https://doi.org/10.31891/2219-9365-2025-84-47
УДК 004.932.056.5(045)

БІЛОУС Віталій
Вінницький національний технічний університет
https://orcid.org/0009-0001-2350-1583
pydev@ukr.net
ПАВЛОВСЬКИЙ Павло
Вінницький національний технічний університет
https://orcid.org/0009-0001-1730-4102
prepod@vntu.net
ЗОРЯ Ірина
Вінницький національний технічний університет
https://orcid.org/0009-0002-7151-3455
iryna.zoria03@gmail.com
ФЕРНЕГА Євгеній
Вінницький національний технічний університет
https://orcid.org/0009-0007-3729-0178
evgeniyfernega@gmail.com
ГРЕСЬ Олександр
Чернівецький національний університет імені Юрія Федьковича
https://orcid.org/0000-0002-8465-193X
e-mail::o.hres@chnu.edu.ua

ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ АВТОМАТИЧНОГО ВИЯВЛЕННЯ ФІШИНГОВИХ САЙТІВ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Дослідження спрямоване на вдосконалення системи автоматизованого виявлення фішингових вебресурсів із використанням нейронних мереж, реалізованих на основі бібліотеки PHP-ML. У роботі здійснено ґрунтовний аналіз сучасних підходів до детекції фішингових сайтів, а також визначено їхні переваги та обмеження з позицій ефективності, адаптивності та практичного застосування.

У межах дослідження запропоновано та програмно реалізовано покращений метод ідентифікації фішингових ресурсів, який базується на застосуванні нейромережевих моделей для автоматизованого аналізу URL-адрес і вмісту вебсторінок. Розроблений підхід характеризується здатністю до адаптації в умовах появи нових векторів атак і врахування змін у динамічному інтернет-середовищі. З метою підвищення точності класифікації використано комплексний аналіз, що охоплює метадані, поведінкові ознаки вебресурсів та особливості структури HTML-коду. Застосування бібліотеки PHP-ML забезпечило ефективну інтеграцію методів машинного навчання у процес розроблення вебзастосунків, що створює передумови для побудови продуктивних та масштабованих систем кіберзахисту.

Експериментальні результати підтверджують результативність запропонованої системи, яка демонструє високий рівень точності виявлення фішингових сайтів за умови зниженого рівня хибних спрацьовувань. Запропонований підхід може бути використаний для підвищення рівня інформаційної безпеки в таких галузях, як електронна комерція, фінансово-банківський сектор та корпоративні інформаційні системи, забезпечуючи багаторівневий захист від фішингових атак у процесі взаємодії з вебресурсами.

Ключові слова: фішингові вебсайти, нейронні мережі, машинне навчання, PHP-ML, автоматизована детекція, інтернет-загрози.

BILOUS Vitalii, PAVLOVSKYI Pavlo, ZORIA Iryna, FERNEHA Yevhenii
Vinnytsia National Technical University
HRES Oleksandr
Yuriy Fedkovych Chernivtsi National University

INCREASING THE EFFICIENCY OF AUTOMATIC DETECTION OF PHISHING SITES BASED ON NEURAL NETWORKS

The research is aimed at improving the system of automated detection of phishing web resources using neural networks implemented on the basis of the PHP-ML library. The work provides a thorough analysis of modern approaches to detecting phishing sites, as well as their advantages and limitations in terms of efficiency, adaptability and practical application.

Within the framework of the research, an improved method of identifying phishing resources was proposed and programmatically implemented, which is based on the use of neural network models for automated analysis of URL addresses and web page content. The developed approach is characterized by the ability to adapt in the conditions of the emergence of new attack vectors and taking into account changes in the dynamic Internet environment. In order to increase the accuracy of classification, a comprehensive analysis was used, covering metadata, behavioral features of web resources and features of the HTML code structure. The use of the PHP-ML library ensured the effective integration of machine learning methods into the process of developing web applications, which creates the prerequisites for building productive and scalable cyber protection systems.

