonstrates that AI and ML represent a significant paradigm shift in network security, offering capabilities superior to conventional methods in analyzing vast data volumes, learning complex behaviors, detecting subtle anomalies, and automating responses. The analysis highlighted key applications in IDPS, phishing prevention, malware analysis, and UEBA that are demonstrably impactful. However, realizing the full potential of AI/ML in security requires addressing the substantial challenges identified in this overview: data scarcity, quality, and bias; the critical balance between false positives and negatives; the persistent threat of adversarial attacks; and the paramount need for explainability, ethical governance, and trust. Explainable AI (XAI) emerges as pivotal for building this trust and enabling effective human-AI collaboration in security contexts.

VII. CONCLUSIONS

The core contribution of this paper lies in synthesizing the current state, key applications, inherent challenges, and future trajectory of AI/ML integration in network security. This integration signifies an evolutionary leap towards intelligent, adaptive defense systems capable of dynamically responding to the ever-evolving cyber threat landscape.

Future research directions, stemming from the challenges analyzed in this review, should prioritize:

1. Adversarial Robustness in Real-Time Systems: designing and validating computationally efficient defense mechanisms capable of providing real-time robustness against adaptive and previously unseen adversarial evasion and data poisoning attacks targeting ML models deployed in dynamic network environments (going beyond static threat models often used in current research [5, 10]).
2. Actionable Explainability for Security Operations: improving the scalability and fidelity of XAI techniques (e.g., [6, 12, 13) to generate actionable and context-aware explanations for complex model predictions (e.g., deep learning IDPS alerts or UEBA anomalies) that are directly usable by security analysts during high-pressure incident triage and response workflows, rather than just providing generic feature importance scores.
3. Balancing Privacy and Utility in Collaborative Security: developing and evaluating advanced privacy-preserving ML frameworks (extending beyond initial concepts in [3, 4]) that effectively balance the trade-offs between strong data privacy guarantees (e.g., measurable differential privacy budgets for sensitive user activity logs or inter-organizational data sharing) and the practical utility (e.g., maintaining high detection accuracy for specific threat types) of collaboratively trained security models.

Continued progress across these areas is essential for effectively harnessing AI/ML to build a more secure and resilient digital future.

REFERENCES

[1]. Sandler, A., Budashko, V. Improving tools for diagnosing technical condition of ship electric power installations // Eastern-European Journal of Enterprise Technologies. – 2022. – №. 5 (119). – P. 25–33. DOI: 10.15587/1729-4061.2022.266267.

[2]. Vaccaro A., Alberto Z. R. User and entity behavior analytics (UEBA) based on combining expert knowledge and machine learning // 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018. – IEEE, 2018. – P. 2798–2806. – DOI: 10.1109/BigData.2018.8622043.

[3]. McMahan B., Moore E., Ramage D., Hampson S., y Arcas B. A. Communication-efficient learning of deep networks from decentralized data // Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA, 20–22 April 2017. – PMLR, 2017. – Vol. 54. – P. 1273–1282. – URL: http://proceedings.mlr.press/v54/mcmahan17a.html.

[4]. Nguyen T. T., Huynh T. T., Le D. H. Federated learning with differential privacy for cybersecurity: A case study on IoT networks // Journal of Network and Computer Applications. – 2023. – Vol. 212. – P. 103–115. – DOI: 10.1016/j.jnca.2023.103115.

[5]. Apruzzese G., Andreolini M., Colajanni M., et al. Adversarial machine learning for network intrusion detection: A comparative study // Computers & Security. – 2023. – Vol. 128. – P. 103–120. – DOI: 10.1016/j.cose.2023.103120.

[6]. Thapaliya S., Srivastava G. Explainable AI for cyber security: Interpretable models for malware analysis // IEEE Transactions on Dependable and Secure Computing. – 2024. – Vol. 21, No. 1. – P. 45–59. – DOI: 10.1109/TDSC.2023.3298765.

[7]. Smith J., Patel R. AI-driven SOAR: Enhancing security orchestration with reinforcement learning // ACM Transactions on Privacy and Security. – 2024. – Vol. 27, No. 2. – P. 1–28. – DOI: 10.1145/3637321.

[8]. Arp D., Spreitzenbarth M., Hubner M., Gascon H., Rieck K. DREBIN: Effective and explainable detection of Android malware in your pocket // Proceedings of the Network and Distributed System Security Symposium (NDSS), San Diego, CA, USA, 23–26 February 2014. – Internet Society, 2014. – URL: https://www.ndss-symposium.org/ndss2014/programme/drebin-effective-and-explainable-detection-android-malware-your-pocket/.

[9]. Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization // Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP), Funchal, Madeira, Portugal, 22–24 January 2018. – SciTePress, 2018. – P. 108–116. – DOI: 10.5220/0006639801080116.

[10]. Papernot N., McDaniel P., Jha S., Fredrikson M., Celik Z. B., Swami A. The limitations of deep learning in adversarial settings // Proceedings of the 1st IEEE European Symposium on Security and Privacy (EuroS&P), Saarbrücken, Germany, 21–24 March 2016. – IEEE, 2016. – P. 372–387. – DOI: 10.1109/EuroSP.2016.36.

[11]. Buczak A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection // IEEE Communications Surveys & Tutorials. – 2016. – Vol. 18, No. 2. – P. 1153–1176. – DOI: 10.1109/COMST.2015.2494502.

[12]. Ribeiro M. T., Singh S., Guestrin C. "Why should I trust you?": Explaining the predictions of any classifier // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 13–17 August 2016. – ACM, 2016. – P. 1135–1144. – DOI: 10.1145/2939672.2939778.

[13]. Lundberg S. M., Lee S. I. A unified approach to interpreting model predictions // Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017 / Ed. by I. Guyon et al. – Curran Associates, Inc., 2017. – P. 4765–4774.

Отримана в редакції 01.10.2025. Прийнята до друку 13.10.2025. Received 01 October 2025. Approved 13 October 2025. Available in Internet 30 December 2025