Baltic Journal of MultiDisciplinary Research - BJMR

Pages: 70-75
Volume-II, Issue-IV (2025)
https://balticpapers.com/index.php/bjmr

The Role of Generative AI in Offensive and Defensive Cyber Operations

[1] Noman Mazher, [2] Zillay Huma
[1] University of Gujrat, Pakistan
[2] University of Gujrat, Pakistan
Corresponding Author: nauman.mazhar@uog.edu.pk

Abstract:

Generative Artificial Intelligence (AI) has emerged as a transformative force in cybersecurity, reshaping both offensive and defensive operations. By leveraging models capable of creating realistic data, simulations, and adaptive attack or defense strategies, Generative AI introduces new paradigms in digital warfare and protection. Offensive applications include automated phishing, deepfake generation, and polymorphic malware creation, while defensive uses involve synthetic data generation for threat modeling, deception technologies, and automated vulnerability patching. This dual-use nature of Generative AI presents a paradox—enhancing system resilience while simultaneously introducing novel attack vectors. The integration of large language models (LLMs), generative adversarial networks (GANs), and diffusion models within cybersecurity infrastructures enables predictive, adaptive, and autonomous responses to complex threats. However, ethical, regulatory, and operational challenges persist, especially concerning attribution, accountability, and misuse prevention. This paper explores the evolving landscape of Generative AI in cyber operations, assessing its implications for future cyber defense frameworks and digital warfare ethics.

Keywords: Generative AI, Cybersecurity, Offensive Operations, Defensive Operations, Artificial Intelligence, Deepfakes, GANs, Cyber Defense, LLMs

I. Introduction

The evolution of Artificial Intelligence (AI) has transformed how societies engage with technology, but the advent of Generative AI has marked a particularly profound shift in the domain of cybersecurity. Generative AI models—capable of producing synthetic yet highly

realistic data, code, and content—have extended beyond creative and analytical tasks into the sphere of cyber operations. The dual-use nature of this technology presents both unprecedented opportunities and serious risks, influencing offensive and defensive strategies alike. In offensive cyber operations, adversaries now harness Generative AI to craft convincing phishing campaigns, generate deepfake audio or video for social engineering, and develop polymorphic malware capable of evading traditional security measures [1]. These AI-driven attacks blur the boundaries between human and machine tactics, increasing the sophistication and scale of modern cyber threats [2].

Conversely, defensive cybersecurity has also entered a new era powered by Generative AI. Security systems enhanced by large language models (LLMs) and generative adversarial networks (GANs) are now capable of simulating complex attack vectors, generating synthetic training data for anomaly detection models, and creating deceptive environments designed to mislead adversarial agents. Automated incident response mechanisms driven by AI can identify, predict, and mitigate attacks in real time. For example, generative models can anticipate potential vulnerabilities before exploitation, enabling proactive patching and strengthening system resilience. The convergence of these technologies represents a fundamental transformation in how security systems evolve. However, the growing sophistication of AI-powered offensive tools raises significant ethical and regulatory questions [3]. Attribution becomes increasingly difficult when AI-generated attacks mimic legitimate traffic or human behavior, complicating legal accountability. Furthermore, as autonomous systems take a greater role in defense and retaliation, issues of control, bias, and unintended escalation come to the forefront.

This paper explores the dual role of Generative AI in modern cyber warfare, emphasizing both its potential to revolutionize cyber defense and the inherent risks it introduces. The discussion highlights key frameworks, models, and ethical considerations that will define the future of AI-driven cybersecurity, ultimately arguing for the need for governance, transparency, and cross-sector collaboration to ensure that generative technologies serve as tools for protection rather than proliferation of cyber threats [4].

Pages: 70-75
Volume-II, Issue-IV (2025)
https://balticpapers.com/index.php/bjmr

II. Generative AI in Offensive Cyber Operations

The offensive use of Generative AI in cyber operations represents one of the most concerning advancements in the digital threat landscape. Attackers can now automate complex social engineering tactics using AI-generated text, audio, and video that are nearly indistinguishable from authentic communication. Large language models (LLMs) can craft personalized phishing emails by analyzing target profiles and contextual data, dramatically increasing success rates. Similarly, deepfake technology allows adversaries to impersonate executives or government officials, manipulating financial transactions or spreading misinformation with alarming precision [5].

Another dimension of offensive Generative AI lies in malware design. Traditional malware detection relies on pattern recognition and signature matching; however, AI-driven malware generation can create polymorphic code that continuously alters its structure to avoid detection. Generative adversarial networks (GANs) are particularly effective in testing and refining such malicious code, allowing adversaries to produce software that evolves autonomously. This adaptability undermines conventional cybersecurity defenses, making reactive strategies insufficient [6].

Beyond direct attacks, Generative AI also supports large-scale disinformation campaigns. AI-generated social media accounts, automated propaganda, and fabricated news content can destabilize organizations or entire societies. These operations are not limited to hacking systems but also target public perception, eroding trust in institutions and digital infrastructure. The accessibility of open-source AI models further exacerbates this issue, allowing even low-resource threat actors to deploy sophisticated tools [7]. While Generative AI enhances the tactical and psychological dimensions of cyber warfare, it also introduces challenges for attribution and accountability. Distinguishing between human-led and AI-generated attacks becomes nearly impossible, complicating international law and cyber deterrence policies. As AI models continue to advance, offensive cyber capabilities risk becoming increasingly

autonomous, raising concerns about the loss of human oversight and the potential for uncontrolled escalation in digital conflicts [8].

