AI in Cybersecurity Today: Detection Systems, Security Frameworks, and Ethical Considerations

Dr. Goldi Soni[1], Mahi Panjwani[2], Charitha NL[3]
[1.] *Assistant Professor Amity University Raipur Chhattisgarh, India*
[2,3.] *Student Amity University Raipur Chhattisgarh, India*

***Abstract*—This review paper carefully examines how artificial intelligence is being applied to strengthen cybersecurity in response to increasingly complex threats reported in the past five years. The paper survey emphasizes on 45 peer-reviewed studies and several industry white papers, with insistence on AI-focused intrusion detection, automated malware classification, and predictive risk assessment. As the cyber threats continue to grow in sophistication, AI-focused solutions for threat hunting, automated malware analysis, and incident response are becoming necessary components of modern security frameworks. The process followed in this study is to conduct a broad review of literature, papers, cases, and reports relating to AI-based cybersecurity solutions, covering IDS, automated malware detection, and risk assessment frameworks with an AI focus. The findings highlight the pros and cons of AI such as better accuracy in detection and automation as well as potential problems involving algorithm bias, data privacy and overreliance on automated processes. The paper states that AI has a huge potential to enhance cybersecurity, future work should research on transparent and explainable models, as well as on fighting always against the most recent cyber threats.**

***Index Terms*—.AI, Threats, Threat detection, Cybersecurity, IDS, Malware, Risk Management, Machine learning, AI-Security Frameworks.**

I. INTRODUCTION

Cybersecurity threats are taking a new turn now, rendering the traditional security systems less effective against modernised attacks. AI enhances security frameworks by providing intelligent threat detection and response mechanisms [1]. With the rapid increase in urbanity of cyber threats, the traditional cybersecurity measures are often inadequate. AI has come out as a notably powerful tool in the cybersecurity landscape, enabling automated threat detection, predictive analysis, and efficient incident response [2][3]. This paper offers an all-inclusive overview of AI applications in the field of cybersecurity, covering intrusion detection systems (IDS), AI-powered security frameworks, and advanced malware detection techniques. Further, it investigates AI's role in risk assessment, incident response, and digital forensics, with an anchor on its practical benefits and potential challenges.

II. LITERATURE REVIEW

AI is now a critical part of the cybersecurity landscape - it has helped improve intrusion detection, malware detection, risk management, and incident response. Leading researchers such as Rajendran and Vyas highlights the significance of artificial intelligence (AI) enabled IDS for real-time threat detection whereas Perumal et al. showing how AI can be used for risk management using frameworks including NIST and FAIR [1][4]. For malware detection, Djenna et al. propose hybrid AI models to tone down the false positives, while Yaseen demonstrates the accuracy of CNNs and RNNs [5][6]. The Iturbe et al. s AI4CYBER framework demonstrates the AI-modulated automation in IRT, and Kumar et al. launched alarms over algorithmic bias and AI-enabled cyber-attacks [7,8]. These findings confirm the disruptive power of AI on cybersecurity and yet reinforce the need to tackle ethical challenges.

The studies included reflect a wide variety of use cases where AI is applied for cybersecurity:

- AI in Threat Detection and Prevention: The studies forefront the role of AI in detecting malicious activities using anomaly detection and behavioural analysis [9][10].
- AI for Incident Response: AI-powered systems ease the automated response mechanisms, reducing the downtime and minimizing damage [11][12].
- Adversarial AI Attacks: A number of papers caution about the misuse of AI and machine learning by malicious actors in order to carry out sophisticated and stealthy attacks, and the need for robust defenses in response. [13,14].

- Ethical and Privacy Concerns: These come from the fact that AI models usually rely on massive-scale datasets which pose privacy challenges, requiring strong privacy-protective process [15,16].

III. AI-POWERED INTRUSION DETECTION SYSTEMS (IDS)

Intrusion detection systems observe network data in order to detect malicious behaviour. AI IDS are able to detect anomalies in real-time, relying on a variety of Machine Learning (ML) algorithms, and cutting down on false positives and response times [17].

A. Machine Learning Algorithms in IDS

- Supervised Learning: Models trained on labelled datasets for anomaly detection. The Support Vector Machines and the Decision Trees are the examples of algorithms that are widely used in today's world [17].
- Unsupervised Learning: It can detect new threats without labelled data and it involves clustering techniques such as K-Means and DBSCAN [4].
- Deep Learning: Neural networks in deep learning are used to perform complex pattern recognition. Examples of such neural networks are CNNs and RNNs [18].

