

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)

Рівень вищої освіти другий (магістерський)

Виконав:
студент 2 курсу, групи АМСЗІм-20-1
Яіцький А. О.
(прізвище, ініціали)

Освітня програма: Адміністративний менеджмент
у сфері захисту інформації
(повна назва освітньої програми)

2022 p.

Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 125 Кібербезпека
(код і повна назва)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Адміністративний менеджмент у сфері захисту інформації
(повна назва)

ЗАТВЕРДЖУЮ

Зав. кафедри _____
(підпис)

« ____ » _____ 2022 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студента Яіцького Андрія Олександровича
(прізвище, ім'я, по батькові)

1. Тема роботи: Дослідження процедур виявлення фішингових повідомлень затверджена наказом по університету від «24» березня 2022 р. №410 Ст.
2. Термін подання студентом роботи до екзаменаційної комісії 26.05.2022 р.
3. Вихідні дані до роботи: методи машинного навчання, алгоритми машинного навчання "Multilayer perceptron", "Logistic regression", "Naive Bayes", "J48", "Bayes Net", методи виявлення фішингових повідомлень, DNS Gets Anti-Phishing Hook, Dark Reading, Phishing E-mail Detection Based on Structural Properties.
4. Перелік питань, що потрібно опрацювати в роботі:
 - 1) Актуальність проблеми фішингу
 - 2) Аналіз признаков фішонгових повідомлень
 - 3) Розробка процедур автоматизації обробки фішингових повідомлень

5. Перелік графічного матеріалу із зазначенням креслень, плакатів, комп'ютерних ілюстрацій: Демонстраційний матеріал у вигляді ppt-презентації

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		(підпис)	(дата)
Основна частина	професор кафедри ІКІ ім. В.В. Поповського Марчук Володимир Степанович		27.05.2022

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання	25.03.2022	Виконано
2	Збір матеріалів для дослідження	29.03.2022	Виконано
3	Розробка 1 розділу	01.04.2022	Виконано
4	Розробка 2 розділу	20.04.2022	Виконано
5	Розробка 3 розділу	25.04.2022	Виконано
6	Оформлення кваліфікаційної роботи	10.05.2022	Виконано

Дата видачі завдання 24 березня 2022 року

Студент  Яіцький А.О.
(підпис) (прізвище, ініціали)

Керівник роботи  професор Марчук В.С.
(підпис) (посада, прізвище, ініціали)

Робота не містить відомостей заборонених до відкритого опублікування

Студент  Андрій ЯІЦЬКИЙ
Керівник  Володимир МАРЧУК

РЕФЕРАТ

Пояснювальна записка: 85 с., 36 рис., 1 табл., 20 джерел.

АТАКИ З ВИКОРИСТАННЯМ ФІШИНГУ, ЗАГРОЗИ, ІНФОРМАЦІЙНА БЕЗПЕКА, ОЗНАКИ ФІШИНГОВИХ ПОВІДОМЛЕНЬ, ФІШИНГ.

Об'єкт дослідження - процес обробки та виявлення фішингових повідомлень.

Предмет дослідження - методи та засоби обробки та виявлення фішингових повідомлень.

Мета кваліфікаційної роботи - підвищення ефективності та оперативності процедур обробки та виявлення фішингових повідомлень.

Методи дослідження - аналіз, спостереження, моделювання, експериментальні дослідження.

У данній кваліфікаційній роботі розглянуті принципи фішингу, типи та схеми побудови фішингу, основні методи та засоби захисту від фішингу.

Виконаний аналіз основних наукових робіт та патентний пошук у галузі фішингу. Розглянути основні ознаки для виявлення фішингових повідомлень. Запропоновані процедури для автоматизації процесу обробки та виявлення фішингових повідомлень.

ABSTRACT

The report contains: 85 pp., 36 figs., 1 tab., 20 sources.

ATTACKS USING FISHING, THREATS, INFORMATION SECURITY,
SIGNS OF FISHING MESSAGES, FISHING.

The object of research is the process of processing and detecting phishing messages.

The subject of research - methods and means of processing and detecting phishing messages.

The purpose of certification work is to increase the efficiency and efficiency of procedures for processing and detecting phishing messages.

Research methods - analysis, observation, modeling, experimental research.

In this attestation work the principles of phishing, types and schemes of phishing construction, basic methods and means of protection against phishing are considered.

The analysis of the basic scientific works and patent search in the field of phishing is executed. Consider the main features for detecting phishing messages. Procedures are proposed to automate the process of processing and detecting phishing messages.

ЗМІСТ

Перелік скорочень, умовних позначень, символів, одиниць і термінів.....	7
Вступ.....	8
1 Застосування та актуальність фішингу	11
1.1 Загальний огляд фішингу.....	11
1.2 Схеми побудови фішингу.....	13
1.3 Фішонгові атаки.....	17
1.4 Методи захисту фішингових атак.....	26
1.5 Аналіз робіт і патентний пошук в області фішинга.....	34
1.6 Статистика фішингових атак та актуальність проблеми фішингу..	40
2 Аналіз та дослідження ознак фішингових повідомлень.....	46
2.1 Ознаки фішингових повідомлень	46
2.2 Дослідження аналізу фішингових повідомлень.....	57
3 Розробка процедур автоматизації та проведення модельного експери- менту з обробки фішингових повідомлень.....	67
3.1 Порівняння алгоритмів машинного навчання та вибір найкращого алгоритму.....	67
3.2 Обговорення та припущення щодо розробки процедур автомати- зації обробки та виявлення фішингових повідомлень.....	71
3.3 Побудова моделі та розробка процедури автоматизації обробки та виявлення фішингових повідомлень	74
Висновки.....	81
Перелік джерел посилання.....	83

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ І ТЕРМІНІВ

ЗМІ – засоби масової інформації
ПЗ – програмеї забезпечення
ТТР – тактики, методи та процедури
ФБР – федеральне бюро розслідувань
2FA – двофакторна автентифікація
APT – постійна серйозна загроза
CVC – код підтвердження картки
DKIM – метод email-автентифікації
DLP – програмне забезпечення для запобігання втраті даних
DMARC – автентифікація повідомлень на основі домену
DNS – система доменних імен
EOP – хмарна служба фільтрації
HTML – мова гіпертекстової розмітки
HTTP – протокол передачі гіпертексту
HTTPS – захищений протокол передачі гіпертексту
IP – міжмережевий протокол
ISP – інтернет-провайдер
MFA – багатофакторна автентифікація
OSINT – розвідка з відкритих джерел
PDF – портативний формат документів
PSI – персональний ключ безпеки
SMS – служба коротких повідомлень
SPF – інфраструктура політики відправника
SSID – назва безпроводної точки доступу Wi-Fi
URL – уніфікований показник ресурсу
VAPT – оцінка вразливості та тестування на проникнення
WIPS – системи запобігання вторгнення у бездротову мережу

ВСТУП

На сьогоднішній день, фішинг, а саме атаки на основі соціальної інженерії, і більш цілеспрямовані атаки фішинга, залишаються одним з найбільш поширених векторів атак які використовуються зловмисниками, і це є проблемою безпеки. Фішинг має на увазі нелегальний процес спроби для того, щоб зловмисник зміг отримати конфіденційну інформацію, таку інформацію як паролі, дані кредитної картки, імена користувачів, логіни, дані від пошти, і так далі. При цьому, зловмисник під час фішингу маскується під людину, яка заслуговує на довіру в тій чи іншій сфері. Публічний випуск наступальних інструментів безпеки та тактик, методів та процедур (ТТР), таких як розкриття вразливостей за допомогою робочого коду експлойту для перевірки концепції, активно використовується зловмисниками у кампаніях. Досить часто просунуті постійні загрози (APT) та інші складні суб'єкти загроз, замість того, щоб концентрувати час і ресурси на дослідженні раніше невідомих вразливостей, відомих як нульовий день, використовують існуючі функції або використовують вже відомі вразливості, які ще не були виправлені.

За останні кілька років, фішинг є найчастішим видом кіберзлочинності. Фішинг призвів до крадіжки мільярдів доларів. Все це відбувається через те, що фішингові зловмисники завжди використовують нові (нульового дня) та витончені методи для того, щоб перехитрити онлайн-покупців. Інтернет надає людям та компаніям новий спосіб зв'язку, що призводить до зміни способу життя людей, наприклад, онлайн-покупки та онлайн-освіта почали замінювати магазини та школи. Також це викликало зрушення в компаніях, які на даний момент сконцентровані на створенні нових інтернет-продуктів та послуг, щоб мати можливість домінувати на ринку. Але, такий зсув дає можливість зловмисникам вчиняти нові види злочинів, види злочинів з використанням комп'ютерів та мереж, які можна назвати як кіберзлочини. При тому, що вкладаються великі зусилля наукових кіл та промисловості, антифішингові дослідження все ще стикаються із низкою проблем. Незважаючи на великий об'єм досліджень теми фішингу в літературі, з точки зору фахівців з інформаційної безпеки, використовується тільки кілька категорій фішингу. Наведемо приклад скомпрометованих ресурсів:

- соціальні мережі (Youtube, Facebook, Myspace);
- аукціонні сайти (eBay);

- онлайн-банки (Приватбанк, Ощадбанк);
- сервіси веб-грошей (PayPal, WebMoney);
- іт-адміністратори (Microsoft, Yahoo, інтернет-провайдери).

Це люди, які працюють у великих компаніях, банках і так далі. Фішингові атаки можуть мати кілька наслідків для компаній. Вони можуть призвести до витоку даних, середня вартість яких може становити 4,80 мільйона доларів. Така втрата грошей визначається низкою факторів, у тому числі збитком для репутації, втратою вартості компанії та руйнуванням бізнесу.

Компанія, яка може зіткнутися із серйозним порушенням, не зможе тримати це просто так. Співробітники, партнери та клієнти матимуть ненадійне уявлення про бізнес, якщо відбудеться витік даних. Публічність, пов'язана із серйозним витоком даних, може зруйнувати репутацію бренду компанії. А це може призвести до втрати співробітників, партнерів та клієнтів.

Що стосується втрати вартості компанії, то нижче наведено приклад найдоржчих фішингових атак за версією The SSL Store [3].

1) У період з 2013 по 2015 рік Facebook та Google були обдурені на 100 мільйонів доларів за допомогою ретельно продуманої афери з рахунками.

2) Crelan Bank у Бельгії втратив 75,8 мільйонів доларів внаслідок шахрайства з генеральним директором.

3) FACC, австрійський виробник деталей для аерокосмічної техніки, втратив 61 мільйон доларів через шахрайство з генеральним директором.

4) Upsher-Smith Laboratories, американська фармацевтична компанія, втратила понад 50 мільйонів доларів протягом трьох тижнів через фішерів, які видавали себе за генерального директора.

5) Ubiquiti Networks, американська компанія, що займається комп'ютерними мережами, не дізналася, що у неї було викрадено 46,7 мільйона доларів, використовуючи електронні листи генерального директора про шахрайство, і була повідомлена про це за допомогою ФБР.

Це приклади того, скільки грошей компанії втратили лише через фішингові атаки. Ці суми не враховують втрати користувачів, партнерів або співробітників. За даних атак, компанії втратили до 40% клієнтів. Якщо провести невеликий аналіз, то можна зробити висновок, що проблема фішингу є досить актуальною. Для того, щоб компаніям, користувачам було простіше обробляти фішингові повідомлення, необхідно розробити припущення щодо автоматизації даного процесу.

У цій кваліфікаційній роботі розглянуто проблему з нової точки зору. Окремі результати роботи, доповідалися на міжнародних наукових конференціях [1, 2]. Замість того, щоб просто виявляти та блокувати фішингові атаки, проведено дослідження ознак фішингових повідомлень, а також яким чином можна полегшити цей процес. На даний момент комунікаційні дослідження фішингу дуже обмежені. Незважаючи на величезну кількість роботи, проведеної на тему фішингу, є мало інформації з якісної точки зору з боку фахівців з безпеки. Мета даної роботи, це розширення поточних досліджень фішингу. При цьому за рахунок аналізу ознак фішингових повідомлень, звужується коло досліджень. Погляд на фішинг з погляду тих, хто запобігає появі фішингових загроз, а не досліджує жертв, може відкрити нові способи скорочення фішингу шляхом виявлення слабких місць. Дане дослідження спрямоване на те, щоб відповісти на конкретне питання про те, які можуть бути способи, припущення для автоматизації обробки фішингових повідомлень.

1 ЗАСТОСУВАННЯ ТА АКТУАЛЬНІСТЬ ФІШИНГУ

1.1 Загальний огляд фішингу

Фішинг можна визначити як форму методу соціальної інженерії, коли мета зловмисника це збір особистої інформації жертви або встановлення шкідливого програмного забезпечення у систему жертви. Ця дія може бути досягнута переконанням користувача клацнути на шкідливе посилання в електронному листі, яка перенаправляє користувача на підроблений веб-сайт, створений з єдиною метою збору конфіденційної інформації, такої як облікові дані. Такі підроблені сайти, які використовуються для фішингу, зазвичай нагадують відомі веб-сайти, такі як банківські сайти, соціальні мережі або навіть власні сайти цільової організації. Зловмисники роблять це для того, щоб у мети не виникало питань у підробленості даного сайту. У деяких випадках деякі сайти можуть виглядати явно підробленими, які з неправильними доменними іменами, і з помилками в граматиці. Зазвичай, звичайні атаки фішингу розсилаються масово сотням, або навіть тисячам одержувачам.

Така ймовірність успіху в основних атаках фішинга може сильно відрізнятися. Щоб досягти більш високих показників успіху, зловмисники можуть вдатися до цільового фішингу або цілеспрямованих атак на певні організації, людей та компанії. Під цільовим фішингом мається на увазі більш цілеспрямована форма фішингової атаки, тобто у даного фішингу є своя ціль і своя жертва. При цільовому фішингу зловмисники не поспішають проводити дослідження, збираючи величезну кількість інформації про свої цілі. Ця інформація використовується для того, щоб атака виглядала більш законною та релевантною, щоб обдурити одержувача, і щоб даний одержувач видав свою інформацію або встановив шкідливе програмне забезпечення на свій комп'ютер. Під час таких цілеспрямованих фішингових атак, зловмисники можуть надіслати електронні листи, які виглядають як законні електронні листи, які на додачу містять очікувані логотипи, і навіть блок підпису. Навіть шкідливе посилання або вкладення можна замаскувати, щоб вони виглядали законними.

Через високу складність дизайну та претексту, користувачам може бути дуже складно виявити та захиститися від цільових фішингових атак. В даному випа-

дку, можна описати попереднє текстове повідомлення як дію зі створення продуманого сценарію, для того щоб переконати цільову жертву розкрити інформацію або виконати дію, цю зловмисник може використовувати як перевагу.

Претекстинг у цьому випадку дає соціальним інженерам перевагу. Якщо зловмисник може надати достатню інформацію у фішинговому електронному листі, що відповідає дійсності, і що вона має законне та авторитетне джерело, то шанси зловмисника на успіх значно зростають.

Розгляд фішингу у межах комунікаційної дисципліни зачіпає суть безпеки. Фішинг можна також інтерпретувати як широкий термін, оскільки в багатьох дослідженнях розглядаються різні аспекти, засоби та стратегії. Найчастіше досліджувані типи фішингових атак, це веб-сайти та електронні листи. Однак, фішинг може відбуватися в контексті багатьох інших форм ЗМІ, наприклад, таких як соціальні мережі та інші служби обміну повідомленнями. Досить часто якщо є можливість для цифрового зв'язку, тобто потенціал для фішингу. Не завжди технічні рішення для зниження загроз фішингу бувають успішними, тому що вони покладаються на кінцевого користувача, який вирішує, як інтерпретувати кожне повідомлення та визначати, які повідомлення є шкідливими. Масштаби фішингу можуть змінюватися від загального фішингу до цільового фішингу.

Цільовий фішинг - це цілеспрямована комунікаційна атака, при якій контекст фішингових повідомлень специфічний для конкретної мети і, отже, не має сенсу для будь-якої іншої мети. Що стосується загального фішингу, то він відноситься до повідомлень, які відносяться до набагато ширшої аудиторії і не адаптовані для кожної мети. В даному випадку є ще один підвид фішингу, який можна назвати як "китобійним", тобто коли цільова фішингова атака спрямована на керівників великих підприємств. Коли ці цілі розміром з "кита" скомпрометовані внаслідок злому, вони можуть принести зловмиснику велику фінансову вигоду або надати додатковий доступ до організації. Тобто, підбиваючи підсумки, можна сказати, що фішинг - це унікальна загроза кібербезпеці, яка включає як технологічні вразливості, так і поєднання психологічних, соціологічних і комунікаційних впливів.

1.2 Схеми побудови фішингу

Тепер розглянемо як працюють фішингові веб-сайти, для цього нижче на рис.1.1 наведено приклад схеми роботи фішингового веб-сайту.



Рисунок 1.1 - Схема роботи фішонгових сайтів

Отже, коли йде мова про налаштування фішингового веб-сайту, як видно з рис.1.1, на даному рисунку три різні об'єкти, це клієнт, проксі-сервер зловмисника, та реальний сайт. Отже, на цьому проксі-сервері зловмисника ми приховали вхід у систему всіх введених даних користувача. Після того, як дані були зареєстровані або отримані на проксі-сервері зловмисника, проксі-сервер перенаправляє клієнта на реальний веб-сайт, показуючи повідомлення «неправильний пароль» або «відбулася помилка, спробуйте ще раз». Таким чином, в більшості випадків фішингові сайти працюють саме так і це не тільки логіни та паролі, але й дані вашої банківської картки. Після того, як була введена вся інформація, включаючи код CVC, сайт показує якусь помилку, але при цьому реєструється вся інформація, яка була введена.

Коли йде мова про фішингові атаки, то необхідно мати поняття як фішингові атаки будуються, тому в даному випадку розглянемо схеми побудови фішингових атак, їх кілька, а саме:

- традиційний (термін життя фішингового сайту - 62 години);
- rock-phish (термін життя фішингового сайту - 172 години);
- fast-flux (термін життя фішингового сайту - 196 годин).

Розглянемо спершу традиційну схему. Ця схема показана на рис. 1.2. нижче.

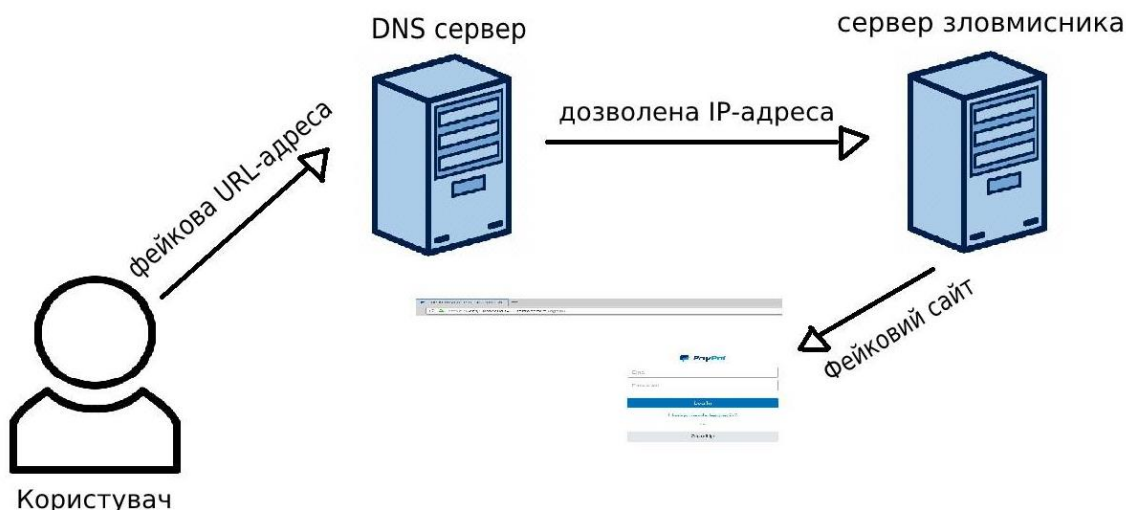


Рисунок 1.2 - Схема традиційної фішонгової атаки

Як видно з рис.1.2, дана схема складається з трьох елементів і трьох різних стрілок. Перша стрілка з підробленою URL-адресою. Ця підроблена URL-адреса має бути якимось чином передана жертві. Жертва натискає на цю URL-адресу і перенаправляється. Спочатку URL-адреса дозволяється DNS сервером, потім дозволена IP-адреса вказує на сервер зловмисника, і далі підроблена веб-сторінка повертається користувачеві. Це досить проста традиційна атака фішингу, вона не дуже складна і не вимагає багато засобів для проведення цієї атаки. Але, крім того, це вразлива фішингова атака, тому що її легко нейтралізувати. Для нейтралізації такої атаки необхідно повідомити інтернет-провайдера, що ця дозволена IP-адреса містить підроблену веб-сторінку або фішинговий веб-сайт, і Інтернет-провайдер проводить розслідування, видаляючи цю IP-адресу. Після цього фішинговий веб-сайт не працює. Тепер розглянемо Rock-phish схему, дана схема показана на рис.1.3 нижче.

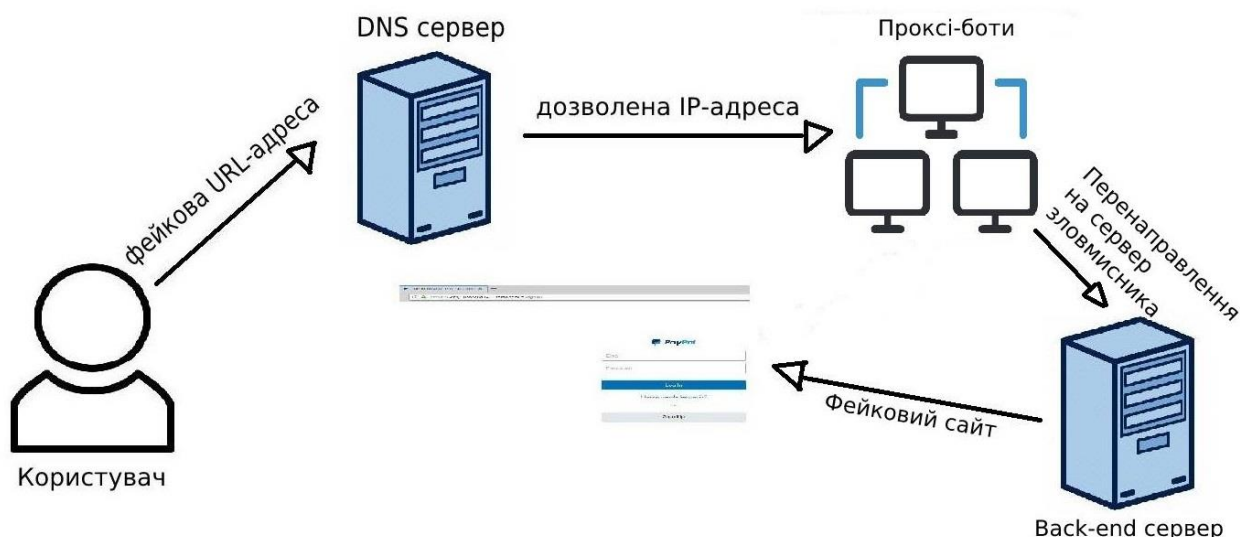


Рисунок 1.3 - Схема Rock-phish атаки

З рис.1.3 вище видно, що якщо дивитися на Rock-phish, то це складніша атака. Ця стратегія атаки намагається уникнути виявлення та максимізує вразливість фішингового сайту. Таким чином, елементи атаки розділяються при додаванні надмірності перед запитом на видалення. Зловмисник спочатку обробляє ряд доменних імен, коротких і, як правило, безглузвих, наприклад щось на зразок "lof80.info". Потім, на етапі розповсюдження, спам приходить електронною поштою, що містить цю довгу URL-адресу. Вона поширюється, і перша її частина є індикатором того, що сайт виглядає універсальним. Такий механізм, як сервер доменних імен (DNS) з підстановними знаками, можна використовувати для вирішення такого варіанту конкретного IP-адреса, який в свою чергу буде вказувати на проксі-бота. Таким чином, кожен домен, який міститься на цьому внутрішньому сервері, зіставляється з динамічним пулом компрометованих машин відповідно до сервера імен, контролюваного зловмисником.

На кожній компрометованій машині працює проксі-система, яка пов'язує запит із серверною системою. І цей внутрішній сервер завантажений великою кількістю іноді навіть до двадцяти фішингових сайтів одночасно. І всі ці веб-сайти доступні з будь-якої з Rock-phish машин. Однак те, який сайт буде досягнутий, залежить від того, яка URL-адреса передається після першої косої межі. І Rock-phish сайти мають спільний хост, тому якщо один з цих веб-сайтів видалається, якщо один з сайтів повідомляється і IP-адреса цього внутрішнього сервера відключений, сайт автоматично переключається. Отже, щоб усунути Rock-phish, потрібно прибрати всіх проксі-ботів. Якщо говорити про те, наскільки вразлива ця

конкретна атака, то вона, звичайно, більш робастна, ніж традиційний фішинг, але якщо дивитися на цих проксі-ботів, то вони завжди мають певний масив ір-адрес, і якщо можна ідентифікувати всі проксі-сервери ботів, які звертаються до внутрішнього сервера, тоді можна нейтралізувати цю схему атаки фішингу.

Але є ще складніша схема, яка називається «Fast-Flux Phishing». Вона представлена на рис.1.4. нижче.

