Хмельницький національний університет
Факультет інформаційних технологій
Кафедра комп'ютерної інженерії та інформаційних систем

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

Галузь знань 12 – Інформаційні технології

Спеціальність 123 –Комп'ютерна інженерія

на тему «Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет»

КвРКІ. 160125.21.01.35 ПЗ

Виконав: студент 2 курсу, група КІ2м-20-1 _____ Шевченко Р.С.
Підпис Ініціали, прізвище

Керівник доктор техн. наук, професор _____ Гурман І.В.
Науковий ступінь, вчене звання Підпис Ініціали, прізвище

До захисту допускаю:
Зав. кафедри КІІС, д.т.н., проф.
Т.О. Говорущенко
19 05 2022 р.

Хмельницький, 2022

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет Інформаційних технологій

Кафедра Комп'ютерної інженерії та інформаційних систем

Освітній рівень магістр

Галузь знань 12 Інформаційні технології

Спеціальність 123 Комп'ютерна інженерія

Освітня програма освітньо-наукова програма «Комп'ютерна інженерія та програмування»

ЗАТВЕРДЖУЮ

Зав. кафедри Т.Ю.Говорущенко

"01" 09 2021 р.

ЗАВДАННЯ
НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ)

Шевченку Роману Сергійовичу

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет

Керівник проекту (роботи) Гурман І.В., д.т.н., професор

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 06.01.2022 р. № 1

2. Строк подання студентом проекту (роботи) на кафедру 03.05.2022 р.

3. Вихідні дані до проекту (роботи) Завдання на дипломне проектування

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)

Аналіз видів та основних характеристик цензури в інтернеті

Модель процесу виявлення з'єднання в комп'ютерних системах

Метод переадресації та токенова економіка

Дослідження ефективності роботи

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень)

6. Консультанти розділів дипломного проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Нормоконтроль	Лисенко С.М., професор кафедри КІІС		
Антиплагіат	Нічепорук А.О., доцент кафедри КІІС		

7. Дата видачі завдання «06» 09 2021р.

КАЛЕНДАРНИЙ ПЛАН

№з/п	Назва етапів (розділів) дипломного проекту (роботи)	Термін виконання етапів проекту (роботи)	Примітка
1	Вибір напряму дослідження та узгодження тематики ДРМ з керівником	05.09.2021	виконано
2	Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження	05.10.2021	виконано
3	Робота над розділом 1 – аналіз відомих моделей, методів за темою; постановка задачі	05.11.2021	виконано
4	Робота над розділом 2 – розробка моделей для вирішення поставленої задачі	05.12.2021	виконано
5	Робота над науковою статтею	05.01.2022	виконано
6	Робота над розділом 3 – розробка методів для вирішення поставленої задачі	15.02.2022	виконано
7	Робота над розділом 4 – проектування та розробка ПЗ для вирішення поставленої задачі, експериментальна частина	05.04.2022	виконано
8	Оформлення пояснювальної записки згідно вимог	15.04.2022	виконано
9	Попередній захист ДРМ	18.04.2022	виконано
10	Захист ДРМ на засіданні ЕК	До 10.05.2022	

Студент _____ Підпис Р.С. Шевченко Ініціали, прізвище

Керівник проекту (роботи) _____ Підпис І.В. Гурман Ініціали, прізвище

РЕФЕРАТ

Тема дипломної роботи: Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі інтернет.

Автор роботи: Шевченко Р.С., студент групи КІ2м-20-1.

Керівник роботи: Гурман І.В., доктор технічних наук, професор кафедри комп'ютерної інженерії та інформаційних систем.

Пояснювальна записка: 85 с., 24 рис., 5 табл., 3 дод., 48 джерел.

ПЕРЕЛІК КЛЮЧОВИХ СЛІВ: : Брандмаузер, VPN, проксі-сервер, цензура, GFW, шифрування, DNS система, coockie-файли.

Об'єктом дослідження є процес мульти-проксування.

Предметом дослідження є методи удосконалення технології мульти-проксування.

Метою дипломної роботи є підвищення ефективності технології мульти-проксування розроблення методів забезпечення конфіденційності в мережі інтернет.

