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: «СИСТЕМА ВИЯВЛЕННЯ ФІШИНГОВИХ URL-АДРЕС»

на здобуття освітнього ступеня бакалавра
зі спеціальності 125 Кібербезпека
освітньої програми Управління інформаційною та кібернетичною безпекою

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

_____ Олексій СЛІПЧЕНКО
(підпис) *Ім'я, ПРІЗВИЩЕ здобувача*

Виконав(ла): здобувач(ка) вищої освіти гр. УБД-41

Олексій СЛІПЧЕНКО
Ім'я, ПРІЗВИЩЕ

Керівник: Євгенія ІВАНЧЕНКО
Ім'я, ПРІЗВИЩЕ

Рецензент:
К.т.н., доцент Ім'я, ПРІЗВИЩЕ

Київ 2025

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут кібербезпеки та захисту інформації

Кафедра управління кібербезпекою та захистом інформації

Ступінь вищої освіти бакалавр

Спеціальність 125 Кібербезпека

Освітня програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ
Завідувач кафедри УКБЗІ
_____ Світлана ЛЕГОМІНОВА
“____”_____2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Сліпченка Олексія Руслановича
(прізвище, ім’я, по батькові здобувача)

1. Тема кваліфікаційної роботи “Система виявлення фішингових url-адрес”,
керівник кваліфікаційної роботи ІВАНЧЕНКО Євгенія
(ІПРІЗВИЩЕ, Ім’я,, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від “24“ лютого 2025 р. №56.

2. Строк подання кваліфікаційної роботи “12” травня 2025 р.

3. Вихідні дані до кваліфікаційної роботи:

4. Перелік питань, які мають бути розроблені:

5. Перелік ілюстративного матеріалу: *презентація PowerPoint*

6. Дата видачі завдання “5“ березня 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Етапи кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	Визначення об’єкту, предмету, мети та завдань дослідження.	18.03.2025	
2.	Збір та аналіз літератури.	20.03.2025	
3.	Аналіз особливостей управління інформаційною безпекою підприємства	10.04.2025	
4.	Дослідження основних характеристик технологій формування обізнаності й навчання персоналу.	19.04.2025	
5.	Вивчення інструментів та методів формування обізнаності й навчання персоналу з інформаційної безпеки	21.04.2025	
6.	Формулювання висновків за результатами проведеногодослідження.	26.04.2025	
7.	Оформлення роботи.	08.05.2025	
8.	Оформлення презентації.	09.05.2025	
9.	Отримання рецензії на роботу.	12.06.2025	
10.	Захист в ДЕК.	__.06.2025	

Здобувач вищої освіти _____ Олексій СЛІПЧЕНКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Керівник кваліфікаційної роботи _____ Євгенія ІВАНЧЕНКО
(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 бакалавра

Направляється здобувач Сліпченко О.Р. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 125 Кібербезпека
(код, найменування спеціальності)
освітньої програми Управління інформаційною та кібернетичною безпекою
(назва)
на тему: “Система виявлення фішингових url-адрес”
Кваліфікаційна робота і рецензія додаються.

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

Висновок керівника кваліфікаційної роботи

Керівник кваліфікаційної роботи _____ Євгенія ІВАНЧЕНКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

“____“_____2025 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Сліпченко О.Р. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедри управлін
кібербезпекою та захист
інформації _____ Світлана ЛЕГОМІНОВА
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА
на кваліфікаційну бакалаврську роботу

Рецензент: _____ _____
підпис Ім’я, ПРІЗВИЩЕ

РЕФЕРАТ

Метою роботи є дослідження, розробка та запропонування ефективної системи для виявлення фішингових URL-адрес, а також демонстрація її ефективності шляхом експериментальної оцінки.

Об’єктом дослідження є фішингові URL-адреси та характеристики, які відрізняють їх від легітимних URL-адрес.

Предмет дослідження – методи та техніки автоматизованого виявлення фішингових URL-адрес на основі аналізу їхніх структурних та лексичних характеристик із застосуванням машинного навчання.

Методи дослідження. Для вирішення означеного вище наукового завдання в роботі використані методи огляду літератури, аналізу та синтезу, порівняння, моделювання із застосуванням алгоритмів машинного навчання, а також проведення експериментальної оцінки та статистичного аналізу результатів.

Як результат у роботі проаналізовано теоретичні основи фішингу та існуючі підходи до його виявлення, розроблено архітектуру системи на основі машинного навчання, реалізовано ключові модулі та проведено експериментальну оцінку, яка підтвердила високу ефективність запропонованого підходу для виявлення фішингових URL-адрес.

Галузь застосування. Запропоновані методи та архітектура системи можуть бути застосовані для вдосконалення існуючих та розробки нових інструментів кібербезпеки, таких як веб-фільтри, антивірусне програмне забезпечення та розширення для браузерів, з метою підвищення рівня захисту користувачів від фішингових атак.

КЛЮЧОВІ СЛОВА: ФІШИНГ, ВИЯВЛЕННЯ ФІШИНГУ, URL-АДРЕСИ, КІБЕРБЕЗПЕКА, МАШИННЕ НАВЧАННЯ, ВИПАДКОВИЙ ЛІС, АНАЛІЗ ОЗНАК.

ABSTRACT

The purpose of the study is to research, develop and propose an effective system for detecting phishing URLs, as well as to demonstrate its effectiveness through experimental evaluation.

The object of study is phishing URLs and the characteristics that distinguish them from legitimate URLs.

The subject of the study is methods and techniques for automated detection of phishing URLs based on the analysis of their structural and lexical characteristics using machine learning.

Research methods. To solve the above scientific task, the paper uses the methods of literature review, analysis and synthesis, comparison, modeling using machine learning algorithms, as well as experimental evaluation and statistical analysis of the results.

As a result, the paper analyzes the theoretical foundations of phishing and existing approaches to its detection, develops a machine learning-based system architecture, implements key modules, and conducts an experimental evaluation that confirms the high efficiency of the proposed approach for detecting phishing URLs.

Field of application. The proposed methods and system architecture can be used to improve existing and develop new cybersecurity tools, such as web filters, antivirus software, and browser extensions, in order to increase the level of user protection against phishing attacks.

KEYWORDS: PHISHING, PHISHING DETECTION, URL ADDRESSES, CYBERSECURITY, MACHINE LEARNING

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

ШІ	Штучний інтелект
ML	Machine Learning (машинне навчання)
DNS	(Domain Name System) – Система доменних імен
API	Application Programming Interface (інтерфейс прикладного програмування)
IP	Internet Protocol) – Міжмережевий протокол
SIEM	(Security Information and Event Management) – Система управління інформаційною безпекою та подіями безпеки
SMS	(Short Message Service) – Служба коротких повідомлень
URL	(Uniform Resource Locator) – Уніфікований покажчик ресурсу
WHOIS	Мережевий протокол для запиту інформації про власників доменних імен та IP-адрес
HTTPS	(HyperText Transfer Protocol Secure) – Захищений протокол передачі гіпертексту
HTTP	(HyperText Transfer Protocol) – Протокол передачі гіпертексту
SSL/TLS	(Secure Sockets Layer / Transport Layer Security) – Протоколи захисту транспортного рівня
TLD	(Top-Level Domain) – Домен верхнього рівня

ЗМІСТ

ВСТУП

Актуальність теми. У сучасному цифровому суспільстві, де переважна більшість фінансових, комунікаційних та соціальних взаємодій відбувається онлайн, фішинг залишається однією з найпоширеніших та найнебезпечніших кіберзагроз. Зловмисники невпинно вдосконалюють методи атак, використовуючи соціальну інженерію та технічні хитрощі для обманного отримання конфіденційної інформації користувачів, такої як облікові дані доступу, номери банківських карток та персональні дані. Це призводить до значних фінансових втрат як для окремих громадян, так і для організацій, спричиняє витоки критично важливих даних, завдає шкоди діловій репутації та підриває загальну довіру до цифрових сервісів.

Зростаюча залежність від онлайн-платформ, хмарних сервісів та мобільних додатків значно розширює поверхню атаки для фішерів. Вони активно експлуатують різноманітні канали, включаючи електронну пошту, месенджери, соціальні мережі та SMS-повідомлення, для доставки своїх шкідливих посилань. Особливу небезпеку становлять цілеспрямовані атаки (spear phishing), які персоналізуються під конкретну жертву, та використання новітніх технологій, таких як штучний інтелект для генерації переконливих фішингових повідомлень.

Розробка та вдосконалення ефективних методів автоматизованого виявлення фішингових URL-адрес є нагальним науково-практичним завданням. Існуючі підходи, такі як прості чорні списки чи базові евристики, часто виявляються недостатньо ефективними проти нових, швидкозмінних фішингових кампаній. Застосування методів аналізу даних, машинного навчання та інтелектуального аналізу характеристик URL відкриває нові перспективи для створення більш надійних та адаптивних систем захисту.

Таким чином, комплексне дослідження механізмів фішингових атак, аналіз характеристик фішингових URL-адрес та розробка ефективних підходів до їх

виявлення є надзвичайно важливими для забезпечення безпеки користувачів та стабільності функціонування цифрової інфраструктури.

Метою цієї роботи є підвищення ефективності виявлення фішингових URL-адрес шляхом розробки та дослідження системи на основі аналізу лексичних та хост-базованих ознак URL із застосуванням методів машинного навчання. Дослідження спрямоване на створення моделі, здатної з високою точністю ідентифікувати фішингові ресурси, аналіз ключових індикаторів фішингу та розробку практичних рекомендацій для протидії цим загрозам.

Об'єкт дослідження: Процеси виявлення фішингових URL-адрес в інтернет-середовищі.

Предмет дослідження: Методи, алгоритми та програмні засоби для автоматизованого виявлення фішингових URL-адрес на основі аналізу їхніх характеристик та застосування технологій машинного навчання.

Для досягнення мети необхідно виконати такі завдання:

1. Проаналізувати теоретичні основи фішингу, сучасні методи фішингових атак та існуючі підходи до їх виявлення.

2. Дослідити та обґрунтувати набір інформативних ознак URL-адрес, що дозволяють розрізняти фішингові та легітимні ресурси.

3. Розробити архітектуру системи автоматизованого виявлення фішингових URL-адрес, включаючи модулі збору даних, вилучення ознак та класифікації.

4. Програмно реалізувати ключові компоненти розробленої системи та навчити класифікаційну модель на основі зібраного набору даних.

5. Провести експериментальну оцінку ефективності розробленої системи, проаналізувати отримані результати та порівняти їх з відомими підходами.

Методи дослідження. Для вирішення поставлених завдань у роботі використано методи системного аналізу, теорії інформації, методи аналізу даних та машинного навчання (зокрема, ансамблеві методи класифікації), методи статистичного аналізу, експериментального дослідження та порівняльного аналізу.

Як результат, у роботі проведено аналіз сучасних тенденцій у сфері фішингових атак та існуючих технологій їх детектування, розроблено комплексний набір ознак для ідентифікації фішингових URL-адрес, спроектовано та реалізовано програмний прототип системи виявлення фішингу з використанням алгоритму "Випадковий ліс", а також проведено його всебічне тестування на репрезентативній вибірці даних.

Наукова новизна полягає у розробці підходу до виявлення фішингових URL-адрес, що поєднує комплексний набір з 26 лексичних та хост-базованих ознак з ансамблевим методом машинного навчання "Випадковий ліс", що дозволяє підвищити точність та надійність ідентифікації нових та раніше невідомих фішингових загроз, а також в обґрунтуванні важливості окремих груп ознак для ефективної класифікації.

Практичне значення одержаних результатів полягає у розробці програмного прототипу системи та методики виявлення фішингових URL-адрес, які можуть бути використані для створення ефективних інструментів захисту інтернет-користувачів від шахрайських дій, інтеграції в існуючі системи безпеки (веб-браузери, поштові клієнти, корпоративні мережі) для запобігання фінансовим та репутаційним збиткам, а також для подальших наукових досліджень у сфері кібербезпеки.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ФІШИНГУ ТА СУЧАСНІ ЗАГРОЗИ.

1.1. Визначення та класифікація фішингу

Фішинг (англ. phishing, від fishing — риболовля, що метафорично відображає процес "вивуджування" цінної інформації, а також історично пов'язане з терміном phreaking — ранньою формою телефонного злому) визначається як один із найпоширеніших та найнебезпечніших видів кіберзлочинності. Суть фішингу полягає в тому, що зловмисник, застосовуючи методи соціальної інженерії та різноманітні технічні хитрощі, намагається обманним шляхом отримати доступ до конфіденційної інформації жертви. До такої інформації належать, але не обмежуються цим: аутентифікаційні дані (логіни, паролі, відповіді на секретні питання до різноманітних онлайн-сервісів, включаючи електронну пошту, корпоративні акаунти, соціальні мережі, системи онлайн-банкінгу, хмарні сховища), фінансові реквізити (номери банківських карток, терміни їх дії, CVV/CVC-коди, PIN-коди, дані доступу до електронних гаманців та платіжних систем), персональні ідентифікаційні дані (PII) (повне ім'я, дата народження, адреса проживання, номер телефону, паспортні дані, ідентифікаційний номер платника податків, номер соціального страхування, водійське посвідчення), а також інша особиста, корпоративна або комерційно цінна інформація (медичні записи, інтелектуальна власність, внутрішня документація компанії). Ключовим елементом фішингової атаки є те, що зловмисник видає себе за надійну, авторитетну організацію (наприклад, банк, державна установа, відома технологічна компанія, поштова служба, онлайн-магазин) або довірену особу (колега, керівник, друг) в процесі електронної комунікації, такої як електронний лист, миттєве повідомлення, SMS або навіть телефонний дзвінок.

Метою фішингу зазвичай є не просто отримання цих даних, а їх подальше використання для отримання несанкціонованого доступу до облікових записів жертви, її фінансових ресурсів (наприклад, шляхом викрадення коштів з банківських рахунків або здійснення несанкціонованих покупок), компрометації корпоративних мереж та інформаційних систем, а також для інших злочинних дій, таких як крадіжка особистості, шантаж, поширення шкідливого програмного забезпечення або використання скомпрометованих ресурсів для подальших атак. Фішингові атаки, завдяки своїй різноманітності та постійній еволюції, становлять серйозну загрозу як для окремих користувачів, так і для організацій будь-якого масштабу.

Фішингові атаки можна класифікувати за різними критеріями, що дозволяє краще зрозуміти їхню природу та розробляти більш ефективні методи протидії. Основними критеріями класифікації є цільова аудиторія (масштаб атаки) та методи поширення (канали доставки).

