МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ
КАФЕДРА УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ ТА ЗАХИСТОМ ІНФОРМАЦІЇ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: “МОДЕЛЬ ОЦІНЮВАННЯ РИЗИКУ СОЦІОІНЖЕНЕРНИХ АТАК У КОРПОРАТИВНОМУ СЕРЕДОВИЩІ ІЗ ВИКОРИСТАННЯМ БАГАТОФАКТОРНОГО ПІДХОДУ”

на здобуття освітнього ступеня магістра
зі спеціальності 125 Кібербезпека та захист інформації
освітньо-професійної програми Управління інформаційною та кібернетичною безпекою

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

Владислав КОРОВІН

(підпис) *Ім’я, ПРІЗВИЩЕ здобувача*

Виконав: Здобувач вищої освіти гр. УБДМ-61
Владислав КОРОВІН

Керівник:
Доктор філософії Михайло ЗАПОРОЖЧЕНКО

Рецензент:
к.т.н., доцент Юрій ПЕПА

Київ 2025

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут кібербезпеки та захисту інформації

Кафедра Управління кібербезпекою та захистом інформації
Ступінь вищої освіти магістр
Спеціальність 125 Кібербезпека та захист інформації
Освітньо-професійна програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ
Завідувач кафедрою УКБЗІ
_____ Світлана ЛЕГОМІНОВА
"____" _____ 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студенту Коровіну Владиславу Петровичу

Тема кваліфікаційної роботи: "Модель оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу "

керівник кваліфікаційної роботи Михайло Запорожченко, *доктор філософії*

затверджені наказом Державного університету інформаційно-комунікаційних технологій від "30" жовтня 2025 р. № 467.

Строк подання кваліфікаційної роботи "25" грудня 2025 р.

1. Вихідні дані до кваліфікаційної роботи:.
2. Перелік питань, які потрібно розробити:
   1. Дослідити теоретичні засади соціоінженерних атак та оцінювання ризику
   2. Проаналізувати підходи до оцінювання ризику соціоінженерних атак
   3. Розробити практичні рекомендації оцінювання ризику соціоінженерних атак у корпоративному середовищі
3. Перелік ілюстративного матеріалу: *презентація*
4. Дата видачі завдання "02" жовтня 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назви етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1.	Визначення об’єкту, предмету, мети та завдань дослідження.	10.10.2025	виконано
2.	Збір та аналіз літератури.	23.10.2025	виконано
3.	Дослідження теоретичних засад соціоінженерних атак та оцінювання ризику	27.10.2025	виконано
4.	Аналіз підходів до оцінювання ризику соціоінженерних атак	10.11.2025	виконано
5.	Розробка практичних рекомендацій оцінювання ризику соціоінженерних атак у корпоративному середовищі	15.11.2025	виконано
6.	Формулювання висновків за результатами дослідження.	22.11.2025	виконано
7.	Оформлення роботи.	04.12.2025	виконано
8.	Оформлення презентації.	14.12.2025	виконано
9.	Отримання рецензії на роботу.	18.12.2025	виконано
10.	Захист в ЕК.	__.01.2026	виконано

Здобувач вищої освіти _____ Владислав КОРОВІН
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Керівник
кваліфікаційної роботи _____ Михайло ЗАПОРОЖЧЕНКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

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

ПОДАННЯ
ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня магістра

Направляється здобувач Коровін В. П. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 125 Кібербезпека та захист інформації
(код, найменування спеціальності)
Освітньо-професійної програми Управління інформаційною та кібернетичною безпекою
(назва)
на тему: "Модель оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу "

Кваліфікаційна робота і рецензія додаються.

Директор ННІКБЗІ _____ Євгенія ІВАНЧЕНКО
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Висновок керівника кваліфікаційної роботи

Здобувач **КОРОВІН Владислав** у кваліфікаційній роботі дослідив теоретичні основи соціоінженерних атак та проблематику оцінювання ризику, проаналізував підходи до оцінювання ризику соціоінженерних атак, базуючись на вивченому матеріалі запропонував практичні рекомендації щодо побудови моделі оцінювання ризику соціоінженерних атак у корпоративному середовищі. **КОРОВІН Владислав** показав високу теоретичну і практичну підготовку, володіння науково-дослідницькими методами, вміння самостійно знаходити шляхи вирішення проблеми дослідження. Результати дослідження апробовані на конференції "Стратегії кіберстійкості: управління ризиками та безперервність бізнесу" 28 лютого 2025 року.

Все це дозволяє оцінити кваліфікаційну роботу здобувача **КОРОВІНА Владислава** на оцінку "добре" та присвоїти йому кваліфікацію "Магістр з кібербезпеки та захисту інформації за освітньо-професійною програмою Управління інформаційною та кібернетичною безпекою".

Керівник кваліфікаційної роботи _____ Михайло ЗАПОРОЖЧЕНКО
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

"____"_____2025 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Коровін В. П. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедри
Управління кібербезпекою та захистом
інформації _____ Світлана ЛЕГОМІНОВА
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА
на кваліфікаційну магістерську роботу

Здобувач вищої освіти Коровін Владислав Петрович
на тему "Модель оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу "

Актуальність. Соціальна інженерія зосереджується на маніпулюванні поведінкою людей з метою отримання несанкціонованого доступу до конфіденційної інформації. Вдосконалюються методи та тактики атак соціальної інженерії у високоінтегрованому цифровому середовищі, тому протидія атакам завжди буде актуальною проблемою для убезпечення корпоративних середовищ організацій, щоб уникнутити матеріальних та репутаційних втрат. Тому тема є актуальною, має важливе теоретичне та практичне значення.

Позитивні сторони

1. У роботі проаналізовано та структуровано основні фактори, які треба враховувати при побудові моделі оцінювання ризику соціоінженерних атак, запропоновано рекомендаціїї щодо захисту корпоративного середовища з детальними рекомендаціями.
2. Кваліфікаційна робота оформлена відповідно до вимог. Виклад матеріалу здійснено відповідно до плану, зроблено логічні висновки. Ключові положення роботи представлено у вигляді рисунків. Автор опрацював значну джерельну базу: 90 публікацій та електронних джерел, в тому числі англомовних.

Недоліки

Доцільно було б скласти опитувальник для виявлення проблемних областей в організації щодо соціоінженерних атак або розробити тренінг для підвищення обізнанності персоналу.

Однак, вищезгадане зауваження не впливає на загальну позитивну оцінку кваліфікаційної роботи.

Висновок: Кваліфікаційна робота виконана на належному науково-методичному рівні і заслуговує позитивної оцінки, а здобувач Коровін Владислав Петрович заслуговує присвоєння кваліфікації "Магістр з кібербезпеки та захисту інформації за освітньо-професійною програмою Управління інформаційною та кібернетичною безпекою".

Рецензент: професор кафедри
Технічних систем кіберзахисту,

к.т.н, доцент _____ *підпис* Юрій ПЕПА

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 76 сторінок, 8 рисунків, 9 таблиць, 91 використане джерело.

Мета роботи - розробка моделі оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу.

Об'єкт дослідження - забезпечення захисту корпоративного середовища від соціоінженерних атак.

Предмет дослідження - модель оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу.

Методи дослідження. Системний аналіз - для дослідження типів соціоінженерних атак; порівняльний аналіз – для визначення ефективних засобів протидії атакам соціальної інженерії; метод експертних оцінок - для визначення ефективних методів підготовки персоналу.

Короткий зміст роботи. Дослідження класифікує основні типи соціоінженерних атак (Spear Phishing, Whaling, Vishing, Smishing, BEC, Clone Phishing, Social Media Phishing, DSD, Waterhole, Tailgating, Deepfake) та аналізує їх психологічні механізми впливу на свідомість, аудиторію й поведінку жертв.

Систематизовано методи соціального впливу (вплив групи, нормативний вплив, норма взаємності, моральний вплив, саморозкриття) та фактори ризику: організаційні, соціальні, культурні, психологічні, поведінкові, усвідомлення безпеки й контенту, з формалізацією багатофакторного підходу для моніторингу.

Обґрунтовано заходи протидії: технічні (машинне навчання, двофакторна автентифікація), організаційні (політики безпеки, обладнання, фільтри) та освітні (SETA-програми, кібергігієна), з інтеграцією NLP і AI для гібридного аналізу поведінки та мови.

Встановлено роль перцептивних факторів обізнаності для зниження вразливості через цільове навчання, залучення керівництва та кількісну оцінку ефективності.

Галузь застосування. Результати можуть бути застосовані в організаціях, підприємствах для зниження ризиків атак соціальної інженерії та формування політик кібербезпеки.

КЛЮЧОВІ СЛОВА: СОЦІАЛЬНА ІНЖЕНЕРІЯ, СОЦІОІНЖЕНЕРНІ АТАКИ, КІБЕРБЕЗПЕКА, РИЗИК, ОЦІНЮВАННЯ РИЗИКУ, КОРПОРАТИВНЕ СЕРЕДОВИЩЕ.

ABSTRACT

Text part of the qualification work for obtaining a master's degree: 76 pages, 8 figures, 9 tables, 91 references.

The purpose of the work is to develop a model for assessing the risk of social engineering attacks in a corporate environment using a multifactorial approach.

Object of research is to ensure the protection of the corporate environment from social engineering attacks.

Subject of research is a model for assessing the risk of social engineering attacks in a corporate environment using a multifactorial approach.

Research methods. System analysis – to study the types of social engineering attacks; comparative analysis – to identify effective countermeasures; expert assessment method – to identify effective methods of staff training.

Brief content of research. The study classifies the main types of social engineering attacks (Spear Phishing, Whaling, Vishing, Smishing, BEC, Clone Phishing, Social Media Phishing, DSD, Waterhole, Tailgating, Deepfake) and analyzes their psychological mechanisms of influence on the consciousness, audience, and behavior of victims.

It systematizes methods of social influence (group influence, normative influence, reciprocity norm, moral influence, self-disclosure) and risk factors: organizational, social, cultural, psychological, behavioral, security and content awareness, with the formalization of a multifactorial approach for monitoring.

Countermeasures are justified: technical (machine learning, two-factor authentication), organizational (policies, equipment, filters), and educational (SETA programs, cyber hygiene), with the integration of NLP and AI for hybrid behavior and speech analysis.

The role of perceptual awareness factors in reducing vulnerability through targeted training, management involvement, and quantitative effectiveness assessment is established.

Field of application. The results can be applied in organizations and enterprises to reduce the risks of social engineering attacks and to develop cybersecurity policies.

KEYWORDS: SOCIAL ENGINEERING, SOCIO-ENGINEERING ATTACKS, CYBERSECURITY, RISK, RISK ASSESSMENT, CORPORATE ENVIRONMENT.

ЗМІСТ

ВСТУП

Актуальність теми. Атаки соціальної інженерії неминучі та ставлять під загрозу цілісність, безпеку та конфіденційність інформації, що використовується на платформах соціальних мереж. Для протидії атакам соціальної інженерії в соціальних мережах використовуються відомі технології: машинне навчання, штучний інтелект та проактивний контроль доступу. Проблематика застосування ефективних рішень протидії атакам та виявлення шкідливих ресурсів на основі штучного інтелекту та машинного навчання для протидії атакам соціальної інженерії є актуальною та затребованою.

Мета роботи полягає у розробці моделі оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу.

Об'єкт дослідження – забезпечення захисту корпоративного середовища від соціоінженерних атак.

Предмет дослідження – модель оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу.

Для досягнення мети в роботі необхідно виконати наступні *завдання:*

1. Дослідити теоретичні засади соціоінженерних атак та оцінювання ризику.

2. Проаналізувати підходи до оцінювання ризику соціоінженерних атак.

3. Розробити практичні рекомендації оцінювання ризику соціоінженерних атак у корпоративному середовищі.

Методи дослідження. Системний аналіз – для дослідження типів соціоінженерних атак; порівняльний аналіз – для визначення ефективних засобів протидії; метод експертних оцінок – для визначення ефективних методів підготовки персоналу.

Наукова новизна дослідження – багатофакторний підхід до оцінки ризиків соціоінженерних атак, структуруючи фактори: організаційні (лідерство, SETA-програми), соціальні (суб'єктивні норми), культурні, психологічні (маніпуляція рисами особистості), поведінкові та усвідомлення безпеки. Новизна полягає в

аналізі піддоменів довіри, обману, мови для атак та інтеграції NLP для протидії, включаючи гібридні моделі з поведінковим аналізом. Запропоновано модель стандартизації класифікації атак, що поєднує AI для передбачення закономірностей.

Практичне значення одержаних результатів. Заходи протидії включають двофакторну автентифікацію, фільтрацію посилань, навчання кібергігієні та машинне навчання для виявлення аномалій, що зменшує ризики. Рекомендовано цільове навчання з демонстрацією маніпуляцій, участь керівництва для культури безпеки та кількісний аналіз факторів для адаптації програм. Модель NLP підвищує безпеку через аналіз мови, інтеграцію з політиками кібербезпеки, антивірусами та звітністю інцидентів. Результати дослідження можуть слугувати підгрунттям для створення потужних моделей, які розпізнають і запобігають складним маніпуляціям.

Галузь застосування. Результати можуть бути застосовані в організаціях, підприємствах для зниження ризиків атак соціальної інженерії та формування політик кібербезпеки.

Апробація результатів. Результати дослідження апробовані на конференції “Стратегії кіберстійкості: управління ризиками та безперервність бізнесу” 28 лютого 2025 року.

РОЗДІЛ 1
ТЕОРЕТИЧНІ ЗАСАДИ СОЦІОІНЖЕНЕРНИХ АТАК ТА ОЦІНЮВАННЯ РИЗИКУ

1.1. Поняття та типологія соціоінженерних атак

Платформи соціальних мереж стають головними цілями для зловмисників, використовуючи їхню доступність та взаємозв'язок (наприклад, Instagram та Facebook взаємопов'язані). Зловмисник може організувати багато зловмисних дій, починаючи від витоків даних, атак соціальної інженерії та крадіжки особистих даних до поширення дезінформації та кібератак [1].

Зокрема, атаки соціальної інженерії є однією з підступних і поширених загроз, що ставить під загрозу конфіденційність і безпеку людини. Ці шкідливі стратегії використовують схильність людини довіряти цифровим ресурсам, коли зловмисник намагається отримати особисту та фінансову інформацію користувача, надсилаючи незаконні URL-адреси. Однією з основних причин атак соціальної інженерії є людська помилка та емоційні реакції на такі фактори, як жадібність, страх, емпатія та цікавість. Зловмисники використовують ці фактори, коли жертва дозволяє та уповноважує зловмисників отримувати її особисту інформацію. Таким чином, захист платформ соціальних мереж від атак соціальної інженерії стає дедалі актуальнішим завданням.

У кібератаці соціальна інженерія слугує першим і вирішальним кроком для отримання несанкціонованого доступу до цільової системи. Вона включає розвідку цифрової системи жертви, тим самим отримуючи важливу інформацію, таку як облікові дані для входу, дані кредитної картки та облікові записи соціальних мереж. Зловмисники використовують різні методи перебору, такі як надсилання фішингових посилань та шкідливих корисних навантажень, для збору особистої інформації вищезгаданої жертви [2].

У «Стан безпеки 2021» Асоціації аудиту та контролю інформаційних систем (ISACA) зазначено, що атаки з використанням соціальної інженерії є

однією з основних причин серйозних порушень безпеки та компрометації даних у різних організаціях [3].

Згідно з даними, наведеними у «Звіті IBM про вартість витоку даних за 2025 рік» різні організації зіткнулися з середніми витратами у розмірі 4,4 млн. доларів через витоки даних, що виникли внаслідок атак соціальної інженерії, але середньосвітова вартість витоку даних у доларах США знизилася на 9% порівняно з минулим роком завдяки швидшій ідентифікації та локалізації [4].

Фінансові втрати, втрата персональних даних, крадіжка особистих даних та вимагання в основному повязано з атаками соціальної інженерії. Типовий життєвий цикл атаки соціальної інженерії включає розслідування, планування, контакт та виконання [5].

Життєвий цикл атаки соціальної інженерії [5].

Розслідування — це початкова фаза, на якій зловмисники збирають інформацію про свою ціль, вивчаючи її цифрові системи та звичну інформацію, щоб виявити вразливості, які згодом можна використати. Для цього зловмисники використовують різні джерела, такі як загальнодоступні дані, акаунти в соціальних мережах, веб-сайти шукачів роботи та профілі в ЗМІ. Головна мета цієї фази — зібрати достатньо інформації про ціль, щоб персоналізувати поверхню атаки, зробивши її більш переконливою. •

Планування — Після успішного збору інформації про ціль зловмисник розробляє проактивну стратегію, яка максимізує поверхню атаки та мінімізує шанси бути спійманим. Це включає ефективну атаку соціальної інженерії, таку як фішингове посилання, видавання себе за іншу особу або телефонний дзвінок під приводом текстового повідомлення.

Контакт — на цьому етапі зловмисник намагається зв’язатися з цільовою особою, використовуючи методи фішингу, вішингу та смішингу. Наприклад, зловмисник може створити фішинговий електронний лист, що містить шкідливе корисне навантаження, або здійснити телефонний дзвінок, видаючи себе за довірену особу. Це найважливіший етап, на якому зловмисники повинні

покладатися виключно на методи соціальної маніпуляції, які встановлюють довіру та переконують цільову особистість вжити бажаних дій.

Виконання — на останньому етапі зловмисники пасивно збирають конфіденційну інформацію від цілі, таку як облікові дані та банківські реквізити, або встановлюють шкідливе програмне забезпечення в їхню систему. Цей етап також вимагає ретельного розгляду можливостей експлуатації, щоб уникнути виявлення з першої лінії захисту (наприклад, системи виявлення вторгнень, брандмауерів та антивіруса). В таблиці 1.1 представлено найведоміші кібератаки на основі соціальної інженерії.

