Міністерство освіти і науки України
Національний технічний університет
«Дніпровська політехніка»

Інститут електроенергетики
Факультет інформаційних технологій
Кафедра безпеки інформації та телекомунікацій

ПОЯСНЮВАЛЬНА ЗАПИСКА
кваліфікаційної роботи ступеню магістра

студента *Шабельника Сергія Павловича*

академічної групи *125м-20-2*

спеціальності *125 Кібербезпека*

спеціалізації[1]

за освітньо-професійною програмою *Кібербезпека*

на тему *Вдосконалення способів фільтрації небезпечного трафіку на основі реєстру цифрових відбитків пристроїв*

Керівники	Прізвище, ініціали	Оцінка за шкалою		Підпис
		рейтинговою	інституційною	
кваліфікаційної роботи	к.т.н., доц. Сафаров О.О.			
розділів:				
спеціальний	к.т.н., доц. Сафаров О.О.			
економічний	к.е.н., доц. Пілова Д.П.			
Рецензент				
Нормоконтролер	ст. викл. Мєшков В.І.			

Дніпро
2022

ЗАТВЕРДЖЕНО:
завідувач кафедри
безпеки інформації та телекомунікацій
_____ д.т.н., проф. Корнієнко В.І.

«_____»_____20___року

ЗАВДАННЯ
на кваліфікаційну роботу
ступеня магістра

студенту *Шабельнику Сергію Павловичу* **академічної групи** *125м-20-2*
(прізвище ім'я по-батькові) (шифр)

спеціальності *125 Кібербезпека*
(код і назва спеціальності)

на тему ***Вдосконалення способів фільтрації небезпечного трафіку на основі реєстру цифрових відбитків пристроїв***

затверджену наказом ректора НТУ «Дніпровська політехніка» від 10.12.2021 № 1036-с

Розділ	Зміст	Термін виконання
Розділ 1	Мережевий трафік, його види. Машинний трафік. Боти: види, застосування, тестування, боротьба з ними. Детальна класифікація та оцінка трафіку. Проблематика анонімізації. Висновки розділу.	10.12.2021
Розділ 2	Аналіз наявних моделей забезпечення анонімізації. Розгляд методів та засобів збору даних про користувачів з висвітленням переваг та недоліків. Метод цифрового відбитку пристрою, аналіз, аргументація, впровадження. Висновок виконання частини.	20.12.2021
Розділ 3	Економічний розрахунок впровадження, капітальних та річних витрат. Оцінка можливих збитків. Висновок розділу.	15.01.2022

Завдання видано _____ _____
(підпис керівника) (прізвище, ініціали)

Дата видачі: 02.09.2021р.

Дата подання до екзаменаційної комісії: 14.01.2022р.

Прийнято до виконання _____ _____
(підпис студента) (прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 97 с., 30 рис., 6 табл., 4 додатка, 54 джерела.

Об'єкт дослідження: методи, засоби та способи фільтрації небезпечного мережевого трафіку.

Мета роботи: вдосконалення наявних способів фільтрації небезпечного трафіку шляхом аналізу наявних засобів з метою визначення їх ефективності для впровадження нових методів задля зменшення рівня небезпечного мережевого трафіку.

Методи розробки: аналіз, порівняння, тестування, дослідження.

У першому розділі було проведено визначення мережевого трафіку. Виконано поділ трафіку за походженням і маршрутом. Проаналізовано та протестовано ботів, зокрема: їх діяльність, вплив на громадську думку, наявні методи та засоби боротьби з ними. Детально розглянуто та проаналізовано злочинний мережевий трафік, фактори що сприяють його зростанню та визначено проблеми які із нього випливають.

У спеціальній частині було проаналізовано наявні моделі забезпечення анонімності, розглянуто методи збору інформації про користувача. Визначено їх основні недоліки, запропоновано методи по їх вирішенню та вдосконаленню. Надані рекомендаційні пропозиції по створенню централізованого реєстру.

У економічному розділі було визначено вартість капітальних та експлуатаційних витрат, оцінені можливі збитки в разі недотримання правил під час впровадження та використання.

Практичне значення роботи полягає у можливості практичного застосування отриманих даних з метою покращення рівня протидії злочинному трафіку.

Наукова новизна полягає у вдосконаленні наявних методів фільтрації мережевого трафіку.

Ключові слова: МЕРЕЖЕВИЙ ТРАФІК, БОТИ, ЗЛОЧИННИЙ ТРАФІК, ФІЛЬТРАЦІЯ ТРАФІКУ, ЦИФРОВИЙ ВІДБИТОК ПРИСТРОЮ, АНОНІМНІСТЬ, АНОНІМІЗАЦІЯ.

ABSTRACT

Explanatory note: 97 p., 30 pics., 6 tables., 4 applications, 54 sources.

Object of research: methods of filtering malicious network traffic.

Objective: improving existing methods of filtering malicious traffic by analyzing available tools in order to determine their effectiveness for the introduction of new methods in order to reduce the level of malicious network traffic.

Research methods: analysis, comparison, testing, research.

In the first section, network traffic was defined. The division of traffic by origin and route was performed. Bots have been analyzed and tested, in particular: their activities, influence on public opinion, available methods and means of combating them. Criminal network traffic, factors contributing to its growth and identifying problems arising from it are considered and analyzed in detail.

The special part analyzed the existing models of anonymity, namely methods of collecting information about the user. Their main disadvantages are determined, methods for their solution and improvement are proposed. Recommendations for creating a centralized register are provided.

The economic section determined the cost of capital and operating costs, estimated possible losses in case of non-compliance with the rules during implementation and use.

The practical significance of research in the possibility of practical application of the obtained data in order to improve the level of counteraction to criminal traffic.

The scientific novelty consists in improving the existing methods of network traffic filtration.

Key words: NETWORK TRAFFIC, BOTS, CRIMINAL TRAFFIC, TRAFFIC FILTERING, DIGITAL FINGERPRINT OF THE DEVICE, ANONYMITY, ANONYMIZATION.

СПИСОК УМОВНИХ СКОРОЧЕНЬ

ООН - Організація Об'єднаних Націй;

ШІ – штучний інтелект;

CAGR – (англ. Compound annual growth rate) – сукупний середньорічний темп зростання;

CDP – англ. Cisco Discovery Protocol;

DDos – англ. (Distributed) Denial-of-service attack;

DEC – (англ. Deep Entity Classification) - глибока класифікація сутностей;

DoS – англ. Denial-of-service attack;

PCI DSS – (англ. Payment Card Industry Data Security Standard) - стандарт безпеки даних індустрії платіжних карток;

WWW – всесвітня мережа «Інтернет».

ЗМІСТ

С.

ВСТУП

Згідно прогнозу ООН, до 2023 року населення світу буде становити вісім мільярдів осіб, а до кінця двадцять першого століття, цей показник сягне десяти мільярдів [1].

Кожного року стрімко зростає кількість людей які отримують доступ до цифрових ресурсів, в тому числі до мережі «Інтернет» [2]. Розвиток мобільних пристроїв, зниження їх ціни та підвищення доступності. Введення в експлуатацію стільникових мереж нового покоління (5G) [3]. Згідно статистики, лише порівнюючи січень 2021 року та січень 2022 року кількість користувач зросла на 7% [2].

Для більшості людей всесвітня мережа «Інтернет» стає засобом не лише для роботи, а також і для проведення вільного часу. Поруч із цим зростає кількість мережевого трафіку. Цей показник вже є рекордним і збільшився майже вдвічі за останні три роки [4].

Країни все частіше використовують інтернет як поле для проведення гібридної війни. Дії направлені на зміну громадської думки, зміна поглядів і навіть вплив на вибори. Це, не що інше, як нова зброя двадцять першого століття [5]. Зловмисники, в свою чергу, використовують «всесвітню павутину» для отримання вигоди шляхом обману, зламу та соціальної інженерії. В цих осіб не стоїть потреба виходити на вулицю і викрадати гроші із гаманців, для цього вони використовують мережу.

Все вищезазначене потребує ретельного контролю та швидкої реакції у випадку своєї появи. Поруч із покращенням засобів для злочинності, мають покращуватись і способи та методи фільтрації направленні на боротьбу з нею, зокрема і на рівні початкової фільтрації інтернет-трафіку.

Об’єктом дослідження є існуючі методи, засоби та способи фільтрації небезпечного мережевого трафіку.

Метою роботи є вдосконалення способів фільтрації небезпечного трафіку, шляхом аналізу наявних засобів з метою визначення їх ефективності для впровадження нових методів задля зменшення рівня небезпечного мережевого трафіку.

Ця робота спрямована на те, аби проаналізувати та обрати найбільш актуальні рішення у сфері фільтрації мережевого трафіку, виходячи із нагальних потреб сьогодення.

1 СТАН ПИТАННЯ. ПОСТАНОВКА ЗАДАЧІ

1.1 Мережевий трафік. Його види та особливості.

Мережевий трафік – це кількість даних, що переміщуються мережею в певний момент часу [6].

За своїм походженням поділяється на:

- Вихідний (находить в зовнішню мережу);
- Вхідний (надходить в внутрішню мережу).

За своїм маршрутом:

- Внутрішній (в середині певної, частіше локальної, мережі);
- Зовнішній (за межами певної локальної мережі).

Зовнішній трафік і є тим самим інтернет-трафіком який генерується як людьми, так і програмно-апаратними засобами. Саме цьому виду трафіку буде приділена найбільша увага в процесі аналізу стану питання та постановки задачі для вирішення можливих проблем та оптимізації роботи з ним.

Трафік, який генерується у мережі інтернет є найбільшим, порівнюючи з іншими. За прогнозом, який був проведений американською компанією «Cisco», його об’єм у 2022 році зросте до 4.8 ЗБайт, що становить приблизно $4{,}617 \times 10^9$ Терабайт [4].

Рисунок 1.1 – Зростання трафіку в Экстабайт (середньорічний місячний показник)

За аналізом рисунка 1.1 можемо спостерігати, що ріст об'єму лише у період з 2017 по 2021 рік відбувся більш ніж у 2.6 рази – з 122 Збайт до 319 Збайт у році 2021. А показник CAGR (сукупний середньорічний темп зростання) становить 26%.

В подальшому, в процесі роботи, в якості додаткових характеристик нами буде виділено наступні додаткові критерії оцінки:

— Створення;

— Кінцева мета.

За генерацією трафік буде розмежований як:

— Машинний;

— Людський.

Машинний трафік – згенеровано виключно програмно-апаратними засобами без, або з мінімальним, впливом людини [7]. Маркером є: автоматичність дії, повторення, об'єм даних (більше ніж людський), час роботи.

Людський – згенеровано людиною без, або з мінімальною, допомогою програмно-апаратних засобів. Маркером є: різноманітність, низька швидкість генерації, періодичність створення.

Рисунок 1.2 – Статистика трафіку за генерацією на 2020 рік.

За кінцевою метою трафік буде розмежований як:

- Білий;
- Сірий;
- Чорний.

Кінцева мета може бути однакова як для машинного, так і для людського трафіку.

Білий – трафік в розрізі якого не відбувається дій, які б могли так чи інакше:

a) порушувати законодавство,
b) порушувати загальноприйнятні норми і правила експлуатації та використання,
c) ускладнювати роботу мережевих пристроїв,
d) створювати протиправні дії та контент,
e) використовувати штучний інтелект як імітацію людини.

В цілому, до цього виду варто віднести звичайне (типове) використання мережі «Інтернет» людиною, або машиною, без вживання зазначених раніше пунктів.

Сірий – трафік, дії, в розрізі якого відбуваються, не є загально прийнятними, але не порушують закон чи будь які інструкції.

В якості прикладу для розуміння можна навести роботу програмних скриптів, які забезпечують захист від рекламного контенту. Вони генерують додатковий об'єм даних, збільшуючи навантаження на мережу, але, при цьому, в їх діях нема чогось протизаконного.

Чорний – трафік, кінцевою метою якого є досягнення, як мінімум, одного з критеріїв (a – e).

Статистичні дані, зображені на рисунках 1.2 та 1.3 є узагальненими і висвітлені шляхом аналізу деяких незалежних звітів міжнародних компаній та організацій які займаються, зокрема, і питанням кібербезпеки [8][9].

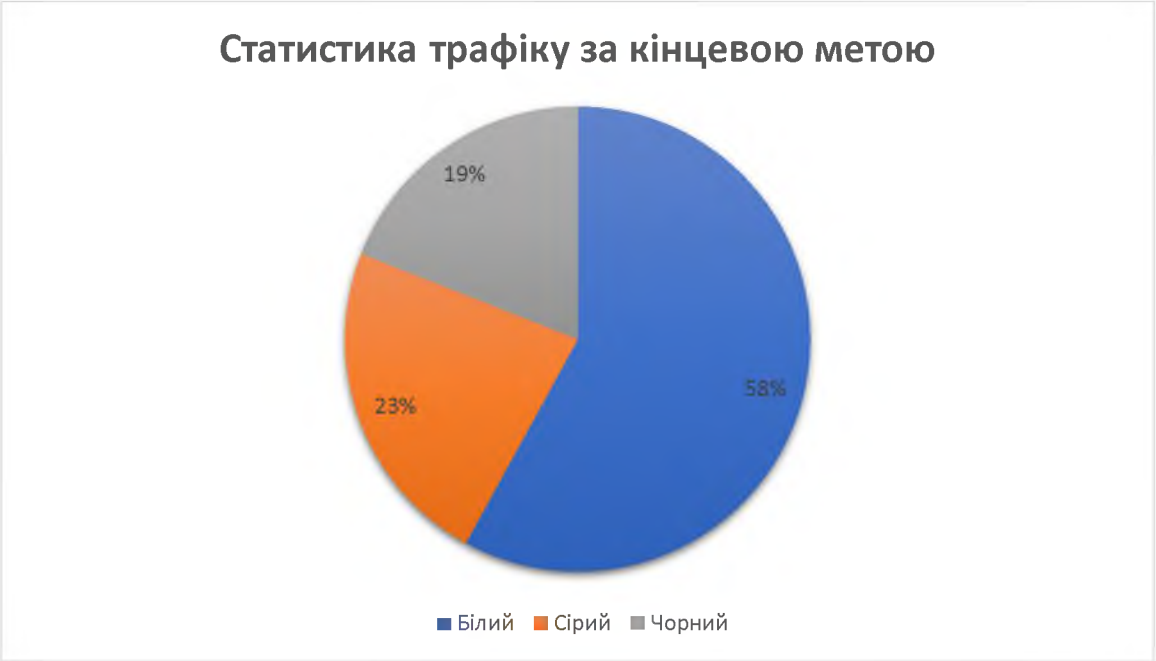

Рисунок 1.3 – статистика трафіку за кінцевою метою

1.2 Використання трафіку машинними засобами

Виходячи із даних, зображених на рисунку 1.2 можемо бачити, що на даний час (на 2021 рік), близько 37% від усього трафіку створюється програмно-апаратними засобами, тобто машинами, а ще частіше – ботами.

Бот – спеціальна програма, або машино-апаратний комплекс, що за певним, раніше створеним, завданням у певний проміжок часу за заданим розкладом з використанням інтерфейсу виконує (імітує) дії подібні людині [10].

Безперечно, не всі дії ботів є шкідливими. Певні скрипти, які виконують машини створенні для спрощення діяльності люди. Зазвичай це рутинні та монотонні дії, які автоматично будуть виконуватись більш якісно та швидше. Кінцева мету таких дій можна охарактеризувати як білу або сіру. Згідно досліджень деяких компаній, зокрема і американської компанії «Imperva», що спеціалізується на розробці програмного забезпечення в сфері кібербезпеки та інформаційної безпеки щорічно проводить аудит поточного стану діяльності ботів у мережі інтернет, таких ботів можна віднести до «Good Bots».

Їх можна поділити на такі категорії:

— Моніторинг-боти – встановлюються для моніторингу показників веб-сайту та надання інформації користувачам. Наприклад, відстеження поточного стану сайту, інструментів/рішень SEO, тощо;

— Чат-боти – боти, які працюють на сайті, або в програмних додатках, з метою забезпечення роботи автоматизованого чату, імітуючи людську розмову;

— Пошук-боти – полегшують пошук інформації, автоматизують цей процес. Як приклад – пошук інформації, що може порушувати авторські права;

— Боти-партнери – створенні для полегшення виконання певних дій.

Інші ж боти, виконують функції, кінцевою метою яких є створення чорного трафіку та виконання, як мінімум, одного завдання з переліку (а – е). Як правило, вони спеціально розроблені для виконання таких задач.

Серед критеріїв їх оцінки можна виділити наступні види:

- Боти для відтворення DDoS-Атак

DDoS - атака на комп'ютерну мережу з наміром зробити ресурси недоступними користувачам [12].

Одним із шляхів вчинення DDoS є надсилання до атакованого пристрою величезну кількість зовнішніх запитів, що значно більше за його можливість опрацювання. Як правило, для використання таких дій і використовуються боти.

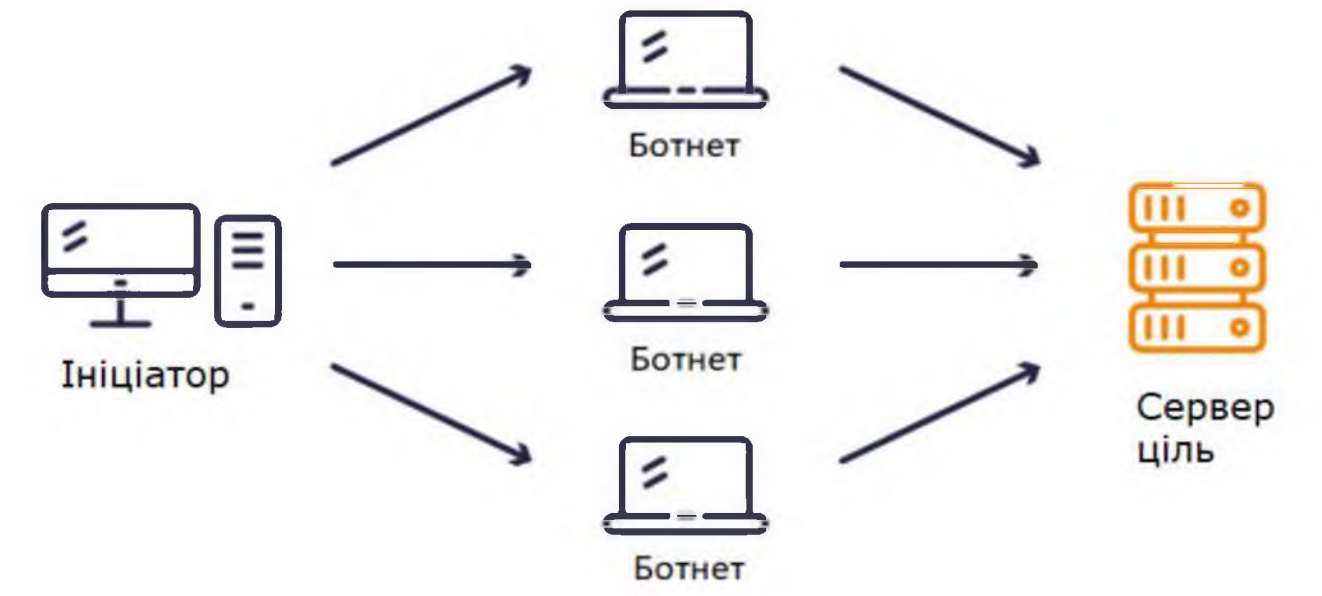

Рисунок 1.4 – схема проведення DDoS-атаки.

Ботнет – мережа, яка складається з великої кількості «Bad Bots» та автономного програмного забезпечення [13].

- Боти-імітатори;

Даний вид також використовується для імітації дії людини у мережі, але уже переслідує інші, злочинні цілі. Як приклад – штучне збільшення певних показників у перегляді реклами.

- Боти у соцмережах;

Головною ціллю є імітація людини у мережі інтернет та соціальних мережах з метою змінення колективної та персональної думки

- Спам-бот;

Боти, дії яких направленні на розсилання спаму, фальшивих повідомлень.

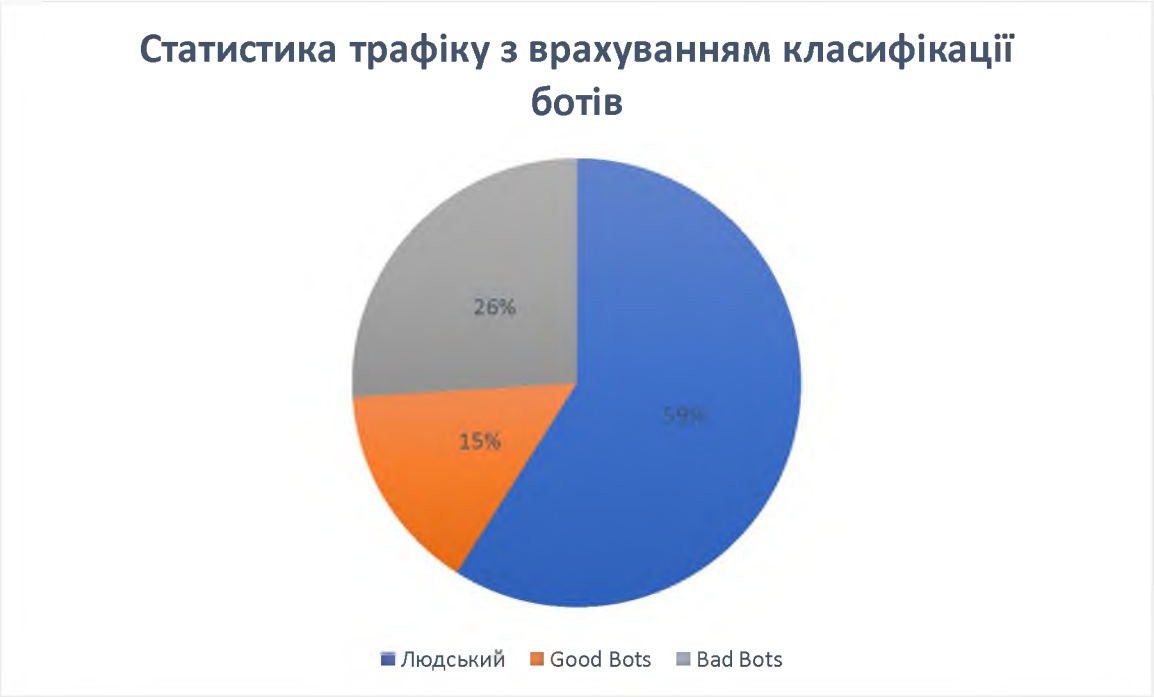

Рисунок 1.5 – Статистика трафіку з врахуванням класифікації ботів

Більш детальний аналіз рисунку 1.2, наведений на рисунку 1.5. На ньому машинний трафік додатково розподіллено на «Good Bots» та «Bad Bots». Якщо підсумовувати на даному етапі, то буде видно, що 26% трафіку створюється ботами, кінцева мета яких може бути класифікована, як мінімум як сіра, а іноді і чорна.

26% від загальної кількості трафіку на даний момент складає близько 1.25 ЗБайт. Для прикладу, за увесь 2017 рік увесь трафік у мережі інтернет склав близько 1,46 ЗБайт.

1.3 Аналіз «Bad Bots»

Враховуючи показники, що наведені у попередньому розділі, варто приділити окрему увагу проблемі з трафіком що генерується так званими «Bad Bots», або ж зловмисними ботами.

