УДК 004.056

Боцанюк І.М., магістрант
Головня О.С., к.пед.н., доцент
Державний університет «Житомирська політехніка»

АДАПТИВНА МОДЕЛЬ ПІДВИЩЕННЯ СТІЙКОСТІ СИСТЕМ ВИЯВЛЕННЯ ФІШИНГУ

Сучасні системи виявлення фішингових повідомлень базуються на поєднанні методів машинного навчання та евристичних підходів[1]. Машинне навчання забезпечує аналіз контексту повідомлення, включаючи семантичні характеристики, тоді як евристичні методи дозволяють оцінювати структурні ознаки, зокрема властивості URL-адрес і метадані. Використання гібридних підходів є доцільним, однак їх ефективна інтеграція залишається проблемною через відсутність універсального механізму узгодження результатів[2].

Основна складність полягає у відмінності форматів вихідних даних. Моделі машинного навчання формують клас разом із ймовірнісною оцінкою, тоді як евристичні алгоритми генерують бальні оцінки критичності ознак. Це ускладнює безпосереднє поєднання результатів і може призводити до суперечливих рішень у випадках, коли різні підходи сигналізують про протилежні висновки.

Додатково, ймовірнісні оцінки моделей не завжди є коректно каліброваними, що впливає на достовірність їх інтерпретації та може спричиняти зміщення у прийнятті рішень[3]. Евристичні підходи, у свою чергу, характеризуються обмеженою гнучкістю та залежністю від заздалегідь визначених правил, що ускладнює їх адаптацію до нових типів фішингових атак. У сукупності ці фактори знижують ефективність систем виявлення.

Метою дослідження є розробка адаптивного методу агрегування результатів машинного та евристичного аналізу електронних повідомлень для підвищення точності виявлення фішингу.

Більшість існуючих підходів використовують статичні схеми інтеграції, такі як фіксовані вагові коефіцієнти, порогові значення або просте голосування[4]. Такі методи не враховують контекст повідомлення та не здатні динамічно реагувати на зміну характеристик вхідних даних, що є критичним у середовищі кіберзагроз, яке постійно еволюціонує.

Запропонований підхід передбачає адаптивне агрегування результатів. На першому етапі виконується нормалізація даних шляхом приведення ймовірнісних і бальних оцінок до єдиної шкали. Для цього

можуть застосовуватись методи калібрування, що забезпечують більш коректну інтерпретацію рівня впевненості моделей машинного навчання.

Другий етап включає аналіз контексту повідомлення, зокрема виявлення ознак, що визначають релевантність кожного підходу. До таких ознак можуть належати структура повідомлення, наявність підозрілих посилань, складність тексту та інші характеристики, що впливають на достовірність оцінок.

На третьому етапі здійснюється агрегування результатів із використанням динамічного зважування. Вагові коефіцієнти змінюються залежно від контексту та рівня впевненості кожного джерела. Додатково може враховуватись узгодженість результатів між підходами, що дозволяє зменшити вплив суперечливих оцінок.

Очікується, що застосування адаптивного підходу дозволить підвищити точність класифікації, зменшити кількість помилок та покращити здатність системи до виявлення нових типів фішингових атак. Така модель є більш стійкою до змін у поведінці зловмисників і може бути інтегрована в існуючі системи кіберзахисту без суттєвих змін їх архітектури.

Подальші дослідження доцільно спрямувати на експериментальну перевірку ефективності методу, аналіз впливу різних стратегій зважування, а також розширення набору ознак, що використовуються у процесі прийняття рішень.

Список використаних джерел

1. A comprehensive survey of AI-enabled phishing attacks detection techniques / Basit A., Zafar M., Liu X. та ін. // Telecommunication Systems. 2021. Vol. 76, No. 1. P. 139–154.

2. Intelligent web-phishing detection and protection scheme using integrated features of Images, frames and text / Adebowale M. A., Lwin K. T., Sanchez E., Hossain M. A. // Expert Systems with Applications. 2019. Vol. 115. P. 300–313.

3. On calibration of modern neural networks / Guo C., Pleiss G., Sun Y., Weinberger K. Q. // Proceedings of the 34th International Conference on Machine Learning. 2017. Vol. 70. P. 1321–1330

4. Taha A. Intelligent ensemble learning approach for phishing website detection based on weighted soft voting. // Mathematics. 2021. Vol. 9, No. 21. P. 2799.