Таблиця 1.1

Деякі з найвідоміших кібератак на основі соціальної інженерії

Джерело	**Компанія**	**Дата**	**Деталі/Пошкодження**	**Метод/інструменти порушення**
[6]	Саудівська Арамко	2021 рік	Хакери заявили, що мають майже 1 терабайт даних Aramco, і вимагали викуп у розмірі 50 мільйонів доларів США.	Фішинговий електронний лист.
[9]	Майкрософт	2021 рік	Кілька користувачів MS Office попалися на фішингову електронну пошту. Кожну жертву обдурили на суму від 100 до 199 доларів США.	Атака на компрометацію ділової електронної пошти (BEC), фішинговий електронний лист
[7 , 8]	Марріотт	2020–2018 рр.	В обох випадках хакери отримали доступ до мільйонів записів гостей. Ці записи містили імена гостей, адреси, контактні номери та зашифровану інформацію про кредитні картки.	Фішинговий електронний лист, скомпрометовані облікові дані двох співробітників Marriott, троян віддаленого доступу (RAT), інструмент пост-експлоатації Mimikatz.
[10]	Твіттер	2020 рік	Хакери зламали 130 облікових записів у Twitter. Кожен обліковий запис мав щонайменше 1 мільйон підписників. Хакери використали 45 впливових облікових записів для просування шахрайської схеми з біткойнами.	Атака SE на основі персонифікації, фішингові атаки.
[11]	Акулячий танк	2020 рік	Ведучий Shark Tank втратив 400 000 доларів США після того, як попався на шахрайський електронний лист.	Фішинговий електронний лист.
[12]	Тойота	2019 рік	Корпорація Toyota Boshoku втратила 37 мільйонів доларів США після того, як стала жертвою атаки BEC.	Фішинговий електронний лист (наприклад, BEC)

Джерело	Компанія	Дата	Деталі/Пошкодження	Метод/інструменти порушення
[13]	Енергетична компанія (база у Великій Британії)	2019 рік	Хакери обдурили та виманили у генерального директора (CEO) 243 000 доларів США.	Видавання себе за дипфейк-фішинг
[14 , 15]	Google та Facebook	2015–2013 рр.	Фішингові електронні листи коштували Google та Facebook понад 100 мільйонів доларів США.	Спис

Найпоширеніші типи соціоінженерних атак представлено в таблиці 1.2.

Таблиця 1.2.

Поширені типи соціоінженерних атак [16]

Фішингова атака	Опис
Спір-фішинг(Spear Phishing)	Спеціально спланована атака, спрямована на конкретну особу. Наприклад, атака на співробітника спрямована на отримання доступу до мережі організації.
Вейлінг(Whaling)	Цільовою метою зазвичай є відома особа. Атака потребує значного часу, щоб знайти можливість або засоби для компрометації кваліфікації цієї особи.
Вішінг(Vishing)	Вішинг або голосовий фішинг – це атака, заснована на пошуковій платформі (SE). Вішинг – це шахрайський дзвінок, спрямований на отримання секретної інформації або облікових даних цільової особи.
Смішінг(Smishing)	Смішинг — це формат текстового повідомлення для атаки вішингу. У смішингу єдина відмінність полягає в тому, що він базується на текстовому повідомленні, а не на дзвінку.
Компрометація ділової електронної пошти/видавання себе за іншу особу (BEC) (Impersonation/business email compromise (BEC))	BEC – це атака, яка вимагає планування та інформації. Під час атаки BES зловмисник видає себе за керівника компанії, стороннього ресурсу або постачальника, щоб отримати секретну інформацію, доступ до мережі організації тощо.
Клонування (Clone)	Клонування фішингу – це фішингова атака на основі електронної пошти. У цих атаках зловмисник знає більшість бізнес-додатків, що використовуються особою або організацією. На основі цих знань зловмисник клонує схожий електронний лист, замаскований під звичайний лист із програми, щоб витягти важливу інформацію або навіть облікові дані від цілі.
Фішинг у соціальних мережах (Social media phishing)	Під час фішингу в соціальних мережах зловмисник спостерігає за соціальними мережами та іншими часто відвідуваними сайтами цільової особи, щоб зібрати детальну інформацію. Потім зловмисник планує атаку на основі отриманої інформації. Зловмисник може використовувати зібрану інформацію, щоб обдурити жертву різними способами.

Фішингова атака	Опис
Розподілена відволікаюча система від спаму (DSD) (Distributed spam distraction (DSD)	Атака DSD виконується у два кроки. На першому кроці жертві надсилають спам-листи з фішинговими електронними листами, що відображають автентичне або надійне джерело, наприклад, новий лист, журнал, програмну компанію тощо. Ці фальшиві електронні листи містять посилання, яке веде жертву на веб-сторінку, що є копією веб-сайту автентичної та надійної компанії. Другий крок залежить від того, як зловмисник планує провести атаку SE, тобто фальшива сторінка може запитати у жертви дані для входу (для отримання додаткової або конфіденційної інформації) для підтвердження особи та подальшого продовження.

Існує кілька типів атак соціальної інженерії, зокрема наступні [17]:

Смішінг (Smishing) є варіацією фішингових атак, які здійснюються через SMS-повідомлення. Головна ідея такого методу полягає у надсиланні жертві текстових повідомлень, що містять посилання на шкідливі ресурси або заклики до негайних дій, спрямованих на отримання конфіденційних даних. Часто зловмисники маскують свої повідомлення під офіційні звернення від банків, поштових служб, державних установ або компаній, що надають різноманітні послуги. Особливо ефективними такі атаки є у періоди підвищеного навантаження на комунікаційні канали, наприклад під час великих знижок або кризових ситуацій, коли люди менше звертають увагу на потенційні загрози.

Вішінг(Vishing) є методом шахрайства, що використовує телефонні дзвінки для отримання конфіденційної інформації. Зловмисники можуть представлятися співробітниками банків, служб підтримки, державних органів або навіть правоохоронних структур, створюючи ситуацію, яка спонукає жертву передати особисті дані або вчинити певні дії, наприклад, здійснити фінансовий переказ. Часто вішінг-атаки супроводжуються психологічним тиском: зловмисники можуть лякати жертву блокуванням рахунків, фінансовими штрафами або іншими негативними наслідками. Використання технологій зміни голосу та підробки номерів телефонів робить такі атаки ще більш складними для виявлення.

Спір-фішинг (Spear Phishing) відрізняється від звичайного фішингу тим, що атака спрямована не на масову аудиторію, а на конкретну особу чи організацію. Основна мета таких атак –отримати доступ до важливої інформації,

зокрема даних про корпоративні мережі, фінансові операції або особистих даних високопосадовців. Такі атаки зазвичай ретельно сплановані, включають персоналізовані звернення та імітують реальні комунікації, які жертва може очікувати отримати. Для підвищення ефективності атак зловмисники можуть використовувати попередньо зібрану інформацію про людину, що робить фальшиві повідомлення ще більш правдоподібними.

Вейлінг (Whaling) це різновидом спір-фішингу, проте його основною мішенню стають керівники вищої ланки, топ-менеджери та інші особи, що мають доступ до критично важливих даних компанії. Такі атаки часто маскуються під офіційні запити від урядових установ, фінансових служб або партнерів, і можуть включати вимоги надати конфіденційну інформацію, підписати документи або здійснити фінансові перекази.

Фішинг (Phishing) є одним із найбільш поширених методів соціальної інженерії, який полягає у спробах шахраїв отримати конфіденційну інформацію користувачів шляхом масового розсилання фальшивих повідомлень, що імітують офіційні звернення від банків, державних установ або великих компаній. Основна особливість фішингових атак полягає у використанні соціального фактору, коли жертва через довіру до джерела інформації сама надає свої особисті дані. Зловмисники можуть змушувати користувачів перейти зашкідливими посиланнями, завантажити шкідливі вкладення або ввести свої паролі на підроблених веб-ресурсах. Така атака може використовуватися як на індивідуальному рівні, так і проти цілих організацій, сприяючи витоку корпоративної інформації. На рисунку 1.1. представлено типовий процес фішингових атак.

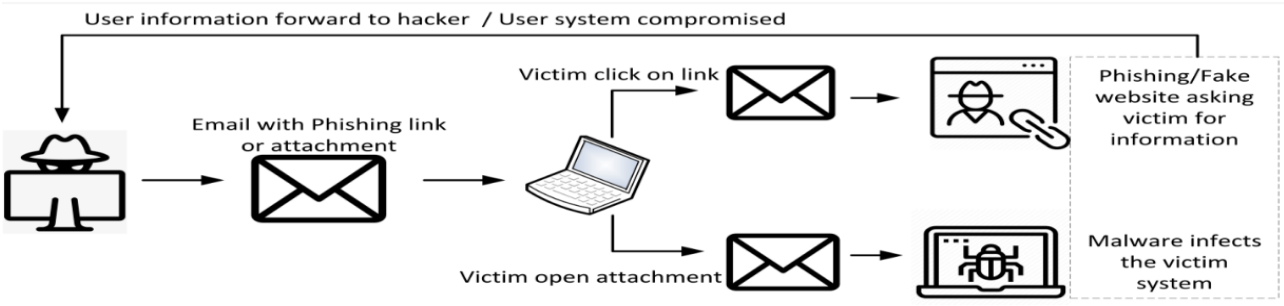

Рис. 1.1. Типовий процес фішингових атак [16]

Фішинг з використанням копій. Це техніка, яка шахрайським шляхом отримує конфіденційну інформацію, надсилаючи високоспеціалізовані електронні листи кільком кінцевим користувачам. Це основна відмінність між фішинговими атаками, оскільки фішингові кампанії зосереджені на розсилці великих обсягів узагальнених електронних листів з очікуванням, що лише кілька людей відповість. З іншого боку, електронні листи з метою фішингу вимагають від зловмисника проведення додаткових досліджень своїх цілей, щоб «обманом» змусити користувачів виконати запитувані дії. Рівень успішності цих фішингових атак значно вищий, ніж у фішингових атак.

Атака типу «Waterhole» (або «питні ями») — це кібератака, під час якої хакери компрометують веб-сайт, який часто відвідує певна група (наприклад, співробітники компанії або представники галузі), щоб заразити їхні пристрої шкідливим програмним забезпеченням, отримуючи доступ до їхніх мереж для шпигунства або крадіжки даних, що названо на честь хижаків, які чатують у «питних ями». Зловмисники заражають легітимні сайти шкідливим кодом, часто використовуючи вразливості нульового дня, щоб доставляти шкідливе програмне забезпечення (наприклад, RAT) через «автоматичні завантаження», коли цілі відвідують сайти, що ускладнює виявлення, оскільки самі сайти не є шкідливими за своєю суттю. Термін «водяна діра» стосується ініціювання атаки проти цільових підприємств та організацій. Зловмисник використовує стратегію соціальної інженерії, яка використовує довіру користувачів до веб-сайтів, які вони регулярно відвідують. Мета цієї атаки не полягає в тому, щоб розповсюдити шкідливе програмне забезпечення на якомога більшій кількості систем. Натомість зловмисники запускають експлойти на відомих та надійних сайтах, які, ймовірно, відвідуватимуть їхні цільові жертви. Це робить техніку «waterhole» ефективною для досягнення цільового корисного навантаження. Ця стратегія була успішно використана для отримання доступу до деяких (нібито) дуже безпечних систем. Підготовка до цього типу атаки починається зі збору інформації, щоб підтвердити, що цілі відвідують веб-сайти та що система дозволяє такі відвідування. Потім зловмисник тестує ці веб-сайти на наявність

вразливостей, щоб впровадити код, який може заразити систему відвідувача шкідливим програмним забезпеченням. Як тільки жертви відвідують скомпрометований сайт, експлойт використовує вразливості програмного забезпечення для встановлення шкідливого програмного забезпечення. Завантажене шкідливе програмне забезпечення може бути у формі трояна віддаленого доступу, який дозволяє зловмисникам отримати доступ до захищеної системи та отримати конфіденційні дані [16].

На рисунку 1.2. представлено етапи атаки «waterhole»

Рис. 1.2. Етапи атаки «waterhole»

Tailgating (тейлгейтинг)— це «слідування за іншою особою» або «прохід за іншим», вид соціальної інженерії та фізичної безпеки (кібербезпеки), коли зловмисник непомітно проходить на захищену територію разом з легітимним працівником, який саме проходить через контроль доступу, використовуючи його пропуск або просто його довіру, щоб отримати доступ до конфіденційних даних чи систем.

Діпфейк – це нещодавня та дуже переконлива техніка, яка використовується для проведення SE-атак. Кіберзлочинці використовують діпфейки для підробки зображень, аудіо та відео для досягнення певної мети. У кібербезпеці діпфейк є зростаючою загрозою. Одним з найвідоміших алгоритмів для створення діпфейкового контенту є генеративно-змагальні мережі (GAN). GAN – це комбінація двох штучних нейронних мереж (ANNs).

Ці штучні нейронні мережі (ANNs) називаються детекторами та синтезаторами. ANNs навчаються з використанням великих наборів даних реальних зображень, аудіо та відеокліпів. Потім синтезаторна ANN генерує контент з діпфейками, а ANN детектора намагається розрізнити справжність контенту. Цикл генерації контенту з діпфейками продовжується доти, доки ANN детектора більше не може ідентифікувати згенерований фальшивий контент як підроблений. Через цей ретельний процес генерації та перевірки, згенерований підроблений контент за допомогою GAN дуже важко ідентифікувати як підроблений. Рисунок 1.3 ілюструє огляд процесу створення матеріалу з діпфейками. Для навчання мережі використовуються два різних обличчя, тобто Обличчя А та Обличчя В. Пізніше мережа використовується для генерації Обличчя А з виразами або аудіо з Обличчя В. Новостворене зображення з оригінальною інтерпретацією Обличчя А за допомогою Обличчя В може бути використане для заплутування або впливу на жертву [16].

Діпфейк був використаний для атаки SE, проведеної на британську енергетичну компанію. В атаці голос з діпфейками був використаний для введення в оману генерального директора компанії [6]. Окрім шахрайства, діпфейк також використовувався в кількох інших злочинних видах діяльності, таких як шантаж, пошкодження репутації, фейкові новини, дезінформація, масова паніка тощо.

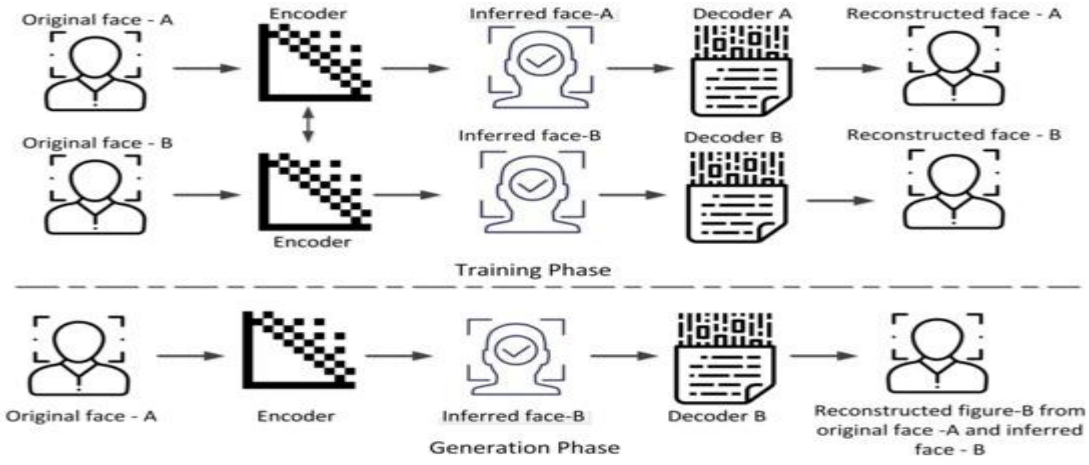

Рис.1.3. Спрощена ілюстрація навчання та створення дипфейкових зображень або відео

1.2. Соціоінженерія у корпоративному середовищі: причини та вразливості психологічного впливу

Несанкціоноване заволодіння інформацією шляхом застосування методів соціальної інженерії відбувається через використання маніпулятивних технік, заснованих на обманних діях та шахрайських схемах. Водночас створення штучних обставину дозволяють зловмисникам імітувати правдоподібні сценарії для введення жертви в оману.

При оцінюванні рівня захищеності інформації в комп'ютерних системах із застосуванням соціоінженерного підходу важливо враховувати форми маніпулятивного впливу. Методи соціальної інженерії спрямовані на моделювання дій порушників, які, наприклад, можуть бути націлені на персонал організації. Соціальний інженер, намагаючись отримати відомості про комп'ютерні системи, використовує контактну інформацію з відкритих джерел, зокрема даніпро користувачів: їхні прізвища, імена, посади. На основі зібраної інформації визначаються потенційні вразливості, через які можуть реалізовуватися загрози соціальної інженерії. Основу методів соціальної інженерії становлять такі аспекти:

- закономірності, що впливають на свідомість людини;
- особливості аудиторії або середовища діяльності;–недостатня обізнаність у термінах і предметних сферах інформаційної безпеки;
- психологічна нестійкість особистості, що проявляється у поведінкових шаблонах і може бути використана для маніпуляції через базові потреби, слабкі сторони, цінності та прагнення [16]. В таблиці 1.3. представлено основні Методи соціального впливу, пов'язані з кібератаками на основі соціальної інженерії.

Таблиця 1.3.

