Application of Machine Learning Algorithms for Anomaly Detection in Cybersecurity Threat Mitigation

Kim Son Lim[1], Shih Yin Ooi[1,2*], Yee Jian Chew[1,2], and Md Shohel Sayeed[1,3]
[1]Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, Melaka 75450, Malaysia
[2]Centre for Advanced Analytics (CAA), COE for Artificial Intelligence, Multimedia University, Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia
[3]Centre for Intelligent Cloud Computing (CICC), COE for Advanced Cloud, Multimedia University, Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia
E-mail: 1181402836@student.mmu.edu.my, syooi@mmu.edu.my, chewyeejian@mmu.edu.my, shohel.sayeed@mmu.edu.my
* Corresponding author

Keywords: artificial intelligence, cybersecurity, threat detection, machine learning, anomaly detection, adversarial attacks, ransomware, incident response, data security, predictive modelling

Received: July 5, 2025

The integration of Artificial Intelligence (AI) into cybersecurity has transformed the landscape of threat detection, analysis, and mitigation. As cyber-attacks become increasingly sophisticated and evasive, traditional rule-based defences are no longer sufficient to identify zero-day exploits and advanced persistent threats. AI-driven approaches, leveraging machine learning and deep learning, enable proactive anomaly detection, behavioural modelling, and predictive analytics that enhance both the accuracy and agility of cyber defence mechanisms.
This paper provides a comprehensive examination of AI applications in cybersecurity, spanning anomaly detection, automated incident response, and adaptive defence frameworks. It also emphasizes the emerging role of AI in vulnerability management, where predictive modelling, natural language processing, and automated remediation are used to identify, prioritize, and mitigate vulnerabilities before they can be exploited. A real-world case study of Panasonic's VERZEUSE™ platform is presented to illustrate the industrial implementation of AI-enhanced cybersecurity. The platform exemplifies how AI-based predictive analytics, threat intelligence integration, and continuous monitoring can strengthen risk management and compliance in complex IT and IoT ecosystems.
The findings demonstrate that AI substantially improves detection accuracy, response speed, and proactive defence capabilities. However, challenges related to data quality, model robustness, interpretability, and ethical deployment must be addressed to ensure trustworthy adoption. The study concludes that the future of cybersecurity depends on harmonizing human expertise with adaptive AI systems to achieve resilient, self-learning defence frameworks.

Povzetek: Članek prikazuje, kako umetna inteligenca izboljšuje kibernetsko varnost, hkrati pa opozarja na ključne izzive pri njeni uporabi.

1 Introduction

1.2 Overview of the role of artificial intelligence in cybersecurity

Cybersecurity has become a strategic and operational priority for organizations globally due to the proliferation of interconnected systems, cloud services, and IoT devices. As the digital landscape expands, so too does the attack surface available to malicious actors. Modern threats range from conventional malware to sophisticated, adaptive techniques such as advanced persistent threats (APTs), zero-day exploits, and ransomware campaigns [1], [2]. These attacks can result in severe consequences, including data breaches, financial loss, operational disruption, and reputational damage.

Traditional cybersecurity approaches, primarily rule-based systems and signature-based detection, lack adaptive and predictive capabilities when confronting rapidly evolving or previously unknown attack vectors. Static defense mechanisms are inadequate in the face of adversaries who leverage automation, obfuscation, and AI-enhanced attack techniques [3], [4].

To meet these challenges, cybersecurity strategies increasingly incorporate Artificial Intelligence (AI) and Machine Learning (ML) to bolster defence capabilities. AI enables real-time threat detection, behavioural analysis, and predictive modelling, providing security systems with the ability to identify anomalies and respond autonomously [1], [5]. Unlike conventional tools, AI-

driven systems can analyse vast amounts of historical and real-time data, detect emerging patterns indicative of malicious activity, and continuously improve their performance over time.

This section presents an overview of the role of AI in cybersecurity, focusing on its application in threat intelligence, intrusion detection, behavioural analytics, and adaptive response mechanisms.

1.3 Importance of detecting and mitigating cyber threats

The failure to detect and mitigate cyber threats in a timely manner can lead to severe consequences for both public and private sector organizations. Cyber incidents such as data breaches may result in the unauthorized disclosure of sensitive information, privacy violations, operational disruptions, financial losses, and reputational damage [6]. In critical infrastructure environments, undetected attacks can compromise essential services, disrupt business continuity, and pose risks to public safety [7].