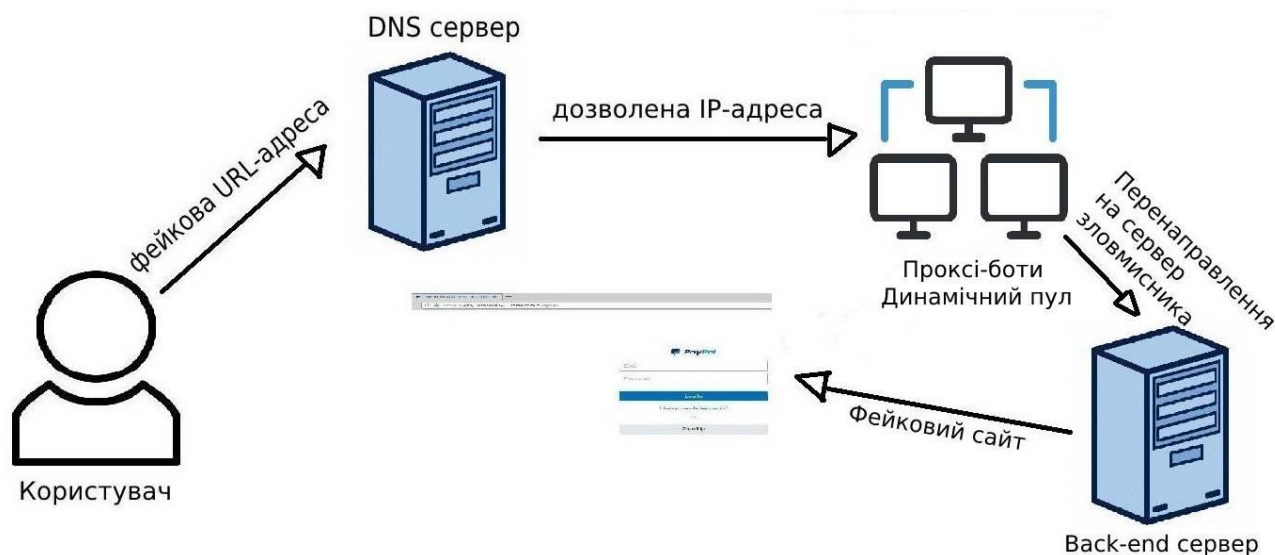


Рисунок 1.4 - Схема Fast-Flux атаки

Отже, як видно із рис.1.4. основна ідея мережі Fast-Flux полягає в тому, щоб мати кілька IP-адрес, пов'язаних з доменним ім'ям, а потім постійно міняти їх у швидкій послідовності. Таким чином, машини в такому механізмі фішингу Fast-Flux, які становлять цей тип мережі, фактично не несуть відповідальності за розміщення та завантаження шкідливого контенту жертв. Завдання у тому, щоб виділити кілька машин, які у ролі серверів, серверів цього шкідливого контенту. Інші просто діють як редиректори, які допомагають маскувати реальну адресу цих систем, які під контролем злочинців.

Злочинці забезпечують максимально можливу доступність критично важливих систем у мережі і навіть розгортають системи балансування навантаження для обробки всіх запитів на завантаження шкідливого контенту. Звичайною практикою є перегляд стану мережі на звичайних міжмережевих екранах, щоб відкинути будь-які невикористані нотатки і переконатися, що їх зловмисний контент, як і раніше, активний і завантажується. Отже, коли йде порівняння схеми Fast-Flux Phishing з Rock-phish, то видно, що, маючи цей динамічний пул проксі-ботів,

то таким чином ускладнюється дія такої фішингової мережі, тому що в більшості випадків цей динамічний пул проксі-ботів постійно розвивається і змінюється. Навіть якщо йде ідентифікація одного IP-адресу і спроба знести цю IP-адресу, то на цьому DNS-сервері прописується інша IP-адреса після того, як буде знесено один з цих проксі-ботів.

Є цікавий приклад лавинної мережі. Це мережа, яка використовувалася як платформа доставки для запуску та управління глобальними атаками шкідливого ПЗ, а також компаніями із залучення «грошових мулів». Наприклад, в Німеччині, збитки завдані цією фішинговою мережею Fast-Flux, оцінюються приблизно в шість мільйонів євро в результаті концентрованих кібератак на онлайн-банкінг. Ця мережа фішингу була відключена тільки за рахунок рішучої підтримки прокурорів і слідчих з тридцяти різних країн. Затримано п'ятьох осіб, зроблено обшуки у 37 приміщеннях, вилучено 39 серверів. Таким чином, просто глянувши на цей приклад, можна скласти уявлення про те, скільки зусиль потрібно щоб відключити таку фішингову мережу Fast-Flux.

Крім того слід звернути увагу, що існують різні типи фішингової мережі Fast-Flux. Одна з фішингових мереж Fast-Flux характеризується декількома окремими нотатками, що реєструють та видаляють реєстрацію своїх IP-адрес як частину DNS-адреси для одного доменного імені. Ці реєстратори мають дуже короткий час життя, в середньому п'ять хвилин, і створюють потік IP-адрес, що постійно змінюється, при спробі доступу до певного домену. А ще є подвійні фішингові мережі Fast-Flux. У цьому типі мережі використовуються компоненти методів встановлення зв'язку між системами жертвами та системами, що контролювані зловмисниками. Зомбі комп'ютери, що входять до цієї мережі є частиною ботнета і використовуються як проксі-сервери, які не дозволяють жертві безпосередньо взаємодіяти з серверами на яких розміщено шкідливе ПЗ або фішинговий веб-сайт.

1.3 Фішингові атаки

Якщо говорити про фішингові атаки, то «фішингові атаки» - це атаки, коли фішингові сайти маскуються під законні веб-сайти з метою крадіжки конфіденційної інформації. Дана дія відбувається, коли зловмисники розсилають SMS або електронні листи зі шкідливими URL-адресами, і обманом змушуючи жертв відк-

рити шкідливе посилання, яке спрямовує їх на підроблений сайт, зовнішній вигляд цього сайту дуже схожий на законний сайт. Коли жертв обманним шляхом змушують надати конфіденційну інформацію, зловмисник може використати цю інформацію для запуску атак соціальної інженерії або крадіжки іншої інформації, пов'язаної з обліковим записом жертви. На рис. 1.5 показана фішингова атака, яка імітує сторінку входу в Google Docs. Зловмисник імітує кожну деталь цього сайту, наприклад, логотип, і навіть зовнішні посилання, через що його важко візуально відрізнити від цього сайту.

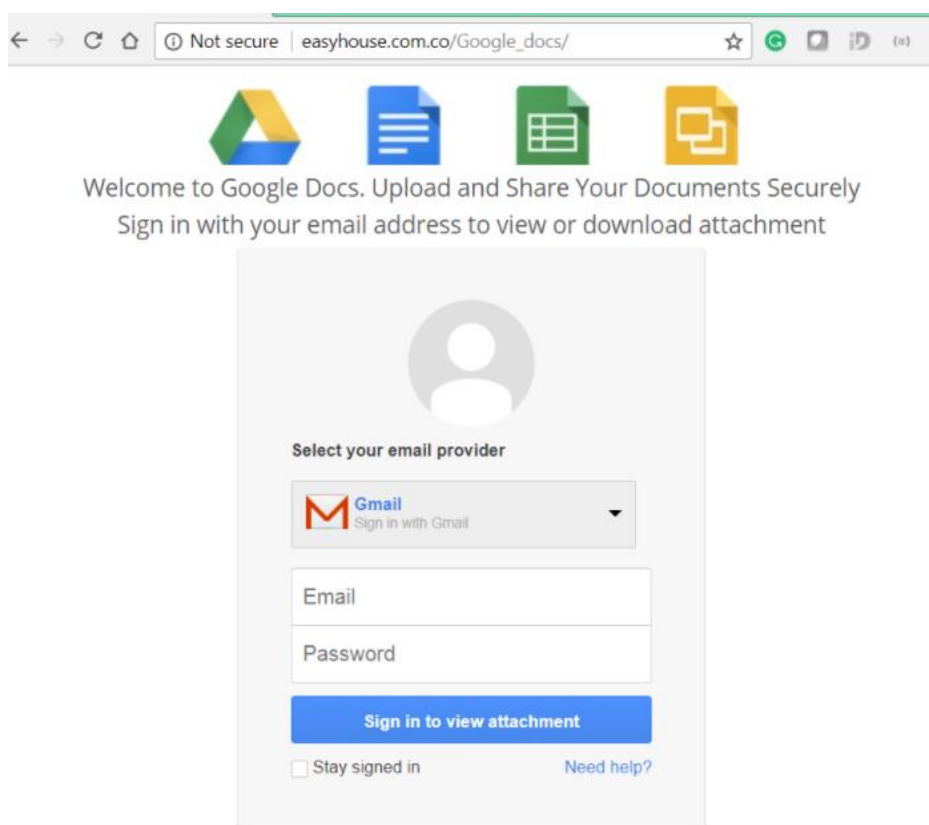


Рисунок 1.5 - Імітація сторінки Google Docs

Дивлячись на рис. 1.5, то звичайному користувачеві не зрозуміло, як саме можна визначити легітимий сайт, і як зрозуміти де фішинговий сайт. Фішинговий сайт можна відрізнити за такими признаками, які наведені нижче.

1) По домену - з рис. 1.5, за візуальною інтерпретацією видно, що це якийсь сервіс, і якщо це сервіс, то передбачається, що він буде знаходитися на google.com домені.

2) По захищеному протоколу передачі гіпертексту, (HTTPS) - з рис. 1.4 видно що в даному випадку замість https позначено not secure, тобто немає серти-

фікатів. У легітимному сайті завжди використовується сертифікат, за яким можна переконатися в автентичності віддаленого сервера, з яким наш браузер з'єднується. Тобто, через ланцюжок сертифікатів, які підписані довіреним сертифікаційним органом, можна переконатися, що цей сертифікат був виданий гуглом саме цьому домену, і цей домен належить гуглу.

3) По вигляду інтерфейсу - деякі сайти оновлюють свій інтерфейс, і якщо виконується вхід на фішинговий сайт, і видно там старий інтерфейс, то це означає, що даний сайт дійсно фішинговий. Тобто зловмисники іноді можуть використовувати застарілі форми.

Існує кілька типів фішингових атак, коли у зловмисників може бути або велика кількість цілей, або лише кілька чітко вибраних цілей, коли це відомо як фішингова атака. Але перш ніж зловмисники зможуть ініціювати будь-які фішингові кампанії, спочатку вони повинні налаштувати інфраструктуру для розміщення та доставки своїх корисних даних або фішингових сайтів. На рис.1.6. показано високорівневі етапи типової фішингової атаки.

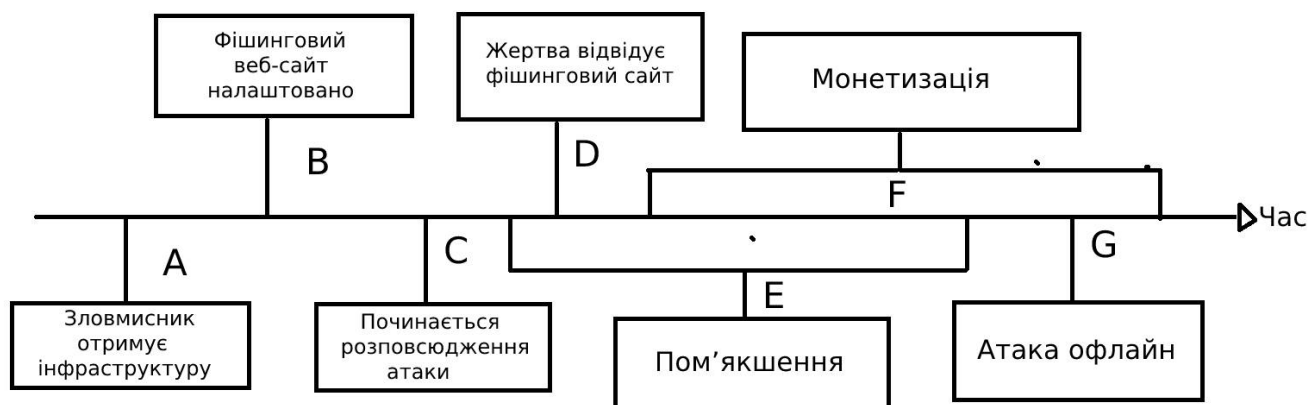


Рисунок 1.6 - Високорівневі етапи фішингової атаки

З рис.1.6 видно, що зловмисник спочатку отримує інфраструктуру (A) і налаштовує фішинговий веб-сайт, який зазвичай містить фішинговий набір, розміщений у цій інфраструктурі (B), який згодом потім використовується або для збору облікових даних, або для надання шкідливого програмного забезпечення для завантаження. Як тільки веб-сайт запрацює, зловмисники почнуть розповсюджувати його серед своїх жертв (C), зазвичай це робиться електронною поштою, після чого жертви починають відвідувати сайт (D).

Залежно від можливостей організації виявляти, що фішингова кампанія продовжується та націлена на їх співробітників, наприклад, за допомогою звітів користувачів, організація може пом'якшувати атаку (E). В оптимальному сценарії дії пом'якшення, будуть зроблені наслідки до (D), коли користувачі не відвідали ще фішинговий сайт, що запобіжить весь майбутній трафік жертв. Але, якщо нічого не вийде, то в такому разі зловмисники отримують тимчасові рамки, щоб почати монетизувати (F) свою атаку за допомогою вкрадених даних або закріпитися в мережі. Фішинговий сайт може вийти з ладу або через видалення, або через дії зловмисників (G). Але в такому разі, як тільки зловмисники отримують дані, що впливають на організацію, або це облікові дані або доступ до мережі, монетизація продовжується, навіть якщо вихідну інфраструктуру вимкнено.

Якщо ж розглядати життєвий цикл фішингової атаки, то загалом життєвий цикл фішингової атаки складається з трьох періодів, періоду підготовки, активного періоду та неактивного періоду. Фішингова атака починається з періоду підготовки і стає активною, коли фішингове посилання публікується в соціальних мережах або відправляється електронною поштою. Як тільки атака виявлена, або заблокована, вона перетворюється на неактивний період. Однак це не означає, що атаку припинено. Зловмисник може перезапустити атаку після внесення ряду модифікацій. Тепер розглянемо всі можливі атаки фішингу які можуть бути. Для початку розглянемо атаку "Фішинг електронної пошти".

Даний фішинг часто називають "обманний фішинг", така атака фішинга є однією з найвідоміших типів атак. Спершу зловмисники відправляють електронні листи користувачам, при цьому видають себе за відомий бренд, а також використовують тактику соціальної інженерії для створення підвищеного відчуття безпосередності, після цього вимагають від користувача щоб він завантажив ресурс або ж клацнув на посилання.

Зазвичай такі посилання ведуть на шкідливі веб-сайти, дані веб-сайти крадуть облікові дані, або встановлюють шкідливий код (шкідливе ПЗ) на пристрій користувача. Що стосується файлів, то в файлах, що завантажуються, зазвичай такі файли у форматі PDF, зберігається шкідливий вміст, який встановлює шкідливе ПЗ, відразу ж як тільки користувач відкриває документ. Розглянемо ознаки, за якими можна визначити фішинг електронною поштою.

- 1) Справжня інформація - в даному випадку можна знайти контактну інформацію, або іншу достовірну інформацію про фейкову організацію. Після цьо-

го, можна знайти адресу електронної пошти відправника, який з неправильним доменом, або має орфографічні помилки.

2) Шкідливий та безпечний код - необхідно пам'ятати про код, який намагається обдурити Exchange Online Protection (EOP). Наприклад, посилання з орфографічними помилками, або завантаження.

3) Скорочені посилання - не рекомендується натискати на якісь скорочені посилання, тому що такі посилання використовуються для обману шлюзів захищеної електронної пошти.

4) Фейковий логотип бренду - необхідно подивитися повідомлення на наявність у даного повідомлення будь-яких логотипів, які виглядають справжніми, тому що дані логотипи можуть мати підроблені шкідливі атрибути HTML.

5) Мало тексту - необхідно ігнорувати електронні листи, в яких є тільки зображення і при цьому дуже мало тексту, тому що в таких випадках за зображенням може ховатися шкідливий код.

Далі йде HTTPS phishing. Досить часто, безпечний протокол передачі гіпертексту, (HTTPS) вважається безпечним посиланням для натискання, тому що він використовує шифрування для підвищення безпеки. Багато легальних організацій використовують HTTPS замість протоколу передачі гіпертексту (HTTP), оскільки такий спосіб може встановити легітимність. Але, зловмисники можуть, і використовують зараз HTTPS у посиланнях, які вони поміщають у фішингові електронні листи. Хоча цей спосіб є частиною фішингової атаки електронною поштою, це трохи модернізований підхід. Розглянемо деякі ознаки, за якими можна визначити фішинг із використанням безпечного протоколу передачі гіпертексту (HTTPS).

1) Гіпертекст - це інтерактивні посилання, які вбудовані в текст, щоб приховати реальну URL-адресу.

2) Скорочення посилання - необхідно переконатися, що це посилання має правильний формат з довгим хвостом, а також містить всі частини URL-адреси.

Розглянемо далі цільовий фішинг (Spear phishing). Цільовий фішинг може використовувати електронну пошту, але в цьому випадку він потребує більш цілеспрямованого підходу. Для збору інформації з опублікованих або загальнодоступних джерел, таких як веб-сайт компанії або соціальні мережі, зловмисники починають з використання аналітики з відкритих джерел (OSINT). Після цього вони націлюються на якихось конкретних осіб в організації, використовуючи при

цьому справжні імена, посадові обов'язки або робочі номери телефонів, щоб одержувач думав, що цей лист надісланий всередині організації іншою людиною. В результаті, тому що користувач який отримав лист думає що це внутрішній запит, він робить ту дію, яка вказана в даному електронному листі. Розглянемо ознаки цільового фішингу.

1) Документи, які захищені паролем - документи, яким необхідно ідентифікатор користувача та пароль для входу в систему, можуть бути спробою вкрасти облікові дані.

2) Запит нестандартного виду - необхідно звертати увагу на внутрішні запити, які надходять від користувачів з інших відділів, або які з погляду посадових обов'язків виглядають нестандартними, або ж незвичайними.

3) Посилання на спільні диски - необхідно бути дуже обережним з посиланнями на документи, які зберігаються на спільних дисках, наприклад таких як "O365", і "Dropbox", тому що дані посилання можуть перенаправити на фейковий шкідливий веб-сайт.

Розглянемо тепер «китобійний» метод атаки, чи шахрайство з генеральним директором. Даний метод атаки є типом корпоративного фішингу, і використовує метод OSINT. Ця атака вже обговорювалася вище, у розділі 1.1. У цьому випадку зловмисники використовують соціальні мережі або корпоративний веб-сайт для того, щоб знайти ім'я генерального директора організації або іншого члена вищого керівництва. Після цього вони видають себе за цю людину, використовуючи аналогічну адресу електронної пошти. Лист поштою може містити запит на грошовий переказ або запит на перегляд документа одержувачем. Ознаки цієї атаки є такими, які наведено нижче.

1) Дивний запит - якщо старший член керівництва ніколи раніше не вступав у контакт, то потрібно бути дуже обережним при здійсненні запитаних дій.

2) Електронна пошта одержувача - безліч людей використовують поштові програми, поштові програми, які пов'язують усі їх адреси електронної пошти, необхідно впевнитися, що кожен запит який виглядає нормально, відправляється не на особисту, а на робочу адресу електронної пошти.

Далі розглянемо атаку "Вішинг". Ця атака відбувається коли зловмисник починає дзвонити за номером телефону і створює підвищене почуття невідкладності, це почуття змушує людину діяти на шкоду своїм інтересам, тобто цю атаку

можна назвати голосовим фішингом. Дані дзвінки зазвичай бувають під час важких або стресових ситуацій. Наприклад, безліч людей отримують фейкові телефонні дзвінки від тих людей, які говорять, що вони податкова служба (IRS), і вказують на те, що вони хочуть провести аудит, і у випадку з цим вони вимагають у людини номер соціального страхування. У цьому випадку дзвінок підвищує паніку у людини, і ця людина може бути легко обдурена таким чином, а зловмисник отримує особисту інформацію. Ознаки такої атаки можна виділити такі, які наведено нижче.

1) Дія яку необхідно зробити - необхідно розуміти, що дзвінок може вимагати особисту інформацію, яка може здатися дуже дивною для типу абонента, що викликає.

2) Час дзвінка - необхідно також перевіряти і час, оскільки час дзвінка може збігатися з сезоном або подією, що може викликати стрес або переживання.

3) Номер людини який дзвонить - необхідно перевіряти номер, даний номер може бути з незвичайного розташування або заблокований.

Розглянемо атаку "Смішинг". Варто відзначити, що злоумисники часто застосовують однакову тактику до різних типів технологій. Ця атака це відправлення текстів з проханням до людини виконати якусь дію. Ця атака є наступним етапом, або ж еволюцією "Вішингу". Досить часто текст містить посилання, натискаючи на яке, шкідливе ПЗ встановлюється на пристрій користувача. Ознаки "Смішингу" можуть бути такими які наведені нижче.

1) Зміна статусу доставки - необхідно завжди шукати електронні листи або переходити безпосередньо на веб-сайт служби доставки, щоб перевірити статус, оскільки текст із проханням, вжити заходів для зміни доставки може містити посилання.

2) Невірний код міста - необхідно також перевіряти код міста, а також порівнювати його зі своїм списком контактів, перед тим як відповідати на текст або робити ті дії, які вказані в листі.

Також розглянемо рибалковий фішинг (Angler phishing). Соціальні мережі стають ще одним популярним місцем для атак фішингу, оскільки зловмисники переміщуються між векторами атак. Так само як і "Смішинг", і "Вішинг", рибальський фішинг, це той випадок, коли зловмисник використовує у додатку для соціальних мереж повідомлення, або функції прямого обміну повідомленнями, щоб

змусити користувача виконати дію. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Посилання на веб-сайти - ні в якому разі не натискати на посилання, яке знаходиться в прямому повідомленні, навіть якщо це посилання виглядає цілком легальним. Можна натиснути в тому випадку, якщо користувач, який відправив це, постійно ділиться таким чином якимись цікавими посиланнями.

2) Дивні спрямовані повідомлення - необхідно стежити за спрямованими повідомленнями від тих користувачів, які досить рідко використовують цю функцію, тому що обліковий запис може бути відтворено, або підроблено.

3) Повідомлення - необхідно дуже обережно ставитися до повідомлень, які вказують на додавання в публікацію, тому що дані повідомлення можуть мати посилання, які ведуть на шкідливі веб-сайти.

Розглянемо атаку фішингу "Фармінг". Ця атака є більш технічною, і таку атаку дуже часто важко виявити. Спочатку зловмисники беруть у контроль сервер доменних імен (DNS), сервер, який переводить URL-адреси з природної мови до IP-адреси. Після того, як користувач вводить адресу веб-сайту, DNS-сервер перенаправляє користувача на IP-адресу шкідливого веб-сайту, яка може бути схожою на легальну. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Невідповідності веб-сайтів - необхідно також приділяти увагу на всі можливі невідповідності які вказують на фейковий веб-сайт, а також приділяти увагу орфографічним помилкам, дивним шрифтам, або невідповідності кольорів.

2) Небезпечний веб-сайт - потрібно шукати веб-сайт із протоколом HTTP, а не HTTPS.

Розглянемо атаку "Спливаючий фішинг". Багато людей використовують блокувальники спливаючих вікон, наприклад "AdBlock", фішинг спливаючих вікон, як і раніше, є ризиком. Шкідливий код може бути розміщений зловмисниками у невеликих вікнах повідомлень, які називаються спливаючими вікнами, вони з'являються, коли люди відвідують веб-сайти. Набагато нова версія спливаючого фішингу застосовує функцію повідомлень веб-браузера. Наприклад, випадок, коли користувач відвідує веб-сайт, браузер пропонує користувачеві "www.buysomestuff.com хоче показувати повідомлення". У цей момент, коли користувач натискає кнопку "Дозволити", спливаюче вікно встановлює шкідливий код. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Повноекранний режим - зазвичай шкідливі спливаючі вікна можуть перевести браузер на повноекранний режим, тобто будь-яка автоматична зміна розміру екрана може бути індикатором.

2) Невідповідності - необхідно перевіряти наявність орфографічних помилок чи нестандартних колірних схем.

Розглянемо атаку "фішинг клонування". Це цільова фішингова атака електронною поштою, дана атака, фішинг-клон, використовує служби, які хтось раніше використовував для запуску поганої дії. Зловмисники знають безліч бізнес-додатків, які змушують людей переходити за посиланнями в рамках їх повсякденної діяльності. Зловмисники досить часто проводять дослідження, щоб дізнатися, які типи послуг регулярно використовує організація, після цього надсилають цільові електронні листи, дані листи походять від цих служб. Наприклад, багато організацій використовують "DocuSign" для відправлення та отримання електронних контрактів, і в даному випадку зловмисники мають можливість створювати фейкові електронні листи для цієї служби. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Особиста інформація - необхідно стежити за електронними листами, які можуть запитувати особисту інформацію, яку постачальник послуг ніколи не запитує.

2) Несприятливий або дивний час - необхідно бути дуже обережним з кожним раптовим електронним листом від постачальника послуг, у тому випадку навіть якщо це частина звичайної роботи, яка відбувається щодня.

Розглянемо фішинг "Злий близнюк". Така фішингова атака використовує фальшиву точку доступу Wi-Fi, часто видаючи дану точку доступу за справжню, яка може перехоплювати дані під час передачі. Якщо ж використовувати цю фейкову точку доступу, то зловмисник може брати участь в атаках підслуховування, або ж "людина посередині". Ця дія дозволяє зловмисникам збирати облікові дані для входу або конфіденційну інформацію, що надсилається через з'єднання. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Необхідний вхід в систему - необхідно розуміти, що будь-яка точка доступу, яка зазвичай не вимагає облікових даних для входу, але в даний момент запитує ці дані, то в такому випадку дана точка доступу викликає сумніви.

2) Небезпечне попередження - необхідно бути дуже обережним з будь-якою точкою доступу, оскільки ця точка доступу може викликати небезпечне попередження на пристрої.

І остання фішинг атака, яка буде розглянута в даному розділі, це "фішинг з водяною дірою (Watering hole attack)". Ця фішингова атака є досить витончена, починається ця атака з того, що зловмисник проводить дослідження веб-сайтів, які часто відвідують користувачі або співробітники компанії, після чого заражають IP-адресу завантаженнями або шкідливим кодом. Веб-сайти в даному випадку можуть бути такими, що надають галузеві новини, або веб-сайти сторонніх постачальників. Коли користувач відвідує веб-сайт, він завантажує шкідливий код. Ознаки цього фішингу можуть бути такими, які наведені нижче.