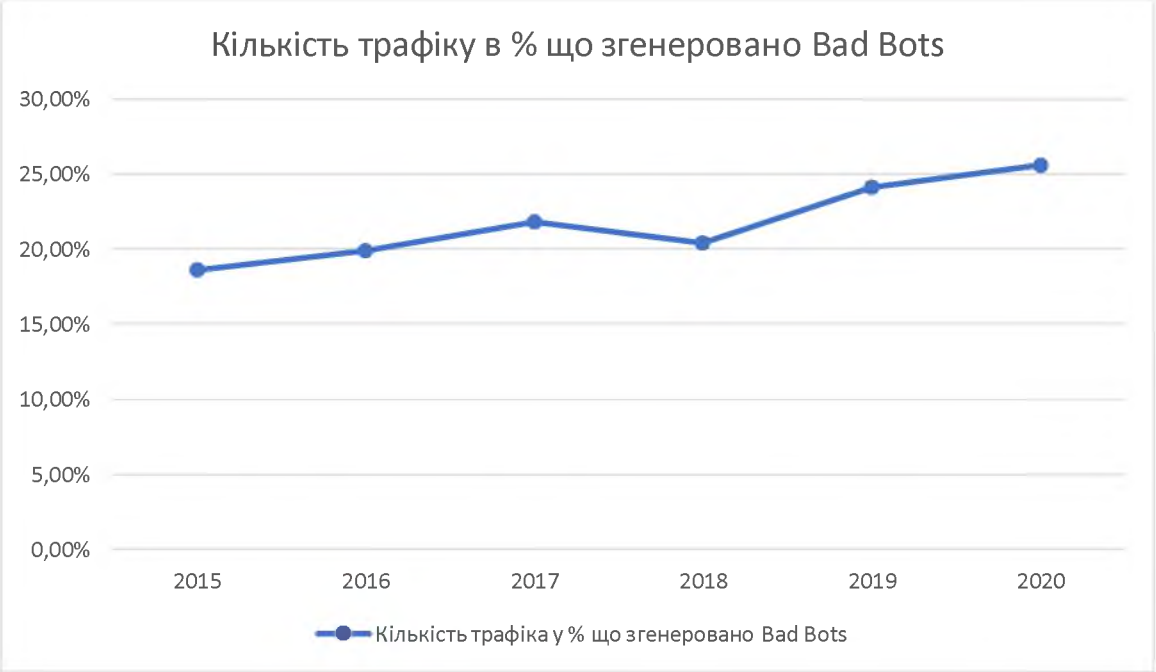

Рисунок 1.6 - Кількість трафіку в (%) що згенеровано Bad Bots

На основі досліджень американської компанії «Imperva» [11], що проводить щорічний аудит діяльності програмно-апаратних комплексів у мережі «Інтернет» представлено рисунок 1.6. Аналізуючи його, можна побачити тенденцію збільшення відсоткового відношення «поганих ботів» до загальної кількості трафіку. Порівнюючи з 2015 роком це зростання становило 7% до загально об'єму.

На це можна виділити декілька причин:

1. Зріст рівня застосування ШІ. Лише за один 2020 рік цей показник зріс на 18%.

2. Епідемія, спричинена вірусом COVID-19, яка змінила стиль роботи більшості людей та компаній у світі, спричинивши підвищення рівня автоматизації та використання програмно-апаратних засобів.

3. Збільшення пристроїв розумного дому, більш ніж у половину з 2017 року по 2020 рік.

4. Загальний приріст автоматизації у світі.

У розділі 1.2 цієї роботи наведено загальні критерії та характеристики на які поділяються зловмисні боти, та на наш погляд варто розкрити це питання більш детально.

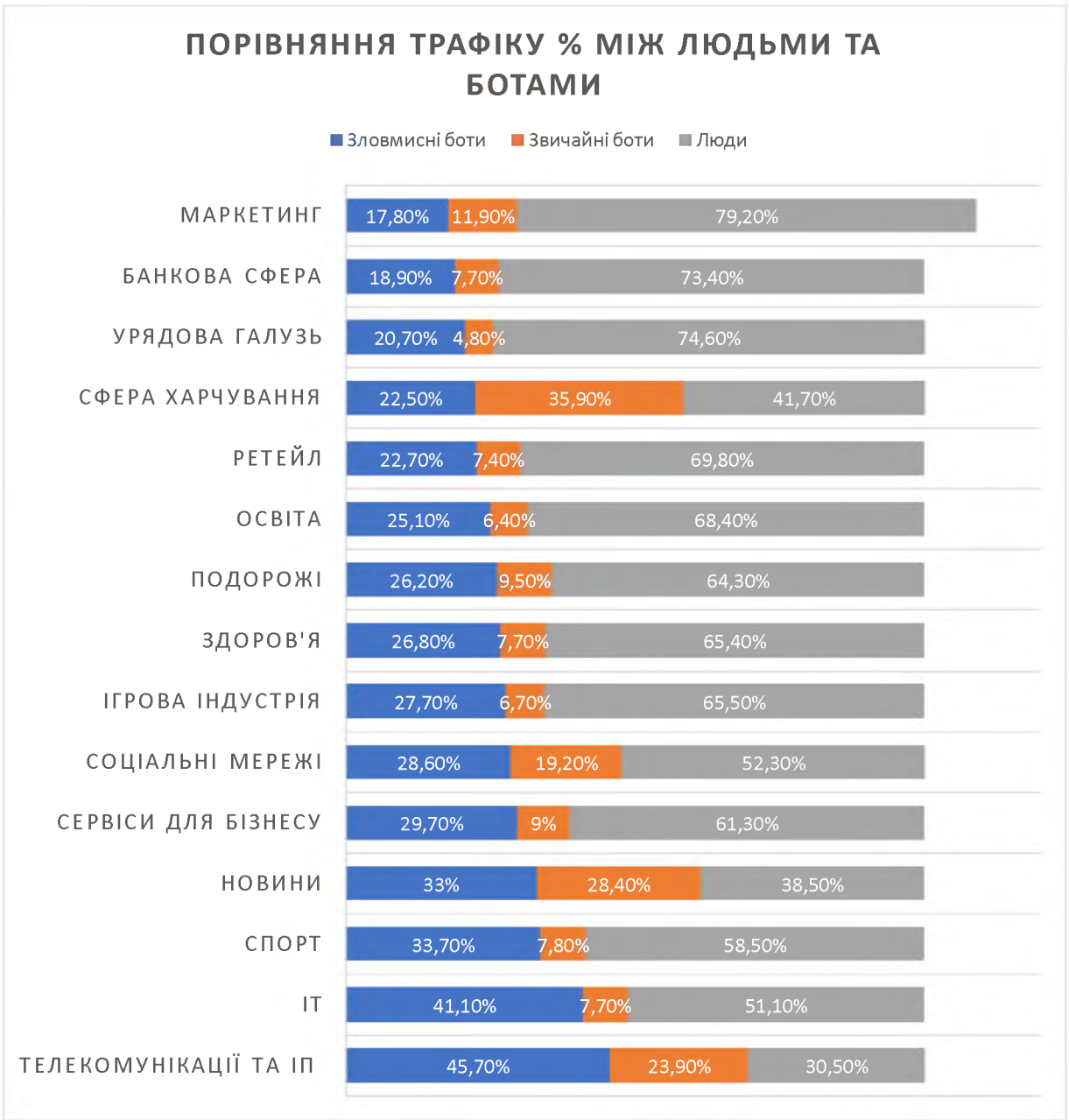

Рисунок 1.7 – Порівняння трафіку у % між людьми та ботами

На рисунку 1.7 наведено порівняльну відсоткову характеристику трафіку залежності від галузі. В деяких галузях, таких як IT, наприклад, доля ботів складає 70% порівнюючи с людськими 30%. А частина зловмисних ботів складає майже половину. А саме – 45,7%. Найкраща ж ситуація у маркетингу та банківській сфері, де, не дивлячись на високий рівень автоматизації зберігається дещо низькій рівень ботів в цілому. Це може бути пояснене високим рівнем витрат на безпекову складову організацій.

Загалом було проаналізовано 15 індустрій.

Таблиця 1.1 – Пояснення до рисунку 1.7

Індустрія	Що віднесено до бізнесу	Злочини ботів
Маркетинг	Маркетингові та рекламні агенції	Копіювання чужого контенту, шахрайство з рекламою, спотворення інформації
Банківська сфера	Банки, крипто обмінники, інвестиційні агенції	Кардінг, копіювання чужого контенту, викрадення акаунтів
Урядова галузь	Владні сайти, сайти з держпослуг	Викрадення аккаунтів, незаконий парсинг даних
Сфера харчування	Служби доставки їжі, онлайн купівлі	Кардінг, викрадення аккаунтів та банківських карт
Ретейл	Електронна комерція, маркетплейси	Копіювання чужого контенту, рекламне шахрайство, спотворення інформації
Освіта	Сайти навчальних закладів, платформи онлайн навчання	Викрадення аккаунтів, копіювання закритої інформаці
Подорожі	Авіалінії, отелі, відповідні сайти	Зниження/підвищення оцінок, викрадення аккаунтів, незаконий парсінг даних.
Здоров'я	Заклади здоров'я та аптеки	Викрадення аккаунтів, копіювання контенту, антивакцинаторська компанія
Ігрова індустрія	Онлайн ігри, гемблінг	Викрадення аккаунтів, створення аккаунтів, рекламне шахрайтсво
Соціальні мережі	Соціальні мережі	Викрадення аккантів, створення аккаунтів, рекламне шахрайство, викрадення банківських карт

Продовження таблиці 1.1

Індустрія	Що віднесено до бізнесу	Злочини ботів
Сервіси для бізнесу	Нерухомість, бізнес метріки та комерційні платформи	Копіювання контенту, викрадення аккунтів
Новини	Сайти з новинами, ЗМІ	Спам, копіювання контенту, рекламне шахрайство
Спорт	Спортивні сайти та сервіси онлайн рахунків	Копіювання контенту
IT	IT сервіси	Викрадення аккаунтів
Телекомунікації та інтернет провайдери	Телекомунікаційні та інші інтернет провайдери	Викрадення аккаунтів, створення аккаунтів

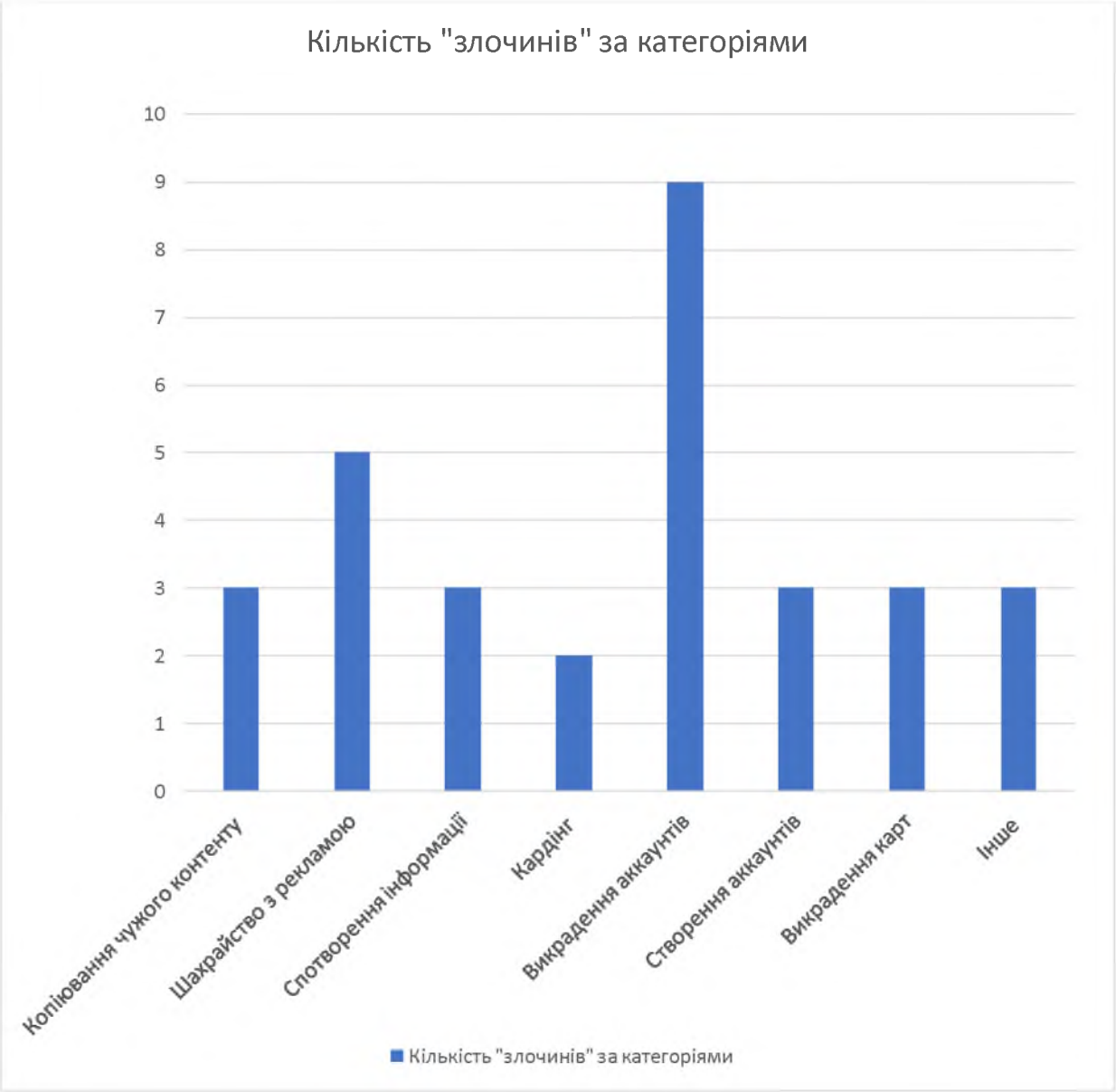

Рисунок 1.8 – узагальнена характеристика злочинних дій ботів

Із обраних протиправних дій, які вказані на рисунку 1.8, та, що виконуються ботами, найбільшу кількість «злочинів» направлення на певну діяльність пов'язану з аккаунтами, при чому на різних ресурсах. Це як і створення Бот-аккаунтів, так і викрадення вже існуючих, яке відбувається у 9 виділених категоріях із 15 наведених. Лише за минулий рік було викрадено близько 500 мільйонів аккаунтів у різних соціальних мережах.

Використання зловмисних ботів під час крадіжки аккаунтів є абсолютно обґрунтованим, оскільки основний перелік задач зводиться до автоматичного підбору даних, якщо використовується метод брут-форс, або швидкого створення фішингових сайтів, посилань якщо отримати данні необхідно саме таким шляхом.

Брут-форс (англ. brute force) - метод зламу шляхом перебору усіх можливих значень варіантів пароля, або ж ключа [14].

Фішинг (англ. Fishing) – метод зламу, шляхом використання соціальної інженерії з метою отримання від жертви її данних [15].

Створення аккаунтів ще одна задача, яка вимагає великої кількості автоматичних монотонних та швидких дій одночасно. За своїми фізичними можливостями людина не здатна відповідати за швидкістю ботам, тому у створенні фейк-сторінок перевагу надають програмно-апаратним засобам.

Не дивлячись на те, що майже усі з вищенаведених категорій мають безпосереднє відношення до грошових операцій, кардінг застосовується лише у 2 категоріях.

Кардінг (англ. Carding) – вид шахрайства з використанням платіжних реквізитів карт, операція з якими відбувається без підтвердження з боку власника [16].

Причиною цього може бути підвищення рівня безпеки банків, та застосування ними стандартів безпеки «Payment Card Industry Data Security» [17] Standard (PCI DSS) в основах якого встановлені вичерпні переліки вимог до безпеки карт. Враховуючи це, можливе використання ботів в цьому напрямі не принесе зловмиснику бажаного результату.

1.3.1 Боти, як інструмент для зміни громадської думки

Одним із критерій із переліку (а – е) який був виділений нами як маркер для поверхневої фільтрації кінцевої мети чорного трафіку , був пункт:

Е - використання штучного інтелекту як імітацію людини.

Аналізуючи накопичені дані, що були зібрані всесвітніми організаціями, та частина з яких була висвітлена на Всесвітньому економічному форумі, можемо прийти до висновку різкого зростання рівня занепокоєння саме з боку кіберзагроз [18].

2014 рік	1) Нерівність доходів
	2) Глобальне потепління
	3) Безробіття
	4) Кліматичні проблеми
	5) Кібератаки

Рисунок 1.9 - п’ять найбільших ризиків для людства у 2014 році

2019 рік	1) Глобальне потепління ↑
	2) Кліматичні пробелми ↑
	3) Стихійні лиха
	4) *Шахрайство або крадіжка даних*
	5) *Кібератаки*

Рисунок 1.10 - п’ять найбільших ризиків для людства у 2019 році

Проводячи паралель між головними загрозами людства у 2014 році (рисунок 1.9) та у 2019 році (рисунок 1.10), варто зауважити наступні показники:

1. Зростання кліматичних проблем;
2. Появу проблеми, яка стосується крадіжки даних;
3. Закріплення критеріїв кібератак.

На цих рисунках варто звернути увагу на появу такого виду загрози як «Шахрайство або крадіжка даних». Вивчаючи показники у період з 2009 по 2019 роки, вперше цей маркер з'явився у 2018 році, де також був віднесений на четверте місце, (в той час як «кібератаки» на третє місце), і у році 2019 залишився на четвертій позиції.

Вивчення штучного інтелекту у 2018 році, вказало на потенціал використання саме цієї здатності інженерної системи до посилення рівня кібератак у світі, що створюють ризики для критичної інфраструктури.

Серед проведеного у 2019 році опитування близько 65% респондентів висловили занепокоєння в питані зростання рівня кібератак, що були направлені на крадіжку їх персональних даних [19] .

Щонайменше 1.1 мільярд персональних даних було викрадено протягом 2018 року, із них [20]:

— Близько 150 мільйонів аккаунтів належали до медичних додатків, а отже, скоріш за все, включали в себе усі фізіологічні данні власника;

— Близько 10 мільйонів були аккаунти від банківських установ та організацій, що зберігали відповідні фінансові дані;

Примітка: не враховуються крадіжка банківських карт.

— Близько 50 мільйонів аккаунтів належали до соціальної мережі «Facebook».

У 2018 році дослідницький відділ американської компанії «IBM» розкрила цільовий штучний інтелект який може приховувати в собі шкідливе програмне забезпечення «WannaCry», і знаходиться він в додатку, що відповідає за створення та проведення відеоконференцій [21]. Активізація його відбувалася

лише в момент фіксації голосу чи обличчя жертви, що допомагало йому бути непоміченим антивірусним програмним забезпеченням, оскільки робота імітувалась під «прикриттям» програми-донора.

Кінцевої метою даного ПЗ було масове копіювання інформації з зараженого пристрою, включаючи як обличчя жертви, так і інші персональні данні (логіни, паролі, дані банківських карток).

Головним питанням, що виникає, є мета збору цієї інформації. З метою дослідження та можливого вивчення даного питання, було виконано пошук у мережі інтернет, з метою вільного віднайдення інформації. Так, без використання спеціального обладнання на одному з веб-ресурсів було знайдено файл формату (.txt) в якому знаходився перелік даних у форматі «логін:пароль» від однієї із соціальних мереж. Враховуючи час створення та знаходження, а також можливе попереднє використання більшість за даних вже не є актуальною, але в деяких випадках зайти на зазначений веб-ресурс вже ж таки вдавалося.

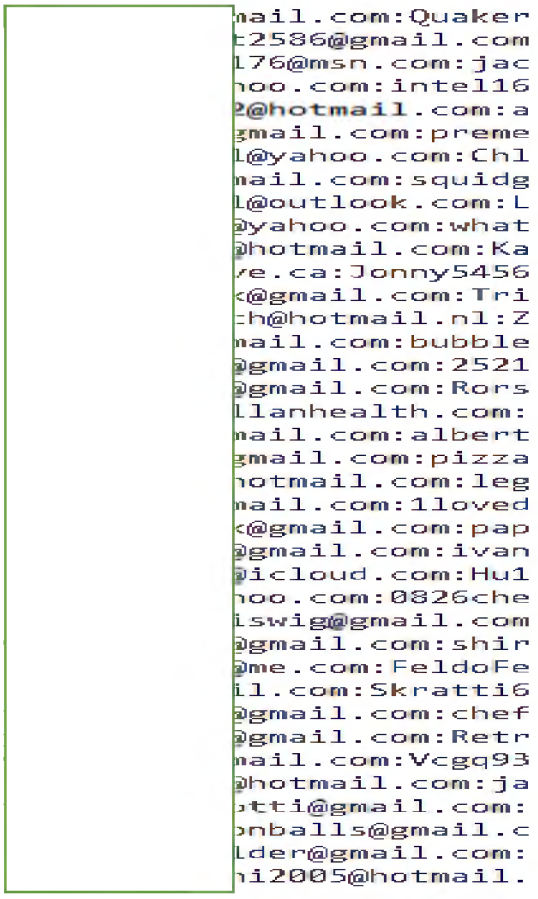

Рисунок 1.11 – приклад знайденої інформації

Цей пошук дає змогу зробити висновки, що сама наявність та факт викрадення не завжди є головною метою розробників ботів.

Окрім цього, є факт саме крадіжки вже готових аккаунтів, про що свідчать дані надані на рисунку 1.8, де обраний критерій зустрівся у дев'яти обраних сферах з п'ятнадцяти. Хоча генерування нових облікових записів є менш ресурсозатратним для програмного забезпечення. Час, що потрібен на створення нових аккаунтів часто складає менше 1 секунди. В той час як на злом, в залежності від методу, може витрачатися в десятки, а то і сотні разів більше часу. Як приклад – злом грубої сили, під час якого відбувається математична перестановка даних з метою підбору паролю. В залежності від його складності термін спроб необмежений. Тобто потреба виникає саме на готовий продукт.

Можливою відповіддю на це питання може стати доклад, що був опублікований у 2017 та 2019 роках Спеціальним комітетом Сенату США з питань розвідки на тему: «Російські активні заходи кампаній та інтерференція на американських виборах 2016 року» [22]. На думку авторів тексту, що складається на більш ніж тисячі листах, частина з яких таємна, уряд Російської федерації, використовуючи «Агентство інтернет досліджень», що знаходиться у Санкт-Петербурзі втрутилось у президентську компанію виборів 2016 році, шкодячи кампанії одного з кандидатів на користь іншого.

Протягом проведення спецоперації було створено, та, що важливо, було використано тисячі аккаунтах у соціальних мережах з метою дестабілізації американського населення та зміни їх думки. Частину із дій виконували так звані «Фабрики тролів» які керувалися живими людьми. Іншу ж частину роботи виконував саме штучний інтелект, імітуючи облікові записи реальних людей.

Фізично, адмініструвати відразу декілька десятків аккаунтів людина не може, так як це потребує значного фізичного та розумового навантаження. У цьому випадку і було застосовано штучний інтелект який керував тисячами аккаунтів реальних людей.

ШІ не потрібно було писати якісь тексти чи відповідати у коментарях, іноді для зміни думки було достатньо просто опубліковувати певний контент та робити автоматичні відповіді у коментарях.

На думку деяких представників наукової спільноти [23] використання, наприклад, Тест Тюрінга з метою встановлення реальності співрозмовника вже не є настільки актуальним як раніше

Ймовірність відрізнити людину від штучного інтелекту на даний час складають 70%, і з кожним роком цей показник буде і надалі знижуватись.

Тест Тюрінга – створений у 1950 році з метою визначеності здатності машин до проявлення поведінки людини яку неможливо відрізнити від реальної.

Якщо показник тесту вже не є сто відсотковим, то можемо зробити висновки, що і не кожна людина зможе якісно оцінити з ким вона спілкується і чи є сторінка реальною.

В 2019 році Міністерством цифрової трансформації України було проведено дослідження на тему «Цифрова грамотність населення України» [24]

Рисунок 1.12 – рівень цифрових навичок у населення України

Отримані дані свідчать про те, що більше половина населення (53% громадян) у віці від 18 до 70 років не володіють базовими цифровими навичками, або володіють на низькому рівні.

