

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ
ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ
ІНФОРМАЦІЇ

КАФЕДРА СИСТЕМ ТА ТЕХНОЛОГІЙ КІБЕРБЕЗПЕКИ

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«Технологія виявлення фішингових атак в електронній пошті або на
інформаційних ресурсах організації»
зі спеціальності 125 Кібербезпека та захист інформації

(код, найменування спеціальності)

освітньо-професійної програми Інформаційна та кібернетична безпека
(назва програми)

*Кваліфікаційна робота містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело*

Ігор ГОРОДЕЦЬКИЙ
(підпис)

Виконав: здобувач вищої освіти групи БСДМ-63
ГОРОДЕЦЬКИЙ Ігор
(прізвище, ім'я)

Керівник д.т.н., професор КАЗМІРЧУК Світлана
(науковий ступінь, вчене звання, прізвище, ім'я)

Рецензент
(науковий ступінь, вчене звання, прізвище, ім'я)

Київ 2025

ЗМІСТ

	Стор.
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	9
ВСТУП	10
1 ДОСЛІДЖЕННЯ ПРОБЛЕМИ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ЕЛЕКТРОННІЙ ПОШТІ ТА НА ІНФОРМАЦІЙНИХ РЕСУРСАХ ОРГАНІЗАЦІЇ.....	12
1.1 Дослідження проблеми фішингових атак у сучасних інформаційних системах.....	12
1.2 Аналіз підходів до виявлення та протидії фішинговим атакам.....	22
1.3 Аналіз існуючих рішень і технологій захисту від фішингових атак.....	32
2 АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ІНФОРМАЦІЙНИХ СИСТЕМАХ.....	40
2.1 Класифікація фішингових атак і вектори їх реалізації.....	40
2.2 Основні принципи побудови систем виявлення фішингових повідомлень.....	47
2.3 Огляд методів аналізу електронних листів і вебресурсів для визначення фішингу.....	53
2.4 Порівняння ефективності існуючих інструментів виявлення фішингових атак.....	56
3 ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ЕЛЕКТРОННІЙ ПОШТІ ТА НА ІНФОРМАЦІЙНИХ РЕСУРСАХ ОРГАНІЗАЦІЇ.....	68
3.1 Розробка архітектури технології виявлення фішингових атак.....	68
3.2 Реалізація механізмів аналізу електронних листів і перевірки URL-адрес.....	73
3.3 Тестування технології та оцінювання її	79
3.4 ефективності.....	
Рекомендації щодо впровадження технології в корпоративне середовище.....	83

ВИСНОВКИ	86
ПЕРЕЛІК ПОСИЛАНЬ	88
ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ (Презентація)	91

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

AI	– Artificial Intelligence
API	– Application Programming Interface
BEC	– Business Email Compromise
DKIM	– DomainKeys Identified Mail
DMARC	– Domain-based Message Authentication, Reporting and Conformance
MFA	– Multi-Factor Authentication
ML	– Machine Learning
NLP	– Natural Language Processing
SEG	– Secure Email Gateway
SIEM	– Security Information and Event Management
SOC	– Security Operations Center
SPF	– Sender Policy Framework
TI	– Threat Intelligence
UEBA	– User and Entity Behavior Analytics
URL	– Uniform Resource Locator
ПК	– Персональний комп'ютер
ШПЗ	– Шкідливе програмне забезпечення

ВСТУП

Актуальність дослідження. Фішингові атаки залишаються одним із найпоширеніших шляхів компрометації облікових записів і доступу до внутрішніх ресурсів організацій. Вони маскуються під легітимні листи та сторінки, використовують соціальну інженерію, підроблені домени, редиректи та шкідливі вкладення/посилання.

З розвитком віддаленої роботи, хмарних сервісів і мобільних клієнтів традиційні периметральні засоби вже не гарантують належного захисту. Тому потрібна технологія виявлення фішингу, що поєднує перевірку автентичності пошти (SPF, DKIM, DMARC), аналіз URL/доменів, контент- та заголовкові ознаки листів, репутаційні показники й політики обробки інцидентів на поштових і веб-ресурсах організації.

Об'єкт дослідження – процес виявлення та аналізу фішингових атак в електронній пошті та на інформаційних ресурсах організації.

Предмет дослідження – технологія виявлення фішингових атак на основі перевірки доменів і джерел надсилання, аналізу посилань і вмісту повідомлень, репутаційних і технічних ознак, а також організаційних політик реагування.

Мета роботи – розробити порядок застосування технології виявлення фішингових атак у корпоративному середовищі та підготувати рекомендації щодо її ефективного впровадження й супроводу.

Наукові завдання:

дослідити сутність проблеми фішингових атак та їх вплив на безпеку електронної пошти і вебресурсів;

проаналізувати підходи до виявлення й запобігання фішингу (перевірка SPF/DKIM/DMARC, фільтрація URL, DNS-фільтрація, SEG/шлюзи безпеки, sandboxing вкладень, user awareness);

оцінити існуючі рішення та практики застосування політик антифішингу в організаціях;

систематизувати методи й засоби побудови технології виявлення фішингових листів і підроблених вебсторінок;

описати порядок реалізації технології: збір і валідація технічних ознак, правила класифікації, дії за результатами, моніторинг та реагування.

Методи дослідження – опрацювання наукової та технічної літератури, аналіз експлуатаційної та нормативної документації, міжнародних стандартів і рекомендацій у сфері електронної пошти та кіберзахисту; порівняльний аналіз інструментів (поштові шлюзи безпеки, DNS/URL-фільтрація, механізми автентифікації пошти); узагальнення практик впровадження політик антифішингу.

Практичне значення одержаних результатів: сформовано порядок застосування технології виявлення фішингових атак для організації, включно з вимогами до налаштування SPF/DKIM/DMARC, політиками перевірки та блокування підозрілих URL, обробкою вкладень, інтеграцією з поштовим шлюзом/веб-проксі, сценаріями реагування (playbooks) і рекомендаціями фахівцям з кібербезпеки щодо впровадження, супроводу та оцінювання ефективності (метрики: частка виявлених фішингових листів, час реагування, зниження кліків по шкідливих посиланнях).

Результати кваліфікаційної роботи апробовані на Всеукраїнській науковій конференції «Актуальні проблеми кібербезпеки», яка відбулася 29 жовтня 2025 року в Державному університеті інформаційно-комунікаційних технологій, м. Київ.

1 ДОСЛІДЖЕННЯ ПРОБЛЕМИ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ЕЛЕКТРОННІЙ ПОШТІ ТА НА ІНФОРМАЦІЙНИХ РЕСУРСАХ ОРГАНІЗАЦІЇ

1.1. Дослідження проблеми фішингових атак у сучасних інформаційних системах

Фішингові атаки залишаються однією з найпоширеніших та найнебезпечніших кіберзагроз для сучасних інформаційних систем. Фішинг можна визначити як вид шахрайства, що поєднує соціальну інженерію та технічні прийоми введення в оману з метою викрадення конфіденційних даних (наприклад, паролів чи фінансової інформації) у користувачів. Зловмисники при цьому маскуються під довірених осіб або організації: надсилають правдоподібні електронні листи, створюють підроблені веб-сайти та доменні імена, аби змусити користувачів повірити, що вони мають справу з легітимною стороною.

Соціальна інженерія в таких атаках використовує психологічний тиск і обман, тоді як технічні засоби фішингу можуть включати шкідливе програмне забезпечення чи перенаправлення на фальшиві веб-сторінки для перехоплення облікових даних. У сучасну цифрову епоху фішингові атаки постійно еволюціонують та адаптуються, знаходячи все нові способи обходити традиційні засоби захисту.

Як відзначає Анти-Фішингова робоча група (APWG), ці атаки є надзвичайно поширеними і адаптивними, а їх розвиток тісно пов'язаний із розширенням цифрових комунікацій та сервісів (рис. 1.1). Тобто, зі зростанням обсягів електронної взаємодії та різноманітності каналів доставки (електронна пошта, веб-сайти, месенджери тощо) фішингові техніки неминуче удосконалюються та урізноманітнюються. Це створює серйозну проблему для організацій, адже їхні співробітники та інформаційні ресурси постійно перебувають під прицілом таких атак [4].



Рис. 1.1. Частка фішингових атак серед усіх кіберзлочинів та їх зростання [7]

Глобальні тренди також демонструють зростання активності фішерів: за даними APWG [4], тільки за 1-й квартал 2025 року було зафіксовано понад 1,003,924 фішингових атак, що є рекордним показником з кінця 2023 року. Хоча частина фішингових листів блокується автоматизованими фільтрами (наприклад, платформа Gmail щодня відсіює понад 100 млн підозрілих повідомлень), значна кількість все ж просочується до скриньок користувачів. У результаті фішинг залишається провідною причиною багатьох інцидентів безпеки.

Отже, сукупно фішинг прямо або опосередковано причетний до значної частки успішних атак на організації.

Фінансові втрати від фішингових атак вимірюються в мільярдах доларів. Особливо небезпечним різновидом є так звані атаки типу Business Email Compromise (BEC) – компрометація ділової електронної пошти (рис. 1.2). У таких схемах шахраї видають себе за керівників компаній або надійних

партнерів і шляхом переконливого листування схиляють жертв (співробітників фінансових відділів, менеджерів тощо) до несанкціонованих переказів коштів чи розголошення цінної інформації. ВЕС-атаки часто не містять явних шкідливих посилань або вірусів, через що їх особливо складно виявити автоматично. У 2023 році було зареєстровано близько 21,500 скарг на ВЕС із загальними збитками понад \$2.9 млрд – це найвитратніший вид кібершахрайства на сьогодні. Для порівняння, прямі втрати від класичних фішингових схем (що включають обман шляхом листів і повідомлень) оцінювались "лише" у десятки мільйонів, тоді як середня ВЕС-афера завдає значно більшої шкоди одній жертві [1].

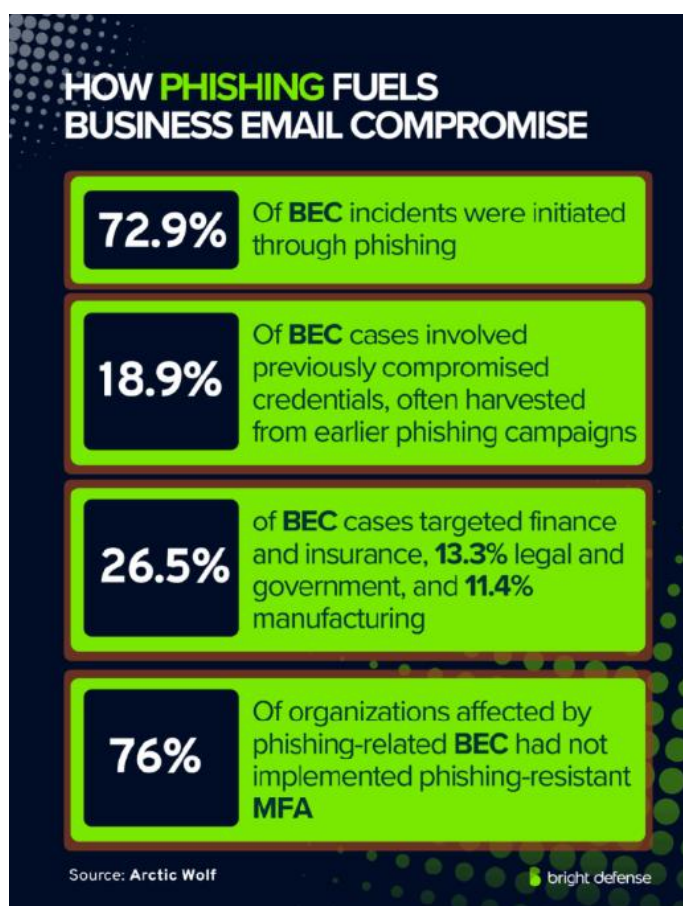


Рис. 1.2. Взаємозв'язок фішингу та компрометації ділової пошти (ВЕС) [7]

Зростає і інтенсивність таких цільових атак: корпорація Microsoft повідомила, що за період з квітня 2022 по квітень 2023 року її служби кіберрозвідки зафіксували 35 мільйонів спроб ВЕС-атак, у середньому ~156 тисяч на добу [18]. Ця статистика вражає і підкреслює глобальний характер

проблеми. До того ж, методи здійснення ВЕС стають все більш масовими – існують навіть підпільні «Phishing-as-a-Service» платформи, що за плату надають злочинцям шаблони листів, хостинг для фішингових сайтів і автоматизацію розсилки. Це дозволяє навіть відносно низькокваліфікованим зловмисникам запускати масштабні кампанії проти бізнесу.

Найтипівішим каналом фішингу залишається електронна пошта. Здебільшого атака починається з того, що жертві надходить правдоподібний лист: наприклад, повідомлення від імені банку, популярного сервісу чи колеги по роботі. У листі може бути посилання на веб-сторінку для «перевірки облікового запису» або вкладений файл (рахунок, резюме тощо) з прихованим шкідливим кодом. Клацнувши на таке посилання, користувач переходить на фальшивий веб-сайт, який візуально копіює легітимний ресурс (банкінг, корпоративний портал, поштовий сервіс тощо). Вводячи там свої логін і пароль, жертва фактично передає їх зловмисникам, а сама сторінка може видати повідомлення про помилку чи зробити непомітне перенаправлення на справжній сайт (рис. 1.3).

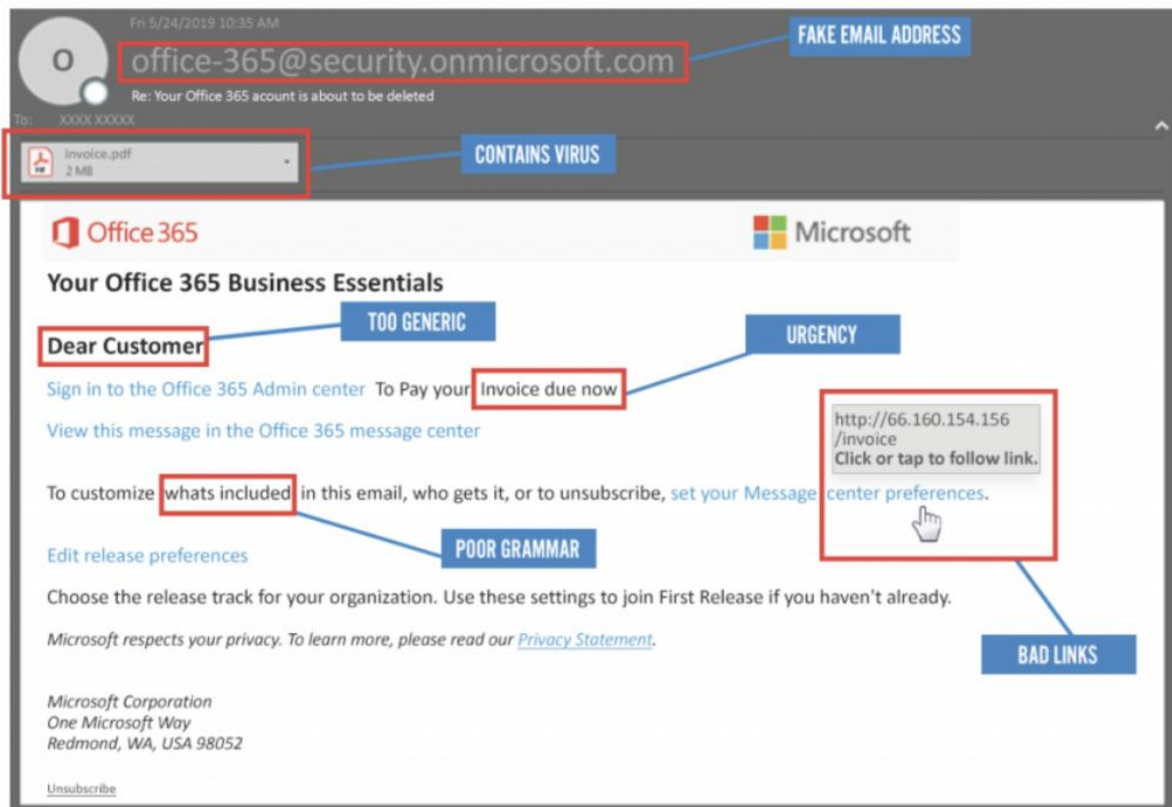


Рис. 1.3. Приклад фішингового листа

В іншому сценарії, відкриття вкладеного файлу призводить до непомітного встановлення малварі (троянця, кейлоггера) на комп'ютер жертви. Такий шкідливий софт може перехоплювати введені паролі або відкривати «бекдор» для подальшого проникнення до системи.

Згідно зі статистикою [8], фішинг став провідним методом початкового проникнення для програм-вимагачів (ransomware). Це означає, що успішна фішингова розсилка може мати наслідком не лише крадіжку паролів, а й повноцінну атаку на інфраструктуру організації.

Окрім масових фішинг-розсилок, що спрямовані на широку аудиторію, все більшого поширення набувають цільові фішингові атаки. Spear phishing (цільовий фішинг) – це атаки, спеціально підготовані для конкретної особи або групи осіб. Зловмисники ретельно вивчають профілі жертв (через соцмережі, публічні джерела, зламану переписку) і складають максимально переконливе повідомлення, враховуючи посаду, коло знайомств та поточний контекст. Наприклад, для фінансового директора можуть надіслати лист нібито від генерального директора з терміновим проханням оплатити рахунок, або для співробітника ІТ – повідомлення від служби підтримки про необхідність негайно змінити пароль. Такий індивідуалізований підхід суттєво підвищує шанс успіху атаки.

За даними Європейського агентства з кібербезпеки (ENISA) [9], державні та військові хакерські групи також активно застосовують спір-фішинг: зокрема, у 2022–2023 рр. спостерігалось зростання кібершпигунських кампаній, що націлені на політиків, урядовців, журналістів і активістів саме через методи фішингу (включно з атаками через соціальні мережі). Таким чином, і кіберзлочинці, і державні кібероператори використовують фішинг як універсальний інструмент для отримання несанкціонованого доступу.

До окремої категорії цільових фішингових загроз відноситься вже згаданий BEC – коли атака спрямована на обман співробітників з метою фінансової вигоди. Також варто згадати «китовий» фішинг (whaling) – різновид spear phishing, націлений на керівників вищої ланки (CEO, CFO) та осіб, що

приймають рішення. У таких випадках зловмисники можуть вивчати публічну активність топ-менеджменту й імітувати ділову комунікацію або екстрені запити від імені VIP-персон, щоб обійти стандартні процедури безпеки.

Крім того, фішингові атаки можуть здійснюватися і через інші канали зв'язку. Смішинг (SMiShing) – це фішинг через SMS або месенджери, коли користувач отримує текстове повідомлення з посиланням (наприклад, «Ваш рахунок заблоковано, перейдіть за посиланням для підтвердження даних»).

Вішинг (vishing) – шахрайство через телефонний дзвінок, де зловмисник, представляючись, скажімо, співробітником банку, виманює конфіденційну інформацію. Хоча основний фокус захисту в організаціях робиться на електронну пошту та веб-трафік, інциденти smishing/vishing теж почастишали, особливо в умовах віддаленої роботи, коли складніше верифікувати особу співрозмовника. Усі ці варіанти мають спільну рису: вони експлуатують людську довірливість і помилки, обходячи традиційні технічні бар'єри кібербезпеки.

Зловмисники постійно вдосконалюють фішингові атаки, застосовуючи як перевірені, так і новаторські підходи. Серед актуальних трендів останніх років можна відзначити такі:

Зловживання брендами і сервісами. Фішингові листи часто маскуються під повідомлення від відомих компаній (банки, поштові служби, ІТ-гіганти). Наприклад, бренд Microsoft вже кілька років поспіль очолює рейтинг найбільш імітованих у фішингових кампаніях у світі. Зловмисники використовують фірмовий стиль і схожі адреси (наприклад, `security-microsoft.com` замість справжнього `microsoft.com`), щоб обдурити отримувача. Лише у 2022 році було зафіксовано понад 30 млн шкідливих повідомлень, що використовували назву чи дизайн Microsoft. Це вказує на масштаб проблеми довіри до брендів: користувачі часто автоматично довіряють листам, якщо бачать знайомий логотип чи тему.

Фішингові сайти на базі легітимних ресурсів. Зростає частка фішингових сторінок, які розміщені не на новостворених доменах, а на скомпрометованих легальних веб-сайтах або хмарних сервісах. Злочинці знаходять вразливі веб-

ресурси компаній та таємно розміщують там фішингові форми. Таким чином, URL такої сторінки може бути на вигляд цілком легітимним (наприклад, домен відомого благодійного фонду), і стандартні фільтри не завжди її блокують. Інший підхід – використання популярних платформ з сертифікатами (Google Sites, Dropbox, Azure, Notion тощо) для хостингу шахрайських сторінок, оскільки трафік до відомих доменів важче відрізнити від легального. Це серйозний виклик для систем виявлення фішингу, бо задіяні ресурси з високим рівнем довіри.

Обхід засобів автентифікації. Багато організацій впроваджують двофакторну аутентифікацію (2FA/MFA), щоб ускладнити використання викрадених паролів. У відповідь зловмисники розробили техніки для обходу і цих бар'єрів. Один зі свіжих прикладів – атаки типу Adversary-in-the-Middle (AiTM), коли шахрайський сайт працює проксі-сервером між жертвою та справжнім сайтом. Користувач вводить OTP-код або підтверджує push-аутентифікацію, і ця інформація в режимі реального часу перехоплюється, дозволяючи зловмиснику увійти в обліковий запис до того, як сесія завершиться (рис. 1.4).



Рис. 1.4. Приклад перехоплення MFA-токену через фішинговий сайт [22]

Такі «проксі-фішинг» комплекти вже виявлялися у диких атаках, зокрема проти користувачів служб Office 365 і Google Workspace.

Інша техніка – «бомбардування» push-повідомленнями (prompt bombing): якщо акаунт захищено мобільним підтвердженням, нападник багаторазово ініціює запит MFA, поки стомлений користувач не підтвердить його помилково. Серед інших методів обходу – крадіжка сесійних токенів (через інфікування комп'ютера або вставлення шкідливого скрипта в браузер). У звіті Verizon 2025 підкреслено, що випадки обходу багатофакторної автентифікації значно почастишали і стали однією з вагомих причин успішних зламів корпоративних мереж.