1) Моніторинг правил брандмауера - необхідно переконатися, що правила брандмауера постійно оновлюються, а також відстежуються для того, щоб запобігти вхідному трафіку зі зламаного веб-сайту.

2) Попередження браузера - необхідно звертати увагу на те, що браузер вказує, наприклад браузер може вказати що на сайті може бути шкідливий код, і що не потрібно переходити на цей сайт, навіть якщо він зазвичай використовується.

1.4 Методи захисту фішингових атак

У минулому розділі були розглянуті одні з основних фішингових атак, у цьому розділі буде детальніше розглянуто якими методами можна забезпечити захист від даних фішингових атак, які були розглянуті. Захист від фішингу має на увазі комплексний набір інструментів та методів, які допомагають заздалегідь виявляти та нейтралізувати фішингові атаки. Це також може включати широке навчання користувачів, призначене для поширення інформації про фішинг, встановлення спеціалізованих антифішингових рішень, інструментів і програм, а також впровадження ряду інших заходів безпеки, спрямованих на запобіжний захист від фішингу, а також надання методів пом'якшення наслідків атак, яким вдається порушити безпеку.

В основному фішингові атаки найбільше поширюються на поштові сервіси, тому щоб убезпечити себе від даної атаки можна навчати користувачів розпізнавати і повідомляти про фішинг, наприклад, про підозрілі адреси електронної по-

шти, звичайне привітання, незвичайне електронне повідомлення або миттєве повідомлення граматичні помилки, посилення за межами організації, спробу створити паніку, щоб як спонукати до поспішних дій. У цьому випадку можна також використовувати сегментування мереж для обмеження доступу до висококонфіденційних даних, це може ускладнити проникнення фішингових атак. У цьому випадку також необхідний аудит середовища кібербезпеки для оцінки вразливостей, виявлення нових загроз та розробки захисних стратегій. Найбільш ефективним методом захисту може бути встановлення технології, яка буде сканувати всю вхідну електронну пошту в режимі реального часу і блокувати користувачів від переходу за посиланнями на підозрілі веб-сайти, пісочниці та сканування всіх вкладень на наявність потенційних загроз безпеці, та виявлення підозрілих URL-адрес ще до того моменту, коли вони будуть доставлені користувачам.

Поштові загрози продовжують розвиватися, тому потрібні все більш потужні технології для того, щоб зупинити фішингові листи та захистити завдання збитків організаціям цільовими фішинговими атаками. Багато компаній намагаються зупинити фішингові електронні листи, навчаючи співробітників способам виявлення підозрілої електронної пошти, проте майже чверть фішингових електронних листів відкривається навіть після того, як співробітників навчили найпоширенішими методами фішингу. Тобто, щоб дійсно зупинити фішингові електронні листи, ефективним методом захисту може бути технологія, яка може унеможливити людську помилку, і автоматично запобігти попаданню фішингових атак у поштові сервіси користувачів.

Ще одним методом захисту від даних атак може бути вибір багаторівневого захисту для запобігання атакам фішингу, оскільки шахрайство з фішинговою електронною поштою продовжує успішно зламувати засоби захисту, і тому все більше компаній застосовують багаторівневий підхід до стратегії безпеки для запобігання атак фішингу. Також, для захисту від фішингових атак електронної пошти можна використовувати Заходи автентифікації DNS, що використовують протоколи DMARC, SPF і DKIM для виявлення та блокування підозрілих повідомлень, а також протоколи двофакторної автентифікації, вони не дозволяють зловмисникам використовувати вкрадену інформацію для доступу до облікових записів.

Тепер розглянемо захист від атак типу "HTTPS phishing". У цьому випадку, навіть якщо веб-сайт має значок замка або HTTPS в адресному рядку, йому все одно не можна довіряти. Також необхідно перевіряти правильність URL-адреси

веб-сайту. Наприклад, потрібно шукати помилки або неправильні домени, такі як домен .net, який зазвичай є доменом .com. Найкраще вводити URL-адресу веб-сайту, яку користувач хоче відвідати, безпосередньо в браузері, а не переходити за посиланням, отриманим в електронному листі. Для захисту від такої атаки, можна використовувати установки таких інструментів, як менеджер паролів, або програмне забезпечення для забезпечення безпеки. Дані інструменти іноді включають функції, які можуть попередити користувача, коли URL-адреса не відповідає законному веб-сайту, або можуть завадити користувачеві відкрити шахрайський сайт. У більшості випадків від даних атак дуже сильно допомагає уважність користувача, тому якщо користувач отримує підозрілий електронний лист з посиланням, необхідно запропонувати йому зателефонувати або написати електронний лист безпосередньо цій людині і запитати, чи відправив він його. Також у цьому випадку може допомогти уникнення використання підстановних сертифікатів у виробничих системах, що збільшує ризик та поверхню атаки, якщо сервер або сертифікат скомпрометовані. Крім цього, для захисту, в компанії може бути розділ навчання з кібербезпеки, присвячений атакам фішингу, необхідно перевірити, що даний розділ безпосередньо охоплює фішинг, пов'язаний з HTTPS. Необхідно розповісти співробітникам про кіберзагрози HTTPS, щоб краще захистити компанію від фішингових схем із використанням SSL-сертифікатів.

Якщо ж користувач не впевнений чи законний веб-сайт чи ні, то він може використовувати швидкий перегляд даних whois для домену. Такий спосіб може допомогти прийняти рішення, чи законний цей сайт чи ні. На даний момент, в сучасні браузери вже вбудований захист від шахрайських сайтів, і вони непогано справляються зі своїм завданням. Більшість антивірусних продуктів та пакетів безпеки додають власний захист від фішингу. Також, для захисту від даної атаки можна використовувати і менеджер паролів, цей спосіб допомагає захистити користувача від атаки. З більшістю таких продуктів користувач зможе відвідати безпечний сайт та увійти в систему одним клацанням миші.

Тепер розглянемо способи захисту від цільового фішингу. До таких методів можна віднести захист електронної пошти, який відстежує атаки з використанням уособлення. Компанії повинні впровадити передове рішення для захисту електронної пошти, таке як безпечний поштовий шлюз, який блокує шкідливе ПЗ, спам та переважну більшість атак, які пов'язані з електронною поштою. Але в такому випадку, оскільки цільові атаки фішингу зазвичай не містять шкідливого ПЗ, кра-

ще додати рівень безпеки, такий як DMARC (перевірка справжності повідомлень на основі домену), який відстежує атаки з використанням уособлення електронної пошти шляхом перевірки справжності електронних листів у базі даних відправників. Також як захист в даному випадку можна використовувати шифрування конфіденційної інформації. Якщо компанія зашифрує свої дані, то злоумисник не зможе отримати доступу до будь-якої цінної чи конфіденційної інформації. У цьому випадку можна використовувати також багатофакторну автентифікацію, вимога додаткового кроку входу в систему крім простого введення пароля ускладнює отримання злоумиснику доступу до даних і систем. Крім того, допоможе також регулярне оновлення системи безпеки. Оновлення систем безпеки з використанням останніх виправлень допомагає блокувати будь-які відомі області, які можуть бути зламані. Для захисту можна проводити навчання персоналу з питань безпеки. Ще одним методом захисту від цільового фішингу може бути провідна служба безпеки електронної пошти, яка запобігає проникненню фішингових листів у корпоративну поштову інфраструктуру. Для того, щоб покращити захист від цільового фішингу, багато компаній намагаються інформувати користувачів про небезпеку цих електронних листів, але майже чверть усіх повідомлень, заснованих на фішингу, все ще відкриваються. А враховуючи, що на сьогоднішній день більше 90% всіх атак починаються з фішингу або цільового фішингу електронною поштою, стає ясно, що компаніям потрібен чудовий захист від цільового фішингу для захисту користувачів і самої компанії.

Тепер розглянемо способи захисту від китобійного фішингу. В даному випадку можна використовувати як захист створення протоколу щодо запобігання китобійного фішингу. Деякі відмінні приклади цього включають перевірку запитів конфіденційної інформації іншими каналами, таким як телефонний дзвінок. Також можна використовувати програмне забезпечення DLP. Програмне забезпечення для запобігання втраті даних (DLP) може блокувати порушення встановлених протоколів. Він також може позначати електронні листи на основі імені та віку домену, схожості відображуваного імені з відомими контактами та підозрілих ключових слів, таких як "переведення з банку". Проти такої атаки можна використовувати хороші програми захисту від спаму і шкідливого ПЗ, які можуть зупинити деякі електронні листи з китобійними атаками на поштовому шлюзі. Крім цього, можна використовувати служби автентифікації DNS, які використовують протоколи DMARC, DKIM і SPF, вони можуть визначити, чи є електронний лист,

надісланий з певного домену, законним або фейковим. Ще одним методом захисту від цієї атаки може бути програмне забезпечення для захисту від уособлення, це ПЗ може блокувати китобійну атаку, визначаючи засновані на соціальній інженерії методи, які є загальними для китобійних електронних листів. Для запобігання цій атаці можна також використовувати технології захисту даних, необхідно впровадити методи роботи з даними, які ускладнюють заподіяння шкоди однією людиною. Використовуйте керування доступом, щоб доступ користувача root був обмежений до необхідного рівня. Крім того, необхідно перевірити, що всі дозволи користувача підходять та необхідні для кожної функції роботи. Впровадити та підтримувати звичайний набір інструментів IT-безпеки, брандмауери, програмне забезпечення для виявлення вторгнень та утиліти сканування шкідливих програм.

Далі розглянемо методи захисту, які можуть запобігти атаці типу "Вішинг". Для того щоб захиститися від цієї атаки більшою мірою важлива увага користувача, необхідно остерігатися телефонного дзвінка від імені державної установи або банку з проханням надати конфіденційну інформацію або гроші. Не потрібно довіряти ідентифікатору абонента, тому що його дуже легко підробити. Необхідно дуже обережно ставитися до довіри, оскільки не можна довіряти тому, хто створює терміновість і переконує користувача передзвонити прямо зараз. Необхідно в такому разі провести маленьке дослідження та зателефонувати за реальними номерами, щоб переконатися, що банк дійсно дзвонить користувачеві. Тобто необхідно перевіряти особу тих, хто запитує інформацію особисто або по телефону, перш ніж розкривати будь-яку інформацію. Велику частину захисту в даному випадку надає навчання співробітників безпеки та їх уважність. У випадку "Вішингу", можна також не брати слухавку.

Далі розглянемо методи захисту для запобігання атаці "Смішинг". Як і у випадку з атакою "Вішинг", для захисту від цієї атаки можна використовувати навчання співробітників та їхню уважність. Крім цього, як захист можна використовувати багатофакторну автентифікацію (MFA). Відкритий пароль може бути марним для атакуючого, якщо зламаний обліковий запис вимагає другого "ключа" для перевірки. Найбільш поширеним варіантом MFA є двофакторна автентифікація (2FA), у якій часто використовується код підтвердження текстового повідомлення. Також є сильні варіанти, включаючи використання спеціальної програми для перевірки (наприклад, Google Authenticator). Можна також використовувати програму захисту від шкідливих програм, такі програми можуть захищати як від

шкідливих програм, так і від фішингових SMS-посилань. Також можна використовувати VPN на своєму мобільному пристрої. Це та річ, яку часто не беруть до уваги смішинг-атак, це збір даних про місцезнаходження. Зловмисники все частіше використовують дані про місцезнаходження для більш точного визначення цілей. Зловмисники можуть використовувати ці дані для надсилання користувачеві повідомлень, які здаються досить локальними. Якщо повідомлення здається більш особистим, воно з більшою ймовірністю викликає реакцію користувача. В даному випадку, VPN допомагає підробити місце розташування, створивши враження, що користувач знаходиться десь в іншому місці. Якщо користувач отримує змішинг повідомлення, засноване на підробленому місці, його набагато легше розпізнати як фейкове.

Далі розглянемо методи захисту для запобігання рибальському фішингу. У цьому випадку можна використовувати підтвердження облікового запису компанії. Перш ніж відповідати будь-кому, хто зв'язався з користувачем, необхідно підтвердити, що обліковий запис є законним. Більшість облікових записів соціальних мереж на таких платформах, як Instagram або Twitter, мають синю галочку поруч з ім'ям облікового запису, що підтверджує законність. Необхідно перевіряти профіль на орфографічні помилки та кількість передплатників. Обліковий запис служби підтримки шановної установи має мати багато передплатників. Крім того, необхідно перевіряти історію профілю, щоб переконатися, що обліковий запис успішно допомагав клієнтам раніше. Технологічні заходи безпеки, такі як брандмауери, шифрування, антивірус, спам-фільтри та сувора автентифікація, не завжди запобігають шахрайству з використанням соціальної інженерії. Скільки б технологій безпеки користувач не застосовував, йому ніколи не вдасться позбавитися найслабшої ланки, людського фактора, тому в даному випадку, уважність користувача та навчання персоналу є хорошим методом захисту від даної атаки.

Далі розглянемо методи захисту для запобігання "фармінгу". Для захисту в цьому випадку можна використовувати перехід лише за безпечними посиланнями, що починаються з HTTPS. Також, уважність відіграє велику роль при захисті від даної атаки, необхідно уникати посилань і вкладень від невідомих відправників, оскільки користувач не може захистити себе від отруєння DNS, необхідно остерігатися шкідливих програм, які дозволяють використовувати фармінг. Також як метод захисту можна використовувати і вибір іншого інтернет-провайдера (ISP). Нові інтернет-провайдери можуть заманити користувача вигідними пропо-

зиціями та величезними швидкостями, але конфіденційність користувача може бути платою за ці переваги. Вибираючи провайдера, необхідно вибирати хорошого провайдера, багато з них за замовчуванням відфільтровують будь-які підозрілі перенаправлення. Це свідчить про те, що користувач ніколи не потрапить на сайт фармінгу. Як метод захисту можна використовувати двофакторну перевірку. На сьогоднішній день багато платформ пропонують двоетапну перевірку для захисту своїх користувачів. Наприклад, соціальні мережі. Якщо користувач використовує двоетапну автентифікацію для своїх облікових записів у соціальних мережах, то зловмиснику буде важче зламати облікові записи. У тому випадку навіть якщо зловмисник отримав облікові дані для входу в систему за допомогою фармінгу, зловмисник не зможе отримати доступ до облікового запису. Можна також використовувати надійний DNS-сервер, що дозволить захистити користувача від отруєння DNS. Хорошим захистом буде надійний пароль для домашнього інтернету. Цей метод дозволить захистити свою домашню мережу від локального отруєння DNS.

Тепер розглянемо методи захисту, які можуть допомогти запобігти атаці типу "Спливаючий фішинг". Для захисту в даному випадку можна також використовувати увагу користувача. Крім цього, захист від даної атаки може надати налаштування браузера, необхідно збільшити налаштування безпеки браузера для того, щоб дана атака не заважала користувачеві. У цьому випадку хорошим захистом може бути антивірусне програмне забезпечення або комплексне рішення для забезпечення безпеки в Інтернеті. Є випадки, що навіть якщо на комп'ютері встановлений антивірус, користувач, швидше за все, зіткнеться з шахрайськими спливаючими повідомленнями на деяких веб-сайтах. Зазвичай це означає, що не заражений комп'ютер, а веб-сайт, який користувач відвідав, тому уважність користувача в даному випадку необхідна.

Тепер розглянемо методи захисту від клонування атаки. Для захисту можна використовувати HTTPS, оскільки цей протокол може забезпечити легітимість сайту та безпеку. Також можна використовувати навчання своїх співробітників безпеки. Методом захисту також може бути використання інструменту захисту, щоб убезпечити свою компанію. Інструмент допоможе занести до чорного списку IP-адреси всіх, хто намагається скопіювати бренд. Крім того, цей інструмент дозволить видалити дані клонованих сайтів. Навчання з кібербезпеки для співробітників, це найкращий спосіб запобігти клонуванню фішингових повідомлень від

іншої жертви в компанії, оскільки кінцевий користувач є останньою лінією оборони перед фішингом. Також можна використовувати брандмауери або рішення для управління погрозами, це метод який буде працювати у фоновому режимі для пошуку несхожих URL-адрес та розбіжностей у відправниках, які можуть вказувати на фішинг-клонування.

Розглянемо методи захисту від атаки "Злий близнюк". Для захисту можна використовувати системи запобігання вторгненням у бездротову мережу (WIPS), щоб виявляти наявність зловмисної точки доступу-близнюка та запобігати підключенню до них будь-яких керованих корпоративних клієнтів. На жаль, для користувачів Wi-Fi практично неможливо виявити зловмисну точку доступу-близнюка, тому що назва безпроводної точки доступу (SSID) здається законним, а зловмисники зазвичай надають інтернет-послуги. Найкращий метод захисту в даному випадку завжди використовувати VPN для інкапсуляції сеансу Wi-Fi на іншому рівні безпеки. Для компаній також можна використовувати свої точки доступу за допомогою персонального ключа безпеки (PSI). Як метод захисту можна використовувати зміни налаштувань входу, щоб вимагати додаткові дані автентифікації. У цьому випадку можна вибрати двоетапну автентифікацію, зловмисникам буде складніше обійти систему безпеки та отримати доступ до облікових записів.

Остання атака - це атака типу "фішинг з водною дірою". Розглянемо методи захисту при цій атаці. Одним із методів захисту може бути оновлення програмного забезпечення. Атаки такого типу часто використовують проломи та вразливості для проникнення на комп'ютер користувача, тому регулярно оновлюючи програмне забезпечення та браузер, можна значно знизити ризик атаки. Також, другим методом захисту можуть бути перевірки безпеки за допомогою інструментів безпеки мережі, щоб спробувати виявити атаки даного типу. Наприклад, системи запобігання вторгненням можуть виявляти підозрілі та шкідливі події в мережі. Програмне забезпечення для управління смугою пропускання дозволить спостерігати за поведінкою користувачів і виявляти аномалії, які можуть вказувати на атаку, такі як передача великих обсягів інформації або велика кількість завантажень. Крім цього, захистом також може бути приховування своєї діяльності в Інтернеті. Зловмисники можуть створювати ефективніші атаки такого типу, якщо вони компроментують веб-сайти, які користувач часто відвідує. Таким чином, користувач повинен приховувати свої дії в інтернеті за допомогою VPN та функції приватно-

го перегляду браузера. Можна також використовувати оцінку вразливостей та тестування на проникнення (VAPT), це може допомогти переконатися, що засоби безпеки користувача забезпечують задовільний захист від загроз, пов'язаних з програмами та браузерами, таких як атаки даного типу. Як ще один метод захисту можна використовувати багатофакторну автентифікацію (MFA) у всіх застосовних кінцевих точках мереж компанії. Цей захист зменшить вплив атак даного типу у випадку, якщо зловмисникам вдасться вкрати облікові дані користувачів.

Зібравши всі описані фішингові атаки, і методи захисту, можна підбити підсумки, що найбільше атакам піддаються поштові сервіси, і сайти, тому в даному випадку необхідно розглянути способи виявлення фішингових повідомлень і сайтів, дані способи не тільки забезпечують безпеку, але і також можуть покращити швидкість роботи при виявленні, оскільки можна застосувати автоматизацію при виявленні фішингових повідомлень та веб-сайтів.

1.5 Аналіз робіт і патентний пошук в області фішинга

Патентна інформація представляє собою технічну та правову інформацію, що міститься в офіційних періодичних публікаціях відомств інтелектуальної власності, що є у всіх країнах, де законодавством передбачено правову охорону об'єктів інтелектуальної власності [4]. На даний момент, загальна кількість патентних документів у світі досягає 40 мільйонів. Як результат, патентна інформація є єдиною всеосяжною добіркою систематизованої технічної інформації. Крім цього, патентні документи можуть класифікуватися за технічними галузями відповідно до єдиної, детально розробленої Міжнародної патентної класифікації, що значно полегшує пошук потрібної інформації. Крім цього, патентна інформація допоможе відстежувати передові досягнення у будь-якій сфері, а технічна інформація, що міститься в патентних документах, дозволить уникнути зайвих витрат на дублювання досліджень, виявити та оцінити технологію для ліцензування та передачі технології, знайти альтернативи технології, бути на сучасному рівні у певній галузі, а також знайти Ідеї для подальших інновацій.

І, нарешті, патентна інформація може використовуватися малими та середніми підприємствами для того, щоб уникнути можливих проблем із порушенням прав власників охоронних документів на об'єкти промислової власності, та оцінити патентоспроможність власних винаходів.

Найважливіша частина патентної інформації складається з офіційних публікацій відомств, які відповідальні за правову охорону об'єктів промислової власності, тобто винаходи, корисні моделі, промислові зразки, знаки для товарів та послуг у кожній державі та регіональних та міжнародних організацій з інтелектуальної охорони власності.

Щоб мати можливість проводити патентні дослідження, необхідні відповідні інформаційні ресурси, впорядковані повні фонди патентної документації. Найбільші в Україні інформаційні ресурси, придатні для проведення патентних досліджень, знаходяться у Фонді патентної документації громадського користування (ФГК) філії Українського інституту промислової власності, Українського центру інноватики та патентно-інформаційних послуг (УкрЦІПП).

Для того щоб розуміти які зусилля та дослідження вже були проведені та зроблені, в даному розділі буде проведений патентний пошук, аналіз патентів (способів, методів, технологій) у галузі фішингу, а саме щодо питань виявлення фішингових повідомлень та веб-сайтів. Після цього результати пошуку будуть описані в цьому розділі. Для пошуку патентів у даному випадку будуть використовуватися такі бази даних патентів як "Patentscope", та "Espacenet".

По результатам пошуку складемо складено список найбільш важливих патентів, щоб детальніше розібрати кожен патент. Список найбільш важливих патентів наведено нижче.

1) Виявлення фішингових сайтів - цей патент має номер "US2021099484A1". У цьому патенті передбачені системи та методи виявлення підозрілих фішингових веб-сторінок. Здійснюється це таким чином, клієнтський пристрій захоплює зображення, що стосується веб-сторінки, доступ до якої здійснюється через клієнтський пристрій, і генерує відбиток веб-сторінки на основі застосування хеш-функції до захопленого зображення. Для кожного фішингового відбитка в базі даних фішингових відбитків, що містить відбитки, пов'язані з відомими фішинговими веб-сторінками, клієнтський пристрій визначає міру подібності між згенерованим відбитком і фішинговим відбитком, порівнюючи згенерований відбиток з фішинговим відбитком клієнтський пристрій ідентифікує веб-сторінку як потенційно фішингову.

2) Метод виявлення фішингових веб-сайтів - цей патент має номер "CN104166725A". У даному патенті наведено спосіб виявлення фішингових веб-сайтів, який застосовується в області комп'ютерних мереж. Метод використо-

ується для вирішення проблем, пов'язаних з тим, що виявлення існуючого фішингового веб-сайту не є активним, а звіти видаються помилково, невдало та невчасно. Спосіб включає в себе етапи, на яких встановлюється вектор ознак, відповідний веб-сторінці, що виявляється, який заснований на візуальному контенті. Вектор ознак порівнюється з вектором ознак заздалегідь встановленого набору векторів ознак, далі за результатом порівняння можна з'ясувати, чи належить веб-сторінка, що виявляється, фішинговому веб-сайту чи ні.

3) Метод та спосіб виявлення фішингових веб-сайтів - цей патент має номер "CN103023874A". Даний винахід здійснює спосіб виявлення фішингових веб-сайтів, який включає чотири етапи, на першому етапі відбувається вилучення тексту каскадної таблиці стилів (CSS), відповідно до різних форм існування та об'єднання всіх витягнутих рядків символів каскадних таблиць стилів (CSS) в одному файлі, на другому етапі відбувається аналіз тексту каскадних таблиць стилів (CSS) для автоматичного аналізу файлу, отриманого на першому етапі, з використанням аналізатора каскадних таблиць стилів (CSS), щоб отримати відповідний набір об'єктів правил. На третьому етапі відбувається вилучення елемента порівняння для вибору атрибуції з безліччю характеристик набору правил і повторного інтегрування для полегшення швидкого обчислення подібності, і останній етап це обчислення подібності для обчислення подібності відповідно до ситуації збігу набору елементів порівняння двох веб-сторінок.

4) Спосіб і пристрій знаходження фішингових сайтів - цей патент має номер "WO2013152610A1". У даному патенті є метод і запропоновано пристрій виявлення фішингових сайтів. Метод виявлення фішингових веб-сайтів включає отримання веб-сайту і визначення тенденції до фішингу. Відповідно до доменного імені виявленого веб-сайту отримується посилання на сторінку в середині веб-сайту, і якщо в результаті виявлення стає відомо, що посилання на сторінку всередині веб-сайту містить посилання на вікно входу в систему, то веб-сайт визначається як фішинговий. Спосіб і пристрій для виявлення фішингових веб-сайтів, запропонований в цьому винаході, може забезпечувати активне виявлення фішингових веб-сайтів.

5) Метод і система виявлення фішингових веб-сайтів - цей патент має номер "CN111669353A". Даний винахід розкриває спосіб і систему виявлення фішингових веб-сайтів. Метод виявлення фішингових веб-сайтів у даному винаході включає етапи, на яких виконується вилучення ознак з URL-адрес пакетних веб-

сайтів, виконується неконтрольований кластерний аналіз вилучених даних URL-адрес, потім визначаються оптимальний параметр та оптимальна модель прогнозування за результатом кластерного аналізу. Нова URL-адреса веб-сайту визначається відповідно до оптимальної моделі прогнозування.

6) Багатоетапний метод і система винаходження фішингових веб-сайтів - цей патент має номер "WO2016201938A1". Даний винахід пропонує багатоетапний спосіб і систему виявлення фішингових веб-сайтів та поєднує засоби як швидкої, так і точної фільтрації.

7) Динамічний метод і пристрій знаходження фішингу - цей патент має номер "US20080163369". Даний патент пропонує метод виявлення спроби фішингу з веб-сайту за допомогою комп'ютера. Цей спосіб включає отримання веб-сторінки з веб-сайту, яка включає в себе машиночитаний код для веб-сторінки. Спосіб також включає встановлення гіперпосилань в коді.

