ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТРИЧНИХ ТА АНСАМБЛЕВИХ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ЕКСПЛУАТОВАНОСТІ ВРАЗЛИВОСТЕЙ У СИСТЕМАХ КІБЕРЗАХИСТУ

Петренко О.Є., Кононенко І.В., Скрильник В.Ю.
Харківський національний університет радіоелектроніки, Харків, Україна

Зростання кількості виявлених вразливостей програмного забезпечення та постійне вдосконалення методів кібератак ускладнюють своєчасне реагування на потенційні загрози в інформаційних системах. Особливої актуальності набуває визначення того, які саме вразливості з найбільшою ймовірністю будуть використані зловмисниками для реалізації кібератак.

Щороку у відкритій Національній базі вразливостей США (NVD) реєструється понад 25 тисяч нових помилок програмного забезпечення [1, 2]. Фахівці підрозділів захисту інформації не встигають своєчасно усувати таку кількість вразливостей, тому виникає потреба у виборі найбільш небезпечних із них. Зазвичай на практиці орієнтуються на базовий бал стандарту CVSS, з огляду на нього, першочергово закривають ті вразливості, які мають оцінку 9,0 балів і вище, що свідчить про «критичний» рівень [3]. Проте базовий бал CVSS показує лише критичність впливу вразливості і не дає відповіді на питання, чи з'явиться для неї реальний експлойт [1, 3]. Статистика свідчить, що зловмисники використовують у реальних атаках менше одного відсотка всіх відомих вразливостей [1, 2, 4]. Як наслідок, відсікання за балом CVSS 9,0 змушує адміністраторів витрачати час на тисячі вразливостей, які ніхто не атакує, і водночас пропускати реальні загрози з меншим балом [1, 2].

Пошук методів виявлення вразливостей, які мають малий відсоток хибних спрацювань є актуальним питанням в сучасних умовах. Для вирішення цієї проблеми доцільно використовувати алгоритми машинного навчання, здатні виявляти закономірності між характеристиками вразливостей та фактом їх практичної експлуатації. Особливий інтерес становить порівняння метричних методів, які оцінюють подібність між об'єктами, та ансамблевих методів, що поєднують результати роботи декількох моделей машинного навчання.

Метою доповіді є порівняння ефективності алгоритмів машинного навчання — метричного методу k-найближчих сусідів та ансамблевого методу Random Forest виявлення вразливостей та дослідження відсотку хибних спрацювань.

Для досягнення мети було вибрано вразливості серед 50 231 з бази NVD за 2022–2023 роки та зіставлено її з переліком реально зафіксованих кібератак CISA KEV. В процесі дослідження ефективності роботи алгоритмів сформовано два варіанти вхідних даних: базовий набір та розширений набір.

Базовий набір складається 8 стандартних параметрів вектору CVSS v3.1 (вектор атаки AV, складність атаки AC, необхідні права доступу PR, потреба у діях користувача UI, зміна області впливу S, а також вплив на

конфіденційність C, цілісність I та доступність A) [3]. Ці параметри являють собою текстові категорії, для обробки яких здійснено перетворення їх у бінарний ряд з 22 бінарних ознак. Розширений набір містить до 22 ознак CVSS та 26 бінарних ознак, які вказують на конкретний тип програмної помилки за класифікатором CWE. Зібрані для навчання дані розділено за часом, а саме моделі навчали на вразливостях за 2022 рік (23 457 записів, серед яких 90 реальних атак, або 0,38%), а перевіряли їхню роботу на вразливостях за весь 2023 рік (26 774 записи, серед яких 120 реальних атак, або 0,45%).

В доповіді наводяться результати порівняння методів машинного навчання: Random Forest з розширеним набором вхідних даних, Random Forest з базовим набором даних, метод k-найближчих сусідів з розширеним набором вхідних даних, модифікації методу k-найближчих сусідів з базовим та розширеним наборами даних, проста таблиця частот за вектором CVSS. Порівняння методів здійснено за трьома практичними критеріями, а саме: площа під кривою точності та повноти, відсоток виявлених реальних атак при перевірці перших 5%, точність, повнота та F1-міра.

Площа під кривою точності та повноти (PR-AUC) — показує загальну якість упорядкування вразливостей за ризиком (для випадкового вгадування цей показник дорівнює частці атак у базі, тобто 0,0045).

Відсоток виявлених реальних атак при перевірці перших 5% вибірки вразливостей показує, яку частку зі 120 реальних хакерських атак за 2023 рік зможуть нейтралізувати фахівці з кіберзахисту, якщо їх ресурсів вистачає на перевірку лише 5% найбільш підозрілих вразливостей (близько 1 338 записів на рік).