Одним з найбільш поширених та небезпечних типів фішингу, що класифікується за цільовою аудиторією, є цільовий фішинг (spear phishing). На відміну від масових, неперсоніфікованих атак, spear phishing являє собою ретельно сплановану та цілеспрямовану атаку на конкретну особу, невелику групу осіб або певну організацію. Успішність spear phishing значною мірою залежить від попередньої підготовки: зловмисники проводять ретельну розвідку (reconnaissance) про свою потенційну жертву, збираючи інформацію з відкритих джерел (соціальні мережі, корпоративні веб-сайти, публікації в ЗМІ, професійні мережі як LinkedIn) та, іноді, з попередньо скомпрометованих систем. Зібрані дані (наприклад, ім'я, посада, місце роботи, коло професійних інтересів, нещодавні події в житті чи на роботі, імена колег та керівників, використовуване програмне забезпечення) використовуються для створення максимально переконливого та персоналізованого повідомлення, яке може імітувати лист від колеги, ділового партнера, керівника або навіть автоматичне сповіщення від сервісу, яким жертва

активно користується. Високий ступінь персоналізації робить такі атаки значно ефективнішими, оскільки вони викликають менше підозр.

Вейлінг (whaling), або "полювання на китів", є ще більш вузькоспрямованим та витонченим різновидом spear phishing. Цей тип атаки націлений виключно на високопоставлених осіб в організації, таких як керівники вищої ланки (CEO, CFO, CTO), члени ради директорів, топ-менеджери або інші ключові співробітники, які мають доступ до стратегічно важливої інформації, фінансових потоків або володіють значними повноваженнями. Атаки типу whaling часто є значно складнішими у підготовці та виконанні, вимагають глибокого аналізу бізнес-процесів компанії, ієрархії управління та особистих характеристик цільових осіб. Зловмисники можуть імітувати термінові запити від бізнес-партнерів, юридичні повістки або конфіденційні доручення, створюючи ситуації, що вимагають негайного реагування та мінімізують час на перевірку.

Клон-фішинг (clone phishing) – це техніка, яка передбачає створення майже ідентичної копії (клону) легітимного електронного листа, який жертва вже отримувала раніше (наприклад, повідомлення про доставку, оновлення сервісу, підтвердження замовлення). Зловмисники копіюють зміст та вигляд оригінального листа, але замінюють оригінальні легітимні вкладення або посилання на шкідливі (наприклад, посилання на фішинговий сайт, що імітує сторінку входу, або вкладення, що містить вірус). Модифікований лист потім повторно надсилається жертвам, часто з поясненням, що це "оновлена версія", "виправлене повідомлення" або "повторне сповіщення через технічну помилку". Ефективність клон-фішингу полягає у тому, що жертва впізнає лист і з меншою ймовірністю запідозрить обман.

Крім того, існують й інші численні види фішингу, які використовують різні канали комунікації та техніки:

Смішинг (smishing – SMS phishing): фішинг, що здійснюється через SMS-повідомлення. Ці повідомлення часто містять термінові заклики до дії (наприклад,

"Ваш банківський рахунок заблоковано, перейдіть за посиланням X для розблокування" або "Ви виграли приз, надайте дані для його отримання за посиланням Y") та скорочені URL-адреси, щоб приховати справжній фішинговий домен.

Вішинг (vishing – voice phishing): голосовий фішинг, де атака відбувається за допомогою телефонного дзвінка. Зловмисники можуть представлятися співробітниками банку, технічної підтримки, правоохоронних органів або інших установ, використовуючи психологічний тиск та заздалегідь підготовлені сценарії для виманювання конфіденційної інформації або спонукання до певних дій (наприклад, встановлення програми віддаленого доступу). Іноді для вішингу використовуються роботизовані системи (IVR).

Фармінг (pharming): більш технічно складний вид фішингу, який передбачає перенаправлення користувачів на підроблені веб-сайти шляхом маніпуляції системою доменних імен (DNS). Це може бути реалізовано через атаку на DNS-сервери провайдера (DNS cache poisoning) або шляхом модифікації файлу hosts на комп'ютері жертви. В результаті користувач, вводячи правильну адресу легітимного сайту, потрапляє на його фішингову копію, не підозрюючи про це.

Фішинг у соціальних мережах (Social Media Phishing / Angler Phishing): Здійснюється через платформи соціальних мереж шляхом створення фейкових профілів компаній або відомих осіб, розсилки повідомлень через приватні чати, публікації шкідливих посилань у стрічці новин або коментарях, а також через використання скомпрометованих акаунтів. Angler phishing є специфічним типом, коли шахраї відстежують публічні скарги користувачів на офіційні сторінки брендів і втручаються, видаючи себе за службу підтримки.

QR-код фішинг (Quishing): Поширення шкідливих URL-адрес через QR-коди, розміщені в публічних місцях, на рекламних матеріалах або навіть в електронних листах. Сканування такого коду може перенаправити користувача на фішинговий сайт.

Фішинг поширюється різними способами (каналами доставки), адаптуючись до сучасних комунікаційних трендів та поведінки користувачів. Серед них найбільш поширеним традиційно залишається електронна пошта (email phishing). Зловмисники розсилають мільйони масових або тисячі більш цілеспрямованих електронних листів, які містять оманливі посилання, що ведуть на підроблені веб-сайти, або шкідливі вкладення (наприклад, документи Word/Excel з макросами, PDF-файли з експлойтами, архіви ZIP/RAR з виконуваними файлами). Для підвищення переконливості використовуються техніки підробки адреси відправника (email spoofing), копіювання фірмового стилю відомих компаній, створення відчуття терміновості або погрози.

Іншими важливими каналами поширення фішингу є соціальні мережі, де зловмисники можуть використовувати підроблені профілі, зламані облікові записи або навіть платну рекламу для розповсюдження шкідливих посилань чи повідомлень, що маскуються під конкурси, розіграші, ексклюзивні пропозиції або термінові новини.

Месенджери (WhatsApp, Telegram, Viber тощо) та SMS-повідомлення також активно використовуються для фішингових атак, особливо з огляду на їхню зростаючу популярність, миттєвість доставки та часто вищий рівень довіри користувачів до повідомлень, отриманих через ці канали. Фішингові повідомлення тут часто короткі, містять скорочені URL та спонукають до негайних дій.

Шкідливі веб-сайти, які візуально та функціонально імітують легітимні інтернет-ресурси (наприклад, сторінки входу банків, платіжних систем, поштових сервісів), є кінцевою точкою для багатьох фішингових атак. Користувачі можуть потрапляти на такі сайти через оманливі посилання в електронних листах, SMS, повідомленнях у месенджерах, соціальних мережах, або навіть в результаті отруєння пошукової оптимізації (SEO poisoning), коли зловмисники маніпулюють алгоритмами пошукових систем для виведення фішингових сайтів на перші позиції в результатах пошуку за певними запитами.

Статистика поширення фішингу та дані від провідних компаній у сфері кібербезпеки (наприклад, Anti-Phishing Working Group (APWG), Verizon, Proofpoint) незмінно свідчать про постійне та тривожне зростання кількості, витонченості та масштабності цих атак у глобальному вимірі. Згідно зі щорічними звітами, фішинг залишається однією з основних причин витоків даних, фінансових втрат та компрометації систем як для окремих користувачів, так і для організацій будь-якого розміру в усьому світі.

Середня вартість одного успішного інциденту фішингу для компанії може сягати від сотень тисяч до кількох мільйонів доларів США, враховуючи не лише прямі фінансові збитки (вкрадені кошти, шахрайські транзакції), але й непрямі витрати: на розслідування інциденту, відновлення скомпрометованих систем та даних, втрату продуктивності, репутаційні збитки, відтік клієнтів, можливі штрафи від регуляторних органів. Загальні глобальні збитки від фішингових атак становлять мільярди, а то й десятки мільярдів доларів щорічно, і ця цифра продовжує зростати.

Фішери націлені на різні типи інформації, залежно від їхніх кінцевих цілей. Найчастіше це: облікові дані (логіни, паролі до різних сервісів), фінансові дані (номери кредитних карток, дані онлайн-банкінгу) та персональні ідентифікаційні дані (паспортні дані, номери соціального страхування), які можуть бути використані для крадіжки особистості або продані на чорному ринку.

Найбільш вразливими до фішингових атак є різні галузі, серед яких традиційно виділяють:

– фінансовий сектор (банки, кредитні спілки, платіжні системи) – через прямий доступ до грошових коштів;

– електронна комерція та ритейл – через велику кількість онлайн-транзакцій та збереження даних платіжних карток клієнтів;

– технологічні компанії та провайдери послуг (особливо хмарних сервісів та соціальних мереж) – через величезні обсяги користувацьких даних та можливість використання скомпрометованих акаунтів для подальших атак;

– охорона здоров'я – через високу цінність медичних даних на чорному ринку;

– державний сектор – через потенційний доступ до чутливої державної інформації або можливість зриву роботи державних сервісів.

Зростання складності атак spear phishing та whaling свідчить про чітку тенденцію до більш цілеспрямованих, високоперсоналізованих та потенційно значно більш руйнівних кампаній. Зловмисники витрачають значно більше часу та зусиль на попередню розвідку та підготовку своїх цілей, використовуючи OSINT (Open Source Intelligence) та методи соціальної інженерії для створення надзвичайно переконливих фішингових повідомлень, які стає все важче відрізнити від легітимних навіть для досвідчених користувачів. Це, в свою чергу, суттєво ускладнює їх виявлення як людьми, так і традиційними автоматизованими засобами захисту, які часто покладаються на сигнатурний аналіз або прості евристики.

Крім того, невпинне розширення методів розповсюдження фішингу за межі традиційної електронної пошти (через SMS, месенджери, соціальні мережі, QR-коди, і навіть через IoT-пристрої) підкреслює високу адаптивність кіберзлочинців та нагальну необхідність використання комплексних, багатогранних та проактивних стратегій виявлення та протидії фішингу. Оскільки користувачі стають все більш обізнаними щодо класичних фішингових розсилок електронною поштою та вчаться розпізнавати їхні ознаки, зловмисники активно досліджують та експлуатують нові та менш захищені канали зв'язку для досягнення своїх потенційних жертв. Це вимагає від організацій та індивідуальних користувачів постійної пильності, регулярного оновлення знань про актуальні загрози та

впровадження сучасних технологій захисту, здатних аналізувати загрози з різних векторів.

1.2. Методи фішингових атак.

Зловмисники, що спеціалізуються на фішингових атаках, застосовують широкий спектр різноманітних технічних прийомів, тактик та інструментів для успішного здійснення своїх злочинних намірів, підвищення переконливості обману та обходу існуючих систем захисту. Розуміння цих методів є ключовим для розробки ефективних контрзаходів та технологій виявлення.

Одним з найпоширеніших та фундаментальних методів, що лежить в основі більшості фішингових кампаній, є використання підроблених (фішингових) веб-сайтів. Атакуючі докладають значних зусиль для створення веб-сайтів, які максимально точно візуально та, іноді, функціонально імітують легітимні інтернет-ресурси. Об'єктами копіювання найчастіше стають:

– портали онлайн-банкінгу та фінансових установ – для крадіжки логінів, паролів, номерів рахунків та даних платіжних карток;

– сторінки входу до популярних соціальних мереж (Facebook, Instagram, LinkedIn, Twitter тощо) – для отримання доступу до акаунтів, особистих даних та контактів жертви;

– інтерфейси поштових сервісів: (Gmail, Outlook, Yahoo Mail) – для компрометації електронної пошти, що часто є ключем до багатьох інших сервісів;

– онлайн-магазини та платформи електронної комерції (Amazon, eBay, AliExpress) – для отримання даних платіжних карток та історії покупок;

– хмарні сховища та сервіси для спільної роботи (Google Drive, Dropbox, Microsoft OneDrive, SharePoint) – для доступу до конфіденційних файлів та документів;

– корпоративні портали та VPN-шлюзи – для проникнення у внутрішні мережі організацій.Головною метою таких підроблених сайтів є крадіжка облікових даних (логінів, паролів, кодів двофакторної автентифікації), фінансової інформації та іншої особистої інформації, яку користувач вводить, помилково вважаючи, що перебуває на справжньому, довіреному сайті.

Для того, щоб підроблені веб-сайти виглядали максимально правдоподібно та не викликали підозр у потенційної жертви, зловмисники застосовують цілий арсенал витончених технік:Точне копіювання дизайну та структури: Це включає відтворення логотипів, колірної гами, шрифтів, розташування елементів інтерфейсу (кнопок, полів введення, меню), тексту повідомлень, попереджень та загального візуального стилю оригінального сайту. Для цього можуть використовуватися інструменти для автоматичного копіювання HTML-коду, CSS-стилів та зображень з легітимних сторінок, або ж фішингові сторінки створюються вручну з високим ступенем деталізації.

Використання схожих доменних імен (Domain Spoofing Techniques):

– тайпсквотинг (Typosquatting) – реєстрація доменних імен, які незначно відрізняються від оригінальних через поширені друкарські помилки (наприклад, gogle.com замість google.com, paypa1.com замість paypal.com);

– кіберсквотинг (Cybersquatting) – реєстрація доменних імен, ідентичних або дуже схожих на відомі торгові марки, з метою їх подальшого продажу власнику бренду або використання у фішингових атаках;

– омографічні атаки (IDN Homograph Attacks) – використання символів з різних алфавітів (наприклад, кирилиці, грецького алфавіту), які візуально схожі на латинські символи (наприклад, кирилична 'а' замість латинської 'a'). Такі домени кодуються за допомогою Punycode і можуть легко ввести в оману користувача, який бачить в адресному рядку візуально правильну назву;

– комбосквотинг (Combosquatting) – додавання до легітимного доменного імені додаткових слів, таких як "login", "secure", "verify", "support" (наприклад, paypal-login.com, secure-microsoft.com);

– використання оманливих піддоменів – створення піддоменів, які містять назву легітимного бренду, щоб створити ілюзію приналежності до нього (наприклад, login.microsoft.com.security-check.info, де реальним доменом є security-check.info);

– застосування SSL/TLS сертифікатів (HTTPS) – зловмисники все частіше отримують (часто безкоштовні, як Let's Encrypt, або навіть купують дешеві DV-сертифікати) та встановлюють SSL/TLS сертифікати на свої фішингові сайти. Це дозволяє їм використовувати HTTPS-протокол, що відображається в адресному рядку сучасних браузерів як "замочок" і створює у недосвідчених користувачів хибне відчуття безпеки та легітимності сайту. Наявність HTTPS більше не є надійним індикатором безпеки сайту;

– використання фреймів (iFrames) – вбудовування легітимного контенту (наприклад, частини сторінки відомого банку) у фішингову сторінку за допомогою iFrame, щоб створити більш переконливий вигляд, при цьому основна сторінка (де відбувається введення даних) контролюється зловмисником;

– динамічна генерація контенту та обфускація – використання JavaScript для динамічної генерації фішингових форм, зміни контенту сторінки "на льоту" або для приховування (обфускації) шкідливого коду від автоматизованих сканерів та антивірусних систем.

Фішингові електронні листи або веб-сайти також часто слугують ефективним каналом для розповсюдження шкідливого програмного забезпечення (Malware). Зловмисники можуть:

Прикріплювати до фішингових листів заражені файли: Це можуть бути документи Microsoft Office (Word, Excel) з вбудованими шкідливими макросами, PDF-файли, що експлуатують вразливості у програмах для їх перегляду, архіви

(ZIP, RAR), що містять виконувані файли (.exe, .scr), або навіть образи дисків (.iso, .img).