8) Методи та системи для виявлення фішингових атак за поштою - цей патент має номер "US08839369". У цьому патенті виявлення фішингових атак електронною поштою ініціюється, коли електронний лист надходить у комп'ютерну систему. Електронна пошта аналізується на наявність ознак, що вказують на атаку електронної пошти фішингу, таких як посилання на зовнішній веб-сайт.

9) Евристичний метод та система виявлення фішингових сайтів - цей патент має номер "CN105653941A". Винахід пропонує евристичний спосіб виявлення фішингового веб-сайту. Спосіб включає етапи захоплення вихідних кодів фрейму домашньої сторінки, що підлягає виявленню веб-сайту, вилучення всіх URL-адрес гіперпосилань, визначення існування повторної URL-адреси гіперпосилання, визначення веб-сайту як безпечного веб-сайту.

10) Метод виявлення фішингових веб-сайтів на основі спільних функцій веб-сайту - цей патент має номер "CN113098887A". Даний винахід відноситься до галузі веб-безпеки та штучного інтелекту, зокрема до способу виявлення фішингових веб-сайтів на основі загальних характеристик веб-сайтів, який є способом виявлення фішингових веб-сайтів на основі характеристик URL-адрес веб-сайтів, вмісту веб-сайтів, інформації про характеристики сторонніх веб-сайтів та спільні характеристики матриці зміни внутрішнього стану при доступі до веб-сайту.

Патенти є важливими, оскільки вони можуть допомогти захистити винахід. Патенти можуть захищати будь-який продукт, дизайн або процес, що відповідають певним вимогам залежно від їх оригінальності, практичності, придатності та

корисності. Проведення такого патентного пошуку може заощадити час і уникнути непотрібних ідей, досліджуючи якесь питання, дізнавшись, чи існує така ідея чи ні. Після ретельного патентного пошуку можна визначити, чи буде ідея автора схожа на деякі патенти, чи ні. Патентний пошук також допоможе покращити ідею, допомагаючи зрозуміти, що ще є. Аналізуючи дані патенти та винаходи, які які зібрані у списку вище, можна сказати, що серед даних патентів є аналоги, і навіть ті, які майже ідентичні. Тим не менш, маючи список даних патентів і винаходів, можна помітити, що є безліч методів і підходів до вирішення питань в області фішингу, а саме щодо виявлення фішингових веб-сайтів і повідомлень. Також, провівши аналіз даних винаходів та патентів, можна розуміти які зусилля і що вже було зроблено в області фішингу. Після аналізу патентів, також можна зробити висновок, що питання пов'язане з автоматизацією виявлення фішингових повідомлень і веб-сайтів є важливим. Далі, буде представлено аналіз відповідних досліджень, робіт, та які дослідження були проведені раніше і який був результат.

У 2018 році Каспар Юристо провів експеримент з моделювання фішингу Електронної Пошти [5], метою якого було вивчити, як створити зведені рекомендації, які компанії могли б легко впровадити і використовувати в якості стандарту для кращого виявлення фішингових листів. Крім того, в той же час враховуйте, як люди реагують на це, щоб вони не змушували випробовуваних відчувати себе ніяково. Вони також хотіли з'ясувати кореляцію між рівнем складності фішингової електронної пошти та частотою переходів. Це було зроблено шляхом створення двох тестових груп "К" і "L", де обидві групи виконували два різних тести з різними рівнями складності. Вони з'ясували, що в більш простих тестах, які мали більш очевидні ознаки, такі як орфографічні помилки, 11% траплялися на фішингові листи, в той час як в більш складних 23% траплялися на приманку. Деякі з випробовуваних також погодилися пройти додатковий тест, де вони будуть спеціально націлені на інше, відоме як китобійний промисел. Роблячи це, точність приманки досягала 100%, але цілісність об'єктів була більшою мірою під загрозою, що обговорювалося.

У дослідженні 2009 року "Школа фішингу: оцінка в реальному світі навчання боротьбі з фішингом" [6]. Дослідження, основна мета якого - навчити користувачів краще приймати рішення при виявленні фішингових і не фішингових листів. Одна з проблем, з якими вони зіткнулися, полягала в тому, що в цілому люди не мотивовані вивчати питання безпеки і відчують, що це другорядне завдання.

Щоб протистояти цьому, був розроблений інструмент і гра під назвою "PhishGuru" і "Anti-Phishing Phil". PhishGuru - це інструмент, який використовується для прямої освітньої відповіді, коли користувач відкриває свої звичайні електронні листи, де в якості антифішингової програми використовується гра в освітніх цілях. Вони виявили, що подібні інструменти можуть ефективно навчити людей уникати фішингових атак.

Ще в 2015 році було проведено кількісне дослідження для вимірювання ефективності виявлення фішингових листів. Тести були розділені на дві групи, перший і другий тести, в кожній частині яких було по десять електронних листів. У кожній частині тестів було п'ять фішингових і п'ять законних електронних листів, і дві різні частини були виконані якомога більш схожими, наприклад, щоб мати однакову кількість електронних листів від компанії X. Потім учасники виконували один з тестів, після чого виконували навчальну частину, а потім переходили до протилежного тесту. Дослідження призвело до того, що, незважаючи на те, що тести були в деякій мірі ефективними, їх було недостатньо, щоб переконатися, що учасник зможе уникнути фішингових листів в цілому [7]. У подальшому дослідженні вони збираються включити більше тестів, і вони також будуть змінені таким чином, щоб після кожного рішення в кожному електронному листі була пряма відповідь.

Ще одна пропозиція наведена в статті під назвою "М. Манро А. Альнаджим. Підхід до боротьби з фішингом, який використовує навчальне втручання для виявлення фішингових веб-сайтів" [8]. Ідея полягає у використанні навчальних втручань для виявлення фішингових веб-сайтів. Навчання кінцевих користувачів є ключовим компонентом для запобігання використанню фішингу. Навчальне втручання, яке запропоноване в цьому дослідженні, дозволяє кінцевому користувачеві спробувати самостійно виявити фішинг, і якщо він допустить помилку, йому негайно буде надана корисна інформація в спробі підвищити компетентність. Основна мета цього дослідження полягала в тому, щоб порівняти ефективність цього підходу до навчання з застарілими методами, такими як надсилання антифішингових електронних листів, які містять поради про те, як виявляти фішингові електронні листи. З цього дослідження був зроблений висновок про значний позитивний ефект використання цього методу, щоб допомогти користувачам правильно оцінювати законні та фішингові веб-сайти.

Здавалося б, поширеною концепцією навчання в спробах поліпшити навички виявлення фішингу у людей є впровадження вбудованого навчання. У дослідженні висувається гіпотеза про те, що це тип навчання, при якому відбувається негайна реакція, коли кінцевий користувач приймає неправильне оціночне рішення. Для подальшого вивчення ефективності вбудованого навчання був проведений великомасштабний експеримент з аналізу реакції кінцевих користувачів на фішинг-розсилку, об'єднаний з вбудованим навчанням. Всупереч тому, що передбачали попередні дослідження, було складніше зробити цей тип навчання досить ефективним у корпоративному середовищі, в якому він був створений. Їх результат показав, що негайного зворотного зв'язку недостатньо для зниження рейтингу кліків. Грунтуючись на кількості часу, який кінцеві користувачі витратили на кожну частину, вони могли припустити, що вони не прочитали весь тренінг [9].

1.6 Статистика фішингових атак та актуальність проблеми фішингу

За дослідженням "IRONSCALES" [10], з березня 2020 року 81% компаній у всьому світі зіткнулися зі збільшенням числа фішингових атак по електронній пошті. Не залучаючи до уваги дійсну загрозу, яку на сьогоднішній день фішинг представляє для компаній і користувачів, майже кожна компанія проводить тільки навчання з питань фішингу для своїх співробітників, один раз на рік. Така відсутність обізнаності є важливим фактором, що сприяє тому, що фішинг залишається типом загрози, яка з найбільшою ймовірністю може призвести до витоку даних. Згідно з "Data Breach Investigations Report" за 2021 рік, близько 25% всіх витоків даних пов'язані з фішингом, 85% витоків даних пов'язані з людським фактором. Така цифра підтверджується дослідженнями які були проведені Центром скарг на злочини в Інтернеті (IC3), який отримав рекордну кількість скарг від американських громадян в 2020 році. У США в 2020 року, 241 342 людини стали жертвами. За цим послідували невіплата, або ж недоставка, вимагання, витік персональних даних, і крадіжка особистих даних [11].

Середня кількість спроб компрометації корпоративної електронної пошти (BEC), яку було отримано за останній рік, значно збільшилася на 15% між другим і третім кварталом року, і все частіше спостерігається зловмисні витoki даних, викликані вкраденими обліковими даними, а не установкою шкідливого ПЗ [12]. За інформацією "IBM" [13], кожна компанія, яка постраждала від витоку даних в

2021 році, була зламана через втрату або крадіжки облікових даних, 17% були зламані в результаті прямої фішингової атаки. Тим не менш, дані "Google Safe Browsing" показують, що на даний момент в Інтернеті майже в 75 разів більше фішингових сайтів, ніж шкідливих сайтів [14].

У першому кварталі 2021 року багато людей зіштовхнулися з небезпечними посиланнями під час використання поштових сервісів. Менш ніж через рік ще більше людей зіткнулися з даними загрозами у третьому кварталі 2021 року. На зображенні нижче показана частота небезпечних URL-адрес в 2021 році.

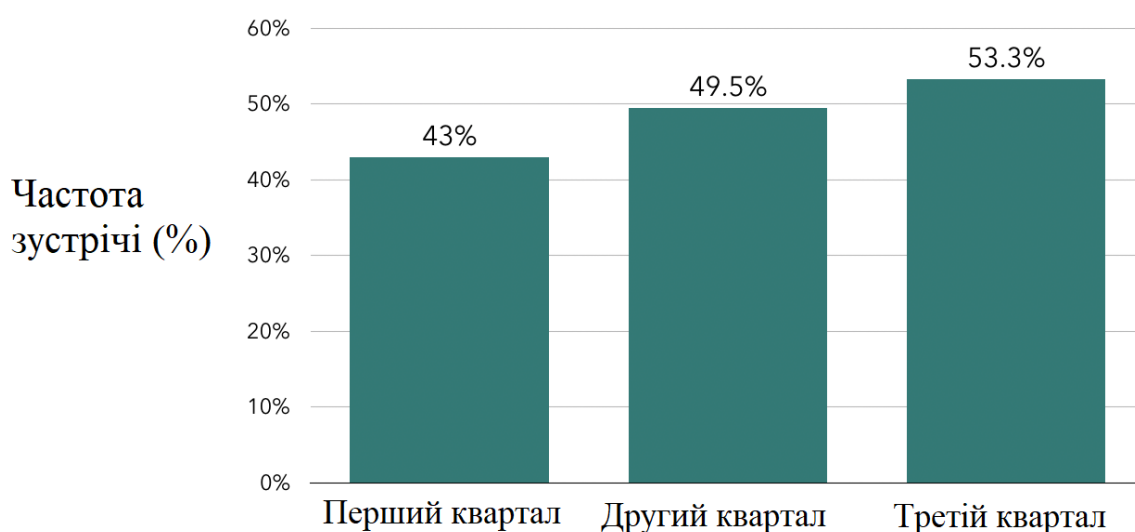


Рисунок 1.7 - Частота небезпечних URL-адрес

З рис. 1.7 видно, що у кожному новому кварталі року частота зустрічей підвищується. Це може означати те, що атаки фішинга нікуди не йдуть, і поступово збільшуються.

За інформацією "SonicWall" [15], загальна кількість атак програм-вимагачів збільшилася більш ніж удвічі в 2021 році, збільшившись за рік на 105% в порівнянні з минулим, 2020 роком. А за інформацією "CrowdStrike", витoki даних які пов'язані з програмами-вимагачами, зросли на 82% у 2021 році. Proofpoint, це звіт який заснований на опитуванні 600 фахівців з безпеки і 3500 співробітників в Австралії, Франції, Німеччині, Японії, Іспанії, Великобританії і США [16], також в даному звіті є дані про змодельованих фішингових атаках, відправлених Proofpoint. Результати опитування про фішинг і програми-вимагачів прийшли до такого висновку, що "співробітники відчувають себе вигорілими, емоційно виснаженими і розсіяними", дані слова позначені в звіті Proofpoint. Нижче наведено

22 статистичні показники зі звіту. Це тривожні статистичні дані про зростання фішингу, та програм-вимагачів, зі звіту "Proofpoint State of the Fish" за 2022 рік [16].

1) Фішинг на основі електронної пошти - 83% компаній заявили, що виступували вдало фішинговою атакою на основі електронної пошти в 2021 році, в порівнянні з 57% в 2020 році.

2) Масовий фішинг - 86% компаній заявили про масові фішингові атаки в минулому році, в порівнянні з 77% роком раніше.

3) Цільові фішингові атаки - у 2021 році 79% компаній зіткнулися з цілеспрямованими атаками, це більше, ніж 66% роком раніше.

4) "Смішинг" - 74% компаній заявили про смішинг-атаки в 2021 році в порівнянні з 61% в 2020 році.

5) "Вішинг" - у 2021 році 69% компаній заявили про вішинг-атаки, в яких використовуються телефонні дзвінки. У 2020 році було 54%.

6) Соціальні атаки - 74% компаній зазнали атак через соціальні мережі в 2021 році в порівнянні з 61% в 2020 році.

Також, для більш чіткої картини, нижче на рис. 1.8. представлена статистика звіту APWG [17], дана статистика буде представлена за четвертий квартал 2021 року. Звіт APWG, це звіт про тенденції фішингової активності, в даному звіті аналізуються фішингові атаки, про які повідомляють APWG її компанії, а також партнери з досліджень, через веб-сайт організації.

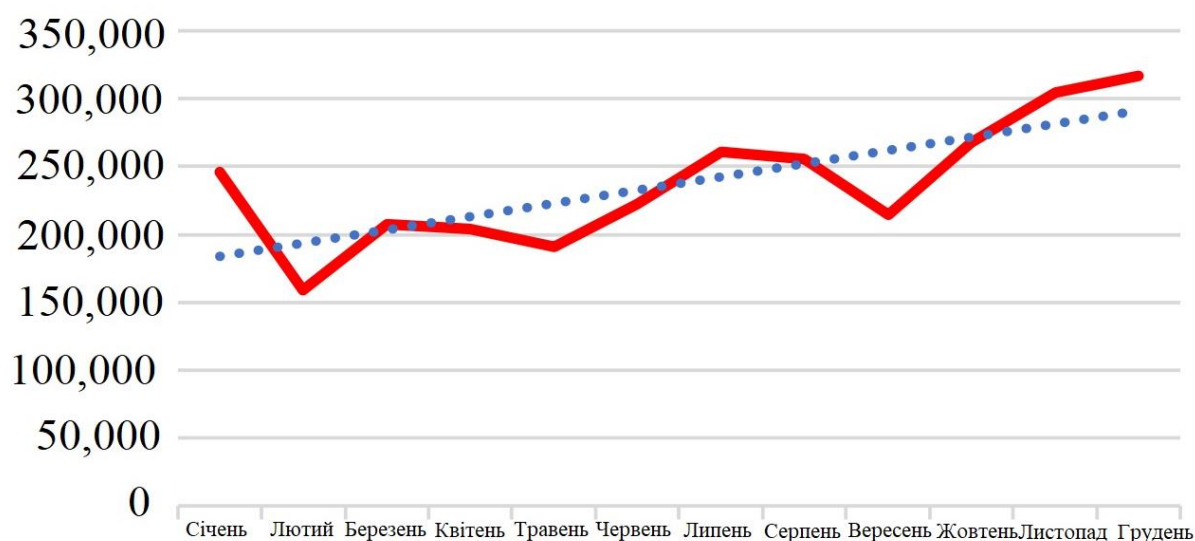


Рисунок 1.8 - Статистика фішингових атак станом на 2021 рік

З даного графіку на рис. 1.8. видно, що в грудні 2021 року APWG зафіксувала 316 747 фішингових атак, це стало найвищим місячним показником за всю історію звітів APWG. Кількість недавніх фішингових атак збільшилася більш ніж втричі з початку 2020 року, коли APWG спостерігала від 68 000 до 94 000 атак на місяць [17]. Далі розглянемо найбільш вразливі галузі промисловості на 2021 рік. На рис. 1.9. представлена кругова діаграма найбільш вразливих галузей промисловості за інформацією "APWG" на 2021 рік.

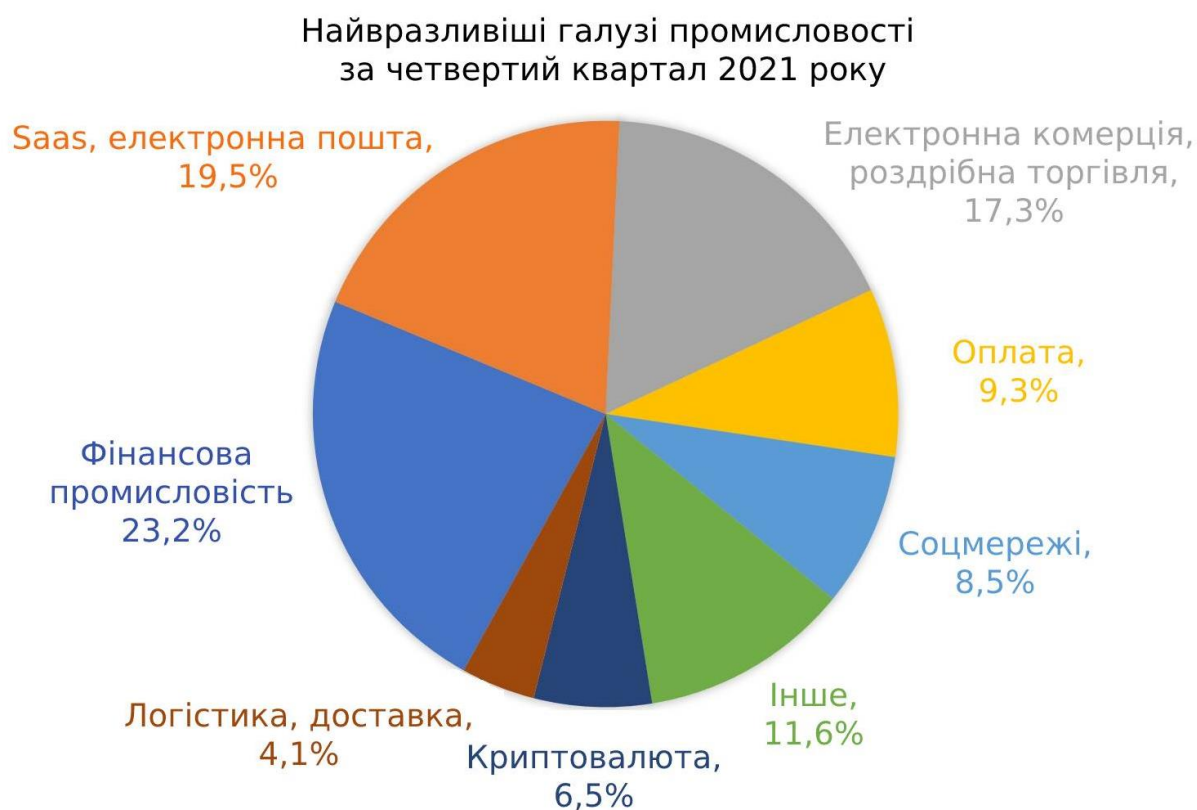


Рисунок 1.9 - Найбільш вразливі галузі промисловості станом на четвертий квартал 2021 року

З рис. 1.9 видно, що в четвертому кварталі 2021 можна помітити, що фішингові атаки на фінансовий сектор, в який входять банки, стали найбільшим набором атак, на які припадає 23,2% всього фішингу. Атаки на провайдерів веб-пошти та програмного забезпечення як послуги (SaaS) теж залишалися популярними, але знизилися з 29,1% усіх атак у третьому кварталі до 19,5% у четвертому кварталі. Що стосується криптовалют, то криптовалютні біржі та постачальники гарантій складають до 6,5% атак. З рис. 1.9 також видно, що електронної комерції роздрібної торгівлі підвищилися, швидше за все це сталося через святкові розпродажі.

Розглянемо тепер атаки з урахуванням електронної пошти. На рис. 1.10 нижче представлено кругову діаграму атак на основі електронної пошти за четвертий квартал 2021 року.

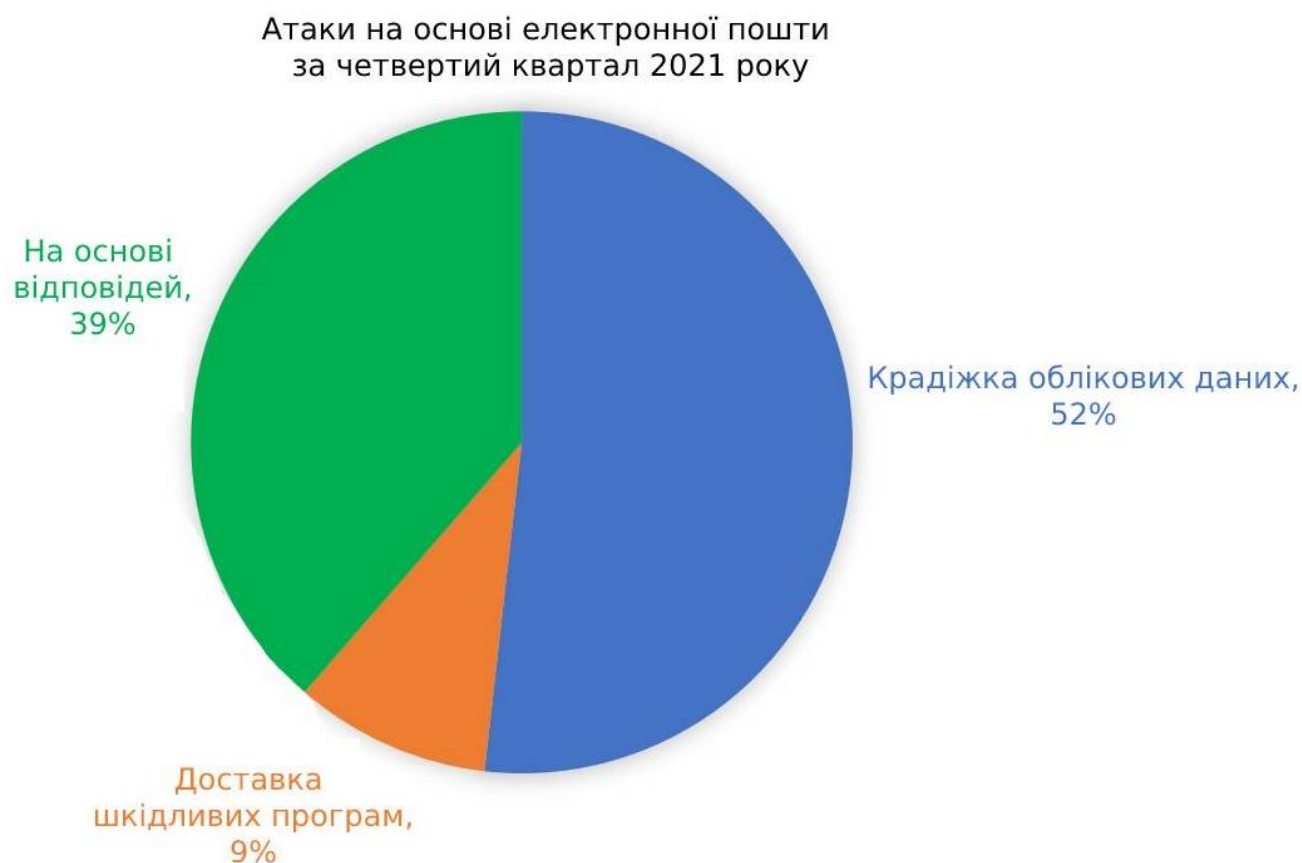


Рисунок 1.10 - Атаки на основі електронної пошти станом на четвертий квартал 2021 року

З рис. 1.10 видно, що 51,8 % листів були фішинговими атаками з крадіжкою облікових даних, 38,6 % це відповідні атаки, 9,6% це листи з доставкою шкідливого ПЗ [17]. Крім цього, в цьому кварталі також було помічено повернення "Emotet", атаки шкідливого ПЗ, зазвичай здійснюваної по електронній пошті. Також, величезною тенденцією шкідливих електронних листів, націлених на корпоративних користувачів, стало різке збільшення кількості електронних листів Qbot, оскільки 59% шкідливих програм які були доставлені поштою в четвертому кварталі, були пов'язані з Qbot.

Тобто, з даної статистики можна сказати, що неважливо будь то компрометація корпоративної електронної пошти, або ж програма-вимагач, або безліч інших типів загроз, електронна пошта завжди залишається каналом №1 для зловми-

сників, який щорічно краде дані і викачує мільярди. Безліч статистик і звітів свідчить про те що більше 90% цільових атак починаються з електронної пошти, і всі вони майже що засновані на взаємодії людей, це робить людей новим периметром компаній, який необхідно захищати. Акцент на захисті цифрових систем протягом цих декількох років свідчить про те що зловмисники стали використовувати поєднання приманок соціальної інженерії по електронній пошті з різними методами атак які доставляються через вкладення або URL-адреси. Враховуючи цю статистику, і те, що на даний момент фішинг на основі електронної пошти є все більш актуальним, у даній роботі з проблем буде приділено увагу фішинговим повідомленням та веб-сайтам.

2 АНАЛІЗ ТА ДОСЛІДЖЕННЯ ОЗНАК ФІШИНГОВИХ ПОВІДОМЛЕНЬ

2.1 Ознаки фішингових повідомлень

Для того щоб проводити дослідження, необхідно дати чітке пояснення, які можуть бути ознаки фішингових повідомлень. Дані ознаки будуть використовуватися в подальшому для побудови критерію для прийняття рішення, що отримано фішингове повідомлення. Під ознаками фішингових повідомлень маються на увазі ті ознаки, які ідентифікують повідомлення як фішингове. Ці ознаки будуть дуже важливими при дослідженні і при розробці припущень щодо автоматизації процедур. Розглянемо ознаки фішингових повідомлень докладно.