B. Examples IDS Powered by AI

- Darktrace: Uses AI algorithms for real-time network monitoring, anomaly detection, and automated response [19].
- Cylance: Focuses on endpoint security using predictive analysis [19].
- IBM Watson for Cybersecurity: Uses NLP and ML to process threat intelligence data [19].

IV. AI IN MALWARE DETECTION AND ANALYSIS

AI enhances malware detection by identifying behavioural patterns and anomalies [20].

A. Techniques in AI Malware Detection

- Behavioural Detection: The AI identifies malware by its Features not just mainly on the Signature [20]. Heuristic-Based Detection: Finds zero-day attacks by observing nonconformities in the operation [20].
- Heuristic Analysis: Identifies behavioural changes in order to detect zero-day attacks [20].
- Deep Neural Networks (DNN): Improve accuracy in classifying malware types [7].
- Hybrid Models: Combine heuristic analysis and DNNs for robust malware detection [7].

B. Case Studies

- Djenna et al. demonstrated the effectiveness of combining behaviours-based DNN with heuristic approaches, reducing false positives and enhancing detection accuracy [5].
- AI-driven botnet detection models effectively identify botnet activity through ML-based anomaly detection [8].

V. RISK MANAGEMENT AND AI FRAMEWORKS

AI enhances cybersecurity risk management through predictive analysis and automated decision-making [4]. AI is widely applied in various cybersecurity frameworks to enhance security operations.

A. AI4CYBER Framework

The AI4CYBER framework provides AI-powered solutions for threat management and incident response. Key components include:

- AI4VUN: AI-enhanced vulnerability identification [7].
- AI4TRIAGE: Root cause analysis and alert management [7].
- AI4SOAR: Automated incident response and adaptive security [7].

The digital forensic framework which is traditional (Figure 1) relies primarily on user inputs and therefore, it becomes circuitous, protracted, and relies on the knowledge of the user. Although the existing products automate file retrieval, file sorting, and group artifacts, the investigator still needs to examine numerous files which is the main reason for delayed investigations. To overcome these limitations, the proposed AI-based digital forensic framework (Figure 2) introduces intelligence into the process by integrating machine learning techniques. Trained on existing datasets, the system is capable of learning from prior cases, reducing redundant work, and prioritizing relevant evidence for investigators. By transforming the traditional steps of digital forensics into "smart" steps, the framework enhances efficiency, minimizes user dependency, and accelerates case analysis while still requiring human oversight for final decision-making.

Fig. 1. Non-AI Digital Forensic Framework [27].

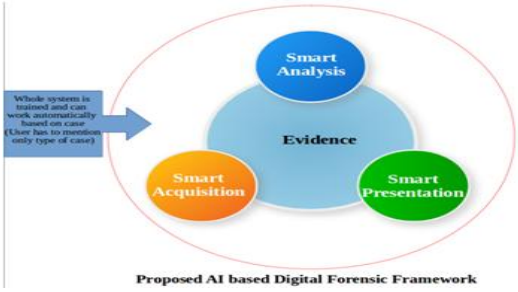

Fig. 2. Proposed AI based Digital Forensic Framework [27].

B. Technique to Reduce Risk

Many frameworks of cybersecurity integrate AI for managing the risk:

- NIST Cybersecurity Framework: Sets the framework for cybersecurity prevention and incident management in entities: for identifying, protecting, detecting, responding and recovering [21].
- FAIR (Factor Analysis of Information Risk) Framework: Applies AI in quantifying cyber risks and offering associated mitigation recommendations [21].
- MITRE ATT&CK Framework: Employs AI to improving the identification of cyber-attacks based on the linkage of stratagems and combat techniques used by attackers [21].

C. Additional AI Security Applications

- Intrusion Detection Systems (IDS): AI algorithms scan network activity to find anomalies which could be potential threats in real time [3, 14].
- Automated Malware Detection: AI systems monitor and evaluate malware activity streamlining false positives and enhancing detection accuracy [22, 23].
- AI in Risk Management: AI augments risk assessment frameworks by offering foresight, clever thoughts on predicting and providing mitigation techniques [24, 25].

VI. ETHICAL CONSIDERATIONS AND CHALLENGES

The gains that AI brings along within the realm of cybersecurity are numerous, however, the technology does raise some ethical and practical dilemmas. A case in point is biased and unfair threat detection due to algorithmic bias, a worrying phenomenon in existing AI systems [15]. Moreover, opaque AI systems tend to act like "black boxes," and it is a serious challenge in today's world to decipher the manner in which these systems make their decisions [26]. An over dependence on AI could make human monitoring almost nil, and that comes with the risk of making oneself an easy target for adversarial attacks. There is also the issue of privacy risk arising from the need of huge troves of data containing personal data, which means the risk of exposing sensitive information to the public is also present. The use of transparent AI, periodic evaluations, and the assurance of human surveillance [26] might help to lower some of these dangers.