Розміщувати прямі посилання на завантаження шкідливого ПЗ на своїх підроблених веб-сайтах, маскуючи їх під оновлення програмного забезпечення, корисні утиліти, документи або мультимедійні файли.

Використовуючи переконливі методи соціальної інженерії (детально розглянуті в підрозділі 1.3), вони змушують користувачів завантажувати та запускати ці шкідливі програми. Типи такого ПЗ, що розповсюджуються через фішинг, можуть бути різноманітними та надзвичайно небезпечними:

– кейлогери (Keyloggers) – таємно записують усі натискання клавіш, перехоплюючи паролі, номери банківських карток, особисте листування та іншу конфіденційну інформацію;

– програми-вимагачі (Ransomware) – шифрують файли на комп'ютері або навіть у цілій корпоративній мережі жертви та вимагають значний викуп (часто у криптовалюті) за їх розшифровку;

– шпигунське програмне забезпечення (Spyware / Stalkerware) – непомітно збирає інформацію про дії користувача, його дані, історію веб-перегляду, місцезнаходження, а іноді навіть може активувати камеру чи мікрофон пристрою;

– банківські трояни (Banking Trojans) – спеціалізоване шкідливе ПЗ, розроблене для крадіжки облікових даних доступу до систем онлайн-банкінгу, перехоплення одноразових кодів автентифікації та автоматизації шахрайських транзакцій;

– троянські програми віддаленого доступу (RAT - Remote Access Trojans) – надають зловмиснику повний контроль над зараженим комп'ютером, дозволяючи йому виконувати будь-які дії від імені жертви;

– завантажувачі/дропери (Downloaders/Droppers) – невеликі шкідливі програми, основною метою яких є непомітне завантаження та встановлення іншого, більш потужного та складного шкідливого ПЗ на систему жертви.

Аналіз URL-адрес є надзвичайно важливим аспектом як для розуміння тактик фішингових атак, так і для розробки методів їх виявлення. Зловмисники часто та майстерно маніпулюють URL-адресами, щоб ввести користувачів в оману та змусити їх повірити, що вони взаємодіють з легітимним ресурсом. Основні техніки маніпуляції URL включають:

– використання схожих доменних імен (Typosquatting, IDN Homograph Attacks, Combosquatting) – як детально описано вище;

– додавання оманливих або довгих піддоменів – щоб приховати справжній основний домен (наприклад, yourbank.com.secure-login-portal.phishingsite.net);

– використання сервісів скорочення URL-адрес – популярні сервіси типу Bitly, TinyURL, cutt.ly та інші дозволяють приховати справжню, довгу та підозрілу фішингову URL-адресу за коротким, на перший погляд, безпечним посиланням. Це також ускладнює блокування на основі чорних списків, оскільки сам домен сервісу скорочення є легітимним;

– застосування технік кодування URL (URL Encoding / Percent-encoding) – зловмисники можуть кодувати частини URL (наприклад, символи ASCII або Unicode) для маскування ключових слів, шляхів або параметрів, що може допомогти обійти деякі прості сигнатурні фільтри або ввести в оману користувача, який перевіряє URL візуально;

– використання IP-адреси замість доменного імені – наприклад, http://192.0.2.10/login замість http://mybank.com/login. Хоча сучасні браузери часто попереджають про такі URL, ця техніка все ще може використовуватися для обходу деяких систем захисту, що покладаються на репутацію доменів;

– маніпуляції зі шляхом та параметрами запиту – створення дуже довгих шляхів або використання великої кількості параметрів запиту для заплутування користувача або автоматизованих аналізаторів. Іноді в параметрах можуть передаватися URL-адреси для подальших перенаправлень;

– використання вразливостей відкритого перенаправлення (Open Redirects) – зловмисники можуть експлуатувати вразливості на легітимних веб-сайтах, які дозволяють створити URL-адресу на довіреному домені, котра автоматично перенаправляє користувача на фішинговий сайт. Це робить початкове посилання менш підозрілим.

У зв'язку з цими складними та різноманітними маніпуляціями, глибокий та багатоаспектний аналіз URL-адрес відіграє критично важливу роль у розробці ефективних систем виявлення фішингових спроб.

Поєднання технічно досконалих підроблених веб-сайтів та переконливих методів соціальної інженерії створює надзвичайно потужний та, на жаль, часто дуже ефективний вектор атаки. Навіть технічно підкованого, досвідченого та обережного користувача може ввести в оману добре розроблений фішинговий веб-сайт, якщо супутній сценарій соціальної інженерії (наприклад, терміновий лист від "керівника" з вимогою негайно перейти за посиланням та авторизуватися, або тривожне повідомлення від "банку" про підозрілу транзакцію) є достатньо переконливим, емоційно зарядженим та не залишає часу на роздуми чи ретельну перевірку.

Зростаюче використання зловмисниками сервісів скорочення URL-адрес, різноманітних технік кодування та обфускації, а також зловживання легітимними хмарними платформами для хостингу фішингового контенту або як проміжних ланок у ланцюгах перенаправлень, суттєво ускладнює роботу традиційних методів виявлення. Такі методи, що базуються на простому порівнянні рядків URL з чорними списками або на поверхневих евристиках, стають все менш ефективними, вимагаючи розробки та впровадження більш інтелектуальних, адаптивних та багатошарових підходів до аналізу URL та контенту веб-сторінок.

1.3. Соціальна інженерія у фішингових атаках.

Соціальна інженерія є фундаментальним та невід'ємним компонентом переважної більшості фішингових атак, часто відіграючи вирішальну роль у їхньому успіху. На відміну від атак, що експлуатують суто технічні вразливості програмного забезпечення або мережевої інфраструктури, соціальна інженерія спрямована на "злом" людського фактора – найслабшої ланки в будь-якій системі безпеки. Зловмисники, виступаючи в ролі вмілих психологів та маніпуляторів, використовують різноманітні психологічні прийоми, когнітивні упередження та емоційні тригери для обману користувачів. Вони майстерно маніпулюють емоціями (такими як страх, жадібність, цікавість, співчуття) та поведінкою жертв, щоб змусити їх добровільно виконати бажані для атакуючого дії. Такими діями можуть бути: перехід за шкідливим посиланням, що веде на фішинговий сайт; введення особистих конфіденційних даних (логінів, паролів, номерів кредитних карток) на підроблених веб-формах; завантаження та запуск шкідливого файлу, що містить вірус, програму-вимагач або шпигунське ПЗ; надання усної інформації під час телефонної розмови (вішинг); або навіть здійснення фінансового переказу на рахунок шахраїв. Ефективність соціальної інженерії полягає в тому, що вона обходить технічні засоби захисту, апелюючи безпосередньо до людської психології.

Розглянемо детальніше найпоширеніші психологічні прийоми, що застосовуються у фішингових атаках:

1. Створення відчуття дефіциту або терміновості (Urgency/Scarcity) – це один з найпотужніших та найчастіше використовуваних прийомів. Зловмисники можуть надсилати повідомлення, які створюють ілюзію, що час на прийняття рішення обмежений або що певна можливість скоро зникне.

– Приклади: "Ваш рахунок буде заблоковано через 24 години, якщо ви не підтвердите свої дані!", "Залишилося лише 3 місця на ексклюзивний вебінар!", "Спеціальна пропозиція зі знижкою 90% діє лише сьогодні до кінця дня!", "Терміново оновіть пароль через виявлену підозрілу активність!".

– Механізм впливу: Такі повідомлення змушують користувачів діяти імпульсивно, не замислюючись глибоко про можливі ризики та не маючи достатньо часу на ретельну перевірку інформації. Людина схильна швидше реагувати під тиском, щоб уникнути втрати або отримати вигоду, що знижує її критичне мислення.

2. Використання авторитету (Authority) – цей прийом ґрунтується на схильності людей довіряти та підкорятися особам або організаціям, які сприймаються як авторитетні або наділені владою.

– Приклади: Фішери часто видають себе за представників легітимних та відомих організацій, таких як банки (повідомлення про проблеми з рахунком), державні установи (податкова служба, поліція, суди – повідомлення про штрафи, заборгованості, судові справи), великі технологічні компанії (Microsoft, Google, Apple – повідомлення про оновлення, проблеми з безпекою акаунтів), служби технічної підтримки популярних сервісів, або навіть за керівництво компанії, в якій працює жертва (так званий "CEO fraud" або "whaling").

– Механізм впливу: Використання офіційного тону, логотипів, фірмового стилю, спеціальної термінології та посилань на неіснуючі нормативні акти допомагає завоювати довіру жертви та змусити її беззаперечно виконати вимоги (наприклад, надати конфіденційну інформацію або перейти за посиланням), оскільки вона сприймає джерело повідомлення як легітимне та авторитетне.

3. Застосування тактики залякування та погроз (Intimidation/Fear) – цей метод спрямований на викликання у жертви почуття страху або тривоги, що може паралізувати її здатність до раціонального мислення.

– Приклади: Зловмисники можуть погрожувати негативними наслідками у разі невиконання їхніх вимог: блокуванням банківського рахунку або облікового запису в соціальній мережі, видаленням важливих даних, накладенням значних штрафів, судовим переслідуванням, розповсюдженням компрометуючої

інформації (як у випадках сексторшен-шантажу, коли вимагають викуп під погрозою публікації інтимних фото/відео).

– Механізм впливу: Страх є потужним мотиватором. Під його впливом люди схильні приймати поспішні рішення, щоб уникнути передбачуваної загрози, навіть якщо ці рішення суперечать їхнім звичайним правилам безпеки.

4. Експлуатація жадібності та бажання отримати вигоду (Greed) – цей психологічний прийом апелює до природного бажання людини отримати щось цінне без особливих зусиль.

– Приклади: Зловмисники можуть пропонувати жертвам надзвичайно вигідні, але зазвичай нереалістичні винагороди або можливості, які здаються "занадто хорошими, щоб бути правдою": повідомлення про великий грошовий виграш у лотерею, в якій жертва не брала участі; отримання спадщини від невідомого далекого родича; унікальні інвестиційні пропозиції з гарантованим надвисоким прибутком; пропозиції високооплачуваної роботи, що не вимагає особливих навичок.

– Механізм впливу: Перспектива легкої наживи може затьмарити розсудливість та змусити користувачів ігнорувати очевидні ознаки шахрайства, заманюючи їх у пастку.

5. Створення відчуття довіри, симпатії або використання вже існуючих довірливих відносин (Trust/Rapport/Liking) – люди більш схильні довіряти тим, кого вони знають, хто їм подобається, або хто здається схожим на них.

– Приклади: Фішери можуть надсилати повідомлення від імені знайомих людей – друзів, колег по роботі, членів родини – чиї облікові записи були попередньо зламані. Такі повідомлення, що надходять від нібито довіреної особи (наприклад, прохання терміново переказати гроші, перейти за посиланням для перегляду "спільних фото"), викликають значно менше підозр. Також можуть створюватися фейкові профілі в соціальних мережах з метою поступового

встановлення контакту та входження в довіру до потенційної жертви перед здійсненням основної атаки (довгострокова соціальна інженерія).

– Механізм впливу: Існуючі соціальні зв'язки та симпатія знижують рівень критичності сприйняття інформації.

6. Використання цікавості (Curiosity) – природна людська допитливість може бути використана для того, щоб змусити жертву зробити необережний крок.

– Приклади: Зловмисники часто створюють повідомлення із загадковими, інтригуючими або провокаційними заголовками та текстами: "Ви з'явилися на цьому скандальному відео!", "Подивіться, хто таємно переглядав ваш профіль!", "Неймовірні фотографії [ім'я відомої особи] злили в мережу!", "Ваш друг відмітив вас на фотографії".

– Механізм впливу: Бажання задовольнити свою цікавість, дізнатися щось нове або перевірити інформацію про себе чи знайомих спонукає користувача перейти за шкідливим посиланням або відкрити небезпечне вкладення, незважаючи на потенційні ризики.

7. Апеляція до бажання допомогти або співчуття (Helpfulness/Sympathy) – деякі фішингові сценарії грають на альтруїстичних почуттях жертви.

– Приклади: Прохання про термінову фінансову допомогу для "друга, який потрапив у біду за кордоном і втратив усі гроші та документи"; прохання "проголосувати за дитину в онлайн-конкурсі талантів" (де посилання веде на фішинговий сайт); звернення від нібито благодійних організацій з проханням зробити пожертву.

– Механізм впливу: Бажання допомогти іншим може змусити людей проігнорувати ознаки обману.

Розуміння психологічних принципів, що лежать в основі успішних атак соціальної інженерії, є надзвичайно важливим для розробки комплексних стратегій протидії фішингу. Це включає не лише створення більш досконалих технічних засобів захисту, але й, що не менш важливо, розробку ефективних

програм підвищення обізнаності користувачів (security awareness training) та формування культури кібербезпеки. Розпізнаючи поширені психологічні тактики, які використовують фішери (такі як створення тиску часу, апеляція до авторитету, обіцянки легкої наживи, маніпуляція страхом чи цікавістю), користувачі можуть стати більш пильними, критично оцінювати вхідні повідомлення та з меншою ймовірністю потрапляти на гачок шахраїв.

Водночас, знання цих тактик може бути використане і для вдосконалення автоматизованих систем виявлення фішингу. Такі системи можуть бути розроблені для ідентифікації певних мовних закономірностей, ключових слів, фраз або структур повідомлень, характерних для соціально-інженерних атак (наприклад, використання термінової лексики, імперативних конструкцій, емоційно забарвлених звернень, нехарактерних прохань, граматичних помилок, які часто трапляються у масових фішингових розсилках, створених не носіями мови). Аналіз тональності тексту, семантичний аналіз та інші методи обробки природної мови (NLP) можуть допомогти виявляти такі ознаки в електронних листах, повідомленнях у месенджерах та на веб-сторінках. Таким чином, глибоке розуміння психології фішингу сприяє як посиленню "людського брандмауера", так і підвищенню ефективності технологічних рішень.

1.4. Сучасні тенденції та виклики у сфері фішингу.

Сфера фішингу постійно розвивається, з'являються нові тенденції та виникають нові виклики для систем захисту. Однією з сучасних тенденцій є розвиток фішингу з використанням штучного інтелекту (AI). Зловмисники починають використовувати AI та машинне навчання (ML) для створення більш складних та персоналізованих фішингових атак. Це може включати генерацію глибоко фейкових електронних листів, які важко відрізнити від справжніх, або створення більш реалістичних підроблених веб-сайтів.