Використання QR-кодів та нестандартних носіїв. Щоб протистояти вдосконаленим антифішинговим фільтрам, злочинці почали застосовувати нові трюки для доставки фішингового контенту. Один з них – вставляння в листи зображень із QR-кодом, який отримувач має відсканувати смартфоном. APWG зафіксувала масовану розсилку саме таких листів: «злочинці щодня розсилають мільйони електронних листів із QR-кодами, що ведуть жертв на фішингові сайти або завантажують шкідливе ПЗ». Цей прийом розрахований на те, що захисні системи пошти не проаналізують вміст зображення, а користувач, скануючи код телефоном, може не помітити підступу. Інший подібний прийом – вкладати фішингове посилання у файли типу PDF або DOC під виглядом документа. Ідея в тому, що вкладення пройде початкову перевірку, а вже при відкритті вмісту користувач клацне на посилання в документі.

Зловживання штучним інтелектом. Розвиток генеративного AI дав у руки зловмисникам нові інструменти для масштабування фішингу. Сучасні моделі на кшталт ChatGPT можуть генерувати грамотно сформульовані, переконливі тексти листів різними мовами, що зменшує кількість очевидних помилок і ознак шахрайства. ENISA у своєму звіті за 2023 [9] рік відзначає безпрецедентне занепокоєння щодо можливого використання deepfake-технологій та чат-ботів зі штучним інтелектом у фішингових і соціоінженерних

атаках. Наприклад, вже зафіксовані випадки, коли зловмисники застосовували «діпфейк»-відео або аудіо голосів керівників, щоб віддавати підроблені накази підлеглим електронною поштою або телефоном. Такі реалістичні підробки суттєво ускладнюють розпізнавання обману. Хоча масового використання AI у фішингу поки що небагато (за оцінками, тільки 0.7–5% виявлених фішингових листів створені за допомогою мовних моделей), експерти прогнозують зростання цієї частки. Організації змушені будуть впроваджувати інструменти для виявлення AI-генерованого контенту та навчати персонал новим методам перевірки достовірності повідомлень.

Проблеми виявлення та протидії фішинговим атакам. Боротьба з фішингом є складним багатоаспектним завданням, оскільки ця загроза лежить на стику технологій та людського фактору. На технічному рівні існує низка рішень для зниження ризиків фішингу. По-перше, це впровадження стандартів захисту електронної пошти, таких як SPF, DKIM та DMARC. Вони дозволяють перевіряти автентичність домену відправника листа і таким чином відхиляти підроблені повідомлення (спуфінг). Міжнародні організації та регулятори останніми роками активно закликають до запровадження DMARC. Так, починаючи з лютого 2024 року, такі гіганти як Google та Yahoo! вимагають від усіх масових розсилальників наявності правильно налаштованого DMARC-домену [19]. Цей крок уже дав результат – за даними компанії Valimail, лише за перші місяці 2024-го понад 500 тисяч доменів зі списку топ-10 мільйонів почали публікувати DMARC-записи для захисту від спуфінгу.

Проте варто зазначити, що багато з них встановили політику `p=none` (тобто фіксують звіти про порушення, але не блокують підроблені листи), що поки не гарантує повного захисту. Очікується, що вимоги будуть посилюватися, і надалі мінімально прийнятним рівнем стане `p=quarantine` або `reject`, коли неавторизовані листи від імені домену будуть відхилятися поштовими серверами.

По-друге, повсюдно рекомендується впровадження багатфакторної автентифікації для доступу до критичних систем. В ідеалі це мають бути методи, стійкі до фішингу (наприклад, апаратні токени, FIDO2-ключі, біометрія

за стандартами WebAuthn), оскільки одноразові SMS-коди чи OTP зі застосунків можуть бути перехоплені. Урядові стандарти переходять на такий підхід: Національний інститут стандартів і технологій США (NIST) зобов'язав федеральні установи використовувати засоби автентифікації, стійкі до фішингу, для захисту доступу співробітників до систем [20]. Це означає відмову від традиційних паролів і OTP на користь, наприклад, смарт-карт або криптографічних ключів, які неможливо підробити через електронний лист. Хоча впровадження подібних технологій потребує часу і ресурсів, багато великих компаній теж рухаються у цьому напрямку, щоб зменшити залежність від людського фактора.

По-третє, технічні засоби фільтрації контенту невпинно удосконалюються. Сучасні поштові шлюзи та системи захисту кінцевих точок використовують штучний інтелект та машинне навчання для аналізу вхідних повідомлень. Вони перевіряють сотні ознак: текст листа на наявність фішингових шаблонів, збіг дизайну з відомими брендами, репутацію посилань і вкладень, характерні помилки або надто загальні звертання. Такі системи можуть виявляти й нові, раніше невідомі загрози (zero-day фішинг) на основі аномалій або схожості з уже баченими атаками.

Наукові дослідження 2023–2024 років [12, 13] фокусуються на вдосконаленні алгоритмів класифікації фішингових URL і email, використанні глибоких нейронних мереж, методів NLP для аналізу тексту листів, а також Explainable AI для інтерпретації рішень моделей. Втім, гонка озброєнь триває: зловмисники також застосовують прийоми для обходу AI-фільтрів (наприклад, вставляють випадковий шум у текст, щоб знизити схожість зі знаними зразками, або генерують унікальні домени на кожен лист).

Нарешті, одним із ключових компонентів захисту залишається навчання та підвищення обізнаності користувачів. Більшість організацій регулярно проводять тренінги з кібербезпеки, симуляції фішингових атак для співробітників, розсилання пам'яток про ознаки підозрілих листів тощо. Мета – сформувати у персоналу настороженість та навички перевірки перед тим, як клікати на посилання чи відкривати файли. Проте, як показує практика, одного

лише навчання недостатньо. Згідно з аналізом Verizon, показник «клікабельності» фішингових посилань залишається високим і тренінги не мали статистично значущого впливу на зменшення частоти помилок користувачів. Інакше кажучи, навіть обізнані працівники в умовах стресу або поспіху можуть стати жертвами добре продуманої атаки. Тому важливо поєднувати методи: безперервно навчати співробітників і одночасно впроваджувати технічні рішення, що підстрахують людський фактор. Відомий принцип кібербезпеки говорить: «Люди – остання лінія захисту». У випадку фішингу цю лінію потрібно укріплювати з обох сторін – знаннями користувачів та інноваційними технологіями виявлення.

Фішингові атаки у сучасних інформаційних системах становлять надзвичайно серйозну проблему, яка продовжує загострюватися. Зловмисники активно експлуатують довіру користувачів і недосконалість цифрових механізмів аутентифікації, завдаючи значних збитків компаніям по всьому світу. Обсяг фішингових інцидентів зростає, атаки стають більш таргетованими та технічно складними, що ускладнює їх своєчасне виявлення. Міжнародна спільнота реагує на виклик: приймаються стандарти і рекомендації (DMARC, стійка до фішингу автентифікація), розробляються все більш розумні системи фільтрації, поширюються програми підготовки користувачів. Однак повністю нейтралізувати цю загрозу поки не вдається – фішинг адаптується до заходів протидії та знаходить нові вектори.

Отже, проблема виявлення фішингових атак залишається актуальною та потребує подальших досліджень і впровадження комплексних рішень, що поєднують передові технології та людську пильність. Це завдання є ключовим кроком до підвищення кіберстійкості організацій у цифрову епоху.

1.2. Аналіз підходів до виявлення та протидії фішинговим атакам

Фішингові атаки через електронну пошту залишаються одним з основних векторів початкового проникнення зловмисників в організації. Це зумовлює необхідність комплексного підходу до їх виявлення та нейтралізації. Сучасні підходи

до протидії фішингу умовно поділяють на технічні, організаційні та поведінкові. У цьому підрозділі проаналізовано кожну з цих груп, виокремлено актуальні методи (з акцентом на фішинг в електронній пошті) – від рішень на основі штучного інтелекту до навчальних симуляцій та наведено порівняльну характеристику їх можливостей і обмежень.

Сучасні системи електронної пошти активно застосовують машинне навчання та штучний інтелект для автоматичного розпізнавання фішингових листів. Наприклад, служба Gmail використовує передові AI-моделі та сигнали від мільярдів кінцевих точок для проактивного блокування понад 99,9% спаму, фішингових спроб і шкідливого ПЗ, не допускаючи їх до скриньок користувачів [11].

Аналогічно, платформа Microsoft Defender for Office 365 [17] навчає ML-моделі на великому масиві даних та аналізує різнотипні сигнали, такі як інфраструктуру відправника, поведінкові індикатори та контекст повідомлення, щоб виявляти фішинг не лише за вмістом листа. Використання великих мовних моделей (LLM) і аналізу тональності (sentiment analysis) дає змогу таким системам розпізнавати наміри зловмисника з точністю до 99,99% при виявленні BEC і схожих загроз. Таким чином, AI-фільтри здатні знаходити підозрілі листи на основі сукупності ознак – від неприродної лексики або емоційного забарвлення до аномалій у заголовках або шаблонах розсилки.

Важливим технічним механізмом протидії фішингу є впровадження стандартів перевірки автентичності e-mail, зокрема SPF, DKIM та DMARC. SPF (Sender Policy Framework) і DKIM (DomainKeys Identified Mail) дозволяють підтвердити, що лист надіслано з авторизованого серверу від імені домену відправника. DMARC (Domain-based Message Authentication, Reporting and Conformance) об'єднує SPF/DKIM і задає політики поводження з неавтентифікованими листами.

Масштабне впровадження DMARC значно ускладнює підробку корпоративних адрес: за даними Google [18], після вимоги до масових розсилників підтримувати DMARC (з 2024 року) кількість неавтентифікованих email у Gmail зменшилася на 265 млрд (на 65%), а протягом пікових фішингових періодів (як-от святковий сезон) користувачі стикалися на 35% менше зі спробами шахрайства [11]. Хоча досі ~87% доменів у світі не мають налаштованого DMARC-запису, динаміка його впровадження прискорюється через регуляторні вимоги (PCI DSS 4.0, EU DORA тощо).

Широке і коректне застосування DMARC вважається “індикатором кіберзрілості” організації та здатне суттєво знизити успішність спуфінгу доменів і фішингу. Водночас слід зважати, що зловмисники адаптуються: вони дедалі частіше реєструють підробні або схожі домени (інколи з увімкненим для них DMARC), аби обходити перевірки, вводячи в оману кінцевих користувачів. Тому технічні методи автентифікації мають доповнюватись аналізом вмісту та відправників на предмет імітації брендів (відстеження доменів-двійників, нелегітимного використання логотипів тощо).

Розвиненим підходом є використання методів аналізу аномалій у метаданих і структурі листів. Машинне навчання дає змогу будувати соціальні граfi комунікацій в організації та виявляти нестандартні взаємодії. Наприклад, алгоритм може навчитися нормальним шаблонам листування (хто з ким зазвичай листується, у який час) і сповістити, якщо відбувається нетипова комунікація, скажімо, якщо раптом генеральний директор пише на пряму молодшому бухгалтеру прохання провести платіж, що є характерною ознакою BEC-атаки.

Інший напрям – профілювання стилю листів за допомогою обробки природної мови. Кожна людина має унікальні мовні риси (словниковий запас, манеру формулювати речення, наявність чи відсутність емодзі тощо). На основі історичних даних ML-моделі можуть навчитися відрізняти стиль написання конкретного співробітника і виявляти лист, що видає себе за повідомлення від нього, але містить лінгвістичні невідповідності. Структурний аналіз заголовків та маршрутів передачі email теж дає цінні сигнали: якщо звичайні листи від даного відправника приходять з певних серверів через 1–2 вузли, а новий лист нібито від нього пройшов через 10 незвичних проміжних хостів – це ознака можливого підроблення. Так само відсутність характерних для законної поштової служби технічних заголовків (або навпаки, їх неочікувана наявність) може видати фішингову розсилку.

Таким чином, базування профілів “нормальної” пошти для кожного користувача з подальшим автоматичним порівнянням нових листів із цими профілями дозволяє алгоритмам аномальної детекції підсвічувати підозрілі повідомлення ще до того, як їх відкриє отримувач.

Технічні рішення дедалі частіше забезпечують перевірку вмісту листа не тільки при його доставці, а й у момент взаємодії користувача. Зокрема, поширені механізми “безпечних посилань” (Safe Links), які перепишують URL у вхідних листах і повторно

перевіряють їх безпечність у момент кліку користувача. Це дозволяє нейтралізувати сценарії, коли спершу шкідливе посилання не значиться в базах та не виявляється фільтрами, але згодом (після доставки листа) ресурс починає розповсюджувати фішинговий контент або зловмисний код. Аналогічно, вкладені файли можуть автоматично аналізуватися у “пісочницях” (sandbox) – ізоляційних середовищах, де емулюється відкриття файлу, щоб перевірити його поведінку. За даними галузевої аналітики, кожен четвертий вкладений файл у електронному листі є шкідливим, тому інтеграція антивірусних сканерів і детонування вкладень у захищеному середовищі – критична складова технічного захисту.

Анти-фішингові веб-фільтри ще один шар безпеки: браузері (наприклад, Microsoft Defender SmartScreen в Edge) та проксі-сервери можуть блокувати переходи на відомі фішингові сайти і попереджати користувача про небезпеку. Компаніям рекомендується використовувати такі рішення для всіх робочих станцій і мереж, аби автоматично зупиняти спроби переходу за сумнівними URL.

Попри те, що зловмисники навчилися часто змінювати інфраструктуру, інтеграція глобальної ТІ-аналітики у технічні засоби (шлюзи безпеки електронної пошти, системи захисту веб-доступу тощо) дозволяє проактивно фільтрувати значну частину фішингового трафіку. Додатково, організації можуть розгортати внутрішні засоби моніторингу мережевого трафіку, щоб виявляти аномальну активність, пов’язану з фішингом, наприклад, масові вихідні листи з облікового запису (що може сигналізувати розсилку атакером) чи нетипові запити з ПК користувача на фішинговий сервер (свідчить про перехід за шкідливим посиланням). Отримані індикатори можуть автоматично блокуватися або слугувати спусковим гачком для реагування.

Організаційні підходи охоплюють впровадження стратегій і політик, що мінімізують вплив успішного фішингу на бізнес. Однією з сучасних парадигм є архітектура нульової довіри (Zero Trust). Її принцип – не довіряти жодному суб’єкту за замовчуванням, незалежно від того, знаходиться він всередині чи поза корпоративною мережею. Стосовно електронної пошти Zero Trust означає, що жоден лист чи вкладення не вважаються “безпечними” апріорі – усі повинні проходити верифікацію, а доступ користувачів до ресурсів має підтверджуватися додатковими перевітками [23].

Практична реалізація цієї моделі включає кілька рівнів контролю. По-перше,

стійка до фішингу автентифікація: організації запроваджують обов'язкову багатофакторну автентифікацію (MFA) для доступу до критичних систем, бажано з використанням методів, що не піддаються підробці через фішинг (наприклад, апаратні U2F-ключі або одноразові токени у додатках). Це значно ускладнює зловмиснику використання вкрадених через фішинг паролів, адже без фізичного ключа або біометрії вхід буде заблоковано.

По-друге, застосовуються політики умовного доступу (Conditional Access) – автоматичне обмеження дій або вимога додаткової перевірки, якщо вхід з облікового запису відбувається з підозрілих умов (новий пристрій, незвична геолокація, невідповідний час тощо).

По-третє, принцип мінімальних привілеїв: навіть якщо зловмисник отримав доступ до облікового запису співробітника, правильно сегментована інфраструктура та обмеження прав не дозволять йому легко “розвинути успіх”. Наприклад, у Zero Trust-мережі сервери і служби аутентифікують та авторизують кожен запит, тож викрадені облікові дані одного сервісу не відкривають автоматично доступ до інших.

На рівні організації важливо встановити чіткі політики безпечного використання email. По-перше, компанії повинні реалізувати протоколи SPF/DKIM/DMARC на власних доменах. Це не лише технічний, а й стратегічний крок – таким чином організація захищає свою репутацію і клієнтів, унеможливлюючи розсилку шахрайських листів від імені її домену.

По-друге, варто застосовувати позначення для зовнішніх листів: багато установ додають автоматичний банер або префікс в тему листа, якщо він надійшов ззовні. Така міра нагадує отримувачам бути більш пильними з повідомленнями від позаорганізаційних відправників, особливо якщо ті видають себе за колег чи керівництво.

По-третє, керівництво має запровадити чіткі процедури підтвердження критичних запитів, що надходять електронною поштою. Згідно з рекомендаціями IBM, будь-які незвичні прохання надати конфіденційну інформацію чи здійснити фінансову транзакцію мають перевірятися співробітниками через незалежний канал – наприклад, телефоном напряму до ініціатора запиту, або особисто. Така практика допомагає зупинити типові атаки на кшталт ВЕС, коли шахрай переконує бухгалтера перевести гроші на підставний рахунок.

Не менш важливим організаційним аспектом є готовність до інцидентів

фішингу. Організація повинна мати план реагування: визначені процедури, як діяти, якщо співробітник повідомив про підозрілий лист або гірший сценарій – клікнув на фішингове посилання і ввів свої дані. План може передбачати негайне сповіщення ІТ-безпеки, тимчасове відключення облікового запису користувача, примусову зміну паролів, сканування пристрою на шкідливе ПЗ, проведення форензики для оцінки масштабів компрометації тощо. Важливо налагодити внутрішню систему повідомлення про фішинг: співробітники мають знати, куди і як пересилати підозрілі листи (окремий адрес служби безпеки або кнопка “Report phishing” у поштовому клієнті). Аналітики з безпеки повинні оперативно перевіряти такі репорти і, у разі підтвердження фішингу, блокувати розсилку для всіх працівників, внесення домену-нападника в чорні списки поштових шлюзів тощо.

Окрім реактивних заходів, превентивна аналітика також належить до організаційних підходів: регулярний аудит поштових логів, оцінка ефективності спам-фільтрів, аналіз тенденцій фішингових загроз з відкритих джерел (звіти ENISA, CERT, тощо) та відповідне коригування політик безпеки. Керівництво організації має забезпечити достатній кадровий і фінансовий ресурс на ці заходи – як показує практика, інвестиції у проактивний захист значно знижують ризик масштабного інциденту.

Згідно зі звітом ЄС, у 2023 році атаки типу «фішинг + компрометація облікових записів» перебували серед лідерів кіберзагроз (30% усіх інцидентів), тож організації, що знехтували побудовою процесів протидії цьому вектору, мають високі шанси стати ціллю зловмисників.

Ефективність політик значною мірою залежить від того, наскільки співробітники їх дотримуються. Тому завдання керівництва – сформувати культуру, за якої кібергігієна є природною частиною робочих процесів. Це досягається регулярною комунікацією “згори вниз” про важливість безпеки, особисту відповідальність кожного та потенційні наслідки нехтування правилами.

Корисною практикою є періодичний випуск внутрішніх бюлетенів чи розсилок з описом нових видів фішингових шахрайств, що з’явилися в індустрії, з порадами як їх розпізнати. Окрім того, варто впроваджувати позитивне підкріплення: відзначати підрозділи або осіб, які демонструють пильність (наприклад, вчасно повідомили про фішингову спробу, що дозволило її нейтралізувати). Культура нульової толерантності до порушень політик (скажімо, спільного використання облікових записів чи обходу

вимог MFA) також має бути чітко артикульована. Загалом, організаційні заходи створюють середовище, у якому технічні рішення та людський фактор можуть максимально ефективно взаємодіяти для захисту від фішингу.

Людський фактор традиційно є найбільш вразливою ланкою в кібербезпеці, тому робота з поведінкою користувачів – ключовий напрям протидії фішингу. Дослідження свідчать, що переважна більшість успішних кібератак стається через помилку чи необережність співробітника, а не через технічні уразливості.

Отже, просвіта і тренування персоналу покликані мінімізувати ймовірність таких помилок. Стандарти безпеки (наприклад, рекомендації NIST, ENISA) наголошують на необхідності регулярних навчальних сесій для всіх співробітників щодо розпізнавання фішингових листів та інших соціально інженерних прийомів. Курс кібергігієни зазвичай охоплює ознаки підозрілих повідомлень (неузгоджені адреси відправників, граматичні помилки, надмірно тривожний або навпаки занадто привабливий зміст, вимога негайних дій тощо) і правила безпечної поведінки (не відкривати вкладень від незнайомих осіб, не переходити за неперевіреними лінками, не вводити паролі на сторонніх сайтах).

Одним з найефективніших поведінкових методів є проведення фішинг-симуляцій – контрольованих розсилок псевдофішингових листів всередині організації з навчальною метою. Подібні симуляції дозволяють “тренувати м’яз” кіберуважності: співробітники отримують досвід розпізнавання фішингу у безпечному середовищі.

Також керівництво отримує метрики рівня ризику – якій частці працівників потрібно покращити навички (через показник кліків за фальшивими посиланнями, частоту повідомлень про підозрілі листи тощо). Для оцінки складності таких навчальних атак використовується, зокрема, шкала NIST Phish Scale, яку впроваджують у програми кібернавчання по всьому світу. Ця методика дозволяє класифікувати кожен тренувальний фішингову приманку за рівнем “складності” для людини, враховуючи наявність або відсутність різних психологічних та технічних підказок. Наприклад, лист без очевидних помилок, стилізований під внутрішню розсилку і з правдоподібним приводом (пре-текстом) матиме високий бал складності.

Аналізуючи результати симуляцій з урахуванням оцінки складності, фахівці можуть більш точно інтерпретувати відсоток кліків: високий відсоток кліків на дуже важку приманку не є провалом, натомість навіть одиничні “провали” на елементарних приманках сигналізують про проблеми в обізнаності персоналу.

Загалом, імітаційні атаки слугують інструментом безперервного навчання: після розсилки усім співробітникам, хто помилився, надсилається роз’яснення, що саме мало насторожити в листі. Доведено, що регулярне проведення таких вправ суттєво знижує рівень успішності реальних фішингових атак, адже користувачі діють більш обережно і автоматично застосовують набуті навички перевірки листів.

Поведінкові підходи мають на меті перетворити кожного співробітника з потенційної жертви на активний елемент системи захисту. Для цього практикується концепція “Stop, Think, Verify”: користувачів привчають перед тим, як виконати дію з листа (перейти за посиланням, відкрити вкладення або відповісти з конфіденційною інформацією), зупинитися і усвідомлено оцінити легітимність запиту. Якщо є найменші сумніви – обов’язково перевірити альтернативним способом.