Recent studies indicate that the global average cost of a data breach has reached millions of dollars per incident, further emphasizing the financial implications of inadequate cyber defences [8]. Consequently, rapid and proactive threat detection has become imperative. Organizations must adopt mechanisms capable of identifying and mitigating cyber threats in their early stages to minimize impact and ensure resilience.

Failure to implement timely countermeasures can erode stakeholder trust, undermine competitive advantage, and in severe cases, jeopardize the survival of the affected entity [9]. Advanced cybersecurity strategies, including AI-enhanced detection systems, are critical in maintaining situational awareness and achieving operational continuity in an increasingly hostile digital environment.

1.4 Objectives

This study aims to examine the diverse applications of AI in the field of cybersecurity, with the goal of enhancing threat detection, response, and mitigation capabilities. By exploring a range of AI-driven methodologies, such as machine learning, behavioural analytics, and anomaly detection. This research seeks to illustrate how AI technologies contribute to strengthening organizational cyber defence mechanisms [9], [10].

1.5 The key area of focus in ai-drive cybersecurity

This study highlights several critical domains in which AI contributes to enhancing cybersecurity capabilities:

1. **Anomaly and behaviour detection**: AI, particularly through ML algorithms, is leveraged to monitor and identify abnormal patterns in network traffic, user behaviour, and system activities. These approaches enable early detection of potentially malicious actions that traditional rule-based systems may overlook [11], [12].
2. **Automated defence and response**: AI supports the automation of incident response, intrusion prevention, and vulnerability management. By enabling faster decision-making and dynamic adaptation to threats, AI improves the speed and effectiveness of cybersecurity operations [13].
3. **Challenges and risks**: Despite its benefits, the deployment of AI in cybersecurity faces several challenges, including the need for high-quality labelled data, the risk of adversarial attacks targeting AI models, and ongoing concerns surrounding ethics, privacy, and regulatory compliance [14].
4. **Emerging techniques**: Advanced AI paradigms such as reinforcement learning, federated learning, and integration with blockchain technologies are gaining traction as future-forward strategies to increase cyber resilience and secure distributed environments [15], [16].

By exploring these dimensions, this research provides insights into how AI technologies can be strategically applied to help organizations address evolving cyber threats and more effectively safeguard digital assets.

2 Cybersecurity threat landscape

The cybersecurity threat landscape is in a state of continuous evolution, presenting organizations across various industries with an expanding spectrum of sophisticated cyber-attacks that pose significant operational and financial risks [2,17]. Prominent categories of cyber threats include:

a. **Phishing:** A social engineering attack vector aimed at deceiving individuals into divulging sensitive data or inadvertently executing malicious software, often without the user's knowledge [18].
b. **Malware:** A broad classification of malicious software—encompassing viruses, worms, Trojans, and spyware—designed to infiltrate, disrupt, or damage computer systems without authorization [19].
c. **Ransomware:** A subclass of malware that encrypts critical data assets and extorts payment from victims in exchange for decryption keys, causing substantial operational disruption and financial loss [20].
d. **Distributed Denial-of-Service (DDoS) Attacks:** Coordinated attempts to incapacitate network services by overwhelming target systems with a flood of illegitimate traffic originating from multiple distributed sources [21].
e. **Insider threats:** Malicious or negligent activities perpetrated by authorized users or employees that result in unauthorized data access, exfiltration, or system compromise [22].
f. **Advanced Persistent Threats (APTs):** Highly sophisticated, often state-sponsored cyber campaigns characterized by prolonged and targeted intrusion efforts designed to covertly

extract sensitive information from high-value organizations [23].

The sophistication and frequency of these cyber threats have escalated markedly, severely impacting critical sectors such as finance, healthcare, government, and infrastructure. The increasing digital dependency of organizations amplifies the potential consequences of these threats, jeopardizing financial stability, operational continuity, and corporate reputation [24].