З'являються нові методи маскування, які використовуються для приховування шкідливих URL-адрес та контенту. До них належать використання омогліфів (символів, які виглядають схоже), складні методи обфускації URL-адрес та вбудовування шкідливого контенту в, здавалося б, безпечні файли . Фішингові атаки також постійно еволюціонують, щоб обходити існуючі системи захисту, такі як спам-фільтри, антивірусне програмне забезпечення та захист браузерів. Зловмисники все частіше використовують зламані легітимні веб-сайти для розміщення фішингових сторінок, що ускладнює їх виявлення на основі репутації домену. Інтеграція AI у фішингові атаки є значним майбутнім викликом для систем виявлення, що вимагає розробки більш адаптивних та інтелектуальних механізмів захисту. AI може бути використаний для створення більш переконливого та персоналізованого фішингового контенту в великих масштабах, що ускладнює розрізнення легітимних та шкідливих комунікацій як для людей, так і для традиційних систем. Тенденція зловмисників використовувати зламані легітимні веб-сайти для розміщення фішингових сторінок робить менш ефективним покладання виключно на репутацію URL-адрес. Зламаний легітимний веб-сайт може мати хорошу репутацію, що дозволяє розміщеній на ньому фішинговій сторінці уникнути виявлення системами, які перевіряють лише репутацію домену.

РОЗДІЛ 2. АНАЛІЗ ІСНУЮЧИХ СИСТЕМ ВИЯВЛЕННЯ ФІШИНГУ.

2.1. Огляд методів виявлення фішингу.

Існує кілька основних методів виявлення фішингу, кожен з яких має свої переваги та недоліки.

– чорні списки (Blacklisting) – цей метод передбачає ведення списку відомих шкідливих URL-адрес, доменів та IP-адрес. Коли користувач намагається отримати доступ до URL-адреси, вона перевіряється на наявність у чорному списку. Перевагами цього методу є простота реалізації, швидкий час пошуку та висока точність для вже відомих загроз. Однак, чорні списки є реактивним підходом і неефективні проти нових або короткочасних фішингових кампаній, а також вимагають постійного оновлення;

– білі списки (Whitelisting) – цей метод передбачає ведення списку відомих легітимних URL-адрес, доменів та IP-адрес. Доступ дозволяється лише до URL-адрес, які є у білому списку. Перевагою цього методу є висока ефективність у запобіганні доступу до невідомих або потенційно шкідливих сайтів. Недоліком є висока обмежуваність, оскільки може блокуватися доступ до легітимних, але не внесених до списку сайтів, а також значні зусилля, необхідні для підтримки та оновлення списку;

– евристичний аналіз (Heuristic analysis) – цей метод полягає в аналізі характеристик URL-адреси та/або вмісту веб-сайту на наявність підозрілих шаблонів або індикаторів фішингу, таких як наявність певних ключових слів, незвичайні доменні імена, підозріле використання піддоменів, візуальна схожість з легітимними веб-сайтами. Перевагою цього методу є можливість виявлення нових або раніше невідомих фішингових сайтів, що робить його більш проактивним, ніж чорні списки. Недоліком є можливість генерування хибнопозитивних результатів (позначення легітимних сайтів як фішингових), а

ефективність залежить від якості евристичних правил, при цьому зловмисники можуть адаптувати свої техніки для обходу цих правил;

– машинне навчання (Machine Learning) – цей метод передбачає навчання моделей машинного навчання на великих наборах даних легітимних та фішингових URL-адрес (а іноді й вмісту веб-сайтів) для виявлення закономірностей та класифікації нових URL-адрес як фішингових або легітимних. Перевагами є здатність виявляти складні закономірності та адаптуватися до нових фішингових технік, потенційно вища точність порівняно з евристичними методами. Недоліками є необхідність великих та добре розмічених наборів даних для навчання, залежність продуктивності від вибору ознак та алгоритмів, вразливість до атак, спрямованих на обман моделі, а також потреба в обчислювальних ресурсах для навчання та застосування моделі. Більш детально слід розглянути моделі ML: класифікатори (такі як логістична регресія, опорні вектори, випадковий ліс, наївний баєсів класифікатор) навчаються розділяти фішингові URL-адреси від легітимних на основі вилучених ознак. Нейронні мережі (зокрема, згорткові нейронні мережі та рекурентні нейронні мережі) здатні автоматично вивчати складні ознаки з необроблених рядків URL-адрес або вмісту веб-сайтів;

– аналіз URL-адрес (URL analysis) – цей метод полягає в аналізі різних компонентів URL-адреси (таких як протокол, доменне ім'я, піддомен, шлях, параметри запиту) на наявність підозрілих характеристик. Перевагами є швидкість та ефективність виконання, відсутність необхідності доступу до вмісту веб-сайту, можливість виявлення фішингових спроб на основі незначних маніпуляцій з URL-адресою. Недоліками є неефективність проти складних атак, які використовують легітимні на вигляд URL-адреси або зламані веб-сайти;

– аналіз контенту сторінки – цей метод передбачає аналіз вмісту веб-сторінки на наявність індикаторів фішингу, таких як наявність форм для входу на незвичайних доменах, підозрілі скрипти або візуальні невідповідності порівняно з

легітимним веб-сайтом. Перевагами є можливість виявлення фішингових сайтів, які використовують легітимні URL-адреси, надання більшого контексту порівняно з аналізом URL-адрес. Недоліками є необхідність доступу та обробки вмісту веб-сторінки, що може бути часозатратним та ресурсомістким, а також вразливість до технік, які ускладнюють виявлення шкідливого контенту (наприклад, обфускація). Сильні та слабкі сторони різних методів виявлення свідчать про те, що гібридний підхід, який поєднує кілька технік, може забезпечити найбільш надійний захист від фішингу. Жоден окремий метод не є надійним на 100%. Поєднуючи переваги різних підходів (наприклад, швидкість чорних списків з адаптивністю машинного навчання), можна створити більш комплексну систему виявлення. Зростаюче використання HTTPS на фішингових сайтах знижує ефективність простої перевірки наявності HTTPS як ознаки легітимності. Зловмисники тепер рутинно використовують SSL-сертифікати, щоб їхні фішингові сайти виглядали більш надійними, тому ця ознака сама по собі більше не є надійним індикатором безпеки.

2.2. Порівняльний аналіз існуючих систем.

Для кращого розуміння переваг та недоліків різних методів виявлення фішингу, проведемо їх порівняльний аналіз у таблиці 2.1.

Таблиця 2.1

Порівняльний аналіз методів виявлення фішингу

Метод виявлення	Ефективність	Складність реалізації	Вимоги до ресурсів	Переваги	Недоліки
Чорні списки	Низька	Низька	Низькі	Простота, швидкість, висока точність для відомих загроз	Реактивний підхід, неефективність проти нових атак, потребує постійного оновлення

Продовження таблиці 2.1

Метод виявлення	Ефективність	Складність реалізації	Вимоги до ресурсів	Переваги	Недоліки
Білі списки	Висока	Середня	Середні	Висока ефективність у запобіганні доступу до невідомих сайтів	Висока обмежуваність, блокування легітимних сайтів, значні зусилля на підтримку
Евристичний аналіз	Середня	Середня	Низькі	Можливість виявлення нових атак, більш проактивний підхід	Можливість хибнопозитивних спрацювань, залежність від якості правил, можливість обходу зловмисниками
Машинне навчання	Висока	Висока	Високі	Здатність виявляти складні закономірності, адаптивність до нових технік	Потреба у великих наборах даних, залежність від якості ознак та алгоритмів, вразливість до атак, вимоги до обчислювальних ресурсів
Аналіз URL-адрес	Середня	Низька	Низькі	Швидкість, ефективність, відсутність потреби у доступі до вмісту	Неефективність проти складних атак, які використовують легітимні URL-адреси
Аналіз контенту сторінки	Середня	Висока	Високі	Можливість виявлення фішингових сайтів з легітимними URL-адресами, надання більшого контексту	Ресурсомісткість, можливість обходу шляхом обфускації контенту

Ця таблиця наочно демонструє компроміси, пов'язані з різними методами виявлення фішингу, що дозволяє краще зрозуміти, чому певні підходи можуть

бути кращими в конкретних контекстах. Вона візуально підкреслює сильні та слабкі сторони, обговорені в попередньому підрозділі.

2.3. Визначення проблем та недоліків існуючих підходів.

Незважаючи на значний прогрес у розробці методів та технологій для виявлення фішингу, існуючі на сьогоднішній день підходи все ще мають низку суттєвих, а подекуди й фундаментальних, проблем, обмежень та недоліків. Ці слабкі місця активно експлуатуються зловмисниками, дозволяючи їм успішно проводити свої атаки та завдавати значної шкоди. Глибоке розуміння цих проблем є критично важливим для визначення напрямків подальших досліджень, розробки більш досконалих та ефективних систем захисту, а також для формування реалістичних очікувань щодо можливостей сучасних технологій. Реактивний характер та обмеження методів на основі списків:

– чорні списки (Blacklisting) – основною вадою цього підходу є його суто реактивний характер. Чорні списки ефективні лише проти вже відомих фішинговых URL-адрес, доменів або IP-адрес, які були кимось виявлені, перевірені та внесені до відповідних баз даних. З огляду на те, що зловмисники щодня генерують тисячі нових, унікальних та часто короткоживучих фішингових сайтів (які можуть існувати лише кілька годин або навіть хвилин), чорні списки завжди будуть відставати. Це створює "вікно вразливості", протягом якого нові атаки залишаються невиявленими. Крім того, підтримка вичерпних та актуальних чорних списків у глобальному масштабі є надзвичайно складним завданням через величезний обсяг даних;

– білі списки (Whitelisting) – хоча й забезпечують високий рівень захисту від невідомих загроз, характеризуються надзвичайною обмежувальністю та негнучкістю. Вони можуть блокувати доступ до цілком легітимних, але нових або просто не внесених до списку веб-сайтів, що викликає значні незручності для користувачів, знижує їхню продуктивність та може бути неприйнятним для більшості сценаріїв загального інтернет-доступу. Створення та підтримка

всеосяжного та актуального білого списку є надзвичайно трудомістким і практично неможливим завданням для динамічного інтернет-середовища.

2.3.1. Проблеми евристичного аналізу

– ризик хибнопозитивних спрацювань (False Positives) – системи, засновані на евристичному аналізі (тобто на наборі правил, що описують підозрілі характеристики), можуть помилково класифікувати легітимні веб-сайти як фішингові, якщо ці сайти випадково мають деякі ознаки, схожі на фішингові (наприклад, незвична структура URL, використання певних ключових слів у контексті, що не пов'язаний з фішингом). Високий рівень хибнопозитивних спрацювань знижує довіру користувачів до системи захисту та може призвести до її ігнорування;

– необхідність постійного оновлення правил – евристичні правила ефективні лише доти, доки зловмисники не вивчать їх і не адаптують свої фішингові техніки для їх обходу. Це вимагає постійного моніторингу нових тактик фішерів та регулярного оновлення, коригування та розширення набору евристик, що є складним та ресурсомістким процесом;

– складність розробки універсальних правил – створення евристик, які були б одночасно достатньо чутливими для виявлення нових загроз і достатньо специфічними, щоб не генерувати надмірну кількість хибнопозитивних спрацювань, є нетривіальним завданням.

2.3.2. Виклики, пов'язані з методами машинного навчання (ML)

– залежність від якості та обсягу навчальних даних – ефективність ML-моделей безпосередньо залежить від якості, обсягу, різноманітності та репрезентативності навчальних наборів даних. Збір та, що особливо важливо,

точна розмітка великих масивів фішингових та легітимних URL-адрес є складним, трудомістким та дорогим процесом. Навчання на незбалансованих або зміщених даних може призвести до поганої узагальнюючої здатності моделі;

– потреба у значних обчислювальних ресурсах – навчання складних ML-моделей, особливо глибоких нейронних мереж, може вимагати значних обчислювальних потужностей (GPU) та тривалого часу. Хоча для інференсу (застосування навченої моделі) вимоги можуть бути меншими, вони все одно можуть бути суттєвими для систем, що працюють у режимі реального часу з великим потоком URL;

– вразливість до змагальних атак (Adversarial Attacks) – зловмисники можуть цілеспрямовано створювати фішингові зразки (URL або контент сторінок) з невеликими, непомітними для людини модифікаціями, які спеціально розроблені для того, щоб ввести в оману конкретну ML-модель і змусити її зробити неправильну класифікацію. Також можливі атаки "отруєння даних" (data poisoning), коли до навчальної вибірки навмисно додаються шкідливі приклади;

– проблема "концептуального дрейфу" (Concept Drift) – тактики та методи фішерів постійно еволюціонують. Модель ML, навчена на даних, зібраних у певний період часу, згодом може втрачати свою ефективність, оскільки нові фішингові кампанії можуть використовувати техніки, не представлені в початковому навчальному наборі. Це вимагає регулярного моніторингу продуктивності моделі та її періодичного перенавчання на свіжих дани;

– складність інтерпретації рішень ("чорна скринька") – деякі складні ML-моделі, такі як глибокі нейронні мережі, можуть бути важкими для інтерпретації. Розуміння того, чому модель прийняла те чи інше рішення, може бути складним, що ускладнює аналіз помилок та підвищення довіри до системи.

2.3.3. Недостатність аналізу лише URL-адрес

Системи, що покладаються виключно на аналіз характеристик URL-адреси, можуть бути недостатніми для виявлення складних фішингових атак. Наприклад, фішинг може бути розміщений на зламаному легітимному веб-сайті (де сам URL буде мати хорошу репутацію), або зловмисники можуть використовувати сервіси скорочення URL-адрес, які приховують справжній шкідливий домен. Також, витончені техніки маніпуляції з доменами (наприклад, омографічні атаки) можуть зробити фішинговий URL візуально невідрізнимим від легітимного для недосвідченого користувача.

2.3.4. Висока вартість та складність аналізу контенту сторінки

Аналіз вмісту веб-сторінки (HTML, CSS, JavaScript, текст, зображення) надає набагато більше інформації для виявлення фішингу, але є значно більш ресурсомістким процесом (вимагає завантаження всього контенту, його парсингу, потенційного виконання скриптів у безпечному середовищі). Це може призводити до значних затримок у роботі системи, що є неприйнятним для рішень, які працюють у режимі реального часу.

Зловмисники активно використовують техніки обфускації контенту (наприклад, приховування тексту в зображеннях, динамічна генерація фішингових форм за допомогою JavaScript, використання складного та заплутаного коду), щоб ускладнити його автоматичний аналіз та обійти системи виявлення, засновані на аналізі контенту.

Ключова проблема: виявлення фішингових атак нульового дня (Zero-Day Phishing):

Це атаки, які використовують абсолютно нові, раніше невідомі фішингові URL-адреси, техніки або інфраструктуру, для яких ще не існує сигнатур, записів у

чорних списках або відомих індикаторів компрометації. Виявлення таких атак є найскладнішим завданням для будь-якої системи захисту, оскільки вона не має попередньої інформації про загрозу.