Довга URL-адреса - це перша ознака, довгі URL-адреси можна використовувати, для того щоб заховати підозрілу частину адресного рядка. Хоча з наукової точки зору є вірний метод прогнозування діапазону довжини, який виправдовує веб-сайт як фішинговий або нефішинговий, але в такому випадку, це критерії, які використовуються з іншими функціями при виявленні підозрілих сайтів. В аналізі Vasnet et al. [18], запропонована довжина ≤ 75 , але не було обґрунтування цієї довжини. В даному випадку, довжина URL-адреси може бути >127 символів, це в разі не фішингових веб-сайтів, і ≤ 127 символів для фішингових веб-сайтів. Дане значення вибирається на основі набору даних, зібраного шляхом ручного порівняння довжини найдовшого нефішингового веб-сайту і фішингового веб-сайту в наборі даних. Нижче на рис. 2.1. наведений приклад довгої URL-адреси.

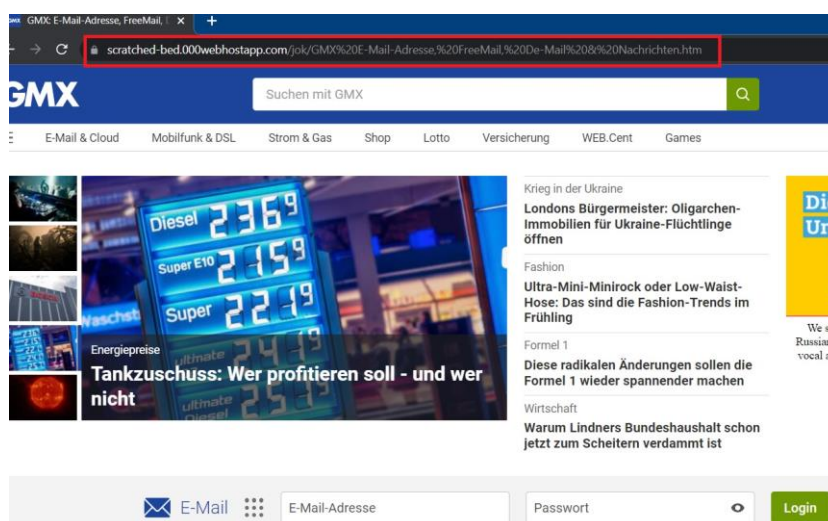


Рисунок 2.1 - Приклад довгої URL-адреси

З рис. 2.1. видно, що даний URL-адрес досить довгий. При першому перегляді на даний URL-адресу, у користувача може виникнути підозри з приводу легітимності даного сайту, оскільки зазвичай посилання не такі довгі. До того ж, в даному випадку в адресі вказується значок замка, і https, але навіть в цьому випадку це не знак того що даний сайт є не фішинговим. Нижче на рис. 2.2. показано результат даної URL-адреси з сайту "Phishtank". "Phishtank", це веб-сервіс, що має на меті мінімізувати фішингові атаки, використовуючи спільну роботу користувачів. Цей сайт зберігає велику кількість фішингових сайтів, сайтів які на даний момент активні. Дані веб-сайти збирають користувачі, після чого завантажують їх на цей веб-сайт. Тобто кожен користувач може знайти для себе на цьому сайті хороший набір фішингових сайтів.



Рисунок 2.2 - Результат URL-адреси

Як видно з рис. 2.2., дана URL-адреса визначається як фішингова. Тобто, таким чином, можна побачити, що довгі посилання дійсно можуть бути небезпечними. В основному, найчастіше в фішингових повідомленнях зустрічаються саме такі подібні посилання.

Точки в URL-адресі - безпечне посилання на веб-сторінку містить не більше п'яти точок. Якщо, можливо, на веб-сторінці більше п'яти точок, то в такому випадку, це може бути розпізнано як фішингова посилання. Наприклад, на рис. 2.3. наведено приклад такої URL-адреси, яка має більше п'яти точок в адресі.

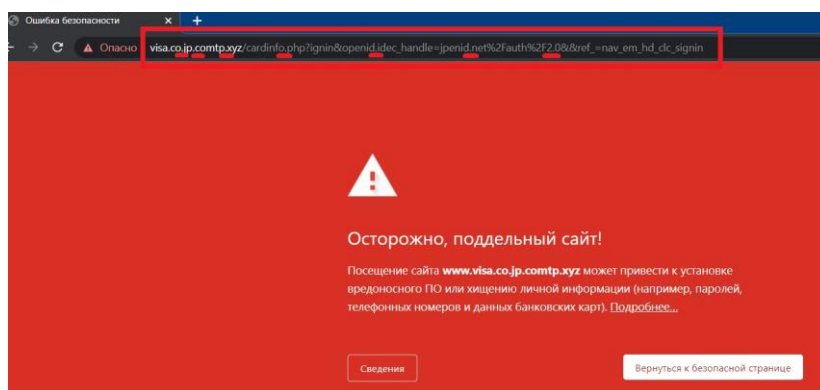


Рисунок 2.3 - Приклад URL-адреси яка має більше п'яти точок

Як видно з рис. 2.3, в даному випадку дана URL-адреса має більше п'яти точок в адресі, Дана адреса має сім точок. Також, при заході на дану веб-сторінку, нам автоматично показує повідомлення про те, що даний сайт небезпечний. Нижче на рис. 2.4. наведено результат даної URL-адреси з сайту "Phishtank".



Рисунок 2.4 - Результат URL-адреси

Як видно з рис. 2.4., даний веб-сайт є фішинговим. Це показує те, що точки в адресі це ще одна ознака за яким можна визначити фішингове повідомлення. До того ж, у користувачів відразу виникає сумнів, якщо в URL-адресі з'являється дуже багато точок.

IP-адреса - деякі веб-сайти розміщені з IP-адресою замість повного домену. Це дуже підозріла дія, тому що безліч законних веб-сайтів більше не використовують цей метод із міркувань безпеки. Також, оскільки більшість веб-сайтів фішингу залишаються в мережі протягом обмеженого часу, цю функцію можна розглядати як одну з дуже важливих функцій виявлення фішингу. Нижче на рис. 2.5,

та рис. 2.6. наведено приклад такої адреси, а також результат цієї URL-адреси з сайту "Phishtank".

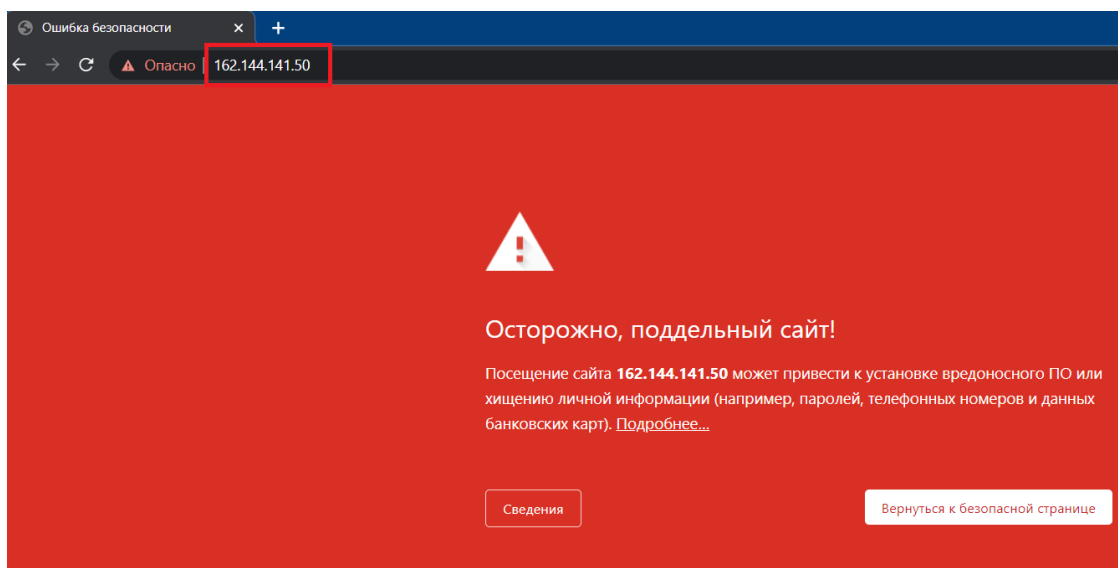


Рисунок 2.5 - Приклад URL-адреси з IP-адресом



Рисунок 2.6 - Результат URL-адреси з IP-адресом

Як видно із рис. 2.5 та рис. 2.6, цей сайт також є фішинговим і дуже підозрілим. Тим не менш, в даному випадку IP-адресу можна також використовувати для отримання більш детальної інформації про цей веб-сайт, наприклад, за допомогою сервісу "Whois", або за допомогою сервісу "GeoIP". "Whois", доступний як у вигляді термінальної програми, так і в багатьох випадках на загальнодоступних веб-сайтах для отримання загальнодоступного реєстратора та іншої інформації для даного домену. Він використовувався в цьому дослідженні для отримання інфор-

мації про ім'я реєстратора веб-сайту, реєстр і дату закінчення терміну дії веб-сайтів. Іншими словами, час життя веб-сайтів. "Geoip", це інструмент, який дозволяє користувачам шукати, де в даний час знаходиться пристрій з певною IP-адресою.

SSL-з'єднання - в даному випадку, це теж може бути ознакою. Необхідно, щоб сайт оплати або електронної комерції був захищений таким чином, щоб дані, що надсилаються з веб-сайту та на веб-сайт, були зашифровані. Також його можна використовувати для підтвердження автентичності веб-сайту за допомогою сертифіката SSL, який включає конкретну інформацію про веб-сайт. Зразок одного з нефішингових веб-сайтів показано на рис. 2.7.

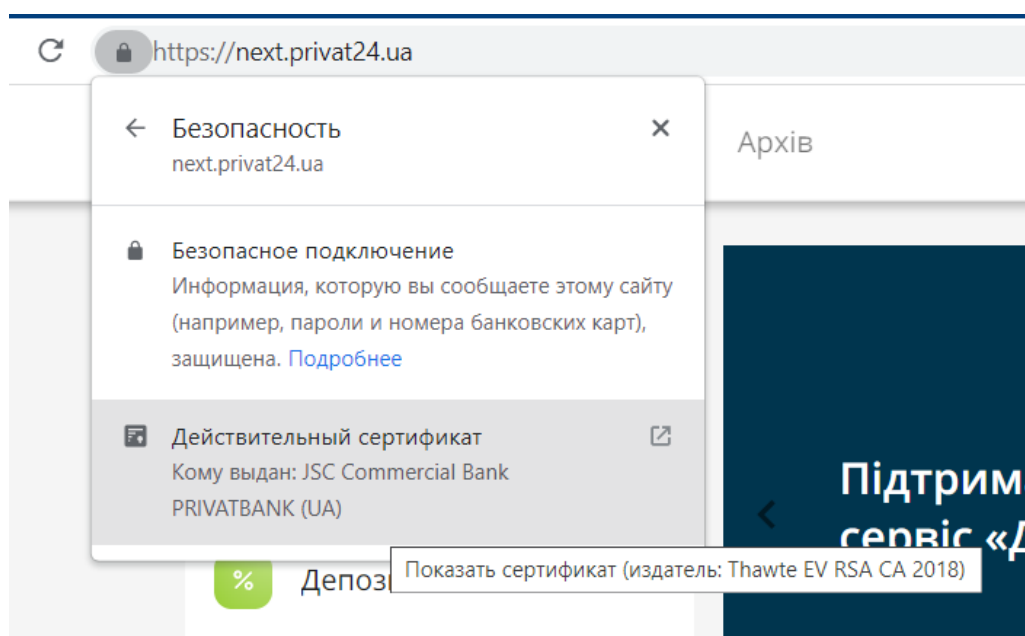


Рисунок 2.7 - Зразок SSL

На рис. 2.7. видно, що цей нефішинговий сайт має дійсний SSL-сертифікат, що свідчить про легітимність даного веб-сайту.

Символ "@" - ще одна ознака, за якою можна визначити фішингове повідомлення. Фішингова URL-адреса може включати цей символ, символ "@", всередині адреси. Зазвичай веб-браузер під час читання адреси ігнорує все, що знаходиться ліворуч від символу "@", тому така адреса як "privat24.ua@fake-bank.ua" насправді буде такою, "fake-bank.ua". Нижче на рис. 2.8. наведено приклад такої URL-адреси.

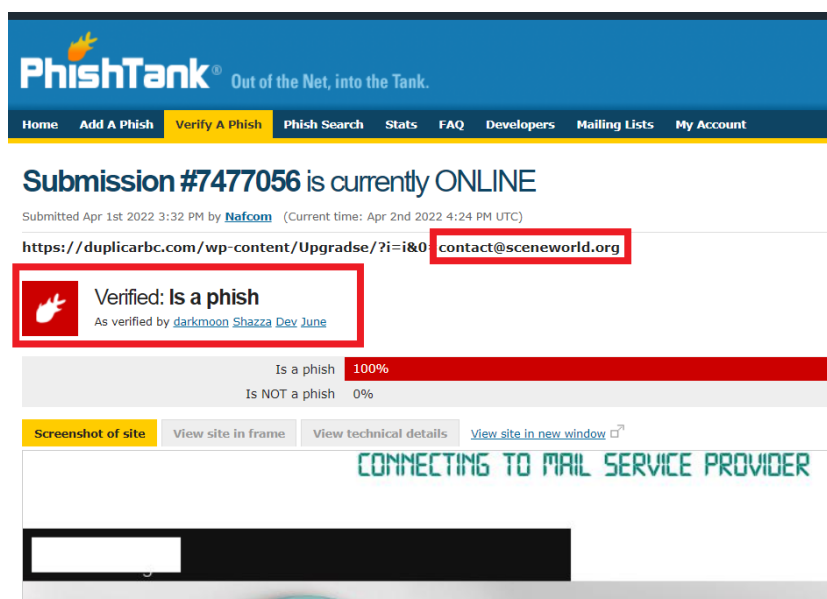


Рисунок 2.8 - Приклад URL-адреси з символом

На рис. 2.8. видно, що це посилання є фішинговою, і в ній присутній символ "@", тобто такий випадок також можна вважати ознакою фішингових повідомлень.

Шістнадцяткове кодування - також є ознакою фішингових повідомлень, зокрема, для фішингу використовуються URL-адреси в шістнадцятковому кодуванні. В інтересах сумісності більшість поштових користувацьких агентів, веббраузерів і HTTP-серверів розуміють основні еквіваленти символів в шістнадцятковому кодуванні. Основна незаконна мета такого кодування полягає в тому, щоб обійти антиспамові фільтри на основі чорних списків, які не обробляють шістнадцяткове кодування символів (по суті, ще одна атака вставкою). Таке кодування також обходить механізми захисту, які забороняють IP-адреси як URL-адреси призначення, виходячи з припущення, що "звичайні" http-посилання будуть використовувати більш знайомі DNS-імена.

Фрейми - також можуть бути ознаками фішингових повідомлень, це досить відомий метод приховування атакованого контенту завдяки однаковій підтримці браузерів [19], а також простому стилю написання коду. На рис. 1.8. нижче наведено приклад фрейма.

```

<title>
<html>
  <head>
    <title>Приклад фрейму</title>
  </head>
  <body topmargin="0" leftmargin="0"
rightmargin="0">
    <iframe src="https://next.privat24.ua/"
width='100%' height='150' frameborder="0">
      <iframe>
        <iframe
src="http://www.yeah.com/" width='100%' height='250%'
frameborder="0">
          <iframe>
            </body>
          </html>
    </iframe>
  </body>
</html>

```

Рисунок 2.9 - Приклад фрейму

З рис. 2.9. видно, що даний випадок описує сценарій, в якому зломисник визначає два кадри. Перший кадр містить інформацію про URL легітимного сайту, тоді як другий кадр займає 0% інтерфейсу браузера, на якому запущений шкідливий код. Сторінка яка пов'язана всередині прихованого фрейму може використовуватися як доставка додаткового контенту, отримання конфіденційної інформації, також при такому випадку, можливий наприклад захоплення екрану і реєстрації клавіш. Результат фрейму який показаний на рис. 2.9, показаний на рис. 2.10 нижче.

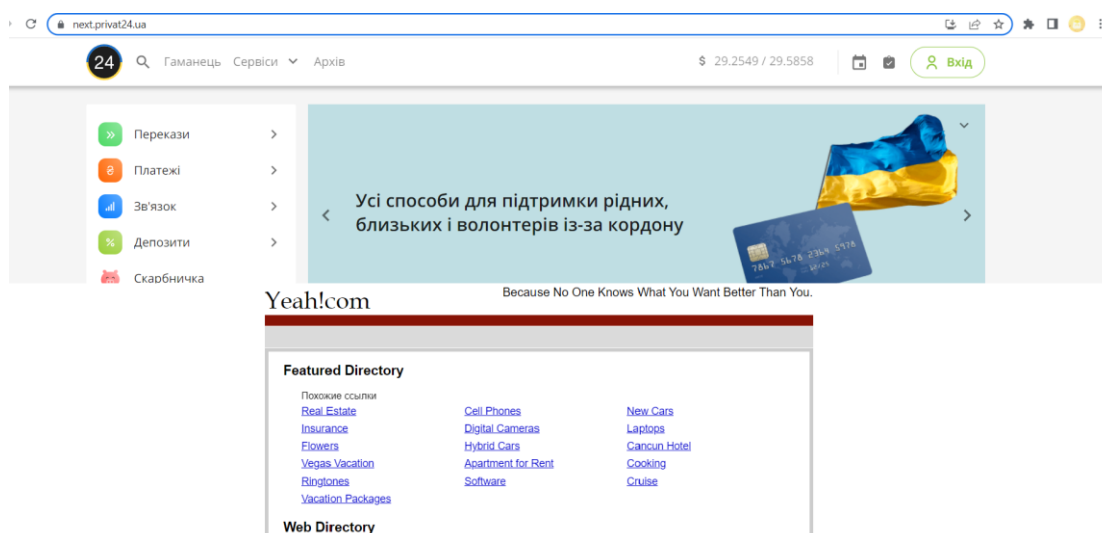


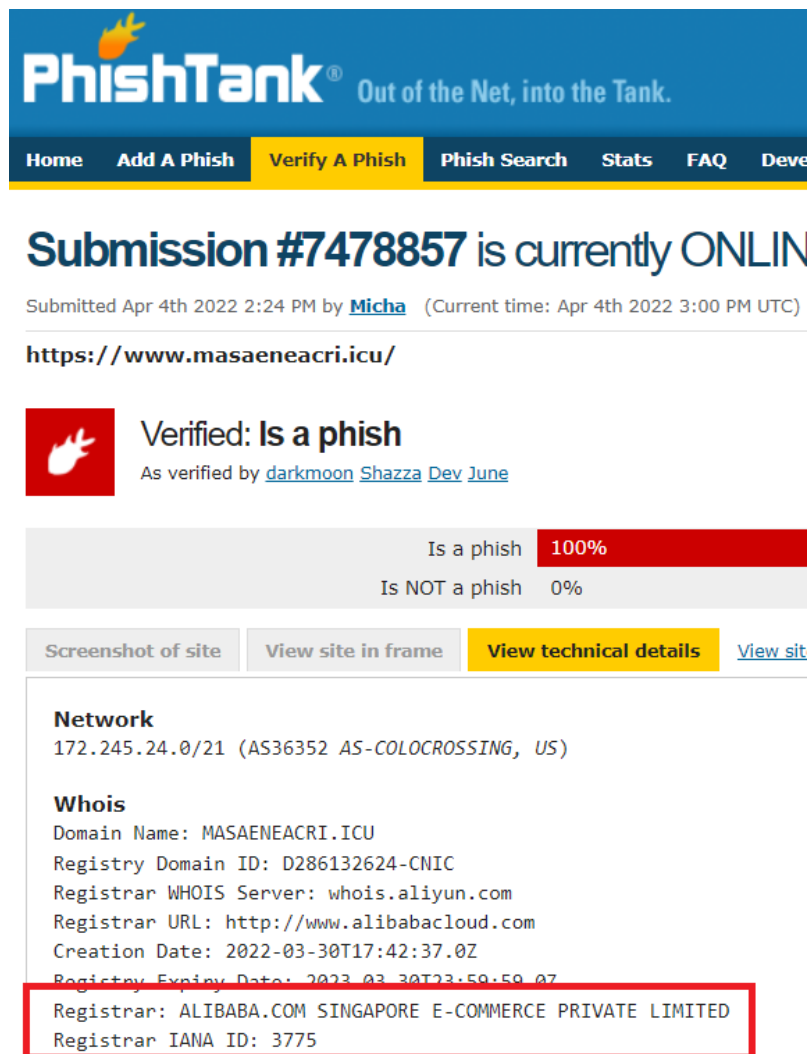
Рисунок 2.10 - Результат фрейму

Перенаправлення - веб-додаток має приймати введення який управляється користувачем, це введення вказує посилання на зовнішній сайт, і використовує це

посилання в перенаправленні. Цей випадок спрощує атаки фішингу. Наприклад, таке посилання як "www.privat24.u/j/43110;hack.com" може перенаправити на hack.com, використовуючи "privat24" як сайт перенаправлення. Через те, що "privat24" це звичайний банківський сайт, користувач перед тим як перейти за посиланням передбачає, що веб-сайт містить "privat24" на початку, і закінчується на "privat24", якби посилання починалося так, "www.hack -privat24.ua", то користувач почав би щось підозрювати.

Кнопка "відправити" - теж може бути ознакою фішингових повідомлень. Зазвичай фішингові сайти досить часто мають у вихідному коді таку кнопку як "відправити". Даний вихідний код може мати адресу електронної пошти зломисника або ж базу даних. У тому випадку якщо ж користувач вирішує що цей сайт цілком легітимний, і він натискає дану кнопку "Відправити", то показується що сторінка не може бути знайдена, або на екрані з'являються будь-які інші повідомлення про помилки. В такому випадку, користувач може припустити, що це якась проблема пов'язана з мережею. Така ознака досить часто трапляється в фішингових повідомленнях.

Реєстратор - теж може бути ознакою. Реєстратор, може резервувати доменні імена, а також призначення IP-адреси для доменних імен. Доменне ім'я це буквенно-цифрові псевдоніми, які можуть використовуватися для доступу до веб-сайтів. Дуже часто, у фішингових повідомленнях зустрічаються підозрілі доменні імена, такі які зовсім не відповідають напрямку компанії, або тому місці, де працює користувач. Нижче на рис. 2.11 наведено приклад доменного імені.



PhishTank® Out of the Net, into the Tank.

Home Add A Phish Verify A Phish Phish Search Stats FAQ Dev

Submission #7478857 is currently ONLINE

Submitted Apr 4th 2022 2:24 PM by [Micha](#) (Current time: Apr 4th 2022 3:00 PM UTC)

<https://www.masaeneacri.icu/>

 **Verified: Is a phish**
As verified by [darkmoon](#) [Shazza](#) [Dev](#) [June](#)

Is a phish	100%
Is NOT a phish	0%

[Screenshot of site](#) [View site in frame](#) [View technical details](#) [View sit](#)

Network
172.245.24.0/21 (AS36352 AS-COLOCROSSING, US)

Whois
Domain Name: MASAENEACRI.ICU
Registry Domain ID: D286132624-CNIC
Registrar WHOIS Server: whois.aliyun.com
Registrar URL: http://www.alibabacloud.com
Creation Date: 2022-03-30T17:42:37.0Z
Registry Expiry Date: 2023-03-30T23:59:59.0Z
Registrar: ALIBABA.COM SINGAPORE E-COMMERCE PRIVATE LIMITED
Registrar IANA ID: 3775

Рисунок 2.11 - Приклад домменного імені

З рис. 2.11 видно, що реєстратор цього сайту це "ALIBABA.COM SINGAPORE E-COMMERCE PRIVATE LIMITED". Тобто, переглянувши дані сайту, можна помітити цей реєстратор, і у користувача можуть виникнути підозри, оскільки він не зможе зрозуміти, яким чином цей реєстратор пов'язаний з його діяльністю, або яким чином цей сайт потрапив до нього. Це ще одна ознака, за якою можна визначати фішингові веб-сайти.

Країна - зазвичай, коли користувач отримує фішингове повідомлення, він не може дивлячись визначити країну даного веб-сайту, але в цьому випадку можна використовувати сервіс "whois". Якщо ж користувач, якому попалося фішингове посилання у фішинговому повідомленні дізнався країну даного фішингового веб-сайту, то він відразу розуміє що це нелегітимний або шкідливий сайт, тому що якщо користувач працює в області, в яку ніяк не можуть входити веб-сайти з ін-

ших країн, і йому трапляється сайт з країни Китай, це може бути ознакою того, що даний сайт швидше за все є фішинговим.

Погана граматика - також є ознакою, оскільки можна відразу визначити за даною ознакою чи є електронний лист шахрайством, якщо він містить погану орфографію, або граматику. Можна припустити, що дані помилки це частина "системи фільтрації", в якій зломисники націлені тільки на самих довірливих людей. Такий випадок також досить часто зустрічається, і є ознакою, але зазвичай користувачі відразу визначають що дане повідомлення фішингове, або пошта автоматично закидає даний лист в спам. Нижче на рис. 2.12 наведено приклад такого фішингового письма з поганою граматикою.