2.1 Emergence of advanced threats and associated challenges

The rising complexity of cyber-attacks is fundamentally reshaping cybersecurity paradigms. Polymorphic malware, capable of dynamically altering its code to evade signature-based detection, exemplifies a formidable challenge to traditional defence mechanisms [25], [26]. Compounding this issue, zero-day vulnerabilities—previously undisclosed security flaws lacking available patches—enable attackers to compromise systems prior to the deployment of defensive countermeasures [27].

Moreover, adversaries are increasingly exploiting automated tools and AI to enhance the precision, scale, and efficacy of cyber-attacks. AI-driven methodologies facilitate the orchestration of sophisticated phishing campaigns, development of adaptive stealth malware, and coordination of complex multi-vector assaults [28], [29]. This technological escalation underscores an urgent imperative for cybersecurity frameworks to adopt adaptive, intelligence-driven defences capable of countering evolving threats.

2.2 Limitations of traditional cybersecurity solutions

Conventional cybersecurity approaches predominantly relying on static rules and signature-based detection face critical limitations in addressing the dynamic and rapidly evolving threat landscape [30]. These reactive methods frequently fail to detect novel or polymorphic attacks that continuously evolve to bypass fixed detection criteria [31]. Consequently, traditional solutions are insufficient for proactive threat identification and mitigation.

AI has emerged as a pivotal enabler, imparting cybersecurity systems with adaptive, predictive, and comprehensive capabilities [32]. AI-driven platforms enhance anomaly detection, early threat identification, and automated response mechanisms, thereby substantially improving organizational security postures [33]. Given the fluidity of cyber threats, the integration of AI is essential to maintain resilient defences and safeguard sensitive information assets [34].

3 Artificial intelligence in threat detection

The increasing sophistication of cyber threats necessitates the deployment of AI to augment threat detection efficacy. AI-powered systems can analyse vast volumes of heterogeneous security data in real-time, discerning novel attack patterns and anomalies beyond the scope of traditional methods [35], [36]. This capability markedly improves detection accuracy and latency, empowering organizations to adopt a proactive cybersecurity stance [37].

3.1 Machine learning for anomaly detection

A cornerstone of AI-enabled threat detection is the application of machine learning algorithms to identify anomalous behaviours indicative of security incidents. Machine learning techniques in cybersecurity are conventionally classified into three categories [38], [39]:

- **Supervised learning:** Models are trained on labelled datasets containing both benign activities and known cyber-attacks, enabling classification of future observations based on learned discriminative features.
- **Unsupervised learning:** Algorithms analyse unlabelled data to detect deviations from normative behavioural baselines, facilitating identification of previously unseen or unknown threats.
- **Semi-supervised learning:** A hybrid approach that leverages both labelled and unlabelled data to enhance detection performance, particularly valuable when labelled data is limited [40].

By leveraging these methodologies, AI-driven cybersecurity systems can detect anomalous activities spanning networks, user behaviours, and host systems, thereby identifying threats potentially missed by signature-based defences [41].

3.2 Behavioural analysis and pattern recognition

Building on anomaly detection, AI systems monitor user and system behaviours—such as login patterns, data access, and network utilization—to establish behavioural baselines. Using clustering and classification techniques, deviations from these baselines are identified as potential indicators of compromise, enabling timely threat detection and response.

3.3 Natural Language Processing (NLP) for phishing detection

AI techniques powered by Natural Language Processing (NLP) have demonstrated efficacy in detecting phishing attempts by analysing linguistic cues within emails and electronic communications. NLP models can identify deceptive language patterns and suspicious contextual signals, facilitating rapid mitigation of phishing campaigns which often serve as entry points for broader cyber-attacks.

4 Real-time threat detection with deep learning

Deep learning—an advanced subset of machine learning utilizing neural networks—has enhanced AI-based threat detection by enabling the identification of polymorphic malware and complex attack vectors through multi-layered feature extraction from network and system telemetry. Organizations employing deep learning approaches benefit from reduced detection latency and improved threat identification accuracy, thus mitigating potential damages [42].

4.1 AI in threat mitigation

Effective cybersecurity extends beyond detection to timely mitigation. AI enhances incident response by automating threat containment and remediation actions.

4.2 Automated incident response and threat remediation

AI outputs that maps to actionable runbooks and remediation steps not just alerts are integrated with existing playbook as part of the SIEM SOAR functions. AI-enabled security orchestration, automation, and response (SOAR) platforms rapidly detect incidents and autonomously execute predefined remediation workflows, such as quarantining infected devices, blocking malicious IP addresses, and isolating suspicious files, thereby reducing dwell time and limiting attack impact.