Методи соціального впливу, що застосовуються під час кібератак на основі соціальної інженерії

Методи соціального впливу	Опис
Вплив групи [19]	Конформність – це варіація поведінки, спрямована на згоду з іншими. Група може впливати на таку поведінку. У соціальних мережах онлайн-групи можуть використовуватися для впливу на жертв, щоб вони піддалися атаці SE. Наприклад, група в соціальних мережах із сотнями підписників може надати підписнику шкідливе посилання та вплинути на жертву, повідомивши, що кожен учасник повинен перейти за посиланням, щоб стати частиною цієї нової групи для майбутніх подій та інформації.
Інформаційний вплив/нормативний вплив [20]	У випадку атак на безкоштовну екшн-систему (SE-атаки) ворог часто використовує певну інформацію та налаштування за допомогою інформаційних/нормативних впливів. Наприклад, інформуючи жертву про якесь безкоштовне програмне забезпечення та переконуючи її встановити його, надаючи інформацію про важливість програмного забезпечення та простоту використання. Такий підхід може бути використаний для того, щоб збити з пантелику або маніпулювати жертвою, змусивши її виконувати певні дії або розкривати інформацію, яка вигідна ворогові.
Теорія соціального обміну/норма взаємності [18 , 20]	Такі методи впливу використовуються у зворотних атаках соціальної взаємодії. Теорія соціального обміну підкреслює, що люди приймають рішення щодо цінності (навмисно чи ненавмисно) стосунків. Наприклад, працюючи в корпоративному середовищі, колеги обмінюються послугами на основі своїх стосунків один з одним.
Моральний вплив/соціальна відповідальність [21]	Атаки з використанням сексуальної орієнтації (SE) використовують моральний вплив або соціальну відповідальність двома способами. Один із них полягає в тому, що ворог використовує корисний характер жертви, щоб отримати інформацію або здобути прихильність для полегшення атаки. Другий спосіб полягає в тому, щоб чинити тиск на жертву з точки зору норм соціальної відповідальності або морального обов'язку під час SE-атаки. Цей тиск морального обов'язку впливає на поведінку жертви, зокрема, якщо жертва не бажає пропонувати будь-яку допомогу. Прикладом може бути онлайн-група допомоги тваринам. Зловмисник може ідентифікувати осіб, які мають високу мотивацію та готові допомогти. Зловмисник може націлитися на цю жертву заради фінансової вигоди або може вигадати історію, щоб націлитися на моральні цінності жертви для отримання інформації.
Саморозкриття/встановлення взаєморозуміння [22 , 23]	Дослідження показують, що в процесі побудови соціальних відносин саморозкриття викликає готовність розкриватися більше людям, які виявляють з нами зв'язки. Супротивники використовують цей метод самозакоханості щодо жертв, які відчувають потребу зв'язатися з кимось особливим.

Існує багатоспособів переконання людини до спонукання здійснити низку дій, в яких зацікавлена сторона спонукання, основні з них редставлено в таблиці 1.4

Таблиця 1.4.

Типи переконання, що використовуються в атаках соціальної інженерії

Види переконань	Опис
Подібність [25]	Схожість інтересів породжує схожість, а несхожість — неприязнь. Злочинець схильний використовувати це як ефективний підхід для завоювання довіри жертви. Наприклад, на платформах соціальних мереж ворог може приєднатися до груп, схожих на групи, до яких приєдналася потенційна ціль. Така схожість може допомогти побудувати довірчі стосунки між хакером і жертвою.
Відволікання уваги/ маніпуляції [24 , 26 , 27]	Дослідження показують, що помірна відволікаюча увага справді сприяє переконанню. Відволікання використовується як ефективний інструмент у маніпулятивних атаках. Прикладом атаки SE на основі відволікання є атака DSD, виділена в таблиці 2 .
Цікавість [28]	Більшість людей за своєю природою допитливі. Під час атаки SE людська допитливість може бути використана багатьма способами. Наприклад, зловмисник може надіслати фішинговий електронний лист або заражений файл як вкладення з цікавою темою листа, наприклад, «вас звільнили», річний звіт про ефективність роботи, список звільнених співробітників тощо.
Переконання з використанням авторитету/ достовірності [29 , 30]	Більшість людей схильні підкорятися вимогам перед авторитарною фігурою. В інтернеті хакери використовують символи та логотипи, що відображають справжність та авторитет. Наприклад, офіційний логотип податкової служби, правоохоронних органів тощо, щоб продемонструвати авторитет та довіру, може бути ефективним підходом до ініціювання SEO-атаки.

Методи впливу на ставлення та поведінку жертв можна додатково розділити на піддомени, як показано в таблиці 1.5.

Таблиця 1.5.

Методи впливу на ставлення та поведінку жертв для проведення атак соціальної інженерії

Методи впливу на ставлення та поведінку	Опис
Враження/ зобов'язання [31 , 32]	Теорія самопрезентації підкреслює той факт, що кожна людина справляє приємне враження, як внутрішньо, так і для інших людей. Людина може докласти багато зусиль для створення бажаного іміджу. Такі зусилля можуть бути можливістю для хакерів здійснювати SE-атаки. Наприклад, на поведінку людини може вплинути, якщо її соціальний імідж знаходиться під загрозою.
Теорія когнітивного дисонансу [33 , 34]	Теорія підкреслює внутрішній конфлікт, коли поведінка та переконання людини не узгоджуються одне з одним. Такий конфлікт може впливати на когнітивні упередження, тобто на прийняття рішень. Зловмисник може використовувати ці когнітивні упередження для вилучення конфіденційної інформації.
Поведінка впливає на ставлення [35]	Дія, за якої людина, погодившись на незначне прохання, з більшою ймовірністю виконає й важливе, відома як ефект «ноги у дверях». По суті, люди створюють імідж, роблячи незначну послугу; щоб підтримувати цей імідж, вони схильні погоджуватися на наступну послугу. Зловмисники, що використовують стратегію SE, можуть використовувати таку поведінку для ініціювання атаки.
Ефект стороннього спостерігача [36]	Ефект свідка визначає людську поведінку, коли людина не бажає допомагати, коли присутні свідки. Під час нападів на жертв, які перебувають у групі, жертви можуть бути

Методи впливу на ставлення та поведінку	Опис
	спокушені втягнути в певну ситуацію та використані пізніше в приватному чаті для отримання особистої або конфіденційної інформації.
Дефіцит/тиск часу [21]	Під час атак самозахопливої загрози зловмисник використовує дефіцит, щоб викликати відчуття паніки або терміновості. Ця паніка/терміновість може вплинути на здатність жертви приймати рішення. Через цю плутанину зловмисник може переконати жертву прийняти рішення, які він вважає бажаними.

У соціальних мережах та віртуальних мережах користувачі демонструють рівень довіри на основі своєї залученості на віртуальних платформах [37]. Чим вища залученість на віртуальних платформах, тим вищий рівень довіри до них. Цей рівень залученості можна виміряти багатьма способами, тобто кількістю друзів або зв'язків, публікацій, груп, на які підписані тощо. Користувачі, які демонструють високий рівень залученості в соціальних мережах або віртуальних мережах, більше схильні до атак SE [38]. В атаках SE довіру та обман можна додатково класифікувати на піддомени, як показано в таблиці 1.6.

Таблиця 1.6.

Піддомени довіри та обману для проведення атак соціальної інженерії

Піддомени довіри та обману	Опис
Довіра/стосунки [39]	Побудова довіри є однією з найважливіших частин атак SE. Зловмисник може використовувати різні засоби для розвитку довірчих стосунків з жертвою. Такі методи, як вплив, переконання, схожість, винагорода тощо, можуть бути використані для побудови довірчих стосунків з ціллю. Згідно з дослідженнями, як тільки встановлені довірчі стосунки, жертва не відчуває вагань бути вразливою перед довіреною особою. Такі стосунки можуть бути ризикованою поведінкою, яка може сприяти атаці SE.
Обман/шахрайство [32 , 40 , 41]	Обман – це навмисна дія, заснована на стратегічній взаємодії обманщика. Теорія міжособистісного обману (ТМОБ) припускає, що люди вважають, що можуть розпізнати обман, але здебільшого це не так. ТМОБ також підкреслює, що зловмисник вживає всіх дій на основі стратегічного плану маніпулювання поведінкою жертви. Атака мізерної обману, заснована на обмані, може спиратися на кілька методів, наприклад, брехня, створення неправдивих вигадок або наративів, часткову правду, ухилення від питань, створення враження непорозуміння тощо.

Мова — це не лише найпоширеніший метод комунікації, але й засіб обробки, генерування та вираження думок. Процес соціальної взаємодії за

допомогою мови дуже схожий на процес, пов'язаний з мовою програмування. Люди сприймають слова як вхідні дані, обробляють їх та генерують відповідь як вихідні дані. Важливо підбирати відповідні слова для пояснення контексту теми чи почуття. Це означає, що розробка та опрацювання інформації для атак SE значною мірою залежить від мови, яка використовується для взаємодії з жертвою, тобто використовується мовне пізнання [42,43]. У таблиці 1.7 наведено огляд піддоменів мови та міркування, пов'язаних з атаками SE.

Таблиця 1.7.

Піддомени мови/міркування, пов'язані з атаками соціальної інженерії

Піддомени мови/міркування	**Опис**
Ефект фреймінгу/когнітивне упередження [44]	Феномен відображення когнітивних упереджень, тобто висловлення думок та прийняття рішень, залежить від способу постановки питання. Це когнітивне упередження, засноване на мовному фреймінгу, призводить до маніпуляції рішеннями. Наприклад, яловичина з етикеткою «75% пісна» є більш бажаною для покупців порівняно з етикеткою «25% жиру». У випадках атак самоізоляції когнітивне упередження жертви використовується за допомогою заздалегідь спланованого мовного фреймворку.
Ускладнення процесу мислення [45 , 46 , 47]	Мова відіграє важливу роль в процесі мислення під час соціальної взаємодії. Ця залежність може створити можливість викликати «сплутаність думок» за допомогою мови. Наприклад, щоб викликати сплутаність думок, зловмисник може залучити свою жертву до заяви з неграматичним або незрозумілим значенням. Такі заяви можуть спокусити жертву діяти на основі презумпції, наприклад, заяву «зворушливо чути» може призвести до того, що жертва торкнеться вуха. Іншим прикладом може бути неповна заяву, наприклад, «Я не чую», яка може спонукати жертву перевірити або налаштувати обладнання.

Висновки до розділу 1

Визначено найпоширеніші типи соціоінженерних атак: Спір-фішинг(Spear Phishing), Вейлінг(Whaling), Вішінг(Vishing), Смішінг(Smishing), Компрометація ділової електронної пошти/видавання себе за іншу особу (Impersonation/business email compromise (BEC), Клонування (Clone), Фішинг у соціальних мережах (Social media phishing), Розподілена відволікаюча система від спаму Distributed spam distraction (DSD), Атака типу «waterhole», Тейлгейтинг (Tailgating), Діпфейк та дано змістовні характеристики.

Доведено, що методи соціальної інженерії використовують закономірності, що впливають на свідомість людини; особливості аудиторії або середовища діяльності; недостатню обізнаність у термінах і предметних сферах інформаційної безпеки; психологічну нестійкість особистості, що проявляється у поведінкових шаблонах і може бути використана для маніпуляції через базові потреби, слабкі сторони, цінності та прагнення.

Проаналізовано методи соціального впливу, що застосовуються під час кібератак на основі соціальної інженерії, які поділяються на вплив групи, інформаційний вплив/нормативний вплив, теорію соціального обміну/норму взаємності, моральний вплив/соціальну відповідальність, саморозкриття/встановлення взаєморозуміння, акцентовано на типах переконання, що використовуються в атаках соціальної інженерії та розкрито сутність технік впливу на ставлення та поведінку жертв для проведення атак соціальної інженерії.

Розкрито піддомени довіри та обману, піддомени мови та міркування для проведення атак соціальної інженерії.

РОЗДІЛ 2
АНАЛІЗ ПІДХОДІВ ДО ОЦІНЮВАННЯ РИЗИКУ СОЦІОІНЖЕНЕРНИХ АТАК

2.1 Формалізація багатофакторного підхіду до оцінювання ризику соціоінженерних атак

Для формалізації багатофакторного підхіду до оцінювання ризику соціоінженерних атак треба визначити основні фактори впливу та структурувати їх для формування стратегії моніторингу та відбиття атак соціальної інженерії. Основними факторами, які можна виділити будемо вважати: організаційні фактори, соціальні фактори, культурні фактори, психологічні фактори, поведінковий фактори, фактор усвідомлення безпеки, фактор контенту.

Організаційний фактор:

Організаційні фактори – це ті, що безпосередньо контролюються організацією та впливають на поведінку та обізнаність працівників щодо безпеки. Цей фактор має п'ять підфакторів:

1. Лідерство [84]: стосується ролі лідерів організації у просуванні культури безпеки, встановленні тону для практик безпеки та ефективному донесенні важливості протидії загрозам соціальної інженерії. Цей підфактор включає цінність організації, мету організації, комунікацію системи ІБ та підтримку керівництвом ініціатив безпеки ІБ.

2. Превентивні контрзаходи [85]: це проактивні заходи, що вживаються організацією для запобігання атакам соціальної інженерії, такі як контроль доступу, кампанії з підвищення обізнаності з питань безпеки та механізми автентифікації. Впроваджуючи превентивні контрзаходи, такі як програми та політики навчання з безпеки, співробітники знайомляться з важливими практиками та знаннями безпеки. Ці заходи надають співробітникам інформацію, необхідну для розпізнавання тактик соціальної інженерії, виявлення

попереджувальних знаків та впровадження превентивних заходів. Як наслідок, співробітники стають більш обізнаними та краще розуміють ризики, пов'язані із соціальною інженерією.

3. Детективні контрзаходи [85]: механізми та системи, що використовуються для виявлення та ідентифікації спроб соціальної інженерії, такі як системи виявлення вторгнень та інструменти моніторингу безпеки. Цей підфактор включає структурні засоби контролю, які є фізичними та технічними заходами, що впроваджуються в організації для захисту від атак соціальної інженерії, такі як безпечний контроль доступу та відеоспостереження. Такі контрзаходи слугують нагадуванням співробітникам про те, що атаки соціальної інженерії все ще можуть відбуватися, незважаючи на вжиті превентивні заходи. Це підсилює необхідність пильного та безпечного підходу. Співробітники стають більш усвідомленими щодо можливості загроз соціальної інженерії та заохочуються залишатися обережними, навіть якщо впроваджено превентивні заходи.

4. Процедурні контрзаходи [85]: це задокументовані процедури та протоколи, які допомагають співробітникам виявляти, повідомляти та реагувати на атаки соціальної інженерії. Визначаючи кроки, що використовуються для перевірки запитів, автентифікації повідомлень або обробки незнайомих ситуацій, ці процедури забезпечують співробітників структурованим підходом до вирішення ризиків соціальної інженерії. Чіткість підвищує їхню обізнаність щодо необхідних запобіжних заходів та дій. Цей підфактор включає організаційну ефективність, забезпечення та просування інтернет-провайдерів та формалізацію робочих процедур, що стосується ступеня, до якого робочі процедури та процеси формально задокументовані та стандартизовані, включаючи ті, що стосуються інформаційної безпеки та обізнаності щодо соціальної інженерії.

5. Програма SETA [86]: розшифровується як програма освіти, навчання та підвищення обізнаності з питань безпеки, яка включає комплексні формальні ініціативи, спрямовані на навчання працівників ризикам соціальної інженерії,

проведення навчання з найкращих практик безпеки та надання винагород або покарань з метою підвищення обізнаності працівників про потенційні загрози.

Організаційна культура пов'язана із соціальними факторами, оскільки працівники можуть бути більш мотивованими дбати про безпеку, якщо їхні колеги так само пильні. Крім того, організації, які надають пріоритет психологічній безпеці, заохочують працівників ставити запитання або перевіряти запити, що пов'язано з такими психологічними факторами, як впевненість та самоефективність.

Соціальний фактор:

Соціальні фактори – це зовнішні та міжособистісні сили та взаємодії, які формують ставлення, переконання та поведінку працівника щодо загроз соціальної інженерії та заходів безпеки. Соціальний фактор складається з п'яти підфакторів:

1. Суб'єктивні норми [85]: відображають сприйняття людиною соціальних очікувань та переконань щодо важливості заходів безпеки проти загроз соціальної інженерії. Якщо працівники сприймають, що їхні колеги або керівництво надають пріоритет обізнаності про соціальну інженерію та дотримуються протоколів безпеки, вони з більшою ймовірністю займуть подібне ставлення та поведінку. Цей підфактор включає соціально-когнітивні фактори, нормативні переконання, особисті норми, описові норми, норми-накази, формальні норми та вплив поведінки однолітків.

2. Афективна відданість [86]: це емоційна прив'язаність, лояльність та ідентифікація людини зі своєю організацією, що може впливати на її зобов'язання захищатися від загроз соціальної інженерії. Така відданість сприяє почуттю відповідальності та причетності до благополуччя організації. Працівники, які емоційно віддані своїй організації, частіше усвідомлюють важливість обізнаності про соціальну інженерію та потенційний вплив порушень безпеки.

3. Зобов'язання щодо продовження роботи [85]: стосується сприйнятих особою витрат та вигод, пов'язаних із перебуванням у поточній організації, що

може вплинути на її мотивацію дотримуватися заходів безпеки. Працівники, які працюють в організації протягом тривалого періоду, накопичують знання та досвід, включаючи розуміння практик безпеки та визнання ризиків соціальної інженерії. Їхня тривалість та безперервність в організації сприяють їхній загальній обізнаності та здатності легше виявляти потенційні загрози.

4. Вплив ЗМІ [86]: стосується впливу засобів масової інформації, включаючи новинні видання та платформи соціальних мереж, на формування сприйняття та розуміння людьми загроз соціальної інженерії. ЗМІ часто повідомляють про інтенсивні інциденти соціальної інженерії, такі як витоки даних, фішингові атаки або випадки крадіжки особистих даних. Цей вплив підвищує обізнаність серед співробітників про потенційні ризики та тактики, що використовуються в атаках соціальної інженерії. Цей підфактор включає кількість публікацій, знаменитість, за якою стежать, прихильність, суспільні очікування щодо захисту інформації та кількість друзів, особливо спільних.