III. Generative AI in Defensive Cyber Operations

In contrast to its offensive potential, Generative AI also serves as a cornerstone for next-generation cyber defense strategies. One of its most significant contributions is the creation of synthetic data for training machine learning models. High-quality, diverse datasets are crucial for detecting anomalies and recognizing emerging attack patterns, yet such data is often limited or sensitive. Generative models can produce realistic but non-identifiable datasets, allowing cybersecurity teams to improve detection systems without compromising privacy or security.

Moreover, AI-driven simulation environments powered by generative models enable proactive threat hunting. By generating simulated attacks, defenders can test the resilience of networks and refine defensive algorithms in advance [9]. This predictive capacity transforms cybersecurity from a reactive to a preventive discipline. Deception technologies also benefit from generative capabilities, as AI can automatically create convincing honeypots and decoy systems that lure attackers into controlled environments, gathering intelligence without exposing real assets [10].

Generative AI further enhances automated incident response. LLM-based agents can analyze large volumes of system logs, detect irregularities, and generate real-time remediation scripts. When integrated with Security Information and Event Management (SIEM) systems, these models accelerate response times and reduce human fatigue in Security Operations Centers (SOCs). Additionally, generative frameworks can assist in vulnerability assessment by simulating potential exploit paths and generating countermeasures automatically, thus closing security gaps before they can be exploited[11]. However, deploying Generative AI in defense also raises ethical and practical challenges. The use of deception and synthetic content, even for protection, must be carefully governed to prevent misuse or unintended consequences. Overreliance on AI automation could also lead to blind spots or exploitation if adversaries succeed in poisoning generative models with manipulated data [12].

Pages: 70-75
Volume-II, Issue-IV (2025)
https://balticpapers.com/index.php/bjmr

IV. Conclusion:

Generative AI stands at the forefront of a new era in cybersecurity, shaping both offensive and defensive paradigms with unparalleled power. While adversaries exploit its creative and adaptive potential to launch sophisticated, automated attacks, defenders harness the same capabilities to predict, deceive, and neutralize threats before they materialize. The convergence of these dynamics underscores the dual-edged nature of Generative AI—capable of fostering security innovation while amplifying risk. As cyber operations grow increasingly autonomous, the future of digital defense will depend not only on technological advancement but also on ethical governance, transparency, and human-centered design to ensure that AI strengthens global security rather than destabilizing it.

REFERENCES:

[1] C. R. Borra, R. V. Rayala, P. K. Pareek, and S. Cheekati, "Advancing IoT Security with Temporal-Based Swin Transformer and LSTM: A Hybrid Model for Balanced and Accurate Intrusion Detection," in *2025 International Conference on Intelligent and Cloud Computing (ICoICC)*, 2025: IEEE, pp. 1-7.

[2] S. Akter, A. Marzan, and N. Mazher, "Expanding the AI Health Frontier: From Public Trends to Genomic and Visual Data Insights," *Pioneer Research Journal of Computing Science,* vol. 2, no. 2, pp. 206-223, 2025.

[3] C. R. Borra, R. V. Rayala, P. K. Pareek, and S. Cheekati, "Optimizing Security in Satellite-Integrated IoT Networks: A Hybrid Deep Learning Approach for Intrusion Detection with JBOA and NOA," in *2025 International Conference on Intelligent and Cloud Computing (ICoICC)*, 2025: IEEE, pp. 1-8.

[4] H. Allam, J. Dempere, V. Akre, D. Parakash, N. Mazher, and J. Ahamed, "Artificial intelligence in education: an argument of Chat-GPT use in education," in *2023 9th International Conference on Information Technology Trends (ITT)*, 2023: IEEE, pp. 151-156.

[5] R. V. Rayala, S. Cheekati, M. Ruzieva, V. Vasudevan, C. R. Borra, and R. Sultanov, "Optimized Deep Learning for Diabetes Detection: A BGRU-based Approach with SA-GSO Hyperparameter Tuning," in *2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, 2025: IEEE, pp. 1-6.

[6] N. Mazher, I. Ashraf, and A. Altaf, "Which web browser work best for detecting phishing," in *2013 5th International Conference on Information and Communication Technologies*, 2013: IEEE, pp. 1-5.

[7] S. Cheekati, R. V. Rayala, and C. R. Borra, "A Scalable Framework for Attack Detection: SWIM Transformer with Feature Optimization and Class Balancing," in *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS)*, 2025: IEEE, pp. 1-7.

[8] B. Namatherdhala, N. Mazher, and G. K. Sriram, "A comprehensive overview of artificial intelligence tends in education," *International Research Journal of Modernization in Engineering Technology and Science,* vol. 4, no. 7, pp. 61-67, 2022.

[9] M. Tayal, A. Singh, S. Kolathaya, and S. Bansal, "A physics-informed machine learning framework for safe and optimal control of autonomous systems," *arXiv preprint arXiv:2502.11057,* 2025.

[10] B. Namatherdhala, N. Mazher, and G. K. Sriram, "Uses of artificial intelligence in autonomous driving and V2X communication," *International Research Journal of Modernization in Engineering Technology and Science,* vol. 4, no. 7, pp. 1932-1936, 2022.

[11] R. V. Rayala, C. R. Borra, V. Vasudevan, S. Cheekati, and J. U. Rustambekovich, "Enhancing Renewable Energy Forecasting using Roosters Optimization Algorithm and Hybrid Deep Learning Models," in *2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)*, 2025: IEEE, pp. 1-6.

[12] B. Othman and N. Mazher, "Data-Driven Degradation Modeling in Batteries Using Sparse Feature Selection," *Journal of Data and Digital Innovation (JDDI),* vol. 2, no. 2, pp. 41-50, 2025.