4.3 AI-Enhanced Intrusion Prevention Systems (IPS)

Integrating AI with Intrusion Prevention Systems (IPS) enables dynamic threat recognition and blocking capabilities beyond static signature rules. Machine learning models continuously analyse network and system data to identify and respond to emerging threats, adapting security postures in real-time.

4.4 Role of AI in vulnerability management

AI assists in proactive vulnerability management by scanning organizational IT assets to identify exploitable weaknesses. Through risk scoring based on exploitability and impact, AI prioritizes remediation efforts, enabling cybersecurity teams to focus resources on the most critical vulnerabilities.

- Automated discovery and classification where AI-driven scanner leverage ML and NLP to analyse system configurations, code repositories, and threat feeds to automatically detect new vulnerabilities even before CVEs are published. Pattern recognition for detecting zero-day or configuration based vulnerabilities that traditional signature-based tools might miss. AI-algorithms cluster similar vulnerabilities, reducing duplicate finding and streamlining analyst review.
- Risk prioritization and predictive scoring where all model combine assets criticality, exploit availability, threat intelligence, and business context to assign dynamic risk score. Predictive models assess which vulnerabilities are most likely to be exploited in the near future, enabling teams to patch what matters the most. The approach goes beyond CVSS scoring by factoring in real-world exploitation trends.
- Intelligence patch and remediate planning where AI can recommend or automate patch deployment sequences by evaluating dependencies, system uptime windows and potential service impacts. ML analyses past remediation patterns to predict patch success probability and minimize operational disruptions. Integration with SOAR platform allows AI to tigger patch workflow to predict patch workflow automatically for low-risk systems.
- Continuous monitoring and adaptive defence continuously ingest logs and telemetry to verify if patched systems remain secure or if exploit attempts persist. Behavioural analytics can flag regression vulnerabilities or partial remediations that conventional checks may overlook. Predictive Analytics for Proactive Cyber defence. Adaptive learning ensures detection models evolve alongside the threat landscape
- Threat intelligence correlation correlates external threat feeds, dark web data, and CVE disclosures to identify vulnerabilities actively discussed or weaponized by attackers. This correlation supports proactive vulnerability mitigation, turning raw data into actionable intelligence.

AI-driven predictive analytics synthesize historical and current threat intelligence to forecast likely attack vectors, empowering organizations to pre-emptively strengthen defences and reduce exposure.

5 Challenges of AI in cybersecurity

Despite its transformative potential, AI adoption in cybersecurity faces notable challenges:

a. **Data quality and availability**
   The effectiveness of AI models hinges on access to extensive, high-quality datasets representative of evolving threats. The dynamic nature of cyber threats complicates data collection, annotation, and timely updating, potentially degrading model performance. Data quality and label scarcity models needs good, representative labelled events, and noisy logs lad to false positives and analyst fatigue.
b. **Adversarial attacks on AI models**
   AI systems themselves are vulnerable to adversarial manipulation, where attackers craft inputs designed to evade detection or cause misclassification, undermining trust in AI-based defences.
c. **Ethical and privacy concerns**

Deployment of AI-driven monitoring and decision-making systems raises significant privacy and ethical considerations, particularly regarding data collection, user consent, and potential biases.

d. Interpretability and trustworthiness
The complexity of AI, particularly deep learning models, often results in opaque decision processes that hinder interpretability and user trust. Explainable AI (XAI) methods are crucial to enhance transparency, accountability, and stakeholder confidence. Integration with existing playbooks, handling telemetry, models, and retraining governed by privacy regulatory regimes and product safety.