Важливим аспектом є зниження страху повідомляти про інцидент: керівництво повинно заохочувати персонал негайно повідомляти, якщо вони випадково клікнули на фішингове посилання або розкрили дані. Реакція на такі повідомлення має бути конструктивною: замість покарання – підтримка і додаткове навчання, адже відкритість інформації дозволяє швидко вжити заходів і запобігти розвитку атаки.

В цілому, поведінкова стійкість до фішингу досягається постійними тренуваннями та комунікацією. Навіть у разі зростання складності атак (наприклад, масове використання зловмисниками генеративного AI для створення правдоподібних фішингових листів) підготовлені користувачі менш схильні діяти імпульсивно.

Порівняльна характеристика технічних, організаційних та поведінкових підходів до виявлення та протидії фішинговим атакам наведена в табл. 1.1. Ця таблиця узагальнює ключові ознаки, переваги й обмеження кожної з груп методів.

Таблиця 1.1.

Порівняння підходів до виявлення та протидії фішинговим атакам

Критерій	Технічні підходи	Організаційні підходи	Поведінкові підходи
Основна спрямованість	Автоматичне виявлення і блокування фішингу засобами технологій	Управлінські рішення, політики і процеси, що створюють	Формування у користувачів навичок та поведінки, які

	та інфраструктури.	середовище для запобігання та стримування атак.	зменшують ймовірність успішного фішингу.
--	--------------------	---	--

Продовження таблиці 1.1.

Порівняння підходів до виявлення та протидії фішинговим атакам

Приклади методів	AI/ML-фільтри електронної пошти; аномалія-детекція; SPF, DKIM, DMARC; сканування URL і файлів; чорні списки фішингових сайтів.	Модель Zero Trust (MFA, мінімальні привілеї); впровадження SPF/DKIM/DMARC на доменах; політики підтвердження платежів; процеси реагування на інциденти; інтеграція Threat Intelligence.	Кібергігієна-тренінги для співробітників; регулярні фішинг-симуляції; інструктаж «думай перед тим, як клікти»; програми заохочення повідомлень про підозрілі листи.
Переваги	Висока швидкодія та масштабованість: тисячі листів перевіряються миттєво. Виявлення відомих та нових загроз завдяки інтелектуальним алгоритмам. Автоматичний захист без участі	Системність: створює багаторівневий захист і чіткі правила. Запобігає атакам ще до їх початку. Забезпечує готовність до інцидентів і швидке реагування.	Адресує соціальну інженерію, проти якої технічні засоби часто безсилі. З часом формує стійку навичку, що захищає і поза межами організації.

	користувача.		
--	--------------	--	--

Продовження таблиці 1.1.

Порівняння підходів до виявлення та протидії фішинговим атакам

Обмеження	Не гарантує відсутності хибнопозитивних чи хибнонегативних результатів. Може бути обійдений атаками без технічних ознак. Потребує постійного оновлення і налаштування.	Залежить від дотримання політик персоналом. Потребує часу та ресурсів на впровадження. Може зустріти опір змін. Не усуває атаку, а лише зменшує її ймовірність та наслідки.	Людський фактор непередбачуваний. Ефективність тренінгів знижується без оновлення. Надмірне навчання може викликати втому або зниження пильності.
-----------	--	---	---

Кожна з розглянутих груп підходів відіграє важливу роль у протидії фішинговим атакам, проте жодна не є всесильною окремо. Технічні засоби забезпечують необхідну автоматизацію та масштаби, перехоплюючи левову частку загроз ще до контакту з користувачем (зокрема, корпоративні поштові фільтри нині затримують понад 99% зловмисних листів).

Водночас організаційні заходи створюють рамки, що ускладнюють реалізацію фішингу та мінімізують збитки у разі його успіху – наприклад, завдяки MFA і Zero Trust зловмисник, навіть викравши пароль, імовірно не отримає доступу до систем, а продуманий план реагування не дасть атаці розвинутися. Поведінкові методи доповнюють технології, охоплюючи ту складову, на яку націлений фішинг – людську довірливість і помилки. Навчений, обачний співробітник може впізнати підробку, яку

не розпізнав автоматичний фільтр, та вчасно її зупинити.

Таким чином, лише поєднання технічних, організаційних і освітніх заходів здатне забезпечити всебічний захист від фішингових атак. Враховуючи зростання креативності зловмисників (зокрема використання ШІ для удосконалення фішингу), стратегія “глибинного захисту” із залученням усіх трьох підходів є обов’язковою умовою кіберстійкості сучасної організації.

1.3. Аналіз існуючих рішень і технологій захисту від фішингових атак

У сучасних умовах зростаючої кількості та складності фішингових атак ефективний захист організацій вимагає впровадження спеціалізованих технологічних рішень. Протягом останніх років з’явилася велика кількість інструментів і платформ, орієнтованих на запобігання, виявлення та реагування на фішинг — від поштових шлюзів і систем на основі штучного інтелекту до сервісів аналітики загроз і навчальних модулів.

Електронна пошта залишається головною ціллю для фішингових атак, тому більшість рішень зосереджені на її захисті. Сучасні засоби захисту пошти можна умовно поділити на дві категорії: традиційні захищені поштові шлюзи (Secure Email Gateway, SEG) та хмарні системи захисту пошти, в тому числі вбудовані механізми платформ електронної пошти і новітні API-рішення. Обидва підходи мають на меті фільтрувати шкідливі та фішингові листи до того, як вони досягнуть користувачів, але реалізують це різними способами.

Захищені поштові шлюзи (SEG) – це класичні рішення, що діють як проміжна ланка між інтернетом та корпоративним поштовим сервером. Вони перенаправляють весь вхідний трафік електронної пошти через себе (шляхом зміни MX-записів домену) та здійснюють багаторівневу перевірку листів: фільтрацію спаму, виявлення фішингових ознак, перевірку вкладень і посилань, карантин підозрілих повідомлень тощо. Такі шлюзи використовують поєднання технік – сигнатурний аналіз, правила (наприклад, блокування відомих підозрілих відправників або ключових слів), машинне навчання для оцінки аномалій, а також інтегровані пісочниці для динамічного аналізу вкладень. Приклади провідних SEG-рішень: Proofpoint Email Security, Mimecast, Barracuda Email Security, Cisco Secure Email, Trend Micro Email Security, Sophos Email, Fortinet FortiMail тощо. За даними аналітичних оцінок 2024

року [24], такі постачальники як Proofpoint та Mimecast стабільно входять до лідерів ринку захисту пошти.

Перевагою SEG є їхня здатність блокувати більшість масових фішингових та спам-атак на периферії, до потрапляння в поштові скриньки. Вони надають адміністраторам гнучкий контроль політик безпеки (напр. блокування виконуваних файлів, перевірка URL у реальному часі, шифрування конфіденційних листів). Водночас, обмеженнями традиційних шлюзів є складність їх впровадження та підтримки (особливо апаратних рішень) і потенційна вразливість до витончених атак, що обходять типові фільтри (наприклад, цілеспрямовані BEC-листи без шкідливих посилань).

Крім того, у хмарних середовищах (Office 365, Gmail) деякі організації надають перевагу вбудованим засобам або API-рішенням, що не потребують зміни MX-записів, – це поступово кидає виклик домінуванню класичних SEG. Дійсно, експерти відзначають тенденцію переходу від шлюзів до хмарних API-рішень: прогнозувалося, що до 2023 року вже 20% антифішингових рішень постачатимуться через інтеграцію по API (проти <5% кількома роками раніше) [24]. Традиційні вендори реагують на цей тренд, додаючи хмарні функції або здійснюючи поглинання нових розробників API-платформ.

Хмарні API-рішення для захисту пошти інтегруються безпосередньо з поштовими сервісами (наприклад, через API Office 365 або Gmail) та забезпечують додатковий шар безпеки поверх вбудованих функцій провайдера. Ці сервіси часто називають CAPES (Cloud Email Security, API-enabled). На відміну від шлюзів, вони не вимагають перенаправлення MX-запису – натомість, вони отримують доступ до поштових скриньок через API для аналізу вхідних листів, й можуть автоматично видаляти або переміщувати небезпечні повідомлення. Серед прикладів – рішення компаній Abnormal Security, Ironscales, Avanan (Check Point), а також Material Security, Barracuda Sentinel, Trend Micro Cloud App Security, Proofpoint CASB/Cloud та інші. Такі платформи зазвичай хмарно-нативні та активно використовують штучний інтелект для виявлення загроз, зокрема складних фішингових прийомів: аналіз поведінкових ознак, побудова профілів комунікації користувачів, виявлення аномалій у змісті листування тощо. Наприклад, Abnormal Security позиціонується як рішення, що за допомогою штучного інтелекту визначає відхилення від нормальної поведінки в електронній пошті і таким чином виявляє спроби соціальної інженерії (включаючи

ВЕС та інші цільові атаки).

Перевагою API-підходу є спрощене розгортання (особливо для організацій, що вже в хмарі) та розширений огляд – такі системи можуть перевіряти не лише зовнішні, а й внутрішні листи або листування між співробітниками, виявляючи компрометацію облікових записів і латеральний фішинг (коли зламаний внутрішній акаунт розсилає фішинг іншим). Обмеженням деяких API-рішень є потенційна затримка: оскільки лист може з'явитися у скриньці одночасно з його аналізом через бекенд-API, існує вікно часу, за яке користувач міг би побачити шкідливий лист, перш ніж система його видалить. Для мінімізації ризиків деякі вендори впроваджують inline-фільтрацію навіть у хмарі, наприклад, технологія INKY дозволяє вставляти себе між поштовим транспортом і сховищем, щоби перевірити лист до доставки.

Глобальні хмарні платформи електронної пошти мають власний вбудований захист, який постійно удосконалюється. Компанія Google, наприклад, використовує сервіс Google Safe Browsing для блокування відомих фішингових URL у Gmail, автоматично позначає підозрілі листи (помітки щодо зовнішніх відправників, попередження про можливе шахрайство), застосовує машинне навчання на великому масиві даних (мільярди листів щодня) для фільтрації спаму і фішингу [11].

Microsoft у складі Exchange Online Protection (EOP) та Microsoft Defender for Office 365 пропонує аналогічні функції: базове фільтрування спаму і зловмисного ПЗ, а в розширених планах – Safe Links (перевірка URL при кліку), Safe Attachments (відправлення вкладень у пісочницю для аналізу), а також захист від імперсонації (спроб видавати себе за іншу особу) через налаштування політик (наприклад, попередження, якщо домен відправника схожий на корпоративний або якщо лист від імені керівника надійшов ззовні) [17].

Перевага вбудованих засобів – простота використання (вони ввімкнені за замовчуванням, не потребують окремого розгорнення) та глибока інтеграція з екосистемою (єдиний інтерфейс адміністрування, журнали в хмарі тощо). Ці засоби ефективно блокують масові й відомі загрози. Втім, аналітики та практика показують, що вбудованого захисту часто недостатньо перед обличчям новітніх, цілеспрямованих фішинг-кампаній. Згідно з дослідженням Trend Micro, лише у 2023 році їхні додаткові засоби виявили понад 45 млн небезпечних електронних листів, які пройшли крізь базові фільтри Microsoft 365 та Google Workspace. Це підтверджує, що особливо витончені загрози здатні оминати стандартні механізми.

Дійсно, у рейтингах лідерів незалежних аналітиків фігурують переважно спеціалізовані вендори: наприклад, у Magic Quadrant для захисту електронної пошти (2024) до лідерів віднесено Proofpoint, Trend Micro, Mimecast, Abnormal Security, Check Point та ін., тоді як Microsoft лише у категорії “challengers”, а Google взагалі не увійшла до переліку (рис. 1.5). Це свідчить, що попри величезні ресурси й дані, основні постачальники поштових платформ дещо поступаються у швидкості впровадження нових антифішингових можливостей спеціалізованим компаніям (зокрема, Gartner відзначив, що Microsoft більше зосереджується на суміжних сферах на кшталт XDR чи AI Copilot, ніж на розвитку суто поштового захисту).



Рис. 1.5. Магічний квадрант для платформ з безпеки поштових сервісів [21]

Отже, комплексна стратегія часто включає використання кількох рівнів захисту: базового (вбудованого) плюс одного або навіть кількох зовнішніх рішень для максимального покриття різних типів атак.

Більшість сучасних продуктів для захисту від фішингу включають компонент динамічного аналізу вмісту – так звані пісочниці (sandbox). Ця технологія передбачає запуск підозрілих файлів або відкриття веб-сторінок у ізольованому віртуальному

середовищі з метою спостереження за їх поведінкою. Якщо файл-вкладення здійснює шкідливі дії (наприклад, встановлює з'єднання з командним сервером, шифрує файли тощо) або веб-сторінка намагається викрасти облікові дані, то система визначає їх як небезпечні і блокує, перш ніж вони зашкодять реальному користувачу.

Традиційно велика частина фішингових атак містила шкідливі документи (наприклад, Word/Excel із макросами, PDF із експлойтом). Поштові захисні шлюзи давно перевіряють хеші файлів за антивірусними базами, але для невідомих зразків потрібна пісочниця. Наприклад, Microsoft Safe Attachments направляє всі вкладення, яких не знає антивірус, у спеціальну середу, де їх відкриває та моніторить. Proofpoint Targeted Attack Protection (TAP) діє подібно, як і Cisco Threat Grid, FortiSandbox та інші продукти. За статистикою, цей підхід виправданий: після того як Microsoft у 2022 році за замовчуванням вимкнула автоматичне виконання макросів у документах Office, число відомих шкідливих вкладень різко зменшилося (на тисячі відсотків), але натомість зросла кількість нових видів шкідливих файлів – тобто атакуючі шукають нові шляхи, які все одно треба аналізувати динамічно. Пісочниці сьогодні часто розгорнуті у хмарі постачальника рішення, що дозволяє масштабувати обчислення і швидко перевіряти великий обсяг файлів паралельно.

Фішингові посилання – інший ключовий елемент атак. Просте блокування відомих шкідливих URL (за базою) не завжди ефективне, адже зловмисники генерують тисячі унікальних посилань (інколи одноразових) або відстрочують активацію фішингового сайту. Тому застосовується техніка URL rewriting – при проходженні через шлюз усі посилання в листі переписуються на проміжний захисний URL. Коли користувач переходить за посиланням, система спочатку перенаправляє його через свою хмару для перевірки: сканує кінцеву сторінку на предмет шкідливого коду, перевіряє репутацію домену, порівнює зі списками відомих фішингових сайтів. За необхідності, сторінка може автоматично бути відкрита в browser sandbox (ізолюваному браузері) і проаналізована на наявність форм логіну, підозрілих скриптів тощо. Наприклад, Microsoft Safe Links виконує таку перевірку кожного разу при кліку, Mimecast URL Protect – схожим чином, Zscaler у межах свого веб-шлюзу також аналізує контент кожної HTTP-запитуваної сторінки на льоту. За даними Gartner, однією з обов'язкових можливостей рішень класу Email Security є саме аналіз та захист URL у листах.

Пісочниця посилань також еволюціонувала у напрямку ізоляції контенту. Деякі

рішення пропонують відкривати всі підозрілі або невідомі посилання у віддаленому ізольованому браузері. Наприклад, технології браузерної ізоляції від Menlo Security, Broadcom (isolation) чи того ж Zscaler дозволяють користувачу переглянути сторінку, але фактично виконання коду відбувається на ізольованому сервері – користувач отримує лише безпечний “стрім” інтерфейсу. Це радикально знижує ризик зараження чи крадіжки даних зі шкідливих сайтів. Однак через високу вартість і вплив на зручність роботи ізоляцію застосовують зазвичай вибірково (для дуже ризикових категорій контенту або для окремих посад).

Отже, “пісочниці” і активний контент-аналіз зараз є стандартною частиною комплексних антифішингових платформ. Вони особливо ефективні проти невідомих раніше атак (нульового дня) з використанням шкідливих файлів або посилань. Недоліком є підвищені вимоги до ресурсів і можлива невелика затримка доставки листів (зазвичай частки секунди або кілька секунд, поки файл сканується). Проте для критичних середовищ ці затримки виправдані, адже дозволяють запобігти інфікуванню мережі організації троянами чи шифрувальниками через фішинговий імейл.

Таблиця 1.2.

Порівняння сучасних рішень для захисту електронної пошти від фішингових атак

Рішення / Платформа	Клас	Основні функції	Підтримка API	Пісочниця / Захист URL	Доступність
Microsoft Defender for Office 365 (P1)	Вбудоване + SEG	Захист пошти, Safe Links, Safe Attachments, виявлення фішингу, захист від BEC	Обмежена	Так	1 місяць безкоштовно
Proofpoint Essentials	SEG	Фільтрація спаму/фішингу, пісочниця, захист URL, аналітика,	Ні	Так	Безкоштовне демо на запит

		карантин			
--	--	----------	--	--	--

Продовження таблиці 1.2.

Порівняння сучасних рішень для захисту електронної пошти від фішингових атак

Abnormal Security	API (CAPE S)	AI-виявлення BEC, аналіз поведінки, виявлення аномалій, інтеграція з O365/Gmail	Так	Частково	Доступне лише для корпоративних клієнтів
Trend Micro Cloud App Security	API + SEG	Захист O365/Gmail, пісочниця, аналіз вкладень і URL, ML, карантин	Так	Так	Безкоштовний тестовий період (до 30 днів)
Google Workspace	Вбудоване	Захист URL, SPF/DKIM/DMARC, автоматичне маркування підозрілих листів	Ні	Ні	Доступне у всіх акаунтах за замовчуванням

З огляду на можливість розгортання в лабораторних умовах та дослідження без додаткових витрат, оптимальним рішенням для подальшого аналізу обрано Microsoft Defender for Office 365 (Plan 1).

Цей продукт входить до складу Microsoft 365, підтримує пробний доступ на 30 днів і включає ключові функції для боротьби з фішингом: перевірку посилань (Safe

Links), динамічний аналіз вкладень (Safe Attachments), виявлення спроб шахрайства (BEC), а також карантин і звіти.

Крім того, Defender for Office 365 інтегрується з іншими сервісами Microsoft — зокрема, системами моніторингу, журналювання та безпеки (Defender XDR, Microsoft Sentinel), що дає змогу зібрати повнішу аналітику в межах експерименту.

Аналіз існуючих рішень свідчить про переорієнтацію ринку від традиційних шлюзів на хмарні технології з підтримкою API та штучного інтелекту. Сучасні платформи комбінують різні механізми захисту — від сигнатурної перевірки до поведінкового аналізу — і дозволяють створити багаторівневу оборону від фішингових атак.

На основі доступності, функціональності та потенціалу для експериментального дослідження у межах магістерської роботи, було обрано рішення Microsoft Defender for Office 365 (Plan 1) як об'єкт для подальшого тестування та моделювання.

2 АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ІНФОРМАЦІЙНИХ СИСТЕМАХ

2.1 Класифікація фішингових атак і вектори їх реалізації

Електронна пошта й веб-сайти виступають ключовими векторами доставки таких атак. Для ефективного аналізу і протидії фішингу важливо розуміти сучасну класифікацію фішингових атак за різними критеріями – за ціллю і мотивом зловмисника, за технікою виконання, за каналом доставки, а також за рівнем таргетованості (персоналізації) атак.

Залежно від мотивів зловмисників фішингові кампанії можуть переслідувати різні цілі. Найпоширеніша мета – викрадення логінів, паролів, токенів сесій тощо, щоб отримати несанкціонований доступ до акаунтів або систем. Оцінки свідчать, що близько 80% фішингових кампаній націлені саме на компрометацію облікових записів користувачів, особливо хмарних сервісів на кшталт Microsoft 365 чи Google Workspace [25]. Для цього фішери створюють копії сторінок входу до популярних сервісів або банків і заманюють жертву ввести там свої дані.

Фінансова вигода та шахрайство. До цієї категорії належать атаки, метою яких є безпосереднє отримання грошей або фінансової інформації жертви. Наприклад, ВЕС – компрометація ділової електронної переписки, коли зловмисник видає себе за керівника або партнера і переконує співробітника здійснити фіктивний платіж. Такі цільові «whaling»-атаки (полювання на «великих риб») можуть призвести до багатомільйонних збитків для організацій (відомі випадки втрат до \$47 млн від однієї атаки).

Установка шкідливого ПЗ. Фішинг часто використовується для потайного доставлення ШПЗ – троянів-завантажувачів, шпигунського ПЗ, програм-вимагачів. Зловмисники розсилають електронні листи з інфікованими вкладеннями або посиланнями на експлойт-сторінки. Метою є зараження

системи жертви і розвиток атаки (наприклад, викрадення конфіденційної інформації або шифрування даних).

Кіберрозвідка та шпигунство. Особливо серед державних АРТ-груп популярні таргетовані фішингові розсилки для збору розвідданих. Так званий spear-фішинг (цільовий фішинг) застосовувався у 65% відомих хакерських груп, причому 96% їхніх цільових атак мали на меті добути розвід інформацію. У таких випадках жертвами стають конкретні посадові особи чи організації, а фішингові листи спеціально адаптуються під них.

Класифікація за рівнем персоналізації (таргетованості). Фішингові атаки варіюються від масових розсилок до вузьконаправлених операцій проти окремих осіб:

Масовий фішинг. Низькоспрямовані «широкі» кампанії, коли одна й та сама приманка розсилається на тисячі адрес. Такі листи зазвичай універсальні, не містять особистої інформації жертви (типові звернення на кшталт «Шановний клієнте») і часто маскуються під повідомлення від відомих брендів (банків, поштових служб, соцмереж). Мета – зачепити хоча б невеликий відсоток отримувачів. Попри примітивність, масовий фішинг досі становить значну частку атак і часто слугує для розповсюдження малварі або збору базових облікових даних.

Цільовий фішинг (spear). Атаки, розроблені під конкретну ціль або групу осіб. Зловмисники збирають інформацію про жертву (через соцмережі, витoki даних, новини компаній) і використовують її для створення правдоподібного підробленого листа. Спеар-фішингові повідомлення можуть виглядати так, ніби надіслані колегою, партнером або керівником жертви, містити її ім'я, посаду, актуальні професійні деталі. Завдяки цьому жертва менш підозріло ставиться до прохання в листі – перейти за посиланням, відкрити файл чи надати інформацію.

