AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-3| Issue-11| 2025 Published: |30-12-2025|

COUNTERING AI-GENERATED CYBERATTACKS: NEW THREATS AND DEFENSE METHODS

https://doi.org/10.5281/zenodo.17959612

Lieutenant Colonel **Niyazov Erkin Shamsievich**,
Associate Professor of the Special Training Cycle, Faculty of Military Education
Komilov Mekhriddin Malikovich
2nd-year Cadet, Faculty of Military Education

Abstract

The paper examines modern cyberattacks supported by artificial intelligence (AI) technologies, as well as methods for countering them. The nature and scale of threats are analyzed, including attack automation, AI-driven phishing, generation of deepfakes, autonomous operations, and the use of AI to bypass traditional security measures. Special attention is given to defense mechanisms such as AI-based anomaly detection, Explainable AI (XAI), behavioral analysis, and innovative cyber defense tactics. The study substantiates the need to develop new approaches to protecting digital systems and to enhance the level of training of cybersecurity specialists.

Keywords

cyberattacks, artificial intelligence, AI threats, defense methods, attack detection, explainable AI (XAI), threat automation, AI security.

1. RELEVANCE OF THE TOPIC

In recent years, artificial intelligence has become not only a tool for defense but also a powerful means of attack automation. OpenAI warns that modern AI models are capable of creating sophisticated exploits that can be used by malicious actors to bypass security systems and attack critical infrastructure.

Reuters

Moreover, most organizations are insufficiently protected against AI-optimized attacks; surveys show that only about 20% of companies consider themselves ready to counter AI-driven bot attacks.

Axios

The growth of new attack methods, including automated phishing, generation of realistic deepfake messages, and the use of AI bots to bypass identification mechanisms, makes this topic particularly important for researchers and cybersecurity practitioners.

TechRadar

2. INTRODUCTION

Cyber threats are undergoing qualitative changes under the influence of artificial intelligence. AI tools help attackers automate reconnaissance, scale phishing campaigns, create realistic fake messages, and identify vulnerabilities faster than traditional security systems can respond. As a result, conventional computer security methods often prove insufficient and require adaptation to new challenges.

Scientific Brigadier

3. NEW AI-GENERATED THREATS

3.1 Attack Automation

AI enables thousands of reconnaissance operations and vulnerability scans to be conducted at high speed, making defense against mass attacks significantly more difficult.

TechRadar

3.2 AI Phishing and Deepfakes

Generative models create convincing phishing emails or fake identities, increasing the likelihood of successful system compromise.

The Guardian

3.3 Autonomous Malicious Campaigns

Reports on the use of AI for autonomous control of offensive operations indicate the evolution of threats—these are no longer just tools, but entire attacks managed by AI systems.

AP News

3.4 Advanced Zero-Day Threats

Attackers use machine learning capabilities to identify and exploit previously unknown vulnerabilities ("zero-day" attacks).

link.springer.com

4. METHODS OF DEFENSE AGAINST AI-DRIVEN CYBERATTACKS

4.1 AI-Based Anomaly and Threat Detection

The use of machine learning and deep learning makes it possible to effectively identify deviations in network and user behavior, significantly accelerating attack detection.

journal.uir.ac.id

4.2 Explainable AI (XAI)

Explainable AI increases model transparency, helping analysts understand why certain activities are considered suspicious and strengthening trust in defensive systems.

arxiv.org

4.3 Behavioral Analysis

Analyzing the behavior of users and machines helps identify synthetic identities and abnormal activity typical of automated attacks.

TechRadar

4.4 Deepfake Countermeasures

Specialized software, such as deepfake detectors, helps block visual and audio manipulations.

Wikipedia

5. DISCUSSION

Research indicates that the use of AI in defense provides significant advantages; however, it also introduces new risks—such as attacks on AI systems themselves through training data poisoning or model evasion. Effective defense requires a combination of algorithmic innovation, organizational measures, and improved specialist training.

cyberrus.info

6. CONCLUSIONS

Artificial intelligence has strengthened both cyber threats and defensive capabilities.

New threats in the form of AI phishing, attack automation, and autonomous cyber activities require innovative defense approaches.

AI- and Explainable AI-based defense systems demonstrate high potential but require continuous adaptation and improvement.

International cooperation, standardization of defense methods, and the training of qualified specialists are essential.

REFERENCES:

The Rise of AI-Powered Cybersecurity Threats and the Evolution of Defense Mechanisms. IJRASET Journal.

IJRASET

AI-Powered Cyberattacks: A Comprehensive Review and Analysis of Emerging Threats. Advances in IT and Electrical Engineering.

Prz Journals

Enhancing Cybersecurity through AI-Powered Security Mechanisms. IT Journal Research and Development.

journal.uir.ac.id

AI will make scam emails look genuine, UK cybersecurity agency warns — The Guardian.

The Guardian

OpenAI warns new models pose 'high' cybersecurity risk — Reuters.

Reuters

Inside the AI-powered assault on SaaS: why identity is the weakest link — TechRadar.

TechRadar

Vastav.AI — deepfake detection system.

Wikipedia

Emerging AI threats in cybercrime: zero-day attacks. Knowledge and Information Systems.

link.springer.com

Explainable Artificial Intelligence Applications in Cyber Security.

arxiv.org

Cyber threats & AI risk analyses (Arkose Labs survey).

Axios