УДК 004.056:004.738.5

А. М. Наконечна
В. В. Гуменюк

ВДОСКОНАЛЕННЯ МЕХАНІЗМІВ ЗАХИСТУ КОРПОРАТИВНОЇ ЕЛЕКТРОННОЇ ПОШТИ ВІД ФІШИНГУ НА ОСНОВІ МАШИННОГО НАВЧАННЯ ТА ПРОТОКОЛІВ SPF, DKIM І DMARC

Вінницький національний технічний університет

Анотація

У роботі розглядається підхід до вдосконалення механізмів захисту корпоративної електронної пошти від фішингових атак на основі використання методів машинного навчання та протоколів автентифікації SPF, DKIM і DMARC. Проведено аналіз сучасних загроз, пов'язаних із поширенням фішингових повідомлень, що становлять одну з основних причин компрометації облікових записів і витоку конфіденційної інформації в корпоративному середовищі. Запропоновано модель виявлення підозрілих електронних листів, яка поєднує результати перевірки SPF, DKIM і DMARC із додатковими характеристиками повідомлень, зокрема параметрами заголовків, доменними ознаками та поведінковими показниками відправника. Використання алгоритмів машинного навчання дозволяє автоматизувати процес класифікації повідомлень і підвищити точність виявлення фішингових атак порівняно з традиційними методами фільтрації. Обґрунтованим результатом є зниження кількості успішних фішингових атак, підвищення рівня інформаційної безпеки корпоративних поштових систем та вдосконалення процесів моніторингу електронної кореспонденції.

Ключові слова: фішинг, корпоративна електронна пошта, машинне навчання, інформаційна безпека, кібербезпека.

Abstract

The paper considers an approach to improving mechanisms for protecting corporate email from phishing attacks based on the use of machine learning methods and the SPF, DKIM and DMARC authentication protocols. An analysis of modern threats associated with the spread of phishing messages, which are among the main causes of account compromise and leakage of confidential information in the corporate environment, is conducted. A model for detecting suspicious emails is proposed that combines the results of SPF, DKIM and DMARC checks with additional message characteristics, including header parameters, domain features and behavioral indicators of the sender. The use of machine learning algorithms makes it possible to automate message classification and increase the accuracy of phishing attack detection compared with traditional filtering methods. The expected result is a decrease in the number of successful phishing attacks, an increase in the level of information security of corporate email systems and an improvement in electronic correspondence monitoring processes.

Keywords: phishing, corporate email, machine learning, information security, cybersecurity.

Вступ

Електронна пошта залишається одним із найважливіших засобів корпоративної комунікації та обміну інформацією між працівниками, партнерами й клієнтами. Разом із зростанням обсягів електронного листування збільшується і кількість кіберзагроз, серед яких особливе місце займають фішингові атаки. Такі атаки спрямовані на отримання конфіденційної інформації, зокрема облікових даних, фінансових реквізитів або службових документів, шляхом надсилання підроблених повідомлень від імені довірених джерел. За даними міжнародних досліджень у сфері кібербезпеки, саме фішинг є одним із найпоширеніших способів початкового проникнення зловмисників до корпоративних інформаційних систем.

Традиційні механізми захисту електронної пошти, засновані на сигнатурному аналізі та статичних правилах фільтрації, не завжди здатні ефективно виявляти сучасні фішингові повідомлення. Зловмисники постійно вдосконалюють методи маскування, використовуючи підроблені домени, соціальну інженерію та автоматизовані засоби генерації контенту. Для підтвердження достовірності відправника широко застосовуються протоколи SPF (Sender Policy Framework), DKIM (DomainKeys Identified Mail) та DMARC (Domain-based Message Authentication, Reporting and Conformance), які дозволяють перевіряти джерело походження повідомлення та виявляти спроби підміни адреси відправника. Однак використання лише цих механізмів не забезпечує повного захисту від складних та цілеспрямованих фішингових кампаній.