5. Нормативне зобов'язання [87] це почуття обов'язку та відповідальності особи дотримуватися заходів безпеки та захищатися від загроз соціальної інженерії через соціальні норми та очікування. Це зобов'язання часто призводить до того, що працівники мають сильне почуття етичної відповідальності перед своєю організацією, оскільки вони вірять у дотримання моральних принципів та діють в найкращих інтересах організації. Такий етичний менталітет впливає на усвідомлення соціальної інженерії, оскільки працівники усвідомлюють важливість захисту конфіденційної інформації, збереження конфіденційності та запобігання атакам соціальної інженерії, які можуть поставити під загрозу цілісність організації.

Соціальне середовище формує ставлення працівників до безпеки. Атаки соціальної інженерії часто використовують довіру або соціальні норми, і якщо працівники працюють у середовищі, де довіра не піддається сумніву або домінують ієрархічні структури, вони можуть мати меншу обізнаність у питаннях кібербезпеки. З іншого боку, соціальне середовище, яке дбає про безпеку, заохочує працівників до колективної пильності. Соціальні фактори

часто посилюють організаційні та культурні фактори. Наприклад, якщо організація сприяє відкритій культурі, де заохочується постановка запитань та перевірка інформації, соціальна підтримка підкріплює таку поведінку.

Культурний фактор:

Культурні припущення та переконання мають значний вплив на обізнаність працівників щодо соціальної інженерії [87]. Культурні норми та цінності формують розуміння людьми того, що вважається прийнятним або неприйнятним з точки зору практики безпеки. Рівень довіри в культурі може впливати на сприйнятливість людей до атак соціальної інженерії, причому культури довіри потенційно є більш вразливими. Культурні орієнтації на колективізм чи індивідуалізм можуть впливати на дотримання працівниками заходів безпеки та їхню готовність повідомляти про підозрілу діяльність. Дистанція влади, або переконання щодо ієрархічних структур, можуть впливати на готовність працівників ставити під сумнів владу або оскаржувати спроби соціальної інженерії. Стилі спілкування, такі як пряме чи непряме спілкування, можуть впливати на здатність працівників виявляти та реагувати на тактику соціальної інженерії. Тому культурні припущення та переконання відіграють вирішальну роль у формуванні обізнаності та реакції працівників на загрози соціальної інженерії. Культурні норми впливають на те, як працівники реагують на владу, довіряють іншим та спілкуються. У деяких культурах повага до влади може призвести до того, що працівники не вагатимуться ставити під сумнів підозрілі інструкції, знижуючи обізнаність. Культурні фактори тісно пов'язані з організаційними та соціальними факторами. Наприклад, у колективістських культурах соціальний вплив відіграє важливу роль у формуванні поведінки, що ускладнює розділення культурної та соціальної динаміки. У високоієрархічних культурах організаційні фактори, такі як політика безпеки, можуть мати на меті чітко заохочувати працівників оскаржувати владу в підозрілих ситуаціях.

Психологічний фактор:

Він охоплює унікальні риси особистості та емоційні реакції людини, які впливають на її поведінку та схильність до тактик соціальної інженерії. Психологічний фактор складається з трьох підфакторів:

1. П'ять головних рис особистості: це відкритість, сумлінність, екстраверсія, доброзичливість та невротизм. Відкритість означає бути неупередженим, мати уяву та бути готовим до нового досвіду. Сумлінність означає бути організованим, відповідальним та старанним. Екстраверсія означає бути комунікабельним, товариським та енергійним. Доброзичливість означає бути схильним до співпраці, співчутливим та уважним до інших. Нарешті, невротизм означає емоційну реактивність та схильність до переживання негативних емоцій. Ці п'ять рис можуть впливати на сприйнятливість працівників до тактик соціальної інженерії на основі їхніх індивідуальних схильностей та схильностей. Цей підфактор також включає страхи працівників, робоче навантаження, стрес та пильність.

2. Довіра: стосується схильності людини покладатися на інших та довіряти їм. Працівники з вищим рівнем довіри можуть бути більш вразливими до атак соціальної інженерії, які експлуатують їхню довіру до інших.

3. Реактність: стосується психологічної реактності, яку відчувають люди, коли вони сприймають свою свободу чи автономію як загрозу. Працівники з вищим рівнем реактності можуть бути більш стійкими до заходів безпеки та менш схильними дотримуватися стратегій запобігання соціальної інженерії.

Психологічні фактори включають емоційні реакції, такі як страх, довіра або терміновість, які соціальні інженери використовують для маніпулювання співробітниками. Усвідомлення підвищується, коли співробітників навчають керувати емоціями та розпізнавати тактики, що експлуатують емоції, наприклад, ті, що створюють хибне відчуття терміновості або викликають страх. Психологічна стійкість часто посилює поведінкові фактори, оскільки співробітники, які можуть керувати емоціями, менш схильні поспішати з виконанням завдань або ігнорувати протоколи безпеки під тиском. Крім того, на

перцептивні фактори впливає психологічний стан співробітників; стрес, наприклад, може знизити здатність людини помічати деталі, що вказують на напад.

Поведінковий фактор:

Це стосується низки елементів, що стосуються дій, реакцій та поведінки людини в конкретних ситуаціях або контекстах. Ці фактори охоплюють видиму поведінку, звички та реакції людей, а також вибір, який вони роблять у своєму повсякденному житті. У цій категорії можна знайти два підфактори:

1. Ставлення: стосується загальної оцінки або сприйняття людиною загроз соціальної інженерії та заходів безпеки. Позитивне ставлення до безпеки може призвести до підвищення обізнаності та проактивної поведінки у захисті від атак соціальної інженерії. Цей підфактор включає фактичну поведінку та переконання співробітників.

2. Намір: це готовність та мотивація людини до поведінки, що усвідомлює безпеку. Намір вживати заходів безпеки та залишатися пильним щодо соціальної інженерії може суттєво вплинути на обізнаність та реагування працівників на такі загрози. Коли працівники мають намір залишатися поінформованими та обізнаними щодо ризиків соціальної інженерії, вони з більшою ймовірністю будуть проактивно шукати ресурси, брати участь у навчальних програмах та бути в курсі нових загроз.

На поведінкові фактори впливають обізнаність у сфері безпеки та організаційні фактори. Якщо працівників постійно навчають застосовувати належні методи безпеки, їхня поведінка відповідатиме цим урокам. Перцептивні фактори також відіграють певну роль, оскільки сильні навички сприйняття можуть посилювати поведінку в галузі безпеки, дозволяючи працівникам виявляти загрози на ранній стадії та діяти відповідно.

Фактор усвідомлення безпеки:

Він охоплює компоненти, пов'язані з розумінням та поглядами користувачів щодо загальних концепцій інформаційної безпеки, а також їх

усвідомленням важливості захисту як фізичних, так і нефізичних аспектів інформаційної безпеки. Фактор усвідомлення безпеки має три підфактори:

1. Обізнаність у сфері інформаційної безпеки: стосується загальних знань та розуміння співробітниками концепцій та передового досвіду інформаційної безпеки. Цей підфактор включає знання соціальної інженерії, обізнаність про технічні рішення безпеки, знання паролів користувачів, знання/досвід фішингу, знання вірусів, знання пошти, досвід користування Інтернетом/вебсайтом та участь співробітників у заходах з інформаційної безпеки.

2. Практики безпеки [88]: це фактична поведінка та дії співробітників щодо інформаційної безпеки, включаючи їх дотримання політик та протоколів безпеки. Цей підфактор включає захисні практики кібербезпеки, такі як вжиття співробітниками проактивних заходів та поведінки для захисту від кіберзагроз.

3. Обізнаність щодо політик безпеки [88]: стосується знайомства співробітників з політиками та процедурами безпеки організації та їх дотримання. Цей підфактор включає сприйняття корисності інтернет-провайдера та обізнаності про нього.

Обізнаність у сфері безпеки надає базові знання про методи соціальної інженерії (наприклад, фішинг, претексти, цькування), що вимагають від працівників пильності. Обізнаність покращує здатність виявляти загрози та реагувати на них, забезпечуючи розуміння працівниками ландшафту ризиків. Обізнаність у сфері безпеки є необхідною умовою для багатьох інших факторів. Наприклад, перцептивні фактори є неефективними без базових знань, що надаються обізнаністю. Вона також впливає на поведінкові фактори, оскільки працівникам необхідно знати, яка поведінка (наприклад, ретельна перевірка електронних листів) є критично важливою для підтримки безпеки.

Фактор контенту:

Він охоплює методи, що використовуються зловмисниками для переконання користувачів та спонукання їх реагувати на їхні погрози. Соціальні інженери використовують ці елементи в електронних листах, публікаціях та

дзвінках, щоб маніпулювати емоціями жертв та викликати певні реакції. Цей фактор має чотири підфактори:

1. Авторитет: стосується використання авторитетної мови або заяв для встановлення довіри та впливу на сприйняття цільовою стороною легітимності відправника. Коли атаки соціальної інженерії надходять від джерела, яке виглядає авторитетним, такого як високопоставлений керівник або довірений відділ в організації, співробітники можуть бути більш схильні вірити в легітимність комунікації. Таке сприйняття легітимності може зробити співробітників більш вразливими до спроб соціальної інженерії, оскільки вони можуть бути менш схильні ставити під сумнів або ретельно перевіряти інформацію чи запити, які вони отримують. Цей підфактор включає адресу відправника та візуальні підказки, такі як логотипи, символи або зображення, що використовуються для посилення довіри, викликання емоцій або створення відчуття знайомства.

2. Взаємність : це коли зловмисник пропонує щось цінне цілі з очікуванням отримати щось натомість. Цей підфактор включає контекст повідомлення та фактори переконання (тобто винагороди, попередження), які включають використання методів переконання, таких як обіцянка винагороди або висвітлення потенційних негативних наслідків, для впливу на прийняття рішень ціллю. Працівники можуть відчувати себе зобов'язаними відповідати взаємністю, виконуючи запити або надаючи інформацію, навіть якщо у них є застереження або сумніви. Це почуття обов'язку може подолати критичне мислення та скептицизм працівників, роблячи їх більш вразливими до маніпуляцій.

3. Соціальний доказ: це використання соціальних сигналів або посилань для створення відчуття відповідності та переконання цільової групи в тому, що інші вже виконали бажану дію. Він спирається на припущення, що якщо велика кількість людей виконала певну дію, це має бути правильний або доречний вибір. Довіра до цифр може зробити працівників більш вразливими до атак соціальної інженерії, які спираються на сприйняття популярності або консенсусу.

4. Дефіцит: експлуатує сприйняття обмеженої доступності або дефіциту продукту, послуги чи можливості, щоб створити відчуття терміновості та заохотити до негайної реакції. Цей підфактор включає імпульсивність, цікавість та сигнали терміновості, такі як термінова мова або терміни, що спонукають до негайних дій, щоб створити відчуття терміновості та ігнорувати раціональне мислення цільової групи.

Зловмисники часто використовують емоційно заряджений контент, щоб викликати негайну, некритичну реакцію у цілі. До поширених емоційних тригерів належать страх (наприклад, погрози блокування облікового запису), терміновість (наприклад, «Дій зараз або втратиш доступ») або жадібність (наприклад, обіцянка винагороди). Ці емоції затьмарюють судження, через що співробітники більш схильні реагувати, не розглядаючи ретельно справжність повідомлення. Емоційні тригери тісно пов'язані з психологічними факторами, оскільки ступінь, до якої співробітники можуть керувати своїми емоціями, визначає рівень їхньої вразливості. Перцептивні фактори також відіграють певну роль, оскільки співробітники повинні знати про поширені методи емоційного маніпулювання, щоб не стати жертвою зловмисних осіб. Крім того, використання влади перетинається з культурними факторами, особливо в ієрархічних організаціях або культурах, де повага до влади глибоко вкорінена. Співробітники, які культурно схильні поважати владу беззаперечно, можуть бути більш вразливими до цієї техніки, якщо їх не навчити перевіряти такі повідомлення. На рисунку 2.1 подана багатофакторна модель оцінювання ризику соціоінженерних атак.

Організаційні фактори	**Соціальні фактори**	**Психологічні фактори**	**Культурні фактори**
Багатофакторна модель оцінювання ризику соціоінженерних атак			
Поведінкові фактори	**Фактори усвідомлення безпеки**		**Фактори контенту**

Рис. 2.1. Багатофакторна модель оцінювання ризику соціоінженерних атак

Підсумовуючи, звертаючи увагу на механізми та взаємозв'язки цих факторів, організації можуть адаптувати свої навчальні програми, щоб допомогти співробітникам розпізнати стратегії, що використовуються для здійснення атак соціальної інженерії. Навчання, що включає реальні приклади та рольові ігри, може бути особливо ефективним для підвищення обізнаності про те, як хакери обманюють та експлуатують співробітників.

2.2. Методи виявлення та протидії загрозам ризику атак соціальної інженерії

Методи виявлення та протидії атакам

Для зменшення ризиків атак соціальної інженерії застосовуються технічні, організаційні та освітні заходи. До них належать:

–використання алгоритмів машинного навчання для виявлення аномальної поведінки;

–автоматичне блокування підозрілих акаунтів;

–двофакторна автентифікація;

–фільтрація шкідливих посилань;

–навчання користувачів основам кібергігієни.

Однак жоден із методів не є універсальним, що зумовлює необхідність комплексного підходу.

Штучний інтелект як засіб протидії загрозам безпеці

Машинне навчання на основі штучного інтелекту пропонує високоефективне рішення для боротьби з такими загрозами, як соціальна інженерія та інші ризики кібербезпеки. Аналізуючи величезні обсяги даних, ці системи можуть передбачати закономірності та виявляти загальні тенденції, які використовують потенційні хакери AI SYSTEMS EXCEL розпізнавання незвичайних дій або поведінки, які можуть вказувати на майбутню атаку.

Крім того, ШІ не тільки допомагає виявляти ці загрози в режимі реального часу, але й надає практичні рекомендації щодо зменшення вразливості та посилення захисту. Завдяки постійному навчанню та адаптації, рішення на основі ШІ відіграють ключову роль у боротьбі з постійно еволюціонуючими кіберзагрозами та забезпеченні надійної безпеки у все більш цифровому світ

На рисунку 2.2 представлено узагальнені можливості штучного інтелекту та машинного навчання в протидії атакам соціальної інженерії.

Можливості штучного інтелекту та машинного навчання
Виявлення загроз (фішинг, шкідливе ПЗ, несанкціонований доступ)
Прогнозування ризиків (аналіз аномалій, попередження інцидентів)
Зменшення людської вразливості (соціальна інженерія, фішинг)
Автоматизація реагування (ізоляція, блокування, контрзаходи)
Запобігання шахрайству (фінансові та транзакційні ризики)

Рис. 2.2 Можливості штучного інтелекту та машинного навчання в протидії атакам соціальної інженерії

1) *Штучний інтелект у виявленні загроз/шахрайства*:

Штучний інтелект (ШІ) значно розширив наші підходи до мислення та вирішення проблем, допомагаючи нам подолати побоювання, пов'язані з кібербезпекою, завдяки своїй здатності виявляти та прогнозувати загрози. Використовуючи великі обсяги даних, ШІ чудово справляється з виявленням закономірностей та аномалій, які можуть залишитися непоміченими традиційними системами. Системи на основі ШІ здатні постійно контролювати мережеву активність, виявляти незвичайну поведінку та позначати потенційні кібератаки, такі як фішинг та шкідливе програмне забезпечення. Ці системи можуть аналізувати підозрілу активність та спроби несанкціонованого доступу, вивчаючи поведінкові закономірності та звички використання своїх

користувачів. Такий проактивний підхід не тільки посилює заходи безпеки, але й зменшує ризик успішних кібератак, що робить ШІ незамінним інструментом у сучасних системах кібербезпеки.

2) *Прогнозні можливості ШІ*:

Штучний інтелект (ШІ) дуже добре справляється з прогнозуванням. Він використовує дані минулого, щоб передбачити можливі ризики та вжити заходів, перш ніж вони перетворяться на серйозні проблеми. Системи ШІ можуть виявляти аномалії, включаючи незвичайні схеми транзакцій, які можуть вказувати на шахрайство або порушення безпеки, шляхом аналізу величезних масивів даних. ШІ також дуже добре виявляє вразливі місця та збої в роботі життєво важливої інфраструктури, генеруючи попереджувальні сигнали, що дозволяють вжити превентивних заходів для зменшення ризиків. Забезпечуючи оперативну оптимізацію та ремонт системи, ця функція скорочує час простою та підвищує ефективність роботи. ШІ є необхідним для збереження надійності та безпеки сучасних технологічних екосистем завдяки використанню його прогнозних можливостей.

3) *Зменшення вразливості людей*:

Хакери часто використовують вразливість людей за допомогою атак соціального інжинірингу — маніпулюючи людьми, щоб отримати несанкціонований доступ до конфіденційних даних, зламати облікові записи або скоїти шахрайство. Ці атаки базуються на психологічній маніпуляції, що ускладнює протидію їм за допомогою традиційних заходів безпеки. Однак штучний інтелект (ШІ) забезпечує ефективні заходи протидії для зменшення цих загроз. ШІ може виявляти спроби фішингу, аналізуючи вміст електронних листів, шаблони повідомлень та метадані для виявлення підозрілих дій. Сучасні алгоритми здатні позначати шахрайські електронні листи та попереджати користувачів про потенційні ризики в режимі реального часу. Системи ШІ можуть надавати автоматизовані відповіді на підозрілі запити, мінімізуючи участь людини в критичних процесах прийняття рішень, де ймовірність помилок є високою.

Поєднуючи аналіз повідомлень з інтелектуальною автоматизацією, ШІ не тільки зменшує ймовірність успішних атак, але й підвищує загальну стійкість кібербезпеки. Цей проактивний підхід допомагає організаціям та окремим особам захищати свої дані та облікові записи від дедалі більш витончених тактик соціальної інженерії.