Experimental results confirm the effectiveness of the proposed system, which demonstrates a high level of accuracy in detecting phishing sites with a reduced level of false positives. The proposed approach can be used to increase the level of information security in such industries as e-commerce, the financial and banking sector, and corporate information systems, providing multi-level protection against phishing attacks in the process of interacting with web resources.

Keywords: phishing websites, neural networks, machine learning, PHP-ML, automated detection, Internet threats

ISSN 2219-9365

. Стаття надійшла до редакції / Received 12.11.2025
Прийнята до друку / Accepted 07.12.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Актуальність обраної тематики зумовлена інтенсивним зростанням кількості фішингових атак, які створюють суттєві загрози для захищеності інформаційних систем і конфіденційності персональних даних користувачів. Із розвитком цифрових технологій та електронної комерції кіберзловмисники постійно вдосконалюють інструменти та підходи, формуючи підроблені вебресурси, що за зовнішніми ознаками практично не відрізняються від легітимних. Основною метою таких атак є несанкціоноване отримання чутливої інформації, зокрема облікових даних, паролів, реквізитів платіжних карток або доступу до корпоративних інформаційних систем. У сучасних умовах фішинг виступає не лише засобом кіберзлочинної діяльності, а й одним із ключових механізмів соціальної інженерії, що підкреслює актуальність цієї проблеми для наукових досліджень.

За даними провідних компаній у сфері кібербезпеки, щорічне зростання кількості фішингових атак за попередні роки становило у середньому 20–30%, а за даними TrendMicro, кількість фішингових атак у 2023 році зросла на 58%. Фінансові збитки, за їх оцінками, у 2024 році склали 3,5 мільярда доларів США.

При цьому традиційні засоби захисту, зокрема антивірусні рішення та стандартні механізми фільтрації, часто виявляються недостатньо ефективними для протидії сучасним фішинговим загрозам через постійну зміну тактик і стратегій зловмисників.

Кіберзловмисники активно застосовують техніки обходу автоматизованих механізмів перевірки, зокрема CAPTCHA, а також скрипти, що динамічно змінюють вміст вебресурсів залежно від IP-адреси користувача, параметрів браузера чи результатів геолокації. Унаслідок цього статичні методи виявлення, які базуються виключно на незмінних ознаках, поступово втрачають свою ефективність та не забезпечують належного рівня протидії сучасним загрозам [1].

Це зумовлює потребу в розробленні нових підходів до виявлення та запобігання фішинговим атакам, які здатні враховувати динамічний характер інтернет-середовища та ідентифікувати навіть незначні ознаки фальшивих вебресурсів [2].

У зазначеному контексті застосування методів машинного навчання та нейронних мереж створює передумови для суттєвого підвищення ефективності систем детекції фішингу. Нейромережеві моделі забезпечують можливість виявлення складних закономірностей, опрацювання великих масивів даних і адаптації до нових типів загроз, що робить їх доцільним інструментом для розв'язання цієї задачі [3]. Водночас значна частина таких рішень характеризується складністю інтеграції у реальні вебзастосунки, що зумовлено використанням мов програмування та платформ, які не завжди сумісні зі стандартними вебтехнологіями.

Для реалізації програмних механізмів розпізнавання фішингових вебресурсів зазвичай застосовуються багаторівневі системи, що інтегрують різні методи аналізу в межах єдиної архітектури. Типова структура такої системи передбачає наявність кількох ключових компонентів [4]:

модуль збору даних забезпечує отримання та первинну обробку відомостей про вебресурс, зокрема URL-адрес, метаданих, HTTP-заголовків, вмісту сторінок та інших релевантних характеристик;

модуль попереднього аналізу виконує базову фільтрацію й евристичну оцінку з метою виявлення очевидних ознак фішингу, наприклад використання підозрілих доменних імен або відсутність SSL-сертифіката;

модуль поглибленого аналізу реалізує методи машинного навчання та нейронні мережі для всебічного дослідження контенту вебсайту, його поведінкових параметрів і структурних особливостей, на основі чого здійснюється класифікація ресурсу як легітимного або фішингового;

модуль ухвалення рішень формує остаточний висновок за результатами аналізу та, у разі виявлення загрози, ініціює сповіщення адміністратора або кінцевого користувача;

модуль самонавчання забезпечує автоматичне оновлення моделей розпізнавання на основі нових даних, отриманих у процесі експлуатації, що сприяє адаптації системи до нових типів атак і зниженню рівня хибних спрацьовувань.