Наукова новизна отриманих результатів: мульти-проксі дослідження стають все більш поширеними. Проблеми дослідження з використанням кількох посередників – визначення запитань дослідження, лідерство, вибір місця та визначення ядра, зберігання даних, хронологія, представлення результатів, числові інструменти та інтерпретація даних. Природа проксі-даних обговорюється з точки зору фізичних проксі та біотичних проксі. Зміни втрат при запалюванні та використання функцій передачі розглядаються як приклади проблем при інтерпретації даних із досліджень із застосуванням кількох проксі. Наголошується на важливості аналізу пилку та аналізу у мульти-проксі дослідженні. Окреслено майбутні напрямки щодо того, як мульти-проксі дослідження можуть сприяти розумінню біотичних реакцій на зміни навколишнього середовища.

– набув подальшого розвитку метод мульти-проксування;

Практична значимість отриманих результатів полягає у:

1) у розробленні схеми мережі мульти-проксі, яка візуалізує взаємозв'язки між вузлами, блоками, смарт-контрактом проксі;

2) розробленні правил для визначення необхідності даних, які забезпечують аналіз конфіденційних даних на предмет відшукання всіх наперед визначених необхідних елементів мульти-проксі; прийняття рішення про достатність або недостатність інформації; виведення візуалізованих підказок, яких елементів даних не вистачає, для забезпечення можливості швидкого доповнення даних; кількісну оцінку достатності даних;

3) розробленні архітектури системи удосконалення технології мульти-проксування, а також у проектуванні та реалізації її окремих модулів.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ACL - список контролю доступу.

DNS - система доменних імен.

HTTP - система доменних імен.

GFW - великий китайський брандмауер.

IP-протокол Інтернету.

ICMP - Протокол керуючих повідомлень Інтернету.

MAC - контроль доступу до медіа.

VPN - віртуальна приватна мережа.

VPS -віртуальний приватний сервер.

ВСТУП

Інтернет та цифрова конфіденційність розглядаються інакше, ніж традиційні очікування конфіденційності. Конфіденційність в Інтернеті в першу чергу стосується захисту інформації користувачів. Професор права Джеррі Кан пояснює, що термін конфіденційність означає простір, рішення та інформацію. З точки зору простору, люди очікують, що в їхні фізичні простори (наприклад, будинки, автомобілі) не будуть вторгненні.

Людям, які не надто піклуються про конфіденційність в Інтернеті, не потрібно досягати повної анонімності. Користувачі Інтернету можуть захистити свою конфіденційність шляхом контрольованого розкриття особистої інформації. Розкриття IP-адрес, профілювання, що не ідентифікує особу, та подібна інформація може стати прийнятним компромісом для зручності, яку в іншому випадку користувачі можуть втратити, використовуючи обхідні шляхи, необхідні для суворого придушення таких деталей. З іншого боку, деякі люди прагнуть набагато більшої конфіденційності.

У цьому випадку вони можуть намагатися досягти анонімності в Інтернеті щоб не дати жодним третім сторонам можливості пов'язувати діяльність в Інтернеті з особистою інформацією користувача. Щоб зберегти конфіденційність своєї інформації, люди повинні бути обережними з тим, що вони викладають і переглядають в Інтернеті. Під час заповнення форм і купівлі товарів відстежується інформація, і оскільки вона не була приватною, деякі компанії розсилають Інтернет користувачам спам і рекламу подібних продуктів.

Існує також кілька урядових організацій, які до певної міри захищають конфіденційність та анонімність особи в Інтернеті. У статті, представленій FTC, у жовтні 2011 року було звернено увагу на ряд вказівок, які допомагають окремому користувачеві Інтернету уникнути можливої крадіжки особистих даних та інших кібератак. Рекомендується, серед іншого, запобігати або обмежувати використання номерів соціального страхування в Інтернеті, бути обережними до електронних листів, включаючи спам, пам'ятати про особисту фінансову інформацію,

створювати надійні паролі та керувати ними, а також використовувати безпечні способи перегляду веб сторінок [1].