4) *Автоматизація реагування на загрози*:

ШІ не тільки виявляє загрози, але й активно реагує на них за допомогою автоматизованих систем реагування на інциденти. Ці системи можуть ізолювати скомпрометовані пристрої або мережі, щоб обмежити потенційний збиток і запобігти подальшому поширенню загроз. Крім того, ШІ може вживати контрзаходів у режимі реального часу, наприклад блокувати підозрілі IP-адреси або відключати точки доступу, пов'язані зі зловмисною діяльністю. Діючи швидко та ефективно, системи реагування на основі ШІ зменшують залежність від ручних втручань, мінімізують збитки та підвищують загальний рівень безпеки. Ця здатність дозволяє організаціям підтримувати надійний захист від постійно еволюціонуючих кіберзагроз, забезпечуючи більшу стійкість і захист критично важливих систем і даних.

Штучний інтелект і машинне навчання в кібербезпеці

Штучний інтелект (ШІ) та машинне навчання (МН) значно покращили кібербезпеку завдяки вдосконаленню механізмів безпеки та підвищенню ефективності та результативності виявлення, запобігання та реагування на загрози. Стандартні підходи до безпеки часто залежать від заздалегідь визначених правил та ручного нагляду, що є недостатнім для протидії зростаючій складності та витонченості сучасних кібератак. ШІ та МН долають ці обмеження, аналізуючи величезні масиви даних у режимі реального часу, що дозволяє їм виявляти закономірності, аномалії та нові загрози з надзвичайною швидкістю та точністю. Ці передові технології не тільки зміцнюють системи захисту, але й адаптуються до нових методів атак, забезпечуючи надійні, проактивні рішення у сфері кібербезпеки. Нижче наведено деякі застосування ШІ та машинного навчання у кібербезпеці:

1) Виявлення загроз:

Аналізуючи мережеву активність, системні журнали та моделі поведінки, системи на базі штучного інтелекту можуть ефективно виявляти такі загрози, як шкідливе програмне забезпечення, спроби фішингу та несанкціонований доступ. Моделі машинного навчання покращують цей процес, постійно адаптуючись і підвищуючи свою точність з часом, забезпечуючи більш точне виявлення загроз і кращу загальну безпеку. Алгоритм контрольованого навчання під назвою Random Forest застосовується як для завдань класифікації, так і для завдань регресії. Він будує багато дерев рішень під час навчання і комбінує їх результати для отримання прогнозів. Ансамбль дерев покращує перенавчання, обсяг і точність. Він використовується для виявлення спам-листів (рис. 2.3).

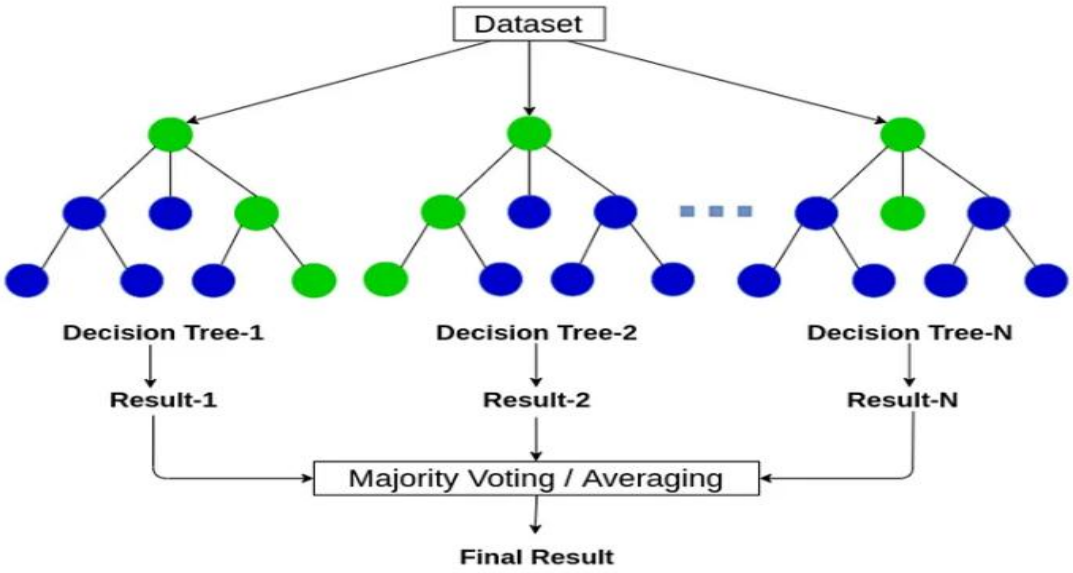

Рис. 2.3. Random Forest – побудова дерева рішень

2) *Запобігання шахрайству*:

Одним з ключових факторів кібербезпеки є запобігання шахрайству, особливо в таких галузях, як банківська справа, електронна комерція та фінанси, де шахрайські дії можуть призвести до значних фінансових втрат, а також шкоди репутації організації. Завдяки виявленню закономірностей та невідповідностей, що вказують на сумнівну діяльність, штучний інтелект (ШІ) та машинне

навчання (МН) стали важливими інструментами у боротьбі з шахрайством в останні роки. Модель виявлення шахрайства представлена на рисунку 2.4.

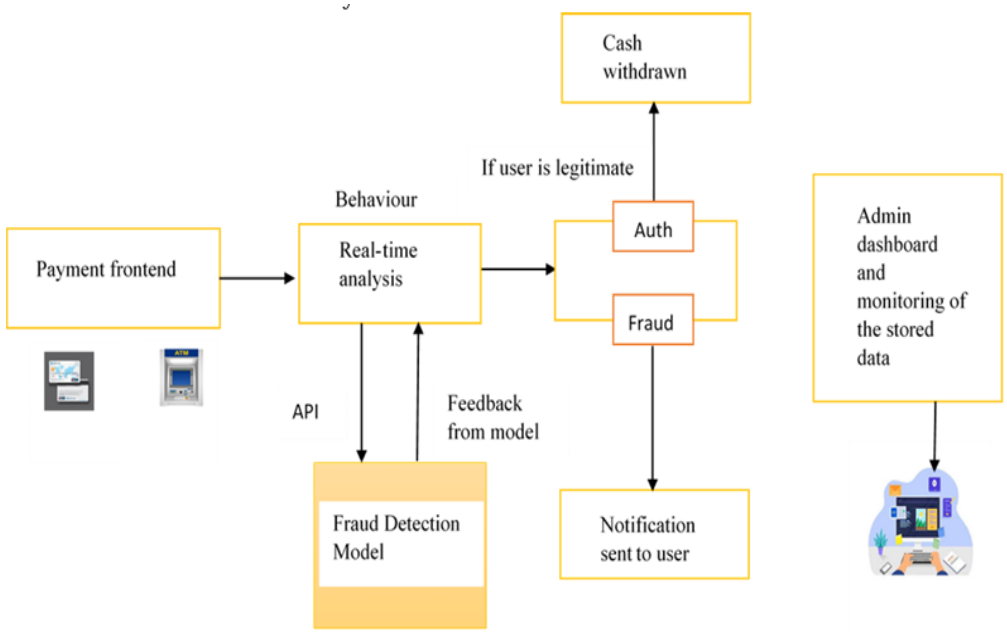

Рис.2.4. Модель виявлення шахрайства [90]

3) *Виявлення аномалій*:

Виявлення аномальних закономірностей або відхилень у даних, що відрізняються від очікуваної поведінки, називається виявленням аномалій. Виявлення аномалій є необхідним у кібербезпеці для виявлення можливих шахрайств, системних помилок або порушень безпеки [89]. ШІ та МН можуть динамічно вивчати закономірності на основі даних, що робить їх більш успішними у виявленні ризиків, які ще не були ідентифіковані. Деякі ключові області виявлення аномалій — це адаптивне навчання, збір даних, розпізнавання закономірностей та аналіз у реальному часі.

4) *Виявлення фішингу та спаму*:

Фішинг та спам є домінуючими ризиками кібербезпеки, які часто використовуються для обману людей з метою розкриття приватної інформації або завантаження шкідливих файлів. Аналізуючи різноманітні ознаки, включаючи вміст електронної пошти, поведінку відправника та шаблони

комунікації, ШІ та МН є невід'ємною частиною виявлення та зупинення спроб спаму та фішингу.

Виклики у впровадженні заходів протидії на основі ШІ та МН

Хоча впровадження заходів протидії на основі машинного навчання (МН) та штучного інтелекту (ШІ) у сфері кібербезпеки є перспективною стратегією, існує кілька перешкод, які необхідно подолати.

Нижче наведено перелік деяких основних викликів, з якими стикаються організації:

- Доступність та якість даних: для ефективної роботи моделей ШІ та МН необхідні величезні масиви даних. Неточні прогнози можуть бути наслідком упереджених масивів даних, низької якості даних або відсутності інформації, що знижує здатність систем виявляти небезпеки.
- Високі витрати на впровадження: для розробки та впровадження рішень на основі штучного інтелекту потрібні значні інвестиції в інфраструктуру, кваліфікований персонал та інструменти. Це може стати значною перешкодою, особливо для малих та середніх підприємств (МСП) з обмеженими фінансовими ресурсами.
- Атаки зловмисників: Кіберзлочинці все частіше використовують ворожі тактики, щоб експлуатувати слабкі місця моделей ШІ. Наприклад, хтось може непомітно змінити вхідні дані, щоб обдурити систему і зробити її марною.
- Складність інтеграції: Інтеграція систем ШІ та МН в існуючі системи кібербезпеки може бути складною і трудомісткою. Може знадобитися перепроектування історичних систем і забезпечення їх сумісності з існуючими інструментами та процесами.
- Відсутність кваліфікованих фахівців: Для ефективного впровадження штучного інтелекту та машинного навчання в кібербезпеці необхідні знання як в області науки про дані, так і в області кібербезпеки. Однак глобальний дефіцит фахівців з необхідними навичками перешкоджає впровадженню.

- Негативні та помилкові позитивні результати: Моделі штучного інтелекту (ШІ) можуть генерувати помилкові позитивні результати, повідомляючи про нешкідливу активність як про загрозу, або помилкові негативні результати, пропускаючи реальні небезпеки. Обидві ситуації можуть призвести до операційної неефективності та зниження довіри до системи.

Як кібербезпека виграє від ШІ та МН

Машинне навчання та штучний інтелект пропонують кілька ключових переваг у кібербезпеці, зокрема:

- Підвищена точність: у міру того, як алгоритми машинного навчання набувають знань і пристосовуються до нових даних, вони поступово підвищують точність своїх виявлень. Це зменшує кількість помилкових негативних результатів (пропущення реальних загроз) і помилкових позитивних результатів (неточне позначення нешкідливої поведінки).

- Пріоритезація загроз: може бути складно вирішити, яким реакціям надати пріоритет, коли існує так багато можливих небезпек. Команди з безпеки можуть ефективніше розподіляти ресурси, використовуючи машинне навчання (ML) для оцінки тривог на основі їх ймовірності та серйозності.

- Швидший аналіз даних: команди з безпеки обробляють величезні обсяги даних з мережевого трафіку, брандмауерів та інших джерел. Системи ML аналізують ці дані набагато швидше, ніж люди, виявляючи тенденції та невідповідності, які можуть вказувати на можливі загрози.

- Підвищена автоматизація: машинне навчання (ML) може використовуватися для автоматизації повторюваних і трудомістких процесів, таких як відбір помилкових спрацьовувань з сигналів тривоги або аналіз файлів журналів. В результаті аналітики з безпеки можуть зосередитися на стратегічних ініціативах.

- Покращене виявлення загроз: методи атак, які використовують кіберзлочинці, постійно змінюються. Проактивна оборона стає можливою завдяки здатності ML аналізувати минулі атаки та виявляти найменші зміни в поведінці, які можуть вказувати на нові небезпеки.

Недостатня обізнаність співробітників та клієнтів у питаннях безпеки значно сприяє шахрайським діям. Для зменшення цих загроз необхідно формувати культуру обізнаності в питаннях безпеки. Це передбачає регулярне поширення відповідних ресурсів та навчальних програм, які інформують зацікавлені сторони про поширені вектори атак та запобіжні заходи.

Використання штучного інтелекту (ШІ) та машинного навчання (МН) є перспективним рішенням для боротьби з шахрайськими діями. Ці технології слугують для виявлення аномалій, ідентифікації нових моделей атак та автоматизації реагування на потенційні загрози. Однак впровадження ШІ та МН у кібербезпеку супроводжується фінансовими та технічними викликами. Організації повинні інвестувати в надійні історичні набори даних та всебічне навчання моделюванню, щоб забезпечити ефективність цих інструментів. Незважаючи на початкові витрати, ці технології забезпечують довгострокові переваги, покращуючи виявлення загроз, скорочуючи час реагування та підвищуючи загальну стійкість безпеки. Важливе поєднання підходів, орієнтованих на людину, з технологічними досягненнями для ефективного протидії загрозам кібербезпеки.

Інтеграція штучного інтелекту (ШІ) та машинного навчання (МН) у кібербезпеку є важливим кроком у боротьбі зі все більш витонченими кіберзагрозами. Зі значним зростанням кількості атак із використанням соціальних інженерних методів та інших зловмисних механізмів, існує гостра потреба у надійних системах на базі ШІ, здатних як виявляти, так і запобігати таким діям. Хоча інформаційні кампанії про соціальну інженерію набули популярності, вони самі по собі є недостатніми для протидії постійно розвиваючимся тактикам зловмисників, які використовують досягнення технологій для вдосконалення своїх методів.

Майбутнє штучного інтелекту в кібербезпеці полягає в просуванні цільових досліджень, які розширюють можливості сучасних технологій, одночасно усуваючи їхні обмеження. Стратегічна інтеграція штучного інтелекту в системи безпеки вимагає пильного нагляду, щоб збалансувати використання

передових можливостей із дотриманням вимог системи та етичні стандарти. Це вимагає ретельного тестування, постійного моніторингу поведінки ШІ та створення механізмів, що забезпечують підзвітність за прийняття рішень на основі ШІ. Створення потужних систем, які розпізнають і запобігають складним маніпуляціям, має бути головним напрямком майбутніх досліджень. Для цього нам потрібна мультидисциплінарна стратегія, яка сприятиме співпраці між урядом, бізнесом та науковими колами з метою встановлення найкращих практик і стандартів. Спільнота кібербезпеки може використовувати революційний потенціал ШІ для створення безпечнішого цифрового середовища, зосередившись на етичному використанні, надійній безпеці та технологічних інноваціях.

Висновки до розділу 2

Визначено, що для формалізації багатофакторного підходу до оцінювання ризику соціоінженерних атак значну роль відіграють основні фактори впливу, які важливо структурувати для формування стратегії моніторингу та відбиття атак соціальної інженерії. Основними факторами визначено: організаційні фактори, які включают: лідерство, превентивні контрзаходи, детективні контрзаходи, процедурні контрзаходи, програму SETA; соціальні фактори, які складаються з п'яти підфакторів: суб'єктивних норм, афективної відданості, зобов'язань щодо продовження роботи, впливу ЗМІ, нормативних зобов'язань; культурні фактори, які грунтуються на культурних нормах та цінностях, формують розуміння людьми того, що вважається прийнятним або неприйнятним з точки зору практики безпеки; психологічні фактори, які передбачають маніпулювання рисами особистості: відкритістю, сумлінністю, екстраверсією, доброзичливістю та невротизмом, довірою та реактністю. Поведінкові фактори стосуються низки елементів: дій, реакцій та поведінки людини в конкретних ситуаціях або контекстах. Фактори усвідомлення безпеки охоплюють компоненти, пов'язані з розумінням та поглядами користувачів щодо

загальних концепцій інформаційної безпеки, а також їх усвідомленням важливості захисту як фізичних, так і нефізичних аспектів інформаційної безпеки. Фактор усвідомлення безпеки має три підфактори: обізнаність у сфері інформаційної безпеки, практики безпеки, обізнаність щодо політик безпеки. Фактор контенту охоплює методи, що використовуються зловмисниками для переконання користувачів та спонукання їх реагувати на їхні погрози. Соціальні інженери використовують ці елементи в електронних листах, публікаціях та дзвінках, щоб маніпулювати емоціями жертв та викликати певні реакції.

Доведено, що для зменшення ризиків атак соціальної інженерії застосовуються технічні, організаційні та освітні заходи. До них належать: використання алгоритмів машинного навчання для виявлення аномальної поведінки; автоматичне блокування підозрілих акаунтів; двофакторна автентифікація; фільтрація шкідливих посилань; навчання користувачів основам кібергігієни.

Зазначено, що машинне навчання на основі штучного інтелекту надає високоефективне рішення для боротьби з такими загрозами, як соціальна інженерія та інші ризики кібербезпеки. Аналізуючи величезні обсяги даних, ці системи можуть передбачати закономірності та виявляти загальні тенденції, які використовують потенційні хакери AI SYSTEMS EXCEL розпізнавання незвичайних дій або поведінки, які можуть вказувати на майбутню атаку.

РОЗДІЛ 3
ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА ОЦІНЮВАННЯ МОДЕЛІ У КОРПОРАТИВНОМУ СЕРЕДОВИЩІ

3.1. Аспекти впровадження моделі оцінювання ризику у корпоративному середовищі

Ефективні кібератаки, засновані на соціальної інженерії, спираються на людські вразливості. Зловмисники використовують людську поведінку, знання, емоції, когнітивні функції, особистісні риси, людську природу тощо. У таблиці 3.1 висвітлено людські вразливості, пов'язані з атаками соціальної інженерії. Одна атака соціальної інженерії може бути здійснена з використанням кількох методів впливу. Один метод впливу може використовувати кілька людських вразливостей. Цей багатовимірний взаємозв'язок між атаками соціальної інженерії, методами впливу та людськими вразливостями робить кібератаки на основі соціальної інженерії складними для фахівців з безпеки.