Точність (Precision), повнота (Recall), F1-міра та загальна кількість тривог при пороговому відсіканні.

Отримані результати дозволили зробити наступні висновки.

По-перше, використання лише базового набору даних не дозволяє порівняти ефективність методів. Так, модифікований метод k-найближчих сусідів має показник PR-AUC = 0,0097, Random Forest – має показник PR-AUC = 0,0099 і показує однаковий результат з таблицею частот за вектором. Розглянуті методи дозволяють знайти близько 14,5–14,6% реальних вразливостей при перевірці перших 5% списку, що помітно ефективніше за сортування за стандартним балом CVSS (10,3%).

По-друге, додавання інформації про тип помилки, тобто застосування розширеного набору даних, знижує ефективності роботи методу к-найближчих сусідів до 4,3% Обумовлено зниження ефективності однаковою вагою усіх бінарних даних, тому застосування розширеної вибірки «розмивають» важливі параметри CVSS. На відміну від методу к-найближчих сусідів метод Random Forest відбирає найбільш небезпечні типи помилок. Завдяки цьому його показник PR-AUC зростає до 0,0134 (майже втричі вищий за випадковий вибір), а при пороговій класифікації видає лише 187 вразливостей із точністю 3,2% проти 4 442 вразливостей із точністю 0,9% у порівнянні простою таблицею частот CVSS $\geq$ 9,0.

Дослідження вразливостей з низьким значенням вектору атаки, а саме віддаленого виконання коду в ядрі PHP — CVE-2026-109618 та CVE-2026-109619 [5,6] показало, що метод Random Forest надає оцінку вразливості вище за 98,1% , тоді як модифікований метод k-найближчик сусідів має рівень 86,8%.

Враховуючи високу складність експлуатації за офіційним звітом розробників [5] бал за простою таблицею частот CVSS дорівнює 8,1, оцінка обох методів знижується до рівня 85,1–86,6% (не потрапляючи у 5-відсоткову квоту перевірки).

З огляду на це, методи, які спираються лише на технічний опис вразливостей не враховують масштаб її розповсюдження [1, 2].

В ході дослідження з'ясовано, що стандартний алгоритм k-найближчих сусідів має низьку ефективність при застосуванні базового набору даних через велику кількість однакових векторів і потребує попереднього групування дублікатів.

При застосуванні розширеного набору ансамблевий метод Random Forest суттєво переважає метод k-найближчих сусідів за точністю виявлення вразливостей та дозволяє в десятки разів зменшити кількість хибних спрацьовувань, порівняно з пороговим правилом CVSS $\geq 9{,}0$.

Для подальшого підвищення точності прогнозування до вхідних даних моделі доцільно додати показники розповсюдження програмного забезпечення та дані від первинних розробників (CNA) [1, 2, 5, 6].

Список літератури

1. Jacobs J., Romanosky S., Edwards B., Adjerid I., Roytman M. Exploit Prediction Scoring System (EPSS). Digital Threats: Research and Practice. 2021. Vol. 2, no. 3. Art. 20. P. 1–17. URL: https://doi.org/10.1145/3436242

2. Jacobs J., Romanosky S., Suciu O., Edwards B., Sarabi A. Enhancing Vulnerability Prioritization: Data-Driven Exploit Predictions with Community-Driven Insights. 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Delft, Netherlands, 3–7 July 2023. P. 194–206. URL: https://doi.org/10.48550/arXiv.2302.14172

3. FIRST.Org, Inc. Common Vulnerability Scoring System v3.1: Specification Document. URL: https://www.first.org/cvss/v3.1/specification-document

4. Cybersecurity and Infrastructure Security Agency (CISA). Known Exploited Vulnerabilities Catalog (catalogVersion 2026.10.29). URL: https://www.cisa.gov/known-exploited-vulnerabilities-catalog

5. The PHP Group. Security advisory: Critical remote code execution vulnerabilities, update now (CVE-2026-109618, CVE-2026-109619). PHP.net News Archive 2026. 21.10.2026. URL: https://www.php.net/archive/2026.php#2026-10-21-5

6. Cybersecurity and Infrastructure Security Agency (CISA). CISA Adds Two Known Exploited Vulnerabilities to Catalog (CVE-2026-109618, CVE-2026-109619). CISA bulletin. 26.10.2026. URL: https://content.govdelivery.com/accounts/USDHSCISA/bulletins/4309083