Розміщення інформації в Інтернеті може бути шкідливим або піддавати людей до зловмисних атак. Деяка інформація, розміщена в Інтернеті, зберігається десятиліттями, залежно від умов надання послуг та політики конфіденційності окремих послуг, які пропонуються в Інтернеті. Це може включати коментарі, написані в блогах, зображеннях і веб сайтах, таких як Facebook і Twitter, тощо. Деякі роботодавці можуть досліджувати потенційного працівника, шукаючи в Інтернеті подробиці їхньої поведінки в Інтернеті, що, можливо, вплине на результат успіху кандидата [2].

Компанії відстежують поведінку користувачів на веб сайтах, щоб потім використовувати цю інформацію, наприклад, надсилаючи рекламу на основі історії перегляду веб сторінок .

Існує багато способів, за допомогою яких люди можуть розголошувати свою особисту інформацію, наприклад, використовуючи соціальні мережі та надсилаючи інформацію про банківські рахунки та кредитні картки на різні веб сайти. Крім того, безпосереднє спостереження за поведінкою користувачів, наприклад за їхніми пошуковими запитами або вмістом профілю Facebook чи інших соціальних мереж, може автоматично оброблятися, щоб зробити висновок про потенційно нав'язливі деталі про особу, такі як сексуальна орієнтація, політичні та релігійні погляди, раса, вживання психоактивних речовин, інтелект, і особистість [3].

Ті, хто стурбований конфіденційністю в Інтернеті, часто посилаються на низку ризиків конфіденційності – подій, які можуть поставити під загрозу конфіденційність – з якими можна зіткнутися під час онлайн-дій[4]. Вони варіюються від збору статистичних даних про користувачів до більш зловмисних дій, таких як поширення шпигунського програмного забезпечення та експлуатації різних форм помилок (помилки програмного забезпечення).

Кілька веб сайтів соціальних мереж намагаються захистити особисту інформацію своїх користувачів, а також надають попередження через угоду про

конфіденційність та умови користування їхнім веб сайтом. Наприклад у Facebook, налаштування конфіденційності доступні для всіх зареєстрованих користувачів: вони можуть заблокувати певним особам доступ до свого профілю, вони можуть вибрати своїх «друзів», а також можуть обмежити доступ до їхніх фотографій, відео та іншої інформації з їхнього профілю. Налаштування конфіденційності також доступні на інших веб сайтах соціальних мереж, таких як Google Plus і Twitter.

Користувач може застосувати такі налаштування при наданні особистої інформації в Інтернеті. Electronic Frontier Foundation створив набір посібників, щоб користувачам було легше використовувати ці налаштування конфіденційності та Zebra Crossing: простий у використанні контрольний список цифрової безпеки – це онлайн-ресурс, який підтримується волонтерами[5].

Наприкінці 2007 року Facebook запустив програму Beacon, в рамках якої записи про оренду користувачів були оприлюднені для друзів. Багато людей були розлючені цим порушенням конфіденційності, і було відкрито судову справу Lane v. Facebook, Inc [6].

Діти та підлітки часто використовують Інтернет (включаючи соціальні медіа) у спосіб, який загрожує їх конфіденційності: це викликає дедалі більше занепокоєння серед батьків. Молодь також може не усвідомлювати, що вся їхня інформація може переглядатися та відстежуватися під час відвідування певного сайту, і що вони мають захищати власну конфіденційність.

Вони повинні бути поінформовані про всі ці ризики. Наприклад, у Twitter загрози включають скорочені посилання, які можуть призвести до потенційно шкідливих веб сайтів або забороненого вмісту.

Загрози з боку використання електронної пошти включають шахрайство з електронною поштою та вкладення, які переконують користувачів встановити зловмисне програмне забезпечення та розкрити особисту інформацію. На сайтах Torrent загрози відносять зловмисне програмне забезпечення, яке ховається у відео, музиці та завантаженні програмного забезпечення.