А це означає що даній категорії буде значно важче розрізнити реальну людину від штучного інтелекту під час використання соціальних мереж. Це є показником того, що використання ШІ під час проведення масових акцій є дієвим.

1.4. Методи боротьби з ботами

За приклад у оцінці методів боротьби веб-сайтів з ботами та фейковими аккаунтами буде взято соціальну мережу «Facebook», що належить американській компанії «Meta Platforms».

Ця соцмережа була одним із головних фігурантів у розслідуванні, яке виконувалось правоохоронними органами США, частина з якого описана у попередньому розділі.

Однією з причиною виникнення сприяння у вибори була якраз роль «Facebook», який є одним з найбільших веб-ресурсів в Америці. За оцінкою дослідників було нараховано наступні дані [25]:

— Кількість аккаунтів які підверглись дії спеціальних рекламних постів - більше десяти мільйонів;

— Кількість власноруч створених аккаунтів – чотириста сімдесят;

— Кількість аккаунтів що використовувались щоденно – сім тисяч;

— Загальна кількість лайків, реакцій та коментарів – триста сімдесят мільйонів;

— Загальна кількість постів - більше восьми мільйонів.

Усі ці дані дають вносять розуміння, що без використання ботів досягти такої кількості дій математично неможливо, у випадку, якщо згідно за доказами, над цим працювала компанія штат якої складає не більше тисячі осіб.

Даний випадок можна розцінювати як перший всесвітньо масштабний прояв використання штучного інтелекту з метою зміни громадської думки. Отже, увесь трафік зловмисних ботів був згенерований з злочинними цілями.

Виникають питання до сучасного стану речей у соціальних мережах. Чи можливо зараз створити дещо схоже, нехай і у інших масштабах, та як і на скільки світові інтернет гіганти готові до боротьби з фейками та ботами.

Згідно з даними 2012 року, вже на той час у «Facebook» було близько восьмидесяти трьох мільйонів нелегітимних акаунтів. Загальна кількість користувачів на той час складала дев'ятсот п'ятдесят п'ять мільйонів. Це означає, що 8,7% від суми не належали той ці іншій особі. Сюди ж варто додати 4,8% дубльованих облікових записів, 2,4% профілів домашніх тварин та 1,5% «небажаних користувачів». Загальна кількість нормальних аккаунтів становить 82,6% від усієї кількості [25].

У тому ж році дослідження направлене на вивчення рекламного ринку в «Facebook» вказало на те, що 80% кліків які відбуваються по банерах виконуються ботами, що, у свою чергу, завдає фінансової шкоди рекламодавцям.

Після подій 2016 року Сенат Сполучених Штатів Америки опублікував список компаній, які повинні дати пояснення на предмет своєї участі у «російській спецопераційї», серед них був і «Facebook».

На питання представників влади «Як саме соціальна мережа протидіє впливу іноземних держав на американські вибори», співзасновник, виконавчий директор і президент «Facebook» Марк Цукерберг відповів: «Ми вже використовуємо штучний інтелект, аби знаходити фейкові акаунти які націлені на розповсюдження політичної агітації. Порівнюючи з роком 2014 ми зробили багато задля того, аби діяти швидко маючи при цьому розвинену інфраструктуру» [26].

Для прикладу варто взяти дослідження, що проводилось в Україні в 2014 році на предмет пошуку ботів на сторінці тогочасного президента.

Так була взята вибірка із коментарів до усіх постів, які знаходились на сторінці вказаного діяча. Із них

- 98% людей (сімдесят сім тисяч) були реальними
- 2% людей (півтори тисячі) були фейками

Але ці 2% зробили двадцять шість тисяч коментарів, що складає 15% від загальної кількості. Тобто кожний шостий коментар залишений ботом.

Отримати будь які конкретні деталі стосовно того як саме штучний інтелект працює, які критерії використовує не є можливим, оскільки це є комерційною інформацією якою обрана соціальна мережа не ділиться.

1.5 Тестування штучного інтелекту соціальної мережі

Як було зазначено, раніше, встановити основні критерії за допомогою яких «Facebook» виявляє фейкові акаунти неможливо. Але у 2019 році була опублікована статистика згідно якої соціальна мережа щоквартально видаляла близько двох мільярдів фейкових аккаунтів на квартал, використовуючи деякий алгоритм штучного інтелекту під назвою «Deep Entity Classification» (DEC) [27].

Вивчаючи документацію цього алгоритму спробуємо виділити головні особливості:

Технологічний гігант розрізняє два типи фейкових акаунтів:

1. Неправильно класифіковані облікові записи користувачів, зокрема особисті профілі для компаній або домашніх тварин, які позначені як Сторінки.

2. Особисті профілі, які займаються шахрайством і розсилкою спаму або іншим чином порушують встановлені умови обслуговування платформи.

— DEC вчиться розрізняти фальшивих і реальних користувачів за моделями підключення в мережі. Він називає ці «глибинними функціями» (їх близько двадцяти тисяч), зокрема вони включають такі речі, як середній вік або гендерний розподіл друзів користувача;

— Система починає аналізувати з використання великої кількості машинних «маркерів». Вони створюються за допомогою поєднання правил та інших моделей машинного навчання. Після того, як ці дані зібрані, за допомогою невеликої партії високоточних прикладів, позначених вручну, система порівнює дії можливого користувача із типовими діями що вчиняють люди які мають бути до нього схожі, а також із базою вже готових шаблонів фейкових користувачів.

Остаточна система класифікації може визначити один із чотирьох типів підроблених профілів:

1) нелегітимні облікові записи, які не представляють особу,
2) зламані облікові записи реальних користувачів,
3) спамери,
4) шахраї, які маніпулюють користувачами

За словами «Facebook», після впровадження DEC кількість фейкових облікових записів на платформі становила близько 5% від загальної кількості щомісячних активних користувачів.

Усі перелічені критерії, на нашу думку, першою чергою націлені саме на протидію шкідливому контенту. Безперечно, навряд хтось ,окрім штучного інтелекту, зможе протистояти самому собі. Але, в технічному описі до алгоритму не вказані види та особливості боротьби на нижчому рівні, на рівні реєстрації та створення профілів. Або використання вже створених.

З цією метою є за доцільне проведення власного експерименту для встановлення початкових критеріїв оцінки щойно створених акаунтів.

В процесі експерименту нами буде використано відразу декілька методів реєстрації аккаунтів кінцевою метою яких буде імітації дій людини.

Основними вимогами до усіх створених профілів буде виділено наступне:

— Реальне ім’я та прізвище;

— Фотографія реальної людини, що буде віддзеркалена;

— Стандартне наповнення профілю та інформації про себе;

— Використання емейлу та мобільного номеру телефону;

— Імітація звичайного користування протягом десяти хвилин після створення.

Загальна кількість протестованих аккаунтів буде складати двадцять два.

Із них:

1. Дванадцять акаунтів буде створено на реальні фізичні SIM-карти:

a) Два аккаунти буде створено за допомогою мобільного телефону, нічого не змінюючи;

b) П’ять аккаунтів буде створено на одному комп’ютері, використовуючи одну IP-адресу, щоразу використовуючи режим «Інкогніто»

c) П’ять аккаунтів буде створено на одному комп’ютері використовуючи різні IP-адреси, щоразу використовуючи режим «Інкогніто»

2. Десять акаунтів буде створено на віртуальні SIM-карти:

a) П’ять аккаунтів буде створено на одному комп’ютері, використовуючи одну IP-адресу, щоразу використовуючи режим «Інкогніто»

b) П’ять аккаунтів буде створено використовуючи різні IP-адреси різних країн, щоразу використовуючи режим «Інкогніто»

3. Шість аккаунтів буде куплено в мережі інтернет. Головною вимогою до них буде максимальна неповність профілю.

Варто зазначити, що всі вищезазначені дії будуть використовуватися лише з метою експерименту і в незалежності від результаті акаунти будуть видалені по закінченню, а SIM-карти і електронні пошти, якщо це буде можливим, будуть відв’язані від сторінок.

Абсолютно в усіх випадках ми будем працювати використовуючи браузерний режим «Інкогніто».

Причиною цьому, на наш погляд, є забезпечення віл запису так званих «куків», особливістю яких є зберігання певної частки інформації отриманих від веб-сайту на нашому (клієнтському) боці і повторна їх відправка при повторному відвіданні сайту, оскільки в такому випадку соціальна мережа відразу розпізнає факт підміни даних, що може вплинути на ризик завчасного блокування сторінок.

Етап 1. День 1.

Рисунок 1.13 – фізичні SIM-карти які будуть використовуватися для першого етапу

1. Використовуємо мобільний телефон марки «Xiaomi» на 2 SIM-карти який знаходиться в стані після скидання до «заводських налаштувань». На ньому встановлена остання версія програмного забезпечення від компанії «Facebook». Раніше будь-які аккаунти в зазначеній соціальній мережі на ньому не створювались.

2. Після вводу даних на мобільний телефон надіслано СМС с кодом підтвердження. Вводимо – вдало. Аккаунт створено. Додаємо фото та певний опис сторінки. Десять хвилин гортаємо стрічку.

3. З цього ж телефону, але вже використовуючи браузер повторюємо дії з попередніх пунктів. Другий профіль створено.

4. На цьому Етап 1 завершено.

Етап 2. День 1.

На другому етапі ми будемо реєструвати аккаунти вже використовуючи комп'ютер.

1. Використовуємо ноутбук марки «Lenovo». На ньому раніше не був встановлений «Facebook». Для отримання СМС с кодом буде задіяний телефон з фізичними клавішами марки «LG».

2. Відкриваємо браузер в режимі «Інкогніто» проводимо реєстрації аналогічно Етапу 1.

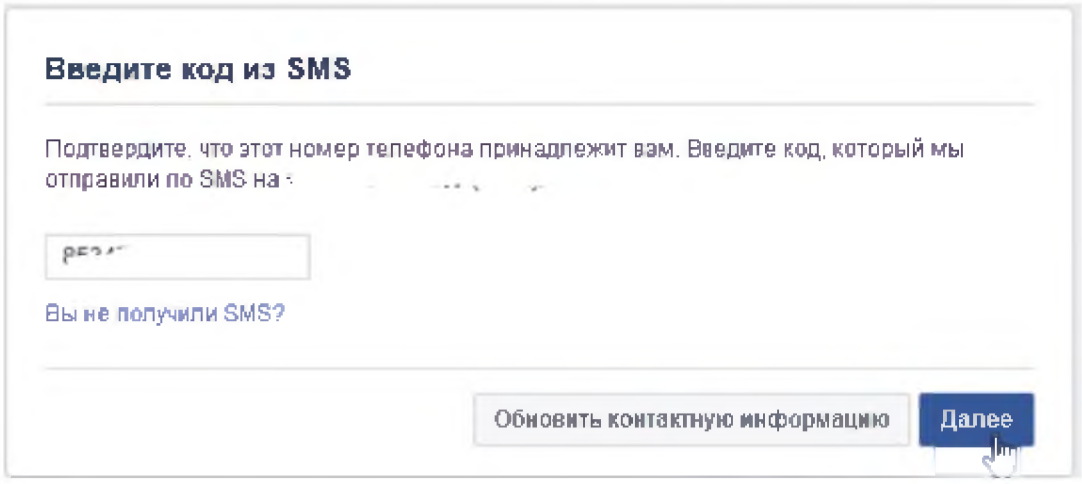

Рисунок 1.14 – вікно для вводу коду підтвердження.

3. Акаунт створено. Повторюємо дії.

4. Повторюємо пункт 2. Другий акаунт створено.

5. Повторюємо пункт 2 другий раз. Після вводу коду отримуємо помилку (рисунок 1.14)

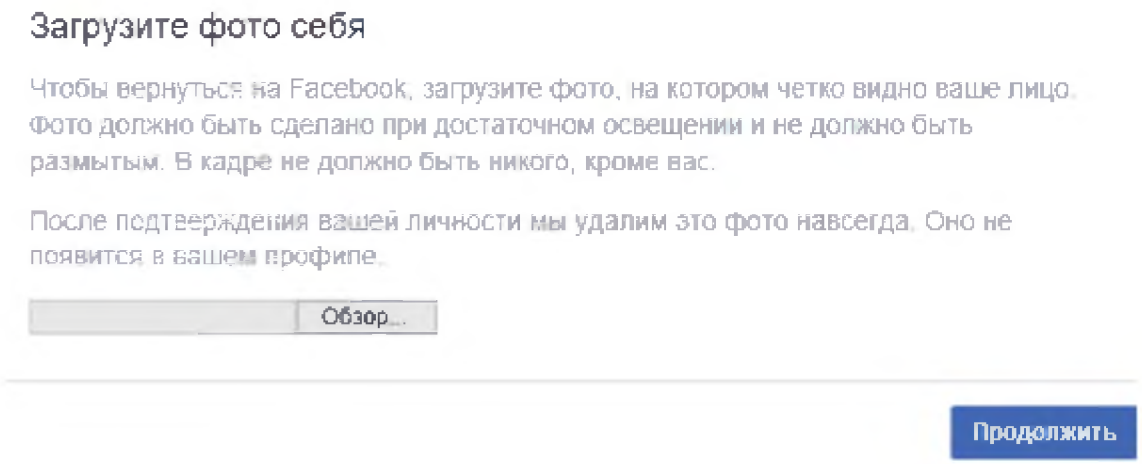

Рисунок 1.15 – помилка яка виникла під час реєстрації.

6. Обираємо попереднє заготовлене фото та відсилаємо його. Нічого не відбувається. Переходимо до наступного етапу.

7. Реєструємо ще один аккаунт. Отримуємо помилку (рисунок 1.15).

Етап 3. День 1.

На третьому етапі ми будемо реєструвати аккаунти використовуючи комп'ютер разом із програмним засобом для зміни IP адреси. Країну яку ми будемо використовувати буде Україна, аби відповідати фізичним SIM-картам.

1. Повторюємо кроки з попереднього етапу.

2. Акаунти зареєстровано. Проблем не виявлено.

Етап 4. День 1.

На четвертому етапі ми будемо реєструвати аккаунти використовуючи комп'ютер але замість фізичних номерів SIM-карт будуть віртуальні, що взяті в оренду.

1. Відтворюємо класичну реєстрацію.

2. Використовуючи сервіс з орендованими номерами вводимо код із СМС повідомлення. Перший аккаунт створено.

3. Під час створення другого аккаунту виникає аналогічна помилка (рисунок 1.15).

4. Створення подальших профілів блокується помилкою

Етап 5. День 1.

На п’ятому етапі ми будемо реєструвати аккаунти використовуючи комп’ютер, віртуальні SIM-картки та програмне забезпечення для зміни IP адреси.

1. Відтворюємо класичну реєстрацію.

2. Під час створення четвертого аккаунта перестали отримувати СМС-коди. Підтримка ресурсу запевнила що проблема на стороні «Facebook».

3. Три профілі створені.

Етап 6. День 1.

У мережі інтернет придбали шість аккаунтів. Три з них буде використовуватись під однією IP адресою. Інша половина – під іншими. Один акаунт – одна персональна адреса.

1. Виконуємо вхід до перших трьох аккаунтів. Проблеми на етапі не виникли.

2. Виконуємо вхід до других трьох аккаунтів. Проблеми на етапі не виникли.

Етап 7. День 2.

У другий день буде перевірено стан тих профілів, які були створені попередньо.

Таблиця 1.2 – підсумки експерименту

Тип створення	За планом створено	Працювало у перший день	Працювало у другий день	Загалом
Телефон, фіз. SIM-карта	2	1	1	1
Комп’ютер Одна IP адреса фіз. SIM-карта	5	2	1	1
Комп’ютер Різні IP адреси	5	5	4	4

Продовження таблиці 1.2

Тип створення	За планом створено	Працювало у перший день	Працювало у другий день	Загалом
Комп'ютер Одна IP адреса віртуальна. SIM-карта	3	1	0	0
Комп'ютер Різні IP адреса віртуальна. SIM-карта	3	3	1	1

На другий день експерименту виникли певні проблеми з деякими з новостворених акаунтів, зокрема ті, що були зареєстровані з однієї IP адреси (окрім телефону) запросили підтвердження (рисунок 1.15). У випадку з застосуванням фізичної SIM-картки ця процедура була пройдена і акаунт розблокований, але розблоковати сторінку, яка була створена з використанням віртуальної «сімки» не вдалось можливим, оскільки доступу до неї ми не маємо.

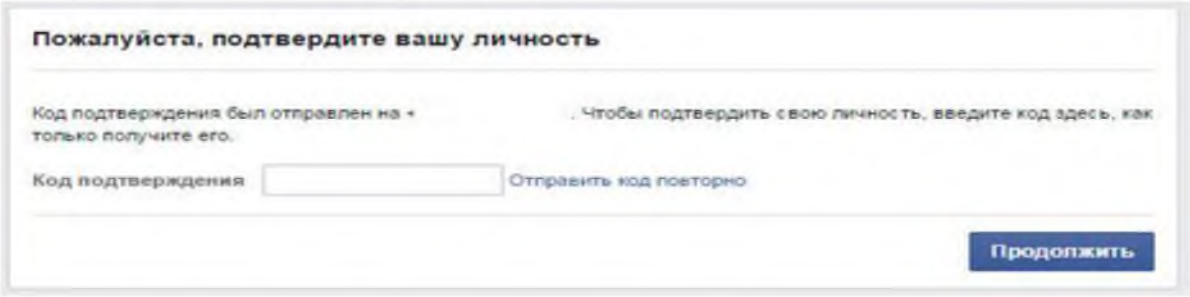

Рисунок 1.15 – підтвердження особистості.

Підводячи результати експерименту варто зауважити, що найкраще себе проявили акаунти які були створені під час використання різних IP адрес, можливо, в майбутньому, користування такими профілями призвело до блокування, але на етапі реєстрації проблем з ними невиникло. Виходячи з цього, можемо зробити висновки, що в процесі боротьби з ботами на «нижчому рівні» «Facebook» в першу чергу звертає увагу на IP-адресу, опускаючи багато інших критеріїв.

1.6 Оцінка мережевого трафіку, що генерується людиною

Не дивлячись на активний розвиток програмно-апаратних машин та збільшення їх ролі у питанні генерації трафіку, станом на 2020 рік 59% від загальної кількості, створюється саме людиною.

Рисунок 1.16 – Кількість трафіку в (%) що згенеровано людиною

Динаміка змін, яка зображена на рисунку 1.16 вказує на те, що доля людського трафіку протягом останніх п’яти років не знижувалась нижче 50%.

Для детального розгляду варто розуміти статистичні тенденції використання зовнішньої мережі (мережі інтернет) пересічним громадянином України.

Дослідження, що проведене у 2021 році з метою оцінки рівня діджиталізації різних світових країн подає наступні дані про нашу державу [28]:

— Загальна кількість користувачів інтернет – 29,47 мільйонів, що є 67,6% від загальної кількості населення;

— За 2020 рік доля активних юзерів зросла на 2 мільйони, тобто плюс 7,3%;

— 96,4% із цих людей використовують для доступу у мережу інтернет мобільні пристрої.

Загалом ці показники є краще за середні у світі. За умови розбиття усіх проаналізованих країн на графік, Україна би входила у першу половину даного переліку.

Найпопулярнішими є сайти категорії онлайн-купівель, ресурси, що містять контент для дорослих, соціальні мережі, та банківські додатки.

Ці всі дані проаналізовані шляхом використання підрахунку метрик при відвідуванні ресурсів які знаходяться у так званій «Всесвітній мережі», або яка більше відома як (англ.) World Wide Web [29]. Особливості побудови якої дають змогу збирати таку статистику.

Але зовнішній інтернет включає в себе і інші категорії, відстежити які куди важче, або майже неможливо, зокрема так звану «Темну мережу» (англ. Dark Web). Вона є складовою загальною всесвітньої «павутини», але доступ до неї можна отримати використовуючи тільки спеціальне програмне забезпечення.

Для зручності у здійснення людиною пошуку, усі сайти індексуються пошуковими машинами, тобто вони вносяться до загального пулу усіх пошукових систем, аби у майбутньому за певними критеріями їх було можливо віднайти.

У ситуації з «Dark Web» відбувається навпаки. Це та частина інтернету яка не індексується пошуковими гігантами і те, куди не є за реальним потрапити без спеціальних попередніх дій.

Як було зазначено раніше, одним із способів отримання доступу є використання спеціалізованого програмного забезпечення (частіше – типу браузер). Одним із таких, як приклад, є «Tor», який публікує власну статистику використання [30].

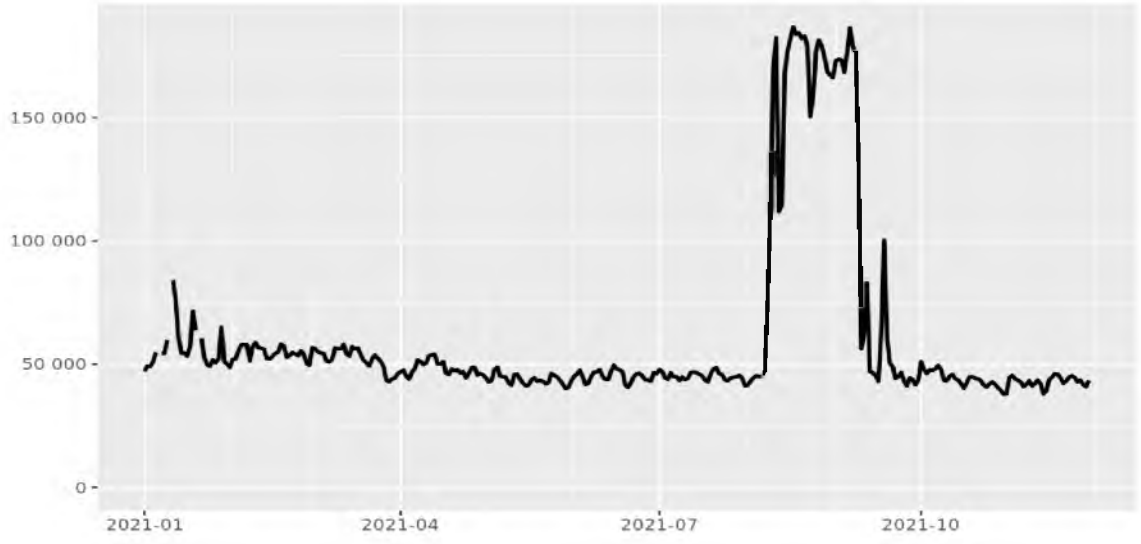

Рисунок 1.17 – Кількість користувачів браузера «Tor»

На рисунку 1.17 зазначена статистика за кількістю користувачів у період з початку 2021 року до грудня місяця. Так середня кількість користувачів саме з України становила близько п’ятдесяти тисяч. Іноді цей показник зростав втричі.

Отримати внутрішні статистичні дані з зазначеної підмережі не є можливим враховуючи специфіку програної реалізації зазначеного продукту.

1.7 Детальна класифікація «чорного» інтернет трафіку

В пункті 1.1 даної роботи нами було запропоновано власну класифікацію видів трафіку, де був виконаний попередній розподіл на:

- «Білий» трафік
- «Сірій» трафік
- «Чорний» трафік

Додатково, були виділені поверхневі критерії за якими можна надати оцінку кожному з цих видів. А саме:

— порушення законодавства,

— порушення загальноприйнятних норми і правил експлуатації та використання електронно обчислюваних машин,

— ускладнення роботи мережевих пристроїв,

— створювати протиправні дії та контент,

— використовувати штучний інтелект як імітацію людини

В зазначеному розділі буде надана більш детальна характеристика кожному із критеріїв які ми використовуємо для оцінки

1) Інтернет трафік кінцевою метою якого ж порушення законодавства

В даному випадку, і надалі буде використовуватися нормативно-правова база України.