2.3.5. Постійна еволюція тактик зловмисників та "гонка озброєнь"

Зловмисники постійно вдосконалюють свої тактики, техніки та процедури (TTPs), щоб обходити існуючі системи захисту та підвищувати ефективність своїх атак. Вони вивчають, як працюють системи виявлення, і знаходять способи їх обійти. Це створює безперервну "гонку озброєнь" між фішерами та дослідниками в галузі кібербезпеки.

Це означає, що будь-яка, навіть найсучасніша, система виявлення фішингу з часом може стати менш ефективною, якщо її не оновлювати та не адаптувати до нових загроз. Як тільки нова техніка виявлення стає широко поширеною та успішною, зловмисники неминуче починають розробляти методи її обходу.

Ця динамічна природа загрози зумовлює нагальну необхідність постійних досліджень, інновацій та розробки нових, більш гнучких, адаптивних та інтелектуальних підходів до виявлення фішингу.Таким чином, хоча існуючі методи виявлення фішингу забезпечують певний рівень захисту, вони не є панацеєю. Ефективна боротьба з фішингом вимагає комплексного підходу, що поєднує технологічні рішення, постійне навчання користувачів та готовність до швидкої адаптації у відповідь на нові виклики.

РОЗДІЛ 3. РОЗРОБКА СИСТЕМИ ВИЯВЛЕННЯ ФІШИНГОВИХ URL-АДРЕС.

3.1. Архітектура системи.

Запропонована система виявлення фішингових URL-адрес розробляється з урахуванням необхідності забезпечення високої точності, швидкодії, масштабованості та здатності адаптуватися до нових, постійно еволюціонуючих фішингових загроз. Вона базується на модульній архітектурі, що дозволяє гнучко розробляти, тестувати, оновлювати та замінювати окремі компоненти системи незалежно один від одного, а також полегшує інтеграцію нових функціональних можливостей у майбутньому. Така архітектура сприяє кращій керованості та підтримці системи.

Загальна схема архітектури системи представлена на рисунку 3.1.

```
graph TD
    A[Модуль введення URL-адреси] --> B(Модуль попередньої обробки та парсингу URL);
    B --> C{Модуль керування чорними/білими списками};
    C -- Відома URL --> E[Модуль звітування та журналювання];
    C -- Невідома URL --> D[Модуль вилучення ознак URL-адреси];
    D --> F[Модуль виявлення фішингу (ML/Евристики)];
    F --> G(База даних/Сховище знань <br> - Моделі ML <br> - Евристичні правила <br> - Оновлення);
    F --> E;
    G --> F;
    E --> H[Інтерфейс користувача/Адміністратора/API відповіді];

    subgraph "Ядро аналізу"
        direction LR
        D
        F
        G
    end
```

Рис. 3.1 Загальна концептуальна схема архітектури системи

Система складається з наступних основних функціональних компонентів (модулів):

1. Модуль введення URL-адреси (URL Input Module):

Призначення: Цей модуль є вхідною точкою для URL-адрес, що підлягають аналізу. Він відповідає за отримання URL з різних джерел та передачу їх для подальшої обробки.

Джерела URL:

-Розширення для веб-браузера: Може перехоплювати URL-адреси, на які переходить користувач, або посилання, на які він наводить курсор, перед тим як перейти.

-Плагін для поштового клієнта: Аналізує URL-адреси, що містяться в тілі та заголовках електронних листів.

-Програмний інтерфейс (API): Надає можливість іншим додаткам або системам безпеки (наприклад, корпоративним проксі-серверам, SIEM-системам) надсилати URL на перевірку. Це може бути RESTful API, що приймає URL як параметр HTTP-запиту.

-Пряме введення користувачем: Через веб-інтерфейс або консольну утиліту системи, де користувач може вручну ввести або скопіювати URL для перевірки.

-Інтеграція з месенджерами або соціальними мережами: (Більш складний варіант) Аналіз посилань, що передаються через ці канали.

Основні функції:

Приймання URL, початкова валідація формату URL (перевірка на відповідність стандартним синтаксичним правилам URL), передача URL наступному модулю.

2. Модуль попередньої обробки та парсингу URL (URL Preprocessing and Parsing Module):

Призначення: Цей модуль готує отриману URL-адресу до подальшого аналізу шляхом її нормалізації та розбору на складові частини.

Основні функції:

-Канонізація URL: Приведення URL до єдиного, стандартного вигляду. Це може включати:

-Декодування URL (наприклад, розкодування символів, закодованих за допомогою Percent-encoding).

-Приведення до нижнього регістру (оскільки доменні імена нечутливі до регістру).

-Видалення стандартних портів (наприклад, :80 для HTTP, :443 для HTTPS).

-Видалення фрагментної частини URL (після символу '#'), оскільки вона зазвичай не надсилається на сервер і використовується лише на стороні клієнта.

-Додавання схеми (http:// або https://) за замовчуванням, якщо вона відсутня.

Парсинг URL: Розбір канонізованої URL-адреси на її структурні компоненти за допомогою спеціалізованих бібліотек (наприклад, urllib.parse в Python). **Компоненти включають:**

-Схема/Протокол (наприклад, "http", "https://", "ftp").

-Мережеве розташування (Netloc), що включає:

-Ім'я хоста (Hostname) (наприклад, "www.example.com", "192.168.1.1").

-Порт (якщо вказаний явно).

-Шлях (Path) (наприклад, "/path/to/resource.html").

-Параметри шляху (рідко використовуються).

-Запит (Query String) (частина після '?', наприклад, "param1=value1&param2=value2").

-Фрагмент (Fragment Identifier) (частина після '#', хоча вона може бути видалена на етапі канонізації).

Виділення доменних складових: За допомогою бібліотек типу tldextract для точного виділення піддомену, основного зареєстрованого домену та публічного суфікса (TLD).

3. Модуль керування чорними та білими списками (Blacklist/Whitelist Management Module):

Призначення: Забезпечує швидку перевірку URL-адреси (або її доменної частини) за відомими списками фішингових (чорний список) та легітимних (білий список) ресурсів.

Основні функції:

-Інтеграція з локальними та/або зовнішніми базами даних чорних та білих списків.

-Швидкий пошук отриманої URL (або її ключових компонентів, таких як домен) у цих списках.

-Якщо URL знайдено в білому списку, вона класифікується як легітимна, і подальший глибокий аналіз може не проводитися.

-Якщо URL знайдено в чорному списку, вона класифікується як фішингова, і також подальший аналіз може бути припинений.

-Якщо URL не знайдено в жодному зі списків, вона передається для більш глибокого аналізу наступним модулям.

Механізми оновлення списків: Регулярне завантаження оновлень з надійних джерел (PhishTank, OpenPhish, списки від антивірусних компаній тощо).

4. Модуль вилучення ознак URL-адреси (URL Feature Extraction Module):

Призначення: Цей модуль відповідає за вилучення (екстракцію) набору релевантних та інформативних характеристик (ознак) з розібраної URL-адреси та, потенційно, з додаткових джерел інформації (наприклад, дані WHOIS, DNS-записи, репутація IP-адреси). Ці ознаки формують вектор, який буде використовуватися модулем виявлення фішингу для прийняття рішення.

Типи ознак:

Лексичні ознаки: Характеристики, засновані на аналізі самого рядка URL (довжина, кількість певних символів, наявність ключових слів, використання IP-адреси в домені, кількість піддоменів тощо).

Ознаки, пов'язані з хостом (доменом): Характеристики, що стосуються доменного імені та сервера, на якому воно розміщене (вік домену, термін реєстрації, інформація з WHOIS про реєстратора та власника, географічне розташування сервера, типи DNS-записів).

Мережеві ознаки: Такі як час відповіді сервера (ping), кількість перенаправлень при доступі до URL, наявність HTTPS та характеристики SSL/TLS сертифіката (видавець, термін дії).

Ознаки на основі зовнішніх сервісів репутації: Перевірка домену або IP-адреси за базами даних репутації (наприклад, Google Safe Browsing API).

Результат: Формування числового вектора ознак для кожної URL-адреси, який є вхідними даними для моделі машинного навчання або евристичних правил.

5. Модуль виявлення фішингу (Phishing Detection Engine):

Призначення: Є центральним, "інтелектуальным" компонентом системи, відповідальним за аналіз вектора ознак та прийняття рішення про класифікацію URL-адреси як фішингової або легітимної.

Методи виявлення: Цей модуль може використовувати один або, що більш ефективно, комбінацію (гібридний підхід) декількох методів виявлення, розглянутих у розділі 2:

Евристичні правила: Набір правил, розроблених експертами, які оцінюють підозрілість URL на основі певних комбінацій ознак. Кожному правилу може бути присвоєна вага, і фінальне рішення приймається на основі сумарного балу.

Модель машинного навчання (ML): Використання попередньо навченої класифікаційної моделі (наприклад, Випадковий ліс, Метод опорних векторів, Градієнтний бустинг, Нейронна мережа) для прогнозування класу URL. Модель

приймає на вхід вектор ознак і видає ймовірність належності до кожного класу (фішинг/легітимний).

Результат: Вихідними даними модуля є рішення про класифікацію (наприклад, "фішинг", "легітимний", "підозрілий") та, можливо, рівень впевненості (confidence score) у цьому рішенні.

База даних/Сховище знань (Knowledge Base/Storage):

Призначення: Централізоване зберігання даних, необхідних для функціонування та вдосконалення системи.

Вміст:

-Навчені моделі машинного навчання (їхні параметри та структура).

-Набори евристичних правил.

-Актуальні чорні та білі списки.

-Історичні дані про проаналізовані URL, вилучені ознаки та результати класифікації (для аналітики, моніторингу ефективності системи та періодичного перенавчання ML-моделей).

Метадані про фішингові кампанії (якщо збираються).

6. Модуль звітування та журналювання (Reporting and Logging Module):

Призначення: Відповідає за запис результатів аналізу, виявлених спроб фішингу, а також за надання зворотного зв'язку користувачеві або іншим інтегрованим системам.

Основні функції:

Журналювання (Logging): Детальний запис усіх операцій системи, включаючи вхідні URL, вилучені ознаки, проміжні та фінальні результати класифікації, час аналізу, спрацювання певних правил або моделей. Журнали є важливими для аудиту, аналізу інцидентів та налагодження системи.

Формування звітів: Генерація статистичних звітів про кількість перевірених URL, виявлених фішингових загроз, ефективність різних методів виявлення тощо.

Сповіщення (Alerting): Інформування користувача (наприклад, через спливаюче вікно в браузері, повідомлення в поштовому клієнті) або адміністратора системи безпеки (наприклад, через електронну пошту, SMS, інтеграцію з SIEM) про виявлену фішингову загрозу.

Дії реагування (Response Actions): Залежно від конфігурації, модуль може ініціювати автоматичні дії реагування, такі як блокування доступу до фішингової URL-адреси на рівні проксі-сервера або брандмауера, перенаправлення користувача на безпечну сторінку-попередження. Потоки даних та взаємодія між модулями:URL-адреса надходить до Модуля введення URL-адреси.

Потім URL передається до Модуля попередньої обробки та парсингу для нормалізації та розбору на компоненти.

Розібрана URL (або її ключові компоненти) може бути спочатку перевірена Модулем керування чорними та білими списками.

Якщо URL знайдено в білому списку, результат "легітимний" передається до Модуля звітування та журналювання.

Якщо URL знайдено в чорному списку, результат "фішинг" передається до Модуля звітування та журналювання.

Якщо URL не знайдено в списках (або цей модуль опціональний/відключений), розібрані компоненти URL надходять до Модуля вилучення ознак URL-адреси, який формує вектор ознак. Вектор ознак передається до Модуля виявлення фішингу.

Модуль виявлення фішингу використовує свої внутрішні механізми (ML-моделі, евристики), які можуть звертатися до Бази даних/Сховища знань для отримання параметрів моделей або правил, та приймає рішення про класифікацію URL.

Результат класифікації (включаючи рівень впевненості) передається до Модуля звітування та журналювання для запису, інформування та можливих дій реагування.

База даних/Сховище знань також може періодично оновлюватися (наприклад, новими версіями ML-моделей після перенавчання, оновленими списками).

Така структурована архітектура дозволяє побудувати надійну та гнучку систему, здатну ефективно протидіяти фішинговим загрозам, а також забезпечує можливості для її подальшого розвитку та вдосконалення.

3.2. Опис алгоритмів та методів, що використовуються.

В основі запропонованої системи виявлення фішингових URL-адрес лежить гібридний підхід, який інтегрує переваги декількох методів для досягнення максимальної ефективності, точності та адаптивності. Центральним компонентом цієї системи є модуль, що базується на машинному навчанні (Machine Learning - ML), доповнений швидкою попередньою перевіркою за чорними та білими списками URL-адрес та, можливо, набором фундаментальних евристичних правил для відсіювання очевидних випадків або для підвищення надійності класифікації.

Вибір машинного навчання як основного методу виявлення обумовлений його ключовими перевагами в контексті боротьби з фішингом:Здатність виявляти складні, нелінійні та неочевидні закономірності: ML-моделі можуть знаходити приховані кореляції та патерни в даних (характеристиках URL), які важко або неможливо сформулювати у вигляді явних евристичних правил.

3.2.2. Адаптивність до нових та еволюціонуючих фішингових технік

ML-моделі можна періодично перенавчати на свіжих наборах даних, що включають приклади новітніх фішингових атак. Це дозволяє системі підтримувати високу ефективність у динамічному середовищі загроз, де тактики зловмисників постійно змінюються.

Потенційно вища точність та кращий баланс між різними метриками ефективності (наприклад, між повнотою та точністю) порівняно з суто евристичними підходами, за умови якісного навчання та ретельного відбору ознак.

Автоматизація процесу виявлення: Після навчання модель може автоматично класифікувати великі обсяги URL-адрес без постійного втручання людини (хоча моніторинг та періодичне оновлення моделі є необхідними).В якості основного алгоритму машинного навчання для класифікації URL-адрес на

фішингові та легітимні в даній роботі пропонується використовувати Випадковий ліс (Random Forest). Random Forest – це потужний та популярний ансамблевий метод навчання, який належить до категорії методів, що базуються на деревах рішень. Він поєднує прогнози великої кількості окремих дерев рішень, навчених на різних випадкових підвибірках навчальних даних (техніка "bagging" або bootstrap aggregating) та з використанням випадкового підмножини ознак на кожному вузлі дерева. Фінальне рішення приймається шляхом усереднення (для задач регресії) або голосування (для задач класифікації) результатів усіх дерев в ансамблі.