Під час використання смартфона до загроз відносять геолокацію, що означає, що телефон може визначити його місцезнаходження та опублікувати його в Інтернеті для всіх. Користувачі можуть захистити себе, оновлюючи антивірусне програмне забезпечення, використовуючи параметри безпеки та конфіденційності, завантажуючи оновлення для встановлених програм, перевіряючи електронну пошту, видаляючи шпигунське програмне забезпечення, керуючи файлами cookie, використовуючи шифрування та блокуючи спливаючі вікна [7-8].

Однак більшість людей не мають уявлення про те, як зробити все перераховане вище. Багато компаній наймають професіоналів для вирішення таких питань, але більшість людей можуть лише зробити все можливе, щоб навчитися цьому самотужки [9].

У 1998 році Федеральна торгова комісія США розглянула відсутність конфіденційності для дітей в Інтернеті і створила Закон про захист конфіденційності дітей в Інтернеті (COPPA).

Закон COPPA обмежує можливості збору інформації від дітей та створення попереджувальних повідомлень, якщо була представлена потенційна шкідлива інформація або вміст. У 2000 році для впровадження політики безпеки в Інтернеті було розроблено Закон про захист дітей в Інтернеті (CIPA).

Політика вимагала вжиття заходів захисту технологій, які можуть фільтрувати або блокувати доступ дітей до Інтернету та фотографій, що можуть бути шкідливими для них. Школи та бібліотеки повинні дотримуватися цих вимог, щоб отримати знижки від програми E-rate [10].

Ці закони, інформаційні кампанії, стратегії батьківського нагляду та Інтернет-фільтри можуть допомогти зробити Інтернет безпечнішим для дітей у всьому світі [11].

Побоювання користувачів Інтернету щодо конфіденційності становлять серйозний виклик. Завдяки розвитку технологій доступ до Інтернету став доступним для використання з будь-якого пристрою в будь-який час. Однак збільшення доступу з кількох джерел збільшує кількість точок доступу для атаки[12].

Під час онлайн-опитування приблизно сім із десяти людей відповіли, що найбільше їх хвилює їхня конфіденційність в Інтернеті, а не поштою чи телефоном. Конфіденційність в Інтернеті повільно, але впевнено стає загрозою, оскільки особисті дані людини можуть потрапити в чужі руки, якщо воно потрапили в Інтернет [13].

Усі веб сайти отримують, а багато хто відстежує IP-адресу комп'ютера відвідувача. Компанії з часом порівнюють дані, щоб пов'язати ім'я, адресу та іншу інформацію з IP-адресую [14]. Існує неоднозначність щодо того, наскільки приватні IP-адреси.

Суд Європейського Союзу постановив, що їх потрібно розглядати як особисту інформацію, якщо веб сайт, який їх відстежує, або третя сторона, наприклад постачальник послуг, знає ім'я або адресу власника IP-адреси, що було б вірно для статичних IP-адрес, а не для динамічних [15].

Правила Каліфорнії стверджують, що IP-адреси потрібно розглядати як особисту інформацію, якщо компанія самостійно може пов'язати їх з іменем та адресую [16-18].

Суд Альберти постановив, що поліція може отримати IP-адреси та пов'язані з ними імена і адреси без ордера на обшук; Поліція Калгарі, Альберта, знайшла IP-адреси, які використовувались для онлайн-злочинів. Постачальник послуг надав поліції імена та адреси, пов'язані з цими IP-адресами [19].

Файли cookie – це дані, що зберігаються на комп'ютері користувача і допомагають в автоматичному доступі до веб сайтів чи веб функцій, або іншої інформацію про стан, необхідну для складних веб сайтів.

Вони також можуть використовуватися для відстеження користувачів, зберігаючи спеціальні дані історії використання в файлах cookie, і такі файли cookie – наприклад, ті, що використовуються Google Analytics – називаються файлами cookie для відстеження.

Файли cookie є загальною проблемою у сфері конфіденційності в Інтернеті. Хоча розробники веб сайтів найчастіше використовують файли cookie для законних технічних цілей, трапляються випадки зловживання. У 2009 році двоє

дослідників відзначили, що профілі соціальних мереж можуть бути підключені до файлів cookie, що дозволить підключити профіль соціальної мережі до «звичок» перегляду користувачем [20].

