УДК 004.056:004.8

AI-ГЕНЕРАТОР АТАК ЯК ІНСТРУМЕНТ ПРОАКТИВНОГО КІБЕР-ТЕСТУВАННЯ ДЕРЖАВНИХ ЦИФРОВИХ СЕРВІСІВ

Товкун Ю.І.
e-mail: yuliia.tovkun@nure.ua
Науковий керівник – д.т.н., проф. Хаханова Г.В.
Харківський національний університет радіоелектроніки, каф. АПОТ
м. Харків, Україна

The paper substantiates the concept of an AI-based attack generator as a proactive cybersecurity testing instrument for governmental digital services. Traditional threat modeling and periodic penetration testing lack scalability and predictive capability in the context of rapidly evolving and AI-assisted cyber threats. The proposed approach employs large language models to generate structured and machine-readable attack scenarios based on the architectural description of public digital platforms. The generator enables dynamic combination of tactics and techniques aligned with MITRE ATT&CK and STRIDE models, facilitating multi-step attack path modeling. Unlike static security assessment methods, the AI-based mechanism supports continuous expansion of the attack surface analysis and risk-oriented prioritization of protection measures. The proposed concept contributes to the transition from reactive to adaptive and proactive cyber resilience in digital government infrastructures.

Стрімка цифровізація державного управління формує новий клас кіберризиків, пов’язаних не лише з наявністю вразливостей, а з еволюцією методів їх експлуатації. Сучасні атаки дедалі частіше характеризуються комбінуванням технік, автоматизацією та адаптивністю до архітектурних особливостей цільових систем. У таких умовах традиційні моделі оцінки безпеки, що базуються на статичному threat modeling або періодичному penetration testing, втрачають достатню прогностичну здатність [1].

Проблема полягає у відсутності механізму системного та безперервного генерування релевантних сценаріїв атак для державних цифрових платформ з урахуванням їх конкретної архітектури. Більшість існуючих інструментів тестування орієнтовані на перевірку відомих класів уразливостей і не забезпечують дослідження нових комбінацій технік, що можуть виникати внаслідок змін конфігурації або інтеграцій[2].

У роботі обґрунтовується концепція AI-генератора атак як інструменту проактивного кібер-тестування державних сервісів. Запропонований підхід полягає у використанні великої мовної моделі для формування структурованих сценаріїв атак на основі опису архітектури системи. Принциповим є перехід від текстових описів до формалізованого сценарію,

що включає: цільовий компонент, передумови реалізації, метод впливу, очікуваний ефект та можливу послідовність подальших кроків.

На відміну від традиційного експертного аналізу, генеративна модель здатна автоматично комбінувати техніки з різних тактик (згідно MITRE ATT&CK) та формувати багатокрокові attack path [3]. Це дозволяє досліджувати не окремі вразливості, а траєкторії компрометації системи. Таким чином, AI-генератор виступає цифровим Red Team-агентом, який забезпечує розширення attack surface аналізу та підвищує варіативність тестування.

Теоретичною основою підходу є інтеграція моделей STRIDE та MITRE ATT&CK з генеративними можливостями LLM. STRIDE забезпечує категоризацію загроз за напрямами порушення конфіденційності, цілісності та доступності, тоді як MITRE ATT&CK дозволяє формалізувати поведінкові аспекти атакуючого. Генеративна модель створює механізм динамічного комбінування технік у межах конкретної архітектури сервісу [4].

Проактивність запропонованого підходу полягає у зміщенні акценту з реагування на інциденти до їх попереднього моделювання. Згенеровані сценарії можуть ранжуватися за рівнем ризику залежно від потенційного впливу та критичності компонентів. Це створює передумови для ризик-орієнтованої пріоритезації заходів захисту.

Наукова новизна полягає у концептуалізації AI-генерації атак як формалізованого інструменту проактивного тестування державних цифрових сервісів та у визначенні структури сценарію атаки, придатної для подальшої автоматизованої симуляції. Практична значущість підходу полягає у можливості його інтеграції в процеси DevSecOps та формуванні механізму безперервного оновлення моделі загроз.

Таким чином, використання генеративних моделей штучного інтелекту дозволяє перейти від статичної до динамічної парадигми кібербезпеки державних сервісів, де тестування стає адаптивним, масштабованим та ризик-орієнтованим процесом.

Список використаних джерел:

1. Taddeo M., Floridi L. How AI can be a force for good in cybersecurity [Електронний ресурс] // Science. 2018. Vol. 361, No. 6405. P. 751–752. URL: https://doi.org/10.1126/science.aat5991 (дата звернення: 25.02.2026).

2. Buczak A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection [Електронний ресурс] // IEEE Communications Surveys & Tutorials. 2016. Vol. 18, No. 2. P. 1153–1176. URL: https://doi.org/10.1109/COMST.2015.2494502 (дата звернення: 25.02.2026).

3. MITRE ATT&CK® framework [Електронний ресурс]. MITRE Corporation. URL: https://attack.mitre.org (дата звернення: 25.02.2026).

4. National Institute of Standards and Technology. Computer security incident handling guide (SP 800-61 Rev. 2) [Електронний ресурс]. Gaithersburg : NIST, 2012. URL: https://csrc.nist.gov/publications/detail/sp/800-61/rev-2/final (дата звернення: 25.02.2026).