

# АНАЛІЗ ЗАСОБІВ ТЕСТУВАННЯ НА ПРОНИКНЕННЯ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ

Вінницький національний технічний університет

## Анотація

У сучасному інформаційному середовищі, тестування на проникнення є одним з найважливіших етапів забезпечення інформаційної безпеки. З допомогою даного підходу можна виявляти та усувати вразливості до того, як зловмисники знайдуть та скористаються ними. З прогресом штучного інтелекту особливо в останні роки, все частіше в основі сучасних засобів тестування на проникнення лежать інтелектуальні алгоритми. Саме за допомогою алгоритмів штучного інтелекту здійснюється автоматизоване виявлення та експлуатація вразливостей. У роботі проведено аналіз основних підходів до використання штучного інтелекту в тестуванні на проникнення, методи машинного навчання для глибокого аналізу мережевого трафіку, використання генеративних алгоритмів для імітації атак. Проведено дослідження переваг та викликів впровадження таких рішень, зважаючи на такі аспекти як: точність, продуктивність та етичність.

**Ключові слова:** тестування на проникнення, штучний інтелект, кібербезпека, машинне навчання, автоматизація атак.

## Abstract

In the modern information environment, penetration testing is one of the most important stages of ensuring information security. This approach allows for the identification and elimination of vulnerabilities before malicious actors can find and exploit them. With the rapid advancement of artificial intelligence, particularly in recent years, intelligent algorithms are increasingly forming the basis of modern penetration testing tools. AI algorithms enable automated vulnerability detection and exploitation. This study analyzes the main approaches to using artificial intelligence in penetration testing, including machine learning methods for deep network traffic analysis and the use of generative algorithms for attack simulation. The research explores the benefits and challenges of implementing such solutions, considering aspects such as accuracy, performance, and ethics.

**Keywords:** penetration testing, artificial intelligence, cybersecurity, machine learning, attack automation.

## Вступ

У сучасності забезпечення кібербезпеки є одним із ключових завдань для державних установ, комерційних організацій та окремих користувачів. Щороку, кількість кібератак та їх складність зростає [1], що викликає необхідність у більш досконалих методах виявлення вразливостей та оцінки рівня захисту інформаційних систем. Такий підхід як тестування на проникнення є одним з найефективніших підходів, що дозволяє достроково змоделювати дії зловмисника та оцінити рівень стійкості системи до атак [2].

Методи тестування на проникнення можуть автоматизовано сканувати вразливості, експлуатувати знайдені прогалини, аналізувати безпеку мережевого з'єднання, а також включає перевірку на стійкість до соціальної інженерії. З основних підходів тестування на проникнення можна виділити: «біла скринька», «чорна скринька» та «сіра скринька». Підхід «білої скриньки» передбачає надання повної інформації про архітектуру системи. «Чорна скринька» – це підхід, який не надає тестувальнику жодних даних про систему. «Сіра скринька» – це метод, що є певним компромісом серед двох попередніх [3].

Такий підхід надає велику низку можливостей для підвищення ефективності процесу завдяки покращенню автоматизації виявлення вразливостей, швидшому аналізу великих обсягів даних та навчанню реагування на аномальні інциденти у разі нових видів атак. Такий процес як машинне навчання дозволяє швидко розпізнавати аномальні патерни, виявляти слабкі місця, які були непомічені під час традиційного пентесту.

## Результати дослідження

Тестування на проникнення (penetration testing) є важливим елементом забезпечення кібербезпеки, оскільки дозволяє виявляти та усувати вразливості в інформаційних системах. З розвитком технологій штучного інтелекту (ШІ) з'явилися інструменти, які інтегрують ці технології для підвищення ефективності та точності тестування на проникнення. З таких можна виділити DeepExploit, Sn1per та Cortex XSOAR.

DeepExploit — це інструмент на базі штучного інтелекту, який автоматизує процес пошуку вразливостей та їх експлуатації. Він інтегрується з Metasploit, популярною платформою для тестування на проникнення, і дозволяє автономно сканувати та використовувати вразливі місця [4]. Використовуючи методи глибокого навчання з підкріпленням, DeepExploit постійно вдосконалюється і вибирає вразливості для експлуатації, замість застосування грубої сили. Це підвищує ефективність тестування та зменшує вплив на системи. З переваг такого інструменту можна виділити автоматизацію процесу тестування на проникнення, можливість інтегруватись з існуючими інструментами, а також використання глибокого навчання для адаптації нових вразливостей. Його недоліками є можливість пропуску вразливостей, які не були виявлені під час навчання моделі, а також потреба в значних обчислювальних ресурсах для навчання та роботи моделі.

Sn1per — це інструмент для автоматизації процесу розвідки та сканування вразливостей. Він використовує різні модулі для збору інформації про цільову систему, включаючи сканування портів, виявлення служб та інші методи [5]. Sn1per може бути корисним для швидкого виявлення потенційних вразливостей та оцінки поверхні атаки. Його основними перевагами є швидке збирання інформації про наявну систему, так само як у випадку з DeepExploit можливість інтеграції з іншими наявними інструментами та автоматизація процесу розвідки та сканування. Недоліком ж є можливість отримання великої кількості даних, які потребують додаткового аналізу, а також цей метод не завжди забезпечує глибокий аналіз вразливостей.