Таблиця 3.1

Зв'язок між кібератаками на основі соціальної інженерії, методом впливу та вразливістю людини.

№	Атака	Методи впливу на жертв	Експлуатація людської вразливості
1	Видавання себе за іншу особу, компрометація ділової електронної пошти (BEC), клонування фішингових даних	Моральний вплив, соціальна відповідальність, подібність, переконання з використанням авторитету/довіри, теорія міжособистісного обману (TMO), ускладнення процесу мислення, допитливість	Бути корисним, милосердним, добрим, намагатися бути прийнятним у соціальних нормах
2	Фішинг під час списування, претекстування, смішинг-фішинг, вейлінг-фішинг, вішинг-фішинг, діпфейк	Подібність, переконання з використанням авторитету/довіри, враження/відданість, поведінка впливає на ставлення, дефіцит, обман/шахрайство, ускладнення процесу мислення	Бути корисним, бути слухняним владі, допомагати природі, панічна недбалість
3	Фішинг у соціальних мережах, програмне забезпечення для залякування, атака зворотним проектуванням, діпфейк	Груповий вплив, інформативний вплив/нормативний вплив, соціальний обмін, теорія/норма взаємності, моральний вплив/соціальна відповідальність	Дружній характер, недбалість, довірливий характер, авторитетність

№	Атака	Методи впливу на жертв	Експлуатація людської вразливості
4	Атака на водопой, діпфейк	Переконання з використанням авторитету/достовірності, ефект фреймінгу/когнітивне упередження, ускладнення процесу мислення	Цікавість, жадібність, хвилювання, страх

Аналіз змісту таблиці 3.1 надає детальне уявлення про функціонування атак соціальної інженерії. Таке зіставлення може значно допомогти у виявленні вразливостей та побудові ефективної інфраструктури безпеки для їх пом'якшення. Виділений взаємозв'язок між людськими вразливостями та атаками соціальної інженерії може бути недостатнім для представлення кожної вікової групи чи людської поведінки; однак, воно може надати окремим особам та організаціям ширше уявлення про розуміння та протидію цим кібератакам, що базуються на соціальній інженерії.

Протидія кібератакам на основі соціальної інженерії

Людську обізнаність можна визначити як ключовий фактор у протидії атакам соціальної інженерії. Методи соціальної інженерії зосереджені на зломі людей, орієнтуючись на когнітивні упередження людини, а не на машині. Методи протидії кібератакам на основі соціальної інженерії можна побачити в таблиці 3.2, галочкою позначено контрзаходи, запропоновані для реалізації

Таблиця 3.2

Методи протидії атакам соціальної інженерії

	Навчання	Політика кібербезпеки	Політика комунікації	Обладнання компанії	Фільтр спаму/ Антивірус/ Брандмауер	Зашифрований зв'язок	Керування паролями/даними	Звіт про інцидент
[48]	✓	✓	✓		✓	✓	✓	
[49]	✓	✓	✓		✓	✓		
[35]	✓				✓			✓
[40]	✓	✓			✓	✓	✓	✓

	Навчання	Політика кібербезпеки	Політика комунікації	Обладнання компанії	Фільтр спаму/ Антивірус/ Брандмауер	Зашифрований зв'язок	Керування паролями/даними	Звіт про інцидент
[50]	✓							
[51]	✓	✓	✓					
[52]	✓	✓	✓	✓				
[53]	✓	✓						
[54]	✓		✓	✓				
[55]	✓	✓	✓					
[56]	✓	✓			✓			
[57]	✓	✓		✓				
[58]	✓	✓						
[59]	✓	✓			✓	✓	✓	
[60]	✓		✓		✓			

Аналіз наукових праць дозволив стверджувати, що більшість дослідників вважають наступні методи дуже ефективними для протидії атакам соціальної інженерії. Аналізуючи таблицю 3.2, можна зробити висновок, що навчання та підвіщення обізнаності людей з питань кібербезпеки та атак соціальної інженерії може відігравати значну роль.

Кілька дослідників підкреслили роль чітко визначених політик для протидії кібератак. Політик, які можна впровадити для запобігання та управління випадками витоку даних або кібератак на основі соціальної інженерії. Організаційні політики далі поділяються на дві основні групи: політика кібербезпеки та політика комунікації. Політики кібербезпеки визначені спеціально для кібератак. Такі політики можуть включати інструкції щодо уникнення незаконного програмного забезпечення, використання особистих пристроїв у мережі компанії, кроки, які слід виконувати у разі кібератак,

документування кібератак, процедури управління персоналом (HR) для привілеїв та доступу сторонніх постачальників, управління доступом до критичних зон, управління паролями, інфраструктуру організаційної безпеки тощо.

Чітко визначена політика кібербезпеки може обмежити багато витоків даних або кібератак. Крім того, слід впроваджувати організаційні політики щодо офіційного (або в деяких випадках неофіційного) спілкування. Причина впровадження окремого набору політик щодо комунікації полягає у великій кількості атак соціальної інженерії, які використовують норми комунікації для проведення кібератак. Організація повинна визначити чіткі політики щодо офіційного спілкування всередині та за межами організації. Такі політики можуть включати процес затвердження та підтвердження підключення особистого пристрою до мережі організації, рівень інформації, який можна передавати в електронних листах, SMS або дзвінках, процедури перевірки автентичності підозрілих електронних листів, SMS або дзвінків, методи комунікації у разі роботи на аутсорсі. Політики організаційної комунікації можуть відігравати важливу роль у запобіганні будь-яким кібератак, заснованим на соціальної інженерії. Важливість та необхідність відповідних антивірусів, брандмауерів, спам-фільтрів та оновлених програмних патчів не можна недооцінювати, і вони повинні бути встановлені як в організаційних, так і в особистих системах. Деякі дослідницькі та оглядові роботи щодо пом'якшення атак соціальної інженерії заохочують організації надавати співробітникам корпоративне обладнання, оскільки обладнання, яким керує IT-відділ організації, можна легко оновити за допомогою програмного забезпечення безпеки та часто перевіряти на наявність шкідливого програмного забезпечення.

Контрзаходи на основі машинного навчання

Машинне навчання (ML) – це ще одна галузь, яка може відігравати важливу роль у протидії кібератакам на основі соціальної інженерії. Для фішингових атак моделі ML можна навчити розпізнавати шаблони та мову в електронних листах, SMS, шкідливих посиланнях і навіть дзвінках за допомогою

обробки природної мови (NLP) [35; 51]. Однак, постійна еволюція характеристик фішингу може викликати занепокоєння у методів на основі ML. Деякі з найважливіших методів ML для протидії кібератакам на основі соціальної інженерії слід обговорити та закцентувати на існучих проблемах.

Глибоке навчання

Підходи на основі глибокого навчання (DL) також можуть відігравати життєво важливу роль у протидії кібератакам на основі соціальної інженерії, оскільки DL є ефективним підходом, що використовується для протидії широкому спектру шкідливих програм, фішингових атак, аналізу трафіку, виявлення спаму, виявлення вторгнень тощо [61;62;63;64;65;66]. Наприклад, глибокі нейронні мережі (DNN) у DL натхненні людським мозком. Зі збільшенням кількості даних, що надходять до DNN, вона поступово стає кращою у виявленні шкідливих діалогів. З цієї причини Google також використовує нейронні мережі для виявлення спам-листів [52]. Коли справа доходить до фішингу, рішення на основі DNN можуть бути дуже ефективними. У [67] автори запропонували гібридну модель, засновану на DNN та довгій короткочасній пам'яті (LSTM), для виявлення фішингових веб-посилань. Автори використовували NLP для вибору функцій та функцій на основі вбудовування символів для моделі DNN-LSTM для виявлення фішингових посилань на веб-сайти. Модель була навчена на двох наборах даних — Ebbu2017 та вторинному наборі даних, який базувався на кількох інтернет-ресурсах. Навіть з огляду на високий рівень виявлення запропонованою моделлю, автори висловили занепокоєння щодо використаних наборів даних. Використані набори даних можуть не мати точного відображення реальних атак на робочому місці. У статтях [68;69] також представлені підходи на основі DL для протидії атакам на основі соціальної інженерії, ініційованим через Twitter, DNS, URL та електронну пошту. Автори представили детальний аналіз походження програм-вимагачів на основі різних тематичних досліджень. Однак відсутність наборів даних, що представляють складні атаки на основі соціальної інженерії, може бути ключем до покращеного DL для протидії кібератакам на основі соціальної інженерії. Це

занепокоєння висвітлено К. Сімаром та ін. у [70]; у своєму дослідженні вони представили підхід на основі DL для структурування неструктурованих даних, згенерованих різними інтернет-джерелами. Однак, обробка дезінформації може бути проблемою в запропонованому підході. Під час перевірки достовірності джерела публікація події чи статті все ще може бути сумнівною. Тим не менш, якщо процес перевірки джерела можна буде вдосконалити, запропонований підхід може відіграти важливу роль у вдосконаленні підходу на основі DL проти атак соціальної інженерії.

Навчання з підкріпленням

Навчання з підкріпленням (RL), ще один аспект ML, також є методом, що використовується для протидії кібератакам на основі соціальної інженерії. У [71] автори запропонували механізм кіберстійкості (CRM) для протидії онлайн-загрозам, включаючи невизначені сценарії реального часу. Модель використовувала архітектуру зворотного зв'язку RL для визначення політик, оскільки система спостерігає за діями в Інтернеті. Однак модель вимагає онлайн-спостережень для вивчення та адаптації невідомих методів атаки, що використовуються в кібератаках на основі соціальної інженерії. У [72] автори використовували жадібний підхід на основі RL. Автори використовували попередньо визначені підходи до атаки та захисту (наприклад, мережу Петрі) для навчання моделі RL. На основі результатів експериментів автори дійшли висновку, що модель поступово покращувала свою продуктивність для виявлення кібератак. Метою статті було висвітлити потенціал RL для протидії кібератакам. Основною проблемою підходу на основі RL є спостереження за необмеженим діапазоном людської поведінки. З плином часу дані, пов'язані з людською поведінкою, зростатимуть експоненціально, що ускладнюватиме відстеження та зберігання інформації [73].

Обробка природної мови

Одним із найзручніших інструментів машинного навчання для протидії кібератакам на основі фішингу є NLP [74]. NLP з ML відіграє важливу роль у протидії фішинговим атакам [75]. Кілька процесів NLP, наприклад, вилучення

інформації, категоризація тексту та машинний переклад, натхненні DL [76]. NLP спирається на п'ять основних ознак для ідентифікації фішингових електронних листів або онлайн-посилань. Ці ознаки - характеристики тіла електронного листа, тема листа, характеристики універсального локатора ресурсів (URL), характеристики прихованого скрипта (тобто JavaScript, спливаюче вікно при кліку тощо) та характеристики відправника. У статті Тіма Репке та ін. [77] використовувався метод вбудованих слів з DL для аналізу тексту електронного листа на предмет ідентифікації людського режиму. Цей метод не використовується для виявлення фішингу, але він може бути корисним для виявлення аномалій у звичайному тексті електронної пошти, що може допомогти у виявленні фішингу електронною поштою, тобто імперсонації, BEC, клонованого фішингу тощо. Єдиною проблемою для моделі NLP на основі машинного навчання (ML) є залежність від поверхневого тексту електронного листа. Якщо структура діалогу або речення змінена, моделі стає важко ідентифікувати це як фішинг [76]. Однак ключовою проблемою для ML для протидії кібератакам на основі соціальної інженерії є відсутність шаблону атаки або методології, яка може ідентифікувати багатовимірний підхід до кібератак на основі соціальної інженерії [66]. Атаки на основі соціальної інженерії використовують людські вразливості, що виходить за рамки традиційних підходів безпеки в інформатиці. Тому психологічне прийняття рішень та когнітивні упередження, пов'язані з кібератаками на основі соціальної інженерії, є постійними проблемами для підходів на основі ML [78]. Щоб досягти вищого рівня виявлення, дослідник повинен дослідити лінгвістичні особливості соціальної інженерії та інтегрувати когнітивні та психологічні фактори в підходи на основі ML.

Маніпулювання людською поведінкою та емоціями під час кібератак є невідомою та складною змінною для експертів з безпеки. Основною причиною цієї змінної загалом можна назвати культуру. Культура може відігравати важливу роль у впливі на людську поведінку, переконання, мораль, рішення та ставлення [79;80]. Навіть з технологічним прогресом у безпеці, люди можуть

бути використані через свої вразливості. Підвищення обізнаності про кібербезпеку шляхом навчання та освіти є важливим. Така обізнаність може допомогти у зменшенні кібератак, заснованих на соціальної інженерії. З іншого боку, деякі дослідження [81;82;83] підкреслюють, що, незважаючи на відповідне навчання та політику, людські вразливості все ще можуть бути використані через атаки соціальної інженерії. Наприклад, не кожна людина, яка працює в організації, має базові знання комп'ютерної безпеки. Навчання співробітника без попередніх знань комп'ютерної безпеки може бути дорогим, трудомістким і не дуже ефективним. Такі співробітники дуже вразливі до кількох атак соціальної інженерії на основі фішингу.

Можна помітити, що існує явний розрив між складними атаками соціальної інженерії та існуючими контрзаходами. Відсутність ефективних підходів до запобігання та уникнення кібератак на основі соціальної інженерії є постійною проблемою для експертів з безпеки. Для ефективної протидії цим атакам соціальної інженерії необхідні багатовимірні контрзаходи, засновані на людських вразливостях та технічних компонентах. Підходи на основі машинного навчання демонструють високу ефективність у протидії кібератакам на основі соціальної інженерії; Однак, все ще існує потреба в подальшому вдосконаленні методів на основі машинного навчання (ML).

3.2. Рекомендації щодо зниження ризику та модель захисних заходів від атак соціальної інженерії

Підхід Agile, відомий своєю гнучкістю, сприяє постійній співпраці між членами команди та зацікавленими сторонами через цикли планування, виконання та оцінки [9]. Ця методологічна основа дозволяє дослідженню оцінити як технічні, так і людські фактори в кібербезпеці, з особливим акцентом на тому, як атаки соціального інжинірингу використовують людські вразливості, та ролі шифрування у запобіганні таким атакам [10].

Цей підхід сприяє швидшому наданню цінності, підвищенню якості результатів та збільшенню адаптивності до мінливих вимог. Сприяння більш оперативному та ітеративному циклу розробки дозволяє командам постійно надавати та вдосконалювати рішення, що відповідають потребам користувачів та мінливим пріоритетам. Оптимальгим є Scrum серед різних Agile-фреймворків через його чітко визначену структуру, яка включає чітко визначені ролі, чітко окреслені артефакти та обмежені в часі заходи, призначені для оптимізації співпраці та ефективності команди.

Scrum особливо підходить для управління складними та динамічними проєктами, оскільки його ітеративні та інкрементальні практики забезпечують постійний прогрес та адаптивність. Розбиваючи роботу на керовані спринти, команди можуть зосередитися на наданні функціональних інкрементів продукту, регулярно враховуючи відгуки для поліпшення кінцевого результату.

Ця ітеративна природа не тільки підвищує продуктивність, але й значно скорочує терміни поставки в порівнянні з традиційними методологіями водоспаду. Інтегруючи обробку природної мови (NLP) з гнучкою методологією при розробці архітектури системи ця комбінація забезпечує гнучкий та ітеративний процес проєктування, який дозволяє архітектурі ефективно використовувати навчальні дані. Основною метою цього підходу є розробка заходів протидії атакам соціального інжинірингу шляхом використання здатності NLP аналізувати та інтерпретувати моделі людської мови, тим самим підвищуючи можливості системи в області безпеки. Архітектура запропонованої моделі запобігання соціоінженерним атакам представлена на рисунку 3.1.

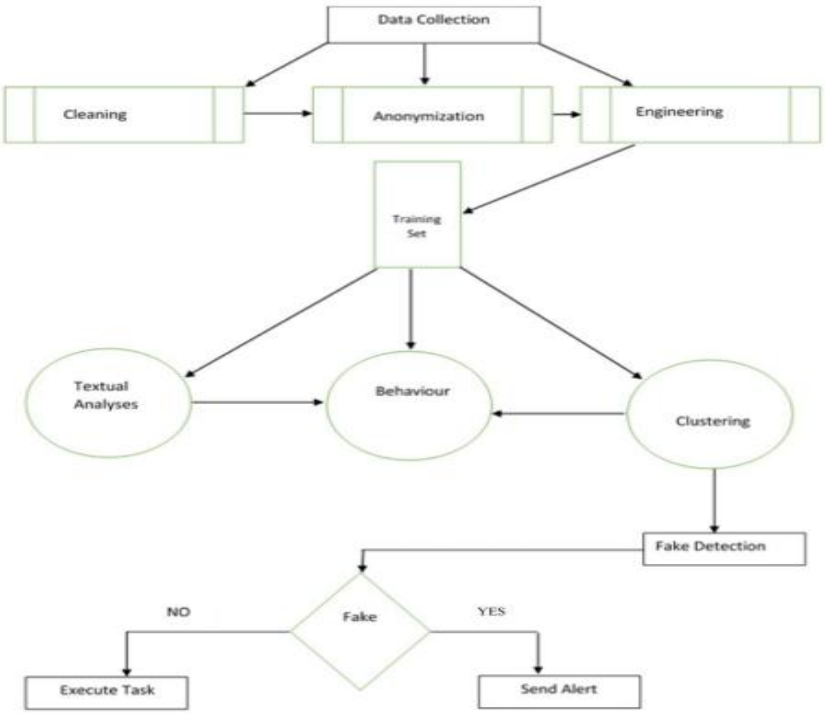

Рис. 3.1. Архітектура моделі запобігання соціоінженерним атакам [90].