While AI provides significant advantages, it also presents challenges:

- Adversarial Attacks: Hackers use adversarial techniques to exploit AI driven systems, injecting malicious inputs to "fool" authoring algorithms [13,22].
- Data Privacy Issues: The vast volumes of data that an AI system works with raises concerns about the leakage of protected information [16].
- Algorithmic Bias: Unreasonably negative slants in the training data might impact the decisions made and the proposed solutions in threat detection, and monitoring these systems becomes vital [15].
- Regulatory Gaps: The absence of adequate regulation for the use of AI within the field of cybersecurity is lamentable [24].

VII. RECOMMENDATIONS FOR FUTURE RESEARCH

Future studies should aim at:

- Adaptive AI Systems: Designing an AI model that can quickly adapt to adversarial variations in cyber threats [13,22].
- Explainable AI (XAI): Establishing transparency and interpretability to AI-based decision-making.

- Ethical AI Frameworks: Create mechanisms to ensure progress in ethical AI development and use.
- Cross-Industry Partnership: Foster cooperation across industries and collaborative research organizations to create resilient AI-based cybersecurity systems.

VIII. COMPARATATIVE ANALYSIS OF KEY STUDIES

Cybersecurity Artificial Intelligence was researched via a set of works, differing in views, methods, and usages. It offers a comparative study of five of the most significant works in such a direction, identifying their aims, limitations, future directions, and contributions. Those works span from ChatGPT's effects on cybersecurity to computer forensic and AI-related defense models in networks.

It reveals similar patterns, including the success of AI in detecting intrusions, malware detection, and real-time threat response, alongside major shortcomings such as the trustworthiness of AI, vulnerability of security, and a necessity for human intervention in forensic analysis. Research pathways for the future cover increasing the transparency of AI, improving in-real-time response capabilities, and utilizing highly sophisticated ML models for enhanced cyber resilience. This table serves as a structured summary to facilitate a deeper understanding of how AI continues to evolve as both a defense mechanism and a potential risk in cybersecurity.

TABLE I. Comparative Analysis of Key Studies

Study	Authors	Year	Objective	Limitations	Future Scope	Outcome
Introduction to ChatGPT: A New Revolution of AI with Machine Learning Algorithms and Cybersecurity	Hadi, Abdulredha, & Hasan	2023	Explore ChatGPT's evolution, structure, and impact on cybersecurity	Security vulnerabilities like adversarial manipulation and data privacy risks not deeply explored	Investigate enhanced AI security measures, reduce biases, improve cybersecurity applications	Demonstrated ChatGPT's role in phishing detection and automated threat analysis
Cyber Defence Based on AI and Neural Network Model in Cybersecurity	Sugumaran et al.	2023	Enhance intrusion detection, malware identification, and vulnerability analysis using AI and neural networks	Adversaries could exploit AI systems for cyberattacks; data privacy remains a challenge	Develop advanced AI algorithms for proactive threat mitigation and improved detection accuracy	Enhanced threat detection using AI-driven systems
AI for Next Generation Cybersecurity: The AI4CYBER Framework	Iturbe, Rios, Rego, & Toledo	2023	Introduce AI4CYBER framework with AI-powered tools for improved cybersecurity	AI trustworthiness, data privacy, and mitigation of AI biases remain challenging	Enhance AI transparency, adversarial resilience, and broaden AI4CYBER's applications	Strengthened cyber resilience and response efficiency using AI tools
Applications of AI to Network Security	Veiga	2018	Explore applications of AI, particularly ML, in detecting and mitigating cyber threats	Limited application beyond supervised ML; heavy reliance on human oversight	Expand use of unsupervised ML for adaptive threat detection and real-time response	Improved threat detection using AI-driven tools like Darktrace
AI-Based Digital Forensics Framework	Rughani	2022	Propose AI-driven framework to reduce human intervention in digital forensic investigations	AI reliability, data bias, and need for human oversight in complex forensic cases	Develop real-time AI evidence analysis algorithms; integrate blockchain for secure evidence management	Reduced investigator workload and improved investigation efficiency

IX. CONCLUSION

Cybersecurity has come on in leaps and bounds with AI used for intelligent threat identification and reducing risk. But there are still questions about ethics and transparency. Future research should address the interpretability of AI models, and in particular algorithmic fairness, as well as a collaboratice environment between AI systems and the cybersecurity professional's needs [26].