У минулому веб сайти, як правило, не повідомляли користувачам про збереження файлів cookie, хоча вони використовуються як способи складання довгострокових записів історії перегляду приватними особами – проблема конфіденційності спонукала законодавців Європи та США вжити заходів у 2011 році файли cookie також можуть бути корисними для комп'ютерної криміналістики [21-22].

У минулі роки більшість комп'ютерних користувачів не були повністю обізнані про файли cookie, але вони усвідомлювали про можливий шкідливий вплив файлів cookie в Інтернеті: нещодавнє дослідження показало, що 58 % користувачів видалили файли cookie зі своїх комп'ютерів принаймні один раз, і що 39 % користувачів щомісяця видаляють файли cookie.

Оскільки файли cookie є основним способом націлювання рекламодавців на потенційних клієнтів, а клієнти видаляють файли cookie, деякі рекламодавці почали використовувати постійні файли cookie Flash але сучасні браузери та програмне забезпечення для захисту від шкідливих програм тепер можуть виявляти, блокувати та видаляти такі файли cookie.

Розробники файлів cookie передбачали, що лише веб сайт, який спочатку розповсюджував файли cookie користувачам, міг отримати їх повертаючи лише ті дані, якими вже володіє. Однак на практиці програмісти можуть обійти це обмеження.

Cookies має свої переваги. Одна з них полягає в тому, що для веб сайтів, які ви часто відвідуєте, файли cookie можуть дозволити вам не вводити пароль при кожному вході в систему.

Файл cookie також може відстежувати ваші вподобання, щоб показувати вам веб сайти, які можуть вас зацікавити. Файли cookie дозволяють використовувати більшість веб сайтів безкоштовно без будь-яких видів оплати. Деякі з цих переваг також розглядаються як негативні.

Наприклад, одним із найпоширеніших способів крадіжки є хакери, які викрадають ваше ім'я користувача та пароль, що зберігаються в файлах cookie. Хоча багато сайтів безкоштовні, вони продають свій простір рекламодавцям.

Ці оголошення, персоналізовані відповідно до ваших уподобань, іноді можуть зависати на комп'ютері або викликати роздратування. Файли cookie в основному нешкідливі, за винятком файлів cookie від третіх сторін. Ці файли cookie створюються не самим веб сайтом, а компаніями, які використовують рекламні веб банери.

Ці сторонні файли cookie небезпечні, оскільки вони беруть ту саму інформацію, що й звичайні файли cookie, наприклад, звички перегляду та часто відвідувані веб сайти, але потім вони передають цю інформацію іншим компаніям.

Файли cookie часто асоціюються зі спливаючими вікнами, оскільки ці вікна досить часто, хоча й не завжди, пристосовані до вподобань людини. Ці вікна викликають роздратування, оскільки кнопка закриття може бути стратегічно прихована в малоймовірній частині екрану.

У найгіршому випадку ці спливаючі оголошення можуть на ввесь екран, і поки хтось намагається їх закрити, вони можуть відкрити в браузері інший небажаний веб сайт.

Файли cookie сприймаються так негативно, тому що їх значення не розуміють і вони залишаються непоміченими, поки хтось просто користується Інтернетом. Ідея про те, що за кожним рухом, який можна зробити під час перебування в Інтернеті, спостерігають, лякає більшість користувачів.

Деякі користувачі вирішують вимкнути файли cookie у своїх веб браузерах [23]. Така дія може зменшити деякі ризики конфіденційності, але може серйозно обмежити або перешкодити функціональності багатьох веб сайтів. Усі великі веб переглядачі мають вбудовану можливість відключення, та не вимагають жодних зовнішніх програм. В якості альтернативи користувачі можуть видаляти будь-які збережені файли cookie.

Деякі браузери (наприклад, Mozilla Firefox і Opera) пропонують можливість автоматичного очищення файлів cookie, коли користувач закриває браузер.

Існують побоювання, що переваги видалення файлів cookie для конфіденційності були перебільшені [24].