Під порушенням законодавства, в нашому випадку, мається на увазі вчинення будь якого діяння з використанням електронно-обчислювальних машин, за вчинення якого передбачена юридична відповідальність.

Такі порушення чітко регламентуються законами и актами, зокрема і Кримінальним Кодексом України, сюди належать [31]:

Стаття 176 ККУ. Порушення авторського права і суміжних прав.

Згідно з статистикою, 61% користувачів мережі інтернет, що проживають на території нашої держави мінімум раз на тиждень переглядають мультимедійний контент. А в переліку найпопулярніших сайтів за січень 2021 року тринадцяте та п'ятнадцяте місце займають ресурси, що, на наш погляд, ретранслюють відео контент не маючи на це юридичного права, оскільки вони не є власниками цього матеріалу. Власники даних ресурсів і є генераторами «чорного трафіку» спонукаючи і надаючи можливість людям до перегляду зазначених матеріалів, та порушення законодавства.

Стаття 190, ч.3. Шахрайство.

Шахрайство, вчинене у великих розмірах, або шляхом незаконних операцій з використанням електронно-обчислювальної техніки.

Визначення цього поняття дає наступне – це заволодіння чужим майном або придбання права на майно шляхом обману чи зловживання довірою.

В якості прикладу буде наведено ресурс «OLX», що являє собою платформу онлайн-оголошень для покупки чи продажу послуг та речей. Та згідно зі статистикою є одним із найпопулярніших сайтів країни, займаючи четверте місце у загальному рейтингу та маючи середній час знаходження на сайті тринадцять хвилин.

За даними одного із досліджень під час використання зазначеного ресурсу, 35% українців стикалися зі зловмисниками під час покупок в інтернеті. Загальна кількість випадків склала дванадцять тисяч триста шістдесят дев'ять, що є 0,4% від загальної кількості.

А ось за даними Департаменту Кіберполіції Національної поліції України, ними було отримано понад тридцять тисяч звернень щодо шахрайства в інтернеті [32].

У всіх вищезазначених даних не враховано ту кількість порушень законодавства, в результаті яких постраждалі особи не звертались як мінімум на лінію підтримку тієї чи іншої платформи або до правоохоронних органів.

Також є статті що надані у розділі XVI ККУ, під назвою – кримінальні правопорушення у сфері використання електронно-обчислювальних машин, систем та комп'ютерних мереж і мереж електрозв'язку, вони включають:

Стаття 361. Несанкціоноване втручання в роботу електронно обчислювальних машин (комп'ютерів), автоматизованих систем, комп'ютерних мереж чи мереж електрозв'язку.

Стаття 361-1. Несанкціоноване втручання в роботу електронно-обчислювальних машин (комп'ютерів), автоматизованих систем, комп'ютерних мереж чи мереж електрозв'язку.

Стаття 361-2. Несанкціоновані збут або розповсюдження інформації з обмеженим доступом, яка зберігається в електронно-обчислювальних машинах (комп'ютерах), автоматизованих системах, комп'ютерних мережах або на носіях такої інформації.

Стаття 362. Несанкціоновані дії з інформацією, яка оброблюється в електронно-обчислювальних машинах (комп'ютерах), автоматизованих системах, комп'ютерних мережах або зберігається на носіях такої інформації, вчинені особою, яка має право доступу до неї.

Стаття 363. Порушення правил експлуатації електронно-обчислювальних машин (комп'ютерів), автоматизованих систем, комп'ютерних мереж чи мереж електрозв'язку або порядку чи правил захисту інформації, яка в них оброблюється.

Стаття 363-1. Перешкоджання роботі електронно-обчислювальних машин (комп'ютерів), автоматизованих систем, комп'ютерних мереж чи мереж електрозв'язку шляхом масового розповсюдження повідомлень електрозв'язку.

Але усі вищезазначені статті більшою мірою пов'язані із, або фізичним втручанням, або використання саме трафіку в цьому випадку менше ніж у випадках передбаченими статтями 190, ч.3 та 176.

2) Порушення загальноприйнятних норми і правил експлуатації та використання електронно обчислюваних машин.

До визначеного переліку входять (але не обмежуються) наступні пункти:

- Розміщення образливих матеріалів у мережі (булінг);
- Створення, надсилання або перегляд порнографічних матеріалів;
- Нехтування конфіденційністю файлів інших людей;
- Надання своїх облікових даних іншим особам
- Видавання себе за когось іншого у мережі;

В незалежності від формату, масштабів та кількості, виконання будь якого із цих пунктів не несуть суспільну користь, а їх використання не є загальноприйнятним.

3) Ускладнення роботи мережевих пристроїв

Цей критерій був виділений для узагальнення DoS та Ddos атак з метою вчинення яких, як правило, використовують значні потужності електронно-обчислювальних машин.

На даний час розрізняють використання двох основних варіантів організації:

— За допомогою Ботнету. Це, як правило, раніше підготовлені пристрої, або зламані пристрої, які одночасно здійснюють запити до атакованого серверу;

— За допомогою попередньої домовленості великої кількості користувачів.

Рисунок 1.18 – Статистика кількості DdoS-атак

В період першого кварталу 2021 року було зафіксовано біля тисячі щоденних атак [33]. Оскільки більшість сучасних систем мають хоча б мінімальний захист, то це провокує зростання потужності з боку нападників, що в свою чергу створює додаткове мережеве навантаження, завдаючи шкоди користувачам.

1.8 Фактори, що сприяють зростанню злочинного трафіку

На сьогоднішній день в Україні не ведеться статистика за зміною показників усіх тих критеріїв, які були відділені нами попередньо, з метою маркування кінцевої мети трафіку. Тому задля висвітлення зазначеної проблеми буде використані дані які можна перевірити за офіційними відкритими джерелами.

А саме - «Кількість кіберзлочинів, за якими була заведена кримінальна справа у період з 2011 по 2019 роки». Що підпадає під критерій (А) вказаного раніше.

Рисунок 1.19 – Кількість відкритих кримінальних справ за злочинами пов’язаними з кіберзлочинністю у період з 2011 по 2018 роки

Стосовно рівня кіберзлочинності та її динаміки варто зазначити значну тенденцію зростання, ще в 2011 році в Україні було зареєстровано сто тридцять один випадок злочинів у сфері використання електронно-обчислювальних машин (комп'ютерів), систем та комп'ютерних мереж і мереж електрозв'язку, у 2012 році – сто тридцять вісім, у 2013 році – п'ятсот дев'яносто п'ять, у 2014 році – чотириста сорок три, у 2015 році – п'ятсот дев'яносто вісім, у 2016 році – вісімсот шістдесят п'ять, у 2017 році – дві тисячі п'ятсот сімдесят три та дві тисячі вісімсот двадцять сім у 2018 році відповідно [34].

Кіберзлочинність – це п'ятий за економічною значимістю вид злочинності в Україна, слідом за такмим видами як: незаконне привласнення майна, хабарництво, корупція та злочини як використовуються з метою підриву конкуренції і маніпуляцією з фінансовою звітністю.

За результатами опитування, на кіберзлочинність припадає 23% випадків шахрайства у світі в Україні цей показник складає – 17%.

Загальна кількість кримінальних справ із початка 2015 року на кінець 2019 склала близько чотирьох тисяч вісімсот, при цьому вироків було винесено лише сто тринадцять. На нашу думку, фактори які вливають на таке співвідношення показників кримінальних проваджень до вироків можна розділити на два ключові показники:

- Юридично-правові;
- Технічні.

До першої категорії можна віднести, зокрема, наступне [35]:

Остіні глобальні зміни до більшості кодексів та нормативно-правових актів, які стосувалися питання відповідальності за злочини в сфері електронно-обчислювальних машин відбувалися в 2012 році. З того моменту минуло майже десять років. За цей час змінилась як кількість, так і характер правопорушень у цій сфері.

Мала кількість та низькій освітній рівень в зазначеній сфері у: слідчих, прокурорів, суддів. На нашу думку, рівень юридичного документування наразі є

доволі низьким. В Україні лише з 2014 року почалось навчання профільних фахівців, як, зокрема, у майбутньому можуть працювати в сфері боротьби з кіберзлочинністю. Їх навчання відбувається на базах юридичних вищих навчальних закладів, але підготовка відбувається лише для лав національної поліції. Профільне навчання прокурорів і суддів наразі не відбувається. Для прикладу в США відбувається підготовка спеціалістів за цими критеріями, оскільки в їх законах закладено поняття «першочерговості дії», де сказано, що від здібностей суб'єкта розслідування залежить чи буде розкрито злочин і чи понесе винний відповідальність.

Низька швидкість проведення експертиз. З практичної точки зору для доказу вичерпного переліку злочинів необхідне проведення судової експертизи. Їх перелік регламентований кримінально-процесуальним кодексом України.

В випадку, що стосується злочинів з електронно-обчислювальними машинами, найчастіше використовується комп'ютерно-технічна та програмно-технічна експертиза від результатів яких залежить продовження процесуальних дій. Її проводять експерти. Але з практичної точки зору кількість спеціалістів досить мала, а більша їх частина не має наявних фахових навичок для детального дослідження, що в свою чергу збільшує час розслідування та частково паралізує роботу відповідних правоохоронних органів.

Цей перелік не є вичерпним. На нашу думку його варто навести для отримання комплексного уявлення про сучасний стан проблеми, та враховуючи тему науково-дослідної роботи, головним питанням є саме технічні критерії, що не дають змогу якісно протистояти «чорному» трафіку.

Серед основних технічних проблем варто виділити наступне [36]:

1. Відсутність загальноприйнятного (на державному рівні) спеціального програмно-апаратного забезпечення яке б автоматизувало протидію зловмисному трафіку.

Серед прикладів варто виділити рішення про блокування переліку російських соціальних мереж та веб-ресурсів, яке було прийняте керівництвом України в 2017 році. Опускаючи аргументацію та причини, сама реалізація

відбувалася шляхом ручного блокування провайдерами зазначеного переліку веб-ресурсів. Загальних технічних вимог виділено не було, окрім головної – створення умов для неможливості доступу до зазначених веб-сайтів. Деякі з інтернет-провайдерів виконали блокування лише самого домену, інші – додали у «чорний список» IP та DNS адреси.

2. Відсутність спеціального програмно-апаратного забезпечення для пришвидшення та автоматизації аналізу даних з пристроїв.

Навіть у випадку встановлення особи або осіб, які генерують злочинний трафік необхідне оперативне отримання та зняття інформації для пришвидшення та оптимізації процесу розгляду правопорушення. Зокрема, використання спеціальних апаратних пристроїв.

Їх використання, у деяких випадках, забезпечує як фізичне зняття інформації так і процес зламу встановлених ключів безпеки та дешифрації крипто-контейнерів в яких зберігається інформація.

Рисунок 1.20 – Приклад пристрою для зняття інформації

Високий рівень анонімізації трафіку під час використання інтернет. Сама суть полягає у використанні користувачами мережі спеціальних програмних та

апаратних засобів з метою видалення або приховування даних з метою запобігання ідентифікації джерела трафіку і кінцевої точки.

«Більшість видів загальнокримінальних злочинів вже здійснюються із застосуванням методів, які використовують кіберзлочинці – це анонімізація, використання зашифрованих каналів зв'язку та передачі даних, а також використання криптовалюти та різних неофіційних платіжних систем для вчинення злочинів, в тому числі – організованих» - заявив в 2019 році, на той час, керівник Департаменту кіберполіції С. Демедюк [37].

Якщо брати до уваги перші два пункти основних технічних проблем, то можливі варіанти їх вирішення полягають у закупівлі спеціального програмно-апаратного забезпечення.

В першому випадку – для організацій, що надають послуги з провайдингу, в другому – для представників правоохоронних органів. Але останнє питання, на нашу думку, є більш комплексним і складним та потребує додаткової уваги.

1.9 Проблематика анонімізації

Як було зазначено раніше, анонімізація це процес видалення або приховування певних або усіх даних в мережі, які виходять в результаті користування [38].

Головною метою є залишитися неідентифікованим під час використання мережі інтернет.

За своєю кінцевою метою, анонімність можна поділити на такі частини:

- Безпекова:

Використання в такому випадку є виправданим та необхідним, оскільки стосується питання безпеки людей, чия професійна діяльність може безпосередньо нести ризики життю та здоров'ю у разі, якщо їх особа буде встановлена. Як правило це журналісти, правозахисники, співробітники правоохоронних органів, свідки, тощо.

Також анонімність використовують і юридичні особи, їх ціль – захист інформаційної системи від, наприклад, Ddos-атаки, яку значно важче виконати не маючи повних відомостей про кінцеву точку.

- Злочинна:

В цьому ж випадку, головною метою користувача є, також, створення умов для приховування своєї особи, але суб'єктивна сторона кардинально протилежна.

Так, наприклад, на думку деяких психологів, анонімне спілкування в мережі інтернет викликає так званий «онлайн-ефект розгальмування» (англ. Online disinhibition effect). Це ослаблення психологічних бар'єрів, що обмежують вихід прихованих почуттів та потреб, що змушує людей поводитися в Інтернеті так, як вони зазвичай не надходять у реальному житті. Розрізняють два види «розгальмовування» [39]:

- Позитивне;
- Токсичне;

В першому випадку люди почуваються вільніше, дають волю прихованим емоціям в інтернет-просторі і навіть виявляють несподівану доброту та великодушність.

В другому – пов'язане з бажанням задоволення потреб що можуть створювати перешкоди для інших. Сюди можна віднести питання булінгу, яке розглядалось раніше.

Але ці показники вирізняються саме з психологічної точки зору. Тим не менш, саме психологічне підґрунтя нерідко є причиною злочинних діянь.

До більш серйозних, аніж психологічний тиск, діянь, з використанням анонімності, можна віднести безпосереднє, вчинення злочинів в мережі інтернет за які передбачена відповідальність.

Примітка: цей перелік надано у розд41ілі 1.7 даної роботи

В цілому, якщо взяти за основу виділений у пункті 1.1 набір критеріїв, які ми виділили з метою класифікації кінцевої мети трафіку (а – е), то до усіх них можна застосувати анонімізацію, а іноді це і взагалі є основою.

На сьогоднішній день виділяють чотири основні методи для організації анонімного трафіку :

- Анонімні мережі;
- Анонімні проксі-сервера;

- Анонімні веб-проксі;
- Анонімні VPN-сервіси.

Анонімні мережі – спеціальні комп'ютерні мережі головною метою створення яких є забезпечення анонімності користувача [40].

Архітектура роботи системи побудована на з'єднані класу P2P (англ. «Peer-to-peer»), принципом дія якого рівноправність усіх підключених пристроїв.

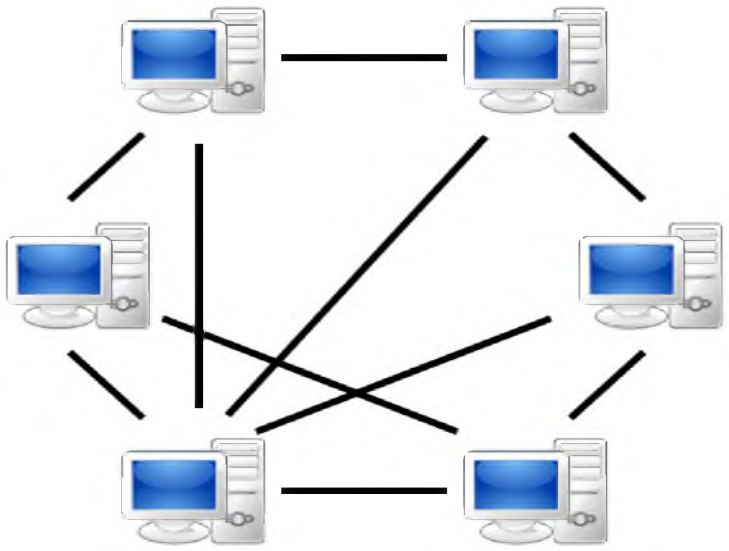

Рисунок 1.21 – Мережа типу peer-to-peer.

Набуття анонімністі в таких мережах відбувається шлязом направлення пакетів даних через інші однорангові вузли. Це не дає змогу отримати достовірну інформацію на предмет того хто саме завантажує файли та генерує інтернет трафік. У випадку, якщо якийсь з пристроїв перестає працювати у мережі – це не паралізує її роботу, а лише знижує швидкість. Тобто, навість у випадку наявності двох машин мережа буде діяти.

Додатковим фактором є використання програмного забезпечення яке надає доступ до мережі та має вбудовану функцію шифрування.

Анонімні мережі поділяються на:

- Децентралізовані

В даному типі мережі будь який пристрій може з'єднуватись з будь яким іншим пристроєм і кожна машина є, певною мірою, сервером який забезпечує

працездатність усієї мережі, отримуючи та надсилаючи інформацію, тим самим генеруючи трафік.

- Гібридні

В гібридних анонімних мережах між пристроями також існують сервери, які використовуються для покращення роботостійкості та забезпечують централізованість роботи усіх підключених машин. Як і в випадку децентралізованих, передача трафіку між вузлами додатково шифрується.

В якості прикладу варто розглянути найбільшу із відомих анонімних гібридних мереж – «TOR» (англ. «The Onion Router»).

У розділі 1.6 нами зверталась увага на дану мережу, де, зокрема, зазначалось, що мінімум, нею користується п'ятдесят тисяч українців щомісяця. Іноді цей показник зростав втричі.

«TOR» є спеціальним вільним програмним забезпеченням з відкритим кодом, типу браузер, що забезпечує створення та реалізацію підключення до, так званої, «лукової мережі» (англ. Onion routing).

Принцип роботи полягає у маршрутизації трафіку через всесвітню мережу серверів та пристроїв (вузлів) з метою анонімності. Запускаючи браузер, користувач створює на своєму пристрої проксі-сервер який підключається до загальної мережі, періодично створюючи певний вузол.

Особливістю цибулевої маршрутизації є технологія, в якій трафік неодноразово шифрується і надсилається через перелік серверів, так званих цибулевих маршрутизаторів кожний з яких видаляє свій шар шифрування, і так до кінцевої точки, в якій буде розшифровано останній шар і отримана сама інформація. Цей принцип дозволяє залишати трафік анонімним навіть, якщо один із вузлів буде скомпрометований, окрім першого та останнього.

Кожний пакет даних проходить через, щонайменше, три вузли і відповідну кількість разів шифрується та розшифровується.

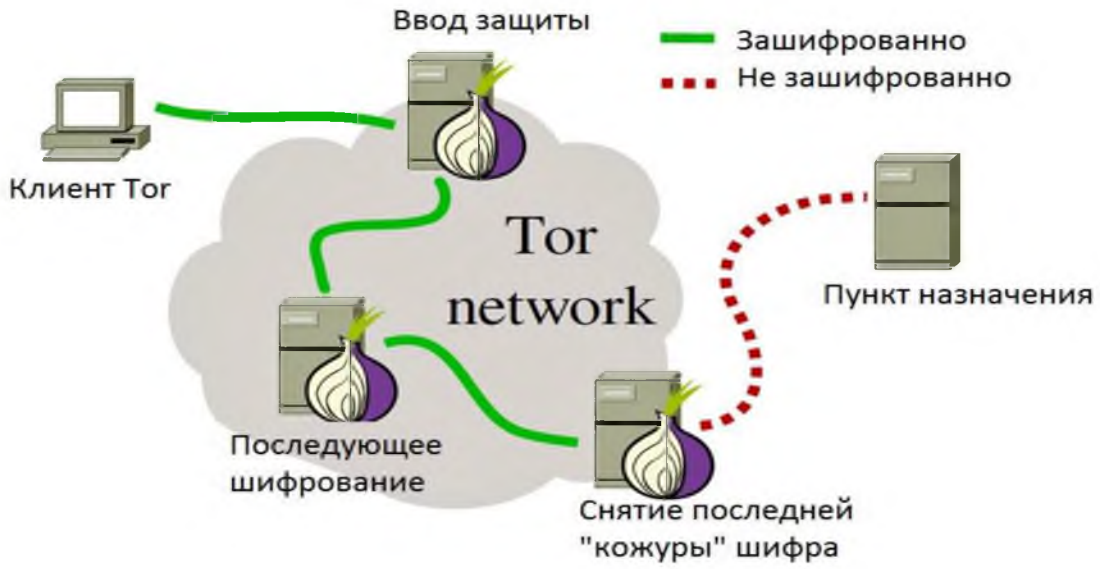

Рисунок 1.22 – Схема мережі

Із головних недоліків варто виділити низьку швидкість з'єднання та передачі даних, що пов'язана з кількістю вузлів та загальним навантаженням системи. Також є ризик можливого втручання в вихідний вузол, дані з якого не шифруються.

Головною вимогою до загальної безпеки анонімних мереж, є посилене вивчення програмного забезпечення та самої мережі на предмет пошуку «бекдорів».

Анонімні проксі-сервери [41].

Проксі сервер – сервер, що виступає в ролі посередника в мережі між користувачем та цільовим кінцевим сервером.

Спочатку відбувається запит користувача до самого проксі-сервера, після чого останній підключається до вказаного, у запиті джерела. Зворотній шлях відбувається у дзеркальному порядку.

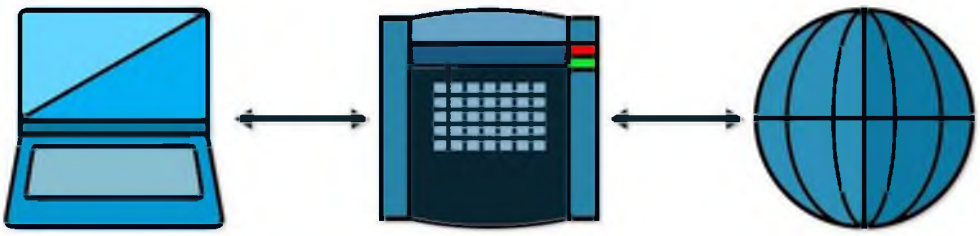

Рисунок 1.23 – Схема проксі-мережі

Технічна складова налаштувань передбачає, що клієнтський пристрій налаштований таким чином, що увесь трафік по певному, раніше зазначеному, протоколу здійснюється не напряму на адресу кінцевого пристрою, а на безпосередньо заданий проксі-сервер. Після отримання запиту, сервер його перевіряє коректність, та не розриваючи з'єднання створює новий запит вже від свого імені. Таким чином, власник кінцевого маршруту отримує інформацію саме про сервер «прокладку», а не про машину з якої запит відбувався від самого початку.

В разі встановлення певних налаштувань, проксі може забезпечувати анонімність доступу до ресурсів, приховуючи відомості про джерело запиту, або спотворюючи їх.

Сервери бувають двох видів:

— Прозорий проксі – варіант налаштувань, при якому трафік перенаправляється засобами маршрутизатора неявно.

— Зворотній проксі – на відміну від прямого відповідає саме за отримання інформації із мережі. Для прикладу, трафік який надходить ззовні, потрапляє на проксі який автоматично та рівномірно розподіляє його між внутрішніми серверами.

З точки зору безпеки, такий варіант анонімізації дійсно забезпечує певну скритність під час використання, застосовуючи «посередника», який у випадку атаки або перевірки видасть інформацію про себе, а не користувача. З іншого боку, якщо не встановлені додаткові налаштування або спеціалізоване програмне забезпечення, то трафік не шифрується, а лише забезпечує додаткову маршрутизацію. Це є засобом анонімності, але не достатньо надійним, порівнюючи з альтернативними методами забезпечення.