До основних переваг програмного підходу до детекції фішингових сайтів належать високий рівень автоматизації, адаптивність до змін у загрозовому середовищі, оперативність аналізу та здатність ідентифікувати складні ознаки, які не піддаються статичним методам оцінювання. Застосування алгоритмів машинного навчання дає змогу зменшити кількість хибнопозитивних результатів і своєчасно виявляти нові різновиди атак, що робить такі системи ефективним засобом захисту як для індивідуальних користувачів, так і для корпоративних мереж.

Важливою перевагою є можливість інтеграції зазначених систем у реальні обчислювальні середовища, що забезпечує їх практичну придатність для розгортання у складі масштабних інфраструктур або як компонент антивірусного програмного забезпечення.

Дослідження потенціалу бібліотеки PHP-ML, яка забезпечує використання методів машинного навчання безпосередньо у середовищі PHP, є актуальним науково-практичним завданням, орієнтованим на спрощення впровадження інтелектуальних алгоритмів у чинні вебплатформи.

Таким чином, розроблення системи автоматизованого виявлення фішингових вебсайтів на основі нейронних мереж із використанням PHP-ML є важливим науково-прикладним напрямом, що поєднує сучасні досягнення у сфері штучного інтелекту та веброзроблення. Реалізація такого підходу дає змогу створити ефективний інструмент протидії фішинговим атакам, який легко інтегрується у вебзастосунки, знижує ризики для користувачів і підвищує загальний рівень безпеки у кіберпросторі.

ОСНОВНА ЧАСТИНА

Формування підходу до виявлення фішінгових сайтів

У контексті зростання кількості фішингових атак актуальним завданням є розробка ефективних методів автоматизованого виявлення фішингових вебресурсів.

Запропонований підхід базується на аналізі сукупності потенційно небезпечних факторів із інтеграцією алгоритмів машинного навчання, зокрема нейронних мереж.

Основна ідея методу полягає у виявленні фішингових ознак шляхом комбінованого аналізу структурних, контентних та поведінкових характеристик вебсайтів, де поєднується статичний і динамічний підходи до аналізу, що дозволяє підвищити точність класифікації.

Статичний аналіз включає дослідження HTML-коду сторінки, метаданих, доменного імені, параметрів URL та SSL-сертифікатів. Динамічний аналіз спрямований на виявлення поведінкових особливостей ресурсу, зокрема використання перенаправлень та взаємодію з користувачем. Для узагальнення результатів застосовується машинне навчання з використанням бібліотеки PHP-ML та навчальних вибірок із відкритих джерел (наприклад, PhishTank).

Запропонований метод може бути використаний для автоматизованого аналізу потенційно шкідливих вебресурсів, інтеграції в корпоративні системи кібербезпеки, моніторингу посилань у електронній пошті та месенджерах, а також у вигляді браузерних розширень для захисту користувачів у режимі реального часу.

Оцінювання ефективності методу здійснюється за такими показниками, як точність класифікації, рівень хибнопозитивних і хибнонегативних спрацьовувань, швидкодія та здатність адаптації до нових типів фішингових атак.

Реалізцію виконано мовою програмування PHP, що забезпечує простоту інтеграції у вебсередовище та сумісність із сучасними вебсерверами (Apache, Nginx). Для класифікації використовується багатошаровий перцептрон (Multilayer Perceptron, MLP), який обробляє характеристики URL, DNS-записи та SSL-індикатори.

Вибір та навчання нейронної мережі

Для реалізації механізму машинного навчання використано бібліотеку PHP-ML, що надає інструменти для побудови та навчання нейронних мереж у PHP-середовищі. Навчання здійснюється на основі відкритих даних із ресурсу PhishTank, який містить актуальні приклади (https://www.phishtank.com/developer_info.php) фішингових URL-адрес (рис.1).