Whaling («китове полювання»). Особлива форма spear-phishing, націлена на «велику рибу» – керівників вищої ланки, топменеджерів, держслужбовців високого рангу. Такі особи мають розширені повноваження та доступ до конфіденційних даних або фінансів, тому їх компрометація особливо цінна.

Атаки типу whaling часто маскуються під ділову кореспонденцію, юридичні запити або екстрені фінансові питання, адресовані саме високопосадовцю. Через нижчу технічну обізнаність деяких керівників і великий обсяг їхньої пошти, whaling-атаки нерідко успішні.

Компрометація переписки (BEC) - це соціотехнічні атаки, які можуть поєднувати злам чужого облікового запису та фішингові методи. Зловмисник може спершу отримати доступ до корпоративної пошти (через фішинг пароля), а потім, використовуючи реальний обліковий запис співробітника, надсилати правдоподібні листи іншим співробітникам або партнерам. Інший сценарій – зловмисник без злomu імпортує у фальшивому листі стиль і підпис реального керівника і просить, наприклад, бухгалтерію здійснити терміновий платіж. BEC-атаки є дуже небезпечними, оскільки експлуатують довіру до внутрішньої або партнерської переписки.

Класифікація за каналами доставки. Сьогодні фішинг реалізується через різноманітні канали взаємодії з жертвою. Найбільш поширені вектори атаки такі:

Електронна пошта. Класичний і наймасовіший канал для фішингу – e-mail. Зловмисники розсилають підроблені листи, що імітують повідомлення від банків, служб підтримки, колег або популярних сервісів. Лист містить фішингове посилання (URL на шахрайський сайт) або шкідливе вкладення (файл із вірусом чи зловмисним скриптом).

В сучасних кампаніях фішингові листи можуть обходити стандартну фільтрацію: зловмисники використовують Phishing-as-a-Service (PhaaS) – готові набори скриптів та шаблонів, які клоном імітують сторінки входу популярних брендів і автоматично розсилають посилання жертвам. Такі платформи значно спростили запуск фішингових атак навіть для малодосвідчених злочинців.

Веб-сайти (фішингові сторінки). Веб є невід’ємною частиною більшості фішингових кампаній, оскільки саме на підроблених сайтах відбувається викрадення даних. Сюди належать фішингові сайти, що імітують легітимні (банківські портали, сторінки входу до пошти чи соцмереж і т.д.), а також загалом шкідливі сайти, які можуть завантажувати малваре. Жертва зазвичай

потрапляє на такий сайт через посилання з електронного листа, SMS або реклами. Важливо, що нині зловмисники активно використовують шифрування і легітимні домени: близько 80% фішингових веб-сайтів мають дійсний HTTPS-сертифікат, аби виглядати правдоподібно і не викликати підозр у користувачів. Для хостингу фішингових сторінок нерідко застосовуються безкоштовні сервіси або скомпрометовані легітимні сайти.

SMS та месенджери (смішинг). Фішинг через текстові повідомлення набирає обертів. Смішинг (SMS-phishing) – розсилка оманливих SMS, які містять коротке посилання або інструкцію. Типові приклади – повідомлення нібито від банку про «блокування картки» чи від пошти про «посилку, що потребує підтвердження», з посиланням на фальшивий сайт. Одержувачі SMS зазвичай менш обізнані щодо небезпек, тому клікають на посилання з телефону.

Телефонні дзвінки (вішинг). Атаки, при яких шахраї телефонують жертві, видаючи себе за представників банку, техпідтримки Microsoft, податкової тощо. Їх мета – виманити конфіденційні дані (паролі, PIN-коди) або переконати зробити переказ коштів. Часто вішинг комбінується з іншими каналами: наприклад, жертві надсилають e-mail із «терміновим попередженням безпеки», а потім телефонують під виглядом підтримки і просять назвати код або встановити програму. Такі гібридні vishing-атаки в останні роки зросли на ~30% – все більше організацій повідомляють про випадки фальшивих дзвінків від імені керівників чи державних службовців. Однією з тактик є TOAD (Telephone-oriented Attack Delivery) – фішингові листи, що містять номер телефону замість посилання: отримувач сам дзвонить на номер, потрапляючи на шахраїв. Таким чином зловмисники обходять захист поштових шлюзів, адже в листі немає явно шкідливого лінку.

Соціальні мережі та інші платформи. Зростає кількість фішингових атак через прямі повідомлення у соцмережах (Facebook, LinkedIn, Twitter) та корпоративних комунікаційних платформах (Slack, Teams). Зловмисники створюють фейкові акаунти, що видають себе за HR-менеджерів, рекрутерів або службу підтримки, і встановлюють контакт з жертвою. Наприклад, схема angler phishing передбачає пошук скарг користувачів у Twitter і надсилання їм

від імені підтримки посилання для «вирішення проблеми», яке веде на фішингову сторінку. За звітами, близько 40% фішингових кампаній зараз виходять за межі однієї лише електронної пошти і охоплюють інші платформи. Зокрема, шахраї можуть спершу написати жертві в месенджері, потім продублювати листом на пошту, а далі навіть зателефонувати – усе задля підвищення довіри. Така мультиканальність розмиває межі між цифровим та фізичним соціальним інжинірингом і потребує комплексного підходу до захисту.

Класифікація за техніками та приманками. Фішингові атаки відрізняються способами обману та технічними прийомами, що застосовуються для введення жертви в оману:

Посилання на фальшиві сайти. Найтиповіша техніка – вбудувати в повідомлення URL, що веде на шахрайський сайт. Такі посилання можуть бути замасковані під текст типу «Перейти в аккаунт» або приховані за кнопкою. Зловмисники використовують різноманітні трюки для правдоподібності: реєструють домени, схожі на справжні (наприклад, `microsoft.com` із цифрою 0 замість літери), вставляють назву бренду у субдомен (`login.paypal.com.phishing-site.tld`), застосовують гідів-редиректи. Останнє – це багатоетапні перенаправлення: спочатку жертва переходить за посиланням на цілком легітимний ресурс, який потім автоматично перенаправляє на фішингову сторінку.

Вкладення зі шкідливим вмістом. Інша популярна техніка – надсилати у фішинговому листі файл-вкладення, відкриття якого запускає атаку. Це можуть бути документи Office (доки, таблиці) з макросами, сценарії у форматі .HTA чи .VBS, архіви .zip/.rar з виконуваними файлами, PDF-документи з небезпечними посиланнями, а останнім часом навіть файли зображень (SVG, ISO тощо), які містять прихований шкідливий код. Вкладення часто маскуються під бізнес-документи: рахунок-фактура, накладна, резюме або під щось цікаве для людини (наприклад, «фото»). Шкідливе вкладення після відкриття може встановити на комп'ютер троян (для крадіжки паролів, як Infostealer), завантажити програму-вимагач або іншим способом скомпрометувати систему. Наприклад, сімейство

трояну QakBot (QBot) масово доставляється через вкладені файли OneNote, PDF із посиланнями, HTML-файли з прихованим JavaScript (техніка HTML-smuggling) тощо.

Приманки з прямим відгуком (response-based). У деяких фішингових атаках немає ні явних посилань, ні виконуваних файлів – жертву спонукають надіслати відповідь або виконати дію поза електронною поштою. Приклад – лист від «директора» до бухгалтера з проханням терміново відповісти і здійснити платіж, або ж класичні «листки щастя» (афери з передоплатою), де жертві обіцяють велику винагороду і просять відповісти для обговорення деталей. Також сюди належать згадані вище сценарії з номером телефону (жертва телефонує сама).

Новітні методи: QR-фішинг, deepfake та ін. У 2024–2025 роках з'явилися нові різновиди фішингу, що використовують сучасні технології. QR-фішинг (квішинг) – це використання QR-кодів як посилань на фішингові сайти. Шахраї, наприклад, розсилають електронні листи з вкладеним PDF, у якому міститься QR-код (нібито для перегляду документа чи оплати); скануючи його, користувач потрапляє на фальшиву веб-сторінку. Атаки з QR-кодами також здійснюються через фізичні носії: розклеюються фальшиві оголошення з QR-кодом, що веде на сайт для «отримання призу» або встановлення додатка.

Adversary-in-the-Middle (AiTM). Окремо варто відзначити техніку «людина посередині» у фішингу, яка стала відповіддю зловмисників на впровадження багатофакторної автентифікації. У таких атаках фішинговий сайт виступає посередником між жертвою та справжнім сервером: користувач вводить на підробній сторінці свої логін, пароль і навіть одноразовий код 2FA, а атакувальник перехоплює їх в реальному часі і негайно використовує на справжньому сайті. Таким чином генеруються валідні сесійні cookie, якими зловмисник може скористатися для входу без пароля і без виклику повторного MFA-підтвердження. Це дозволяє обходити навіть стійкі механізми захисту. Експерти називають AiTM-фішинг однією з причин успішних зломів облікових записів, захищених MFA, адже нападники фактично здійснюють крадіжку сесії

користувача. На чорному ринку доступні фішингові комплекти EvilProху, Мигаєпа та інші, які автоматизують ці атаки.

Фішингова кампанія зазвичай розгортається у кілька етапів. Спочатку жертві надходить шахрайський електронний лист. Сучасні фільтри здатні відсіювати значну частину спаму ще на цьому кроці. Якщо ж лист минув фільтрацію і потрапив у скриньку – все залежить від дій користувача. В типовому сценарії можливі три розгалуження:

(1) Перехід за посиланням. Жертва клікає URL з листа, у результаті в браузері відкривається фальшива веб-сторінка. Найчастіше це форма для введення облікових даних (логіну, пароля, номера карти тощо) – після заповнення ці дані непомітно пересилаються зловмиснику. Інший варіант – сторінка експлойта, яка без відома користувача завантажує та запускає шкідливий код (так званий drive-by download).

(2) Відкриття вкладення. Жертва завантажує з листа файл (документ, архів, скрипт) і відкриває його, після чого на комп'ютері виконується вбудований шкідливий код. Це призводить до встановлення малварі або іншого компрометування системи.

(3) Прямий відгук. Жертва може відповісти на інструкцію в листі – наприклад, надати запитовані дані у відповіді або зателефонувати за вказаним номером. У такому разі шахраї отримують потрібну інформацію безпосередньо через продовження комунікації.

Приклад схеми реалізації фішингової атаки зображено на рис. 2.1:

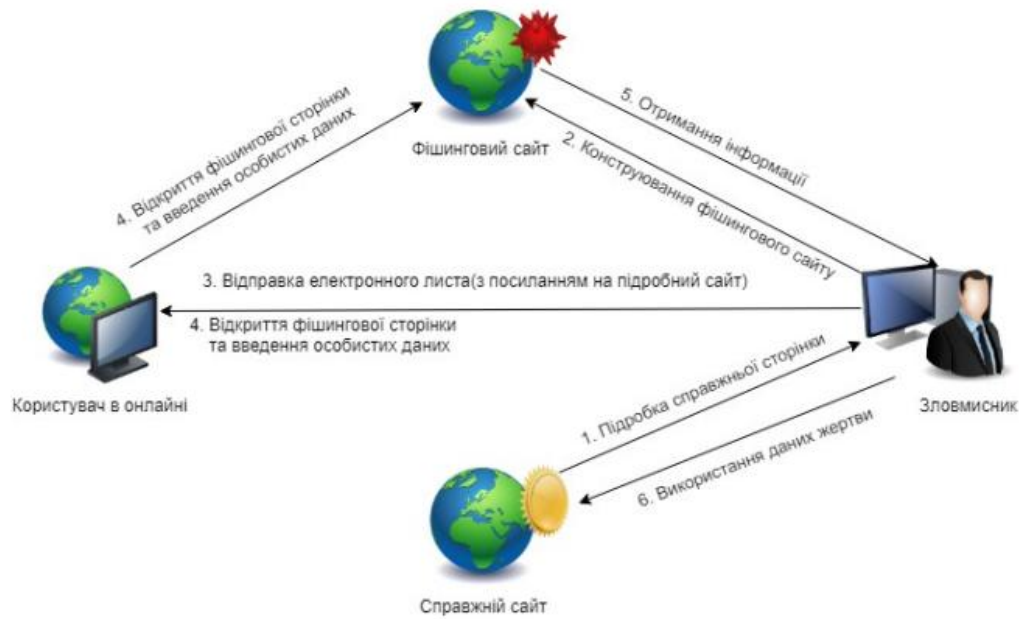


Рис. 2.1. Схема реалізації фішингової атаки

З погляду захисника, усі три гілки розвитку подій небезпечні: навіть якщо користувач не перейшов за посиланням, але зателефонував чи переписується з нападниками, його слід вважати скомпрометованим. У кінцевому підсумку, успіх фішингової атаки залежить від того, чи виконає користувач потрібну дію (клік, запуск файлу, надання інформації). Тому ключовими є як автоматичні бар'єри (спам-фільтри, веб-фільтри), так і навчання користувачів розпізнавати загрози.

Це підкреслює критичну необхідність наскрізного моніторингу і виявлення фішингових атак на всіх етапах – від перехоплення підозрілих листів і блокування фішингових сайтів до аналізу аномальної фінансової активності, що може вказувати на ВЕС.

Сучасні фішингові атаки надзвичайно різноманітні за способом реалізації та каналами, однак переважна їх більшість так чи інакше пов'язана з електронною поштою та вебом як основними векторами. Класифікація фішингу за різними ознаками – цілями, типом жертви, технічними прийомами, каналами доставки – дозволяє краще зрозуміти природу цих загроз і побудувати багаторівневий захист. Статистика останніх років підтверджує актуальність проблеми: обсяги фішингових атак не спадають, навпаки, вони еволюціонують (особливо з появою інструментів на базі ШІ). Фішинг залишається

найпоширенішим методом соціальної інженерії і першим кроком у більшості складних зломів. Тому в наступних підрозділах особливу увагу буде приділено методам виявлення фішингових атак саме в каналах електронної пошти та веб-трафіку, а також засобам нейтралізації фішингових загроз до того, як вони спричинять небажані наслідки.

2.2. Основні принципи побудови систем виявлення фішингових повідомлень

Сучасні системи виявлення фішингу можуть бути впроваджені на різних рівнях інфраструктури організації. Один з поширених підходів – інтеграція в поштові шлюзи або сервери електронної пошти (inline deployment). Система аналізує вхідні листи ще до доставки в поштову скриньку, блокуючи або позначаючи підозрілі повідомлення. Альтернативою для хмарних сервісів є API-інтеграція, коли зовнішня система підключається через API до поштового середовища (наприклад, Microsoft 365, Google Workspace) і перевіряє листи в режимі post-delivery (після доставки) або pre-delivery. Також застосовуються гібридні рішення – одночасно inline та API: листи скануються перед доставкою, а вже доставлені можуть бути відкликані або позначені заднім числом, якщо нові дані вказали на фішинг.

Окрім електронної пошти, схожі принципи використовуються для захисту веб-трафіку. Проксі-сервери та шлюзи веб-доступу можуть аналізувати HTTP/HTTPS трафік, щоб виявляти спроби переходу на фішингові сайти. Наприклад, якщо користувач натискає на посилання з email, спеціальний проксі або модуль Safe Links перенаправляє запит через систему перевірки URL у реальному часі. Це дозволяє заблокувати перехід на шкідливий сайт або ізолювати сесію. Деякі рішення (як-от Cloudflare Area 1) позиціонуються як хмарні сервіси захисту електронної пошти, що діють проксісно: листи організації переадресовуються через хмарний сервіс, де відбувається їх предоставна перевірка з застосуванням ШІ та пісочниць. Важливою тенденцією є модульність та гнучкість розгортання – систему можна вписати практично в будь-яку існуючу інфраструктуру, щоб забезпечити багатоешелонний захист (наприклад, попереднє сканування на шлюзі, динамічний аналіз URL при кліці та пост-фактум моніторинг уже доставлених листів).

Для точного виявлення фішингу системи використовують комплекс різнорідних даних з кожного повідомлення. Наприклад, текст листа (вміст повідомлення) – аналізується зміст електронного листа (тіло і тема) на наявність ознак соціальної інженерії: термінові заклики, погрози або прохання негайно виконати дію, мовні звороти, характерні для шахрайства, неправдоподібні історії тощо. Також враховуються лінгвістичні особливості: граматичні помилки, невідповідність стилю корпоративному листуванню або, навпаки, надто бездоганна мова (що може вказувати на згенерований ШІ текст). Сучасні системи, як-от Microsoft Defender чи Proofpoint, використовують обробку природної мови (NLP) для семантичного аналізу тексту: визначають тональність, наявність фінансових термінів, імітацію стилю відомих осіб тощо.

Заголовки листа і метадані протоколів – аналізуються технічні атрибути: адреса відправника («From»), поле «Reply-To», домен відправника та його схожість на легітимний (наприклад, підміна домену typ0-squatting), ланцюжок серверів (Received headers), час відправлення. Перевіряються результати автентифікації: SPF, DKIM, DMARC. Налаштування DMARC є важливим для блокування спроб спуфінгу доменів – впровадження DMARC допомагає відсіяти листи, що видають себе за домени організації. Також враховуються аномалії, наприклад невідповідність між полем «From» і фактичним доменом в «Return-Path», відсутність легітимних заголовків поштових клієнтів або незвична конфігурація полів. Метадані можуть включати геолокацію IP відправника, наявність IP у відомих блок-листах, показники репутації домену та IP. Сумарно ці сигнали формують профіль технічної довіри листа. Якщо лист походить начебто від керівника, але насправді надісланий з зовнішнього серверу з поганою репутацією та не проходить SPF/DKIM – це сильний індикатор фішингу.

Вбудовані URL-адреси (гіперпосилання) – практично всі фішингові листи містять посилання на шахрайські сайти (для крадіжки облікових даних) або на завантаження шкідливого ПЗ. Система видобуває всі URL з тіла листа (включно зі згорнутими посиланнями, кнопками, зображеннями як посилання). Для кожного URL проводиться перевірка репутації (звірка з базами відомих фішингових доменів, чорними списками), аналіз самого тексту URL (пошук схожості на відомі легітимні домени, наприклад «micros0ft.com») та динамічний аналіз. Динамічний аналіз може включати звернення до URL у ізольованому середовищі (пісочниці) або через URL sandboxing API, щоб перевірити, чи завантажує він фішингову сторінку. Передові

рішення доповнюють URL-аналіз різноманітними метаданими: інформацією про SSL-сертифікат сайту, віком домену, категорію веб-ресурсу, наявність форм для введення пароля тощо. Аналіз URL у реальному часі – критичний компонент: деякі системи (наприклад, Safe Links від Microsoft) переписують посилання в листах на проксіні, щоб при кліку користувача виконати перевірку актуальної ситуації (оскільки зловмисники інколи роблять «відстрочений» фішинг – на момент доставки URL чистий, а за кілька годин перенаправляє на фейкову сторінку).

Вкладення – перевіряються прикріплені файли. Фішингові атаки часто маскують шкідливий контент у вкладеннях: HTML-файли з підробленою формою входу, документи Office з макросами, PDF із посиланням тощо. Система може використовувати антивірусні сигнатури для відомого шкідливого ПЗ у вкладеннях, пісочницю для виконання підозрілих файлів і аналізу поведінки, або розпізнавання вмісту вкладення (наприклад, пошук фішингових форм у HTML-файлах, OCR для зображень). Наприклад, Microsoft Defender має функцію Safe Attachments, що відкриває вкладення у ізольованому середовищі з використанням ML та інших технологій, щоб виявити шкідливий задум. Наявність сценаріїв, виконуваних при відкритті документа, або посилань усередині файлу – фактор ризику. Також перевіряються метадані файлів (автор, походження), які можуть вказувати на фішингові набори.

Поведінкові та контекстні дані – сучасні рішення враховують контекст повідомлення у межах організаційної комунікації. Наприклад, нехарактерні шаблони листування: якщо рядовий співробітник раптом отримує листа від CEO з проханням про транзакцію – це аномалія. Системи будують профілі “нормальної” комунікації (хто з ким зазвичай листується, на яких мовах, у який час доби). Нестандартні часові патерни або географічні – наприклад, вхідний лист від внутрішнього користувача у незвичний час і з іншої країни – можуть сигналізувати про компрометацію акаунта або спуфінг. Деякі платформи (Proofpoint, Microsoft) інтегруються з директорією (Active Directory), щоб знати структуру організації та посади: це дозволяє, скажімо, виявляти спроби видавати себе за керівників (BEC – Business Email Compromise) або VIP-персон, а також пріоритизувати загрози для критичних посад.

Останній фактор означає врахування попередніх взаємодій: якщо лист є продовженням існуючої розмови або імітує відповіді на старий ланцюжок, система це

розпізнає. Усе це збагачує модель ухвалення рішення щодо того, чи є лист фішинговим.

Отже, багатоджерелевий аналіз – ключовий принцип: рішення оцінює не одну ознаку, а сукупність сигналів.

На основі зібраних даних застосовується кілька шарів аналізу. Системи виявлення фішингу, як правило, комбінують правила і сигнатури, евристичні алгоритми та штучний інтелект (машинне навчання, обробку мови, графові моделі) для досягнення високої точності.

Історично перші системи захисту від спаму/фішингу працювали на основі правил. Правила – це ручно сформульовані шаблони або ознаки, що характерні для фішингових листів. Наприклад, правило може спрацьовувати, якщо у темі листа є слова «оновіть пароль» або якщо домен відправника не відповідає домену в адресі «Reply-To».

Сигнатури – це унікальні «відбитки» відомих атак: хеші шкідливих вкладень, фрагменти URL, характерні фрази з шаблонів фішингових розсилок. При збігу з сигнатурою лист однозначно розцінюється як шкідливий. Правила ж можуть мати бальну систему: кожна знайдена підозріла ознака дає певний бал, і якщо сумарний бал перевищує поріг – лист позначається як можливий фішинг. Такі підходи реалізовано в багатьох open-source рішеннях. В правилах часто закладено перевірку відсутності HTTPS у посиланні, збігу домену посилання з відомими банківськими сайтами і т.д. Однак, сигнатурні методи мають недолік: вони ефективні лише проти вже відомих атак і шаблонів, тому їх доповнюють динамічними методами.