6 Emerging trends in AI-driven cybersecurity

As AI continues to evolve, several innovative trends are shaping the future of cybersecurity, offering new methods to defend against increasingly sophisticated threats.

a. Reinforcement learning for adaptive defence
Reinforcement learning (RL) has emerged as a promising approach for developing autonomous cybersecurity systems capable of adapting to dynamic and evolving threats. By interacting with simulated or real network environments, RL agents learn optimal defence policies through trial-and-error, continuously refining their strategies based on feedback or rewards. In cybersecurity contexts, RL can be applied to intrusion detection, malware containment, and network traffic control, where decisions, such as isolating suspicious nodes or adjusting firewall rules must be made dynamically under uncertainty. For example, RL-based intrusion detection systems can automatically adjust detection thresholds or response actions in real time, minimizing false positives while improving reaction speed to new attack patterns. This adaptability enables proactive and context-aware defence, aligning cybersecurity operations with evolving attack behaviours.

b. Federated learning for distributed cybersecurity
Federated learning (FL) addresses a critical challenge in cybersecurity, i.e., how to collaboratively improve threat detection models without compromising sensitive or proprietary data. In a federated setup, multiple organizations or devices train local AI models on their own data and share only model updates, not the raw data itself. This decentralized learning paradigm enhances privacy and compliance with data protection regulations while enabling cross-organizational intelligence sharing. In practice, FL supports distributed intrusion detection systems and IoT security frameworks where data from endpoints or edge devices can collectively improve anomaly detection models. By aggregating global intelligence without centralizing data, federated learning strengthens the overall robustness of cyber defences against emerging and distributed attack vectors.

c. Integration of AI with blockchain
Combining AI with blockchain technology enhances the security, transparency, and integrity of AI models and data, mitigating risks of tampering and ensuring trustworthiness.

d. AI-driven deception technologies
AI-powered honeypots and deception environments emulate realistic systems to lure and analyse attacker behaviours, providing actionable intelligence for improving defensive strategies.

e. Ethical and regulatory implications regulatory requirements
Compliance with data protection regulations such as the EU's GDPR and the US CCPA imposes strict constraints on AI-based cybersecurity solutions regarding personal data handling and user privacy.

f. Ethical considerations
Organizations must balance AI deployment with respect for privacy, consent, and civil liberties, ensuring transparent communication and adherence to ethical standards.

g. Development of guidelines and standards
Standards bodies like IEEE and NIST are actively developing ethical frameworks and best practices to guide responsible AI use in cybersecurity, emphasizing transparency, fairness, and accountability.

6.1 Case study: Panasonic's AI-driven cybersecurity ecosystem

Example of AI with Security adopted by the Panasonic. AI transforms vulnerability management from a reactive process to a proactive, predictive, and continuous function. Panasonic's use of AI platforms like VERZEUSE™ exemplifies how enterprises can apply advanced analytics to strengthen cyber resilience, reduce exposure, and maintain compliance in complex IT/OT ecosystems.

- It uses AI to simulate attack paths and predict the most probable exploitation vectors.
- It automates reporting and risk assessment, enabling faster patch development and deployment cycles.
- It integrates with Vehicle SOC operations, ensuring continuous monitoring and adoptive response.

VERZEUSE – Vehicle Risk and Vulnerability Analysis System a proprietary AI-driven platform that automatically identifies, evaluates, and prioritises potential vulnerabilities in connected vehicles. The key components of VERZEUSE

- TARA (Threat Analysis & Risk Assessment) design where the tool leverages an internal threat

intelligence database (Panasonic's collected example of threats, mitigation, previous incidents) to map threats, vulnerability and control appropriate of the ECU type. Generates an SO/SAE-21434-compliant output: threat scenarios, risk assessments, required countermeasures virtualization extension for implementation in modern cockpit system with virtualization or hypervisor containers.

- Runtime integrity checker to ensure monitoring functions themselves aren't tampered with during operation. Includes a "heartbeat" mechanism with digital signatures so that logs and status reports are validated and external verification (Vehicle SOC or equivalent) can detect if tampering has occurred. If the integrity module detects tampering or abnormal shutdown of security components, it triggers alert/notification outside the vehicle. This has been "adopted as an in-vehicle product" in some cases and highly evaluated by OEMs.
- Threat evaluation and security Test Assistance toolkit for testing, SIRT (Security Incident Response Team) to handle vulnerabilities post shipment, using attack path analysis. After vehicles are shipped, known vulnerabilities (from internal or external sources) are fed into the system. It uses data like: SBOM (Software Bill of Materials) for each ECU, connection topology (how ECUs are networked or communicate) to understand how vulnerabilities might be exploited or escalate across ECUs. Panasonic Newsroom Global+1. It also uses Panasonic's Threat Intelligence data to understand current exploit patterns. The algorithm estimates possible attack routes & impact across the whole vehicle (not just isolated ECUs), calculates risk scores, and helps prioritize which vulnerabilities to remediate first. Ensures alignment with ISO/SAE 21434 for vulnerability management and risk assessment. Shown to reduce time for analysis significantly.