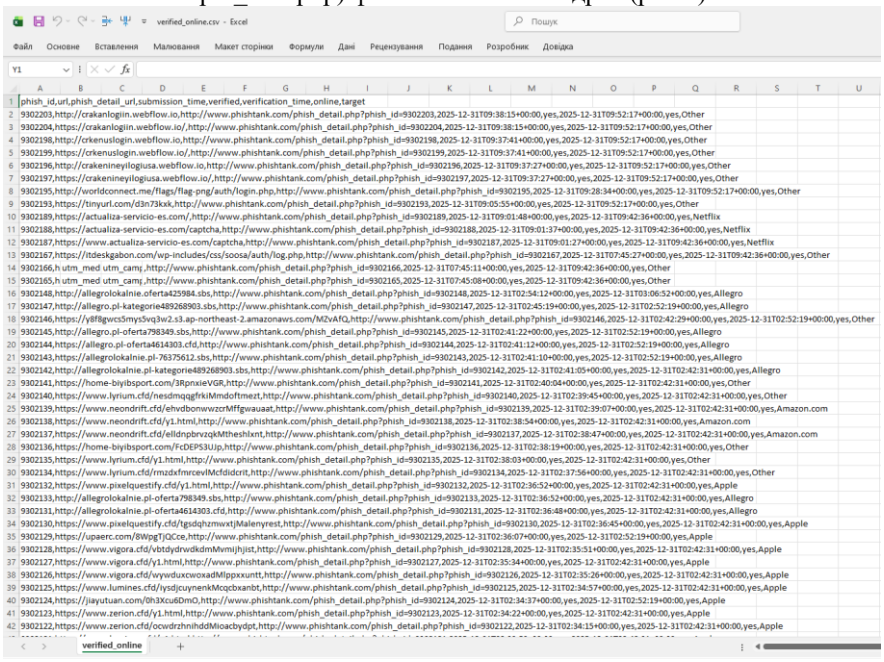

Рис. 1. Відкриті дані PhishTank для навчання нейронної мережі

Для навчання обрано багатошаровий перцептрон із алгоритмом зворотного поширення помилки (Backpropagation), який ефективно працює з нелінійними залежностями та добре підходить для задач класифікації. Використання кількох прихованих шарів дозволяє моделі виявляти складні патерни, характерні для фішингових атак.

Етапи реалізації процесу [5]:

- імпорт набору даних у форматі CSV;
- виконання попередньої обробки з метою перетворення текстових атрибутів у числові ознаки;
- формування навчальної та тестової вибірок;
- проєктування й параметричне налаштування архітектури нейронної мережі;
- навчання моделі з подальшою оцінкою її точності.

З метою досягнення балансу між точністю та швидкістю навчання застосовано оптимізацію швидкості навчання, підбір кількості нейронів і шарів, а також механізми запобігання перенавчанню, зокрема контроль продуктивності на валідаційній вибірці та регуляризацію.

Основну увагу приділено здатності мережі коректно ідентифікувати фішингові вебресурси на основі характеристик URL-адрес, наявності захищеного з'єднання та інших метаданих.

Для розв'язання задачі класифікації фішингових сайтів використано нейронну мережу типу Multilayer Perceptron (MLP), яка належить до класу багатошарових нейронних мереж і здатна моделювати складні нелінійні залежності. Архітектура MLP включає вхідний, приховані та вихідний шари, між нейронами яких формуються зважені зв'язки.

У межах дослідження було обрано мережу з трьома основними рівнями [5]: вхідним шаром із семи нейронів, що відповідають кількості вхідних ознак; двома прихованими шарами з п'ятьма та трьома нейронами відповідно; а також вихідним шаром з одним нейроном, який визначає клас сайту – фішинговий або легітимний.