У зв’язку з цим актуальним напрямом досліджень є інтеграція технологій машинного навчання в системи захисту корпоративної електронної пошти. Алгоритми машинного навчання здатні аналізувати великі обсяги даних, виявляти приховані закономірності та автоматично класифікувати повідомлення за рівнем загрози. Поєднання ознак автентифікації SPF, DKIM і DMARC із характеристиками вмісту листів, метаданими та поведінковими показниками дає змогу підвищити точність виявлення фішингових атак і зменшити кількість помилкових спрацьовувань. Тому розроблення та вдосконалення механізмів захисту корпоративної електронної пошти на основі методів машинного навчання є важливим завданням для підвищення рівня інформаційної безпеки сучасних організацій.

Результати досліджень

Фішингові атаки залишаються одним із найнебезпечніших видів кіберзагроз для корпоративних інформаційних систем, оскільки вони спрямовані на отримання конфіденційних даних шляхом використання методів соціальної інженерії та підробки електронних повідомлень. За даними звітів у сфері кібербезпеки, фішинг і компрометація облікових даних залишаються важливими чинниками початкового доступу зловмисників до інформаційних систем [1]. Крім того, витоки даних призводять до значних фінансових і репутаційних втрат для організацій [2]. Сучасні фішингові кампанії активно використовують підроблені домени, компрометовані поштові сервери та автоматизовані засоби генерації текстового контенту, що значно ускладнює їх виявлення традиційними механізмами фільтрації [3]. У зв’язку з цим актуальним є застосування додаткових методів аналізу повідомлень, здатних автоматично визначати ознаки потенційної загрози та адаптуватися до нових схем атак.

Важливим елементом захисту корпоративної електронної пошти є використання протоколів SPF (Sender Policy Framework), DKIM (DomainKeys Identified Mail) та DMARC (Domain-based Message Authentication, Reporting and Conformance), які забезпечують перевірку достовірності джерела повідомлення та цілісності його вмісту. Протокол SPF дозволяє перевірити, чи має поштовий сервер право надсилати повідомлення від імені певного домену [4]. Протокол DKIM забезпечує криптографічне підтвердження незмінності листа під час передавання [5]. Протокол DMARC об’єднує результати SPF і DKIM для формування політики обробки підозрілих повідомлень [6]. Незважаючи на ефективність цих механізмів, зловмисники можуть використовувати легітимні домени або скомпрометовані облікові записи, що дозволяє обходити окремі перевірки автентичності. Тому для підвищення точності виявлення фішингових атак доцільним є поєднання результатів SPF/DKIM/DMARC-автентифікації з інтелектуальними методами аналізу.

Порівняльну характеристику основних механізмів захисту корпоративної електронної пошти наведено в таблиці 1.

Таблиця 1 – Порівняльна характеристика механізмів захисту корпоративної електронної пошти від фішингу

Механізм	Основне призначення	Переваги	Обмеження
SPF	Перевіряє право поштового сервера надсилати листи від імені домену.	Просте налаштування; виявляє підміну домену.	Не аналізує зміст листа; має обмеження при пересиланні.
DKIM	Підтверджує цілісність листа та домен відправника цифровим підписом.	Захищає від зміни повідомлення; підвищує довіру до домену.	Не визначає фішинговий зміст; залежить від коректних DNS-записів.
DMARC	Узгоджує результати SPF і DKIM та задає політику обробки підозрілих листів.	Забезпечує звітність; зменшує ризик доменного спуфінгу.	Не блокує фішинг із легітимних або скомпрометованих акаунтів.
Традиційна фільтрація	Використовує сигнатури, чорні списки та статичні правила.	Ефективна для відомих загроз; швидко працює.	Слабко адаптується до нових атак; можливі хибні спрацювання.
Машинне навчання	Класифікує листи за технічними, текстовими та поведінковими ознаками.	Виявляє складні шаблони; може оновлюватися на нових даних.	Потребує навчальних даних і контролю якості моделі.
Інтегрований підхід	Поєднує SPF/DKIM/DMARC із ML-аналізом електронних листів.	Підвищує точність виявлення та якість моніторингу.	Складніше впровадження й налаштування в корпоративній інфраструктурі.