Анонімні веб-проксі – виступає певним аналогом звичайного проксі-серверу, які схожі за принципом дії – увесь трафік проходить через додатковий пристрій. Але відмінність полягає в тому, що частіше за все це є спеціальний веб-додаток, який має бути встановлений на веб-сервері. В випадку, якщо останні умови не виконані – неможливо забезпечити маршрутизацію тим самим цей засіб

є надійним лише у випадку застосування його на власних або довірених веб-ресурсах.

Анонімні VPN сервіси – VPN (англ. Virtual Private Network) – віртуальна приватна мережа – технологія, що розширює певну приватну мережу через загальнодоступну або мережі із меншим рівнем довіри [42].

Принцип роботи складається з створення так званого VPN-тунелю, який створюється мінімум між двома пристроями та дозволяє отримувати доступ до закритих мереж, або забезпечувати додатковий рівень безпеки.

Як правило, безпека такої мережі забезпечена реалізацією шифрування, внаслідок чого створюється закритий канал обміну та передачі. Навіть у випадку компрометації та зламу цього тунелю буде отримано шифрований набір даних що не дасть змогу ідентифікувати сам трафік.

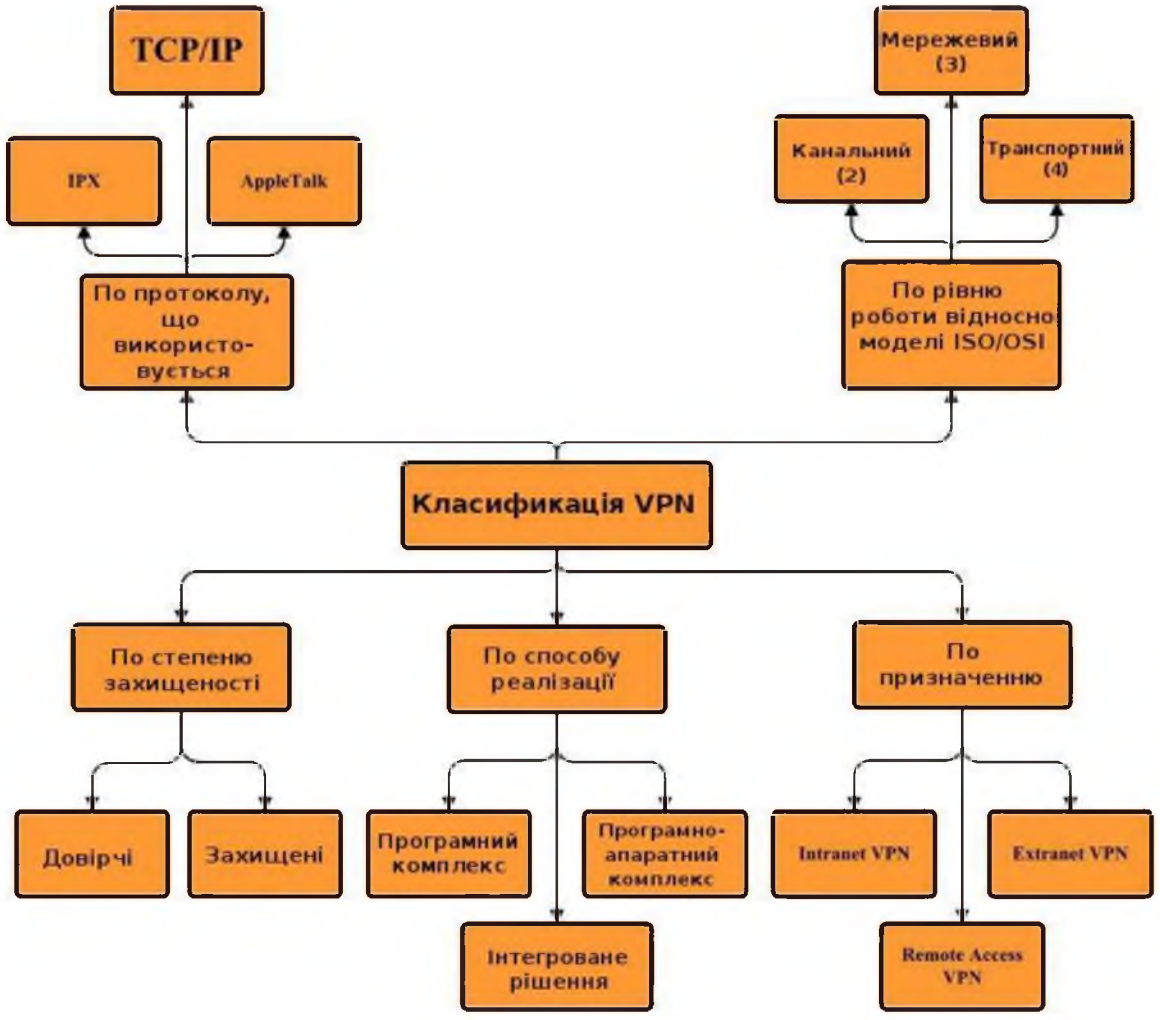

Рисунок 1.23 – поділ VPN за критеріями

Віртуальні приватні мережі поділяються на [43]:

— Захищені – створюються з метою захисту інформації та додаткового тунелю поверх ненадійних мереж, таких, зокрема, як мережа Інтернет.

— Довірчі – використовуються коли треба передавати дані за певними маршрутами усередині великої мережі. Оскільки використовується тунелювання, це зменшує ризик перерозподілу трафіку. В такому випадку на перше місце виставляється саме автоматизація мережевої передачі, і вже потім безпеки.

За методом реалізації:

— За допомогою спеціального програмно-апаратних засобів;

— За допомогою спеціального програмного забезпечення;

— Інтегроване рішення – наявність фізичного програмно-апаратного комплексу.

За призначенням:

— Intranet VPN - використовується для об'єднання в єдину захищену мережу кількох розподілених пристроїв однієї організації, що обмінюються даними відкритими каналами зв'язку.

— Remote-access VPN - використовується для створення захищеного каналу між сегментом корпоративної мережі та певним користувачем який перебуває за ним. Як приклад – віддалена робота.

— Extranet VPN - використовується для мереж, до яких підключаються «зовнішні» користувачі. Як приклад – клієнти.

— Internet VPN - використовується інтернет-провайдерами для надання доступу до інтернету, якщо по одному фізичному каналу підключаються кілька користувачів.

— Client/server VPN – використовується коли VPN будується між вузлами, що знаходяться, як правило, в одному сегменті мережі, наприклад, між робочою станцією та сервером.

Головною особливістю зазначеного виду забезпечення анонімізації є наявність певного серверу який буде забезпечувати тунелювання від першої до

кінцевої точки. Тобто для використання цього методу потрібна або фізична наявність такого серверу, або оренда пропонованих.

В незалежності від обраного методу для анонімізації трафіку, кінцева мета досягається усіма з вищезазначених засобів.

1.10 Висновки розділу

В результаті виконання першого розділу, було надано коротку характеристику та інформацію про види та типи трафіку. Проаналізовано поточний стан та динаміку числових змін.

Проведено розділення трафіку за походженням, маршрутом. Виділено власні критерії на предмет кінцевої мети та генерації. Надано визначення та перелік критерій «чорному (злочинному) трафіку».

У підрозділі використання трафіку машинними засобами, було проаналізовано трафік, що генерується машинами. Виділені поняття «ботів» з переліком розмежування їх за категоріями та діями. Визначено основні показники за якими можна провести їх категоризацію на «Good Bot» та «Bad Bot»

У наступному підрозділі було детально проаналізовано саме категорію поганих (зловмисних) ботів, з виділенням додаткових критеріїв, аналітики міжнародних досліджень. Були зроблені висновки на предмет їх діяльності та можливі засоби захисті від них.

У підрозділах 1.4 та 1.5 було проаналізовано методи та засоби які використовують світові мережеві гіганти у боротьбі з ботами. Проведено власний експеримент, кінцевою метою якого було отримання інформації стосовно алгоритмів протидії злочинним машинам, які генерують левову кількість злочинного трафіку. Отримані висновки допоможуть більш точно визначити потребу для подальшої частини роботи.

У підрозділі 1.7 більш детально було розкрите поняття «злочинного трафіку» яке було виділене нами попередньо. Надані конкретні дані що, на нашу думку, розкривають кожний із підпунктів (а-е) з виділенням причин за якими відбувалося сортування.

Окрема увага була приділена факторам, що сприяють зростанню злочинного трафіку (підрозділ 1.8) в мережі інтернет. Були проаналізовані юридичні, практичні та технологічні проблеми що провокують збільшенню показників. Виділено проблематику анонімізації, дослідженню якої було присвячено підрозділ 1.9 з детальною аналітикою сучасних методів та засобів які цю саму анонімність і забезпечують. Були виділені сильні та слабкі сторони кожної технології та надана загальна характеристика для розуміння стану проблеми.

Враховуючи усі вищезазначені підрозділи, зокрема і підрозділ 1.9, у другому розділі даної роботи, ставимо задачу:

— Запропонувати та аргументувати програмно-апаратний засіб або метод, мета якого буде покликана на зменшення рівня анонімізації з кінцевою метою покращення наявного рівня фільтрації небезпечного трафіку.

2 СПЕЦІАЛЬНА ЧАСТИНА

2.1 Аналіз наявної моделі забезпечення анонімізації.

Для отримання уявлення про те, як саме та яким засобом варто зменшити рівень анонімізації, тим самим підвищивши фільтрацію трафіку, необхідно проаналізувати принцип дії наявних методів.

В підрозділі 1.9 було виділено основні чотири основні методи, зокрема:

- Анонімні мережі;
- Анонімні проксі-сервера;
- Анонімні веб-проксі;
- Анонімні VPN-сервіси.

Під час проведення аналізу нашою головною метою буде отримання інформації щодо змін, які завдають задані програмно-апаратні та програмні засоби під час їх використання.

Для усереднення результатів, під час використання ми не будемо застосовувати спеціальні налаштування, а лише опиратися на принцип «меншої кількості дій». Чим менше дій варто виконати для досягнення мети тим краще для нашого аналізу. Головна мета – досягнення певного рівня анонімності шляхом зміни первинних налаштувань.

Базові налаштування.

На першому етапі буде виділено базові налаштування нашого пристрою. В даному випадку це персональний комп'ютер, який складається з:

Системний блок чорного кольору, характеристики компонентів:

1. Процесор – AMD Ryzen 3700X.
2. Материнська плата – ASUSTek PRIME X570-P.
3. Дві плашки оперативної пам'яті по 16 (ГБ) кожна, класу DDR-4, загальним об'ємом 32 (ГБ).
4. Відеокарта Radeon RX 570
5. SSD накопичувач на 256 (ГБ).
6. HDD накопичувач на 1500 (ГБ).

Два екрани чорного кольору:

1. Екран Samsung – діагональ 24”, роздільна здатність 1920x1080.
2. Екран LG - діагональ 24”, роздільна здатність 1920x1080.

В якості програмного забезпечення для доступу до мережі інтернет, буде використовувались інтернет-браузер «Google Chrome», версія - 96.0.4664.93.

Операційна система: Microsoft Windows 10 Pro.

Будь яких додаткових налаштувань чи програмних засобів які б могли вносити зміни на пристрій не встановлено.

Під час тестування, щоразу буде здійснюватися підключення до веб-ресурсу «2IP», який буде надавати типову характеристику стосовно нашого пристрою. Це буде відповідати тим же умовам, в яких знаходиться будь який інший веб-ресурс коли ми його відвідуємо.

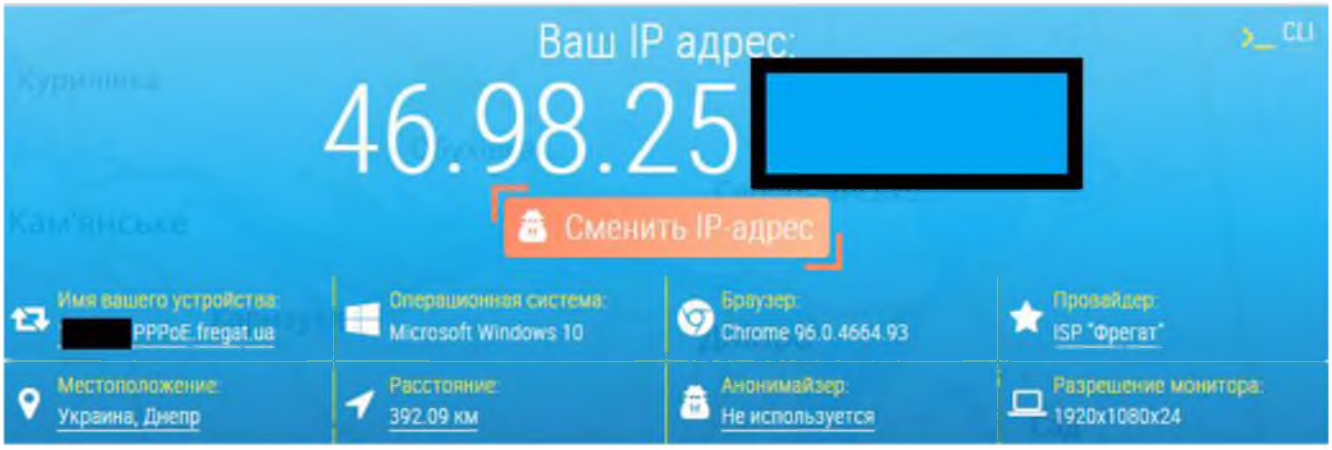

Рисунок 2.1 – базові налаштування пристрою

На рисунку 2.1 зображені базові налаштування пристрою на даний час без внесення змін.

Веб-ресурс чітко визначив IP-адресу та супутні критерії (провайдер, назву пристрою в мережі), операційну систему, бразуер та його версію.

Окрім визначення інформації про пристрій, вищезазначений веб-сервіс також пропонує виконання тесту на приватність та використання засобів анонімності. Його особливість полягає у перевірці інформації про пристрій за переліком певних критеріїв та наданні підсумкової оцінки на предмет того, чи має місце застосування засобів та методів анонімізації.

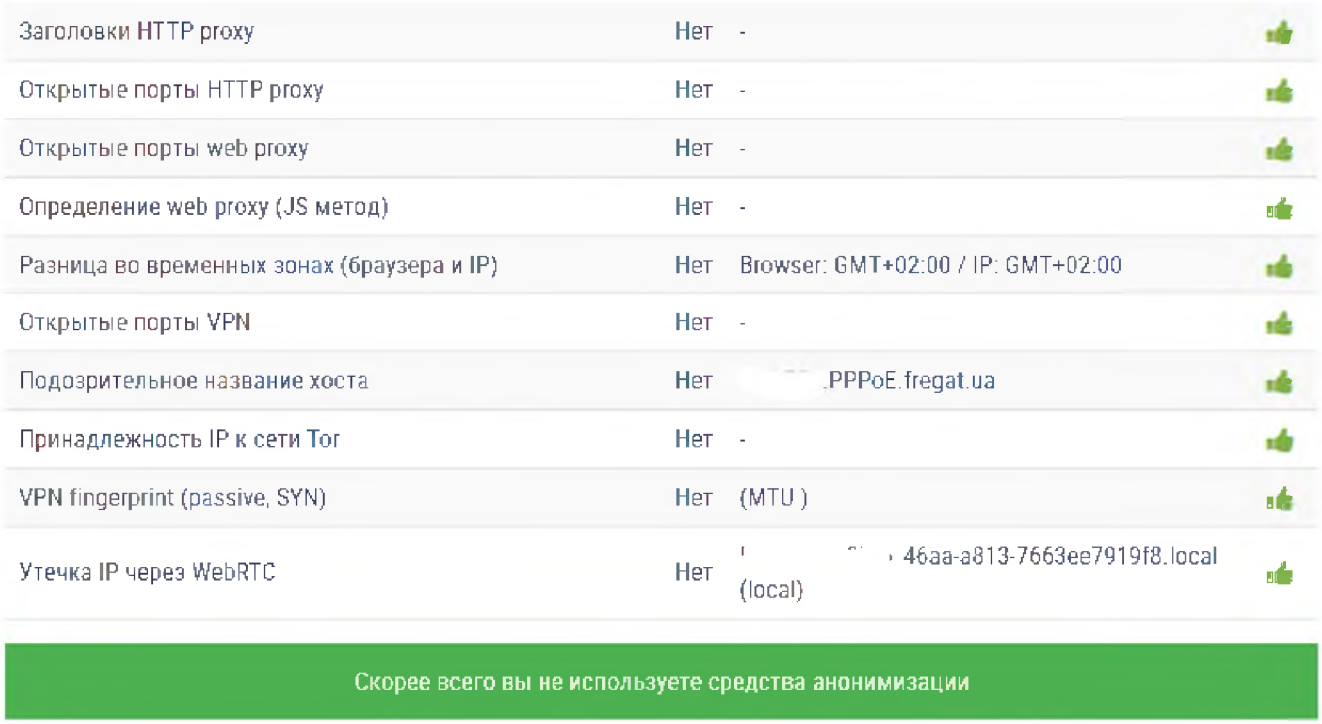

Рисунок 2.2 – результаті перевірки

Результати перевірки, що зображені на рисунку 2.2 вказають на те, що веб-ресурс не зміг виявити у нас будь яких маркерів які б свідчили про те, що ми використовуємо методи анонімізації.

Це важливо з боку того, що певні сервіси, наприклад, не дають змогу використовувати їх через відсутність уявлення про місце походження запиту.

Тестування шляхом використання анонімної мережі.

Серед переліку усіх доступних мереж, для підключення буде використовуватись програмне забезпечення «TOR», яке, як було описано попередньо, шляхом лукової маршрутизації забезпечить підключення до анонімної мережі.

Усі налаштування пристрою залишаються без змін. З офіційного веб-ресурсу завантажується та встановлюється програмне забезпечення інтернет-браузер «TOR». Версія - Tor Browser 10.5.8.

Після встановлення здійснюємо перехід за адресою, попередньо зазначеного, веб ресурсу.

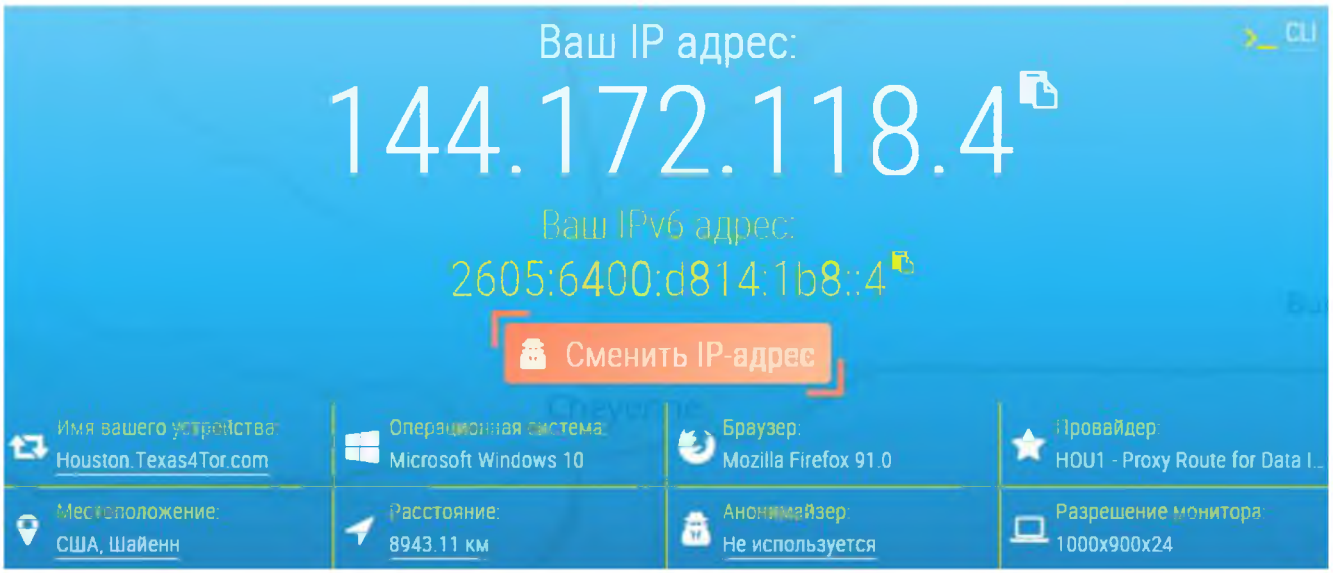

Рисунок 2.3 – отримані налаштування після використання «TOR».

На рисунку 2.3 зазначені налаштування які ми отримали після того, як під'єдналися до анонімної мережі, використовуючи спеціальне програмне забезпечення. В результаті отриманих даних можемо спостерігати, що були змінені всі налаштування окрім одного – операційної системи. Наше місцерозташування було змінено на США, відповідно змінена і IP-адреса. Ще одним важливим фактором є зміна веб-браузеру. Як зазначалося раніше, використовується програмне забезпечення яке має назву «TOR», але сервіс визначив як «Mozilla Firefox», що є назвою іншого браузеру. Можливою причиною цьому можна назвати факт того, що «TOR» побудований на тому ж ядрі що і «Мозілла».

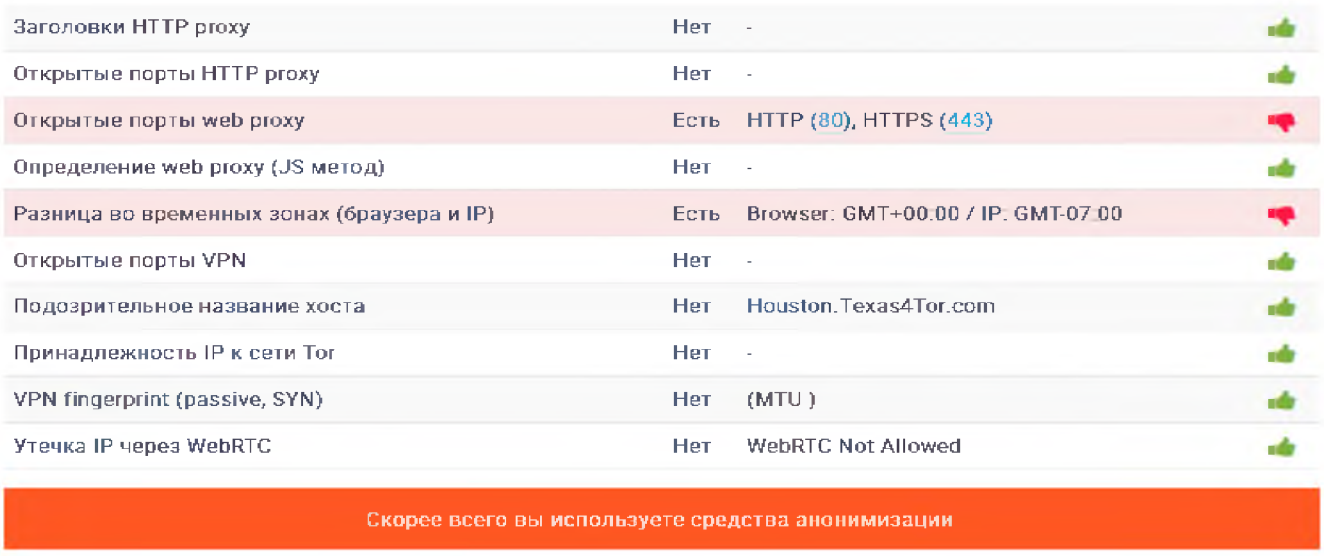

Рисунок 2.4 – результаті перевірки на анонімність

Та результаті перевірки на анонімність, що зазначені на рисунку 2.4, виявили, що «скоріш за все» ми використовуємо засоби анонімізації. Серед маркерів які сприяли цьому визначенню є наявність відкритих веб-проксі портів 80 та 443 які є характерними під час використання спеціальних програмних засобів. Та різниця у часовому поясі, що склала 7 годин, так як фізично сервер до якого ми під'єднані знаходиться в США.