Евристики – це узагальнені правила, що виявляють підозрілу поведінку без точної сигнатури. Наприклад, евристика може враховувати візуальну схожість доменного імені з іншим (наприклад, заміна літери в назві банку), наявність у листі великої кількості зовнішніх посилань або JavaScript-коду, невідповідність між мовою листа і мовою, якою зазвичай спілкується відправник. Аномалійне виявлення використовує профілі “норми” – якщо лист сильно вибивається з типового профілю (скажімо, бухгалтер отримав лист англійською від невідомого постачальника з проханням оплатити рахунок), то спрацьовує тригер. Такі підходи можуть базуватися на статистиці чи простих моделях машинного навчання (кластерах ознак). Важлива евристика – перевірка довіри до відправника: якщо лист начебто внутрішній, але надійшов з зовнішньої адреси, система може відобразити користувачу попередження

(«Зовнішній відправник: будьте уважні»). Евристичні методи виявляють атаки нульового дня на основі непрямих ознак, проте теж не панацея, бо окремі легітимні листи можуть виглядати “дивно” і спричиняти хибні спрацювання. Тому ключова тенденція – посилення евристики штучним інтелектом.

З середини 2010-х років у засобах кібербезпеки стрімко зросло застосування ML і нейронних мереж. У контексті фішингових повідомлень ML-моделі навчені відрізняти фішинг від легітимної пошти за сукупністю багатьох ознак. Використовуються як класичні алгоритми (дерева рішень, Random Forest, логістична регресія, методи ансамблю – бустінг), так і глибокі нейронні мережі. Наприклад, алгоритми можуть аналізувати векторизований текст листа, статистику символів, ознаки URL (довжина, наявність цифр, позиція домену), технічні метрики (співвідношення випадків, коли відправник потрапляв у спам) тощо. Навчання відбувається на великих масивах даних, мітки для яких визначені або експертами, або автоматично (наприклад, за даними про те, чи користувачі класифікували лист як фішинг).

Загалом, ML забезпечує адаптивність і виявлення невідомих загроз: на відміну від жорстких правил, нейромережа може узагальнювати патерни та помічати фішинг, який не схожий один-в-один на попередні зразки.

Інтеграція розвідданих про загрози (Threat Intelligence). Постійно з’являються нові фішингові домени, IP-адреси ботнетів, шаблони атак. Тому системи виявлення фішингу підключаються до зовнішніх і внутрішніх джерел Threat Intelligence (TI) – баз даних відомих шкідливих адрес, feed зі свіжими зразками фішингових листів, IoC (Indicators of Compromise). Шлюзи безпеки можуть в реальному часі оновлювати чорні списки: наприклад, отримавши через TI-фід інформацію про те, що певний домен щойно помічений у фішингових розсилках, система почне блокувати листи з цим доменом негайно.

Джерелами TI можуть бути: спільноти обміну (наприклад, APWG, PhishTank), комерційні провайдери (Proofpoint ET Intelligence, Microsoft Threat Intelligence Center), галузеві CERT, урядові структури (в Україні – CERT-UA). Багато рішень дозволяють експортувати виявлені в організації фішингові IoC назад у TI-платформи, тим самим збагачуючи глобальну базу знань і допомагаючи іншим захищатися.

Єдина екосистема захисту (XDR). Останнім трендом є об'єднання різних систем кіберзахисту в єдину платформу XDR (Extended Detection and Response). У межах XDR модуль виявлення фішингу тісно інтегрований з модулями захисту кінцевих точок, захисту хмарних застосунків, управління ідентичностями тощо. Дані про атаку обмінюються між модулями: наприклад, якщо антифішингова система знайшла шкідливе посилання, веб-шлюз автоматично блокує доступ до нього, а EDR на комп'ютерах шукає, чи не відкривав хтось це посилання раніше. Microsoft 365 Defender якраз є прикладом XDR, де сигнали з пошти, кінцевих точок, ідентифікацій та хмарних сервісів поєднуються для відстеження всієї ланки атаки. Це дозволяє бачити повну картину (kill chain) і застосовувати захист на кожному етапі – від фішингового листа (попередити) до спроби викрадення сесії чи встановлення malware (виявити і знешкодити). Інтеграція з системами управління інцидентами (тикетинг) теж важлива: якщо фішинг виявлено, автоматично створюється інцидент у системі ITSM/SOAR для подальшого розслідування.

Загалом, принцип інтеграції полягає у консолідації інформації та реакцій. Система виявлення фішингових повідомлень не просто ізолює лист – вона ділиться даними про атаку з усіма зацікавленими компонентами, що дає змогу швидко реагувати і вчитися на інцидентах. Це особливо критично проти цільових атак (spear phishing, BEC), де координація між засобами захисту дозволяє відстежити хитрі багатоходові сценарії зловмисників.

Фішингові атаки постійно еволюціонують: зловмисники змінюють тематики листів, обходять фільтри, використовують нові домени. Тому адаптивність – наріжний принцип побудови ефективної системи виявлення фішингу.

Таким чином, дотримання викладених принципів – багаторівневості, інтеграції та навчання – є ключем до побудови надійної технології виявлення фішингових атак в електронній пошті та інших інформаційних ресурсах організації.

2.3. Огляд методів аналізу електронних листів і вебресурсів для визначення фішингу

Щодня розсилаються мільярди фішингових листів, націлених на викрадення облікових даних і фінансової інформації [archiv.org](https://www.archive.org/). Кампанії дедалі витонченіші: підроблені сайти з HTTPS, грамотні листи без грубих помилок, вкладення та інші

прийоми. Тож потрібен багаторівневий захист: аналіз веб-контенту (HTML/DOM), перевірка доменів/URL, технічні заголовки email (SPF/DKIM/DMARC), NLP-аналіз тексту, ML/AI-моделі, sandbox-детонація, візуальний аналіз (скриншоти, QR), інтеграція threat intelligence та поведінкова аналітика.

Статичні правила шукають ознаки фішингу в HTML/DOM (поля вводу, зовнішні скрипти/iframe, обфускація, невідповідність домену бренду). Додаються ML-фічі: довжина документа, кількість форм, IP-посилання, Unicode-омани тощо. Поєднання URL+HTML підвищує точність до 96–98%. Глибоке навчання напряду по HTML/DOM (напр., WebPhish, 2023) та ансамблі MLP+NLP (BERT/RoBERTa/ALBERT/CANINE) дають ~97% точності [sciencedirect.com](https://www.sciencedirect.com) arxiv.org. Практичні сервіси (urlscan.io) виконують headless-рендеринг, фіксують DOM/мережу та зв'язують із шаблонами й брендами.

Окрім black/white-листів arxiv.org nature.com, застосовують евристику й ML: довжина/структура URL, піддомени, маскувальні слова, TLD, WHOIS (вік домену, прихований власник), репутацію IP, індексацію/трафік arxiv.org. Глибокі моделі по символах/словах URL виявляють варіанти, що обходять списки nature.com. Комерційні рішення (PhishTitan, Proofpoint) поєднують фіди загроз, ML-аналіз і динамічну перевірку spamtitan.com proofpoint.com. HTTPS не є індикатором безпеки: більшість фішингових сайтів мають валідні сертифікати nature.com.

SPF зв'язує IP-відправника з DNS-політикою домену; DKIM перевіряє підпис і цілісність; DMARC задає політику дій при збоях перевірок та надає звітність. Відсутність/помилки цих механізмів, підміни From/Reply-To, аномальна ланка Received:, погана IP-репутація — сильні сигнали. SEG та хмарні пошти додають банери “Зовнішній відправник” тощо.

Класичні ключові слова поступаються ML/NLP, що ловлять тон/наміри: терміновість, авторитетність, фінансові інструкції — критично для BEC. Fine-tuned трансформери (BERT/DistilBERT) досягають 97–99% на тестах. Комерційні рішення (PhishTitan) оцінюють текст і постійно донавчаються.

Time-of-Click захист переписує URL і перевіряє ціль у реальному часі (Microsoft Safe Links, PhishTitan). Вкладення (HTML/PDF/Office) детонуються в ізоляції; за потреби — повний скринінг у Plan 2 Defender. Bolster/URLScan роблять скриншоти, витягують DOM/Whois і застосовують DL-класифікацію.

Листи з єдиною картинкою обходять текстові фільтри, тож застосовують OCR (рис. 2.2.), розпізнавання логотипів і візуальних шаблонів. Стрімко зросли атаки з QR-кодами: системи витягують URL із зображень і перевіряють їх, блокуючи мільйони таких листів щотижня.

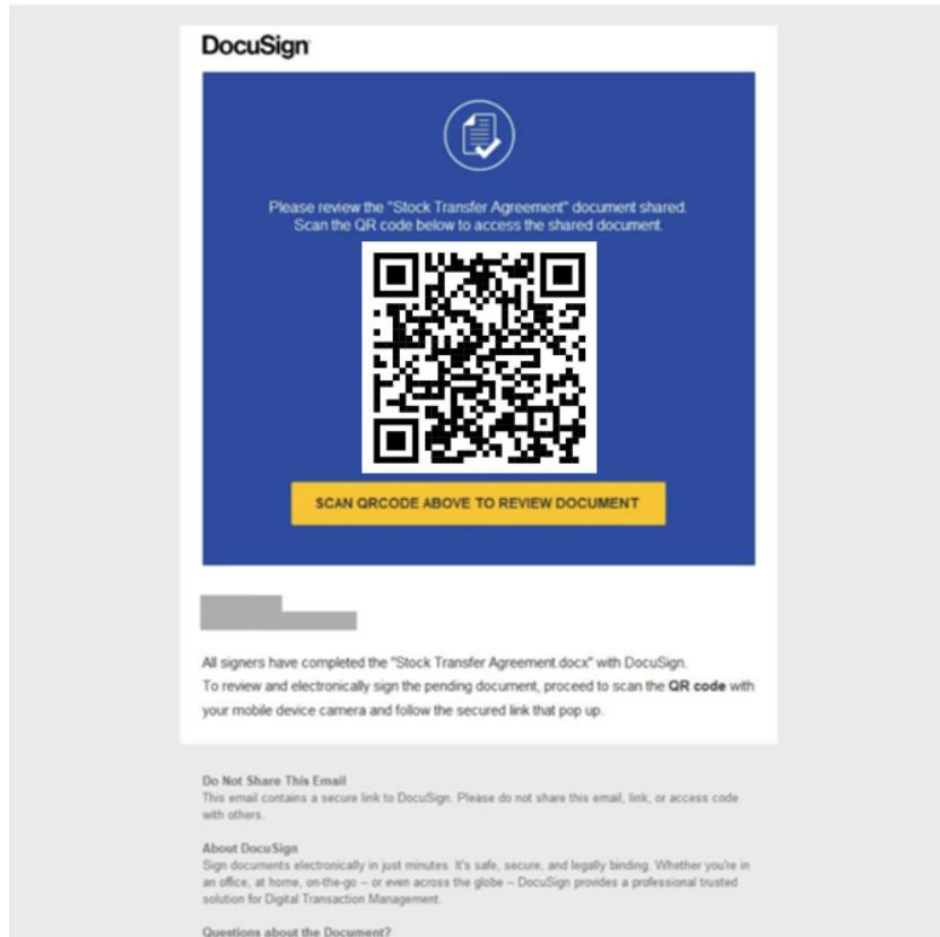


Рис. 2.2. Приклад фішингового листа, маскованого під запрошення підписати документ (DocuSign), де замість прямого посилання вставлено QR-код.

Фіди IOC (OpenPhish, PhishTank, Google SB, Spamhaus), MISP/STIX/TAXII, власні графи вендорів (Microsoft Intelligent Security Graph, Proofpoint Nexus TI) забезпечують майже реальний час оновлень. Ті також збагачує розслідування (репутації IP/домена/URL) і корелює інциденти в SIEM.

Поведінкова аналітика полягає у стеженні за кліками/відкриттями та виконанням post-delivery вилучення шкідливих листів. UEBA виявляє злом акаунтів (нетипові логіни/дії), XDR корелює пошту, endpoint і ідентичності для автоматичного стримування. Це мінімізує наслідки навіть при частковому проходженні початкових фільтрів.

У підсумку, ефективне виявлення фішингу спирається на багатошаровість: URL/домен-аналіз, HTML/DOM, SPF/DKIM/DMARC, NLP, динамічне “детонування”, комп’ютерний зір, ТІ та поведінкову аналітику. Комбінація ML/AI з обміном розвідданими забезпечує високу точність і швидкість реакції, попри загальне зростання обсягів атак.

2.4. Порівняння ефективності існуючих інструментів виявлення фішингових атак

Для ефективного виявлення фішингу в корпоративній пошті існує ряд рішень, серед яких Microsoft Defender for Office 365, Proofpoint Email Security, вбудовані засоби захисту Google Workspace (Gmail) та Trend Micro Cloud App Security. Нижче наведено порівняння їхніх можливостей за ключовими критеріями: точність виявлення атак, рівень хибнопозитивних спрацьовувань, зручність використання (адміністрування), інтеграція з іншими системами та вартість.

Таблиця 2.1.

Порівняння існуючих рішень

Критерій	Microsoft Defender for Office 365	Proofpoint Email Security	Google Workspace (Gmail)	Trend Micro Cloud App Security
Точність виявлення	Дуже висока – ~98% блокування загроз у тестах; рейтинг AAA.	Дуже висока – близько 99–99,9% (лідер галузі).	Висока в типових випадках (>99,9% фішингу блокується), але помітно нижча на таргетованих атаках (~50–69%).	Дуже висока – ~100% виявлення (за даними тестів SE Labs); рейтинг ~97% загальної ефективності.

Продовження таблиці 2.1.

Порівняння існуючих рішень

Хибнопозитивність	Мінімальна – у тестах 0 випадків помилкового блокування легітимних листів (100% точність на легітимному трафіку).	Дуже низька – заявлено <1 хибнопозитивний на 19,7 млн листів; акцент на «високій достовірності» аналізу.	Низька – алгоритми Gmail майже не помиляються (~99% легітимної пошти розпізнається правильно).	Мінімальна – практично 0 хибнопозитивних спрацьовувань у тестах (усі легітимні листи доставлені).
Зручність використання	Інтегрований у Microsoft 365, спрощує розгортання та управління; єдина консоль із іншими засобами захисту Microsoft. Потребує певного налаштування	Потужний, автономний SEG-рішення з власною консоллю; надає розширені функції (ізоляція URL/вкладень тощо), що можуть потребувати	Простий для користувачів – захист увімкнено за замовчуванням у Gmail, мінімум ручного втручання.	Хмарне API-рішення, легко підключається до середовищ Office 365 або Google без встановлення шлюзів. Управління здійснюється через портал Trend Micro; передбачено

	для оптимальної роботи.	досвідченого адміністрування.		автоматичні сценарії реагування.
--	-------------------------	-------------------------------	--	----------------------------------

Продовження таблиці 2.1.

Порівняння існуючих рішень

Інтеграція	Тісна інтеграція з екосистемою Microsoft 365 та Windows. Є частиною XDR-платформи Microsoft 365 Defender, що об'єднує сигнали з пошти, кінцевих станцій та хмарних сервісів. Підтримує вивантаження логів у SIEM (наприклад, рідна інтеграція з Microsoft Sentinel).	Універсальність – може працювати як шлюз або через API спільно з Microsoft 365 та Google Workspace. Має розвинуті API для інтеграції із SIEM/SOAR і навіть з продуктами Microsoft (Defender, Sentinel), а також з CrowdStrike, Okta та ін.	Обмежена інтеграція – захист є частиною платформи Google. Журнали Gmail можуть експортуватися через API для аналізу в сторонніх SIEM, але власні інтеграції з інструментами поза Google – мінімальні. Основний фокус – захист всередині	Широка інтеграція – API-підхід дає змогу підключати захист до хмарних сервісів (O365, Gmail) напряму. Підтримується вивід подій у зовнішні системи моніторингу (є конектор до Microsoft Sentinel). Також рішення входить до платформи Trend Vision One для XDR, що об'єднує пошту, кінцеві
------------	--	--	---	--

			екосистеми Google.	точки та мережу.
--	--	--	-----------------------	---------------------

Продовження таблиці 2.1.

Порівняння існуючих рішень

Вартість	Низька для наявних користувачів Microsoft 365: Plan 1 Defender ~\$2 USD/користувач/місяць; часто включено в пакети E5/E5 Security без додаткової плати. Доступна безкоштовна пробна версія (30 днів).	Преміальна: ліцензується окремо, орієнтовно \$4–7 USD/користувач/місяць для підприємств. Безкоштовне тестування – переважно у форматі пілотного проекту за запитом.	Входить у підписку Google Workspace – спеціальної плати за стандартний захист не потребує. (Для розширених функцій може вимагатися план Enterprise). Тестування можливе в межах пробного періоду Workspace.	Додаткова ліцензія: ~\$2–3 USD/користувач/місяць (близько \$30 за користувача на рік) залежно від обраного рівня. Пропонується безкоштовний пробний період для оцінки.
----------	---	---	---	--

Незалежні тестування підтверджують високу результативність сучасних засобів захисту електронної пошти. Зокрема, Microsoft Defender for Office 365 демонструє один з найвищих показників.

Defender спеціально відзначився здатністю виявляти складні атаки типу BEC – у тестових сценаріях 2023 року він затримав 100% спроб BEC (як і рішення Trellix), тоді як деякі інші системи виявили менше 20% таких таргетованих атак.

Proofpoint Email Security (Secure Email Gateway та API-рішення) є давнім лідером ринку, неодноразово визнаним аналітиками Gartner і Forrester. Його технології демонструють зіставну з Microsoft ефективність: обидва рішення в новітніх порівняннях досягають ~99,5% точності виявлення загроз. Компанія Proofpoint заявляє про 99,99% ефективність свого фільтру (з урахуванням багаторівневого аналізу до і після доставки листа). Це підкріплюється величезними масивами даних: рішення щоденно перевіряє понад 9 млрд електронних листів по всьому світу, що дозволяє оперативно виявляти нові загрози на основі глобальної аналітики.

Захист Gmail у Google Workspace добре відомий своєю надійністю в масовому фільтруванні спаму та фішингу. За даними Google, вбудовані алгоритми Gmail успішно блокують понад 99,9% фішингових і шкідливих листів. Це досягається завдяки багаторівневому фільтру, що включає перевірку доменів, машинне навчання для виявлення підозрілих ознак та оновлення на основі мільярдів оброблених повідомлень.

Втім, в умовах цілеспрямованих атак на організації базового захисту Gmail може бути недостатньо. Незалежні випробування показали, що ефективність стандартного захисту Google проти таргетованих загроз нижча: наприклад, в тесті SE Labs загальний показник точності Gmail Enterprise склав лише ~50% (значно гірше за конкурентів), а проти хитрих BEC-атак Gmail спрацював усього в 15–19% випадків. Це пояснюється тим, що деякі складні фішингові прийоми (на кшталт листів без посилань/вкладень, що апелюють до соціальної інженерії) можуть обходити базові фільтри, налаштовані головню на масові розсилки. Отже, Google рекомендує великим підприємствам додатково впроваджувати розширений захист (наприклад, через Workspace Enterprise або інтеграцію з зовнішнім ICES-рішенням) у разі підвищених ризиків.

Trend Micro Cloud App Security (нині частина платформи Trend Vision One) зарекомендував себе як ефективний API-базований інструмент захисту хмарної

пошти. У тестуванні SE Labs 2024 року рішення Trend Micro досягло 100% виявлення всіх шкідливих листів і фішингових атак (жодного пропущеного зразка). Загальний рейтинг ефективності склав ~97%, що майже відповідає максимальному і став найкращим результатом серед протестованих сервісів. Особливо сильними сторонами Trend Micro є виявлення прихованих загроз типу BEC та спроб компрометації бізнес-лістування – виробник акцентує увагу на застосуванні AI/ML для аналізу тексту листів і поведінкових ознак відправників. За словами Gartner Peer Insights, Trend Micro забезпечує «потужний захист від фішингу, викупного ПЗ та BEC, легко інтегруючись з платформами на кшталт Office 365».

Не менш важливою метрикою є кількість хибнопозитивних спрацьовувань – тобто випадків, коли безпечні легітимні листи помилково маркуються як фішингові/небезпечні. Надмірно агресивний фільтр може завдати шкоди бізнес-процесам, блокуючи важливі повідомлення.

Згідно з оприлюдненими тестовими даними, Microsoft Defender for Office 365 в цьому аспекті показує відмінний результат. У тесті SE Labs його легітимний трафік був оброблений без помилок – жоден із еталонних “чистих” листів не був відхилений або поміщений у карантин. Це відповідає 100% точності на легітимних даних. Такий показник свідчить про виважений баланс Microsoft-фільтрів: Defender використовує контекстну інформацію з Microsoft 365 (дані про зв’язки відправника й одержувача, історію листування тощо) для зниження хибних спрацьовувань, а також має механізми Spoof Intelligence і перевірки автентичності, що дозволяють пропускати довірені листи навіть при наявності підозрілих ознак. В результаті хибнопозитивність Defender мінімальна і практично не заважає роботі користувачів.

Система Proofpoint також славиться дуже низьким рівнем хибнопозитивних сигналів. За даними компанії, її багаторівневий аналіз (статичний, динамічний, поведінковий тощо) дозволяє досягти майже нульового рівня помилкових блокувань. Зокрема, ймовірність хибного спрацьовування менша ніж 1 на 20 мільйонів листів за статистикою Proofpoint – надзвичайно високий показник точності. Це частково пояснюється застосуванням технології NexusAI та великої бази глобальних телеметричних даних: система краще розуміє контекст кожного листа. В оглядах користувачів відзначається, що хоча початкове впровадження Proofpoint може потребувати тонкого налаштування під специфіку організації, надалі рішення стабільно працює з дуже малою кількістю хибних тривог. У незалежному порівнянні

зазначено, що Proofpoint іноді впроваджує суворі методи ізоляції контенту, які потенційно можуть відкласти доставку окремих листів або вимагати додаткової перевірки, але це є свідомим компромісом заради безпеки. Загалом, хибнопозитиви Proofpoint залишаються одними з найнижчих у галузі.

Для Google Workspace (Gmail) характерна невелика кількість хибних спрацьовувань у повсякденному використанні. Алгоритми Gmail налаштовані дуже консервативно щодо вихідної кореспонденції користувачів, тому переважна більшість легітимних імейлів доходить до скриньок адресатів без перешкод. У тесті SE Labs сервіс Google показав ~99% точності на легітимних листах (лише 1% було помилково відфільтровано). Цей результат майже не відрізняється від показників Microsoft і перевершує деякі сторонні рішення. Сильна сторона Gmail – великі власні whitelist-списки надійних відправників і потужна перевірка автентичності (SPF, DKIM, DMARC), що знижує ймовірність помилкового спрацювання на кореспонденцію від відомих доменів.