Навчання нейронної мережі здійснюється за допомогою алгоритму зворотного поширення помилки (Backpropagation), який забезпечує корекцію вагових коефіцієнтів шляхом мінімізації різниці між фактичним і очікуваним результатами. Процес навчання включає обчислення похибки, оновлення ваг з урахуванням швидкості навчання та передачу помилки на попередні шари мережі для послідовного уточнення параметрів [6].

На вхід мережі подаються числові характеристики, отримані шляхом попереднього перетворення текстових полів CSV-файлу.

До них належать довжина URL-адреси, що є індикатором потенційної фішингової активності, кількість символів «/», яка відображає складність структури адреси, наявність протоколу HTTPS, а також часові характеристики, пов'язані з моментом подання ресурсу на перевірку. Додатково враховуються бінарні параметри, що відображають статус верифікації та доступності сайту, а також числове представлення цільової організації, отримане шляхом кодування першого символу відповідного поля.

Архітектура нейронної мережі.

1. Задача бінарної класифікації веб-сайтів:

- клас 0 – легітимний сайт;
- клас 1 – фішинговий сайт.

Модель реалізована у вигляді багатошарової перцептронної нейронної мережі (MLP) з одним вихідним нейроном, значення якого інтерпретується як імовірність належності сайту до фішингового класу (рис.2).

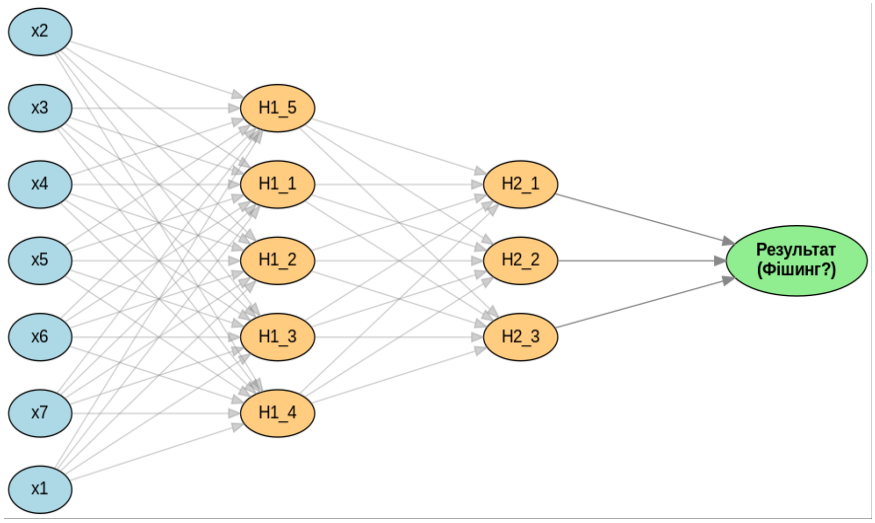

Рис. 2. Реалізація моделі

MLP складається з 4 шарів:

Шар	Кількість нейронів	Функція активації
Вхідний	7	-
Прихований 1	5	Sigmoid
Прихований 2	3	Sigmoid
Вихідний	1	Sigmoid

Вхідний вектор:

$$\mathbf{x} = (x_{1,} x_2, \dots x_7)$$

3. Ініціалізація параметрів.

Вагові коефіцієнти $w_{ij}^{(l)}$ та зсуви $b_j^{(l)}$ для всіх шарів $l \in \{1,2,3\}$ ініціалізуються малими випадковими значеннями.

Задаються гіперпараметри навчання:

- оптимізатор: Stochastic Gradient Descent (SGD);
- швидкість навчання: $\eta = 0.01$;
- кількість епох: 100;
- розмір батчу: 32 (частка навчальної вибірки);
- функція втрат: Binary Cross-Entropy.

4. Пряме поширення сигналу (Forward Propagation).