Тестування шляхом використання проксі-серверу.

В переліку стандартних налаштувань операційної системи Windows 10 нами попередньо задано такі налаштування:

Адреса - 195.201.100.236;

Порт – 10018.

Задані параметри застосовуються до усього трафіку що виходить з системи. Це означає що будь який трафік буде проходити, щонайменше, через один проксі-сервер та таким же чином буде повертатися.

Для перевірки застосовується інтернет-браузер «Google Chrome».

Рисунок 2.5 – отримані налаштування після використання «Proxy-серверу».

На рисунку 2.5 вказані отримані нами результати після підключення до «Proxy-серверу». Першим фактом є різниця між заданою IP-адресою, яку ми початково вказували для примусової маршрутизації, та тією що було отримано в результаті перевірки даних. Можливою причиною може бути внутрішні налаштування серверів які під час підключення виконують додаткову маршрутизацію.

Тим не менш, серед ідентифікаторів які змінилися можна виділити: IP-адресу, місцерозташування. Інші характеристики, такі як версія браузеру та показники роздільна здатності монітору.

Результати тесту на перевірку анонімності відповідають даним зазначеними на рисунку 2.5, тобто «використання засобів анонімності не виявлено», що може свідчити про те, що використання зазначеного методу є доволі непоганим рішенням у випадку, коли не стає потреба в визначенні того, чи користується суб'єкт додатковими засобами.

Тестування шляхом використання VPN-сервісу:

Для тестування зазначеного методу буде використовуватись спеціальне програмне забезпечення «IPVanish VPN» від американської компанії «Highwinds Network Group».

Після його встановлення та запуску, за допомогою внутрішнього функціоналу маємо змогу обрати одну країну із переліку до серверу якої буде здійснено підключення. Або обрати функцію «автоматичне визначення». В такому випадку програмне засіб автоматично підключить до серверу. Шляхом автоматичного визначення відбулося під'єднання до серверу, що дислокується у місті Варшава (Польща).

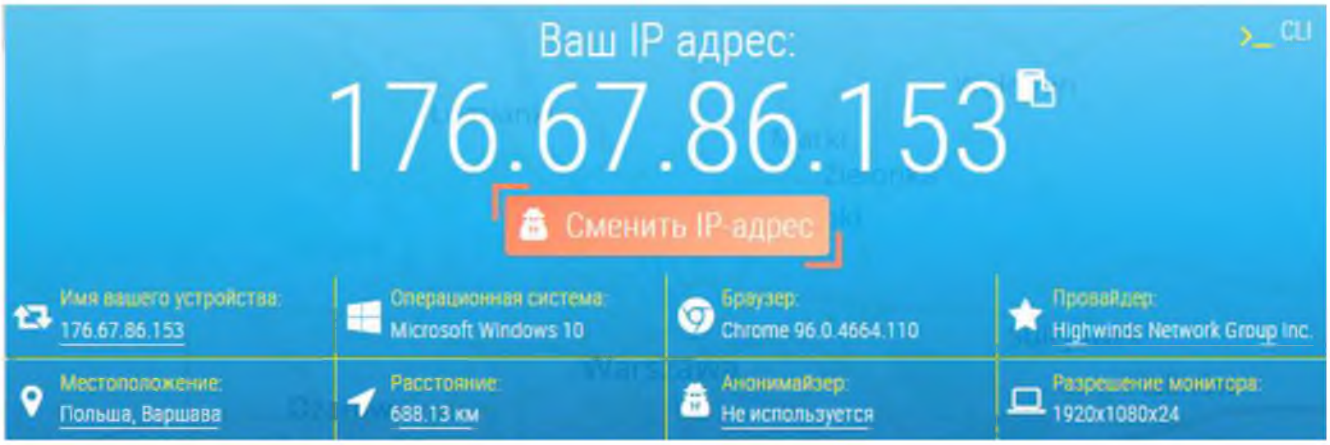

Рисунок 2.6 – отримані налаштування після використання «VPN-серверу».

На рисунку 2.6 вказані отримані результати після підключення до серверу за допомогою технології VPN. Як і у випадку з використанням проксі-серверу, відбулися зміни у IP-адресі та місцерозташуванні. Версія браузера та роздільна здатність монітору залишилася без змін.

Аналогічно, результати перевірки на анонімність відповідають даним з рисунку 2.2 - «використання засобів анонімності не виявлено». Тобто можемо зробити висновки, що суть використання як Proxy-серверу так і VPN-серверу, є схожою.

Під час вивчення проблематики анонінмізації у підрозділі 1.9 було додатково виділено такий метод, як «веб-проксі». Оскільки його реалізація потребує встановлення саме з боку сервера, враховуючи відсутність технічної можливості для повноцінного тестування та аналізу даного методу, та враховуючи той факт, що за своєю технічною суттю вони тотожні до звичайних проксі-серверів, його аналіз не проведений.

Таблиця 2.1 – підсумки тестування методів забезпечення анонімності

	Анонімні мережі (TOR)	Proxy-сервер	VPN-мережі
Зміна IP-адреси	+	+	+
Зміна місцерозташування	+	+	+
Зміна версії браузеру	+	-	-
Зміна визначеної роздільної здатності монітору	+	-	-
«Тест на анонімність»	-	+	+

В таблиці 2.1 зазначено підсумкові результаті проведеного тестування визначених методів забезпечення анонімізації. Виходячи з цього можемо зробити такі висновки: усі методи забезпечують зміну IP-адреси і місцерозташування. Заміна версії браузер відбувається лише у випадку застосування підключення до анонімних мереж, але саме підключення відбувається за допомогою програмного забезпечення типу браузер. Також цей метод забезпечив зміну даних про роздільну здатність монітору, на відміну від використання VPN та Proxy.

2.2 Розгляд методів збору інформації про користувачів

Під час виконання першого розділу нами було проведене певне тестування функціоналу, яке забезпечує штучний інтелект на прикладі соціальної мережі «Facebook». Розробники використовують автоматизований функціонал на основі штучного машинного інтелекту під назвою Deep Entity Classification. В результаті проведення, було отримано висновки (таблиця 1.2), що більшою мірою увага приділяється IP-адресі, далі йдуть ідентифікатори SIM-карток і останнім, серед критеріїв, є сам пристрій з якого виконувалась реєстрація.

Підсумки, отримані в результаті аналізу моделей забезпечення анонімності також, більшою мірою, висвітлюють забезпечення перемінності саме IP-адреси. Важко сказати, які саме чинники є ключовими у даному вибору, але це дає змогу до припущення, що другорядні фактори, а саме такі, як мінімум, як версія браузера, роздільна здатність екрану, тощо, не розглядаються так детально.

Тим не менше, більшість веб-сайтів та програмних засобів збирають певні данні про своїх користувачів. Як правило, це відбувається з метою покращення персоналізації реклами або рекомендацій.

Із основних видів збору даних варто виділити два:

- Збір за допомогою програмного забезпечення.
- Збір за допомогою веб-сайтів.

В якості першого прикладу варто виділити бізнес-екосистему, створену американською компанією «Google» [44]. На даний час два основні ринки смартфонів поділяються на девайси від компанії «Apple» на основі операційної системи «iOS» та ті, що виходять під брендами різних виробників на основі операційної системи «Android». Одним із ключових власників останньої є вищезгадана компанія «Google». Окрім самої операційної системи, на пристроях за замовчуванням буде встановлено такі програмні продукти, як: магазин застосунків «Google Play Market», браузер «Google Chrome», онлайн-мапу «Google Maps», тощо. Використання телефоном не є можливим без створення аккаунту на платформі американського розробника.

Використовуючи, як сам пристрій так і його програмне забезпечення, користувач має погодитися з політикою конфіденційності, пункти якої, зокрема, включають наступне: «Google збирає дані користувача для того, щоб наші сервіси були більш зручними. Якщо Ви не ввійшли в обліковий запис «Google», ми реєструємо інформацію, яку збираємо за допомогою унікальних ідентифікаторів, пов'язаних з браузерами, програмами та пристроями.

Також збираються такі дані [45]:

— Пошукові запити, які виконує користувач;

— Відео та контент які переглядаються;

— Аудіо інформація, що передається під час голосового керування;

— Покупки через внутрішні сервіси;

— Користувачі з якими користувач спілкується;

— Дії користувача на веб-ресурсах, які використовують сервіси «Google»;

— Історія перегляду веб-сторінок;

— Дані про місцерозташування;

— IP-адреса;

— Данні з датчиків на пристрої;

— Інформація про об’єкти навколо користувача

Усі ці дані користувач має змогу переглядати, частину з них видаляти.

В якості другого випадку деякі дані збирають вже самі веб-сайти. Як правило, це відбувається за допомогою так званих «HTTP-cookie». Це невеликий набір даних, що створюється веб-сервером під час того, як користувач переглядає певний веб-сайт. «Cookie» розміщуються на самому пристрою за допомогою якого виконується доступ до мережі інтернет.

Існують такі типи «кук-файлів» [46]:

— Сеансові — зберігаються лише під час однієї сесії браузера;

— Постійні – зберігаються на пристрої без термінів видалення.

— Первинні — встановлюються саме тим веб-ресурсом який відвідує користувач під час користування мережі Інтернет.

— Сторонні – встановлюються стороннім веб-ресурсом під час користування певним веб-сайтом. Приклад: веб-ресурс google.com, а самі «cookie» встановлені від іншого ресурсу «notgoogle.com». Як правило такі файли використовують рекламодавці.

— Супер-cookie – використовуються для ідентифікації як веб-ресурси доменів верхнього рівня.

— «Evercookie» - «cookie» які є можливим відновити навіть у випадку видалення шляхом задавання певних команд використовуючи у браузері мову програмування «JavaScript».

Головна мета – отримання веб-серверами інформації для швидкої ідентифікації користувача або його налаштувань. Для прикладу, якщо є певний веб-ресурс з можливістю тонкого налаштування, і персоналізації, то після його повторного закриття та відкриття за допомогою цього файлу усі попередньо встановлені налаштування буде відтворено.

Деякі веб-ресурси використовують «Cookie» як метод для аутентифікації користувача. Під час першого вдалого входу на ресурс використовуючи логін/пароль, генерується унікальний ідентифікатор який записується в «кукі-файл». З однієї сторони це є пришвидшенням та додатковою оптимізацією дій, з іншої – у випадку перехоплення такого файлу третя особа зможе використовувати його в власних цілях. А вбачаючи рівень довіри (до «кук-файлів» він вище аніж до звичайних данних) така інформація є критично важливою.

Як правило в «Cookie» файлах зберігається наступна інформація:

- Данні для входу на веб-ресурс;
- Персональні налаштування сайту (персоналізація);
- Версію операційної системи та веб-браузера;
- Кліки та переходи користувача між веб-сторінками на одному ресурсі;
- IP-адресу, місцерозташування, дата і час входу на веб-ресурс.

В 2018 році Європейський Союз посилив вимоги у своїй директиві «О конфіденційності і електронних засобах зв'язку» зобов'язавши веб-ресурси та розробників програмного забезпечення, що надає доступ до мережі Інтернет повідомляти користувачів про «кук-файли» та іншу інформацію яка буде збиратися в процесі використання. У разі відмови збір такої інформації не є допустимим [47].

Це означає, що особа може не погодиться зі збором, таким чином кожного разу при відвідуванні збір даних буде повторно запитуватись і можливе збирання буде відбуватися лише з боку отримувача доступу. У випадку, коли дані збираються з боку розробника програмного забезпечення та операційної системи, відмова від збору даних призведе до неможливості подальшого використання, наприклад, пристрою до моменту поки рішення не буде змінене. Тобто розробник в будь якому випадку буде отримувати та збирати дані.

2.3 Використання методу збору цифрових відбитків пристроїв.

Розгляд наявних методів збору інформації, що був виконаний у попередньому розділі, на наш погляд, висвітлює відразу декілька проблем, які відіграють свою роль під час фільтрації трафіку, зокрема:

— Дані створюються лише з метою персоналізації, без будь-яких дій на предмет порівняння;

— Часто, інформація зберігається саме з боку користувача, який за потреби може її спотворити або видалити (зміна кук-файлів);

— Можлива відмова від збору даних;

Примітка: важливо уточнити, що будь-яка особа має право на відмову збору персональних даних стосовно себе та свого приватного життя (окрім випадків передбачених законодавством), але має залишатися можливість ідентифікації за публічними даними. Мова йде не про збереження історії браузеру, умовно (хоча, як ми бачимо з деяких прикладів це відбувається), а про збереження даних про пристрій та його специфікацію.

— Відсутність централізованих налаштувань та вимог до них;

— Використання спеціальних режимів «інкогніто», під час використання браузеру.

Усі вищезазначені пункти вказують на те, що для покращення методів фільтрації варто розглянути додаткові методи збору та подальшого аналізу даних.

В якості прикладу, нами запропоновано використання технології, яка забезпечить централізоване збирання інформації стосовно самого пристрою (не користувача).

Цифровий відбиток пристрою (англ. Fingerprint) – технологія, що дозволяє, збирати інформацію про програмне та апаратне забезпечення віддаленого обчислювального пристрою з метою його подальшої ідентифікації [48].

Головною особливістю є те, що цифрові відбитки пристрою можна використовувати для повної або часткової ідентифікації окремих пристроїв навіть у випадках, якщо;

— Постійні файли «cookie» не можуть бути прочитані (або вони відключені), чи якщо їх зберігання відбувається на пристрої саме отримувача послуг.

— IP-адреса прихована, або змінена (змінюється);

— Використання різних програмно-апаратного забезпечення – браузерів.

Тобто, їх варто використовувати в якості додаткових методів під час збору інформації про користувача, так як ті проблеми які не вирішує збір даних в випадку застосування, наприклад «cookie», виправляє метод цифрового відбитку.

2.4 Тестування методу збору цифрових відбитків.

На думку деяких експертів, в результаті зняття «фінгепринту» пристрою, можна визначити його унікальність щонайменше на 94%[49].

Для отримання точної інформації, та обґрунтування подальшого використання, ми, використовуючи незалежні ресурси проведемо тестування нашого програмно-апартного пристрою з метою порівняння його з, вже готовими, базами інших для виявлення власного коефіцієнту унікальності.

В процесі тестування буде використано три веб-ресурси:

1) http://fp.virpo.sk/
2) https://amiunique.org/
3) https://coveryourtracks.eff.org/

Усі вищеозначені ресурси розроблені різними експертами, або міжнародними організаціями.

В результаті використання першого ресурсу отримаємо свій набір даних, що наведений у таблиці 4

Таблиця 2.2 – цифровий відбиток пристрою, що тестується

Назва характеристики	Значення
timezone	-120
screenSize	1920,1080
availSize	1920,1040
colorDepth	24
pixelRatio	1
userAgent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/96.0.4664.110 Safari/537.36
cookiesEnabled	true
mathtan	-1.4214488238747245
dateFormat	01.01.1970, 03:00:00
touchCompatibility	0,false,false
languages	ru-RU,ru-RU,,,
localStorage	true
sessionStorage	true
userData	false
indexedDB	true
doNotTrack	undefined
hardwareConcurrency	16
cpuClass	undefined
platform	Win32

Продовження таблиці 2.2

Назва характеристики	Значення
plugins	>Plugin 0: PDF Viewer, internal-pdf-viewer, Portable Document Format, application/pdf, pdf;;>Plugin 1: Chrome PDF Viewer, internal-pdf-viewer, Portable Document Format, application/pdf, pdf;;>Plugin 2: Chromium PDF Viewer, internal-pdf-viewer, Portable Document Format, application/pdf, pdf;;>Plugin 3: Microsoft Edge PDF Viewer, internal-pdf-viewer, Portable Document Format, application/pdf, pdf;;>Plugin 4: WebKit built-in PDF, internal-pdf-viewer, Portable Document Format, application/pdf, pdf
iePlugins	empty
webGLVendor	Google Inc. (AMD)
webGLRenderer	ANGLE (AMD, Radeon RX 570 Series Direct3D11 vs_5_0 ps_5_0, D3D11-27.20.1034.6)
adBlock	false
installedFontsJs	Arial;Arial Black;Arial Narrow;Arial Unicode MS;Book Antiqua;Bookman Old Style;Calibri;Cambria;Cambria Math;Century;Century Gothic;Century Schoolbook;Comic Sans MS;Consolas;Courier;Courier New;Garamond;Georgia;Helvetica;Impact;Lucida Bright;Lucida Calligraphy;Lucida Console;Lucida Fax;Lucida Handwriting;Lucida Sans;Lucida Sans Typewriter;Lucida Sans Unicode;Microsoft Sans Serif;Monotype Corsiva;MS Gothic;MS PGothic;MS Reference Sans Serif;MS Sans Serif;MS Serif;Palatino Linotype;Segoe Print;Segoe Script;Segoe UI;Segoe UI Light;Segoe UI Semibold;Segoe UI Symbol;Tahoma;Times;Times New Roman;Trebuchet MS;Verdana;Wingdings;Wingdings 2;Wingdings 3;
canvasFp	15f07fd550c5a1298275c02e6ef3e8e4
audio	124.90862639580155

Примітка: В даній таблиці не наведені мережеві налаштування, зокрема IP та MAC адреси.

Під час тестування другого веб-ресурсу отримуємо схожі результати, але тут вже відбувається порівняльна характеристика з загальною базою, що, згідно з інформації з ресурсу, за останні сім днів складає близько тринадцяти тисяч пристроїв.

Підсумковий результат показав, що наш пристрій є унікальним з, близько, тринадцяти тисяч інших пристроїв. Серед унікальних даних (менше 5%), які були порівняні з загальною базою є наступне:

Таблиця 2.3 – порівняльна характеристика унікальності відбитків

Назва характеристики:	Значення	(%) Від загальної бази
User agent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/96.0.4664.110 Safari/537.36	4,86
Content language	ru-RU,ru;q=0.9	0,07%
List of fonts (JS)	Agency FB, Algerian, Arial, Arial Black, Arial Narrow and 167 others	0,07%
Navigator properties	64	4,86%
Screen available Left	-1920	0,22%
WebGL Vendor	Google Inc. (AMD)	1,09%
WebGL Renderer	ANGLE (AMD, Radeon RX 570 Series Direct3D11 vs_5_0 ps_5_0, D3D11-27.20.1034.6)	0,07%
Media devices	-	0%
Connection	downlink : 10, effectiveType : 4g, rtt : 50	4,13%

Третій веб-ресурс має базу у, приблизно, двісті двадцять тисяч пристроїв, що були протестовані за останні сорок п’ять днів. В результаті наш пристрій є унікальним, порівнюючи з іншими. Оскільки даний ресурс не дає більш детальної характеристики окрім набору даних що вказані в таблиці 4, то зробити будь які висновки стосовно точності визначення не є можливим.

В разі прорахунку, за умови наявності двохсот мільйонів пристроїв, заданий нами пристрій буде ідентифікований з точністю більше 99%, що підтверджує наявні дані експертів та вказує на значну ефективність використання обраного методу.

2.5 Дослідження варіантів збору інформації методом цифрового відбитку

На сьогоднішній день, згідно з останніми дослідженнями, найпопулярнішими програмно-апаратними пристроями для доступу в мережу інтернет в Україні є: мобільні смартфони, ноутбуки та персональні комп'ютери. Тому під час дослідження можливих варіантів збору інформації про пристрої, буде виділено такі групи:

1) Збір інформацій на ноутбуках та персональних комп'ютерах.
2) Збір інформації на мобільних пристроях:
    a) Збір на мобільних пристроях, операційна система яких є IOS;
    b) Збір на мобільних пристроях, операційна система яких є Android.

2.5.1 Збір інформацій на ноутбуках та персональних комп'ютерах.

Програмно-апаратні машини, що працюють на основі операційної системи «Windows» від американської компанії «Microsoft» мають певну кількість персоніфікованих налаштувань. За допомогою браузеру та внутрішніх плагінів ця інформація може буде зібрана та проаналізована.

Сам збір може відбуватися двома видами:

- Прихований – «невидимий» запит на клієнтську машину;
- Активний – клієнт сам надає дозвіл на зчитування даних;

Перший варіант, найчастіше, реалізований на певних рівнях моделі OSI.

Модель OSI - це концептуальна модель, яка характеризує та стандартизує комунікаційні функції телекомунікаційної або обчислювальної системи без урахування внутрішніх структур та технологій. Мета - сумісність різноманітних комунікаційних систем зі стандартними протоколами зв'язку. На сьогоднішній день має сім рівнів протоколів [50]:

1. Фізичний рівень.
2. Канальний рівень.
3. Мережевий рівень.
4. Транспортний рівень.

5. Сеансовий рівень.
6. Рівень представлення.
7. Прикладний рівень.

Серед доступних для зчитування протоколів варто виділити наступні:

На другому рівні доступне зчитування через протокол CDP - це закритий протокол другого рівня, розроблений компанією «Cisco Systems», що дозволяє виявляти підключене (безпосередньо або через пристрої першого рівня) мережеве обладнання Cisco, його назву, версію IOS і IP-адреси

На третьому рівні за допомогою мережевих протоколів IPv4, IPv6, ICMP, IEEE 802.11. Оскільки сам рівень є мережевим, то його побудова має включати застосування, щонайменше, один із зазначених протоколів. В нашому випадку, наприклад, використовується IPv4, IPv6. IEEE 802.11 ж включає набір стандартів для комунікації в бездротовій мережі WLAN різних діапазонів. Ці стандарти, зокрема, визначають основи функціонування апаратних пристроїв мережевого стандарту «Wi-Fi». Тобто інформацію можливо буде отримати у випадку підключення пристрою до бездротової мережі.

На четвертому рівні зчитування відбувається через мережевий протокол TCP. В цьому випадку використовується властивість, що кожна операційна система має власну реалізацію стеку протоколів TCP/IP, налаштування якого можна знайти в заголовку мережевих пакетів і які відрізняються від параметрів інших операційних систем. Серед цих параметрів вісім критеріїв загальним розміром у 67 (Біт), унікальність та яких дозволяє створювати сигнатуру відповідних розмірів. Як мінімум, це надає змогу до швидкої ідентифікації операційної системи пристрою, що на майбутніх рівнях забезпечить діяльність інших протоколів у відповідності до визначених показників.

На п’ятому рівні реалізовано на основі мережевих протоколів SNMP та NetBIOS у випадку їх використання апаратним пристроїм.

На сьомому рівні застосовуються мережеві протоколи прикладного рівня, зокрема: SMB, FTP, HTTP, Telnet, TLS, DHCP.

На сьогоднішній день описано, щонайменше, три варіанти отримання пасивного відбитку пристрою через використання протоколів четвертого та сьомого рівня.

Отримання відбитку через протокол HTTP відбувається виконанням запиту на отримання таких даних:

1. SETTINGS frame (фрейм налаштувань)

Таблиця 2.4 – налаштування першого фрейму згідно нормативного документу RFC 7540

Ім'я параметру	Область застосування
SETTINGS_HEADER_TABLE_SIZE	Дозволяє відправнику інформувати віддалену кінцеву точку про максимальний розмір використовуваної таблиці стиснення заголовка.
SETTINGS_ENABLE_PUSH	Для вимкнення пуш-повідомлень серверу
SETTINGS_MAX_ _CONCURRENT_STREAMS	Вказує максимальну кількість одночасних потоків
SETTINGS_INITIAL_WINDOW_SIZE	Вказує початковий розмір вікна відправника (в байтах) для управління потоком
SETTINGS_MAX_FRAME_SIZE	Вказує на розмір найбільшого корисного навантаження кадру який відправник готовий отримати, в байтах
SETTINGS_MAX_HEADER_LIST_SIZE	Довідковий параметр інформує партнера про максимальний розмір список заголовків, який відправник готовий прийняти, у байтах.