Виклики та обмеження сучасних підходів

Наявні механізми захисту стикаються з низкою проблем, зокрема:

- швидка адаптація зловмисників до нових захисних систем;
- складність аналізу людської поведінки;
- етичні та правові обмеження збору даних;
- висока кількість хибнопозитивних спрацювань.

Ці виклики обмежують ефективність автоматизованих систем виявлення атак.

Які методології використовувалися для оцінки ефективності факторів, що впливають на обізнаність співробітників про атаки соціальної інженерії?

Значна кількість досліджень (63%) зосереджена на методології, яка грунтується на опитуванні співробітників організації.

Опитування широко використовуються для збору інформації про фактори впливу, оскільки дані можна зібрати від великої кількості учасників. Цей підхід також, як правило, є економічно ефективним і може бути застосований для

кількох організацій, секторів або демографічних груп, що робить його ідеальним для масштабних досліджень.

Кількісні методи слугують для виявлення загальних тенденцій, вони можуть надмірно спрощувати проблему, надмірно зосереджуючись на числових даних.

Дослідження, які грунтуються на якісних методах, спрямовані на вивчення та розуміння факторів, що впливають на обізнаність працівників щодо соціальної інженерії, з більш суб'єктивної та глибшої точки зору.

Якісні дослідження дозволяють провести глибший аналіз того, як співробітники інтерпретують та реагують на потенційні атаки соціальної інженерії, надаючи розуміння, яке кількісні методи можуть не охопити. Наприклад, інтерв'ю та спостереження можуть виявити контекстуальні фактори, такі як організаційна культура чи психологічні рушійні сили, що впливають на обізнаність співробітників. Однак якісні дослідження зазвичай проводяться з меншими вибірками, що може обмежувати узагальнюваність результатів. Крім того, інтерпретаційний характер якісних методів створює ризик упередженості дослідників.

Дослідження зі змішаними методами поєднують як кількісні, так і якісні методи. Перевага цього підходу полягає в тому, що він дозволяє дослідникам доповнити широту кількісних даних глибиною якісних висновків, забезпечуючи більш цілісне уявлення про фактори, що впливають на обізнаність співробітників щодо соціальної інженерії. Наприклад, хоча опитування може виявити конкретні фактори, які корелюють з більшою обізнаністю, дані подальших інтерв'ю або спостережень можуть вказати, чому ці фактори є значними, а також виявити будь-які основні механізми, які можуть бути присутніми. Однак, незважаючи на свої переваги, підхід зі змішаними методами може бути складним для впровадження, оскільки він часто вимагає значного часу, ресурсів та досвіду для збору та інтеграції даних, отриманих обома методами.

Ключові аспекти оцінювання ризиків соціоінженерних атак

Вартість реагування: це сприйняті негативні наслідки або витрати, пов'язані з вжиттям заходів проти загроз соціальної інженерії. Це впливає на обізнаність працівників щодо соціальної інженерії, роблячи їх більш скептичними та обережними.

Уявна вразливість описує переконання працівника у власній вразливості до атак соціальної інженерії. Коли працівники сприймають себе як вразливих до атак соціальної інженерії, вони з більшою ймовірністю будуть мотивовані посилити свою обізнаність та вжити відповідних запобіжних заходів. Цей підфактор включає фактичну вразливість, уявний ризик, уявну загрозу та управління ризиками.

Сприйняття серйозності: стосується сприйняття працівником серйозності та потенційної шкоди, спричиненої атаками соціальної інженерії. Сприйняття серйозності атак соціальної інженерії як високої створює відчуття терміновості серед працівників. Вони розуміють, що наслідки становлення жертвою таких атак можуть бути значними як для них самих, так і для організації. Цей підфактор включає ризиковану поведінку.

Ефективність реагування: це віра працівника в ефективність його/її контрзаходів проти атак соціальної інженерії. Коли працівники вважають, що певна поведінка або дії можуть ефективно захистити їх від атак соціальної інженерії, вони більш мотивовані до такої поведінки. Цей підфактор включає впевненість працівників у зобов'язанні організації захищати їхню інформацію та вирішувати ризики соціальної інженерії.

Обробка інформації: стосується того, як співробітники сприймають, інтерпретують та обробляють інформацію, пов'язану із загрозами соціальної інженерії. Початковий етап усвідомлення соціальної інженерії полягає у зосередженні уваги на відповідних сигналах та стимулах. Ефективна обробка інформації включає уважність до потенційних тривожних сигналів, таких як підозрілі електронні листи, незнайомі запити або незвичайна поведінка. Співробітники, які краще усвідомлюють своє оточення та свідомо обробляють

вхідну інформацію, мають кращі можливості для виявлення спроб соціальної інженерії. Цей підфактор включає стиль прийняття рішень, який є кращим підходом співробітника до прийняття рішень, та те, як він впливає на їхню реакцію на загрози соціальної інженерії.

Перцептивні фактори впливають на здатність співробітників розпізнавати попереджувальні ознаки атак соціальної інженерії. Таким чином, здатність сприймати ці сигнали підвищує обізнаність та час реакції, роблячи співробітників менш вразливими до обману. Сприйняття тісно пов'язане з обізнаністю щодо безпеки. Без базових знань (обізнаності щодо безпеки) співробітники можуть навіть не знати, на які сигнали звертати увагу, а перцептивні здібності можуть бути не повністю реалізовані. Це також пов'язано з поведінковими факторами, які свідчать про те, що постійні звички ретельного перегляду комунікацій можуть з часом загострити перцептивну обізнаність.

Перцептивний фактор та обізнаність щодо безпеки вважаються найважливішими факторами обізнаності працівників щодо тактик соціальної інженерії.

Коли працівники мають чітке сприйняття загроз соціальної інженерії та підвищену обізнаність щодо безпеки, вони більш схильні до поведінки, усвідомленої безпекою, повідомляють про підозрілу діяльність та активно беруть участь у захисті конфіденційної інформації, тим самим знижуючи ризик успішних атак соціальної інженерії. Пильна увага приділяється соціальним та психологічним факторам, оскільки ці фактори допомагають зрозуміти індивідуальну та міжособистісну динаміку, яка впливає на сприйнятливість до атак соціальної інженерії. Аналогічно, культурні та поведінкові фактори привертають майже однакову увагу через їхню роль у підвищенні обізнаності працівників щодо соціальної інженерії. Організаційні та змістовні фактори були визначені як два найменш впливові фактори у підвищенні обізнаності працівників щодо соціальної інженерії, що може бути пов'язано з їхньою природою як нелюдських факторів, які більше зосереджуються на структурних

аспектах заходів безпеки та інформаційного контенту, ніж на індивідуальній чи міжособистісній динаміці.

Кількісна методологія широко застосовується у вивченні обізнаності працівників щодо соціальної інженерії завдяки її здатності надавати об'єктивні та вимірювані дані. Зокрема, опитування є найпопулярнішим засобом, що використовується для виявлення факторів у цій галузі досліджень. Опитування пропонують кілька переваг, включаючи ефективний збір даних з великої вибірки, можливість вимірювати різні аспекти обізнаності та структурований підхід, який дозволяє проводити систематичний аналіз та вивчення взаємозв'язків між різними факторами. Крім того, опитування дозволяють дослідникам зібрати уявлення про сприйняття, ставлення та поведінку учасників, пов'язані із соціальною інженерією, надаючи цінну інформацію для розуміння та підвищення обізнаності в цій галузі.

Теорія мотивації захисту стверджує, що люди мотивовані захищати себе на основі свого сприйняття серйозності загрози, вразливості, ефективності реагування та самоефективності.

Соціально-когнітивна теорія підкреслює важливість спостережливого навчання, імітації та моделювання у зміні поведінки. Працівники можуть отримати знання про соціальну інженерію через навчальні програми та інформаційні кампанії. Такий вплив може призвести до більшої саморегуляції, оскільки працівники розмірковують та змінюють свою поведінку для покращення практики безпеки.

Крім того, концепція взаємного детермінізму підкреслює, як індивідуальні дії, такі як повідомлення про підозрілі електронні листи, взаємодіють з організаційною культурою, впливаючи на загальний рівень обізнаності.

Теорія культурних вимірів дає уявлення про те, як культурні цінності формують поведінку та сприйняття. У культурах, що надають пріоритет індивідуалізму, працівники можуть відчувати сильніше почуття особистої відповідальності за захист інформації, тим самим підвищуючи свою обізнаність про загрози соціальної інженерії. Крім того, в культурах, що характеризуються

високим рівнем уникнення невизначеності, працівники можуть проявляти більшу проактивність у вирішенні потенційних загроз, зумовлену дискомфортом від неоднозначності.

Виявлення та аналіз перцептивних, психологічних, соціальних, культурних, організаційних факторів дозволять організаціям розробляти більш цілеспрямовані навчальні програми, покращувати політику організаційної безпеки та створювати середовище, де працівники більш стійкі до маніпуляцій.

Практичні наслідки та практичні рекомендації, що узгоджуються з виявленими факторами, такі:

Організаціям слід зміцнювати перцептивну обізнаність шляхом впровадження індивідуальних програм навчання з безпеки, які покращують сприйняття ризиків та розпізнавання загроз співробітниками. Це навчання повинно використовувати реальні приклади атак соціальної інженерії та включати інтерактивні симуляції, де співробітники повинні виявляти потенційні загрози та реагувати на них.

Щоб сприяти культурі, яка насамперед ставиться до безпеки, керівництво повинно моделювати сильну поведінку в галузі безпеки та сприяти відкритому середовищу, де співробітники почуваються комфортно, повідомляючи про підозрілу діяльність, не боячись осуду. Використання впливу колег може бути ефективним шляхом створення груп безпеки під керівництвом колег та організації командних завдань для покращення колективної обізнаності про соціальну інженерію.

Крім того, організації повинні забезпечувати навчання психологічній стійкості, щоб допомогти співробітникам справлятися з емоційними маніпуляціями, стресом та тактикою терміновості, що часто використовується в таких атаках. Постійне навчання з питань безпеки має вирішальне значення, використовуючи різні формати, такі як відео, вікторини та інформаційні бюлетені, щоб співробітники були залучені та інформовані.

Також важливо інтегрувати культурну чутливість у навчальні програми. Ініціативи з питань безпеки повинні поважати та враховувати культурні цінності

співробітників, такі як повага до влади та довіра громади. У культурно колективістському середовищі заохочення лідерів або шанованих осіб до просування поведінки в галузі безпеки може підвищити сприйнятливість.

Розпізнавання загроз, пов'язаних з контентом, є важливим, оскільки співробітники повинні бути навчені виявляти маніпулятивний контент, такий як фальшива терміновість та тактика залякування, у фішингових електронних листах та повідомленнях, з чіткими інструкціями щодо перевірки комунікацій.

Перспективні напрями розвитку протидії атакам соціальної інженерії

***Багаторівневий аналіз*:** Багаторівневий аналіз може дослідити, як індивідуальні характеристики (наприклад, вік, освіта, попередній досвід атак), організаційні практики (наприклад, культура безпеки, участь керівництва) та контекстуальні впливи (наприклад, галузеві стандарти, специфічні для країни правила) взаємодіють, формуючи обізнаність щодо соціальної інженерії. Такий підхід забезпечить глибше розуміння того, як фактори на різних рівнях впливають на поведінку співробітників та їх вразливість до атак.

***Міжкультурні та галузеві порівняння*:** Порівняльні дослідження різних культурних контекстів, галузей та розмірів організацій можуть виявити специфічні культурні, поведінкові та організаційні відмінності в рівнях обізнаності. Наприклад, майбутні дослідження можуть з'ясувати, як культурні цінності впливають на сприйнятливість до певних векторів атак та як локалізовані навчальні програми можуть ефективно вирішувати ці вразливості.

***Вплив програм підвищення обізнаності на організаційну безпеку*:** Майбутні дослідження можуть оцінити прямий вплив програм підвищення обізнаності на результати організаційної безпеки. Це може включати вивчення того, чи призводить підвищена обізнаність до помітного зниження рівня успішності атак соціальної інженерії, або визначення типів навчання (наприклад, сценарного чи теоретичного), які є найефективнішими для зменшення кількості інцидентів.

Висновки до розділу 3

Доведено, що основними методами протидії атакам соціальної інженерії являються навчання персоналу, застосування політик кібербезпеки, політик комунікації, забезпечення обладнанням організацій, встановлення фільтрів спаму, антивірусів, брандмауєрів, встановлення зашифрованого зв'язку, керування паролями та даними, звітування про інциденти.

Запропонована модель, яка грунтується на розробці заходів протидії атакам соціального інжинірингу шляхом використання здатності NLP аналізувати та інтерпретувати моделі людської мови, тим самим підвищуючи можливості системи в області безпеки. Розроблення гібридних моделей виявлення атак, досягнення інтеграції поведінкового аналізу; підвищення рівня обізнаності користувачів; стандартизації підходів до класифікації атак соціальної інженерії полягають в основі моделі.

Визначено ключові фактори, що формують обізнаність співробітників про атаки соціальної інженерії, особливо перцептивні фактори та фактори обізнаності щодо безпеки для пом'якшення наслідків соціальної інженерії та запобігання порушенням безпеки.

Кількісні методи передбачають збір та аналіз числових даних для кількісної оцінки впливу різних факторів на обізнаність співробітників про атаки соціальної інженерії. За допомогою кількісних підходів можна отримати уявлення про поширеність та значення різних факторів у формуванні розуміння співробітниками та реакції на загрози соціальної інженерії.

Організації можуть вжити кількох заходів для підвищення цієї обізнаності:

- співробітникам слід пропонувати цільове навчання, яке пояснює та демонструє емоційні маніпуляції, що використовуються зловмисними особами або організаціями.

- активна участь керівництва є важливою для просування культури, що ставить безпеку на перше місце.

Доведено, що багаторівневий аналіз може виявити, як індивідуальні, організаційні та контекстуальні фактори, які формують вразливість.

ВИСНОВКИ

У кваліфікаційній роботі вирішено актуальне науково-прикладне завдання, сутність якого полягає в обґрунтуванні оцінювання ризику соціоінженерних атак у корпоративному середовищі із використанням багатофакторного підходу.

Розв’язавши поставлені дослідницькі завдання, в результаті наукового пошуку вдалося дійти таких висновків:

1. Виявлено основні типи соціоінженерних атак, зокрема спір-фішинг (Spear Phishing), вейлінг (Whaling), вішінг (Vishing), смішінг (Smishing), компрометація ділової електронної пошти (BEC), клонування (Clone Phishing), фішинг у соціальних мережах, розподілена спам-відволікаюча атака (DSD), waterhole-атака, тейлгейтинг (Tailgating) та дипфейк, з детальним описом їх особливостей. Ці атаки базуються на використанні психологічних закономірностей, що впливають на свідомість, специфіки аудиторії чи середовища, низького рівня знань про інформаційну безпеку та індивідуальних слабкостей, таких як базові потреби, цінності й поведінкові шаблони.

2. Проаналізовано ключові методи впливу в кібератаках: вплив групи, інформаційний та нормативний вплив, теорія соціального обміну (норма взаємності), моральний вплив та соціальна відповідальність, саморозкриття для встановлення довіри. Особливу увагу приділено типам переконання, технікам маніпуляції ставленнями й поведінкою жертв, а також піддоменам довіри, обману, мови та міркування, які активно застосовуються зловмисниками.

3. Систематизовано фактори, що визначають ризики соціоінженерних атак: організаційні (лідерство, превентивні, детективні та процедурні заходи, програми SETA), соціальні (суб'єктивні норми, афективна відданість, зобов'язання щодо роботи, вплив ЗМІ, нормативні аспекти), культурні (норми та цінності безпеки), психологічні (відкритість, сумлінність, екстраверсія, доброзичливість, невротизм, довіра, реактність), поведінкові (реакції в конкретних ситуаціях), усвідомлення безпеки (обізнаність, практики, політики) та контенту (методи переконання в листах, постах, дзвінках). Ці фактори

слугують основою для моніторингу та формують стратегію захисту, з використанням кількісних методів для аналізу їх поширення та впливу на обізнаність персоналу.

4. Для мінімізації ризиків рекомендовано технічні рішення (машинне навчання для виявлення аномалій, блокування підозрілих акаунтів, двофакторна автентифікація, фільтрація посилань), організаційні інструменти (політики кібербезпеки та комунікацій, обладнання, спам-фільтри, антивіруси, брендмауери, шифрування, керування паролями, звітність інцидентів) та освітні програми (навчання кібергігієні). Штучний інтелект і NLP застосовуються для аналізу поведінки, мови та передбачення атак, з акцентом на гібридні моделі та стандартизацію класифікації.

5. Визначено перцептивні фактори та аспекти обізнаності безпеки як ключові для зниження вразливості, з використанням багаторівневого аналізу (індивідуальний, організаційний, контекстуальний рівні). Пропонуються заходи: цільове навчання з прикладами маніпуляцій, залучення керівництва для культури безпеки, адаптація програм до культурних і галузевих особливостей, оцінка ефективності через кількісні показники (зниження атак).