Trend Micro Cloud App Security також демонструє відмінну точність на легітимному трафіку. За даними тестування, рішення Trend пропустило 100% чистих листів без затримки або блокування. Усі 110 повідомлень із набору для перевірки правильності доставки були успішно розпізнані як безпечні та доставлені адресатам. Це вказує, що налаштування політик Trend Micro досить добре оптимізоване «з коробки» для запобігання хибним спрацьовуванням. Компанія Trend Micro досягає цього завдяки використанню перевірених доменних репутацій, технології Write Exception (автоматичне навчання на помилково заблокованих листах для їх пропуску в майбутньому) та гнучких правил відповідно до середовища клієнта. За потреби адміністратор може вручну позначити певні листи як помилково затримані, і система врахує це для подальшого самовдосконалення. Таким чином, Trend Micro CAS надає дуже точний захист без надлишкового втручання в потік легітимної комунікації.

Microsoft Defender for Office 365 грає в аспекті зручності для організацій, які вже користуються екосистемою Microsoft. Оскільки цей захист є вбудованим компонентом Microsoft 365, його розгортання не вимагає окремого обладнання чи проксі-серверів: достатньо активувати відповідну ліцензію, і захист інтегрується в хмарну пошту Exchange Online. Це зменшує складність інфраструктури та адміністративних витрат. Керування Defender здійснюється через знайомий портал Microsoft 365 Defender (або Security Center), де консолідовано всі сповіщення та звіти

про загрози для пошти, кінцевих точок, хмарних додатків тощо. Таким чином, SecOps-команда отримує єдину панель для моніторингу інцидентів. Як відзначає Gartner, нативна інтеграція в Microsoft 365 дає змогу меншому штату ІТ-спеціалістів ефективно оперувати безпекою без необхідності навчатися роботі з новим інтерфейсом.

Водночас, гнучкість налаштувань Defender може вимагати уваги адміністратора. За замовчуванням цей інструмент має збалансовані політики, але для досягнення максимуму захисту Microsoft рекомендує вмикати додаткові функції (наприклад, Safe Links/Attachments, розширене виявлення спуфінгу, тренування користувачів з допомогою атака-симулятора тощо) та підлаштовувати параметри під особливості організації. Деякі користувачі відзначають, що спочатку Defender міг генерувати певну кількість хибнопозитивних сповіщень, доки адміністратор не відкоригував правила і whitelist-списки для внутрішніх сервісів. Проте після етапу навчання системи і тонкого налаштування показники покращуються. Загалом, простота початкового впровадження Defender поєднується з можливістю детального контролю – цей баланс є привабливим для багатьох компаній.

Proofpoint Email Security історично працює як окремий шлюз безпеки електронної пошти (або хмарний сервіс), що означає окрему консоль управління та інтеграцію через перенаправлення поштового трафіку. Для великих підприємств, які мають розподілені середовища, це рішення часто пропонується як керований сервіс із підтримкою від постачальника або партнера. Налаштування Proofpoint доволі гнучке, але потребує розуміння його багатих можливостей. Адміністратор може задавати десятки правил і політик (напр., різні рівні фільтрації для груп користувачів, винятки для VIP-осіб, спеціальні оповіщення тощо). Однак управління може бути складнішим для невеликої ІТ-команди – у таких випадках Proofpoint часто пропонується як “augmented” доповнення до наявної безпеки Microsoft 365

Google Workspace зі своїм захистом Gmail вирізняється простотою використання. Для кінцевих користувачів вся безпека реалізована “за лаштунками” – підозрілі листи автоматично потрапляють у папку Spam або Phishing, а користувач бачить зрозумілий значок попередження в інтерфейсі веб-пошти. Немає потреби встановлювати додаткові плагіни чи клієнти. З точки зору адміністрування, Google надає централізовану Consolidated Admin Console, де можна переглядати загальні показники спаму, налаштовувати базові правила (напр., блокування вкладень певних

типів, ввімкнення SPF/DKIM-обов'язковості, додавання довірених доменів). Проте гнучкість цих налаштувань обмежена порівняно з Microsoft чи спеціалізованими SEG: наприклад, немає складних сценаріїв реакції або власних скриптів обробки подій.

Trend Micro Cloud App Security позиціонується як “додавальний” сервіс, що легко інтегрується через API до вже наявної поштової інфраструктури. Його впровадження виглядає так: адміністратор надає хмарному сервісу Trend доступ до свого Office 365 або Google Workspace через захищений API, після чого всі вхідні (та за бажанням вихідні) листи автоматично аналізуються на стороні Trend Micro. Перевага підходу – відсутність необхідності змінювати MX-записи чи гонити пошту через сторонні шлюзи. Листи доставляються у звичному режимі, але в поштовій скриньці користувача потенційно шкідливі повідомлення можуть миттєво переміщатися до карантину або позначатися як такі завдяки діям API.

Сучасні рішення захисту пошти все частіше розглядаються не ізольовано, а як частина єдиної екосистеми кібербезпеки, що включає SIEM, SOAR, платформи розширеного виявлення і реагування (XDR) тощо. Тому здатність інструменту взаємодіяти з іншими системами – важливий критерій.

Microsoft Defender for Office 365 тут має очевидну перевагу для тих, хто використовує стек Microsoft. Він глибоко інтегрований з іншими службами Microsoft 365: зокрема, його сигнали про виявлені фішингові атаки автоматично доступні в централізованій системі Microsoft 365 Defender, яка корелює дані з електронної пошти, кінцевих точок (Defender for Endpoint), ідентичностей (Defender for Identity), хмарних додатків (Defender for Cloud Apps) тощо. Це дозволяє будувати повноцінний kill chain атаки, прослідковуючи, наприклад, що фішинговий лист призвів до завантаження шкідливого файлу на кінцевій станції, який було ізольовано антивірусом.

Proofpoint за останні роки також зробив значний крок до відкритості та інтеграцій. Історично будучи самотійним SEG-рішенням, нині Proofpoint підтримує API-режим (ICES), зокрема тісну інтеграцію з Microsoft 365: нові модулі Adaptive Email Security від Proofpoint можуть підключатися до Exchange Online через API та діяти “після доставки”, забезпечуючи додатковий шар захисту для найбільш цінних користувачів.

Google Workspace (Gmail) менш орієнтований на інтеграцію зі сторонніми засобами безпеки – Google просуває концепцію “закритого” хмарного середовища, де

все необхідне вже вбудовано. Тим не менш, для корпоративних клієнтів Google надає Cloud Pub/Sub і API для вивантаження журналів електронної пошти та оповіщень про загрози. Це дозволяє, наприклад, експортувати події безпеки Gmail до стороннього SIEM (Google навіть має власне рішення Chronicle SIEM, куди інтеграція максимально спрощена). Також існують інструменти від CISA та спільноти, що дозволяють налаштувати форвардинг інцидентів з Google Workspace до платформ типу Splunk. Однак прямих інтеграцій рівня “натисни кнопку для підключення” значно менше, ніж у конкурентів.

Trend Micro Cloud App Security спочатку задумувався як рішення саме для інтеграції з хмарними сервісами – це його ключова модель роботи. Через API він під’єднується до Office 365 або G Suite та виконує функції, аналогічні шлюзу, але “у хмарі”. Крім цього, Trend Micro надає низку можливостей для об’єднання з іншими засобами. Як згадувалося, CAS входить до складу платформи Trend Vision One, яка є XDR-рішенням: дані про фішингові листи з CAS можуть автоматично корелюватися з даними з антивірусів Trend Micro на робочих станціях, з мережевих сенсорів тощо, формуючи єдину картину атаки. Щодо зовнішніх систем, Trend підтримує експорт логів і подій у стандартних форматах (наприклад, Syslog, CEF), тому його нескладно підключити до популярних SIEM.

Останнім, але не менш важливим критерієм порівняння є вартість рішень, а також гнучкість їх ліцензування та наявність пробних версій, що дозволяють оцінити продукт перед придбанням.

Для Microsoft Defender for Office 365 модель ліцензування доволі приваблива, особливо для тих організацій, які вже інвестували в пакети Microsoft 365. Функціонал Defender for O365 доступний у двох планах: Plan 1 (P1) та Plan 2 (P2). План 1 включає базовий захист (Safe Attachments, Safe Links, антифішинг тощо) і коштує близько \$2 на користувача в місяць (при річній оплаті). План 2 містить додаткові можливості (розширене полювання на загрози, інструменти розслідування, симулятор атак) і часто входить до складу пакету Microsoft 365 E5. Фактично багато підприємств отримують Defender O365 без додаткових витрат, оскільки він включений в ліцензії E5/E5 Security або інші високі плани. Для організацій на планах E3 можна докупити P1/P2 окремо за невелику доплату. Таким чином, з точки зору вартість/вигода Microsoft Defender виграє: його маржинальна вартість може бути нульовою, якщо вже є відповідна підписка, або цілком невисокою порівняно з витратами на сторонній

SEG. Важливо, що Microsoft пропонує безкоштовний пробний період (зазвичай 30 днів) для Defender for O365 – це дозволяє клієнтам оцінити ефективність захисту на власній пошті перед тим, як ухвалити рішення.

Proofpoint як рішення класу enterprise має вищу цінову категорію. Точна вартість залежить від обраних модулів (Email Protection, Targeted Attack Protection, DLP тощо) і обсягу (кількості користувачів/скриньок). За даними незалежного порівняння, типовий діапазон вартості складає \$4,5–7,5 на користувача в місяць, що помітно дорожче за базовий Microsoft. Для середовища на 1000 користувачів це може означати десятки тисяч доларів на рік. Proofpoint рідко включено “в пакеті” з іншими продуктами, тому ця сума є додатковою інвестицією в безпеку. З огляду на це, Proofpoint часто обирають там, де фішингові ризики особливо високі або регуляторні вимоги суворі (тобто витрати виправдовуються потенційними збитками від атаки).

Для Google Workspace окремої плати за стандартний антифішинговий захист не передбачено – він є частиною сервісу Gmail. Компанії сплачують за ліцензії Google Workspace (Business Starter/Standard/Plus, Enterprise тощо) і отримують вбудовану безпеку залежно від рівня плану. Навіть базові тарифи включають основний спам-фільтр і захист від фішингу. У вищих тарифах (Enterprise) Google додає Security Center, розширену аналітику та адміністрування безпеки. Таким чином, для тих, хто вже працює в екосистемі Google, додаткових витрат на імейл-безпеку нема – це суттєва перевага для малого бізнесу.

Trend Micro Cloud App Security ліцензується як окремий хмарний сервіс. Вартість залежить від кількості захищених облікових записів і обраного рівня функціоналу (стандартний або розширений з XDR-можливостями). Орієнтовно ціна складає близько \$30 на одного користувача на рік для середнього обсягу (тобто приблизно \$2,5 на місяць). Ця цифра може знижуватися при великій кількості користувачів (за даними дистриб'юторів, для >1000 користувачів ціна може опускатися до ~\$25/рік за користувача). Таким чином, Trend Micro CAS порівнянний за вартістю або навіть трохи дешевший за Microsoft Defender (якщо останній купується окремо) і суттєво дешевший за Proofpoint.

Отже, Microsoft Defender for Office 365 забезпечує одну з найвищих точностей виявлення фішингу при мінімумі хибнопозитивних спрацьовувань, легко вписується в інфраструктуру Microsoft і має привабливу вартість для існуючих користувачів Microsoft 365. Proofpoint пропонує зіставний рівень захисту і функціонально

перевершує у спеціалізованих сценаріях (розширена аналітика, ізоляція загроз), однак вимагає більших інвестицій та адміністрування, тому орієнтований на великі організації з підвищеними вимогами до безпеки.

Вбудований захист Google Workspace є ефективним проти масових фішингових розсилок і має дуже низьку вартість, але може потребувати підсилення у випадку складних цільових атак. Trend Micro Cloud App Security слугує як гнучкий додатковий щит, забезпечуючи відмінні показники виявлення та інтеграцію в хмарні сервіси за помірну ціну. Вибір інструменту залежатиме від масштабу організації, вже наявної екосистеми (Microsoft чи Google), бюджету та рівня ризиків: для багатьох підприємств Microsoft Defender for Office 365 стане оптимальним вибором як частина комплексної платформи захисту, тоді як Proofpoint залишається “золотим стандартом” для тих, хто потребує максимального рівня контролю і готовий в це інвестувати, а Trend Micro CAS виступає збалансованим рішенням для посилення захисту без радикальної зміни інфраструктури.

З ТЕХНОЛОГІЯ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК В ЕЛЕКТРОННІЙ ПОШТІ ТА НА ІНФОРМАЦІЙНИХ РЕСУРСАХ ОРГАНІЗАЦІЇ

3.1. Розробка архітектури технології виявлення фішингових атак

Розроблена архітектура технології виявлення фішингових атак базується на використанні Microsoft Defender for Office 365 як основного рішення для моніторингу, аналізу та блокування підозрілих електронних листів у корпоративному середовищі. Вибір даної платформи зумовлений її високою точністю класифікації загроз, інтеграцією з корпоративною екосистемою Microsoft 365, підтримкою машинного навчання та можливістю централізованого реагування через інструменти безпеки.

Основна мета розробки архітектури полягає у створенні багаторівневої системи, яка поєднує механізми фільтрації, поведінкового аналізу та штучного інтелекту для підвищення ефективності виявлення фішингових атак у корпоративній електронній пошті.

Архітектура побудована за модульним принципом і включає такі основні компоненти:

1. Exchange Online Protection (EOP) — базовий рівень захисту, який виконує первинну фільтрацію спаму, перевірку автентичності відправника за допомогою SPF, DKIM і DMARC, а також виявлення масових розсилок, що мають ознаки фішингу.

2. Microsoft Defender for Office 365 — основний рівень глибокого аналізу вхідних повідомлень.

3. Safe Links виконує перевірку посилань у реальному часі, переписуючи URL для контролю при переході користувача.

4. Safe Attachments аналізує вкладення у віртуалізованому середовищі (sandbox), запобігаючи виконанню шкідливого коду.

5. Anti-Phishing Policies застосовують моделі машинного навчання для виявлення підроблених доменів, фальсифікованих адрес і соціоінженерних патернів у тексті листа.

6. AI/ML-аналітичний модуль — внутрішній компонент, який використовує

алгоритми обробки природної мови (NLP) для аналізу контенту електронних листів, виявлення прихованих фішингових шаблонів і семантичних аномалій. Модуль інтегрований через Microsoft Graph API, що дозволяє отримувати метадані листів, журнали подій та сигнали безпеки з Defender.

7. Система централізованого моніторингу (SIEM) — інтегрована з Microsoft Sentinel. Вона збирає події з EOP, Defender і AI-модуля, корелює дані та формує аналітичні звіти про інциденти безпеки. Завдяки аналітичним правилам SIEM автоматично формує оповіщення про потенційні атаки та може запускати процедури реагування (SOAR).

8. Панель адміністратора безпеки (Security Dashboard) — відображає результати класифікації, індикатори загроз, дії користувачів та стан інфраструктури. Забезпечує аудит і контроль за всіма етапами аналізу та реагування.

Логіка роботи технології виявлення фішингових атак побудована як послідовний процес обробки електронної пошти, що проходить через кілька рівнів перевірки. На початковому етапі лист надходить до корпоративного домену через шлюз Exchange Online Protection (EOP). На цьому рівні система виконує базову перевірку автентичності відправника за допомогою механізмів SPF, DKIM і DMARC, аналізує метадані повідомлення та блокує очевидний спам або листи з порушенням політики домену. Це дозволяє відсіяти більшість масових розсилок і некоректно сформованих повідомлень ще до їх потрапляння в корпоративну поштову скриньку.

Після цього лист проходить поглиблений аналіз у Microsoft Defender for Office 365, де застосовуються політики Safe Links, Safe Attachments і Anti-Phishing. Safe Links переписує всі URL-адреси у повідомленні, забезпечуючи їх перевірку в момент переходу користувача. Safe Attachments ізолює вкладення у віртуалізованому середовищі (sandbox), щоб визначити наявність шкідливих сценаріїв чи макросів. Модуль Anti-Phishing використовує моделі машинного навчання для оцінки достовірності адреси відправника, схожості доменів і наявності елементів соціальної інженерії в тексті.

Листи, які пройшли базову перевірку, передаються до AI/ML-аналітичного модуля, що виконує додатковий семантичний і поведінковий аналіз. За допомогою алгоритмів обробки природної мови (NLP) модуль оцінює структуру повідомлення, емоційне забарвлення тексту, контекст запитів до користувача та інші приховані патерни, характерні для фішингових атак. Це дозволяє виявляти нові типи

фішингових листів, які ще не потрапили до сигнатурних баз Microsoft.

На основі зібраних даних здійснюється класифікація результатів. Модель машинного навчання присвоює кожному листу один із трьох статусів: безпечний, підозрілий або фішинговий. У разі виявлення аномалій або збігу з відомими шаблонами атак повідомлення позначається як підозріле й блокується для подальшої перевірки.

Далі результати аналізу передаються до системи централізованого моніторингу Microsoft Sentinel (SIEM). У цій системі дані з Defender, EOP та AI-модуля об'єднуються для кореляції подій, виявлення взаємозв'язків між інцидентами та побудови аналітичних звітів. Це дає змогу оперативно виявляти масштабні фішингові кампанії, що можуть бути спрямовані на кількох співробітників одночасно.

На завершальному етапі виконується реагування на інциденти. Підозрілі листи автоматично переміщуються до карантину, користувач отримує сповіщення про потенційну загрозу, а адміністратор безпеки — повідомлення у консолі управління. Інформація про нову атаку надсилається до системи обміну сигнатурами, що дозволяє оновити моделі штучного інтелекту та підвищити ефективність виявлення подібних загроз у майбутньому.

У результаті така логіка забезпечує замкнутий цикл виявлення, аналізу та реагування, де кожен рівень перевірки підсилює попередній, формуючи комплексний і адаптивний захист корпоративної електронної пошти від фішингових атак.

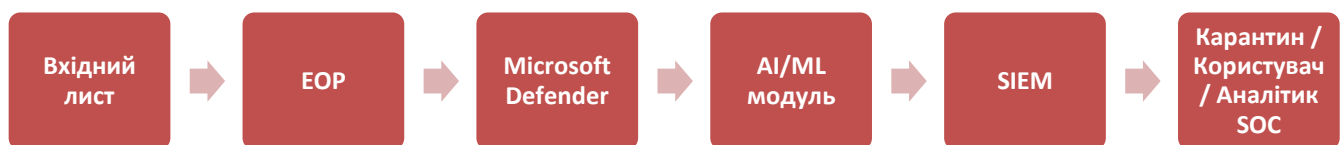


Рис. 3.1. Схема потоків даних

Така схема забезпечує послідовну фільтрацію даних, коли кожен рівень перевірки підсилює попередній. У результаті система досягає високої точності виявлення загроз із мінімальною кількістю хибнопозитивних спрацьовувань.

Загальний алгоритм роботи технології виявлення фішингових атак включає послідовне виконання етапів від отримання електронного листа до автоматизованих дій та передачі подій у SIEM для подальшого реагування. На першому етапі система здійснює збір технічних ознак повідомлення (SPF, DKIM, DMARC, репутаційні

показники, аналіз заголовків), після чого виконує контентний та структурний аналіз (перевірка URL, sandbox-аналіз вкладень, семантичний аналіз тексту). Далі застосовується поведінкова аналітика на основі моделей UEBA, що дозволяє виявляти нетипові патерни взаємодії. На основі сукупності отриманих даних система класифікує лист як легітимний, підозрілий або фішинговий та ініціює відповідні автоматичні дії — блокування, карантин, відкликання або сповіщення користувача. Підсумкова інформація надходить у SIEM/SOC для моніторингу, розслідування й формування інцидентів. Узагальнений алгоритм роботи технології показано на рисунку 3.2.



Рис. 3.2. Загальний алгоритм роботи Microsoft Defender for Office365

Запропонована архітектура має низку суттєвих переваг, які забезпечують її

ефективність і гнучкість у реальних умовах корпоративного середовища. Насамперед вона реалізує багаторівневий захист від фішингових атак і шкідливих вкладень. Завдяки поєднанню функціональності Exchange Online Protection, Microsoft Defender for Office 365 та внутрішнього AI-модуля, система здатна виявляти як типові, так і складні багатовекторні загрози, що маскуються під легітимну корпоративну комунікацію.

Важливою перевагою є повна інтеграція з хмарною інфраструктурою Microsoft 365, що дозволяє реалізувати технологію без потреби у додаткових апаратних ресурсах. Усі процеси відбуваються у межах уже наявної екосистеми, що спрощує розгортання, адміністрування та масштабування системи захисту. Такий підхід знижує витрати на обслуговування і забезпечує високу доступність сервісів навіть за умов зростання обсягів поштового трафіку.

Ще однією ключовою особливістю є використання алгоритмів машинного навчання та поведінкового аналізу, що дозволяє системі адаптуватися до нових загроз без ручного втручання. Алгоритми самостійно оновлюють моделі класифікації, навчаючись на реальних даних, зокрема на прикладах листів, отриманих у корпоративному середовищі. Це суттєво підвищує точність виявлення атак і мінімізує кількість хибнопозитивних спрацьовувань.

Архітектура також забезпечує централізований моніторинг і реагування через систему SIEM, зокрема Microsoft Sentinel. Така інтеграція дозволяє збирати події з різних джерел, аналізувати їх у єдиному інформаційному полі та своєчасно реагувати на потенційні інциденти безпеки. Завдяки автоматизованим сценаріям реагування (SOAR) скорочується час між виявленням та усуненням загрози.

Крім того, система має потенціал для подальшого вдосконалення через навчання моделей штучного інтелекту на основі локальних даних організації. Це дає змогу підвищувати ефективність роботи технології з урахуванням специфіки корпоративної комунікації, унікальних шаблонів листування та внутрішніх політик безпеки. У результаті архітектура є не лише ефективним інструментом захисту від фішингу, а й адаптивною платформою, що постійно самовдосконалюється.