Отже, найбільш ефективним для корпоративного середовища є інтегрований підхід, який поєднує перевірку автентичності відправника з інтелектуальним аналізом змісту та поведінкових ознак електронного повідомлення.

Запропонований підхід передбачає використання алгоритмів машинного навчання для класифікації електронних повідомлень на основі сукупності технічних і поведінкових характеристик. До набору

ознак можуть входити результати SPF, DKIM та DMARC-перевірок, параметри заголовків повідомлень, репутація домену відправника, частота використання певних ключових слів, наявність підозрілих URL-адрес, а також статистичні характеристики текстового вмісту [7]. Для навчання моделі можуть застосовуватися алгоритми Random Forest, Gradient Boosting, Support Vector Machine або нейронні мережі, які демонструють високі показники точності під час аналізу великих обсягів поштового трафіку [8]. Отримані результати класифікації можуть використовуватися для автоматичного блокування небезпечних повідомлень або їх направлення на додаткову перевірку.

Додатковою перевагою використання машинного навчання є можливість постійного оновлення моделі на основі нових даних про виявлені інциденти. Це дозволяє адаптувати систему до змін у тактиці зловмисників та зменшувати кількість помилкових спрацьовувань порівняно з традиційними правилами фільтрації [8]. Інтеграція такого підходу в корпоративну поштову інфраструктуру сприяє підвищенню рівня захисту інформаційних ресурсів організації, зменшенню ризику компрометації облікових записів користувачів та вдосконаленню процесів моніторингу електронної кореспонденції.

У результаті проведеного дослідження встановлено, що використання лише стандартних механізмів автентифікації електронної пошти SPF, DKIM та DMARC забезпечує ефективне виявлення випадків підміни домену відправника, проте не дозволяє повною мірою протидіяти сучасним фішинговим кампаніям, які використовують легітимні або скомпрометовані облікові записи. Аналіз існуючих підходів показав, що значна частина шкідливих повідомлень проходить базові перевірки автентичності, що потребує застосування додаткових механізмів інтелектуального аналізу електронної кореспонденції.

Дослідження продемонструвало доцільність використання методів машинного навчання для автоматизованого виявлення фішингових повідомлень на основі комплексного аналізу технічних та поведінкових ознак. До найбільш інформативних характеристик належать результати SPF/DKIM/DMARC-перевірок, особливості заголовків електронних листів, репутація домену відправника, наявність підозрілих URL-посилань, структура текстового вмісту та поведінкові параметри взаємодії користувачів із повідомленнями. Використання алгоритмів класифікації дозволяє виявляти приховані закономірності, характерні для фішингових атак, та своєчасно ідентифікувати потенційні загрози навіть у випадках, коли повідомлення успішно проходять стандартні перевірки автентичності.

Отримані результати підтверджують, що інтеграція механізмів машинного навчання з протоколами SPF, DKIM і DMARC сприяє підвищенню ефективності захисту корпоративної електронної пошти, зменшенню кількості помилково пропущених фішингових повідомлень і підвищенню рівня інформаційної безпеки організації. Запропонований підхід забезпечує можливість адаптації до нових типів атак завдяки регулярному оновленню моделей на основі накопичених даних, що робить систему більш стійкою до змін у тактиці кіберзлочинців та перспективною для впровадження в сучасних корпоративних мережах.