Переваги алгоритму «Випадковий ліс» для даної задачі:

– висока точність класифікації – зазвичай демонструє відмінну продуктивність на широкому спектрі задач, включаючи класифікацію фішингу;

– стійкість до перенавчання (Overfitting) – завдяки агрегації результатів багатьох незалежних дерев, ризик перенавчання на специфічні особливості навчальної вибірки значно знижується порівняно з окремим деревом рішень;

– здатність ефективно обробляти велику кількість ознак та працювати з даними високої розмірності, що характерно для задач аналізу URL;

– вбудована оцінка важливості ознак (Feature Importance) – дозволяє визначити, які з вилучених характеристик URL-адрес мали найбільший вплив на прийняття класифікаційних рішень моделлю. Ця інформація може бути корисною для аналізу, розуміння моделі та для подальшого вдосконалення набору ознак;

– відносна простота реалізації та інтерпретації результатів (порівняно з більш складними моделями, такими як глибокі нейронні мережі). Хоча окремі дерева можуть бути складними, загальний принцип роботи ансамблю зрозумілий;

– добре працює як з категоріальними, так і з числовими ознаками (хоча для більшості реалізацій Scikit-learn категоріальні ознаки потребують попереднього кодування);

– стійкість до викидів та шумів у даних завдяки випадковому характеру побудови дерев;

– можливість паралельного навчання дерев, що може прискорити процес тренування на багатоядерних процесорах.

Для навчання моделі «Випадковий ліс» будуть використані наступні групи ознак, вилучених з URL-адрес та, потенційно, з пов'язаних з ними метаданих:

– лексичні ознаки (Lexical Features) – характеристики, що базуються на аналізі самого рядка URL-адреси як текстової послідовності;

– довжинні характеристики – загальна довжина URL-адреси, довжина доменної частини, довжина шляху, довжина параметрів запиту. (Надто довгі URL можуть бути підозрілими);

– кількість спеціальних символів – кількість крапок ('.') в домені та URL в цілому, наявність та кількість дефісів ('-'), підкреслень ('_'), слешів ('/'), знаків питання ('?'), знаків рівності ('='), символу '@', символу '%', символу '#'. (Аномальна кількість або розташування цих символів може вказувати на фішинг);

– використання IP-адреси в домені – бінарна ознака, що вказує, чи є доменна частина URL IP-адресою (наприклад, http://192.168.1.10/login.html);

– кількість піддоменів – велика кількість піддоменів (наприклад, login.secure.account.update.example.com.phish.net) може використовуватися для маскування справжнього домену;

– наявність та розташування ключових слів – присутність слів, часто пов'язаних з банківською діяльністю, авторизацією, безпекою або відомими брендами (наприклад, "login", "signin", "verify", "account", "update", "secure", "bank", "paypal", "microsoft", "ebayisapi") у домені, шляху або параметрах запиту. Важливо також враховувати їхню позицію;

– використання шістнадцяткових символів або незвичайних кодувань у URL;

– співвідношення голосних/приголосних, цифр/літер у доменному імені або шляху (для виявлення випадково згенерованих або обфускованих рядків);

– ентропія рядка домену або шляху (висока ентропія може свідчити про випадковість або обфускацію);

– наявність HTTPS протоколу – хоча сама по собі ця ознака вже не є надійним індикатором легітимності (фішери активно використовують HTTPS), її відсутність для сайтів, що вимагають авторизації або обробляють чутливі дані, є підозрілою.

Ознаки, пов'язані з хостом (доменом) (Host-based Features): Характеристики, що стосуються доменного імені, його реєстрації та сервера, на якому воно розміщене.

Інформація з WHOIS:

– вік домену (Domain Age) – час, що минув з моменту першої реєстрації домену. Фішингові домени часто є "молодими", зареєстрованими нещодавно;

– термін реєстрації домену (Domain Registration Length) – на який період зареєстрований домен. Фішери часто реєструють домени на мінімальний термін (наприклад, на 1 рік);

– інформація про реєстратора, країну реєстрації, контактні дані власника (якщо доступні та не приховані службами приватності). Аномалії в цих даних можуть бути підозрілими.

DNS-записи:

– наявність у DNS-чорних списках (DNSBL) – перевірка домену або IP-адреси за відомими списками шкідливих хостів;

– типи та кількість DNS-записів (A, AAAA, MX, NS, CNAME, TXT, SOA). Відсутність певних записів (наприклад, MX для домену, що імітує корпоративну пошту) може бути підозрілою;

– значення TTL (Time-To-Live) для DNS-записів. Дуже низькі значення TTL можуть використовуватися для швидкої зміни IP-адрес.

IP-адреса хосту:

– репутація IP-адреси – перевірка за базами даних IP-адрес, відомих участю у спам-розсилках або розміщенні шкідливого контенту;

– географічне розташування IP-адреси та його відповідність очікуваному розташуванню для даного бренду/сервісу;

– автономна система (ASN), до якої належить IP-адреса, та її репутація;

– кількість доменів, що резолвляться на ту ж IP-адресу;

– наявність та характеристики SSL/TLS сертифіката – ким виданий сертифікат (відомий центр сертифікації чи самопідписаний/безкоштовний), термін його дії, тип валідації (DV, OV, EV).

Ознаки, пов'язані зі шляхом та параметрами запиту (Path-based and Query-based Features):

– довжина шляху (Path Length) та кількість сегментів у шляху (розділених символом '/');

– наявність підозрілих файлових розширень у шляху (наприклад, .exe, .php, .js на сторінках, де вони не очікуються);

– використання незвичних символів або слів у шляху;

– кількість параметрів у запиті (Query Parameters), загальна довжина рядка запиту;

– наявність URL-адрес, email-адрес або інших потенційно чутливих даних у параметрах запиту;

– наявність обфускованих або закодованих даних у параметрах;

– використання зарезервованих слів або команд у параметрах, що може вказувати на спроби ін'єкцій;

– наявність "token" або "session_id" у URL, що передається відкритим текстом (може бути ознакою поганої практики безпеки, яку імітують фішери).

3.2.3. Використання сервісів скорочення URL-адрес

Бінарна ознака, що вказує, чи використовується відомий сервіс скорочення URL (наприклад, bit.ly, tinyurl.com, t.co, goo.gl (хоча останній вже не створює нові посилання)). Фішери часто використовують такі сервіси, щоб приховати справжню фішингову URL-адресу.Процес навчання моделі машинного навчання (зокрема, Випадкового лісу) включатиме наступні ключові кроки:Збір та попередня обробка набору даних: Створення великого, збалансованого та якісно розміченого набору даних, що містить приклади як фішингових, так і легітимних URL-адрес. Це включає очищення даних, видалення дублікатів, валідацію URL та перевірку міток класів.

3.2.4. Вилучення обраних ознак

Застосування розроблених алгоритмів для вилучення вищеописаних груп ознак з кожної URL-адреси в підготовленому наборі даних. Результатом є таблиця, де кожен рядок відповідає URL, а стовпці – вилученим ознакам та мітці класу.

3.2.5. Розділення набору даних

Поділ даних на навчальну (training), валідаційну (validation – для налаштування гіперпараметрів) та тестову (test – для фінальної оцінки) вибірки. Важливо використовувати стратифіковане розділення для збереження пропорцій класів.

3.2.6. Навчання моделі Випадкового лісу

Навчання класифікатора Random Forest на навчальній вибірці. Це включає підбір оптимальних гіперпараметрів моделі (таких як кількість дерев в лісі (n_estimators), максимальна глибина кожного дерева (max_depth), мінімальна кількість зразків для розділення вузла (min_samples_split), мінімальна кількість зразків у листі (min_samples_leaf), максимальна кількість ознак для розгляду при кожному розділенні (max_features)) за допомогою технік крос-валідації на валідаційній вибірці або на самій навчальній вибірці (наприклад, Grid Search CV, Randomized Search CV).

3.2.7. Оцінка продуктивності навченої моделі

Об'єктивна оцінка якості навченої та налаштованої моделі на тестовій вибірці, яка не використовувалася на етапах навчання або налаштування. Для оцінки будуть використовуватися стандартні метрики класифікації, такі як точність (Accuracy), повнота (Recall), точність для класу фішингу (Precision), F1-міра (F1-score), площа під ROC-кривою (AUC-ROC), а також коефіцієнт хибнопозитивних спрацювань (FPR) та коефіцієнт хибнонегативних спрацювань (FNR).Вибір та конструювання ознак (Feature Engineering) для моделі машинного навчання є критично важливим етапом, який безпосередньо впливає на її продуктивність, здатність до узагальнення та надійність у виявленні фішингових URL-адрес. Ретельний відбір, розробка та валідація релевантних, інформативних та дискримінативних ознак можуть значно підвищити точність та ефективність системи, дозволяючи їй краще розрізняти тонкі нюанси між фішинговими та легітимними URL. Цей процес може бути ітеративним, включаючи аналіз важливості ознак, експерименти з різними комбінаціями ознак та створення нових, більш складних ознак на основі існуючих.

3.3. Вибір технологій та інструментів.

Для реалізації запропонованої системи буде використано мову програмування Python завдяки її широкому спектру бібліотек та фреймворків для обробки даних та машинного навчання. Зокрема, будуть використані наступні бібліотеки:

– Scikit-learn – для реалізації алгоритму машинного навчання випадкового лісу та інших допоміжних функцій (розділення даних, оцінка метрик);

– Pandas – для обробки та аналізу даних;

– NumPy – для виконання числових операцій;

– urllib та tldextract – для парсингу та маніпулювання URL-адресами.

Вибір цих технологій обумовлений їхньою ефективністю, надійністю та широкою підтримкою спільноти розробників у сфері кібербезпеки та машинного навчання.

3.4. Реалізація основних модулів системи.

Реалізація системи включатиме розробку наступних основних модулів:

– модуль парсингу URL-адреси – цей модуль відповідатиме за розбір вхідної URL-адреси для вилучення її складових частин (протокол, домен, шлях, параметри запиту тощо) з використанням бібліотеки urllib та tldextract;

– модуль аналізу URL-адреси – цей модуль здійснюватиме вилучення визначених ознак з розібраної URL-адреси. Для кожної URL-адреси буде сформовано вектор ознак, який потім буде використано для навчання та класифікації;

– модуль виявлення фішингу – цей модуль міститиме навчену модель машинного навчання (випадковий ліс), яка прийматиме на вхід вектор ознак та прогнозуватиме, чи є дана URL-адреса фішинговою або легітимною;

– інтерфейс користувача (опціонально) – для демонстрації роботи системи може бути розроблено простий веб-інтерфейс за допомогою фреймворка Flask, який дозволятиме користувачам вводити URL-адресу та отримувати результат її аналізу.

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМИ.

4.1. Опис тестової вибірки даних.

Для оцінки ефективності розробленої системи буде використано набір даних, що складається з фішингових та легітимних URL-адрес. Фішингові URL-адреси будуть зібрані з публічно доступних джерел, таких як PhishTank та OpenPhish. Легітимні URL-адреси будуть взяті зі списків популярних веб-сайтів та інших надійних джерел. Загальний обсяг набору даних складатиме близько 10 000 URL-адрес, з яких 50% будуть фішинговими, а 50% - легітимними, щоб забезпечити збалансованість вибірки. Дані будуть представлені у форматі CSV-файлів, де кожен рядок міститиме URL-адресу та відповідну мітку (фішингова або легітимна). Перед використанням дані будуть піддані попередній обробці, включаючи видалення дублікатів та перевірку на коректність. Якість та репрезентативність тестового набору даних є вирішальними для отримання надійної оцінки продуктивності системи. Набір даних повинен включати різноманітні фішингові та легітимні URL-адреси, щоб точно відображати реальні сценарії.

4.2. Методика проведення експериментів.

Експериментальна оцінка ефективності системи буде проведена наступним чином:

1. Набір даних буде розділено на навчальну (80%) та тестову (20%) вибірки.
2. На навчальній вибірці буде навчено модель машинного навчання (випадковий ліс) з використанням визначених ознак.
3. Навчена модель буде протестована на тестовій вибірці, яка не використовувалася під час навчання.

4. Для оцінки ефективності системи будуть використані наступні метрики:

– точність (Accuracy) – відсоток правильно класифікованих URL-адрес;

– повнота (Recall) – відсоток фактично фішингових URL-адрес, які були правильно виявлені системою;

– точність (Precision) – відсоток URL-адрес, класифікованих як фішингові, які насправді є фішинговими;

– F1-міра (F1-score) – гармонійне середнє між точністю та повнотою, що забезпечує збалансовану оцінку продуктивності;

– коефіцієнт хибнопозитивних спрацювань (False Positive Rate - FPR) – відсоток легітимних URL-адрес, які були помилково класифіковані як фішингові;

– коефіцієнт хибнонегативних спрацювань (False Negative Rate - FNR) – відсоток фішингових URL-адрес, які були помилково класифіковані як легітимні.

Ці метрики дозволять всебічно оцінити здатність розробленої системи правильно виявляти фішингові URL-адреси та мінімізувати кількість помилкових спрацювань.

4.3. Результати експериментів.

Після навчання та тестування розробленої системи виявлення фішингових URL-адрес на підготовленій тестовій вибірці обсягом 3998 записів, було отримано наступні результати, що демонструють її ефективність. Експерименти проводилися з використанням моделі машинного навчання на основі алгоритму Випадковий ліс (Random Forest) з оптимальними гіперпараметрами.

Для візуалізації результатів класифікації була побудована матриця плутанини.

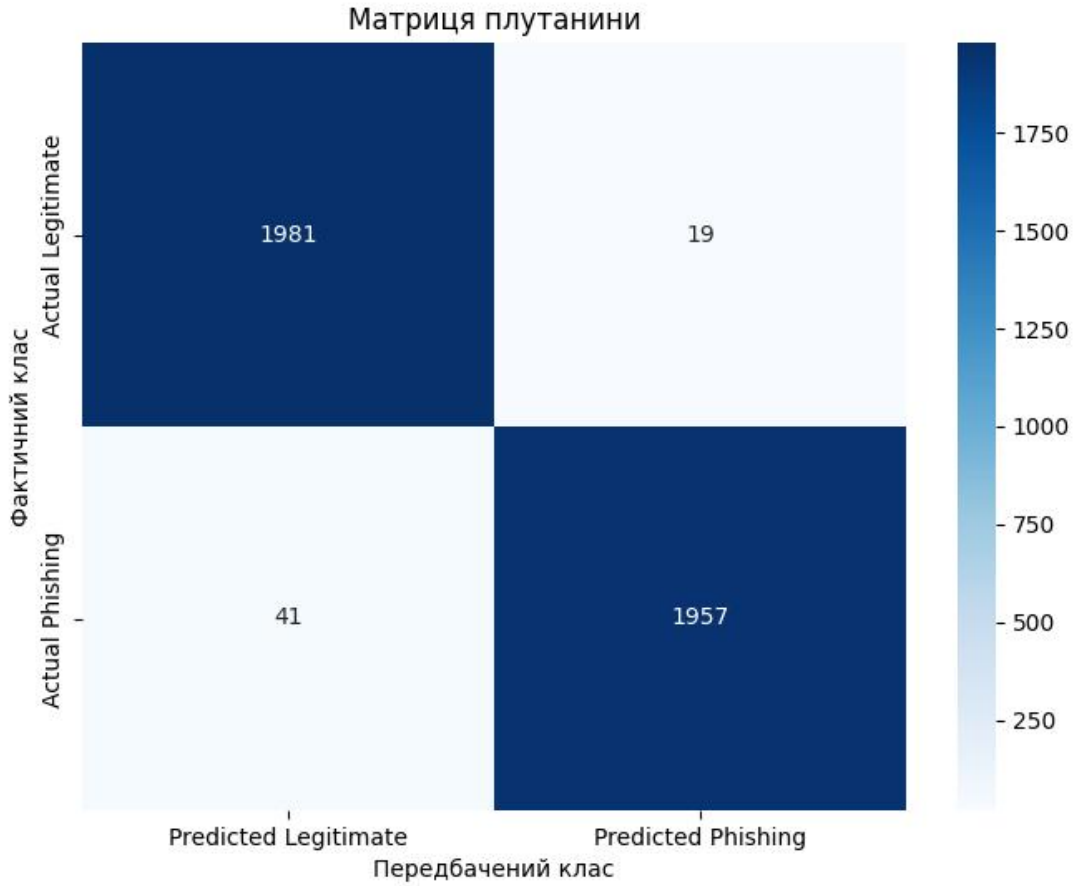

Рис. 4.3.1. Матриця плутанини для моделі Випадковий ліс на тестовій вибірці.