4.1. Перший прихований шар (7->5):

$$z_j^{(1)} = \sum_{i=1}^{7} w_{ij}^{(1)} x_i + b_j^{(1)}$$

$$a_j^{(1)} = \sigma(z_j^{(1)})$$

4.2. Другий прихований шар (5->3):

$$z_k^{(2)} = \sum_{j=1}^{5} w_{jk}^{(2)} a_j^{(1)} + b_k^{(2)}$$

$$a_k^{(2)} = \sigma(z_k^{(2)})$$

4.3. Вихідний шар (3->1):

$$z^3 = \sum_{k=1}^{3} w_k^{(3)} a_k^{(2)} + b^{(3)}$$

$$\hat{y} = \sigma(z^{(3)})$$

де $\hat{y} \in (0, 1)$ – прогнозована ймовірність фішингу.

5. Обчислення функції втрат.

Для кожного навчального прикладу використовується Binary Cross-Entropy:

$$L(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

де:

- $y \in \{0, 1\}$ – істинний клас;
- $\hat{y}$ – вихід мережі.

6. Зворотне поширення помилки (Backpropagation).

6.1. Вихідний шар.

Локальний градієнт:

$$\delta^{(3)} = \hat{y} - y$$

(для Sigmoid + Binary Cross-Entropy похідна спрощується).

6.2. Другий прихований шар

$$\delta_k^{(2)} = \sigma'\left(z_k^{(2)}\right) \sum \delta^{(3)} w_k^{(3)}$$

6.3. Перший прихований шар

$$\delta_j^{(1)} = \sigma'\left(z_j^{(1)}\right) \sum \delta_k^{(2)} w_{jk}^{(2)}$$

7. Оновлення ваг (SGD).

Оновлення виконується для кожного батчу:

7.1. Оновлення ваг:

$$w_{ij}^{(l)} \leftarrow w_{ij}^{(l)} - \eta \cdot a_i^{(l-1)} \delta_j^{(l)}$$

7.2. Оновлення зсувів:

$$b_{ij}^{(l)} \leftarrow b_{ij}^{(l)} - \eta \cdot \delta_j^{(l)}$$

8. Критерій класифікації.

Після завершення прямого проходу результат інтерпретується таким чином:

$$class = \begin{cases} 0, & \hat{y} < 0.5 \\ 1, & \hat{y} \geq 0.5 \end{cases}$$

9. Цикл навчання.

Навчальна вибірка розбивається на батчі розміром 32.

Для кожного батчу виконується:

- пряме поширення;
- обчислення втрат;
- backpropagation;
- оновлення ваг.

Повний прохід по всій вибірці утворює одну епоху.

Процес повторюється 100 епох.

10. Результат навчання.

Після завершення навчання модель:

- зберігає оптимізовані вагові коефіцієнти;
- використовується для класифікації нових веб-сайтів;
- повертає бінарне рішення щодо наявності фішингу.

Приховані шари виконують основну функцію обробки інформації та формування складних взаємозв'язків між ознаками. Перший прихований шар здійснює початкову трансформацію вхідних даних, використовуючи сигмоїдальну функцію активації, що забезпечує моделювання нелінійних залежностей. Другий прихований шар агрегує результати попередньої обробки та формує компактне внутрішнє подання, необхідне для побудови остаточного прогнозу [7]. Сигмоїдальна функція також застосовується для забезпечення плавного переходу між значеннями вихідної змінної в діапазоні від 0 до 1.

Вихідний шар призначений для обчислення ймовірності належності аналізованого ресурсу до класу фішингових сайтів. Значення, що близькі до одиниці, свідчать про фішингову природу ресурсу, тоді як значення, наближені до нуля, вказують на його легітимність.

Для оптимізації процесу навчання було використано такі параметри: кількість епох, що становить 100 і забезпечує стабільність навчання; швидкість навчання зі значенням 0,01, обрану для уникнення перенавчання; а також розмір батчу, встановлений на рівні 32 з метою досягнення балансу між точністю класифікації та швидкістю обчислень.

Після завершення навчання нейронна мережа демонструє здатність з високою точністю класифікувати нові вебресурси на основі сформованого набору ознак.

Аналіз ознак фішингових сайтів

Першочергову увагу приділено аналізу структури URL, оскільки адреса ресурсу є одним із перших індикаторів потенційної загрози. Фішингові сайти часто використовують надмірно довгі або складні URL-адреси, піддомени, що імітують легітимні бренди, а також ключові слова на кшталт "secure", "login", "verify", спрямовані на формування довіри користувача.