Висновки

У результаті проведеного дослідження встановлено, що фішингові атаки залишаються однією з найбільш поширених і небезпечних загроз для корпоративних поштових систем, оскільки можуть призводити до компрометації облікових записів користувачів, несанкціонованого доступу до інформаційних ресурсів та витоку конфіденційних даних. Аналіз сучасних методів захисту показав, що протоколи SPF, DKIM і DMARC ефективно забезпечують перевірку автентичності відправника та цілісності повідомлень, однак самостійно не гарантують виявлення всіх видів фішингових атак.

Обґрунтовано доцільність поєднання механізмів SPF/DKIM/DMARC-автентифікації з методами машинного навчання для комплексного аналізу електронних повідомлень. Запропонований підхід дозволяє враховувати не лише результати перевірки поштових протоколів, а й додаткові характеристики листів, зокрема особливості заголовків, доменні ознаки, текстовий вміст та поведінкові параметри відправника. Це забезпечує більш точну класифікацію повідомлень та підвищує ефективність виявлення потенційно небезпечної кореспонденції.

Отримані результати свідчать про перспективність використання інтелектуальних методів аналізу в системах корпоративної електронної пошти. Інтеграція алгоритмів машинного навчання з існуючими засобами поштової автентифікації сприяє зменшенню кількості успішних фішингових атак, підвищенню рівня інформаційної безпеки організації та вдосконаленню процесів моніторингу електронної кореспонденції. Запропонований підхід може бути використаний як основа для створення адаптивних систем захисту, здатних ефективно протидіяти сучасним кіберзагрозам.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Verizon. 2025 Data Breach Investigations Report (DBIR). URL: https://www.verizon.com/business/resources/reports/dbir/ (дата звернення: 19.05.2026).

2. IBM Security. Cost of a Data Breach Report 2025. URL: https://www.ibm.com/reports/data-breach (дата звернення: 19.05.2026).

3. ENISA. ENISA Threat Landscape 2025. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025 (дата звернення: 19.05.2026).

4. Kitterman S. Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1. RFC 7208. URL: https://datatracker.ietf.org/doc/html/rfc7208 (дата звернення: 19.05.2026).

5. Crocker D., Hansen T., Kucherawy M. DomainKeys Identified Mail (DKIM) Signatures. RFC 6376. URL: https://datatracker.ietf.org/doc/html/rfc6376 (дата звернення: 19.05.2026).

6. Kucherawy M., Zwicky E. Domain-based Message Authentication, Reporting, and Conformance (DMARC). RFC 7489. URL: https://datatracker.ietf.org/doc/html/rfc7489 (дата звернення: 19.05.2026).

7. Salloum S., Gaber T., Vadera S., Shaalan K. Phishing Email Detection Using Natural Language Processing Techniques: A Literature Survey. Procedia Computer Science. 2021. Vol. 189. P. 19–28. DOI: https://doi.org/10.1016/j.procs.2021.05.077.

8. Kyaw P. H., Thinn T. T., Tun W. T., Maw A. H. A Systematic Review of Deep Learning Techniques for Phishing Email Detection. Electronics. 2024. Vol. 13, No. 19. Article 3823. DOI: https://doi.org/10.3390/electronics13193823.

Наконечна Аліна Максимівна - студентка групи 1КІТС-22б, Факультет менеджменту та інформаційної безпеки, Вінницький національний технічний університет, м. Вінниця, e-mail: a66812210@gmail.com

Науковий керівник: Гуменюк Вячеслав Володимирович - асистент кафедри менеджменту та безпеки інформаційних систем, Вінницький національний технічний університет, м. Вінниця, e-mail: hvv@vntu.edu.ua.

Nakonechna Alina Maksymivna - student of group 1KITS-22b, Faculty of Management and Information Security, Vinnytsia National Technical University, Vinnytsia, e-mail: a66812210@gmail.com

Scientific supervisor: Humeniuk Viacheslav Volodymyrovych - Assistant of the Department of Management and Security of Information Systems, Vinnytsia National Technical University, Vinnytsia, e-mail: hvv@vntu.edu.ua.