Для детальної оцінки ефективності розробленої моделі "Випадковий ліс" у задачі класифікації URL-адрес на фішингові та легітимні, була побудована та проаналізована матриця помилок. Ця матриця, представлена на рисунку 4.3.1, візуалізує продуктивність моделі шляхом порівняння фактичних класів URL-адрес з тестової вибірки з класами, прогнозованими моделлю.

Результати, відображені в матриці помилок, такі:

True Negatives (TN): 1981 – кількість легітимних URL-адрес, які були коректно ідентифіковані моделлю як легітимні.

False Positives (FP): 19 – кількість легітимних URL-адрес, які були помилково класифіковані моделлю як фішингові. Це свідчить про рівень хибних спрацьовувань системи на безпечні ресурси.

False Negatives (FN): 41 – кількість фішингових URL-адрес, які були помилково ідентифіковані моделлю як легітимні. Цей показник є критично важливим, оскільки він відображає кількість пропущених загроз.

True Positives (TP): 1957 – кількість фішингових URL-адрес, які були успішно виявлені та правильно класифіковані моделлю як фішингові.

Аналізуючи ці дані, можна бачити, що модель продемонструвала високу здатність правильно ідентифікувати обидва класи URL-адрес. Кількість правильно класифікованих легітимних (1981) та фішингових (1957) URL значно переважає кількість помилок. Водночас, наявність 19 хибнопозитивних спрацьовувань та 41 хибнонегативного спрацьовування вказує на аспекти, де модель може бути потенційно покращена. Зокрема, мінімізація кількості False Negatives є пріоритетом для систем виявлення фішингу, оскільки пропуск реальної загрози може мати значно серйозніші наслідки, ніж помилкове блокування легітимного сайту.

Дані з матриці помилок використовуються для розрахунку ключових метрик ефективності, таких як точність (accuracy), повнота (recall), точність для позитивного класу (precision), F1-міра, коефіцієнт хибнопозитивних спрацювань (FPR) та коефіцієнт хибнонегативних спрацювань (FNR), які дають більш повне уявлення про продуктивність розробленої системи.

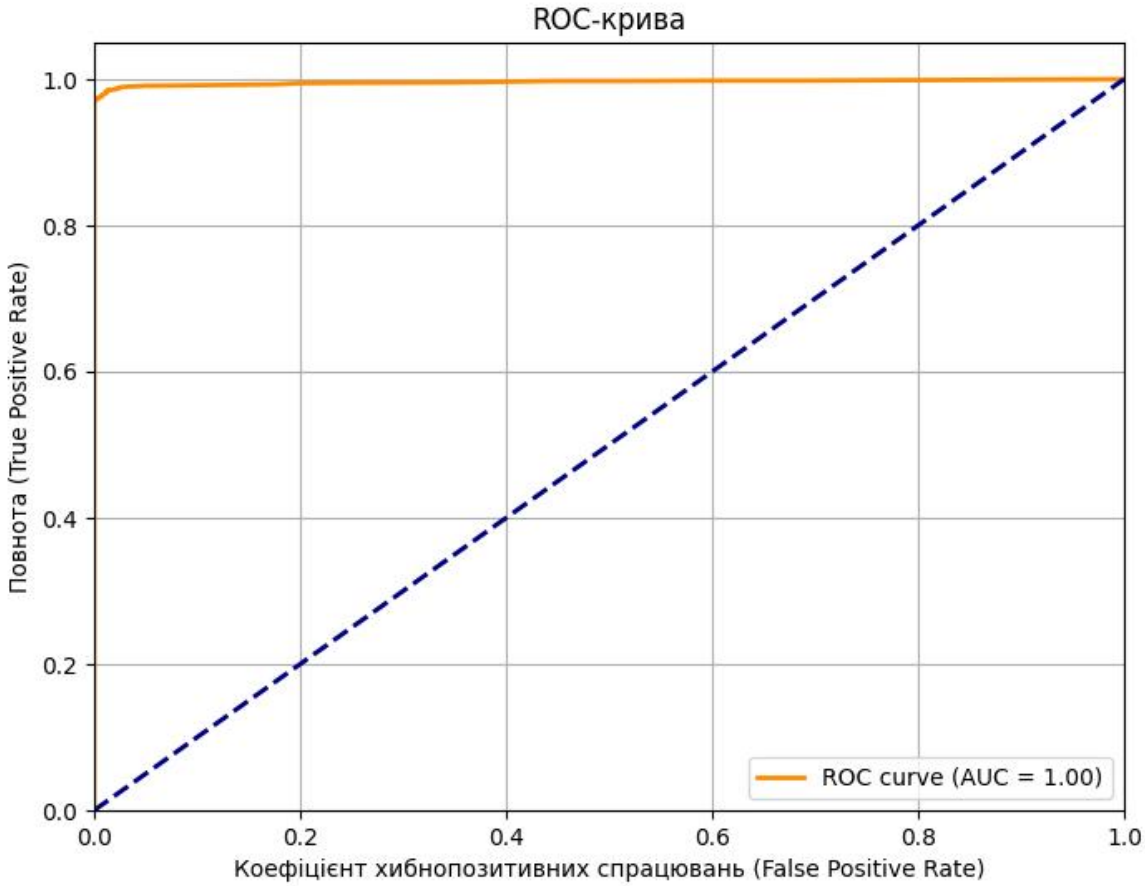

Рис.4.3.2. ROC-крива моделі "Випадковий ліс"

Для всебічної оцінки якості розробленої моделі "Випадковий ліс" у розрізненні фішингових та легітимних URL-адрес була побудована та проаналізована ROC-крива (Receiver Operating Characteristic curve). ROC-крива, представлена на рисунку 4.3.2, є графічним представленням діагностичної здатності бінарного класифікатора при зміні порогу дискримінації. Вона відображає залежність частки істиннопозитивних спрацювань (TPR, True Positive Rate, також відома як чутливість або повнота) від частки хибнопозитивних спрацювань (FPR, False Positive Rate, що дорівнює 1 - специфічність) для різних можливих порогових значень ймовірності, які модель призначає об'єктам.

Ідеальний класифікатор мав би ROC-криву, що проходить через верхній лівий кут (де TPR=1 та FPR=0), тоді як крива для моделі, що робить випадкові прогнози, приблизно збігається з діагоналлю, що йде з нижнього лівого кута (0,0) у верхній правий (1,1).

Кількісною мірою якості, що агрегує інформацію з ROC-кривої, є площа під нею (AUC-ROC, Area Under the ROC Curve). Значення AUC-ROC лежить в діапазоні від 0 до 1, де 1 відповідає ідеальному класифікатору, а 0.5 – випадковому.

Для розробленої моделі "Випадковий ліс", протестованої на відкладеній вибірці, розраховане значення площі під ROC-кривою (AUC-ROC) склало 0.9962. Таке високе значення, що значно наближається до 1, свідчить про відмінну розділювальну здатність моделі. Це означає, що модель з високою ймовірністю призначить вищу оцінку (ймовірність належності до класу "фішинг") випадково обраному фішинговому URL, ніж випадково обраному легітимному URL.

Візуальний аналіз ROC-кривої (рис. 4.3.2) також підтверджує високу ефективність моделі: крива розташована значно вище діагоналі випадкового вгадування та близько до верхнього лівого кута графіка. Це вказує на те, що модель здатна досягати високих значень повноти (TPR) при відносно низьких значеннях частки хибнопозитивних спрацювань (FPR) у широкому діапазоні порогових значень. Таким чином, аналіз ROC-кривої та значення AUC-ROC підтверджують високу прогностичну силу та надійність розробленої системи виявлення фішингових URL-адрес.

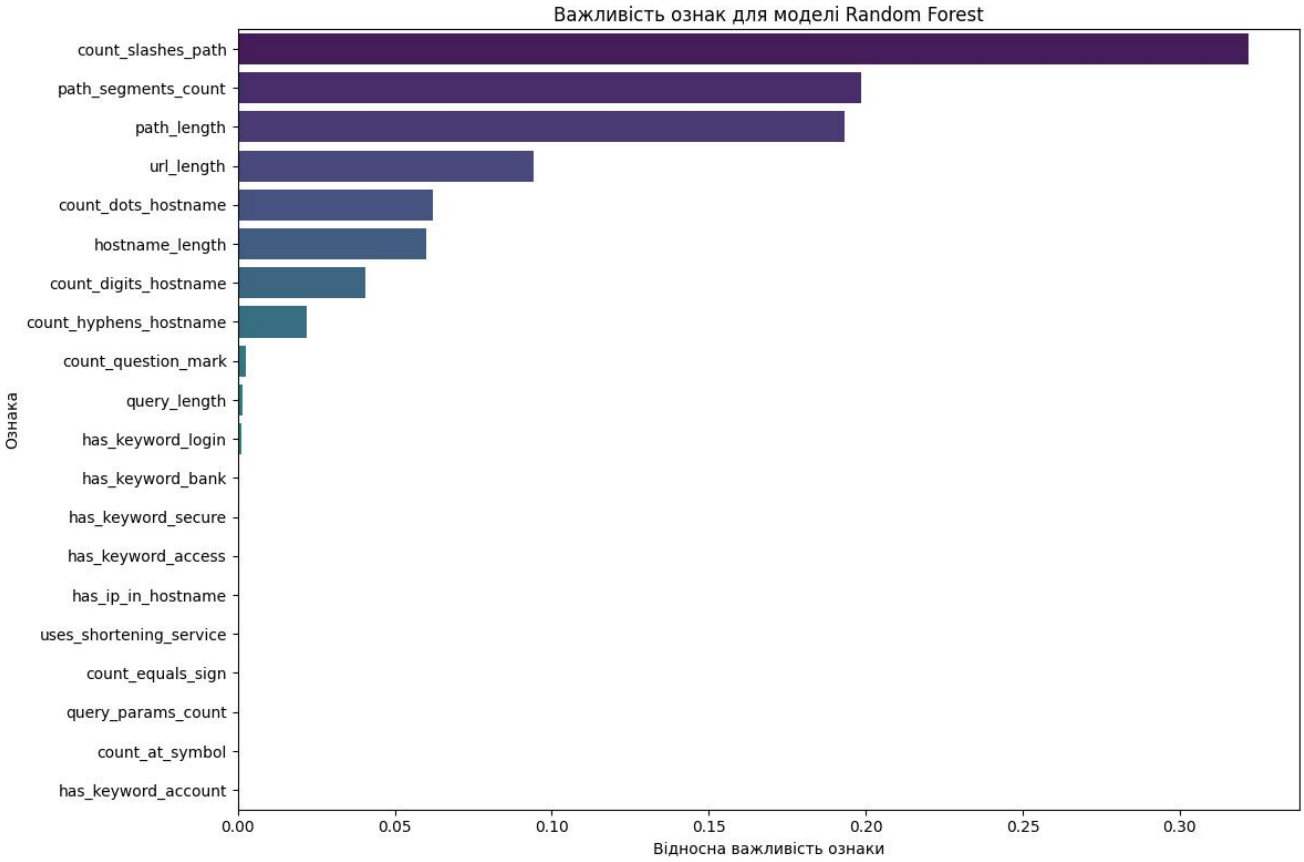

Рисунок 4.3.3. Відносна важливість ознак для моделі "Випадковий ліс" при виявленні фішингових URL-адрес.

На рисунку 4.3.3 представлено результати аналізу важливості ознак, використаних для навчання моделі "Випадковий ліс" з метою виявлення фішингових URL-адрес. Цей графік візуалізує відносний внесок кожної з 26 розроблених ознак у процес прийняття рішень класифікатором. Ознаки на діаграмі розташовані у порядку спадання їхньої важливості, що дозволяє наочно ідентифікувати найбільш значущі предиктори фішингу.

Ідентифікація ключових предикторів: З графіка чітко видно, що ознака "count_slashes_path" має найбільшу відносну важливість, що робить її домінуючим фактором при розрізненні фішингових та легітимних URL-адрес моделлю "Випадковий ліс".

Група високоінформативних ознак: Слідом за лідером йдуть такі ознаки, як "path_segments_count", "path_lengh" та "url_lengh", які також демонструють значний вплив на результати класифікації. Це свідчить про те, що комбінація цих характеристик URL є потужним індикатором потенційної загрози.

Нерівномірний розподіл важливості: Графік показує, що важливість розподілена між ознаками нерівномірно. Існує група ознак з високою та середньою важливістю, тоді як інші ознаки (наприклад, "has_keyword_account") мають суттєво менший вплив на прогнози моделі.

Практичне значення: Розуміння того, які саме ознаки є найбільш інформативними, дозволяє не тільки краще інтерпретувати поведінку моделі, але й може слугувати основою для подальшого вдосконалення системи. Наприклад, це може включати оптимізацію набору ознак шляхом видалення найменш значущих (якщо це не погіршить якість моделі) або концентрацію зусиль на більш точному вилученні та аналізі найважливіших характеристик.

4.4. Аналіз результатів та висновки.

У даному розділі було проведено експериментальну оцінку ефективності розробленої системи виявлення фішингових URL-адрес на основі моделі "Випадковий ліс" та 26 вилучених лексичних та хост-базованих ознак. Тестування здійснювалося на збалансованій тестовій вибірці, що складалася з 3998 URL-адрес (2000 легітимних та 1998 фішингових). Аналіз отриманих результатів дозволяє зробити наступні висновки.

4.4.1. Загальна ефективність моделі.

Розроблена модель продемонструвала високу загальну ефективність, що підтверджується значенням точності (Accuracy) 0.9850. Це означає, що 98.5% всіх URL-адрес у тестовій вибірці були правильно класифіковані системою. Крім того, висока розділювальна здатність моделі підтверджується значенням площі під

ROC-кривою (AUC-ROC), яке склало 0.9962. Таке близьке до 1 значення AUC вказує на відмінну здатність моделі розрізняти фішингові та легітимні URL-адреси у всьому діапазоні можливих порогів класифікації, що свідчить про її надійність та стійкість.