Through the platform, Panasonic has transitioned from periodic vulnerability scanning to predictive, model-based vulnerability management, enhancing both speed of remediation and accuracy or risk prioritization, particularly viral for IoT and automotive systems where downtime directly impacts product safety.

From another example, Panasonic has taken several concrete steps involving SOC, SIEM, and SOAR for automation orchestration especially in its VERZEUSE™ programs and vehicle factory operations. The Key experiment in SOC services includes:

a. Vehicle SOC + McAfee partnership
Panasonic and McAfee built a Vehicle Security Operation Center to handle monitoring of connected vehicles. The data architecture involves Automotive Intrusion Detection Systems onboard vehicles sending telemetry / log / alert data to a backend SOC + SIEM system for analysis and visualization. This indicates Panasonic is operationalizing SIEM for post-shipment / runtime vehicle security monitoring.

b. VERZEUSE™ for SIEM (Operation / Post-shipment)
As part of VERZEUSE™, they have a module for SIEM. It "automatically detects and analyses cyber-attacks on shipped vehicles" to "accelerate and streamline security monitoring. It supports threat intelligence from Panasonic Group's broader CI sources.

c. VERZEUSE™ for SIRT (Vulnerability Management after shipment)
For vulnerabilities discovered post-shipment, Panasonic uses VERZEUSE for SIRT to analyse risk using SBOM (software bill-of-materials), ECU connectivity, and their own threat intelligence, to prioritize remediations.

d. Factory / Manufacturing SOC Operations
Panasonic has had SOCs operating for factories (since ~2016) to protect factory networks and systems. They analyse communications in factory environments for anomalies & malware intrusions.

e. Proof of Concepts (PoCs) on log analytics & anomaly detection
Panasonic's Annual Reports mention PoCs to "analyse huge data logs and to detect abnormalities. These to be preliminary experiments to build capability in monitoring, detection and response.

7 Conclusion

AI has become indispensable in fortifying cybersecurity defences against increasingly complex cyber threats. Leveraging machine learning, deep learning, and related AI techniques, organizations achieve improved detection accuracy, faster response times, and adaptive protection.

Looking forward, AI's role in autonomous cybersecurity systems and cross-domain intelligence integration is poised to expand, contingent on addressing challenges such as adversarial robustness, model explainability, and ethical deployment. A multidisciplinary approach involving technologists, policymakers, and industry stakeholders is imperative to harness AI's potential while safeguarding privacy and civil rights.

Disclosure statement

We declare there are no conflicts of interest regarding the research, authorship, or publication of this article.

References

[1] S. A. Shaikh, N. A. Shaikh, and M. N. Shaikh, "A review of artificial intelligence techniques in cybersecurity," *2020 3rd International Conference on Computing, Mathematics and Engineering

Technologies (iCoMET)*, Sukkur, Pakistan, 2020, pp. 1–5, doi: 10.1109/iCoMET48670.2020.9073885.

[2] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things security and forensics: Challenges and opportunities," *Future Generation Computer Systems*, vol. 78, pp. 544–546, Jan. 2018, doi: 10.1016/j.future.2017.07.060.

[3] M. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating deep learning approaches to intrusion detection," *Procedia Computer Science*, vol. 132, pp. 485–490, 2018, doi: 10.1016/j.procs.2018.05.198.

[4] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," *2010 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 2010, pp. 305–316, doi: 10.1109/SP.2010.25.

[5] IBM Security, "Cost of a Data Breach Report 2023," IBM Corporation, 2023. [Online]. Available: https://www.ibm.com/reports/data-breach

[6] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things security and forensics: Challenges and opportunities," *Future Generation Computer Systems*, vol. 78, pp. 544–546, Jan. 2018, doi: 10.1016/j.future.2017.07.060.

[7] H. A. Mohammed and I. M. Husien, "A Deep Transfer Learning Framework for Robust IoT Attack Detection," *IJCAI*, vol. 48, no. 12, Sep. 2024, doi: 10.31449/inf.v48i12.5955.