Cortex XSOAR (раніше відомий як Demisto) — це платформа для оркестрації, автоматизації та реагування на інциденти безпеки. Вона використовує штучний інтелект для автоматизації процесів реагування на інциденти, включаючи тестування на проникнення [6]. Cortex XSOAR дозволяє об'єднувати різні інструменти безпеки та автоматизувати робочі процеси для швидкого та ефективнішого реагування на загрози. Цей інструмент також має можливості інтеграції з широким спектром інструментів безпеки та використовує ШІ для покращення ефективності та точності. Його недоліками є складність налаштування та інтеграції в існуючу систему та потреба в навчанні персоналу для ефективного використання платформи.

PentestGPT — це інструмент для автоматизації тестування на проникнення, який використовує великі мовні моделі (LLM), зокрема ChatGPT, для надання покрокових рекомендацій та виконання дій, спрямованих на виявлення вразливостей у системах [7]. З основних характеристик PentestGPT можна виділити інтерактивний режим роботи, який надає користувачам покрокові інструкції під час тестування на проникнення. Також важливим фактором є рішення для CTF-завдань. Засіб здатен вирішувати завдання «HackTheBox» середньої складності та інші CTF-виклики, демонструючи свою ефективність у практичних сценаріях. Даний інструмент має модульну архітектуру, він складається з трьох взаємодіючих модулів, кожен з яких відповідає за певні підзадачі тестування на проникнення, що значно покращує ефективність [8]. PentestGPT автоматизує процес тестування, що дозволяє фахівцям зосередитись на більш складних аспектах аналізу. Так само він може інтегруватись з іншими інструментами безпеки, тому він являється гнучким та розширюваним. Даний засіб використовує LLM, завдяки цьому він може обробляти великі обсяги інформації та надавати точні рекомендації. З недоліків можна зазначити необхідність у значних обчислювальних ресурсах, тому він може бути обмежений для деяких користувачів.

Ці інструменти демонструють, як штучний інтелект може бути інтегрований у процеси тестування на проникнення для підвищення ефективності та точності виявлення вразливостей. Однак, важливо враховувати, що автоматизація не може повністю замінити експертний аналіз, і поєднання автоматизованих інструментів з ручним тестуванням забезпечує найбільш повний підхід до забезпечення безпеки.

## Висновки

Використання штучного інтелекту в тестуванні на проникнення відкриває нові можливості для підвищення рівня кібербезпеки, автоматизації процесів аналізу загроз та зменшення часу на пошук вразливостей. Подальші дослідження в цій сфері можуть бути спрямовані на розробку більш точних моделей машинного навчання, а також на створення механізмів захисту від використання таких технологій у шкідливих цілях.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Кількість Інститут масової інформації. Кількість кібератак на Україну у 2024 році зросла на 70%. 2024. URL: <https://imi.org.ua/news/u-2024-rotsi-kilkist-kiberatak-na-ukrayinu-zroslo-na-70-i65931>. (дата звернення: 18.03.2025).
2. BDO Ukraine. Ідентифікація вразливостей: тестування на проникнення як засіб захисту. 2024. URL: <https://www.bdo.ua/uk-ua/insights-2/information-materials/2024/identifying-vulnerabilities-penetration-testing-as-a-means-of-protection>. (дата звернення: 22.03.2025).
3. Datami. Що таке тестування на проникнення? 2024. URL: <https://datami.ee/ua/blog/what-is-penetration-testing/>. (дата звернення: 22.03.2025).
4. М.С. Колощук, О.Ю. Дячук, О.О. Окуньова, О.В. Пірог. Інструменти штучного інтелекту для автоматизації тестування на проникнення. 2024. URL: [https://library.ztu.edu.ua/e-copies/VISNUK/94\\_II/121.pdf?utm\\_source=chatgpt.com](https://library.ztu.edu.ua/e-copies/VISNUK/94_II/121.pdf?utm_source=chatgpt.com). (дата звернення: 22.03.2025).
5. Sn1per Security. Attack Surface Management. 2024. URL: <https://sn1persecurity.com/wordpress/>, вільний. (дата звернення: 22.03.2025).
6. Palo Alto Networks. Cortex. 2024. URL: <https://www.paloaltonetworks.com/cortex/cortex-xsoar>. (дата звернення: 22.03.2025).
7. Gelei Deng , Yi Liu , Víctor Mayoral-Vilches, Peng Liu , Yuekang Li , Yuan Xu , Tianwei Zhang, Yang Liu, Martin Pinzger, Stefan Rass. PENTESTGPT: Evaluating and Harnessing Large Language Models for Automated Penetration Testing. 2024. URL: <https://arxiv.org/pdf/2308.06782>. (дата звернення: 22.03.2025)
8. Abstracta. PentestGPT: Revolutionize Penetration Testing with AI. 2025. URL: <https://abstracta.us/blog/security-testing/pentestgpt-penetration-testing-with-ai/>. (дата звернення: 22.03.2025)

**Абрамюк Дмитро Андрійович** — студент групи ІБС-21б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: dima.abramiuk@gmail.com

**Abramiuk Dmytro A.** – student 1BS-21b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: dima.abramiuk@gmail.com

**Науковий керівник : Куперштейн Леонід Михайлович** — к. т. н., доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця, e-mail: kupershtein.lm@gmail.com

**Supervisor : Kupershtein Leonid M.** — Ph.D., Associate Professor of the Department of Information Protection, Vinnytsia National Technical University.