Процес профілювання (також відомий як «відстеження») збирає та аналізує декілька подій, кожна з яких відноситься до однієї сутності для того, щоб отримати інформацію (особливо про моделі діяльності), що стосується вихідної сутності. Деякі організації займаються профілюванням перегляду веб сторінок, збираючи URL-адреси відвідуваних сайтів.

Отримані профілі потенційно можуть зв'язуватися з інформацією, яка ідентифікує особу, що здійснила перегляд.

Деякі веб орієнтовані організації маркетингових досліджень можуть використовувати цю практику законно, наприклад, для створення профілів «типових користувачів Інтернету». Такі профілі, які описують середні тенденції великих груп користувачів Інтернету, а не конкретних осіб, можуть виявитися корисними для аналізу ринку . Хоча зведені дані не є порушенням конфіденційності, деякі люди вважають, що початкове профілювання порушує їхні права.

Профілювання стає більш спірним питанням відносно конфіденційності, коли узгодження даних пов'язує профіль особи з особистою інформацією особи.

Актуальність роботи полягає в удосконалені технологій мульти-проксування для забезпечення конфіденційності в мережі Інтернет.

Метою дипломної роботи є підвищення конфіденційності в мережі Інтернет за допомогою технології мульти-проксування.

Поставлена мета досягається розв’язанням таких основних задач:

- аналіз відомих методів та рішень для забезпечення конфіденційності за допомогою мульти-проксі технології;
- моделювання процесу мульти-проксування;
- розроблення методу оцінювання конфіденційності інформації;
- розроблення методу виконання мульти-проксування.

Об’єктом дослідження є процес мульти-проксування.

Предметом дослідження є процес забезпечення конфіденційності.

Наукова новизна отриманих результатів:

1) в розроблені моделі блоку мульти-проксування, фрагменту проксі та вузла, представлені у формалізованому та схематичному вигляді, а також моделі процесу мульти-проксування;

2) розроблено метод оцінювання достатності забезпечення конфіденційності даних;

3) набув подальшого розвитку метод мульти-проксування.

Практична цінність отриманих результатів. В результаті виконаного наукового дослідження розроблені схеми мережі мульти-проксування яка візуалізує взаємозв'язки між процесами.

Для розв'язання поставлених задач використовуються основні положення загальної теорії систем, системного аналізу (ієрархічності, декомпозиції та ін.), теорії моделювання процесів.

Внаслідок проведення моделювання процесу мульти-проксування, використано теоретико-множинні підходи, алгебру систем, апарат модельно-орієнтованих підходів, методи концептуального моделювання, принципи побудови баз знань та формування логічного висновку, евристичні оцінки.

За темою дипломної роботи опублікована одна стаття у фаховому науковому виданні.

Шевченко Р.С. Забезпечення конфіденційності передачі інформації в мережі Інтернет // Міжнародна науково-практична конференція «Сучасні інформаційні технології 2022» ISM–2021 (Одеса, 19-20 травня 2022).

1 АНАЛІЗ ВИДІВ ТА ОСНОВНИХ ХАРАКТЕРИСТИК ЦЕНЗУРИ В МЕРЕЖІ ІНТЕРНЕТ

1.1 Коротка історія GFW

Цензура існує вже багато років за різними правилами і нормами, оскільки Інтернет став спільною комунікаційною платформою, візьмемо найвідоміший приклад: GFW, також відомий як Великий брандмауер Китаю, є найбільшою розгалуженою і складною системою інтернет-цензури та моніторингу по всій країні та по всьому світу.

Це поєднання апаратного та програмного забезпечення яке спрямоване на розрізнення та блокування мережевого трафіку і відправленню в чорний список. Цей небажаний веб-контент містять пошукові системи, наприклад, Google і DuckDuckGo, соціальні медіа та веб-сайти соціальних мереж, наприклад, Twitter, Facebook, Instagram, YouTube тощо.

GFW використовує кілька методів і модулів щоб заборонити громадянам доступ до заблокованого вмісту. В останні кілька десятиліть більшість протоколів прикладного рівня, таких як HTTP і DNS використовуються безпосередньо на основі протоколів транспортного рівня, наприклад, TCP і UDP.