[8] M. Alazab, S. Venkatraman, P. Watters, and M. Alazab, "Zero-day malware detection based on supervised learning algorithms of API call signatures," in *Proceedings of the Ninth Australasian Data Mining Conference - Volume 121*, in AusDM '11. AUS: Australian Computer Society, Inc., 2011, pp. 171–182.

[9] L. Idouglid, S. Tkatek, and K. Elfayq, "Multilayer Perceptron-Based Defense Mechanisms for Securing Industrial IoT in Industry 4.0 Environments," *IJCAI*, vol. 49, no. 33, Aug. 2025, doi: 10.31449/inf.v49i33.6944.

[10] R. M. Alguliyev, Y. E. Imamverdiyev, and L. A. Sukhostat, "Cyber-attacks detection in critical information infrastructures using artificial intelligence approaches: A review," *IEEE Access*, vol. 9, pp. 14014–14043, 2021, doi: 10.1109/ACCESS.2021.3052133.

[11] M. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating deep learning approaches to intrusion detection," *Procedia Computer Science*, vol. 132, pp. 485–490, 2018, doi: 10.1016/j.procs.2018.05.198.

[12] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," *2010 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 2010, pp. 305–316, doi: 10.1109/SP.2010.25.

[13] N. Idika and B. Bhargava, "Extending attack graph-based security metrics and aggregating their application," *IEEE Trans. Dependable and Secure Computing*, vol. 9, no. 1, pp. 75–85, Jan.-Feb. 2012, doi: 10.1109/TDSC.2011.44.

[14] R. M. Alguliyev, Y. E. Imamverdiyev, and L. A. Sukhostat, "Cyber-attacks detection in critical information infrastructures using artificial intelligence approaches: A review," *IEEE Access*, vol. 9, pp. 14014–14043, 2021, doi: 10.1109/ACCESS.2021.3052133.

[15] H. Huang, Z. Wang, and X. Huang, "A blockchain-based privacy-preserving incentive mechanism in crowdsensing applications," *IEEE Access*, vol. 7, pp. 62996–63006, 2019, doi: 10.1109/ACCESS.2019.2916556.

[16] Y. Liu, R. Wang, and Y. Chen, "Federated learning for cyber security: Concepts, applications, and future directions," *IEEE Trans. Ind. Informatics*, vol. 18, no. 5, pp. 2926–2939, May 2022, doi: 10.1109/TII.2021.3085284.

[17] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of Things security and forensics: Challenges and opportunities," *Future Generation Computer Systems*, vol. 78, pp. 544–546, Jan. 2018, doi: 10.1016/j.future.2017.07.060

[18] J. Kumar, A. Santhanavijayan, B. Janet, B. Rajendran and B. S. Bindhumadhava, "Phishing Website Classification and Detection Using Machine Learning," *2020 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2020, pp. 1-6, doi: 10.1109/ICCCI48352.2020.9104161.

[19] S. Axelsson, "The base-rate fallacy and its implications for the difficulty of intrusion detection," in *Proc. ACM Conf. Computer and Communications Security*, 1999, pp. 1–7, doi: 10.1145/319709.319710.

[20] R. Richardson and M. M. North, "Ransomware: Evolution, mitigation and prevention," *Int. Manage. Rev.*, vol. 13, no. 1, p. 10, 2017

[21] H. Djuitcheu, M. Debes, M. Aumüller and J. Seitz, "Recent review of Distributed Denial of Service Attacks in the Internet of Things," *2022 5th Conference on Cloud and Internet of Things (CIoT)*, Marrakech, Morocco, 2022, pp. 32-39, doi: 10.1109/CIoT53061.2022.9766655.

[22] S. Rakhi; H. K. Sampada; Arun Balodi; P. C. Shobha; Roshan Kumar, "Insider Threat Detection and Prevention," in *Emerging Threats and Countermeasures in Cybersecurity* , Wiley, 2025, pp.241-262, doi: 10.1002/9781394230600.ch12.

[23] A. Alshamrani, S. Myneni, A. Chowdhary and D. Huang, "A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1851-1877, Secondquarter 2019, doi: 10.1109/COMST.2019.2891891.