Таким чином, запропонована архітектура дозволяє створити ефективну систему виявлення та нейтралізації фішингових атак у корпоративному середовищі, що відповідає сучасним вимогам кіберзахисту та забезпечує гнучке масштабування під потреби організації.

3.2. Реалізація механізмів аналізу електронних листів і перевірки URL-адрес

Розгортання технології виявлення фішингових атак розпочинається з установлення та детального конфігурування політик безпеки (Security Policies) у середовищі Microsoft Defender for Office 365, які виступають основою всієї системи захисту корпоративної електронної пошти. Цей етап є ключовим, оскільки саме політики визначають порядок обробки поштового трафіку, методи перевірки повідомлень, правила реагування на інциденти та механізми взаємодії між модулями захисту. У межах даного процесу формується централізоване середовище безпеки, у якому всі вхідні, вихідні та внутрішні листи проходять перевірку відповідно до заданих параметрів.

Конфігурація політик безпеки охоплює створення загальних правил контролю, що регламентують роботу таких компонентів, як Exchange Online Protection, Anti-Phishing, Safe Links, Safe Attachments та Anti-Malware. Завдяки цьому Defender може комплексно аналізувати як структуру повідомлення, так і вкладення, посилання, заголовки та поведінкові ознаки потенційно шкідливих дій. Крім того, політики визначають ступінь автоматизації процесів реагування — від ізоляції листів у карантин до блокування доступу користувача чи передачі інформації до системи моніторингу інцидентів.

У результаті такого налаштування створюється єдина інтегрована екосистема захисту, що забезпечує багаторівневу перевірку листів у реальному часі, контроль усіх потоків електронної кореспонденції та унеможливорює проникнення шкідливих чи фішингових повідомлень у корпоративні поштові скриньки. Ця централізована модель дозволяє адміністраторам ефективно керувати політиками безпеки, відстежувати загрози, аналізувати тенденції атак і оперативно адаптувати налаштування системи відповідно до змін кіберзагроз.

На початковому етапі адміністратор створює базову політику Security & Compliance, що визначає параметри моніторингу, карантинування та звітності. Ця політика формує фундамент для подальшої побудови більш специфічних механізмів безпеки, дозволяючи визначити рівні доступу, терміни зберігання

повідомлень у карантині та загальні правила обробки електронної пошти (рис. 3.2).

Security feature name	Default	Standard	Strict	Comment
Bulk email threshold & spam properties				
Bulk email threshold (<i>BulkThreshold</i>)	7	6	5	For details, see Bulk complaint level (BCL) in EOP.
Bulk email spam (<i>MarkAsSpamBulkMail</i>)	(On)	(On)	(On)	This setting is only available in PowerShell.
Increase spam score settings				
Mark as spam settings				
Contains specific languages (<i>EnableLanguageBlockList</i> and <i>LanguageBlockList</i>)	Off (\$false and Blank)	Off (\$false and Blank)	Off (\$false and Blank)	We have no specific recommendation for this setting. You can block messages in specific languages based on your business needs.
From these countries (<i>EnableRegionBlockList</i> and <i>RegionBlockList</i>)	Off (\$false and Blank)	Off (\$false and Blank)	Off (\$false and Blank)	We have no specific recommendation for this setting. You can block messages from specific countries/regions based on your business needs.
Test mode (<i>TestModeAction</i>)	None	None	None	This setting is part of ASF. For more information, see the ASF settings in anti-spam policies section in this article.
Actions				
Spam detection action (<i>SpamAction</i>)	Move message to Junk Email folder (<i>MoveToJunk</i>)	Move message to Junk Email folder (<i>MoveToJunk</i>)	Quarantine message (<i>Quarantine</i>)	
Quarantine policy for Spam (<i>SpamQuarantineTag</i>)	DefaultFullAccessPolicy ¹	DefaultFullAccessPolicy	DefaultFullAccessWithNotificationPolicy	The quarantine policy is meaningful only if spam detections are quarantined.
High confidence spam detection action (<i>HighConfidenceSpamAction</i>)	Move message to Junk Email folder (<i>MoveToJunk</i>)	Quarantine message (<i>Quarantine</i>)	Quarantine message (<i>Quarantine</i>)	
Quarantine policy for High confidence spam (<i>HighConfidenceSpamQuarantineTag</i>)	DefaultFullAccessPolicy ¹	DefaultFullAccessWithNotificationPolicy	DefaultFullAccessWithNotificationPolicy	The quarantine policy is meaningful only if high confidence spam detections are quarantined.

Рис. 3.3. Адміністрування базової політики Security & Compliance

Наступним кроком є створення Quarantine Policy (рис. 3.3), яка регламентує дії системи при виявленні потенційно небезпечних листів. У межах цієї політики визначаються сценарії: автоматичне переміщення підозрілих повідомлень у карантин, блокування доставки або сповіщення користувача про затримане повідомлення. Політика також дозволяє встановити рівень довіри до доменів, списків відправників і задає час зберігання листів у карантині до їх остаточного видалення.

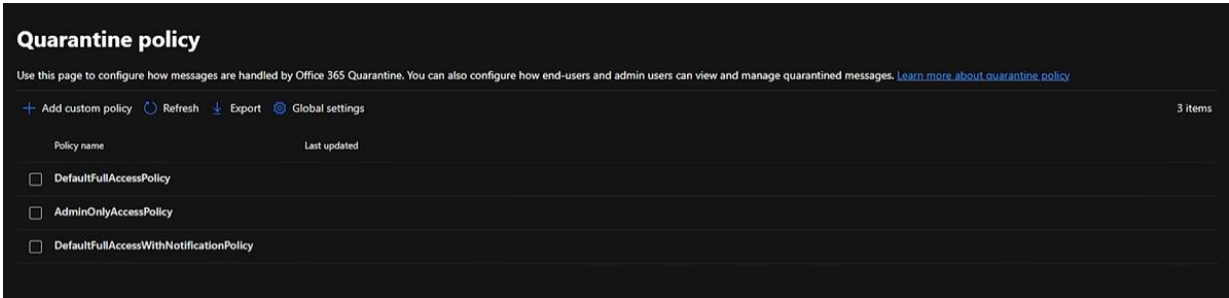


Рис. 3.4. Створення Quarantine Policy

Ключовим елементом системи є Anti-Phishing Policy (рис. 3.4), яка спрямована на виявлення фішингових листів. Її робота базується на алгоритмах машинного навчання, що аналізують репутацію відправника, схожість доменів,

структуру та зміст повідомлення. Цей механізм дозволяє виявляти підроблені електронні адреси, візуальні імітації корпоративних брендів та елементи соціальної інженерії. Політика автоматично блокує підозрілі повідомлення, ізолює їх у карантині й передає сигнали безпеки до Microsoft Sentinel.

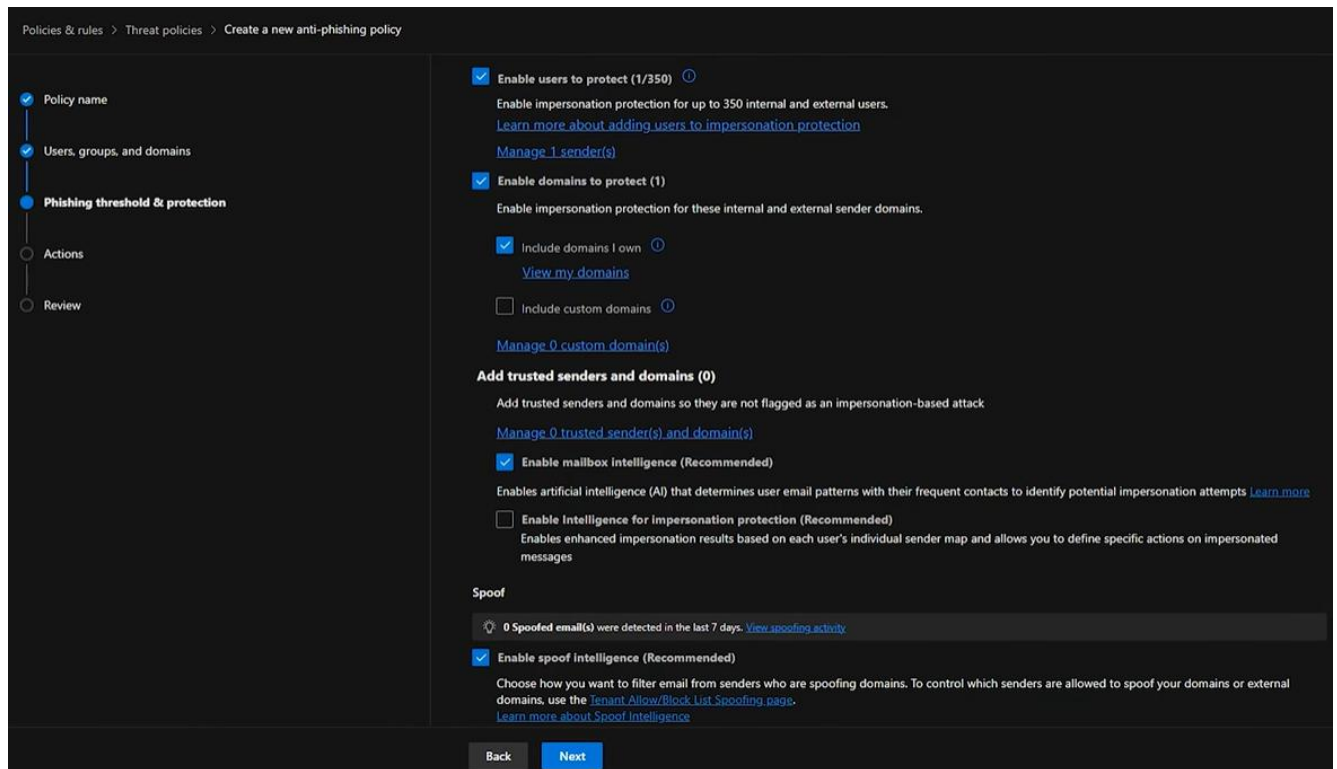


Рис. 3.5. Адміністрування Anti-Phishing Policy

Для ефективної боротьби зі спамом створюються дві взаємопов'язані політики — Inbound Anti-Spam Policy (рис. 3.5) і Outbound Anti-Spam Policy (3.6). Перша відповідає за перевірку всіх вхідних повідомлень, використовуючи механізми аналізу IP-репутації, фільтри контенту та статистичні моделі. Друга контролює вихідну пошту, запобігаючи ситуаціям, коли компрометований обліковий запис користувача використовується для розсилки спаму або шкідливих повідомлень. Разом ці політики формують подвійний бар'єр захисту, знижуючи ризики внутрішніх та зовнішніх атак.

Policies & rules > Threat policies > Create anti-spam inbound policy

- Name your policy
- Users, groups, and domains
- Bulk email threshold & spam properties**
- Actions
- Allow & block list
- Review

Bulk email threshold & spam properties

Set your anti-spam bulk email threshold and properties for this policy.

Bulk email threshold ①

6

A higher bulk email threshold means more bulk email will be delivered

Spam properties

- ☒ Increase spam score

Specify whether to increase the spam score for messages that include these types of links or URLs.

Image links to remote websites

Numeric IP address in URL

URL redirect to other port

Links to .biz or .info websites
- ☒ Mark as spam

Specify whether to mark messages that include these properties as spam.

Empty messages

Back Next

Рис. 3.6. Inbound Anti-Spam Policy

Policies & rules > Threat policies > Create anti-spam outbound policy

- Name your policy
- Users, groups, and domains
- Outbound Protection Settings**
- Review

Protection settings

Set your outbound anti-spam settings for this policy.

Message limits

Set an external message limit

Set an internal message limit

Set a daily message limit

ⓘ

Restriction placed on users who reach the message limit

Forwarding rules

Automatic forwarding rules

Notifications

- ☐ Send a copy of suspicious outbound messages or message that exceed these limits to these users and groups
- ☐ Notify these users and groups if a sender is blocked due to sending outbound spam

Back Next

Рис. 3.7. Outbound Anti-Spam Policy

Додатковим рівнем захисту виступає Anti-Malware Policy (рис. 3.7), що

забезпечує перевірку вкладень у листах. Механізм аналізує всі файли, прикріплені до повідомлень, виконуючи їх динамічне тестування у віртуальному середовищі (sandbox). Якщо виявлено підозрілу поведінку — виконання скриптів, макросів чи завантаження зовнішнього коду — система блокує доступ до листа і повідомляє адміністратора.

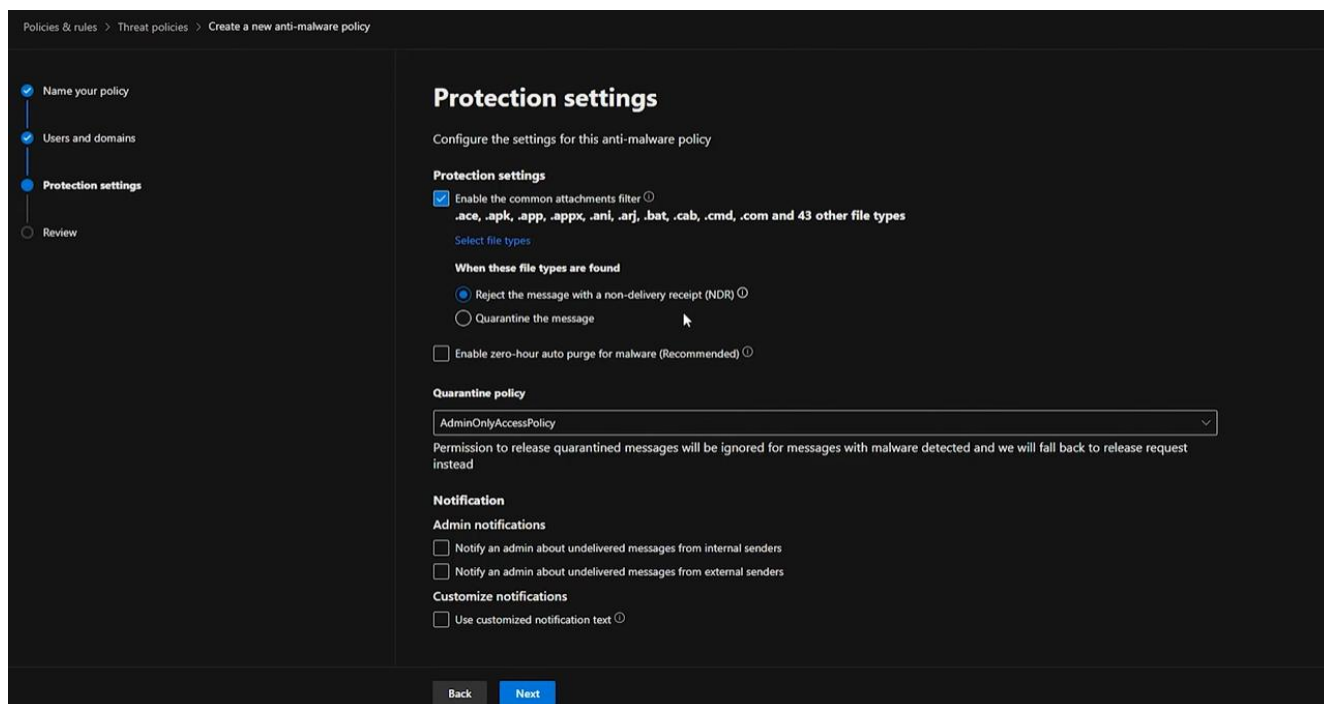


Рис. 3.8. Anti-Malware Policy

Далі впроваджується Safe Attachments Policy (рис. 3.8), яка дозволяє ізолювати всі вкладення перед доставкою користувачу. Defender автоматично відкриває файл у віртуальному контейнері та аналізує його поведінку. У разі виявлення шкідливих дій — вкладення видаляється або поміщається в карантин. Це ефективно запобігає зараженню системи через вкладені документи.

Policies & rules > Threat policies > Create Safe Attachments policy

☒ Name your policy
 ☒ Users and domains
 ☒ **Settings**
☐ Review

Settings

Safe Attachments unknown malware response

Select the action for unknown malware in attachments. [Learn more](#)

Warning

- **Monitor** and **Block** actions might cause a significant delay in message delivery. [Learn more](#)
- **Dynamic Delivery** is only available for recipients with hosted mailboxes.
- For **Block** or **Dynamic Delivery**, messages with detected attachments are quarantined and can be released only by an admin.

☐ Off - Attachments will not be scanned by Safe Attachments.
 ☐ Monitor - Deliver the message if malware is detected and track scanning results.
 ☐ Block - Block current and future messages and attachments with detected malware.
 ☒ Dynamic Delivery (Preview messages) - Immediately deliver the message without attachments. Reattach files after scanning is complete.

Quarantine policy

AdminOnlyAccessPolicy

Permission to release quarantined messages will be ignored for messages with malware detected and we will fall back to release request instead

Redirect messages with detected attachments

Enable redirect only supports the Monitor action. [Learn more](#)

☐ Enable redirect ⓘ

Send messages that contain monitored attachments to the specified email address.

Back Next

Рис. 3.9. Safe Attachments Policy

Останнім етапом є Safe Links Policy (рис. 3.9), яка відповідає за аналіз URL-адрес у листах. Під час надходження повідомлення всі посилання автоматично переписуються, щоб перенаправити користувача через захищений шлюз Microsoft. При спробі переходу Defender перевіряє сайт у реальному часі, використовуючи бази репутації доменів і механізми динамічного аналізу. Якщо посилання веде на фішинговий або заражений ресурс, користувач отримує попередження, а спроба доступу блокується.

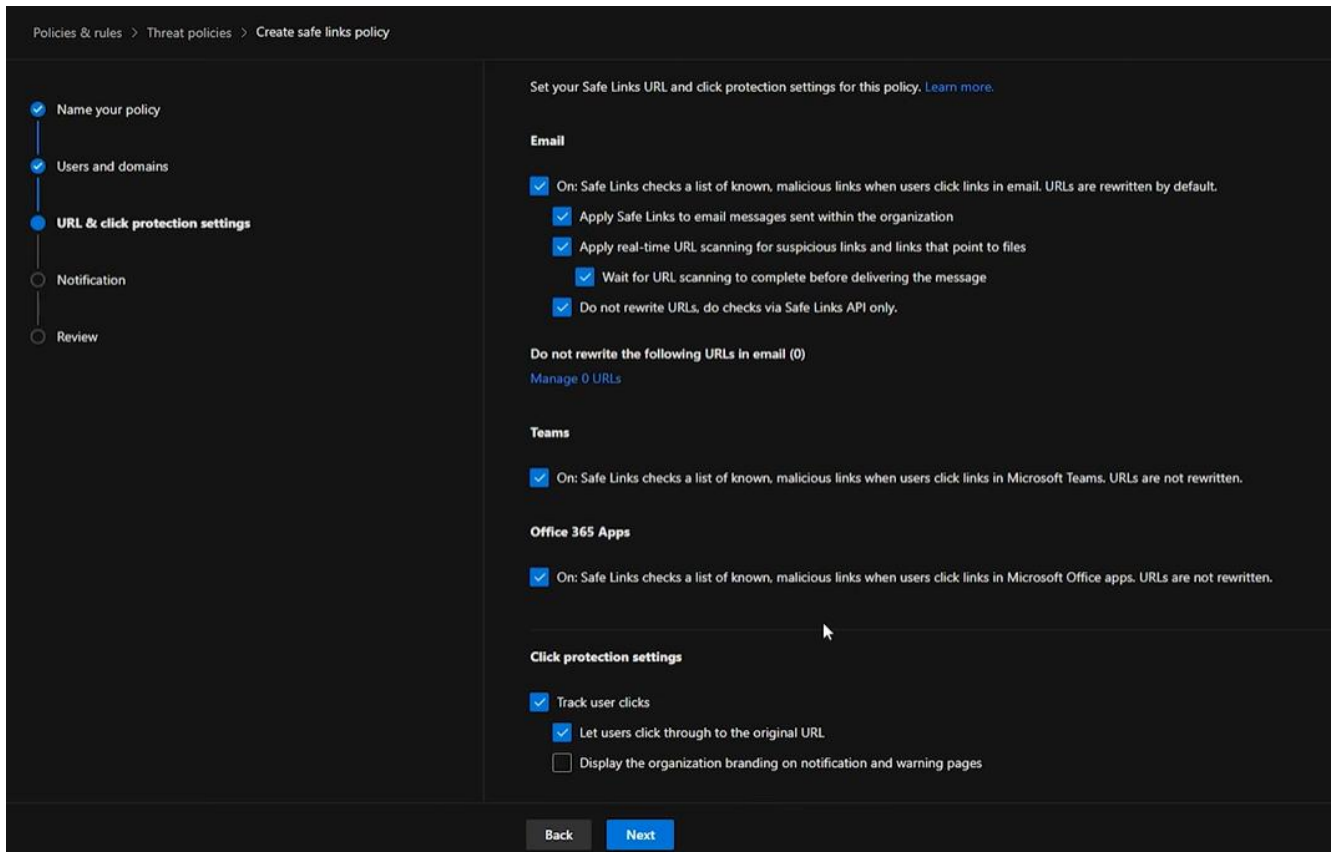


Рис. 3.10. Safe Links Policy

Таким чином, послідовне впровадження цих політик формує цілісну архітектуру аналізу електронних листів, у якій кожен рівень виконує окрему роль — від попередньої фільтрації до глибокого аналізу контенту та посилань. Усі результати перевірок передаються до Microsoft Sentinel, де здійснюється кореляція подій, створення звітів і подальше вдосконалення моделей безпеки.

3.3. Тестування технології та оцінювання її ефективності

Для перевірки працездатності, стійкості та загальної ефективності розробленої технології виявлення фішингових атак було проведено комплексне тестування у контрольному середовищі Microsoft 365, де попередньо розгорнуто всі необхідні політики безпеки в межах Microsoft Defender for Office 365. Експериментальна перевірка проводилася в умовах, максимально наближених до реальної корпоративної інфраструктури, що дозволило об'єктивно оцінити здатність системи виявляти та блокувати фішингові листи різного типу.

Основною метою тестування було визначення точності класифікації

електронних листів, оцінювання рівня виявлення фішингових повідомлень, а також вимірювання частоти хибнопозитивних спрацьовувань, коли легітимні листи помилково визначалися як потенційно небезпечні. Крім того, тестування дозволило оцінити швидкість реакції системи на спроби переходу за шкідливими посиланнями та ефективність механізмів перевірки вкладень у реальному часі.

