УДК 004.056.5

Е. С. Гуменчук
А. В. Грицак

ОСНОВНІ ПІДХОДИ ДО ВИЯВЛЕННЯ ФІШИНГОВИХ ЗАГРОЗ У ВЕБОРІЄНТОВАНИХ СИСТЕМАХ

Вінницький національний технічний університет

Анотація

У роботі проаналізовано еволюцію методів виявлення фішингових атак у сучасних веборієнтованих інформаційних системах. Розглянуто обмеження класичних сигнатурних підходів та евристичного аналізу в умовах зростання поліморфних загроз. Обґрунтовано доцільність застосування методів машинного навчання, зокрема ансамблевих моделей та глибинних нейронних мереж, для детекції семантичних аномалій. Запропоновано концепцію гібридної архітектури системи захисту, що поєднує експертні правила для швидкої фільтрації та інтелектуальні алгоритми для аналізу невідомих векторів атак.

Ключові слова: фішингові атаки, машинне навчання, веборієнтовані системи, кібербезпека, штучний інтелект, NLP.

Abstract

This paper analyzes the evolution of phishing attack detection methods in modern web-oriented information systems. The limitations of classical signature-based approaches and heuristic analysis in the context of growing polymorphic threats are examined. The feasibility of using machine learning methods, particularly ensemble models and deep neural networks, for detecting semantic anomalies is substantiated. A concept of a hybrid security system architecture is proposed, combining expert rules for rapid filtering and intelligent algorithms for analyzing unknown attack vectors.

Keywords: phishing attacks, machine learning, web-oriented systems, cybersecurity, artificial intelligence, NLP.

Вступ

Фішингові загрози залишаються домінуючим вектором кібератак на веборієнтовані інформаційні системи, демонструючистійкутенденціюдозростанняскладностітаперсоналізації. За даними актуальних досліджень, зловмисники все частіше використовують технології генеративного штучного інтелекту для створення переконливого контенту (Deepfakes, text-generation), що дозволяє обходити традиційні спам-фільтри. Сучасні фішингові кампанії характеризуються мультиканальністю, охоплюючи не лише електронну пошту, а й месенджери, соціальні мережі та SMS (smishing), а також активним застосуванням методів соціальної інженерії для маніпуляції користувачами. Така еволюція загроз, зокрема поява поліморфних URL-адрес та використання легітимних хмарних сервісів для хостингу шкідливого контенту, обумовлює необхідність перегляду існуючих парадигм захисту [1].

Класичні методи виявлення фішингу

Традиційні підходи до виявлення фішингу базуються на статичному аналізі та детермінованих правилах. Сигнатурний метод, який передбачає порівняння URL-адрес та хеш-сум файлів із базами даних відомих загроз (чорними списками), залишається ефективним для блокування вже ідентифікованих атак, проте є безсилим проти атак «нульового дня» (zero-day). Евристичні методи розширюють можливості детекції шляхом аналізу типових ознак: наявності IP-адрес замість доменних імен, використання скорочувачів посилань, помилок у сертифікатах SSL/TLS та підозрілих лексичних конструкцій. Однак, ці методи часто генерують високий рівень хибнопозитивних спрацювань (False Positives) та потребують постійного оновлення правил, що знижує їх ефективність в умовах автоматизованої генерації фішингових сторінок [2].

Інтелектуальні системи на основі машинного навчання

Інтеграція методів машинного навчання (ML) та глибинного навчання (DL) дозволяє виявляти неявні патерни атак, які неможливо формалізувати простими правилами. Системи на базі алгоритмів Random Forest, Support Vector Machines (SVM) та Gradient Boosting демонструють високу точність у класифікації веб-ресурсів за сукупністю ознак (структурні особливості URL, вік домену, наявність JavaScript-обфускації). Особливу увагу привертають підходи з використанням обробки природної мови (NLP). Моделі на архітектурі трансформерів (BERT, RoBERTa) здатні аналізувати семантику тексту повідомлень, виявляючи маніпулятивні наміри та терміновість, характерні для Business Email Compromise (BEC) атак. Використання згорткових нейронних мереж (CNN) дозволяє аналізувати візуальну схожість веб-сторінок, ефективно протидіючи атакам із клонуванням інтерфейсів легітимних брендів [3].

Гібридні та інтегровані підходи

Найвищу ефективність демонструють гібридні архітектури, що поєднують швидкість класичних методів із адаптивністю штучного інтелекту. У таких системах попередній відбір здійснюється за допомогою полегшених евристик та списків репутації (Allow/Block Lists), після чого підозрілий трафік передається на глибинний аналіз нейронним мережам [4]. Критичним аспектом впровадження таких систем є реалізація механізмів пояснюваного штучного інтелекту (Explainable AI, XAI). Інструменти, такі як SHAP або LIME, дозволяють інтерпретувати рішення моделі, вказуючи на конкретні ознаки (наприклад, гомогліфи в URL або специфічні фрази), що стали причиною класифікації об’єкта як загрози. Це підвищує довіру адміністраторів безпеки до автоматизованих рішень та спрощує розслідування інцидентів [5].

Висновки

Проведений аналіз дозволяє стверджувати, що ефективна протидія сучасним фішинговим загрозам неможлива без використання інтелектуальних систем аналізу даних. Перехід від реактивних сигнатурних методів до проактивних гібридних систем дозволяє виявляти поліморфні та таргетовані атаки в режимі реального часу. Перспективним напрямом подальших досліджень є розробка полегшених моделей глибокого навчання для розгортання на стороні клієнта (в браузерних розширеннях), а також удосконалення методів захисту самих ML-моделей від змагальних атак (adversarial attacks).

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Basit, A., Zafar, M., Liu, X., Javed, A. R., Jalil, Z., & Kifayat, K. (2021). A comprehensive survey of AI-enabled phishing attacks detection techniques. *Telecommunication Systems*, *76*(1), 139-154.
2. Ayibam, J. N. (2025). Artificial Intelligence in Public Procurement: Legal Frameworks, Ethical Challenges, and Policy Solutions for Transparent and Efficient Governance. Alkebulan: A Journal of West and East African Studies, 5(2), 54-69..
3. Gupta, B. B., Arachchilage, N. A., & Psannis, K. E. (2018). Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems*, *67*(2), 247-267.
4. Saha, I., Sarma, D., Chakma, R. J., Alam, M. N., Sultana, A., & Hossain, S. (2020, August). Phishing attacks detection using deep learning approach. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1180-1185). IEEE.
5. Tang, L., & Mahmoud, Q. H. (2021). A survey of machine learning-based solutions for phishing website detection. *Machine Learning and Knowledge Extraction*, *3*(3), 672-694.

Гуменчук Едуард Сергійович – студент групи 1КІТС-24м, факультет менеджменту та інформаційної безпеки, Вінницький національний технічний університет, Вінниця, e-mail: hasesed1@gmail.com

Науковий керівник: ***Грицак Анатолій Васильович*** – кандидат технічних наук, провідний фахівець Центру ІТ та захисту інформації, факультет менеджменту та інформаційної безпеки, Вінницький національний технічний університет, Вінниця.

Eduard S. Humenchuk student of group 1KITS-24m, Faculty of Management and Information Security, Vinnytsia National Technical University, Vinnytsia.

Supervisor: ***Anatolii V. Hrytsak*** Candidate of Technical Sciences (Ph.D.), Leading Specialist of the Center for IT and Information Protection, Faculty of Management and Information Security, Vinnytsia National Technical University, Vinnytsia.