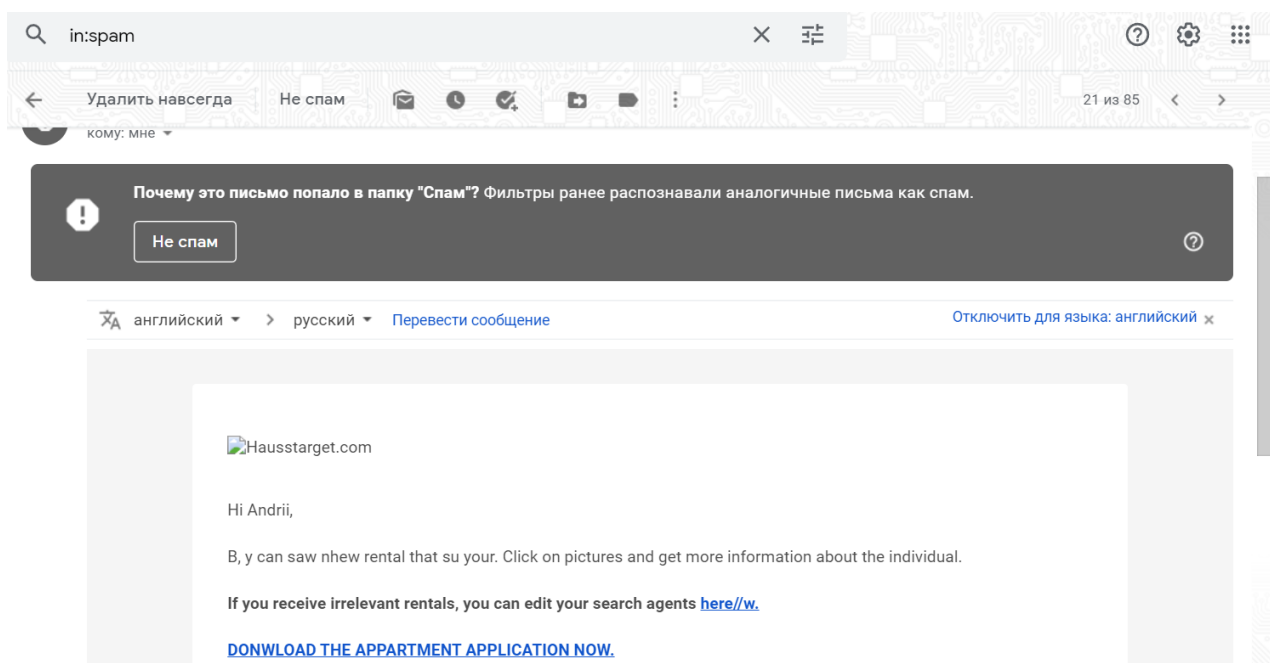


Рисунок 2.12 - Приклад фішингового письма з поганою граматикою

Відчуття терміновості - ознака фішингових повідомлень. Дуже часто в фішингових повідомленнях зустрічається текст який створює відчуття терміновості, що необхідно прямо зараз перейти за посиланням, або ж інший якийсь варіант. Штучне відчуття терміновості однаково ефективно при фішингу на робочому місці. Зломисник повідомлений про те що користувачі швидше за все кинуть всі свої справи якщо начальник відправить йому електронного листа з дуже життєво важливим проханням. Нижче на рис. 2.13 наведено приклад такого типу листа.

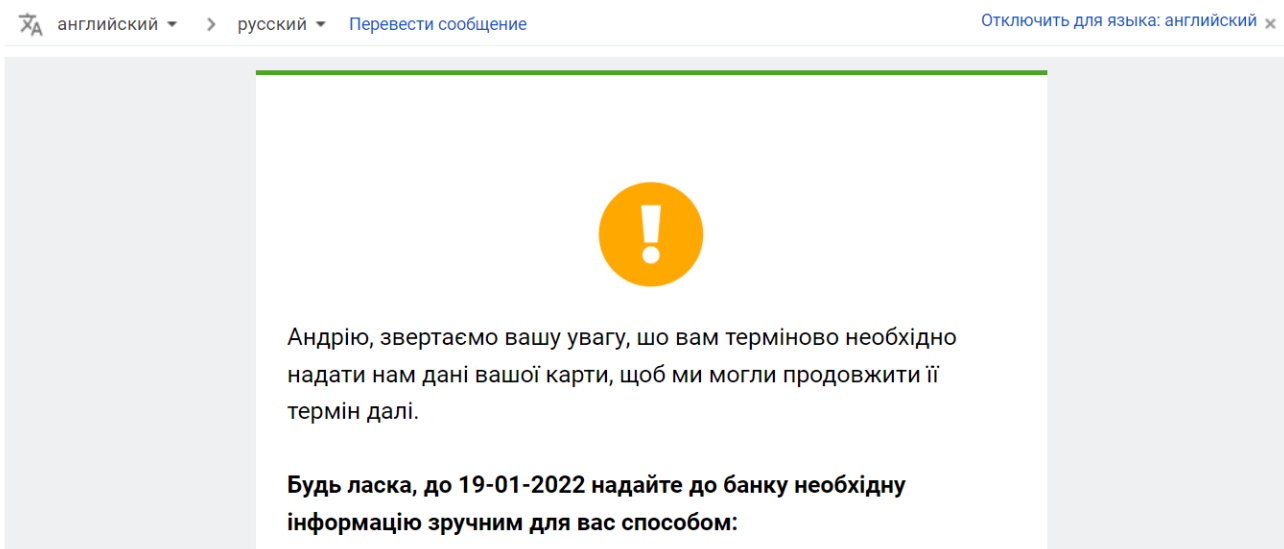


Рисунок 2.13 - Приклад фішингового письма с відчуттям терміновості

Також можна виділити і такі ознаки фішингових повідомлень:

- наявність піддомену;
- фавікон значок веб-сайту;
- префікс та суфікс;
- кількість косих рисок;
- довжина піддомену;
- використання нестандартного порту;
- посилання у тегах;
- кількість перенаправлень на сайт;
- вимкнення правої кнопки миші;
- використання спливаючих вікон;
- вік домену;
- посилання що вказують на сторінку.

Виділивши ознаки фішингових повідомлень, можна розподілити їх за категоріями і зробити схему. Нижче на рис.2.14. представлена класифікація ознак фішингових повідомлень.

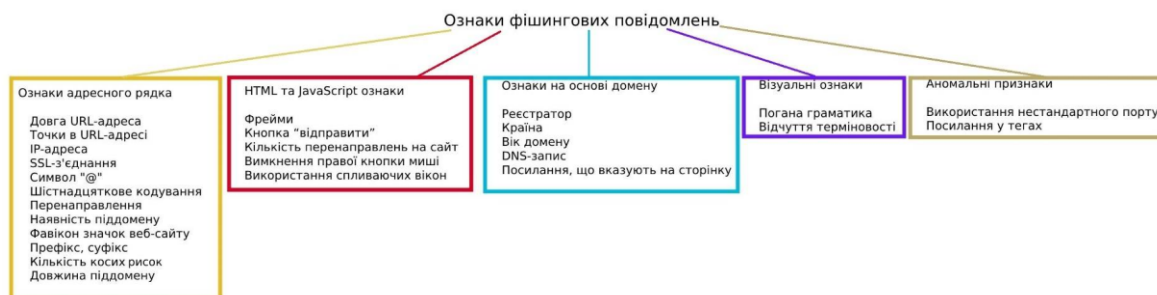


Рисунок 2.14 - Класифікація ознак фішингових повідомлень

На рис. 2.14. вище були зібрані і розподілені всі ознаки фішингових повідомлень. Дані ознаки, необхідні для того, щоб можна було ідентифікувати фішингові повідомлення. Як видно з рис. 2.14, всі ці ознаки служать дуже хорошим ідентифікатором, за даними ознаками можна визначити чи є повідомлення фішинговим чи ні. Деякі з даних ознак будуть надалі використовуватися в ході дослідження, для того щоб можна було ідентифікувати фішингові повідомлення. Також, дані ознаки можуть послужити і звичайним користувачам, а також компаніям, наприклад переглянувши дану схему, користувач може визначити для себе, що може бути фішинговим повідомленням, і як за допомогою даних ознак визначити таке повідомлення.

2.2 Дослідження аналізу фішингових повідомлень

Для того щоб розуміти які є підходи до вирішення завдання виявлення фішингових повідомлень, в даному розділі буде проведено дослідження щодо цього питання. В ході дослідження, будуть зібрані фішингові і не фішингові URL-адреси, після чого на основі зібраних адрес будуть зроблені діаграми частот. Для даного дослідження передбачалося дослідити пару фішингових і чистих веб-сайтів. Для того щоб визначити фішингові і не фішингові сайти, були взяті ознаки з розділу 2.1, а саме IP-адреса, реєстратор, час життя, країна. Буде підрахована частота зустріч даних ознак, наприклад скільки посилань мають реєстрацію на один рік, або на два роки, тобто в даному дослідженні буде проведено також частотний аналіз. Завдання даного дослідження, це зрозуміти за даними частотним діаграм яку ознаку можна використовувати для виявлення фішингових повідомлень. Тобто, за яким з цих (Registrar, Lifetime, Country) ознак можна розділити фішингове і довірене повідомлення. В даному випадку необхідно вирішити яку ознаку можна використовувати і які значення можуть стати розрахункової кордоном. Дослі-

дження фішингових повідомлень було зроблено за допомогою декількох інструментів. Конкретні інструменти можна побачити в наступному списку нижче.

1) Whois - у контексті цього дослідження для визначення інформації про доменне ім'я сайту.

2) Geoip - у контексті цього дослідження для визначення місцезнаходження веб-сайтів використовувався "geoIP".

3) Ping - інструмент для відправки Ехо-запитів ICMP до джерел в інтернеті або локальних мережах. У контексті цього дослідження, він використовувався для визначення деяких IP-адрес за відомими URL-адресами, щоб потім ці IP-адреси можна було використовувати для визначення того, де розміщені веб-сайти.

Для дослідження було відібрано 33 фішингові URL, і 20 нефішингові URL. Дані URL представлені нижче на рис.2.15 та рис.2.16.

URL	IP address	Registrar	Lifetime	Country
https://updateatitbox.weebly.com/	74.115.50.110	WEEBLY.COM		16 USA
https://erformance.serveftp.com/?checkid=a@abc	8.23.224.108	SERVEFTP.COM		21 USA
http://eggheadspooled.info/	N/A	EGGHEADSPOOLED.INFO		1 GERMANY
https://f002.backblazeb2.com/file/offals-tigenshly-vitalizers/index.html	104.153.233.176	BACKBLAZEB2.COM		8 USA
https://actualisation.ddns.net/particuliers/a04a3	8.23.224.108	DDNS.NET		21 USA
Ow3egG6dtWYckdb74RpAJVOWR6x	162.241.121.34	RUSSELLSTREAMS.COM		1 USA
https://usdtmarkets.com/	161.117.253.185	USDTMARKETS.COM		1 Singapore
https://tsfdffdfdfdf.000webhostapp.com/help-support.html	153.92.0.100	000WEBHOSTAPP.COM		6 Cyprus
https://show.servemp3.com/?checkid=a@abc	8.23.224.108	SERVEMP3.COM		22 USA
http://worldbankgroupe.com	N/A	WORLDBANKGROUPE.COM		1 USA
https://lsquaredwebdesigns.com/awe/lopassw/mweb/login	50.116.92.29	LSQUAREDWEBDESIGNS.COM		4 USA
http://prevencionvialbcpzonaseguras.esc-pons.com	68.66.226.98	ESC-PONS.COM		1 USA
http://meusacessocaixas.com/Error/	20.40.202.21	MEUSACESSOCAIXAS.COM		1 USA
http://www.saisoncard-info.com/	153.122.190.3	SAISONCARD-INFO.COM		1 JAPAN
http://roommates.roomvoid.com//TO/abuse@microsoft.com	23.238.16.58	ROOMVOID.COM		5 USA
http://check.app-hayes-fbbsmsnhd.fwwwu4gyil-zng4p7vv54dp.p.runcloud.link/	138.68.166.111	runcloud.link		2 GERMANY
https://bit.ly/3jXa3WU?Confirmation	67.199.248.11	googledomains		6 USA
https://aroueaom.zh2fh.cn/	107.150.7.48	zh2fh.cn		1 USA
https://aroueaom.d5f1s.cn/	155.94.182.235	d5f1s.cn		1 USA
https://eternal.servehttp.com/?checkid=a@abc	8.23.224.108	SERVEHTTP.COM		22 USA
https://12mo7.trk.elasticemail.com/tracking/click?d=vmSxXbiXG-	79.137.96.34	ELASTICEMAIL.COM		12 FRANCE
https://myaccount2.godaddydites.com/	72.167.191.83	N/A		3 USA
https://check.app-hayes-fbbsmsnhd.fwwwu4gyil-zng4p7vv54dp.p.runcloud.link/	191.235.228.35	Wild West Domains, LLC		1 BRAZIL
https://www.etc-meisa1.xyz	182.16.74.146	Web Commerce Communicatio		1 USA
https://discord.giveaway.com/steam	95.181.163.46	Alena Pavlenko		3 USA
https://recovery-pagesconfrim.ga/	162.241.194.21	N/A		3
https://linktr.ee/btesecurebts365	151.101.130.133	lnstra Corporation Pty Ltd		7 USA
https://hugobosspolska.top/	196.196.223.107	NameSilo, LLC		1 Estonia
https://mercari.aruv.shop/	155.94.184.151	NameSilo		1 USA
https://amzo.jp.tyaxing.net.cn/	45.87.230.148	北京阿尊科技有限公司		1 CHINA
https://youdubashenome.bounceme.net/	8.23.224.108	Dan Durrer		20 USA
	average		6,225806452	

Рисунок 2.15 - Приклад фішингових URL

Як видно з рис.2.15 вище, на цьому рисунку показані фішингові URL-адреси разом з обраними ознаками. У ході збору даних URL, які показані на рис.2.15., було визначено, що збір таких URL посилань займає велику кількість часу, з урахуванням того, що всі ці URL посилання були зібрані використовуючи інструменти "Whois", "Geoip", та "Ping".

URL	IP address	Registrar	Lifetime	Country
https://www.testteacher.online/	142.250.74.147	hostinger.com		2 INDIA
https://www.delectablesbydanette.com/	35.208.54.48	godaddy.com		7 USA
https://www.google.com	142.250.74.46	MarkMonitor Inc.		31 USA
http://www.cinquantegresnordsouslestropiques.fr/	213.186.33.5	ns18.ovh.net		3 FRANCE
https://facebook.com	157.240.26.35	RegistrarSafe, LLC		31 USA
https://www.pharmacie-moustier.fr/	87.98.154.146	http://www.ovh.com		3 FRANCE
https://www.yahoo.com	98.137.11.163	MarkMonitor Inc.		28 USA
http://www.fahrrad-werkzeug.com/	104.165.95.142	son ghe		1 CHINA
https://www.reddit.com	151.101.193.140	MarkMonitor Inc.		17 USA
https://www.feber.se	46.21.97.151	ns1.loopla.se		16 SWEDEN
http://www.tecnicaschaar.com	217.160.0.22	http://ionos.com		12 SPAIN
https://www.sweclockers.com	158.174.189.9	Ascio Technologies, Inc. Danmar		22 SWEDEN
https://www.swagelok.com/	23.78.57.222	Eurodns S.A.		29 USA
https://www.twitter.com	104.244.42.65	CSC Corporate Domains, Inc.		22 USA
https://www.bloomberg.com	151.101.193.73	CSC Corporate Domains, Inc.		29 USA
https://www.ovh.co.uk/	198.27.92.7	www.nominet.uk		26 USA
https://www.canvas.com	172.67.7.29	GoDaddy.com, LLC		26 USA
https://www.blogger.com	142.250.74.73	www.icann.org		23 USA
https://www.unsplash.com	151.101.65181	eNom, LLC		9 USA
https://www.myspace.com	63.135.90.70	Network Solutions, LLC		27 USA
		average		18,2

Рисунок 2.16 - Приклад нефішингових URL

Як видно з рис. 2.15. і рис. 2.16, дані сайти були зібрані з бази даних перевірених фішингових посилань, а саме з "Phishtank.org", де нові антивірусні компанії та окремі дослідники завантажують нові фішингові посилання, і вони перевіряються. Також, в даному випадку було пораховано середнє арифметичне часу життя, як видно із зображень вище, середнє арифметичне часу життя фішингових URL дорівнює 6,225806452, а середнє арифметичне часу життя нефішингових URL дорівнює 18,2. Деякі доброякісні URL-адреси були зібрані за допомогою сайту "Alexa.com", даний сайт відстежує популярні веб-сайти.

В даному випадку, різноманітність технік запуску фішингових сторінок обмежена можливостями хостингу і реєстратора доменних імен, за останні роки мало що змінилося, тільки те що сертифікати стали використовуватися частіше. Зловмисники все ще реєструють доменні імена які співзвучні атакованому сервісу, на максимально короткий термін, зазвичай на один рік. Веб-сторінка розміщується локально або використовується зламаний сервер. Більшість атак зазвичай використовують протокол HTTP без TLS і сертифіката. Тобто, в даному дослідженні було достатньо враховувати наступні атрибути, доменне ім'я, дані Whois, геолокацію, тип мережевого протоколу (HTTP або HTTPS) і, також, наявність сертифіката з його параметрами.

Після того як були зібрані дані URL, а також розрахували середнє арифметичне їхнє життя, наступний етап дослідження, це проведення дослідницького аналізу даних. Для цього, на основі даних зібраних URL, були зроблені діаграми частот. Нижче на рис.2.17 показана частотна діаграма країн для фішингових URL.

Phising - country

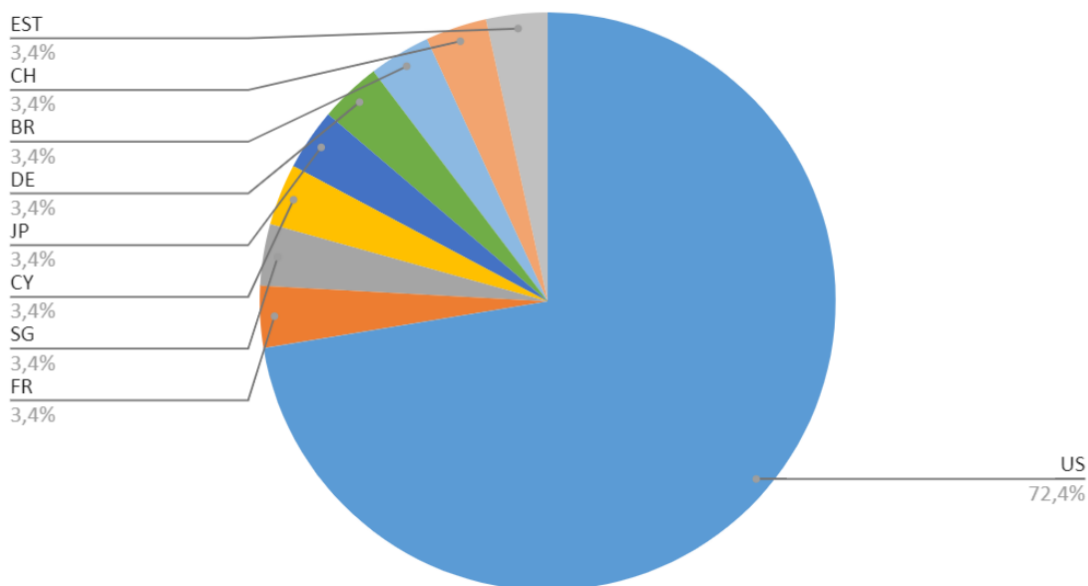


Рисунок 2.17 - Частотна діаграма країн для фішингових URL

З рис.2.17 вище видно, що 72,4% фішингових URL знаходяться у США, тобто це означає, що найбільша кількість фішингових URL має саме ця країна. Також, із рис.2.17 видно, що решта країн мають 3,4% фішингових URL, це досить низький показник, якщо порівнювати результат знайдених URL, які знаходяться в США. Тобто можна відзначити, що на даний момент найбільше доменів знаходяться у США. Також, на рис.2.18 нижче представлена частотна діаграма часу життя для фішингових URL.

Phising - lifetime

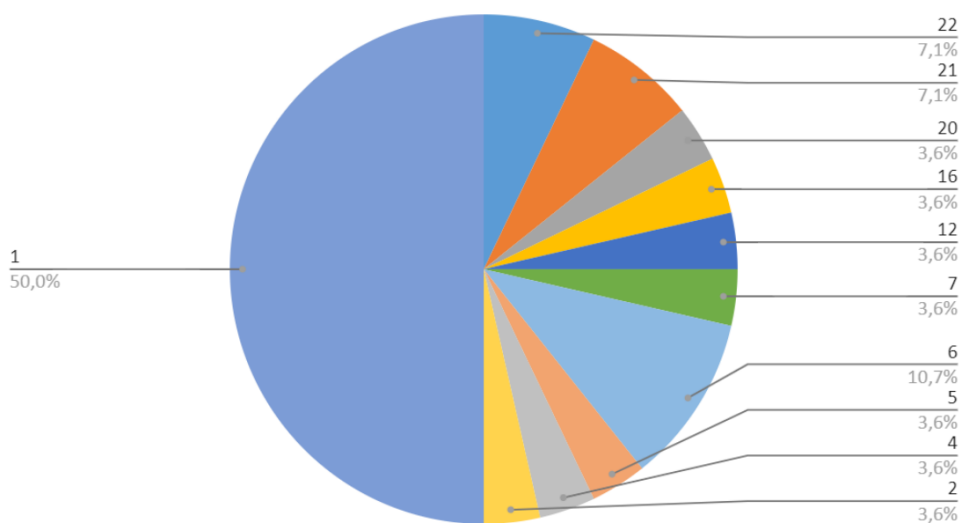


Рисунок 2.18 - Частотна діаграма часу життя для фішингових URL

З рис.2.18 вище видно, що велика кількість фішингових URL-адрес, а саме 50%, мають час життя тільки на один рік. Тобто, дана частотна діаграма показує, що більшість URL реєструють лише на один рік, і лише мала кількість фішингових доменів реєструється на більший термін. Нижче на рис.2.19 представлена частотна діаграма імені реєстратора для фішингових URL.

Phising - registrar

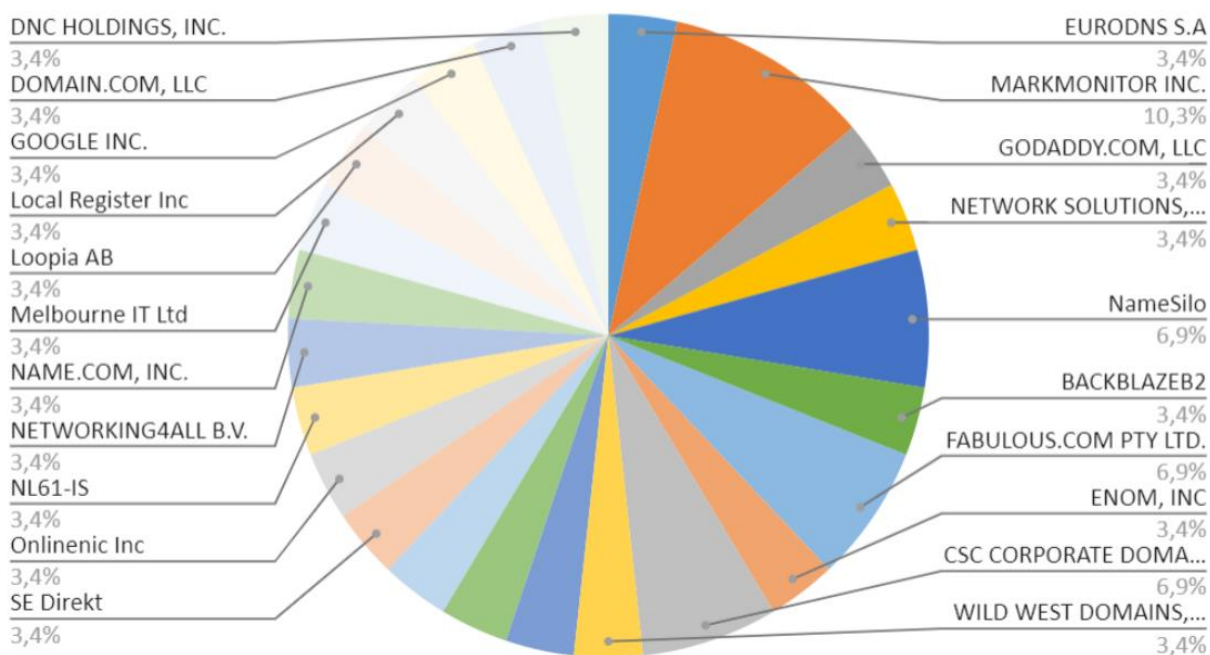


Рисунок 2.19 - Частотна діаграма імені реєстратора для фішингових URL

З рис.2.19 вище видно, що одне ім'я реєстратора має частоту зустрічі яка дорівнює 10,3%. Також на рис.2.19 видно, що тільки три імені реєстратора мають частоту зустрічі яка дорівнює 6,9%. Тобто, це означає те, що в даному випадку, найбільш поширене ім'я реєстратора це те ім'я, яке має частоту зустрічі яка дорівнює 10,3%. Нижче на рис.2.20 показано частотну діаграму країн для нефішингових URL.

Clean - country

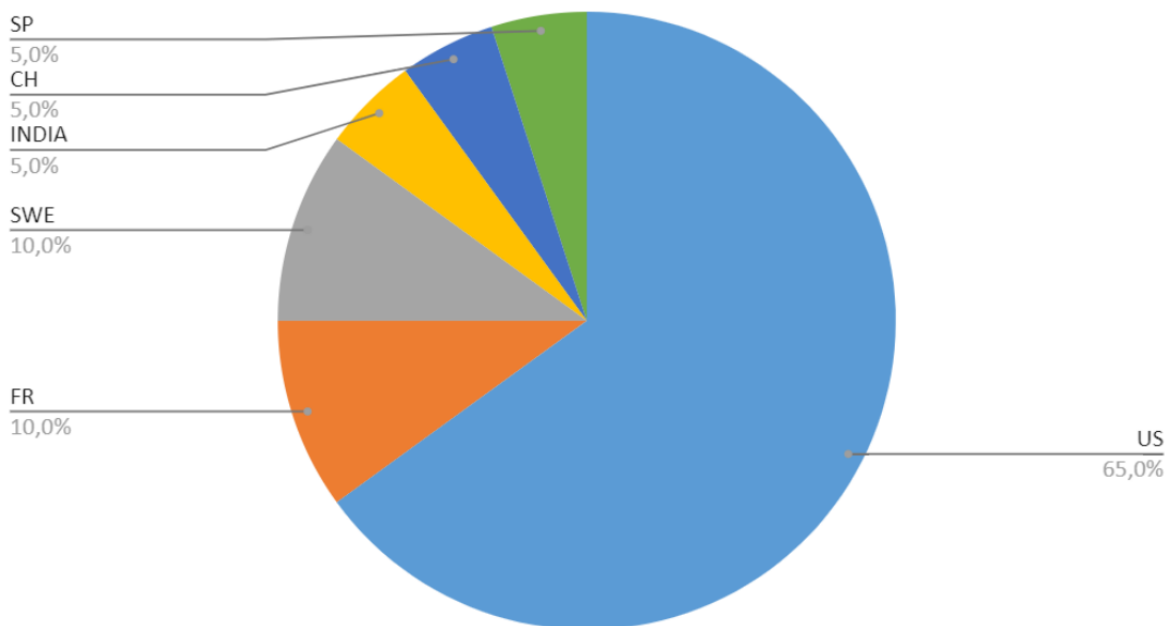


Рисунок 2.20 - Частотна діаграма країн для нефішингових URL

З рис.2.20 вище видно, що незважаючи на те, що в даному випадку це частота зустрічі країн нефішингових URL, країна США все ще має більший відсоток зустрічі. Якщо ж порівнювати частоту зустрічі країн нефішингових URL, і частоту зустрічі країн фішингових URL, то у випадку нефішингових URL, відсоток частоти зустрічі інших країн буде більшим, оскільки в даному випадку це чисті сайти, і також тому що нефішингових URL було зібрано трохи менше ніж фішингових URL. На рис.2.21 нижче представлена частотна діаграма часу життя для нефішингових URL.