[24] A. Aktayeva *et al.*, "Cybersecurity Risk Assessments within Critical Infrastructure Social Networks," *Data*, vol. 8, no. 10, p. 156, Oct. 2023, doi: https://doi.org/10.3390/data8100156.

[25] V. L. S., Chandradeva, “Polymorphic Malware Detection Using Machine Learning,” *Iit.ac.lk*, 2021, doi: https://doi.org/2019739.

[26] T. Holz, M. Steiner, F. Dahl, E. Biersack, and F. C. Freiling, "Measurements and mitigation of peer-to-peer-based botnets: A case study on Storm worm," in *Proc. 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, 2008.

[27] F. El Husseini, H. Noura, O. Salman and A. Chehab, "Advanced Machine Learning Approaches for Zero-Day Attack Detection: A Review," *2024 8th Cyber Security in Networking Conference (CSNet)*, Paris, France, 2024, pp. 297-304, doi: 10.1109/CSNet64211.2024.10851751.

[28] Alexandru-Raul Matecas, P. Kieseberg, and S. Tjoa, “Social Engineering with AI,” *Future Internet*, vol. 17, no. 11, pp. 515–515, Nov. 2025, doi: https://doi.org/10.3390/fi17110515.

[29] A. S. Elmaghraby and M. M. Losavio, "Cybersecurity challenges in Smart Cities: Safety, security and privacy," *Journal of Advanced Research*, vol. 5, no. 4, pp. 491–497, 2014, doi: /10.1016/j.jare.2014.02.006.

[30] M. Roesch, "Snort - lightweight intrusion detection for networks," in *Proc. 13th USENIX Conf. System Administration*, 1999, pp. 229–238.

[31] G. Creech and J. Hu, "A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns," *IEEE Trans. Comput.*, vol. 63, no. 4, pp. 807–819, Apr. 2014, doi: 10.1109/TC.2013.13.

[32] N. H. Motlagh, M. Bagaa and T. Taleb, "UAV-Based IoT Platform: A Crowd Surveillance Use Case," in *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128-134, February 2017, doi: 10.1109/MCOM.2017.1600587CM.

[33] S. Srinivas *et al.*, “AI-Augmented SOC: A Survey of LLMs and Agents for Security Automation,” *Journal of Cybersecurity and Privacy*, vol. 5, no. 4, p. 95, Nov. 2025, doi: https://doi.org/10.3390/jcp5040095.

[34] S. Zeijlemaker, Y. K. Lemiesa, S. L. Schröer, A. Abhishta, and M. Siegel, “How Does AI Transform Cyber Risk Management?,” *Systems*, vol. 13, no. 10, p. 835, Sep. 2025, doi: https://doi.org/10.3390/systems13100835.

[35] T. T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," in *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, pp. 56-76, Fourth Quarter 2008, doi: 10.1109/SURV.2008.080406.

[36] L. Tang and Q. H. Mahmoud, “A Survey of Machine Learning-Based Solutions for Phishing Website Detection,” *Machine Learning and Knowledge Extraction*, vol. 3, no. 3, pp. 672–694, Aug. 2021, doi: https://doi.org/10.3390/make3030034.

[37] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[38] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006. Available: https://link.springer.com/book/9780387310732

[39] A. B. Nassif, M. A. Talib, Q. Nasir and F. M. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," in *IEEE Access*, vol. 9, pp. 78658-78700, 2021, doi: 10.1109/ACCESS.2021.3083060.

[40] O. Chapelle, B. Scholkopf and A. Zien, Eds., "Semi-Supervised Learning (Chapelle, O. et al., Eds.; 2006) [Book reviews]," in *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 542-542, March 2009, doi: 10.1109/TNN.2009.2015974.

[41] D. E. Denning, "An Intrusion-Detection Model," in *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222-232, Feb. 1987, doi: 10.1109/TSE.1987.232894.

[42] Lekhetho Joseph Mpoporo, Pius Adewale Owolawi, and C. Tu, “Deep Reinforcement Learning Algorithms for Intrusion Detection: A Bibliometric Analysis and Systematic Review,” *Applied Sciences*, vol. 16, no. 2, pp. 1048–1048, Jan. 2026, doi: https://doi.org/10.3390/app16021048.