Також, фішингові вебресурси нерідко застосовують спеціалізовані методи маскування, зокрема використання доменних імен, що візуально імітують легітимні, а також реалізацію IDN-гомографічних атак, за яких символи латинського алфавіту підмінюються схожими знаками з інших писемностей (наприклад, кириличними символами у доменах англомовних ресурсів). Програмні засоби, орієнтовані на виявлення фішингових сайтів, здійснюють аналіз URL-адрес на наявність подібних аномалій, оцінюють довжину доменного імені, факт використання числових IP-адрес замість традиційних доменів, що нетипово для легітимних ресурсів, а також ідентифікують потенційні орфографічні відхилення, характерні для атак типу тайпосквотингу.

Важливим фактором є також наявність та тип SSL-сертифіката. Відсутність HTTPS-з'єднання або використання самопідписаних сертифікатів суттєво підвищує ризик фішингової природи сайту, оскільки більшість сучасних легітимних ресурсів використовують перевірені сертифікати [8].

Окремо враховується дата реєстрації домену, отримана за допомогою WHOIS-запитів. Короткий термін існування домену є характерною ознакою фішингових сайтів, які створюються для тимчасового використання. Аналіз перенаправлень дозволяє виявляти складні схеми маскування справжньої URL-адреси, що широко застосовуються у фішингових кампаніях.

Додатково реалізовано перевірку наявності ресурсу у чорних списках (DNSBL, Google Safe Browsing тощо), що дозволяє швидко ідентифікувати вже відомі загрози.

Аналіз контенту та метаданих

Важливою складовою процесу визначення фішингової природи вебресурсу є аналіз низки ключових характеристик, зокрема HTML-контенту сторінки, її метаданих та зовнішніх посилань [9]. З цією метою було

розроблено метод аналізу HTML-коду, оскільки структура вебсторінки може містити приховані індикатори фішингової активності.

Під час дослідження HTML-контенту враховується наявність форм для введення конфіденційної інформації, зокрема облікових даних користувачів, а також напрям передавання введених даних. Передавання інформації на сторонні домени або використання небезпечних способів обміну даними розглядається як ознака потенційної загрози. Окрему увагу приділено використанню тегу <iframe>, який часто застосовується для прихованого розміщення підозрілих елементів.

Аналіз зовнішніх посилань включено з огляду на те, що фішингові ресурси нерідко містять гіперпосилання на домени, не пов'язані з легітимною організацією. Невідповідність доменів, на які ведуть посилання, основному домену сайту, а також використання доменних імен, схожих на офіційні, але з незначними відмінностями, розглядаються як потенційні індикатори фішингу.

Враховано аналіз метаданих вебсторінки, зокрема заголовків і метаописів, які можуть відображати справжнє призначення ресурсу. Використання формулювань, характерних для соціальної інженерії, аналізується на предмет наявності ключових слів, типових для фішингових схем. Оцінювання ризику здійснюється з урахуванням частоти та контексту їх використання [10].

Після обробки усіх визначених параметрів сформовано модель, яка агрегує результати окремих перевірок і обчислює інтегральний показник ризику для кожного вебресурсу. Кожному критерію присвоюється відповідний ваговий коефіцієнт залежно від його значущості, на основі чого сайт класифікується як фішинговий або безпечний. Застосування комплексного підходу з використанням кількох ознак дозволяє суттєво знизити ймовірність хибнопозитивних результатів та підвищити загальну точність виявлення загроз.

Покращення та адаптивність методу

Подальше вдосконалення передбачає створення модульної архітектури, що дозволить окремо оновлювати компоненти збору даних, аналізу ознак та навчання моделі. Реалізовано механізм повторного навчання на нових даних, що забезпечує адаптацію до еволюції фішингових атак.