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Aldawood, H. and Skinner, G. Contemporary Cyber Security Social Engineering Solutions, Measures, Policies, Tools and Applications: A Critical Appraisal. International Journal of Security (IJS), (2019). Vol.10 (1). URL: https://www.cscjournals.org/manuscript/Journals/IJS/Volume10/Issue1/IJS-151.pdf

2. Abass, I. A. M. Social Engineering Threat and Defense: A Literature Survey. Journal of Information Security. 2018.Vol. 9. (04), 257-264. URL: https://www.researchgate.net/publication/327698025_Social_Engineering_Threat_and_Defense_A_Literature_Survey

3. State of Cybersecurity 2021. URL: https://www.isaca.org/resources/infographics/state-of-cybersecurity-2021-infographic

4. Cost of a Data Breach Report 2025. URL:https://www.ibm.com/reports/data-breach

5. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. URL: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

6. Saudi Aramco Confirms Data Leak after Reported Cyber Ransom. URL: https://www.bloomberg.com/news/articles/2021-07-21/saudi-aramco-confirms-data-leak-after-reported-cyber-extortion

7. Marriott Discloses Data Breach Possibly Affecting over 5 Million Customers. URL:https://edition.cnn.com/2020/04/01/business/marriott-hack-trnd/index.html

8. Marriott Data Breach FAQ: How Did It Happen and What Was the Impact? URL:https://www.csoonline.com/article/3441220/marriott-data-breach-faq-how-did-it-happen-and-what-was-the-impact.html

9. Widespread Credential Phishing Campaign Abuses Open Redirector Links. URL:https://www.microsoft.com/security/blog/2021/08/26/widespread-credential-phishing-campaign-abuses-open-redirector-links/

10. Twitter Hack: Staff Tricked by Phone Spear-Phishing Scam. URL: https://www.bbc.com/news/technology-53607374

11. Shark Tank Host Barbara Corcoran Loses $380,000 in Email Scam. URL:https://www.forbes.com/sites/rachelsandler/2020/02/27/shark-tank-host-barbara-corcoran-loses-380000-in-email-scam/?sh=73b0935a511a

12. Toyota Parts Supplier Hit by $37 Million Email Scam. URL: https://www.forbes.com/sites/leemathews/2019/09/06/toyota-parts-supplier-hit-by-37-million-email-scam/?sh=733a2c6e5856

13. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. URL:https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

14. Google and Facebook Duped in Huge 'Scam'. URL: https://www.bbc.com/news/technology-39744007

15. Facebook and Google Were Conned out of $100m in Phishing Scheme. URL:https://www.theguardian.com/technology/2017/apr/28/facebook-google-conned-100m-phishing-scheme

16. Siddiqi MA, Pak W, Siddiqi MA. A Study on the Psychology of Social Engineering-Based Cyberattacks and Existing Countermeasures. *Applied Sciences*. 2022; 12(12):6042. URL: https://doi.org/10.3390/app12126042

17. Гарасимчук, О., Оліярник, Ю., Нестор, А., Наконечний, Т. Психологічні методи шахрайства в кіберпросторі та способи їм протидіяти. *Кібербезпека: освіта, наука, техніка.* 2025. 2(30), 511–529. URL: https://doi.org/10.28925/2663-4023.2025.30.990

18. Irani, D.; Balduzzi, M.; Balzarotti, D.; Kirda, E.; Pu, C. Reverse social engineering attacks in online social networks. In Proceedings of the Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA), Berlin, Germany, 7–8 July 2011.

19. Myers, D. Social Psychology, 10th ed.; Mc Graw Hill: New York, NY, USA, 2012; pp. 266–304.

20. Handbook of Social Resource Theory, 2012th ed.; Springer: New York, NY, USA, 2012; pp. 15–32.

21. Wang, Z.; Zhu, H.; Liu, P.; Sun, L. Social engineering in cybersecurity: A domain ontology and knowledge graph application examples. Cybersecurity 2021, 4, 31.

22. Collins, N.L.; Miller, L.C. Self-disclosure and liking: A meta-analytic review. Psychol. Bull. 1994, 116, 457–475.

23. Hacking Human Psychology: Understanding Social Engineering Hacks URL: https://www.relativity.com/blog/hacking-human-psychology-understanding-social-engineering/

24. Social Engineering Attack Escalation. URL: https://appriver.com/blog/201708social-engineering-attack-escalation

25. Norton, M.; Frost, J.; Ariely, D. Less is more: The lure of ambiguity, or why familiarity breeds contempt. J. Pers. Soc. Psychol. 2007, 92, 97–105.

26. Guadagno, R.E.; Cialdini, R.B. The Social Net: The Social Psychology of the Internet, 1st ed.; Oxford University Press: New York, NY, USA, 2009; pp. 91–113.

27. Robert, O.; Timothy, B. Distraction increases yielding to propaganda by inhibiting counterarguing. J. Pers. Soc. Psychol. 1970, 15, 344–358.

28. Siadati, H.; Nguyena, T.; Gupta, P.; Jakobsson, M.; Memon, N. Mind your SMSes: Mitigating social engineering in second factor authentication. *Comput. Secur*. 2017, 65, 14–28.

29. Priester, J.; Petty, R. Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. Pers. Soc. Psychol. Bull. 1995, 21, 637–654.

30. Mitnick, K.D.; Simon, W.L.; Wozniak, S. The Art of Deception: Controlling the Human Element of Security, 1st ed.; Wiley: Hoboken, NJ, USA, 2003; pp. 59–71.

31. Leary, M.R. Self-Presentation Impression Management And Interpersonal Behavior, 1st ed.; Routledge: London, UK, 1996; pp. 25–35.

32. Montañez, R.; Golob, E.; Xu, S. Human cognition through the lens of social engineering cyberattacks. Front. Psychol. 2020, 11, 1755–1773.

33. Metzger, M.J.; Hartsell, E.H.; Flanagin, A.J. Cognitive dissonance or credibility? A comparison of two theoretical explanations for selective exposure to partisan news. Commun. Res. 2020, 47, 3–28.

34. Social Engineering as a Threat to Societies: The Cambridge Analytica Case. URL: https://thestrategybridge.org/the-bridge/2018/7/18/social-engineering-as-a-threat-to-societies-the-cambridge-analytica-case

35. Lahcen, R.A.M.; Caulkins, B.; Mohapatra, R.; Kumar, M. Review and insight on the behavioral aspects of cybersecurity. Cybersecurity 2020, 3, 10.

36. You, L.; Lee, Y.H. The bystander effect in cyberbullying on social network sites: Anonymity, group size, and intervention intentions. Telemat. Inform. 2019, 45, 101284.

37. Sherchan, W.; Nepal, S.; Paris, C. A survey of trust in social networks. ACM Comput. Surv. 2013, 45, 1–33.

38. Molodetska, K.; Solonnikov, V.; Voitko, O.; Humeniuk, I.; Matsko, O.; Samchyshyn, O. Counteraction to information influence in social networking services by means of fuzzy logic system. Int. J. Electr. *Comput. Eng*. 2021, 11, 2490–2499.

39. Albladi, S.M.; Weir, G.R.S. Predicting individuals' vulnerability to social engineering in social networks. *Cybersecurity* 2020, *3*, 7.

40. Campbell, C.C. Solutions for counteracting human deception in social engineering attacks. Inf. Technol. People 2019, 32, 1130–1152.

41. Burgoon, J.K.; Buller, D.B. Interpersonal deception theory. Commun. Theory 1996, 6, 203–242.

42. Handoko, H.; Putri, D.A.W. Threat language: Cognitive exploitation in social engineering. In Proceedings of the International Conference on Social Sciences, Humanities, Economics and Law (ICSSHEL), Padang, Indonesia, 5–6 September 2018.

43. Dorr, B.J.; Bhatia, A.; Dalton, A.; Mather, B.; Hebenstreit, B.; Santhanam, S.; Cheng, Z.; Shaikh, S.; Zemel, A.; Strzalkowski, T. Detecting asks in SE attacks: Impact of linguistic and structural knowledge. arXiv 2020, arXiv:2002.10931.

44. Rodríguez-Priego, N.; Bavel, R.V.; Vila, J.; Briggs, P. Framing effects on online security behavior. *Front. Psychol.* 2020, *11*, 2833–2844.

45. Yasin, A.; Fatima, R.; Liu, L.; Wang, J.; Ali, R.; Wei, Z. Understanding and deciphering of social engineering attack scenarios. *Secur. Priv.* 2021, *4*, e161.

46. Handoko, H.; Putri, D.A.W.; Sastra, G.; Revita, I. The language of social engineering: From persuasion to deception. In Proceedings of the 2nd International Seminar on Linguistics (ISL), Padang, West Sumatra, Indonesia, 12–13 August 2015.

47. Comment of NLP and Social Engineering Hacking the Human Mind Article. URL: https://www.hellboundhackers.org/articles/read-article.php?article_id=8%78
48. Hughes-Larteya, K.; Li, M.; Botchey, F.E.; Qin, Z. Human factor, a critical weak point in the information security of an organization's Internet of things. *Heliyon* 2021, *7*, 6522–6535.
49. Parthy, P.P.; Rajendran, G. Identification and prevention of social engineering attacks on an enterprise. In Proceedings of the International Carnahan Conference on Security Technology (ICCST), Chennai, India, 1–3 October 2019.
50. Alkhaiwani, A.H.; Almalki, G.A. Saudi human awareness needs. A survey in how human causes errors and mistakes leads to leak confidential data with proposed solutions in Saudi Arabia. In Proceedings of the National Computing Colleges Conference (NCCC), Taif, Saudi Arabia, 27–28 March 2021.
51. Spear Phishing: Top Threats and Trends. URL: https://assets.barracuda.com/assets/docs/dms/spear-phishing_report_vol6.pdf
52. Sushruth, V.; Reddy, K.R.; Chandavarkar, B.R. Social engineering attacks during the COVID-19 pandemic. *SN Comput. Sci.* 2021, *2*, 78.
53. Washo, A.H. An interdisciplinary view of social engineering: A call to action for research. *Comput. Hum. Behav. Rep.* 2021, *4*, 100126.
54. Alsulami, M.H.; Alharbi, F.D.; Almutairi, H.M.; Almutairi, B.S.; Alotaibi, M.M.; Alanzi, M.E.; Alotaibi, K.G.; Alharthi, S.S. Measuring awareness of social engineering in the educational sector in the kingdom of Saudi Arabia. *Information* 2021, *12*, 208.
55. Aldawood, H.; Skinner, G. Reviewing cyber security social engineering training and awareness programs—Pitfalls and ongoing issues. Future Internet 2019, 11, 73.
56. Fan, W.; Lwakatare, K.; Rong, R. Social engineering: I-E based model of human weakness for attack and defense investigations. Int. J. Comput. Netw. Inf. Secur. 2017, 9, 1–11.
57. Bakhshi, T. Social engineering: Revisiting end-user awareness and susceptibility to classic attack vectors. In Proceedings of the 13th International Conference on Emerging Technologies (ICET), Islamabad, Pakistan, 27–28 December 2017.

58. Sillanpää, M.; Hautamäki, J. Social engineering intrusion: A case study. In Proceedings of the 11th International Conference on Advances in Information Technology (IAIT), Bangkok, Thailand, 1–3 July 2020.
59. What Is Social Engineering? A Definition + Techniques to Watch for. URL: https://us.norton.com/internetsecurity-emerging-threats-what-is-social-engineering.html (accessed on 16 September 2021).
60. What Is Social Engineering and How to Prevent It. URL: https://www.avast.com/c-social-engineering
61. Network Intrusion Detection Techniques Using Machine Learning. URL: https://www.researchgate.net/publication/349392282_Network_Intrusion_Detection_Techniques_using_Machine_Learning.
62. Here's How Cyber Threats Are Being Detected Using Deep Learning. URL: https://techhq.com/2021/09/heres-how-cyber-threats-are-being-detected-using-deep-learning.
63. Peng, T.; Harris, I.; Sawa, Y. Detecting phishing attacks using natural language processing and machine learning. In Proceedings of the IEEE 12th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 31 January–2 February 2018.
64. Tsinganos N.; Sakellariou G.; Fouliras P.; Mavridis I. Towards an automated recognition system for chat-based social engineering attacks in enterprise environments. In Proceedings of the 13th International Conference on Availability, Reliability and Security (ICARS), Hamburg, Germany, 27–30 August 2018.
65. Siddiqi M.; Pak W. An agile approach to identify single and hybrid normalization for enhancing machine learning based network intrusion detection. IEEE Access 2021, 9, 137494–137513.
66. Lansley M.; Polatidis N.; Kapetanakis S.; Amin K.; Samakovitis G.; Petridis M. Seen the villains: Detecting social engineering attacks using case-based reasoning and deep learning. In Proceedings of the Twenty-seventh International Conference on Case-Based Reasoning (ICCBR), Otzenhausen, Germany, 28–30 September 2019.
67. Ozcan, A.; Catal, C.; Donmez, E.; Senturk, B. A hybrid DNN–LSTM model for detecting phishing URLs. *Neural Comput. Appl.* 2021, *9*, 1–17.

68. Vinayakumar, R.; Alazab, M.; Jolfaei, A.; Soman, K.P.; Poornachandran, P. Ransomware Triage Using Deep Learning: Twitter as a Case Study. In Proceedings of the Cybersecurity and Cyberforensics Conference (CCC), Melbourne, Australia, 8–9 May 2019.
69. Vinayakumar R.; Soman K.P.; Poornachandran P.; Mohan,V.S.; Kumar A.D. ScaleNet: Scalable and Hybrid Framework for Cyber Threat Situational Awareness Based on DNS, URL, and Email Data Analysis. J. Cyber Secur. Mobil. 2019, 8, 189–240.
70. Ketha, S.; Srinivasan, S.; Ravi, V.; Soman, K.P. Deep Learning Approach for Intelligent Named Entity Recognition of Cyber Security. In Proceedings of the the 5th International Symposium on Signal Processing and Intelligent Recognition Systems (SIRS'19), Trivandrum, India, 18–21 December 2019.
71. Huang, Y.; Huang, L.; Zhu, Q. Reinforcement learning for feedback-enabled cyber resilience. *Annu. Rev. Control* 2022, *23*, 273–295.
72. Bland, J.A.; Petty, M.D.; Whitaker, T.S.; Maxwell, K.P.; Cantrell, W.A. Machine learning cyberattack and defense strategies. *Comput. Secur.* 2020, *92*, 101738.
73. Rawindaran, N.; Jayal, A.; Prakash, E.; Hewage, C. Cost benefits of using machine learning features in NIDS for cyber security in UK small medium enterprises (SME). Future Internet 2021, 13, 186.
74. Sallouma S.; Gaber, T.; Vadera S.; Shaalan K. Phishing email detection using natural language processing techniques: A literature survey. Procedia Comput. Sci. 2021, 189, 19–28.
75. Fang, Y.; Zhang, C.; Huang, C.; Liu, L.; Yang, Y. Phishing email detection using improved RCNN model with multilevel vectors and attention mechanism. IEEE Access 2019, 7, 56329–56340.
76. Gutierrez, C.N.; Kim, T.; Corte, R.D.; Avery, J.; Goldwasser, D.; Cinque, M.; Bagchi, S. Learning from the ones that got away: Detecting new forms of phishing attacks. IEEE Trans. Dependable Secure Comput. 2018, 15, 988–1001.
77. Repke, T.; Krestel, R. Bringing back structure to free text email conversations with recurrent neural networks. In Proceedings of the European Conference on Information Retrieval (ECIR), Grenoble, France, 25–29 March 2018.

78. Lan, Y. Chat-oriented social engineering attack detection using attention-based Bi-LSTM and CNN. In Proceedings of the 2nd International Conference on Computing and Data Science (CDS), Stanford, CA, USA, 28–30 January 2021.
79. Cano J. The human factor in information security: The weakest link or the most fatigued? Inf. Syst. Audit. Control Assoc. 2019, 5, 1–7.
80. Bian J.; Li L.; Sun J.; Deng J.; Li, Q.; Zhang X.; Yan L. The influence of self-relevance and cultural values on moral orientation. Front. Psychol. 2019, 10, 292.
81. Bada M.; Sasse A.M.; Nurse J. Cyber security awareness campaigns: Why do they fail to change behavior? In Proceedings of the International Conference on Cyber Security for Sustainable Society (ICCSSS), Coventry, UK, 26 February 2015.
82. Mortan E.A. Cyber Security and Supply Chain Management: Risk, Challenges, and Solutions, 1st ed.; World Scientific Publishing: Singapore, 2021; pp. 62–63.
83. Alkhalil Z.; Hewage C.; Nawaf, L.; Khan, I. Phishing attacks: A recent comprehensive study and a new anatomy. Front. Comput. Sci. 2021, 3, 563060.
84. Frauenstein E.D., Flowerday S.. Susceptibility to phishing on social network sites: a personality information processing model. Comput. Secur., 94 (2020), Article 101862
85. Shahbaznezhad H., Kolini F., Rashidirad M.. Employees' behavior in phishing attacks: what individual, organizational, and technological factors matter? J. Comput. Inf. Syst., 61 (6) (2021), pp. 539-550
86. Mohammed S., Apeh E. A model for social engineering awareness program for schools. 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), IEEE (2016)
87. Stewart J., Dawson M. How the modification of personality traits leave one vulnerable to manipulation in social engineering. Int. J. Inf. Priv. Secur. Integr., 3 (3) (2018), pp. 187-208.
88. Wahyudiwan D.D.H., Sucahyo Y.G., Gandhi A. Information security awareness level measurement for employee: case study at ministry of research, technology, and higher education. 2017 3rd International Conference on Science in Information Technology (Icsitech), IEEE (2017).

89. Capuano N., Fenza G., Loia V., and Stanzione C., "Explainable Artificial Intelligence in CyberSecurity: A Survey," IEEE Access, vol. 10, pp. 93575–93600, 2022, doi: 10.1109/ACCESS.2022.3204171

90. Esite J., Happy F., Asheshemi's O. Human Vulnerabilities in Cybersecurity: Analyzing Social Engineering Attacks and AI- Driven Machine Learning Countermeasures Human Vulnerabilities in Cybersecurity: Analyzing Social Engineering Attacks and AI- Driven Machine Learning Countermeasures. Journal of Science and Technology 30(1):72–84. DOI:10.20428/jst.v30i1.2597

91. Rahman, M. M., Kshetri, N., Sayeed, S. A. and Rana, M. M. (2024) AssessITS: Integrating Procedural Guidelines and Practical Evaluation Metrics for Organizational IT and Cybersecurity Risk Assessment. Journal of Information Security, 15, 564-588. doi: 10.4236/jis.2024.154032