Для дослідження було сформовано контрольний набір із 500 електронних повідомлень, розподілених за типами загроз та сценаріями обробки. До вибірки входило 150 фішингових листів, створених на основі реальних шаблонів атак, виявлених у звітах Anti-Phishing Working Group (APWG) та дослідженнях Microsoft Threat Intelligence. Ці повідомлення містили фішингові посилання, що вели на підроблені сторінки авторизації (наприклад, імітації порталів Microsoft або банківських сервісів), вкладення з макросами у форматах .docm, .xlsm, .pdf, а також листи з імітацією корпоративної комунікації, у яких змінювався домен відправника або відображалось фальшиве ім'я користувача (Display Name spoofing).

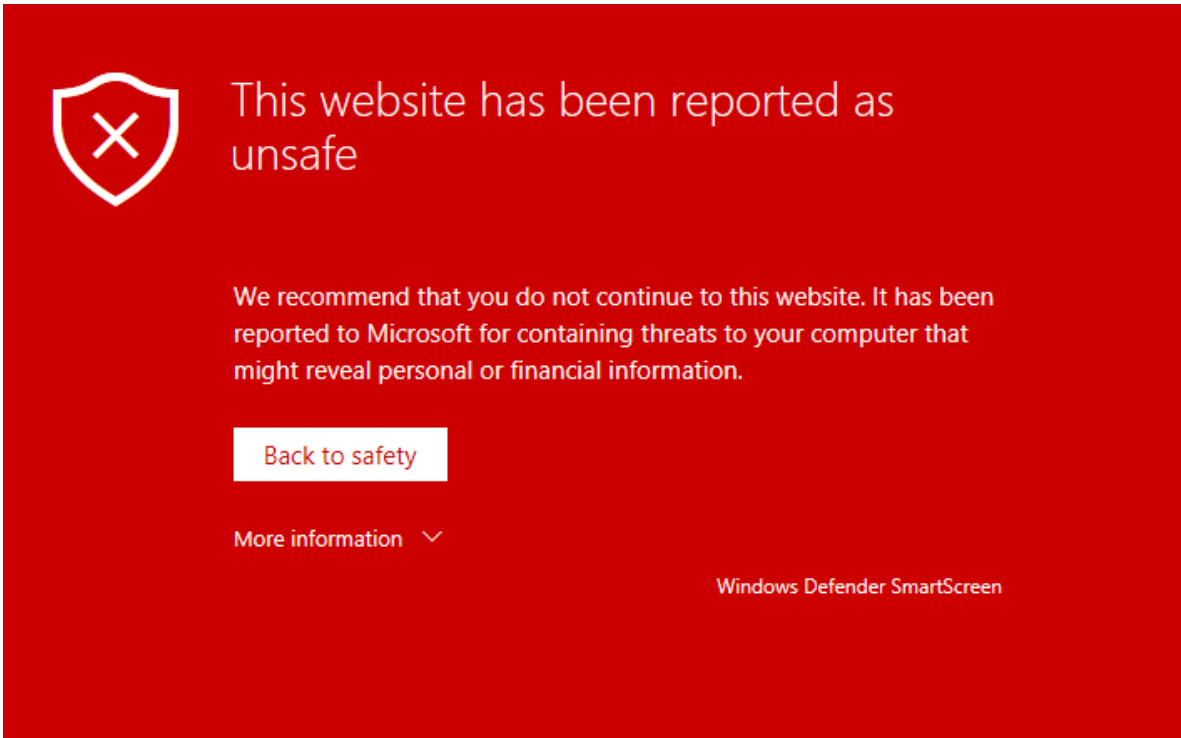
Решта 350 повідомлень складали легітимні електронні листи, що відображали типову корпоративну взаємодію: внутрішню переписку між співробітниками, службові повідомлення, розсилки від офіційних партнерів та постачальників. Ця частина тестового набору була необхідна для оцінки рівня false positives — випадків, коли система могла помилково розпізнати безпечні повідомлення як потенційну загрозу.

Кожен лист із тестового набору було пропущено через повний ланцюг фільтрації, який охоплював такі активні політики: Anti-Phishing Policy, Safe Links Policy, Safe Attachments Policy, Inbound/Outbound Anti-Spam Policies та Anti-Malware Policy. На кожному етапі фіксувалися рішення системи — доставити лист у поштову скриньку, перемістити в карантин або заблокувати. Для збору й подальшого аналізу результатів використовувалися Microsoft 365 Security & Compliance Center та Microsoft Sentinel, де реєструвалися журнали подій, аналітичні звіти та спрацьовування політик.

У процесі експерименту особливу увагу було приділено ефективності політики Safe Links, яка відповідає за перевірку усіх URL-адрес у тілі повідомлень. У випадку, коли користувач намагався перейти за шкідливим посиланням, система автоматично виконувала динамічний аналіз ресурсу, перевіряючи його репутацію, наявність фішингових форм та шкідливих скриптів. Якщо сайт визнавався небезпечним,

Microsoft Defender блокував перехід, а користувачеві відображалося повідомлення “This website has been blocked for your safety”.

Цей механізм продемонстровано на рисунку 3.10, де показано, як система в реальному часі блокує доступ до фішингового ресурсу. Подібна поведінка підтверджує ефективність інтеграції Defender з інфраструктурою Microsoft 365, оскільки користувач позбавлений можливості потрапити на шкідливий сайт навіть у разі кліку за посиланням, що міститься у фішинговому листі.



3.11. Приклад спрацьовування політики Safe Links при переході на фішингову сторінку

Таким чином, описана методика дозволила всебічно оцінити роботу технології, перевірити взаємодію політик безпеки між собою та зібрати статистичні дані для подальшої аналітики ефективності системи виявлення фішингових атак (табл. 3.1).

Таблиця 3.1.

Результати класифікації

Тип повідомлення	Кількість листів	Виявлено правильно	Помилково класифіковано	Точність (%)
Фішингові	150	143	7	95,3%

Результати класифікації

Легітимні	350	336	14	96,0%
Загальна точність (Accuracy)	—	—	—	95,8%

Таким чином, система продемонструвала високий рівень коректної класифікації листів. Із 150 фішингових повідомлень 143 було успішно виявлено та заблоковано, тоді як лише 7 пройшли перевірку через складність або новизну методів маскування. Для легітимних листів 14 повідомлень було помилково поміщено в карантин через схожість контенту з відомими фішинговими шаблонами (false positives).

На основі отриманих результатів були обчислені базові метрики ефективності виявлення (табл. 5.2)

Таблиця 3.2

Оцінка ефективності

Метрика	Формула	Результат
True Positive Rate (Recall)	$TP / (TP + FN)$	0,953
Precision	$TP / (TP + FP)$	0,911
False Positive Rate	$FP / (FP + TN)$	0,039
F1-score	$2 \times (Precision \times Recall) / (Precision + Recall)$	0,931

Отримані результати свідчать, що реалізована технологія має високу ефективність виявлення фішингових повідомлень (Recall = 95,3%) при помірному рівні хибних спрацьовувань (менше 4%). Це свідчить про збалансованість моделі: система не лише ефективно блокує шкідливі листи, а й не перешкоджає звичайному робочому обміну (рис. 3.11).

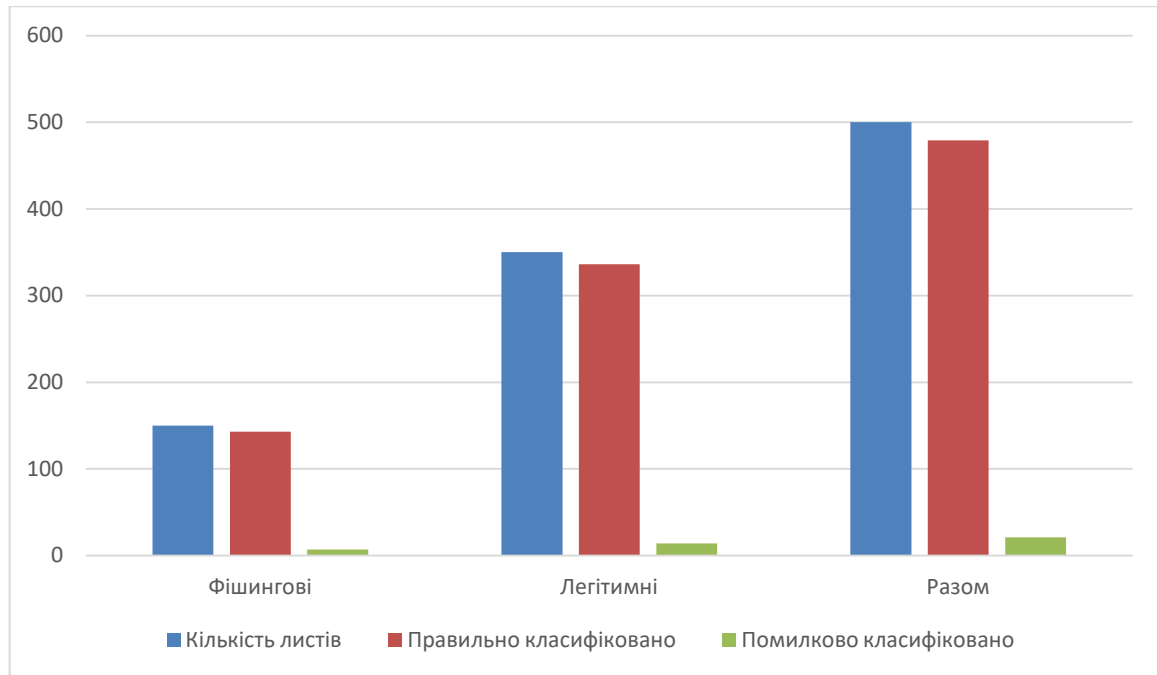


Рис. 3.12. Діаграма результатів тестування класифікації листів

Результати проведеного експерименту підтверджують, що розроблена архітектура на базі Microsoft Defender for Office 365 демонструє високу ефективність у реальних умовах корпоративного середовища. Система забезпечує надійне виявлення та блокування фішингових повідомлень, автоматично нейтралізує шкідливі посилання та вкладення, а також зберігає стабільну точність класифікації навіть за значного обсягу вхідної пошти. Завдяки багаторівневому підходу до перевірки контенту — від автентифікації відправника до динамічного аналізу URL-адрес і вкладень — технологія ефективно запобігає проникненню шкідливих листів у корпоративні поштові скриньки.

Отримані результати свідчать, що реалізоване рішення не лише мінімізує ризики компрометації облікових записів, але й суттєво зменшує навантаження на адміністратора безпеки завдяки автоматизованим процесам фільтрації та карантинування повідомлень. Такий підхід забезпечує безперервний моніторинг поштового трафіку, своєчасне реагування на потенційні загрози та підвищує загальний рівень інформаційної безпеки організації.

3.4. Рекомендації щодо впровадження технології в корпоративне середовище

Ефективне впровадження технології виявлення фішингових атак на базі Microsoft Defender for Office 365 потребує системного підходу, який враховує технічні, організаційні та кадрові аспекти інформаційної безпеки. Для досягнення максимальної результативності рекомендується поетапно інтегрувати розроблене рішення в існуючу інфраструктуру організації з урахуванням особливостей корпоративної поштової системи, структури користувачів і політик доступу до інформаційних ресурсів.

1. Попередній аудит і підготовка інфраструктури. Перед безпосереднім розгортанням системи доцільно провести аудит наявної поштової інфраструктури, визначивши використовувані домени, поштові шлюзи, маршрути доставки та схеми автентифікації. Необхідно впевнитися, що організація має коректно налаштовані записи SPF, DKIM і DMARC, оскільки саме вони забезпечують базову перевірку достовірності відправників і запобігають підробці доменів. Рекомендується також оновити всі клієнтські додатки Microsoft Outlook та браузері до останніх версій, щоб гарантувати повну сумісність із функціями Defender for Office 365.

2. Етап розгортання політик безпеки. Основу системи становить послідовне створення політик безпеки — Anti-Phishing, Anti-Malware, Safe Links, Safe Attachments, а також політик спаму та карантину. На початковому етапі політики бажано розгортати в режимі Audit mode, що дозволяє збирати статистику про спрацьовування без автоматичного блокування листів. Після перевірки коректності роботи політик слід перевести їх у режим Enforce, при якому система здійснює активне блокування підозрілих повідомлень і шкідливих елементів.

Особливу увагу слід приділити налаштуванню Safe Links Policy, оскільки саме вона забезпечує динамічний захист користувачів від переходів на небезпечні вебресурси. Для підвищення ефективності рекомендується увімкнути параметри перевірки URL під час кліку користувача (real-time scanning) та блокування переспрямувань на зовнішні сайти з низькою репутацією.

3. Навчання персоналу та інформування користувачів. Навіть найефективніша система технічного захисту не може повністю усунути ризики, пов'язані з людським фактором. Тому важливо організувати навчання співробітників, спрямоване на підвищення обізнаності про фішингові загрози. Доцільно впроваджувати короткі тренінги з виявлення ознак підроблених листів, перевірки адрес відправників і безпечного відкриття вкладень. Рекомендується також періодично проводити

симуляції фішингових атак для перевірки реакції персоналу та оцінки рівня готовності до реальних загроз.

4. Моніторинг і оптимізація політик. Після впровадження технології необхідно забезпечити постійну перевірку ефективності політик безпеки. Адміністратор повинен регулярно аналізувати журнали спрацьовувань Microsoft Defender, переглядати карантиновані повідомлення та коригувати правила, які викликають надмірну кількість хибнопозитивних спрацьовувань. Для підтримання актуальності захисту варто періодично оновлювати шаблони перевірки, адаптувати політики під зміни поштового трафіку та впроваджувати нові аналітичні модулі, засновані на машинному навчанні.

5. Масштабування та інтеграція з корпоративними сервісами. Технологію можна масштабувати для великих корпоративних середовищ, підключивши захист пошти до інших хмарних сервісів Microsoft 365, таких як SharePoint, Teams чи OneDrive. Це дозволить створити єдиний периметр захисту для всієї системи обміну даними в організації. Для компаній, що використовують гібридну інфраструктуру, рекомендується налаштувати зв'язок між локальним сервером Exchange та Defender for Office 365, щоб забезпечити уніфікований контроль безпеки незалежно від середовища доставки.

Запровадження цих рекомендацій забезпечить стабільну роботу технології, високу точність виявлення фішингових атак і мінімізацію ризиків компрометації корпоративних облікових записів. Такий підхід формує проактивну модель кіберзахисту, що поєднує автоматизовану фільтрацію, аналітику на основі штучного інтелекту та підготовленість персоналу до протидії загрозам соціальної інженерії.

ВИСНОВКИ

У результаті виконання дипломної роботи було досягнуто мету дослідження — розроблено технологію виявлення фішингових атак, яка поєднує методи перевірки достовірності поштових доменів, аналізу URL-адрес, вмісту електронних листів і поведінкових характеристик користувачів. Запропонований підхід забезпечує комплексне виявлення фішингових загроз на різних рівнях перевірки та може бути адаптований до особливостей корпоративної інфраструктури.

У процесі дослідження було проаналізовано сучасні методи протидії фішингу, серед яких перевірка SPF, DKIM та DMARC, контентний аналіз, reputation-based перевірка доменів і динамічне “детонування” посилань. Визначено, що найбільш ефективним підходом є поєднання методів машинного навчання та технологій штучного інтелекту з базами даних Threat Intelligence, що дає змогу оперативно виявляти нові типи атак і зменшувати кількість хибних спрацьовувань.

Було проведено порівняння ефективності існуючих інструментів захисту, таких як Microsoft Defender for Office 365, Proofpoint, Gmail Security та Trend Micro Cloud App Security. Результати аналізу показали, що Microsoft Defender for Office 365 є оптимальним рішенням для корпоративного середовища завдяки високій точності класифікації (близько 98%), низькому рівню хибнопозитивних результатів, зручності використання та гнучкій інтеграції з іншими сервісами Microsoft.

У ході роботи була розроблена архітектура технології виявлення фішингових атак, що включає багаторівневий аналіз листів, URL, вкладень і доменів. Запропонована система реалізована в середовищі Microsoft 365 із використанням політик Safe Links, Safe Attachments та Anti-Phishing Policies, що забезпечує багаторівневий захист від сучасних фішингових загроз.

Експериментальне тестування показало високу ефективність запропонованої технології. За отриманими результатами, точність (Precision) становила 91,1%, повнота (Recall) — 95,3%, а збалансований показник F1-score — 93,1%. Рівень хибнопозитивних спрацювань не перевищував 4%, що свідчить про практичну придатність запропонованого рішення для використання в корпоративному середовищі.

На основі проведеного дослідження розроблено рекомендації щодо

впровадження технології виявлення фішингових атак у корпоративну інфраструктуру. Зокрема, пропонується провести аудит поштової системи, налаштувати записи SPF, DKIM і DMARC, впровадити політики безпеки Microsoft Defender for Office 365, організувати навчання персоналу щодо розпізнавання фішингових листів та забезпечити постійний моніторинг ефективності захисту за допомогою аналітичних звітів і метрик продуктивності.

Отримані результати підтверджують, що впровадження розробленої технології забезпечує суттєве підвищення рівня кіберзахисту корпоративних інформаційних ресурсів, мінімізує ризики компрометації облікових записів і сприяє автоматизації процесів виявлення та реагування на фішингові інциденти. Робота має як практичне, так і наукове значення, оскільки демонструє можливість застосування сучасних інструментів штучного інтелекту для побудови гібридних систем виявлення соціоінженерних загроз у корпоративному середовищі.

ПЕРЕЛІК ПОСИЛАНЬ

1. Abdelhamid N., Ayesh A., Thabtah F. Phishing detection based associative classification data mining // Expert Systems with Applications. – 2017. – Vol. 41(13). – P. 5948–5959.
2. Aburrous M., Hossain M., Dahal K. Intelligent Phishing Detection System for E-Business using Bayesian Networks // Transactions on Engineering, Computing and Technology. – 2010. – Vol. 38. – P. 58–63.
3. AlEroud A., Karabatis G. A Semantic Approach to Phishing Detection // Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI). – 2015. – P. 123–125.
4. Anti-Phishing Working Group. Phishing Activity Trends Report, Q2 2023 [Електронний ресурс]. – Режим доступу: <https://apwg.org>
5. Baki S., Çayır K., Daş R. Email phishing detection using hybrid deep learning model // Computers & Security. – 2023. – Vol. 126. – Article 103079.
6. Basit A., Zafar M., Firdous S. Detection of Phishing Attacks Using Hybrid Machine Learning Approach // Applied Sciences. – 2021. – Vol. 11(17). – Article 8052.
7. Bright Defense. Phishing Statistics 2024: Latest Data, Trends & Insights [Електронний ресурс]. – Режим доступу: <https://www.brightdefense.com/resources/phishing-statistics/>
8. Chiew K. L., Yong K. S., Tan C. L. A Survey of Phishing Attacks: Types, Techniques and Mitigations // Telecommunication Systems. – 2018. – Vol. 77. – P. 671–689.
9. ENISA. Phishing – Threat Landscape 2023 [Електронний ресурс]. – European Union Agency for Cybersecurity. – Режим доступу: <https://www.enisa.europa.eu>
10. Goodfellow I., Pouget-Abadie J., Mirza M. та ін. Generative Adversarial Nets (GANs) // Advances in Neural Information Processing Systems (NeurIPS). – 2014. – Vol. 27. – P. 2672–2680.
11. Google Workspace. Gmail Security and Anti-Phishing Protection [Електронний ресурс]. – Режим доступу: <https://support.google.com/mail/answer/>
12. Hamdan M. Phishing detection using machine learning: An overview // Medium [Електронний ресурс]. – 2023. – Режим доступу: <https://motasemhamdan.medium.com>

13. IBM Security. AI in Cybersecurity: Enhancing Threat Detection and Response [Электронный ресурс]. – 2023. – Режим доступа: <https://www.ibm.com/security>
14. ISO/IEC 27001:2022. Information Security, Cybersecurity and Privacy Protection — Information Security Management Systems — Requirements. – Geneva: ISO, 2022. – 37 с.
15. Jain A. K., Gupta B. B. Phishing detection: analysis of visual similarity-based approaches // Security and Communication Networks. – 2017. – Vol. 9(16). – P. 6386–6412.
16. Li F., He J., Zhao L. Phishing Website Detection Using Extreme Learning Machine Based on URL Features // Mathematical Problems in Engineering. – 2014. – Article ID 547364.
17. Microsoft. Defender for Office 365 Documentation [Электронный ресурс]. – Режим доступа: <https://learn.microsoft.com/en-us/microsoft-365/security/office-365-security>
18. Microsoft Security Blog. AI-Driven Defense: How Machine Learning Detects Modern Phishing Campaigns [Электронный ресурс]. – 2023. – Режим доступа: <https://www.microsoft.com/security/blog>
19. National Institute of Standards and Technology (NIST). NIST SP 800-53 Rev. 5. Security and Privacy Controls for Information Systems and Organizations. – Gaithersburg: NIST, 2020. – 487 с.
20. National Institute of Standards and Technology (NIST). NIST SP 800-61r2. Computer Security Incident Handling Guide. – Gaithersburg: NIST, 2012. – 79 с.
21. Proofpoint. State of the Phish 2024 Report [Электронный ресурс]. – Режим доступа: <https://www.proofpoint.com>
22. PwnedLabs. Cybersecurity Research & Phishing Simulation Insights [Электронный ресурс]. – Режим доступа: <https://pwnedlabs.io/>
23. Ramesh G., Mahalakshmi V., Divya R. Efficient Phishing URL Detection with Machine Learning // International Journal of Computer Applications. – 2014. – Vol. 99(2). – P. 1–6.
24. Rao R. S., Pais A. R. Detection of Phishing Websites Using an Efficient Feature-Based Machine Learning Framework // Neural Computing and Applications. – 2019. – Vol. 31. – P. 3851–3873.
25. Trend Micro. Cloud App Security Overview [Электронный ресурс]. – Режим

доступу: <https://www.trendmicro.com>

26. Zhang Y., Hong J., Cranor L. F. CANTINA: A Content-Based Approach to Detecting Phishing Web Sites // Proceedings of the 16th International Conference on World Wide Web (WWW 2007). – Banff, Canada, 2007. – P. 639–648.