2. WINDOW_UPDATE frame (фрейм оновлення)

Використовується для реалізації керування потоком. Коли вперше встановлюється з'єднання, створюються нові потоки розміром 65535 Байт, які шляхом запуску команди SETTINGS_INITIAL_WINDOW_SIZE можуть змінюватись, оскільки налаштування відрізняються, і клієнтські машини надсилають дані повторно, але вже після внесення власних налаштувань.

3. PRIORITY frame (фрейм першочерговості)

Надсилається для того, аби встановити пріоритет потоку.

Певні клієнти після підключення до серверу надсилають свій заданий перелік пріоритетних кадрів, які ще не були запущені. Набір цих кадрів дає можливість для додаткової ідентифікації пристрою.

Альтернативним методом отримання відбитку пристрою за допомогою браузеру є використання скриптів, що вбудовані з боку надавача послуг і реалізовані, як приклад, за допомоги мови програмування JavaScript.

В якості прикладу буде розглянуте рішення «Fingerprintjs» [51]. Воно має відкритий програмний код. У випадку впровадження методу збору відбитків пристроїв, на основі даного скрипту можлива побудова власного, включаючи усі необхідні потреби, або використання вже готового рішення.

Метод дії полягає у встановленні до агенту JavaScript коду, в результаті чого агент буде надсилати запити до API програмного продукту з метою ідентифікації браузеру.

Кожний такий запит генерує унікальну хеш-суму, яка не буде повторюватись у 99,5% випадків. Цей показник і є унікальним для кожного користувача.

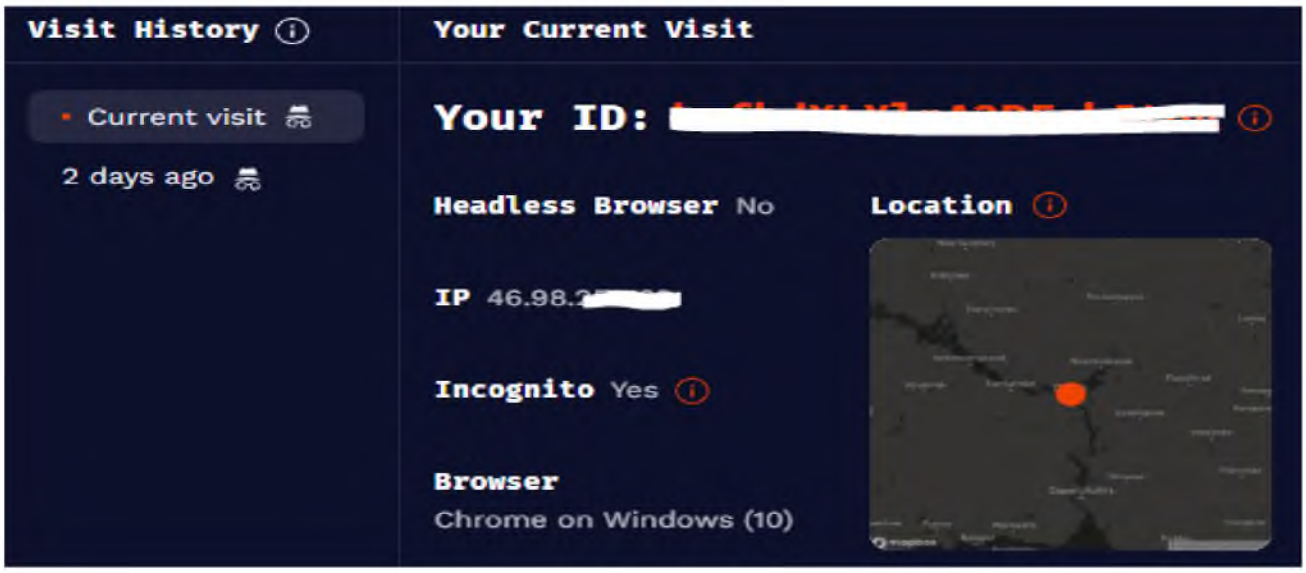

Рисунок 2.7 – персональний хеш.

На рисунку 2.7 наведено приклад такої ідентифікації використовуючи наш пристрій. Так, при першому відвідуванні сайту, нам був наданий унікальний ідентифікатор. Який, зокрема, визначив використання браузерного режиму

«Інкогніто», що передбачає автоматичне видалення «cookie» після завершення сеансу. Відвідавши цей самий ресурс повторно, знову використовуючи вищесказаний браузерний режим, скрипт знову зміг нас ідентифікувати і надати історію попередніх відвідувань з вказаними на попередній і теперішній час даними.

Принцип дії полягає в використанні, щонайменше, п'яти різних критеріїв, підсумковий результат яких генерується у хеш-суму. До таких критеріїв належать:

User Agent – програмне забезпечення яке діє від імені користувача. Являє собою ідентифіковану строку клієнтського додатку, що використовує певний мережевий протокол. При відвідувані веб-ресурсу з боку користувача надсилається певна інформація, як правило, назва і версія програмного продукту, дані про операційну систему, апаратний пристрій, мову, тощо.

Canvas – елемент HTML5, призначений для створення растрових двомірних зображень за допомогою скриптів (як правило – JavaScript).

В 2014 році був виділений додатковий метод зняття цифрового відбитку пристрою на основі Canvas, який надав змогу використовувати ідентифікацію користувача за допомогою елементу HTML5, замість файлів «cookie».

Принцип дії: коли користувач відвідує веб-сторінку, скрипт спочатку створює текст з обраним шрифтом і поступово змінює його, додаючи розмір, фонові кольори. Після цього методом «ToDataURL» отримує пікселі «полотна» в форматі «dataURL», що являє собою представлення у двійковому форматі Base64. Отриманий хеш закодований пікселів і є тим самим відбитком пристрою.

— Встановлена часова зона;

— Встановлені плагіни та данні про них;

— Встановлені мова та мови;

Кожний із запропонованих методів, а саме: реалізований на певних рівнях моделі OSI та використання створених на JavaScript скриптів, має як свої переваги так і недоліки.

Переваги методу використання рівнів моделей OSI:

— Робота на нижчому рівні операційної системи, відсутність можливості до блокування;

— Важкість внесення змін з боку користувача;

— Загальнодоступність використання;

— Відсутня можливість до блокування з боку браузерів

Недоліки методу використання рівнів моделей OSI:

— Нижчій, порівняно з другим методом, рівень ідентифікації

Примітка: За деякими дослідженнями максимальна точність – 90%, в випадку з використанням скриптів – цей показник може зростати до 96%.

— Доречне використання разом із іншими методами

— Потреба у створенні персональних ідентифікаторів користувачів (створення хешів, за замовчуванням, відсутнє);

Переваги методу створених на JavaScript скриптів:

— Точність даних;

— Простота та універсальність у встановлені;

— Можливість порівняння отриманих даних;

Недоліки методу створених на JavaScript скриптів:

— Наявна певна можливість до блокування з боку браузерів та закритих операційних систем;

— З п’яти наявних критеріїв для оцінки, щонайменше, три можуть бути спотворені користувачем.

— Потреба у наявності власних засобів для зберігання у випадку самостійної розробки і впровадження;

— Необхідність у стабільних оновленнях, в випадку внесення змін з боку розробників програмних продуктів;

Ідеальними умовами в використанні, на нашу думку, є застосування обох методів для отримання більш комплексної оцінки для порівняння. Але, якщо ж питання постає саме в обрані одного методу, то перевага має бути надана саме використанню скриптів. Причиною цьому є швидкість та простота у

налаштуванні, вищий рівень точності, можливість власних налаштувань та програмування.

2.5.2 Збір інформації на мобільних пристроях

Збір інформації на мобільних пристроях варто поділити за такими напрямами:

— Збір за допомогою використання програмного забезпечення та/або програмних продуктів (додатків);

— Збір за допомогою програмних продуктів типу браузер;

— Збір за допомогою мережевих налаштувань.

Збір за допомогою використання програмного забезпечення та/або програмних продуктів (додатків) першим чином залежить від наявності дозволів з боку користувача та операційної системи на зберігання та отримання даних, які можуть бути використані з метою ідентифікації. В даному випадку мова не йде про персональні дані, такі як контакти, історія листувань, тощо. Лише про технічну інформацію, що можливо отримати під час використання додатку.

Як було зазначено раніше, на сьогоднішній день найпопулярнішими є смартфони на двох типах операційних систем:

— IOs – закрита операційна система, що розроблюється та підтримується американською корпорацією «Apple».

— Android – відкрита операційна система, що розроблюється та підтримується американською корпорацією «Google».

Головна різниця полягає в політиках компаній по відношенню до своїх систем.

В випадку використання ОС від компанії «Apple», єдиним шляхом отримання додатків є їх пряме завантаження з застосунку «App Store». Для того, аби туди потрапити, розробники програмних продуктів проходять чітку перевірку, яка, зокрема, включає в себе наступне: додатки мають використовувати лише власний програний код, виконання стороннього коду не є допустимим, окрім випадків коли додаток створений для навчання. Це означає, що у випадку

влаштування у сам додаток скрипту для отримання цифрового відбитку, такий додаток може бути вилучений з «маркету».

Окрім цього, згідно з останніми оновленнями політики конфіденційності був введений спеціальний фреймворк (App Tracking Transparency), метою якого є донесення до користувача інформації про те, як використовуються дані користувача, а також які програма має запити. Перелік, для прикладу, зображений на рисунку 2.8.

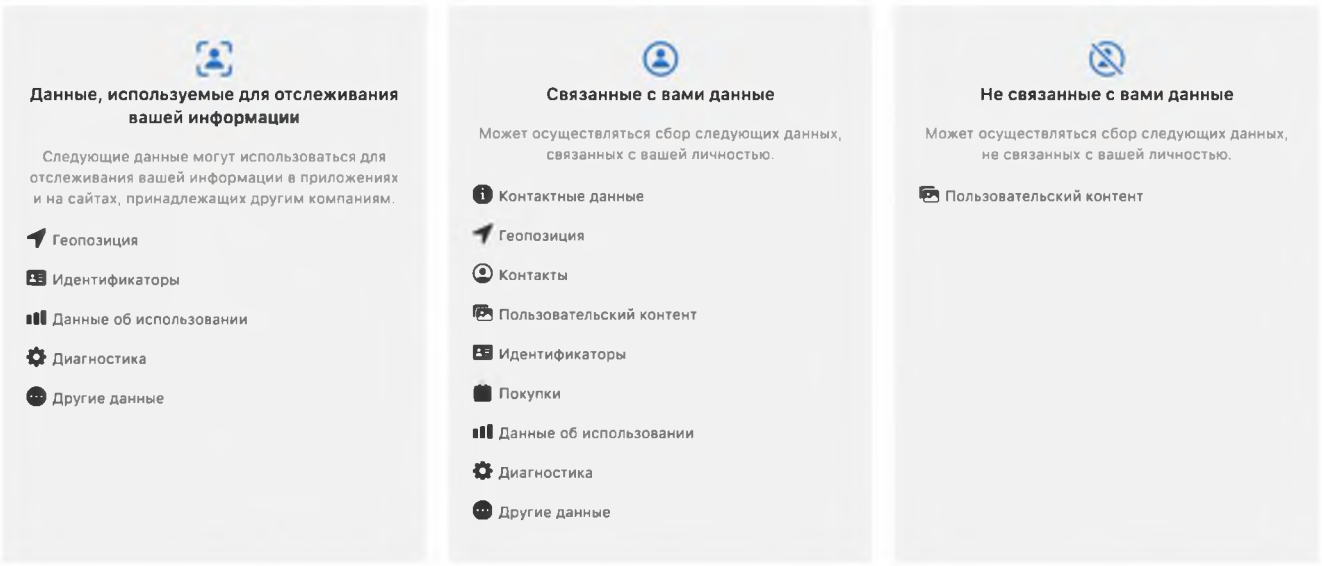

Рисунок 2.8 – перелік дозволів.

Як і будь який програмний продукт, за умови звичайного користування, є можливість отримати інформацію про, як мінімум, IP-адресу користувача, IMEI, та інформацію про сам пристрій. Але користувач може обмежити або заборонити збір таких даних, що призведе до неможливості зняття відбитку пристрою, або отриманої інформації буде замало для унікальної ідентифікації особи.

Примітка: IMEI - серійний номер мобільного пристрою який встановлюється заводом-виробником та є унікальним для кожного мобільного телефону. Він служить для точної і повної ідентифікації телефону в мережах форматів GSM та UMTS: автоматично передається апаратом у мережу оператора при підключенні. Двох телефонів з однаковим IMEI не існує.

В свою чергу, використання смартфону, який працює на основі операційної системи «Android», встановлення додатків можливе також з використанням сторонніх сервісів.

Як і у випадку з аналогічною ОС, для стабільного функціонування необхідна згода користувача на передачу програмному продукту доступу до своїх даних, але використання певних системних налаштувань не потребує згоду. Так, наприклад, за використанням «android.telephony.TelephonyManager» будь-який розробник може отримати дані про IMEI телефону. Сама операційна система не приховує цю інформацію і дає доступ до неї шляхом виклику через стандартні запити.

Окрім цього, є цілий перелік інших команд, які на програмно-консольному рівні можуть отримати данні про пристрій з метою його подальшої ідентифікації. Використання самого IMEI може бути також корисним для прив'язки до конкретного користувача замість хеш-суми.

В цілому, порівнюючи обидві операційні системи: і в першому і в другому випадку можливо отримати данні про технічну складову пристрою як із дозволом користувача, так і без. Нюансом є той факт, що ця інформація не завжди є вичерпною та достатньою для створення якісного відбитку.

Збір за допомогою програмних продуктів типу браузер або за допомогою мережевих налаштувань передбачає схожу механіку дії, як і у випадку з використанням ноутбуків та персональних комп'ютерів. Ініціалізація відбувається з боку надавача послуг, який однаковим методом зчитує дані для створення цифрового відбитку з будь-яких пристроїв.

Наприклад, деякі із брузерних скриптів для ідентифікації смартфонів на операційній системі «Android» використовують маркери: DeviceID, GSF ID, Android ID. Усі вони є унікальними, але при певних подіях можуть бути спотворені користувачем, у випадку, якщо це робиться навмисно.

DeviceID - набір цифр і букв, який ідентифікує кожен окремий смартфон або планшет. Зберігається на мобільному пристрої та може бути отриманий будь-якою програмою, яка завантажена та встановлена на девайс.

GSF ID (Google Services Framework Identifier) - унікальне число з 16 символів. Автоматично генерується під час першої авторизації в обліковому записі «Google». Є постійним. Після генерації можливе лише його видалення.

Android ID - 64-бітне число, що генерується випадковим чином під час першого запуску пристрою. Залишається незмінним протягом усього часу використання девайсу.

До додаткових критеріїв, які можливо використовувати для ідентифікації також належать: Android IMEI, ID GSM Модулю, Wi-Fi Mac-адреса, Android BlueTooth ID.

Головною особливістю цих параметрів є їх стійкість навіть у випадку скидання налаштувань операційної системи.

Підсумовуючи вищевикладене, в залежності від типу апаратного пристрою, особливостей його операційної системи, та обраного методу збору цифрових відбитків, надійність та точність залежить від комплексності застосування.

Є необхідність у створенні певного реєстру, в якому мали би зберігатися усі отримані дані з метою їх подальшої передачі або порівняння.

2.6 Створення реєстру цифрових відбитків пристроїв

Як було зазначено в підсумках попереднього розділу, головною метою збору цифрових відбитків є можливість їх подальшого порівняння та зберігання. В обраному розділі нами будуть описані вимоги та сценарії можливої реалізації реєстру «фінгерпринтів».

Ми будемо розглядати два напрями:

— Централізована реалізація – створення єдиного всеукраїнського (або масштабнішого реєстру), куди усі надавачі послуг будуть надавати отримані відбитки пристроїв.

— Місцева реалізація – кожний веб-ресурс або додаток має взяти на себе обов'язки по зберіганню, надсиланню та обробці отриманих відбитків.

Надалі будуть надані загальні правила та вимоги до таких реєстрів. Звичайно, у випадку впровадження цей перелік може бути змінений и розширений, але на даний час будуть виділені основні аспекти які варто попередньо враховувати.

За умови, що кожний веб-ресурс буде використовувати власний алгоритм та сценарій отримання відбитків, в результаті є ризик отримати безліч ідентифікаторів, що будуть належати одному пристрою. Будь-яка подальша дія буде неможлива до тих пір, поки дані не будуть централізовані, за умови, що це є можливим.

Загальні правила та вимоги:

1. Цифрові відбитки пристроїв (далі – відбитки) є такими же даними, що і інші ідентифіковані маркери (номера телефонів, IP-адреси, тощо).

2. Відбитки відносяться до інформації із обмеженим доступом.

3. Кожному пристрою має надаватися один унікальний ідентифікатор, в незалежності від обраного методу збору відбитків. Цей ідентифікатор не має повторюватись та співпадати з іншими. У випадку часткової зміни відбитку на апаратному рівні має змінюватись і ідентифікатор. Він має використовуватись як основний маркер для будь-яких дій з відбитками.
В якості основи на мобільних пристроях пропонується застосовувати IMEI, за необхідністю може бути застосована частина іншого незмінюваного ідентифікатора. Приклад: IMEI та певна числова частина DeviceID. Для немобільних пристроїв, варіантом є створення хеш-суми, алгоритм генерації якої буде загальним. Пропонований алгоритм хешування - MD5. Усі отримані дані, що стосуються саме апаратної складової пристрою мають бути переведені у текстовий формат, зміст якого і буде захешований.

4. Зміна мережевих адрес не має впливати на ідентифікатор.

5. На одному накопичувачі має зберігатися не більше 500000 унікальних даних.

6. Надавач послуг має забезпечувати цілодобову безперебійну діяльність реєстру. Окрім форс-мажорних випадків.

7. Дозволяється групування інформації, якщо достовірно відомо, що вона відноситься до одного і того ж пристрою (чи користувача). В такому випадку, при передачі, такі дані мають бути зазначені як додаткові і не впливати на сам відбиток.

8. Обмін даними між надавачами послуг має відбуватися лише у форматі ідентифікаторів без повної розшифровки даних з зазначенням першочергового джерела отримання відбитку.

9. У випадку отримання самостійної інформації, або доповнення для певного ідентифікатора має відбуватися його виправлення

10. Формат зберігання має передбачати доступну візуалізацію та швидку роботу під час пошуку. Рекомендується використовувати табличний загальнодоступний формат.

11. Доступ до повних даних повинні мати лише окремі співробітники, які несуть відповідальність за можливе розголошення цієї інформації.

12. Розголошення має відбуватися лише на законних підставах, зокрема: запити правоохоронних органів, рішення судів, обмін із міжнародними організаціями, тощо.

13. Отримувачі послуг мають право на отримання інформації про свій унікальний ідентифікатор, у випадку запиту, з перерахуванням усіх критеріїв що стали відомими в процесі його збору.

14. Термін передачі та законних відповідей на запити не має перевищувати чотирнадцять днів із моменту їх отримання, якщо інше не передбачене законодавством.

15. Отримувач послуг має право відмовитись від збору його даних, в такому випадку користування вказаним ресурсом не має бути можливим до моменту повторного погодження.

16. Надавачі послуг мають право на обмеження доступу на основі ідентифікаторів пристроїв, якщо дії їх власників не відповідають угодам та загальноприйнятим нормам та правилами користування та отримання послуг.

17. Відбитки пристроїв мають використовуватись виключно в цілях додаткової фільтрації та можливого обмеження доступу без використання з метою покращення та персоналізації реклами. У випадку, якщо така потреба існує, користувач має окремо погоджувати даний запит. Отримувач послуг має право

відмовитись від використання його відбитків в таких цілях без отримання зустрічної відмови в надані послуг.

18. Термін зберігання даних – три роки.

Вимоги для централізованої реалізації:

1. Централізований реєстр не може бути реалізованим на основі будь якого з наявних надавачів послуг. Це має бути незалежна відокремлена юридична особа.

Примітка: Для України рекомендовано створити єдиний реєстр з дислокацією у будь-якому з міст, населення якого перевищує пів мільйона осіб. З забезпеченням можливості доступу до швидкісної мережі «Інтернет» (щонайменше 1 Гб/сек).

2. Усі відбитки пристроїв мають зберігатися на окремих накопичувачах або пристроях, які пройшли відповідну атестацію Державної служби спеціального зв'язку та захисту інформації України та зберігаються в приміщеннях, що отримали сертифікат КСЗІ. До інформації мають бути застосовані загальні засади конфіденційності, доступності, цілісності.

3. Додатково, має бути встановлене обладнання з відповідними, як до основного, вимогами, головною метою якого буде зберігання резервного копіювання реєстру.

4. Резервне копіювання реєстру має відбуватися щонайменше один раз на тиждень.

5. Видалення інформації можливе лише в письмово аргументованих випадках. Після видалення, інформація має зберігатися у кошику, щонайменше чотирнадцять днів, із забезпеченням можливості до її відновлення.

6. Централізований реєстр має право на витребування та отриманні повної інформації стосовно відбитків пристроїв від усіх організацій, що надають свої веб-послуги на території країни, та за умовами правил зберігають такі відбитки. Процес має бути автоматизованим та відбуватися, щонайменше, один раз на дев’яносто шість годин.

7. Централізований реєстр повинен всіляко допомагати та підтримувати місцеві реєстри.

8. В випадку наявності централізованого реєстру, будь які законні запити, або рішення на предмет відбитків пристроїв мають надходити саме до юридичної особи, яка відповідає за функціонування реєстру.

9. В випадку наявності централізованого реєстру, взаємодія із міжнародними компаніями надавачами веб-послуг відбувається за використання юридичної особи, яка відповідає за функціонування реєстру.

Вимоги для місцевої реалізації:

1. Надавач послуг має встановлювати відповідне програмно-апаратне забезпечення у випадку, якщо показник щоденних відвідувань його веб-ресурсу (або додатку) перевищує десять тисяч пристроїв.

2. Відповідальність за конфіденційність, доступність, цілісність несе надавач послуг який зберігає у себе відбитки пристроїв.

3. Надавач послуг має право самостійно обирати, або створювати, будь-який із методів, що передбачає збір відбитків пристроїв. Головними вимогами є: дотримання загальних правил і норм та забезпечення можливості зберігання максимально великого обсягу даних. В якості порівняння рекомендується використовувати перелік даних отриманий в результаті збору методом скриптів. Приклад – Fingerprint.js. Результат, що буде отриманий має бути, щонайменше, на 75% повним порівняно з запропонованим скриптом.

4. В випадку відсутності централізованого реєстру, всі вимоги щодо його функціонування перекладаються на місцеві реєстри.

5. Якщо надавач веб-послуг не виконує зазначених вимог – до нього можуть застосовуватись методи впливу аж до обмеження роботи на території країни.

На нашу думку, встановлювати чіткі рамки, щодо особливостей та технічних вимог до апаратного забезпечення є доцільним лише у випадках створення централізованого реєстру з урахуванням економічної складової та вартості додаткових послуг. У ситуації, коли мова йде про місцеві реєстри, рішення про тип, склад та технічні характеристики апаратного та програмного забезпечення повинні приймати надавачі послуг особисто виходячи із власних потреб із головною умовою в дотриманні раніше встановлених вимог.

2.7 Висновок рішень спеціальної частини

В результаті виконання спеціальної частини роботи з використанням власних програмно-апаратних пристроїв було проведено аналіз наявної моделі забезпечення анонімізацації.

Розглянуті методи збору даних про користувачів мережі «Інтернет», під час якого було виділено два основні: збір за допомогою програмного забезпечення та збір за допомогою веб-сайтів.