Clean - lifetime

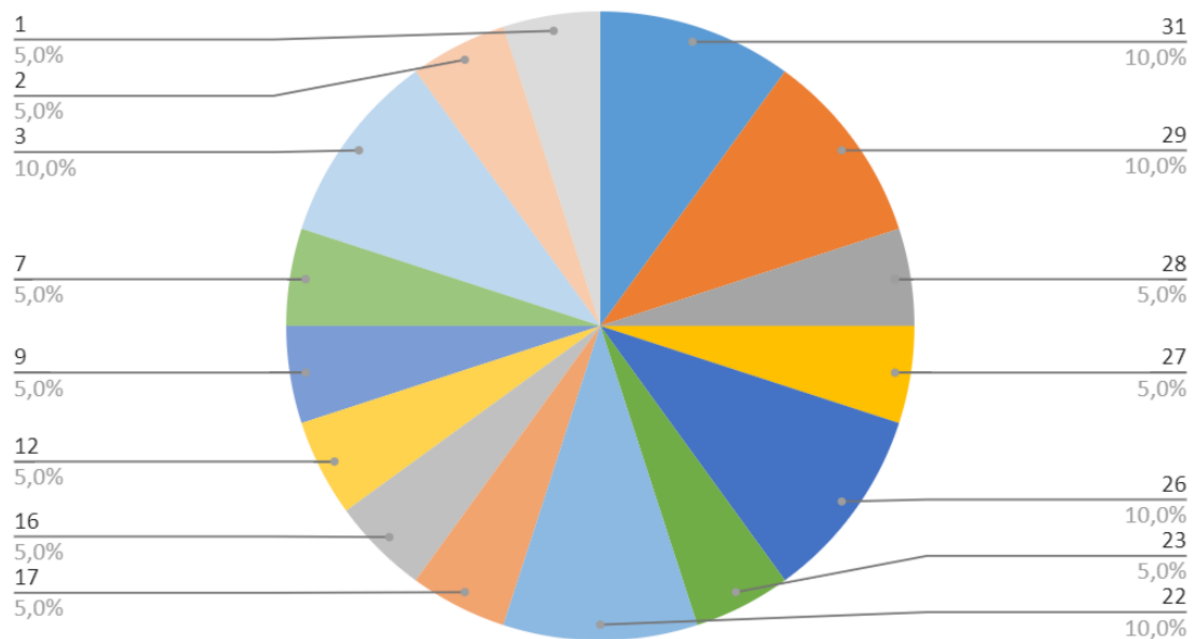


Рисунок 2.21 - Частотна діаграма часу життя для нефішингових URL

З рис.2.21 вище видно, що якщо порівнювати цю частотну діаграму з частотою зустрічі часу життя фішингових URL, то в даному випадку, велика кількість нефішингових URL реєструють на більш тривалий час життя. Також, із рис.2.21 можна відмітити, що тільки невелика кількість нефішингових URL-адрес зареєстровані на один рік або два роки, на відміну від частоти зустрічі часу життя фішингових URL-адрес. Нижче на рис.2.22 представлено частотну діаграму імені реєстратора для нефішингових URL.

Clean - registrar

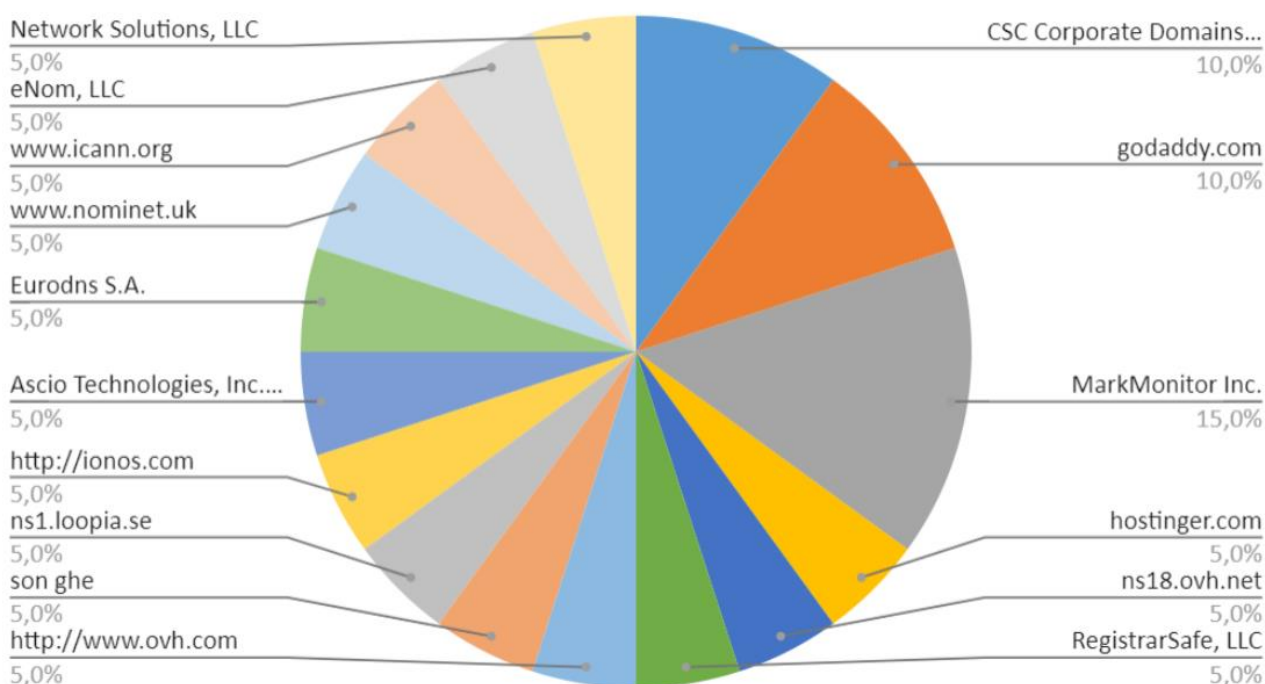


Рисунок 2.22 - Частотна діаграма імені реєстратора для нефішингових URL

З рис.2.22 вище видно, що в даному випадку є схожі імена реєстратора як і у випадку з частотною діаграмою імені реєстратора для фішингових URL. З рис.2.22 також видно, що найбільший відсоток зустрічей має ім'я реєстратора "MarkMonitor Inc.", схоже ім'я реєстратора також має великий відсоток зустрічей у випадку з фішинговими URL-адресами. В даному випадку, на діаграмах були показані частоти зустрічальності значень трьох атрибутів, а саме, імені реєстратора(Registrar), періоду реєстрації домену (Lifetime), і геолокації хоста (Country) для нефішингових і фішингових URL-адрес. Зробивши аналіз даних, можна помітити, що в даному випадку, атрибути реєстратор і країна буде складно використовувати для класифікації, оскільки серед найбільш часто зустрічаються значень багато збігів. Відмінності в даних двох атрибутах більш тонкі і можуть бути виявлені у великих наборах даних. Але атрибут реєстрації домену, який відповідає періоду, на який зареєстрований домен, показує, що серед фішингових повідомлень є тенденція реєструвати домени на набагато маленький термін. Наприклад, з рис.2.18 можна помітити, що 50% фішингових доменів були зареєстровані на один рік, на відміну від нефішингових доменів, це можна помітити переглянувши рис.2.21, з даного зображення видно що деякі домени реєструють на 29 років., і всього лише 5% реєструють на 1 рік.

Що стосується TLS, то більшість фішингових URL не використовують протокол HTTPS. Тобто, це один з основних критеріїв, який можна використовувати для виявлення фішингових атак. Зловмисники можуть реєструвати сертифікати для підроблених компаній, наприклад "LockerGoga", файли якого були підписані цифровими сертифікатами, виданими "Sectigo RSA Code Signing CA" підробленим компаніям "Alina Ltd, Kitty's Ltd", "Mikl Limited" і "AB Simba Limited".

У цьому дослідженні було зроблено порівняння 30 фішингових URL з 20 нефішингових URL, це дослідження можна віднести до ручного аналізу фішингових повідомлень. В ході порівняння було виявлено, що зловмисники зазвичай розміщують сайти в популярних доменних зонах таких як, .com, .net, .org, .info, а також в зонах, де придбання домену дешево, таких як, .site, .xyz, .online, .top, club, .live. Термін життя підроблених доменів в два рази менше, ніж у нефішингових URL. На першому місці за частотою розміщення фішингових і нефішингових URL знаходяться США. Це тому, що в даному дослідженні було зібрано лише невелику кількість фішингових повідомлень. В результаті в даному дослідженні були отримані імена реєстраторів. Якщо ім'я реєстратора добре відомо, швидше за все, це нефішинговий веб-сайт, а якщо воно невідомо, швидше за все, це фішинговий веб-сайт. Висновок, до якого призвело дане дослідження, полягає в тому, що цих даних недостатньо для того щоб отримати остаточне уявлення про те, в якій країні знаходиться найбільше фішингових сайтів. Можна спробувати отримати кращу і більш повну картину всього цього, зібравши більше даних з набагато більшого набору даних. Беручи до уваги, що з великою кількістю даних можна було б побачити більш чіткі відмінності між фішинговими і чистими фішинговими повідомленнями на основі імен реєстраторів і країн. Виходячи з даного висновку, в подальшому розділі буде обговорюватися питання автоматизації даного процесу, і будуть запропоновані процедури автоматизації процесу обробки і виявлення фішингових повідомлень.

Як доповнення до дослідження, можна запропонувати зважений алгоритм, який обчислює, чи є URL нефішинговим або фішинговим. Тоді як зважений алгоритм працює наступним чином, існує набір атрибутів, що визначають вагу запису. Чим більше атрибутів відповідає запису, тим більша вага присвоюється цьому запису. Чим вище ступінь виконання ознак, тим більша вага присвоюється запису. Наприклад, якщо подивитися на час життя веб-сайтів, чим коротше час життя веб-

сайту, тим більшу вагу буде присвоєно веб-сайту. І різні ваги в залежності від того, в якій країні розміщений веб-сайт і так далі.

Різні атрибути дають різні ваги, час життя веб-сайтів дає вагу, так що чим менше час життя веб-сайту, тим більшу вагу йому присвоюється. Це тому, що більшість фішингових веб-сайтів, здається, мають дуже короткий термін життя.

Ім'я реєстратора дає вагу в залежності від того, наскільки відомий реєстратор. Якщо реєстратор сайту відомий, присвоюється вага зменшується. Якщо ім'я реєстратора веб-сайту менш відомо, присвоюється вага збільшується.

І нарешті, країна, в якій розміщений веб-сайт, також має значення, однак, згідно з висновками даного дослідження, цей атрибут має найменше значення. Частково це пов'язано з тим, що як фішинговий, так і чистий веб-сайт мають спільні країни і частоту їх появи, а приймаючі країни можуть бути підроблені через VPN. Таким чином, якщо країна, в якій розміщений веб-сайт, менш поширена, веб-сайту присвоюється трохи більшу вагу, якщо країна поширена, веб-сайту присвоюється трохи меншу вагу.

3 РОЗРОБКА ПРОЦЕДУР АВТОМАТИЗАЦІЇ ТА ПРОВЕДЕННЯ МОДЕЛЬНОГО ЕКСПЕРИМЕНТУ З ОБРОБКИ ФІШИНГОВИХ ПОВІДОМЛЕНЬ

3.1 Порівняння алгоритмів машинного навчання та вибір найкращого алгоритму

Мета даного підрозділу, це порівняти існуючі алгоритми, і зрозуміти який алгоритм найбільш краще справляється із завданням виявлення фішингових поси-лань. Завдання тут полягає в тому, щоб точно передбачити, що вийде з великих блоків даних. Для цього випадку, використовується комп'ютерна програма "WEKA". Використання WEKA тут дуже зручно, тому що WEKA дуже швидко дає результат в алгоритмах використання машини. Weka, це цінний інструмент для оцінки методів класифікації та кластеризації. Даний інструмент також реалізує методи оцінки за допомогою свого інструменту Experimenter. Завдяки отриманим результатам можна порівняти алгоритми, які використовуємо. В даному випадку, була відібрана база даних фішингових і нефішингових сайтів, з урахуванням ознак які вже обговорювалися раніше.

Класифікація, це функція інтелектуального аналізу даних, яка присвоює елементам набору даних цільові категорії або класи. Мета класифікації полягає в тому, щоб зробити точні прогнози для кожного значення в даних. У моделі класифікації були розглянуті такі функціональні алгоритми, як Multilayer Perceptron, J48, байєсівські алгоритми та логістична регресія.

Багатошаровий персептрон (Multilayer Perceptron, MLP) - це доповнення до нейронної мережі з прямим зв'язком. Цей алгоритм складається з трьох типів шарів, перший шар - це вхідний шар, другий - вихідний шар, і третій - прихований шар. Вхідний шар отримує вхідний сигнал обробки. Необхідне завдання, передбачення та класифікація, виконується вихідним шаром. Довільна кількість прихованих шарів, що поміщаються між вхідним та вихідним шарами, є справжнім обчислювальним механізмом MLP. У MLP дані передаються в прямому напрямку до вихідного рівня, від вхідного рівня.

J48 - даний алгоритм машинного навчання, це алгоритм класифікації дерева рішень з машинним навчанням, такий алгоритм дуже корисний для категоричного

та безперервного вивчення даних. Дерево рішень, це алгоритм навчання на основі дерева, є одним з найбільш часто використовуваних алгоритмів для класифікації інтелектуального аналізу даних. Алгоритм, що має деревоподібну модель, приймає деякі рішення для досягнення бажаних результатів. Але він містить деякі умовні твердження, щоб прийти до такого висновку [20]. Цей алгоритм можна використовувати у всіх деревах рішень, задачах класифікації та регресії. Дерево рішень використовується для поділу набору даних на ще більш дрібні набори із застосуванням певних правил. Іншими словами, в даних, розділених на дрібні частини, робляться ще простіші кроки. Крім того, дерево рішень, яке можна візуалізувати, набагато простіше для розуміння.

Логістична регресія - це завдання категоризації залежних змінних, що використовується в задачі лінійної класифікації. Мета тут полягає в тому, щоб отримати аналіз незалежних змінних, які використовуються в наборі даних. Подібно феномену всіх алгоритмів регресії, цей алгоритм являє собою прогнозуючий аналіз. Результат, який буде отримано, береться з бінарної змінної. Ці змінні дорівнюють 0 (правда) і 1 (істина). Для даного порівняння, було досліджено кілька алгоритмів з урахуванням всіх можливих налаштувань параметрів.

Bayes Network learning (Bayes Net) - даний алгоритм це навчання байєсівської мережі з використанням різних алгоритмів пошуку та показників якості. Тобто базовий клас для класифікатора байєсівської мережі. Цей алгоритм надає структури даних (структури мережі, умовні розподіли ймовірностей) та засоби, загальні для алгоритмів навчання байєсівської мережі.

Наївний байєсовський алгоритм (NaiveBayes) - це один із простих і ефективніших алгоритмів класифікації. Цей алгоритм дозволяє створювати швидкі моделі машинного навчання, здатні робити швидкі прогнози. Це ймовірнісний алгоритм, тобто даний алгоритм передбачає на основі ймовірності об'єкта.

Для того щоб оцінити роботу алгоритмів, були використані такі метрики як ІПП (Істинний Позитивний Показник) та ЧХР (Частота Хибнопозитивних Результатів), також було використано ІНП (Істинний негативний показник), ЛНП (Ложний негативний показник), точність та Recall. Отже, точність (Accuracy) є одним із загальновідомих заходів, що застосовуються для оцінки. Точність, це відношення правильно визначених фішингових та легітимних ресурсів до загальної кількості ресурсів з огляду на помилки класифікації, як показано у формулі 3.1 нижче:

$$\text{Точність} = \frac{(\text{ІН} + \text{ІП})}{(\text{ІН} + \text{ЛП}) + (\text{ІН} + \text{ПН})} \quad (3.1)$$

де ІН – кількість правильно класифікованих легітимних ресурсів;

ІП – кількість правильно класифікованих фішингових ресурсів;

ЛП – кількість неправильно класифікованих фішингових ресурсів, які класифіковані як легітимні;

ПН – кількість неправильно класифікованих легітимних ресурсів, які класифіковані як фішингові.

Наступна метрика оцінки алгоритмів, це Відгук. Відгук, це відношення ІП до суми ІП та ПН. Відгук обчислюється за формулою 3.2, яка показана нижче:

$$\text{Відгук} = \frac{\text{ІП}}{(\text{ІП} + \text{ПН})} \quad (3.2)$$

Наступна метрика оцінки це ІНП. Дана метрика це відношення ІН до суми ІН і ЛП, ця метрика показана у формулі 3.3 нижче:

$$\text{ІНП} = \frac{\text{ІН}}{(\text{ІН} + \text{ЛП})} \quad (3.3)$$

І остання метрика це ЛНП. ЛНП, це ставлення ПН до суми ПН та ІП. Дана метрика обчислюється за формулою 3.4, яка наведена нижче:

$$\text{ЛНП} = \frac{\text{ПН}}{(\text{ПН} + \text{ІП})} \quad (3.4)$$

Для того щоб зробити порівняння, в даному випадку необхідно використувати набір даних фішингових та нефішингових посилань, з урахуванням деяких основних ознак. Для цього порівняння був обраний набір посилань, який має 11053 значень, і 32 ознаки. Цільові змінні, які є результатом, це значення -1 і 1. 1 це законна URL, і -1 це фішингова URL. Набір даних включає набір даних значень. Після вибору набору даних, цей набір був завантажений у програму "Weka", після цього програмою був проведений результат. Результат порівняння алгоритмів показано на таблиці 3.1 нижче.

Таблиця 3.1 - Порівняння алгоритмів

Метрика	Multilayer Perceptron	NaiveBayes	Bayes Net	J48	Logistic Regression
ІПП	0,970	0,922	0,950	0,954	0,948
ЧХР	0,072	0,117	0,099	0,057	0,098
ІНП	0,928	0,883	0,901	0,943	0,902
ЛНП	0,030	0,078	0,050	0,046	0,052
Точність	95.133	90.465	92.8352	94.8978	92.7628
Відгук	0,951	0,905	0,928	0,949	0,928

В даному випадку, дивлячись на таблицю 3.1. можна помітити, що найкращий глобальний результат був досягнутий алгоритмом багат шарового перцептрона. Нейронна мережа дала точність 95.133% видавши 7 хибно позитивних результатів (тобто законні повідомлення, які позначені як фішингові повідомлення) і 11 помилково негативні результати (фішингові повідомлення, які позначені як законні повідомлення). Нейронна мережа створює неперевірювану модель, і дану модель можна застосувати до антифішингового фільтру. Що стосується алгоритмів дерева рішень, то J48 показав результат з точністю 92.8352%. У випадку байєсівських імовірнісних алгоритмів був використаний байєсівський мережевий Класифікатор і наївний байєсівський Класифікатор. З точки зору результатів, як наївний байєсівський, так і байєсівський чистий алгоритми працювали досить добре, незважаючи на те, що кількість помилкових спрацьовувань була вищою, ніж у алгоритму J48 і алгоритму багат шарового перцептрона. В даному випадку, найгірший результат показав наївний байєсівський Класифікатор, показавши точність в 90.465%. Що стосується алгоритму логістичної регресії, то з таблиці можна помітити, що даний алгоритм за значеннями майже схожий на алгоритм Bayes Net, але точність у даного алгоритму гірше. Проте, проаналізувавши дані результати, можна припустити, що для подальшої роботи, і розробки припущень щодо процедур автоматизації, алгоритм який показав найвищу точність може стати в нагоді в подальшій роботі.

3.2 Обговорення та припущення щодо розробки процедур автоматизації обробки та виявлення фішингових повідомлень

Тепер, коли обраний алгоритм, необхідно дати можливі припущення щодо автоматизації процесу обробки фішингових повідомлень. У даній роботі буде використовуватися метод машинного навчання, оскільки використання даного методу дає більше точності, і більш кращий результат. Машинне навчання - це галузь штучного інтелекту (AI) та комп'ютерних наук, яка зосереджується на використанні даних та алгоритмів, щоб імітувати те, як люди навчаються, поступово підвищуючи точність. Підхід до машинного навчання починається зі збору набору даних. Набір даних включає різні функції та ярлики (безпечні та шкідливі) для великої кількості URL-адрес. Потім Класифікатор навчається на зібраному наборі даних. Основне припущення полягає в тому, що шкідлива URL-адреса має суттєво інший розподіл функцій, ніж безпечна URL-адреса. В результаті хороший підхід до машинного навчання повинен вміти розрізняти безпечні та шкідливі URL-адреси на основі цих функцій.

Виявлення фішингових повідомлень, це проблема класифікації, тому в даному випадку потрібні помічені дані, які мають зразки як фішингові домени, так і законні домени на етапі навчання. Набір даних, які будуть використовуватися для навчання це дуже важливі дані для створення успішного механізму виявлення. В даному випадку, необхідно використовувати зразки, класи яких точно відомі. Простими словами, це означає що зразки, помічені як фішинг, повинні бути абсолютно точно визначені як фішинг, а зразки які позначені як законні, повинні бути визнані законними. У поганому випадку, даний спосіб не буде працювати коректно, якщо ж будуть використані зразки, в яких немає впевненості.

Для цього випадку створюються деякі загальнодоступні набори даних для фішингу, за допомогою "PhishTank". Дані джерела даних зазвичай використовуються в академічних дослідженнях. Після того як були отримані необроблені дані для фішингових і законних сайтів, наступний крок це обробка цих даних і витяг з них значущої інформації для виявлення фішингових повідомлень. Набір даних, який буде використовуватися для машинного навчання, повинен складатися з цих функцій. Тобто, в даному випадку необхідно обробити необроблені дані, зібрані з "Alexa", "Phishtank" або ж інших ресурсів даних, і після цього створити новий на-

бір даних для навчання нашої системи алгоритмам машинного навчання. Значення ознак повинні бути обрані відповідно до потреб і цілей.

Фішингові домени є однією з проблем класифікації. Це означає, що необхідні помічені екземпляри для створення механізму виявлення. У данному випадку є два класи, це фішингові повідомлення і легальні повідомлення.

Коли йде обчислення функції, які були вибрали для потреб і цілей даної роботи, то набір даних виглядає так, як показано на рис. 3.1 нижче.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Index,UsingIP,LongURL,ShortURL,Symbol@,Redirecting//,PrefixSuffix-,SubDomains,HTTPS,DomainRegLen,Favicon,NonStdPort,H											
2	0	1	1	1	1	-1	0	1	-1	1	1	-1
3	1	1	0	1	1	1	-1	-1	-1	1	1	-1
4	2	1	0	1	1	1	-1	-1	1	1	1	-1
5	3	1	0	-1	1	1	-1	1	1	1	1	0
6	4	-1	0	-1	1	1	-1	1	1	1	1	-1
7	5	1	0	-1	1	1	1	-1	1	1	1	-1
8	6	1	0	1	1	1	-1	-1	1	1	1	-1
9	7	1	0	-1	1	1	1	-1	1	1	1	1
10	8	1	1	-1	1	1	-1	1	1	1	1	0
11	9	1	1	1	1	-1	0	1	1	1	1	1
12	10	1	1	-1	1	1	-1	1	1	1	1	-1
13	11	-1	1	-1	1	-1	0	1	1	1	1	-1
14	12	1	1	-1	1	1	-1	1	1	1	1	-1
15	13	1	1	-1	1	1	-1	1	1	1	1	1

Рисунок 3.1 - Приклад набору даних з використанням ознак

Як видно з рис.3.1, набір даних це набір чисел, і набір ознак. Цей приклад має дуже великий набір даних, що складається з безлічі чисел. В даному прикладі було вибрано 32 ознак, після чого вони були обчислені. Таким чином, це приклад того як створюється набір даних, який буде використовуватися на етапі навчання алгоритму машинного навчання.

Розглянемо які можуть бути варіанти щодо процедур автоматизації обробки та виявлення фішингових повідомлень. Першим припущенням розробки процедур автоматизації може бути програма, чи скрипт написаний мовою програмування Python. У разі, якщо ж автоматизація буде зроблено з допомогою скрипта, необхідно розробити скрипт з урахуванням алгоритму, і навіть з урахуванням ознак фішингових повідомлень. Даний скрипт вирішить задачу автоматизації, а також дозволить підвищити точність, час та ефективність визначення та обробки фішингових повідомлень.

Для успішного застосування даного скрипта, необхідна наявність даного скрипта у користувача який хоче провести обробку фішингових повідомлень. Що

стосується ознак, то в даному випадку функції будуть витягнуті з URL-адрес за допомогою "Features.py", які складаються з декількох функцій.