Поєднання класичних методів аналізу безпеки (SSL, чорні списки, аналіз URL і контенту) з нейронною мережею дозволяє створити динамічну систему, здатну виявляти як відомі, так і нові фішингові загрози з високою точністю та швидкодією.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У роботі досліджено актуальну проблему фішингових атак та розроблено систему їх автоматизованого виявлення з використанням нейронних мереж на базі бібліотеки PHP-ML. Дослідження охоплює повний життєвий цикл проєкту – від аналізу предметної області та обґрунтування необхідності застосування інтелектуальних методів до практичної реалізації програмного засобу та оцінки його економічної ефективності.

На початковому етапі виконано комплексний аналіз фішингових атак, їхніх основних типів і етапів реалізації. Розглянуто сучасні методи протидії фішингу та обґрунтовано доцільність розроблення нових підходів у зв'язку зі зростанням складності та різноманітності фішингових схем. Отримані результати сформували теоретичну основу для подальшого вдосконалення алгоритмів виявлення загроз.

Розроблено та оптимізовано засіб аналізу фішингових вебресурсів, а також підвищено його ефективність за рахунок застосування методів машинного навчання. Було обрано та налаштовано нейронні мережі, які забезпечили зростання точності класифікації потенційно небезпечних сайтів і зменшення кількості хибних спрацьовувань. Реалізовано практичну складову системи, що включає користувацький інтерфейс, базу даних для зберігання результатів аналізу та клієнтську частину для інтерактивної перевірки URL-адрес. Використання нейронних мереж дозволило впровадити удосконалений метод, здатний забезпечувати високоточний аналіз загроз у режимі реального часу. Також розроблено інструкцію з експлуатації системи, що спрощує її впровадження та використання.

Узагальнюючи отримані результати, можна стверджувати, що поставлені цілі роботи досягнуто в повному обсязі. Розроблена система автоматичного виявлення фішингових сайтів демонструє високу точність і ефективність та відповідає сучасним вимогам у сфері кібербезпеки. Проведені теоретичні та практичні дослідження, доповнені економічною оцінкою, підтверджують готовність системи до впровадження та подальшої експлуатації.

Запропоноване рішення має практичну цінність для підвищення рівня захисту користувачів від фішингових атак і характеризується значним потенціалом для подальшого розвитку та масштабування у реальних умовах функціонування систем кібербезпеки.

ISSN 2219-9365

References

1. Rittinghouse J. A Simple Guide to Understanding Internet Threats and Taking Common Sense Actions. Infinity Publishing, 2017. 168 p.

2. Atkins C. Phishing Attacks: Advanced Attack Techniques. Scotts Valley : CreateSpace, 2018. 160 p.

3. Ola A. Identity Theft: Managing Risk of Private Information Online (Safety & Security Key Books). South Carolina : CreateSpace Independent Publishing Platform, 2017. 156 p.

4. Banerjee J. S., Bhattacharyya S., Obaid A.J., Jeh W.S. Intelligent Cyber-Physical Systems Security for Industry 4.0 (Chapman & Hall/CRC Computational Intelligence and Its Applications). Second edition. London: Chapman and Hall/CRC, 2022. 284 p.

5. Kofler M., Gebeshuber K. Hacking and Security: The Comprehensive Guide to Penetration Testing and Cybersecurity (Rheinwerk Computing). Massachusetts : Rheinwerk Computing, 2023. 1141 p.

6. Goslin K., Hofmann M. Applied User Data Collection and Analysis Using JavaScript and PHP. London : Chapman and Hall/CRC, 2021. 346 p.

7. Gunikhan S. Phishing and Communication Channels: A Guide to Identifying and Mitigating Phishing Attacks. New York : Apress, 2021. 240 p.

8. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Third edition. Sebastopol : O'Reilly Media, 2022. 861 p.

9. Pallavi R. Securing the Web: Machine Learning for CSRF Vulnerability Detection: Harnessing Machine Learning Algorithms for Accurate Detection and Prevention. Saarbrucken: LAP LAMBERT Academic Publishing, 2023. 64 p.

10. Howard R. Cybersecurity First Principles: A Reboot of Strategy and Tactics. New Jersy : Wiley, 2023. 400 p.