За результатами цих розглядів було виявлено відразу декілька проблем, які відіграють свою роль під час фільтрації трафіку та прийняте рішення про розгляд додаткових методів збору та подальшого аналізу даних. Результатом стало тестування методу збору цифрових відбитків пристроїв. Було проведено детальне вивчення наявних засобів збору інформації методом цифрового відбитку на базі стаціонарних та мобільних пристроїв із вивченням та описом усіх переваг та недоліків. Надалі було проведено тестування на базі власної електронно-обчислювальної машини.

Результатами стала потреба у впроваджені реєстру цифрових відбитків пристроїв. В підрозділі 2.6 була надана можлива модель, де може використовуватись як всеукраїнських реєстр так і реєстр на базі певного невідомого надавача послуг. Були сформовані попередні вимоги та правила в питані впровадження.

3 ЕКОНОМІЧНИЙ РОЗДІЛ

Під час виконання економічного розділу буде визначено вартість впровадження «Всеукраїнського реєстру цифрових відбитків пристроїв з урахуванням витрат на його створення та підтримку протягом року роботи.

У розділі роботи, для розрахунку буде використовуватись метод, при якому обов'язковим є виконання наступних пунктів:

- розрахунок капітальних витрат на придбання і налаштування програмно-апаратних засобів;
- розрахунок річних експлуатаційних витрат на обслуговування об'єкта;
- визначення річного економічного ефекту від впровадження об'єкта;
- визначення показників обраного рішення;
- висновок про економічну доцільність обраного рішення;

3.1 Розрахунок витрат

Для розрахунку капітальних витрат (К) буде використана наступна формула:

$$К = К_{пр} + К_{зпз} + К_{пз} + К_{аз} + К_{навч} + К_{н}, \text{ грн.} \quad (3.1)$$

де $К_{пр}$ – вартість розробки проекту ІБ та залучення зовнішніх консультантів;

$К_{зпз}$ – вартість закупівлі ПО;

$К_{пз}$ – вартість розробки політики безпеки інформації;

$К_{аз}$ – вартість закупівлі апаратного забезпечення та допоміжних матеріалів, тис. грн;

$К_{навч}$ – витрати на навчання технічних фахівців;

$К_{н}$ – витрати на встановлення обладнання та налагодження системи ІБ;

Розрахунок показників:

$К_{пр}$ – точна сума буде визначена шляхом тендерної процедури, оскільки реєстр планується створити державним. На даний час орієнтовна сума витрат на

розробку проекту та залучання зовнішніх консультантів складає близько 1000000 грн.

$К_{зпз}$ буде розраховано як витрати на розробників, які мають самостійно створити та впровадити два методи отримання відбитків пристроїв:

- Кількість розробників – 20 людей;
- Заробітна плата: 35000 грн;
- Термін впровадження – 3 місяці;
-

$$К_{зпз} = 35000 \cdot 20 \cdot 3 = 2100000 \text{ грн} \quad (3.2)$$

$К_{пз}$ – 20000 грн (заробітна плата одного фахівця);

$К_{аз}$ щомісячні витрати 10000 грн · 3 місяці розробки :

$$К_{аз} = 10000 \cdot 3 = 30000 \text{ грн} \quad (3.3)$$

$К_{навч}$ – 0 грн., співробітники не потребують додаткового підвищення кваліфікації.;

$К_{н}$ – 0 грн., вартість за встановлення і налаштування врахована у пункті $К_{зпз}$;

Таким чином, загальні капітальні витрати (К) складають:

$$К = 100000 + 2100000 + 20000 + 30000 = 3150000 \text{ грн} \quad (3.4)$$

3.2 Експлуатаційні витрати

Далі, необхідно визначити річні витрати на користування системою ІБ (С). Розрахунок буде відбуватися за наступною формулою:

$$С = С_{н} + С_{а} + С_{з} + С_{єв} + С_{ел} + С_{о} + С_{тос}, \text{ грн.} \quad (3.5)$$

де $С_н$ – витрати на навчання персоналу;

$С_а$ – річний фонд амортизаційних відрахувань;

$С_з$ – річний фонд заробітної плати інженерно-технічного персоналу;

$С_{єв}$ – єдиний соціальний внесок;

$С_{ел}$ – вартість електроенергії, що споживається апаратурою системи ІБ;

$С_о$ – витрати на залучення сторонніх організацій;

$С_{тос}$ – витрати на технічне й організаційне адміністрування;

Розрахунок показників:

$С_н$ – 0 грн., співробітники не потребують додаткового навчання;

$С_а$ – 50000 грн;

$С_з$ (вказано з урахуванням $С_{єв}$) на міс. – складається з таких компонентів:

Загальна кількість співробітників 30 осіб, із них:

Заробітна плата керуючому складу – 30000 · 3 особи;

Заробітна плата основному складу – 18000 · 20 осіб;

Заробітна плата допоміжному складу – 9000 · 7 осіб;

$С_{єв}$ – єдиний соціальний внесок (22%). ;

$С_{ел}$ – 84000 грн., за рік;

$С_о$ – складається з таких компонентів:

- Щорічний тендер з метою вдосконалення наявних методів отримання відбитків пристроїв – 150000 грн;
- Щорічний тендер з метою вдосконалення надання консультації місцевим реєстрам – 250000 грн;

$$С_о = 150000 + 250000 = 400000 \text{ грн} \quad (3.6)$$

$С_{тос}$ – 10000 грн., додаткові витрат;

Таким чином, загальні річні витрати на користування системою ІБ(С):

$$С = 50000 + (30000 \cdot 3 + 18000 \cdot 20 + 9000 \cdot 7) \cdot 12 + 84000 +$$

$$+400000 + 10000 = 6700000 \text{ грн} \quad (3.7)$$

3.3 Оцінка можливих збитків

У разі впровадження, «Всеукраїнським реєстром цифрових відбитків» не буде проводиться економічна діяльність, головною метою його створення є покращення загальної взаємодії між державними органами та надавачами послуг, а також правоохоронними органами, в питанні зростання загального рівня кібербезпеки держави.

Тому, задля розрахунку можливих збитків буде враховано можливий ризик від штрафних санкцій (адміністративної відповідальності), які перебачені за недотримання вимог у питанні захисту персональних даних, що регламентовано Законом України «Про захист персональних даних» [52]. Надивлячись на те, що ризик такого витоку є мінімальним і складає приблизно 0,2-0,5%, це все одно варте уваги.

Адміністративна відповідальні передбачена ст. 188-39 Кодексу України про адміністративні правопорушення [53]. А саме:

Недодержання встановленого законодавством про захист персональних даних порядку захисту персональних даних, що призвело до незаконного доступу до них або порушення прав суб'єкта персональних даних,

- тягне за собою накладення штрафу на громадян від ста до п'ятисот неоподатковуваних мінімумів доходів громадян і на посадових осіб;

- громадян - суб'єктів підприємницької діяльності - від трьохсот до однієї тисячі неоподатковуваних мінімумів доходів громадян.

Згідно з п. 5 підрозділу 1 розділу XX Податкового кодексу України від 02.12.2010 №2755-VI якщо норми інших законів містять посилання на неоподатковуваний мінімум доходів громадян, то для цілей їх застосування використовується сума в розмірі 17 гривень.

Згідно інформації, що була надана у розділі 1.6 даної роботи, користувачами мережі «Інтернет» в Україні є близько 29,47 мільйонів осіб. Згідно правил наведених у пункті 2.6 роботи, максимальна кількість даних на одному накопичувачі має становити не більше п'ятсот тисяч. Тобто показник максимально можливих записів у реєстрі становить саме таке число.

Вартість штрафу (B_1) за одну особу може складати:

$$B_1 = 300 \cdot 17 = 5100 \text{ грн}$$

В такому випадку, максимально допустимий штраф(B_2) буде складати:

$$B_2 = 5100 \cdot 500000 = 2550000000 \text{ грн}$$

Враховуючи ймовірність (близько 0,2%) показник(B_3) буде становити:

$$B_3 = 2550000000 \cdot 0{,}02 = 51000000 \text{ грн}$$

Але, варто зазначити, що можливість такого штрафу є лише у випадку визнання цієї інформації персональними даними, що має відбуватися згідно до нормативних документів якими передбачене визначення критеріїв відношення даних до персональних.

3.4 Висновки економічного розділу

В результаті проведення економічних розрахунків, з метою оцінки вартості та рентабельності впровадження «Всеукраїнського реєстру цифрових відбитків пристроїв», було розраховано показники капітальних витрат (К) та щорічних витрат (С). Капітальні витрати на запровадження обраних програмно-апаратних засобів склали 3150000 (три мільйони сто п'ятдесят тисяч) гривень. Щорічні витрати, включаючи витрати на оновлення та підтримку складатимуть 6700000 (шість мільйонів сімсот тисяч) рівень.

Для підрахунку можливих збитків, враховуючи специфіку підприємства, було розраховано витрати, які можуть статися через виток персональних даних, які мають зберігатися у реєстрі. У випадку, якщо перелік технічних даних, який потраплятиме до реєстру буде визначений як «персональні дані», то за умови недодержання встановленого законодавством про захист персональних даних порядку захисту персональних даних, що призвело до незаконного доступу,

адміністративна відповідальність буде передбачена штрафом, що з урахуванням ймовірності такого випадку становить показник B_3, а саме – 51000000 (п’ятдесят один мільйон) гривень.

Розрахувати доцільність впровадження на даному етапі не є можливим, оскільки ефект, більшою мірою, є соціально значущим, та для отримання точних даних потребує всебічного окремого вивчення з урахуванням показників як приватних надавачів послуг так і державних інститутів.

ВИСНОВКИ

В результаті виконання кваліфікаційної роботи, було проаналізовано наявний стан мережевого трафіку. Виділено та аргументовано власні критерії поділу трафіку за своїм походженням та маршрутом. Висвітлено проблему зростання машинного трафіку, де зокрема, окрема увага приділялась питанню ботів: їх діяльності, питанню наявних засобів та методів боротьби з ними.

Було приділено увагу проблемі з трафіком, що генерується так званими «Bad Bots», або ж зловмисними ботами, кількість яких невпинно зростає. Окремо було розглянуто приклад втручання «машин» як інструменту для зміни громадської думки з вивченням прикладу виборів в Сполучених Штатах Америки. Провівши аналіз вищезазначеного випадку було прийняте рішення про проведення власного тестування захисту від ботів на прикладі однієї із найбільших соціальних мереж «Інтернету». Результати вказали на недостатній рівень початкової фільтрації на моменті реєстрації аккаунту, що частково висвітлило наявну проблему у фільтрації трафіку.

З метою більш глибокого вивчення окрема увага була приділена факторам, що сприяють зростанню так званого «злочинного» трафіку. В результаті була висвітлена проблема анонімізації.

Головною задачею поставлено «Запропонувати та аргументувати програмно-апаратний засіб або метод, мета якого буде покликана на зменшення рівня анонімізації з кінцевою метою покращення наявного рівня фільтрації небезпечного трафіку».

У спеціальній частині роботи виконано аналіз наявної моделі забезпечення анонімізації з тестуванням на базі власної електронно-обчислювальної машини. Розглянуті методи збору інформації про користувачів з висвітленням наявних проблем та розумінням у потребі до впровадження нових методів та засобів. Увага була приділена методу отримання цифрових відбитків пристроїв. Висновки та результати тестування, отримані в результаті вивчення вищезазначеного методу вказали на можливість до виправлення більшості проблем які наявні у

альтернативних способах збору інформації. Після огляду наявних засобів збору інформації методом цифрового відбитку на базі стаціонарних та мобільних пристроїв із вивченням та описом усіх переваг та недоліків було виявлено аргументовану потребу у впроваджені такого методу. Результатами стала потреба у впроваджені реєстру цифрових відбитків пристроїв. Окремо був наданий перелік правил та рекомендацій у випадку створення реєстру або як всеукраїнського, або на базі певних надавачів послуг.

Під час виконання економічного розділу було розраховано показники капітальних витрат та щорічних витрат у разі впровадження «Всеукраїнського реєстру відбитків пристроїв».

ПЕРЕЛІК ПОСИЛАНЬ

1. Population [Електронний ресурс] // United Nations. – 2019. – Режим доступу до ресурсу: https://www.un.org/en/global-issues/population.
2. Тенденции развития Интернет-аудитории, 2020. – (GfK).
3. Ericsson Mobility Report [Електронний ресурс] // Ericsson. – 2021. – Режим доступу до ресурсу: https://www.ericsson.com/en/reports-and-papers/mobility-report.
4. Cisco Visual Networking Index: Forecast and Trends, 2017–2022. // Cisco. – 2021. – С. 4.
5. Яворська Г. М. Гібридна війна як дискурсивний конструкт / Галина Михайлівна Яворська. // Стратегічні пріоритети. – 2016. – №4. – С. 41–47.
6. Network Traffic [Електронний ресурс] // Techopedia. – 2015. – Режим доступу до ресурсу: https://www.techopedia.com/definition/29917/network-traffic.
7. Воробієнко П.П. Телекомунікаційні та інформаційні мережі / П.П. Воробієнко, Л. А. Нікітюк, П. І. Резніченко. – Київ: САММІТ-Книга, 2020. – 708 с. – (САММІТ-Книга)
8. Global - 2021 Forecast Highlights [Електронний ресурс] // VNI Complete Forecast Highlights. – 2021. – Режим доступу до ресурсу: https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2021_Forecast_Highlights.pdf.
9. FitzGerald J. Business Data Communications and Networking / J. FitzGerald, A. Dennis, A. Durcikova, 2020. – С. 280–290.
10. Nguyen N. Essential Cyber Security Handbook / Nam H Nguyen., 2018.
11. Bad Bot Report 2021. // Imperva. – 2021. – №1.
12. Bhattacharyya D. K. DDoS Attacks: Evolution, Detection, Prevention, Reaction, and Tolerance / D. K. Bhattacharyya, J. K. Kalita // DDoS Attacks: Evolution, Detection, Prevention, Reaction, and Tolerance / D. K. Bhattacharyya, J. K. Kalita., 2016. – С. 5.

13. Кибербезопасность в условиях электронного банкинга. Практическое пособие – Київ: Прометей, 2020. – 522 с. – (Прометей).
14. Ендою Х. Безопасность веб-приложений / Хофман Ендою. – Мінськ: Питер, 2021.
15. Редько М. М. Інформатика та компютерна техніка.: Навчальних посібник для І-ІІ р.а. / М. М. Редько. – Вінниця, 2017. – 567 с.]
16. Science and Practice, Actual Problems, Innovations – Амстердам, Нідерланди, 2021. – 518 с.
17. Payment Card Industry (PCI) Data Security Standard [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: https://www.pcisecuritystandards.org/documents/PCI_DSS_v3-2-1_RU.pdf.
18. The Global Risks Report 2019. // World Economic Forum. – 2020. – С. 114.
19. The 4 Trends That Prevail on the Gartner Hype Cycle for AI, 2021 [Електронний ресурс] // Gartner. – 2021. – Режим доступу до ресурсу: https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021.
20. Утечки данных 2019: статистика, тенденции кибербезопасности и меры по снижению рисков взлома [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: https://vc.ru/services/103616-utechki-dannyh-2019-statistika-tendencii-kiberbezopasnosti-i-mery-po-snizheniyu-riskov-vzloma.
21. Stoecklin M. DeepLocker: How AI Can Power a Stealthy New Breed of Malware [Електронний ресурс] / Marc Ph. Stoecklin. – 2018. – Режим доступу до ресурсу: https://securityintelligence.com/deeplocker-how-ai-can-power-a-stealthy-new-breed-of-malware/.
22. Russian active measure; Campaigns and Interference in the 2016 u.s. Election volume 2: Russia's use of social media with additional views – Washington: Committee Sensitive, 2018.
23. Гэри Маркус, Эрнест Дэвис. Rebooting AI: Building Artificial Intelligence We Can Trust. / Интеллектуальная Литература, 2021. — 304 с.

24. Міністерство цифрової трансформації України. Цифрова грамотність населення України / Міністерство цифрової трансформації України. – 2021. – С. 211.
25. Facebook, Inc. – California, 2017. – 59 c. – (United States Securities and Exchange Commission).
26. Facebook CEO Zuckerberg testifies at congressional hearings [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: https://www.reuters.com/live/zuckerberg-testimony.
27. Deep Entity Classification: Abusive Account Detection for Online Social Networks / [T. Xu, G. Goossen, H. K. Cevahir та ін.]. – 2019. – С. 1–18.
28. DIGITAL 2021: UKRAINE [Електронний ресурс]. – 2021. – Режим доступу до ресурсу: https://datareportal.com/reports/digital-2021-ukraine.
29. Shah F. Dark Web / Fahad Shah., 2018. – 9 с.
30. Tor Project. Tor Users [Електронний ресурс] / Tor Project. – 2022. – Режим доступу до ресурсу: https://metrics.torproject.org/userstats-relay-country.html.
31. Кримінальний Кодекс України [Електронний ресурс]. – 2001. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2341-14#Text.
32. З початку 2020 року до кіберполіції надійшло понад 25 тисяч звернень щодо Інтернет-шахрайства [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: https://cyberpolice.gov.ua/news/z-pochatku--roku-do-kiberpolicziyi-nadijshlo-ponad--tysyach-zvernen-shhodo-internet-shaxrajstva-6472/.
33. (Worldwide Infrastructure Security Report, WISR) [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: https://www.netscout.com/report/.
34. Єдиний державний реєстр судових рішень [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: https://reyestr.court.gov.ua/.
35. Махницький О. В. Боротьба з кіберзлочинністю: вітчизняний та зарубіжний досвід та напрямки діяльності / О. В. Махницький, О. О. Косиченко // ІІ Міжнародна науково-практична конференція / О. В. Махницький, О. О. Косиченко. – Дніпро, 2018. – (ДДУВС). – (2). – С. 277–281

36. Андрусенко С. В. Боротьба з кіберзлочинністю – проблема транснаціонального масштабу / С. В. Андрусенко // Протидія злочинності: проблеми практики та науково-методичне забезпечення / С. В. Андрусенко. – Одеса, 2016. – С. 16–18.
37. Сергій Демедюк: кіберполіція запроваджує навчальний курс для оперативних працівників поліції [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: https://cyberpolice.gov.ua/news/sergij-demedyuk-kiberpolicziya-zaprovadzhuye-navchalnyj-kurs-dlya-operatyvnyx-praczivnykiv-policziyi-8584/.
38. Басиня Є. Сетевая информационная безопасность и анонимизация / Є. Басиня. – Новосибірськ: НГТУ, 2018. – 75 с.
39. Suler J. The Online Disinhibition Effect / John Suler // Cyberpsychology & Behavior / John Suler. – Нью-Джерсі, 2018. – С. 321–326.
40. A branch hash function as a method of message synchronization in anonymous P2P conversations / [A. Kobusińska, J. Brzeziński, M. Maciejewski та ін.]. – Варшава, 2018. – 495 с.
41. Шраго А. О. Протидія порнографії як засіб забезпечення інформаційної безпеки / Альона Олексіївна Шраго // Економічна та інформаційна безпека: проблеми та перспективи / Альона Олексіївна Шраго. – Дніпро, 2018. – С. 262–266.
42. Семеріков С. О. Фундаменталізація навчання інформатичних дисциплін у вищій школі / С. О. Семеріков. – Кривий Ріг, 2018. – 339 с. – (Мінерал).
43. Blokdyk G. VPN Network a Complete Guide / Gerardus Blokdyk. – Варшава, 2019. – 278 с. – (Emereo Pty Limited).
44. Google about company [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: https://about.google/intl/ALL_ru/.
45. Політика конфіденційності Google [Електронний ресурс]. – 2021. – Режим доступу до ресурсу: https://www.gstatic.com/policies/privacy/pdf/20210701/7yn50xee/google_privacy_policy_ru.pdf.

46. Casey E. Digital Evidence and Computer Crime: / Eoghan Casey. – Каліфорнія: ELSIVIER, 2021. – 810 с. – (3).
47. Посібник з європейського права у сфері захисту персональних даних – Страсбург: Агенція ЄС з основоположних прав, 2018. – 436 с.
48. Гусев П. Д. The Review of Existing Digital Fingerprinting Algorithms [Електронний ресурс] / П. Д. Гусев. – 2018. – Режим доступу до ресурсу: https://bit.mephi.ru/index.php/bit/article/viewFile/68/74.
49. Eckersley P. How Unique Is Your Web Browser? [Електронний ресурс] / Peter Eckersley. – 2018. – Режим доступу до ресурсу: https://coveryourtracks.eff.org/static/browser-uniqueness.pdf.
50. Cao Y. Browser Fingerprinting via OS and Hardware Level Features [Електронний ресурс] / Y. Cao, S. Li, E. Wijmans. – 2018. – Режим доступу до ресурсу: http://yinzhicao.org/TrackingFree/crossbrowsertracking_NDSS17.pdf.
51. Fingerprintjs [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: https://dev.fingerprintjs.com/docs.
52. Закон України «Про захист персональних даних» [Електронний ресурс]. – 2010. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2297-17#Text.
53. Кодекс України про адміністративні правопорушення [Електронний ресурс]. – 1984. – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/80731-10#Text.
54. Методичні рекомендації до виконання кваліфікаційних робіт бакалаврів спеціальності 125 Кібербезпека/ Упоряд.: О.В. Герасіна, Д.С. Тимофєєв, О.В. Кручінін, Ю.А. Мілінчук – Дніпро: НТУ «ДП», 2020. – 47 с.

ДОДАТОК А. Відомість матеріалів кваліфікаційної роботи

№	Формат	Найменування	Кількість листів	Примітка
1	А4	Реферат	2	
2	А4	Список умовних скорочень	1	
3	А4	Зміст	1	
4	А4	Вступ	2	
5	А4	1 Розділ	45	
6	А4	2 Розділ	30	
7	А4	3 Розділ	6	
8	А4	Висновки	2	
9	А4	Перелік посилань	5	
10	А4	Додаток А	1	
11	А4	Додаток Б	1	
12	А4	Додаток В	1	
13	А4	Додаток Г	1	

ДОДАТОК Б. Перелік документів на оптичному носії

1 Пояснювальна записка Шабельник С.П.docx

2 Пояснювальна записка Шабельник С.П.pdf

3 Презентація Шабельник С.П.pptx

ДОДАТОК В. Відгуки керівників розділів

Відгук керівника економічного розділу:

Керівник розділу _____ _____
(підпис) (ініціали, прізвище)

ДОДАТОК Г. ВІДГУК

на кваліфікаційну роботу студента групи 125м-20-2
Шабельника Сергія Павловича
на тему: «Вдосконалення способів фільтрації небезпечного трафіку на основі реєстру цифрових відбитків пристроїв»

Пояснювальна записка складається зі вступу, трьох розділів і висновків, викладених на 97 сторінках.

Метою кваліфікаційної роботи є вдосконалення існуючих способів фільтрації небезпечного трафіку та розробка рекомендацій щодо створення реєстру цифрових відбитків пристроїв.

Для досягнення поставленої мети в кваліфікаційній роботі вирішуються наступні задачі: аналіз особливостей мережевого трафіку та моделей забезпечення анонімності, дослідження методів збору інформації про користувача. Запропоновано методи вдосконалення існуючих методів.

Розроблено рекомендації по створенню реєстру цифрових відбитків пристроїв.

Практичне значення результатів кваліфікаційної роботи полягає у отриманих результатах дослідження властивостей небезпечного трафіку, ботів, методів забезпечення анонімності та рекомендацій, необхідних для створення цифрового реєстру відбитків пристроїв.

За час дипломування Шабельник С.П. проявив себе фахівцем, здатним самостійно вирішувати поставлені задачі та заслуговує присвоєння кваліфікації магістра за спеціальністю 125 Кібербезпека, освітньо-професійна програма «Кібербезпека».

Кваліфікаційна робота заслуговує оцінки «відмінно».

Керівник
кваліфікаційної роботи
доц. каф. БІТ, к.т.н. Олександр САФАРОВ