AI has revolutionized cybersecurity through faster and more efficient threat detection and response s olutions [1,2,11,12]. There are further hurdles to face, but ongoing investigation and cooperation will result in robust AI-driven security frameworks. Ethical concerns, model transparency and regulatory standards can add significantly to the potential of AI in cyber security.

REFERENCES

[1] R. M. Rajendran and B. Vyas, "Cyber Security Threat and Its Prevention by Artificial Intelligence Technology." 2023

[2] Dendy Jonas, Natasya Aprila Yusuf, and Achani Rahmania Az Zahra, "The Improvement of Security Frameworks Using Artificial Intelligence in Cybersecurity." 2023

[3] Irshaad Jada and Thembekile O. Mayayise, “The Effect of Artificial Intelligence on Organisational Cyber Security: A Result of a Systematic Literature Review. A. P. Perumal, et al., “Risk Assessment of Artificial Intelligence Systems in Cybersecurity,” 2024

[4] 2024 [5] A. Djenna et al., “Artificial Intelligence-Based Malware Detection, Analysis, and Mitigation.” 2023

[5] A. Djenna et al., “Artificial Intelligence-Based Malware Detection, Analysis, and Mitigation.” 2023

[6] A. Yaseen, “AI-Driven Threat Detection and Response: A Paradigm Shift in Cybersecurity.” Dec 2023

[7] E. Iturbe et al., “AI4CYBER- Artificial Intelligence for Next-Generation Cybersecurity.” 2023

[8] S. Kumar et al., “Artificial Intelligence-Revolutionizing Cyber Security in the Digital Era.” 2023

[9] Syed Minhajul Hassan and Dr. Javed Wasim, “Study of Artificial Intelligence in Cyber Security and The Emerging Threat of AI-Driven Cyber Attacks and Challenges.” 2023

[10] Sarah Al-Mansoori and Mohamed Ben Salem, ‘The Role of Artificial Intelligence and Machine Learning in Shaping the Future of Cybersecurity- Trends, Applications, and Ethical Considerations.”

[11] Musadaq A. Hadi, Mohammed Najm Abdulredha, and E. Hasan, “Introduction to ChatGPT: A New Revolution of Artificial Intelligence with Machine Learning Algorithms and Cybersecurity.” December 2023

[12] Krishna Shree Achuthan et al., “Advancing Cybersecurity and Privacy with Artificial Intelligence- Current Trends and Future Research Directions.” December 2024.

[13] Ankush Mehra and Sumit Badotra, “Artificial Intelligence Enabled Cyber Security.” 2021

[14] Dr. D. Sugumaran et al., “Cyber Defence Based on Artificial Intelligence and Neural Network Model in Cybersecurity.” 2022

[15] A. P. Veiga, “Applications of Artificial Intelligence to Network Security.” 2018

[16] Yinglan Zhao et al., “Privacy Crisis in the Age of Artificial Intelligence and its Countermeasures.” 2022

[17] A. Hebbar and S. A. Kumar, “Artificial Intelligence in Cyber Security.” 2021

[18] M. S. Alzboon et al., “The Two Sides of AI in Cybersecurity: Opportunities and Challenges.” 2024

[19] S. Zeadally et al., “Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity.” 2020

[20] S. K. Hassan and A. Ibrahim, “The Role of Artificial Intelligence in Cyber Security and Incident Response.” 2023

[21] S. Dambe et al., “The Role of Artificial Intelligence in Enhancing Cybersecurity and Internal Audit.” 2023

[22] Unyime Ufok Ibekwe et al., “A Critical Review of The Intersection of Artificial Intelligence and Cybersecurity.” Nov 2023

[23] Radhika Nautiyal et al., “Artificial Intelligence Indulgence in Protection of Cybercrime.”- June 2023

[24] Eider Iturbe et al., “Towards Trustworthy Artificial Intelligence- Security Risk Assessment Methodology for Artificial Intelligence Systems.” Dec 2023

[25] Gaurav Rawat et al., “Use of Artificial Intelligence in Modern Warfare and National Security.” 2021
[26] R. Calderon, “The Benefits of Artificial Intelligence in Cybersecurity.” 2019
[27] P. H. Rughani, “Artificial Intelligence-Based Digital Forensics Framework.” 2018
[28] M. E. Bonfanti, “Artificial Intelligence and Cybersecurity: A Promising but Uncertain Future.” 2020
[29] Walaa Mohamed et al., “Cybersecurity in the Era of Artificial Intelligence: Risks and Solutions.” 2024
[30] Fadi Al-Ayed, “Contemporary Cybersecurity Challenges in Metaverse Using Artificial Intelligence.” 2022