Для того щоб витягти функцій потрібно відправка HTTP-запитів і виконання іншої частини очищення з використанням bs4 і отримання даних з домену з використанням бази даних whois, для цього також потрібно буде створити окремий клас, де можна буде використовувати його багатопотокові можливості, тобто розпаралелювання або прискорення запитів. Тобто, спочатку починається вилучення функцій, які допомагають визначити характер URL-адреси, за яким слід попередня обробка даних, така як обробка пропущених значень, видалення постійних функцій, і видалення викидів. Після цього слід дослідницький аналіз даних, щоб дізнатися більше про функції, їх важливості, зв'язку з метою. Після даної дії йде розробка ознак, яка пізніше може допомогти моделі більш точно виявити закономірності. Після цього вже йде масштабування і моделювання функцій. В даному випадку, дослідницький аналіз даних працює таким чином.

Спершу йде поділ набору даних на навчений і тестовий набір, щоб уникнути витоку даних. Тільки набір даних навчання повинен бути присутнім в навчанні моделі як інформація, а не весь набір даних, це потрібно для того щоб уникнути дисперсії. Якщо під час навчання моделі задіяний весь набір даних, то він дає хороший прогноз для цього набору даних, але не для нових запитів, що надходять під час виробництва. Таким чином, в повній програмі, у нас може бути код з ознаками, і код з виявленням фішингових повідомлень, крім цього, в даній програмі також може бути і набір даних необхідний для машинного навчання.

Ще одним способом вирішити задачу автоматизації обробки фішингових повідомлень можна шляхом розробки коду мовою Java, а також шляхом розробки коду на Python додавши при цьому взаємодію з веб-додатком, на якому і можна буде здійснювати обробку фішингових повідомлень. Тим не менш, оскільки єдиний спосіб вирішити задачу автоматизації обробки фішингових повідомлень є спосіб написання скрипта, або написання коду, то в даному випадку було прийнято рішення, що в наступному підрозділі буде запропоновано скрипт мовою Python, який дозволяє обробляти фішингові повідомлення, і виявляти фішингове повідомлення чи ні.

3.3 Побудова моделі та розробка процедури автоматизації обробки та виявлення фішингових повідомлень

Оскільки рішення яке було прийнято в попередньому підрозділі це написання скрипта, то зараз буде детально показана робота запропонованого скрипта, який дозволяє обробляти та визначати фішингові та нефішингові повідомлення. Інструменти, які були використані для написання скрипта, наведені нижче:

- pyCharm;
- colab.research.google.com;
- excel;
- база наборів даних Kaggle.

У цьому списку три інструменти, які використовувалися при написанні скрипта. PyCharm є інтегрованим середовищем розробки для мови програмування Python. В даному випадку цей інструмент використовувався на початку написання скрипта. У ході написання, з'ясувалося, що написання скрипта використовуючи даний інструмент є складним, оскільки є складнощі з налаштуванням CUDA і встановленням потрібних бібліотек, також є проблеми з завантаженням набору даних. Тому, було прийнято рішення, що для запуску та подальшого продовження написання коду буде використовуватися "colab.research.google.com", оскільки в "Google colab" бібліотеки йдуть звідти, а також немає труднощів із завантаженням набору даних. "Google colab" є продуктом Google, який дозволяє писати і виконувати код Python через браузер, цей інструмент дуже добре підходить для машинного навчання та аналізу даних. Веб-сайт "<https://www.kaggle.com/>" використовувався для того, щоб мати доступ до наборів даних. Kaggle це платформа для фахівців за даними. Kaggle дозволяє користувачам співпрацювати з іншими користувачами та публікувати набори даних.

Тепер розглянемо роботу запропонованого скрипту. Основні бібліотеки, що використовуються в даному скрипті, це такі бібліотеки як "Sci-kit Learn", "BeautifulSoup", "tdlextract", "pythonssl", та "matplotlib". Набір даних, який використовується в даному випадку, має 420465 спостережень, даний набір має безліч url, який позначений поганими, і хорошими. У наборі даних перший стовпець, це url, який представляє фактичні посилання, а другий стовпець це мітка, яка містить як погані, так і хороші URL-адреси. Цей набір даних було взято з веб-сайту "<https://www.kaggle.com/>", цей веб-сайт має безліч наборів даних. Для початку ви-

користовується пакет "Pandas" для завантаження набору даних. Після цього завантажуються набір даних. Після успішного завантаження набору даних використовується команда, щоб переглянути структуру набору даних. На зображенні вище набір даних має два стовпці. Після цього йде процес очищення набору даних, щоб підготувати його до використання моделлю під час навчання.

Очищення набору даних є видаленням шуму з набору даних. Тобто шум, це зайві символи в текстових даних, розділові знаки, а також повторювані слова. Видалення шуму з набору даних дозволить моделі зосередитись лише на найважливішій інформації в наборі даних. Це також підвищить і продуктивність моделі. Після такого очищення модель зможе робити більш точні прогнози. У цьому випадку в процесі очищення спочатку розділяються тексти, а потім видаляються повторення набору даних. Після цього з кожного URL йде видалення "com". Таким чином, була створена функція Python для очищення набору даних.

Функція називається "createTokens". Ця функція потрібна для поділу тексту на косу межу, тире, а також точку. Це гарантія того, що в даному випадку буде чистий текст. Ця функція видаляє зайві слова в тексті. Наприкінці функція видаляє слово "com", використовуючи метод `total_Tokens.remove('com')`, після чого функція повертає чистий текст. Далі, після даного очищення йде додавання ознак, і міток у набір даних. У даному випадку, ознаки, це унікальні точки даних у даному наборі даних, дані точки використовуються як вхідні дані для моделі під час навчання. Ознаки є стовпцем URL, який є вхідним стовпцем. Також ознаки додаються використовуючи код з ознаками, де вказані такі ознаки як:

- довгий URL;
- використання IP-адреси;
- перенаправлення з використанням "///";
- url-адреса із символом "@";
- піддомен, а також кілька піддоменів;
- вік реєстрації домену;
- фавікон;
- наявність нестандартного порту;
- посилання у тегах;
- наявність SSL;
- надсилання інформації використовуючи електронну пошту;
- неправильна URL-адреса;

- префікс або суфікс;
- переадресація сайту;
- вимкнення правої кнопки миші;
- наявність токена "HTTPS" у доменній частині адреси;
- спливаюче вікно;
- фрейм.

Якщо розглядати мітки, то мітка є результатом моделі після прогнозування. Результат представлений за допомогою шпальти міток. Результат моделі може виявитися поганим, або хорошим.

Після цього йде процес побудови моделі. Для початку цього процесу, спочатку йде додавання потрібних пакетів машинного навчання. Функції даних пакетів описані нижче.

1) Багатошаровий перцептрон, це алгоритм навчання Scikit, який використовуватиметься для навчання моделі. Даний алгоритм дозволить моделі зрозуміти закономірності та відносини у наборі даних. Після цього, дана модель отримає досить хороші знання та розуміння, яке необхідно для прогнозування фішингових повідомлень.

2) `train_test_split` - дана функція, це вибір моделі Sklearn для поділу масивів даних на два підмножини, тобто дані навчання, навчальні дані, і на тестові дані.

3) `TfidfVectorizer` - цей пакет робить можливим для моделі розуміння текстових даних та управління ними. Текст є досить великою проблемою для машин, машини що неспроможні споживати текст у необробленому вигляді. Тому в цьому випадку йде перетворення тексту на вектори чисел, які машина зможе прочитати, а також зрозуміти.

4) `TfidfVectorizer` використовується для того, щоб зробити перетворення необроблених текстових даних у вектори чисел, дані числа представляють вихідний текст. Цей текст перетворюється на основі частоти появи кожного слова у текстових даних.

У рядку `"vector = TfidfVectorizer(tokenizer=createTokens)"` йде перетворення текстових даних на вектори чисел. Тобто йде конвертація тексту за допомогою `TfidfVectorizer`, а також передача `createTokens` як параметра. Функція `"createTokens"` використовується для чищення тексту. Після цього в рядку `"X =`

`vector.fit_transform(urls_spisok)"` йде процес збереження векторів чисел у нову змінну.

Наступний етап, це розділення набору даних. В даному випадку використовується `"train_test_split"`, щоб розділити набір даних на два набори. Один набір даних для навчання моделі, другий набір даних буде для тестування моделі. У рядку де `"train_test_split"` використовується `"test_size=0.2."`, це співвідношення, яке застосовується при розділенні набору даних. Дане співвідношення гарантує, що 80% набору даних використовуватиметься для навчання моделі, а решта 20% використовуватиметься для тестування моделі. Наступний етап, це побудова моделі. Для побудови моделі використовується алгоритм Багатошаровий перцептрон, ініціалізується даний алгоритм у рядку `"logregs = MLPClassifier(max_iter=1000)"`. йде процес заповнення алгоритму. Залежно від алгоритму і набору даних, заповнення може тривати досить довго, а може тривати й швидко. Наприклад, при перевірці алгоритму дерева рішень, на заповнення пішло чотири години. Все це залежить від набору даних і обраного алгоритму після заповнення, буде отриманий висновок, що модель навчена з потрібними параметрами, в даному випадку це кількість ітерацій. Після того, як було здійснено розрахунок точності, слід етап виявлення фішингових повідомлень. Для того щоб зробити виявлення використовується кілька URL-адрес. після обробки посилання, і після того, як дана модель класифікувала, чи є URL-адреса фішинговою чи ні, виходить результат у вигляді тексту який визначає фішингове посилання чи ні. Виявлення відбувається у даних рядках `"X_detect = vector.transform(X_detect),Novoe_detect= logregs.predict(X_detect)"`. У даному рядку, використовується метод `"vector.transform"` необхідний для перетворення тексту у вектори чисел, після цього застосовується метод `"logregs.predict"`, для того щоб робити прогнози. Далі, рядок `"print(Novoe_detect)"` використовується для виведення результату.

Далі йде процес порівняння алгоритмів. Для початку йде завантаження набору даних, потім використовуються такі алгоритми як `"Multilayer Perceptron"`, `"NaiveBayes"`, `"J48"`, `"Logistic Regression"`. Для початку, у рядку `"train_test_split"` також йде розділення даних, і після цього йде заповнення. При порівнянні використовується такий самий спосіб як і при виявленні фішингових повідомлень, тобто спочатку йде поділ даних, потім побудова моделей, заповнення алгоритму, і показ точності алгоритму на основі набору даних, це показується в подібних рядках коду `"Tocnost_NaiveBayes = accuracy_score(Y_test, y_pred ), print("Точність`

Naive Bayes=", accuracy_score(Y_test, y_pred))". Після того, як точність всіх алгоритмів була розрахована, то в рядку "import matplotlib.pyplot as plt" йде підключення потрібного пакета для побудови графіка з результатами порівняння. У рядку "fig = plt.figure()" йде побудова графіка. Далі, у рядку "ax = fig.add_axes(нуль, нуль, один, один)" йде додавання до графіка області Axes. Після цього йде додавання написів, а також результату точності, і у рядку "plt.show()" йде виведення графіка з результатами порівняння.

Що стосується ознак, то в цьому випадку, використовується безліч функцій, які є ознаками. Усередині цих функцій йде умова if, і повернення значень 0, 1, тобто після того, як всі функції виявлені, створюється вхідний вектор, в якому 1 показує, що функція присутня, а 0 показує, що функція відсутня. Після всіх функцій, у рядку "def main(url):" йде функція, яка перевіряє наявність ознак. Нижче на рис.3.2, показано процес побудови моделі, та результати точності побудованої моделі.



▼ Поділ набору даних для навчання та тестування у співвідношенні 80 на 20

```
[45] X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Побудова моделі

```
logregs = MLPClassifier(max_iter=1000)
logregs.fit(X_train, y_train)
```

MLPClassifier(max_iter=1000)

```
[47] # Точність моделі
print("Точність моделі=", logregs.score(X_test, y_test))
```

Точність моделі= 0.9841722854458754

Рисунок 3.2 - Побудова моделі та результати точності

Як видно з рис.3.2 вище, показник точності становить 98,417%. Можна сказати, що в даному випадку це дуже хороший показник точності, цей показник означає, що дана модель була добре навчена. Тобто результат показує, що дана модель багато чому навчилася з набору даних на етапі навчання, і цю модель можна використовувати для виявлення фішингових повідомлень. Також нижче на рис.3.5, показано результат розробленої процедури автоматизації виявлення фішингових та нефішингових URL-адрес.

```

+ Код + Текст
Виявлення посилань, використовуючи нашу модель

[65] X_detect = ["https://forms.zohopublic.eu/brownhgbv144756786/form/LOGINYOURBTACCOUNTCORRECTLY/formperma/3cTeLJHbpAGgKnEgZLQWekwyRok8PCVufqjc1",
               "https://forms.zohopublic.eu/brownhgbv144756786/form/LOGINYOURBTACCOUNTCORRECTLY/formperma/3cTeLJHbpAGgKnEgZLQWekwyRok8PCVufqjc1rQCYnA",
               "next.privat24.ua",
               "https://storageapi2.fleek.co/82c2d680-6f92-48cf-aab9-0830ccb62867-bucket/index.html",
               "https://surendrabudhathoki.com.np/mels",
               "http://webmail-109752.weeblysite.com/"]

[66] X_detect = vector.transform(X_detect)
     Novoe_detect = logregs.predict(X_detect)

[67] print(Novoe_detect)

['bad' 'bad' 'good' 'bad' 'bad' 'bad']

[46] # Можна також використовувати тут базу фейкових сайтів
     X_detect1 = ["www.buyfakebillsonlinee.blogspot.com",
                  "www.unitedairlineslogistics.com",
                  "www.stonehousedelivery.com",
                  "www.silkroadmeds-onlinepharmacy.com",
                  "https://www.youtube.com/",
                  "https://wikipedia.org/",
                  "https://sex.com/",
                  "https://flirty.com/",
                  "https://welsfarto.com/"]

```

Рисунок 3.3 - Виявлення та обробка повідомлень використовуючи модель

На рис.3.3 вище показано розроблений процес автоматизації виявлення фішингових та нефішингових URL. Як видно із рис.3.3, у разі виявлення фішингових і нефішингових URL досить точно, оскільки з рис.3.2. було видно, що показник точність побудованої моделі дорівнює 98,417%. Також, найкраще виявлення тут дає комбінування хорошого та великого набору даних, функцій ознак, а також найкращого алгоритму. Крім цього, за результатами розробленого скрипту, на рис.3.4 нижче, показано результат порівняння побудованої моделі, з іншими алгоритмами машинного навчання.

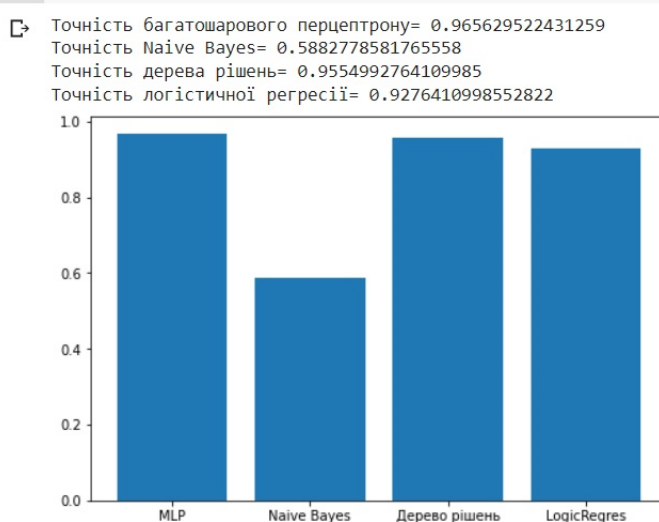


Рисунок 3.4 - Точність алгоритмів у графічному вигляді

Як видно з рис.3.4 вище, можна побачити, що порівнявши алгоритми, наша модель має більш високу точність, що також означає, що дана модель дозволяє визначати фішингові повідомлення більш точніше і краще. Тим самим отримані результати показують, що мета автоматизації була успішно досягнута, тобто даний скрипт дозволяє проводити обробку та виявлення фішингових повідомлень. Крім цього, цей скрипт також підвищує ефективність за рахунок точності, а також покращує час виявлення. Використовуючи цей скрипт, скорочується час, який буде витрачено на ручне визначення фішингових повідомлень, оскільки для ручного способу, необхідно використовувати різні інструменти такі як whois, geoip, ping, а також перевіряти веб-сайти візуально. Дані дії займають багато часу, через що ручний спосіб більш довгий, і не може бути точним. У цьому випадку, цей скрипт має процедуру автоматизації, що означає те, що використовувати його буде краще, ніж робити все вручну. Тим не менш, подальша робота буде спрямована на покращення роботи даного скрипта, а також додавання до цього скрипту інших корисних функцій.

ВИСНОВКИ

Ця робота допомогла зрозуміти основні принципи та проблеми фішингу. Детальний аналіз фішингу, який був проведений у цій роботі, дозволив зрозуміти, які фішингові атаки є на даний момент, які схеми фішингу, а також які методи захисту від фішингових атак є найбільш ефективними, Як і наскільки проблема фішингу є актуальною. Незважаючи на всі труднощі в фішингу, на даний момент фішинг залишається проблемою, оскільки користувачі все ще знаходяться під цією методикою соціальної інженерії. Сьогодні користувачі дуже часто використовуються поштовими службами, і тим самим стикаються з фішинговими атаками. Питання фішингу є дуже важливим, оскільки це показує статистику нападів, захист від фішингових атак, а також процес ідентифікації фішингових повідомлень, вимагає значного покращення, а також у деяких аспектах пріоритетних цікавих подій та розвитку. У більшості випадків дуже хороший захист фішингу є підготовка працівників, але навіть у цьому випадку можуть бути випадки, коли користувач може натиснути на шкідливе посилання. Існує кілька інструментів, які допомагають захистити від фішингу, наприклад, інструменти у браузері, такі як "Avira Browser Safety", цей інструмент видаляє шкідливі сайти з пошуку, а також блокує шкідливі сайти під час відвідування, також цей інструмент захищає конфіденційність користувача. Але це лише основні засоби захисту, які не можуть повністю блокувати сайти. У цьому випадку ця робота показує, що при виявленні фішингових повідомлень метод машинного навчання та використання його алгоритмів показують найкращий результат і найкраще справляються з виявленням.

У цій кваліфікаційній роботі, перший розділ представляє загальний аналіз фішингу. Також розглядаються схеми побудови фішингу, їх відмінності та особливості. Крім цього, в даному розділі також розглянуті можливі атаки фішингу та методи їх захисту. Також, у першому розділі розглядається актуальність фішингу, та статистика фішингу. Статистика показує, що на даний момент проблема фішингу залишається актуальною, і довго буде мати значення, оскільки протягом багатьох років може бути все більше нових фішингових атак. У другому розділі цієї роботи було зроблено дослідження, в якому були відібрані фішингові, та не фішингові повідомлення, на основі цих даних було пороховано кругові діаграми з урахуванням ознак. У ході цього дослідження виявилось, що цей метод обробки

та ідентифікації фішингових повідомлень досить довгий, та не зовсім точний. Також, в даному дослідженні з'ясувалося, що цих ознак не зовсім достатньо для отримання остаточної відповіді, і що можна було б досягти кращого успіху, якщо мати більш точне визначення фішингових повідомлень, а також мати трохи більше ознак. У третьому розділі порівнювалося порівняння алгоритмів машинного навчання. Під час цього порівняння було визначено, що для кращого та точного визначення фішингових повідомлень краще використовувати алгоритм "Multilayer perceptron". Крім цього, в даному розділі також було проведено обговорення того, як краще здійснити автоматизацію виявлення фішингових повідомлень. Під час обговорення з'ясувалося, що найкращим способом для автоматизації буде написання скрипта мовою Python. У заключному підрозділі третього розділу, були показані результати написаного скрипту, який включає автоматизацію виявлення фішингових повідомлень використовуючи машинне навчання і кращий алгоритм "Multilayer Perceptron". В результаті, цей скрипт показав хороші результати щодо визначення фішингових повідомлень, а також точності. Як висновок, можна сказати, що в цій роботі була розроблена процедура автоматизації та обробки фішингових повідомлень, крім цього, у цій роботі також була досягнута хороша точність визначення фішингових повідомлень, також, використання даного скрипту дозволить тим користувачам які збирають інформацію про фішингові і не фішингові повідомлення, краще спростити собі процес збору даних, оскільки даний скрипт включає в себе автоматизацію, більш точне визначення, і краще час. Тобто, немає необхідності використовувати інші інструменти які були використані в другому розділі даної роботи. Результати цього скрипту показують, що цей скрипт може добре послужити як визначення фішингових і не фішингових повідомлень. Даний скрипт буде також надалі поліпшуватися, і допрацьовуватися, включаючи до нього нові функції, крім цього, в даному скрипті також, будуть виправлені деякі недоробки. Крім того, цей скрипт може використовувати не тільки користувачі, які збирають інформацію про сайти, а також регулярні користувачі, які хочуть мати більш точну обробку та визначити фішингові повідомлення.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Яіцький А. О. Ефективні методи та засоби захисту фішингових атак / А. О. Яіцький, М. Г. Сахаров. – Варшава: Problems of science and practice, tasks and ways to solve them. – 2022. – С. 417-419.
2. Яіцький А. О. Застосування хмарних технологій в космосі / А. О. Яіцький, М. Г. Сахаров. – Київ: Priority directions of science and technology development. – 2021. – С. 337-339.
3. The SSL Store: The 12 Most Costly Phishing Attack Examples [Електронний ресурс] // SSL Store. – 2021. – Режим доступу до ресурсу: <https://www.thesslstore.com/blog/the-dirty-dozen-the-12-most-costly-phishing-attack-examples/>.
4. Matthew L Cole. The Complete Guide to Patents, Copyrights, and Trademarks: What You Need to Know Explained Simply. / Cole L Matthew // США: Atlantic Publishing Group Inc., 2008. – 288 с.
5. How to Conduct Email Phishing Experiments. [Електронний ресурс] / Kaspar Jüristo // Computer Science. – 2018. – Режим доступу до ресурсу: <https://1library.net/document/q7x98rky-how-to-conduct-email-phishing-experiments.html>.
6. School of phish: A real-world evaluation of anti-phishing training. [Електронний ресурс] / Ponnuram Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, Theodore Pham // DBLP. – 2009. – Режим доступу до ресурсу: https://www.heinz.cmu.edu/~acquisti/papers/Acquisti_ReaIWorld_Evaluation_of_Anti-Phishing_Training.pdf.
7. An Article of a Follow-up Study of Detecting Phishing Emails. [Електронний ресурс] / Ernst Bekkering, Dan Hutchison, Laurie Werner // Researchgate. – 2015. – Режим доступу до ресурсу: https://www.researchgate.net/profile/ErnstBekkering/publication/267705517_A_Follow-up_Study_of_Detecting_Phishing_Emails/links/557ed71608aeb61eae260c2f/A-Follow-up-Study-of-Detecting-Phishing-Emails.pdf.
8. An Article of Anti-Phishing Approach that Uses Training Intervention for Phishing Websites Detection. [Електронний ресурс] / Abdullah Alnajim, Malcolm Munro // Researchgate. – 2009. – Режим доступу до ресурсу:

https://www.researchgate.net/publication/220841443_An_AntiPhishing_Approach_that_Uses_Training_Intervention_for_Phishing_Websites_Detection.

9. Deanna D. Caputo, Shari Lawrence Pfleeger, Jesse D. Freeman, M. Eric Johnson. Going Spear Phishing: Exploring Embedded Training and Awareness / D. Deanna Caputo, Pfleeger Lawrence Shari, Freeman D. Jesse, Johnson M. Eric // CИИА: IEEE, 2013. – 38 с.

10. Robust Email Security Requires Alignment Between Security Practitioners and Decision Makers [Электронный ресурс] // Osterman Research.– 2020. – Режим доступа до ресурсу: <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RWlYx>.

11. DBIR 2021 Data Breach Investigation Report [Электронный ресурс] // Verizon. – 2021. – Режим доступа до ресурсу: <https://www.verizon.com/business/resources/reports/2021-data-breach-investigations-report.pdf>.

12. Abnormal Security Quarterly BEC Report. [Электронный ресурс] / Evan Reiser // Abnormal Security. – 2020. – Режим доступа до ресурсу: https://info.abnormalsecurity.com/rs/231IDP139/images/AS_Qtrly_BEC_Report_Q3_2020.pdf.

13. Cost of a Data Breach Report 2021 [Электронный ресурс] // IBM Security. – 2020. – Режим доступа до ресурсу: <https://www.ibm.com/downloads/cas/OJDVQGRY>.

14. Google Safe Browsing Service of Google that helps protect devices. [Электронный ресурс] // Google. – 2022. – Режим доступа до ресурсу: <https://transparencyreport.google.com/safe-browsing/overview>.

15. 2022 SonicWall cyber threat report. [Электронный ресурс] / Bill Conner // SonicWall. – 2022. – Режим доступа до ресурсу: <https://www.infopoint-security.de/media/2022-sonicwall-cyber-threat-report.pdf>.

16. 2022 State of the Phish. [Электронный ресурс] / Proofpoint. – 2022. – Режим доступа до ресурсу: <https://www.proofpoint.com/sites/default/files/threat-reports/pfpt-us-tr-state-of-the-phish-2022.pdf>.

17. Anti-Phishing Working Group. [Электронный ресурс] / David Jevans // APWG . – 2003. – Режим доступа до ресурсу: <https://apwg.org/trendsreports/>.

18. Detection of Phishing Attacks: A Machine Learning Approach. [Электронный ресурс] / Ram Basnet, Srinivas Mukkamala, Andrew H. Sung //

Researchgate,2008. Режим доступу до ресурсу:https://www.researchgate.net/publication/226420039_Detection_of_Phishing_Attacks_A_Machine_Learning_Approach.

19. Madhav P. Namdev. Phishing Attacks and Detection / P. Namdev Madhav - Англія:, LAP Lambert Academic Publishing, 2015. – 68 с.

20. Oluwatobi Ayodeji Akanbi, Iraj Sadegh Amiri, Elahe Fazeldehkordi: A Machine-Learning Approach to Phishing Detection and Defense. / Oluwatobi Ayodeji Akanbi, Iraj Sadegh Amiri, Elahe Fazeldehkordi // CIHA: Syngress Media, 2015. – 100 с.