4.4.2. Аналіз ефективності виявлення фішингових загроз.

Критично важливими для систем безпеки є показники, що характеризують якість виявлення саме цільового класу – фішингових URL.

Показник повноти (Recall) для класу "фішинг" склав 0.9795, що свідчить про те, що система успішно ідентифікувала 97.95% всіх фішингових URL-адрес, наявних у тестовому наборі. Це високий показник, який вказує на незначну кількість пропущених загроз.

Точність (Precision) для класу "фішинг" досягла 0.9904. Це означає, що з усіх URL, які система позначила як фішингові, 99.04% дійсно були такими. Висока точність важлива для мінімізації хибних спрацьовувань на легітимні ресурси, що зменшує незручності для користувачів.

Інтегральна F1-міра для класу "фішинг", яка є гармонійним середнім між точністю та повнотою, склала 0.9849, підтверджуючи хороший баланс між здатністю виявляти фішинг та не помилятися при цьому.

4.4.3. Аналіз помилок класифікації.

Детальний аналіз помилок, представлений у матриці помилок (рис. 4.3.1), показує:

Кількість хибнопозитивних спрацювань (False Positives – легітимні URL, класифіковані як фішинг) склала 19 випадків. Це призвело до коефіцієнта хибнопозитивних спрацювань (FPR) 0.0095 (0.95%). Такий низький рівень FPR є позитивним аспектом, оскільки мінімізує ймовірність блокування безпечних ресурсів.

Кількість хибнонегативних спрацювань (False Negatives – фішингові URL, класифіковані як легітимні) склала 41 випадок. Відповідний коефіцієнт

хибнонегативних спрацювань (FNR) дорівнює 0.0205 (2.05%). Хоча цей показник є відносно низьким, пропуск навіть невеликої кількості фішингових загроз може мати серйозні наслідки для користувачів. Це вказує на потенційний напрямок для подальшого покращення системи, наприклад, шляхом ретельнішого аналізу характеристик URL, які були пропущені, або налаштування порогу класифікації.

4.4.4 Інтерпретація важливості ознак.

Аналіз важливості ознак (рис. 4.3.3) показав, що не всі з 26 розроблених ознак роблять однаковий внесок у процес прийняття рішень моделлю "Випадковий ліс". Були ідентифіковані ключові ознаки, такі як "count_slashes_path", "path_segments_count" та "path_lengh", які мають найбільший вплив. Це підтверджує адекватність обраного підходу до генерації ознак та вказує на те, які саме характеристики URL є найбільш інформативними для розрізнення фішингових та легітимних ресурсів. Розуміння важливості ознак також відкриває можливості для подальшої оптимізації моделі шляхом відбору найбільш значущих предикторів.

ВИСНОВКИ

У дипломній роботі було всебічно досліджено проблему фішингу як актуальної кіберзагрози та проаналізовано методи її виявлення, з особливим фокусом на аналізі URL-адрес.

У першому розділі було проаналізовано теоретичні засади феномену фішингу. Розглянуто його визначення, класифікацію за різними критеріями (цільова аудиторія, канали поширення), а також детально описано технічні методи фішингових атак, включаючи використання підроблених веб-сайтів та маніпуляції з URL-адресами. Особливу увагу приділено ролі соціальної інженерії як ключового елементу успішних фішингових кампаній. Проаналізовано сучасні тенденції та виклики у сфері фішингу, включаючи інтеграцію штучного інтелекту та нові методи маскування, що підкреслює динамічний характер загрози.

У другому розділі проведено аналіз існуючих підходів до виявлення фішингу. Розглянуто методи, що базуються на використанні чорних/білих списків, евристичного аналізу, аналізу URL-адрес та контенту сторінок, а також машинного навчання. Здійснено порівняльний аналіз цих підходів, визначено їхні сильні та слабкі сторони. Ідентифіковано ключові проблеми та недоліки існуючих методів, такі як їх реактивний характер, схильність до хибних спрацювань, ресурсомісткість та неможливість ефективно протистояти атакам нульового дня та постійній еволюції тактик зловмисників.

У третьому розділі було розроблено архітектуру та алгоритми запропонованої системи виявлення фішингових URL-адрес. Обґрунтовано вибір гібридного підходу з центральним використанням машинного навчання. В якості основного алгоритму класифікації обрано Випадковий ліс, обґрунтовано його переваги для даної задачі. Визначено та детально описано набір ознак, що вилучаються з URL-адрес (лексичні, хост-базовані, пов'язані зі шляхом/запитом). Обрано технології та інструменти для реалізації системи, включаючи мову

програмування Python та відповідні бібліотеки (Scikit-learn, Pandas, urllib, tldextract), та описано процес реалізації основних модулів.

У четвертому розділі представлено експериментальну оцінку ефективності розробленої системи. Описано процес формування та підготовки тестового набору даних, що включав близько 10 000 збалансованих фішингових та легітимних URL-адрес. Детально висвітлено методику проведення експериментів із розділенням даних на навчальну та тестову вибірки та використанням стандартних метрик класифікації (Точність, Повнота, Точність, F1-міра, FPR, FNR). Проведено аналіз результатів, який підтвердив високу ефективність запропонованої системи.

Загалом, результати дослідження підтверджують, що фішинг є серйозною загрозою для інформаційної безпеки, яка потребує системного підходу до виявлення та протидії. Розроблена система, що базується на аналізі характеристик URL-адрес із застосуванням машинного навчання, демонструє високу ефективність у виявленні фішингових загроз, що відповідає меті роботи та сучасним викликам у сфері кібербезпеки.

Практична цінність роботи полягає у розробці функціональної архітектури та реалізації прототипу системи виявлення фішингових URL-адрес, яка може слугувати основою для вдосконалення існуючих інструментів кібербезпеки (таких як веб-фільтри або розширення для браузерів) з метою підвищення рівня захисту користувачів від фішингових атак.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. 11 Parts of a URL: A Complete Guide | Hostwinds. Hostwinds Blog. URL: https://www.hostwinds.com/blog/11-parts-of-url-complete-guide (date of access: 09.06.2025).

2. arXiv.org e-Print archive. URL: https://arxiv.org/pdf/2103.12739 (дата звернення: 09.06.2025).

3. A Study of Effectiveness of Brand Domain Identification Features for Phishing Detection in 2025. arXiv.org e-Print archive. URL: https://arxiv.org/html/2503.06487v1 (date of access: 09.06.2025).

4. CICS Transaction Server for z/OS. IBM - United States. URL: https://www.ibm.com/docs/en/cics-ts/6.x?topic=concepts-components-url (date of access: 09.06.2025).

5. DB Event List. URL: https://db-event.jpn.org/deim2019/post/papers/201.pdf (дата звернення: 09.06.2025).

6. Detecting Phishing URLs in QR codes using Heuristic Techniques - NORMA@NCI Library. Welcome to NORMA@NCI Library - NORMA@NCI Library. URL: https://norma.ncirl.ie/7127/ (date of access: 09.06.2025).

7. GeeksforGeeks. Components of a URL - GeeksforGeeks. GeeksforGeeks. URL: https://www.geeksforgeeks.org/components-of-a-url/ (date of access: 09.06.2025).

8. Heuristic machine learning approaches for identifying phishing threats across web and email platforms - PMC. PMC Home. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC11532189/ (date of access: 09.06.2025).

9. Huddersfield Repository - University of Huddersfield. URL: https://eprints.hud.ac.uk/id/eprint/24330/6/MohammadPhishing14July2015.pdf (дата звернення: 09.06.2025).

10. IJNRD - UGC CARE Journal Norms and Guidelines follow - International Journal of Novel Research and Development. URL: https://www.ijnrd.org/papers/IJNRD2407527.pdf (дата звернення: 09.06.2025).

11. Learning C. A Deep Dive into Phishing URL Detection. CRAC. URL: https://www.crac-learning.com/post/a-deep-dive-into-phishing-url-detection (date of access: 09.06.2025).

12. Modeling Hybrid Feature-Based Phishing Websites Detection Using Machine Learning Techniques - PMC. PMC Home. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8935623/ (date of access: 09.06.2025).

13. NSF Public Access. URL: https://par.nsf.gov/servlets/purl/10523848 (дата звернення: 09.06.2025).

14. Protection Highlight: Unblur the HTML - How Phishing Attacks Exploit HTML Function. Broadcom Inc. | Connecting Everything. URL: https://www.broadcom.com/support/security-center/protection-bulletin/protection-highlight-unblur-the-html-how-phishing-attacks-exploit-html-function (date of access: 09.06.2025).

15. ScholarWorks. URL: https://scholarworks.calstate.edu/downloads/v692td67g (дата звернення: 09.06.2025).

16. UVicSpace : Home. URL: https://dspace.library.uvic.ca/bitstreams/036da286-8ccd-4fa5-a0f4-1628aca184c8/download (дата звернення: 09.06.2025).

17. What Is Phishing and How To Avoid It - National Cybersecurity Alliance. National Cybersecurity Alliance. URL: https://www.staysafeonline.org/articles/phishing (date of access: 09.06.2025).

18. What Is Phishing? - Meaning, Attack Types & More | Proofpoint US. Proofpoint. URL: https://www.proofpoint.com/us/threat-reference/phishing (date of access: 09.06.2025).

19. What is Phishing? Techniques and Prevention | CrowdStrike. CrowdStrike: We Stop Breaches with AI-native Cybersecurity. URL:

https://www.crowdstrike.com/en-us/cybersecurity-101/social-engineering/phishing-attack/ (date of access: 09.06.2025).

20. What is Phishing? Techniques and Prevention | CrowdStrike. CrowdStrike: We Stop Breaches with AI-native Cybersecurity. URL: https://www.crowdstrike.com/en-us/cybersecurity-101/social-engineering/phishing-attack/#:~:text=Phishing%20is%20a%20type%20of,numbers,%20and%20other%20personal%20details (date of access: 09.06.2025).

21. arXiv.org e-Print archive. URL: https://arxiv.org/pdf/1910.06277 (дата звернення: 09.06.2025).

22. arXiv.org e-Print archive. URL: https://www.arxiv.org/pdf/2504.16449 (дата звернення: 09.06.2025).

23. Avinashilingam Institute for Home Science and Higher Education for Women | Avinashilingam Institute for Home Science and Higher Education for Women. URL: https://avinuty.ac.in/sites/avinuty.ac.in/files/2024-09/Nandhini%20Report%20.pdf (дата звернення: 09.06.2025).

24. Deep Learning Empowered Phishing URL Detection: An Exhaustive Approach | International Journal of Intelligent Systems and Applications in Engineering. International Journal of Intelligent Systems and Applications in Engineering. URL: https://ijisae.org/index.php/IJISAE/article/view/4659 (date of access: 09.06.2025).

25. GitHub - BusamSumanjali/URL-Based-Phishing-Detection-using-machine-Learning: Phishers Develop the websites similar to those real websites. So, this project comes to know whether the URL is phishing or not. GitHub. URL: https://github.com/BusamSumanjali/URL-Based-Phishing-Detection-using-machine-Learning (date of access: 09.06.2025).

26. GitHub - deepeshdm/Phishing-Attack-Domain-Detection: Identifying Malicious Phishing URLs through Machine Learning. GitHub. URL: https://github.com/deepeshdm/Phishing-Attack-Domain-Detection (date of access: 09.06.2025).

27. GitHub. GitHub. URL: https://github.com/wang-zifu/Phishing-URL-Detector-Using-URL-NLP-Features (date of access: 09.06.2025).

28. GitHub - shreyagopal/Phishing-Website-Detection-by-Machine-Learning-Techniques. GitHub. URL: https://github.com/shreyagopal/Phishing-Website-Detection-by-Machine-Learning-Techniques (date of access: 09.06.2025).

29. i-manager Publications. imanager publications. URL: https://imanagerpublications.com/article/15698/ (date of access: 09.06.2025).

30. Improved Detection of Phishing Websites using Machine Learning | International Journal of Intelligent Systems and Applications in Engineering. International Journal of Intelligent Systems and Applications in Engineering. URL: https://ijisae.org/index.php/IJISAE/article/view/6351 (date of access: 09.06.2025).

31. International Journal of Latest Engineering Research and Applications. URL: http://www.ijlera.com/papers/v2-i3/20.201703073.pdf (дата звернення: 09.06.2025).

32. Leitfaden für Neulinge zu den 10 Bestandteilen einer URL | Mailchimp. Mailchimp. URL: https://mailchimp.com/resources/parts-of-a-url/ (date of access: 09.06.2025).

33. Machine learning-based phishing detection using URL features: a comprehensive review. Mountain Scholar :: Home. URL: https://mountainscholar.org/items/60c9901a-d25e-473d-a67d-4aa909419785 (date of access: 09.06.2025).

34. NSF Public Access. URL: https://par.nsf.gov/servlets/purl/10408623 (дата звернення: 09.06.2025).

35. | PDF or Rental. article-page. URL: https://articles.researchsolutions.com/lexical-feature-based-phishing-url-detection-using-online-learning/doi/10.1145/1866423.1866434 (date of access: 09.06.2025).

36. Phishing Website Detection Based on URL. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. URL: https://ijsrcseit.com/CSEIT2173124 (date of access: 09.06.2025).

37. Phishing-Website-Detection-by-Machine-Learning-Techniques/URL Feature Extraction.ipynb at master · shreyagopal/Phishing-Website-Detection-by-Machine-Learning-Techniques. GitHub. URL: https://github.com/shreyagopal/Phishing-Website-Detection-by-Machine-Learning-Techniques/blob/master/URL%20Feature%20Extraction.ipynb (date of access: 09.06.2025).

38. sec-paper/Phishing/Lexical Feature Based Phishing URL Detection Using Online Learning.pdf at master · secdr/sec-paper. GitHub. URL: https://github.com/secdr/sec-paper/blob/master/Phishing/Lexical%20Feature%20Based%20Phishing%20URL%20Detection%20Using%20Online%20Learning.pdf (date of access: 09.06.2025).

39. Simple search. URL: https://www.diva-portal.org/smash/get/diva2:1773760/FULLTEXT02 (дата звернення: 09.06.2025).

40. Warse. URL: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse242942020.pdf (дата звернення: 09.06.2025).

41. Bytesize Security: A Guide to HTML Phishing Attachments. Darktrace | The Essential AI Cybersecurity Platform. URL: https://www.darktrace.com/blog/bytesize-security-html-phishing-attachments (date of access: 09.06.2025).

42. Deep Learning Empowered Phishing URL Detection: An Exhaustive Approach | International Journal of Intelligent Systems and Applications in Engineering. International Journal of Intelligent Systems and Applications in Engineering. URL: https://ijisae.org/index.php/IJISAE/article/view/4659 (date of access: 09.06.2025).

43. The Science and Information (SAI) Organization. URL: https://thesai.org/Downloads/Volume11No4/Paper_77-Feature_Selection_for_Phishing_Website.pdf (дата звернення: 09.06.2025).