Хоча ці протоколи застосування забезпечують стандартні способи передачі ресурсів, вони вразливі до атак «людина в середині» через недостатні міркування безпеки. Весь мережевий трафік знаходиться в простому тексті між джерелом і місцем призначення. Відсутність безпеки дає GFW шанс виявити вміст у певних протоколах. У перші роки з 2002 року, GFW починає розробку системи фільтрації ключових слів, щоб заблокувати доступ до вибраних цільових веб-сайтів, включаючи деякі пошукові системи та веб-сайти соціальних мереж які можуть поширювати величезну кількість інформації [25].

Основні методи, включаючи виявлення ключових слів HTTP, скидання TCP-з'єднання, викрадення DNS та DNS отруєння. Деякі прості проксі TCP і HTTP не працюють через ці методи. GFW також використовує блокування IP-адрес, щоб

запобігти доступу користувачів мережі до отруєних веб-сайтів за правильними IP-адресами.

Пізніше деякі інструменти обходу цензури розробляються для ухилення від інтернет-спостереження, наприклад, VPN, HTTP проксі і HTTPS проксі, а також розробляються анонімні однорангові системи зв'язку, наприклад, проект Tor. Після тривалого періоду гонки озброєнь між системами обходу GFW і цензури, GFW починає блокувати і фільтрувати конкретний мережевий трафік, протокол VPN.

Крім того, як тільки будуть виявлені функції обходу, ці постачальники послуг, будуть заблоковані IP- адресою GFW. Крім того, GFW має здібності, щоб виявляти приховані методи обходу [26]. Виявилося, що GFW використовує механізми зондування наприклад, надсилання випадкових двійкових даних кожні 15 хвилин для виявлення та блокування Мостів Tora.

В останні роки вона починає приймати виявлення сертифікатів HTTPS і індифікації імені сервера (SNI) проти трафіку HTTPS до заблокованих веб-сайтів, але він не використовувався у великих масштабах. Причиною того, що механізми працюють, є те, що адреси цільових серверів все ще можуть бути ідентифіковані.

Хоча GFW може блокувати шкідливі веб-сайти та обмежувати поширення насильницької і кримінальної інформації, деякі прибуткові і корисні сайти також блокуються. Там є деякі заміни цих заблокованих веб-сайтів, наприклад, Китай має свій пошук називається Baidu, але він не дуже точний при пошуку іноземних новин, і літератури. Тому існування GFW створює бар'єр для користувачів мережі.

1.2 Категорії обхідних систем

Існують деякі системи обходу цензури, спрямовані на допомогу людям, всередині цензурований домен отримує юридичну та корисну інформацію з заблокованих веб-сайтів. Більшість систем з клієнт-серверною архітектурою. Наприклад VPN, який працює на мережевому рівні використовує практичні програми, включаючи OpenVPN і OpenConnect.

Ще проксі-сервери працюють на транспортному шарі, включаючи Shadowsocks і V2Ray. Клієнт-серверні системи мають інтуїтивно зрозумілу, просту структуру і просту в розгортанні та використанні, але недоліком є очевидно, що всі клієнти покладаються на існування серверів, як тільки сервери приймаються вниз по GFW, рішення полягає в тому, щоб купити і налаштувати інший екземпляр хмари, щоб настроїти сервер.

Ще одним недоліком є те, що якщо серверні вузли шкідливі ,конфіденційність і безпеку під час перегляду важко гарантувати. Іншим типом систем є анонімні системи зв'язку, які використовують маршрутизації цибулі для побудови безпечного тунелю передачі даних між ними, наприклад, проектом Tor.

У порівнянні з клієнто-серверними системами, такого роду системи можуть уникнути однієї точки відмови, оскільки якщо GFW блокує одну мережеву схему, системи можуть перекидатися на інші тунелі [27].

Крім того, системи також можуть запобігти ідентифікації та забезпечити конфіденційність оригінаторів запиту, через вузли в ланцюзі діють як транспортний експедитор, і мають обмежений спектр знань, таких як знати тотожності попередніх і наступних вузлів.

Вузли виходу, у мережі Tor, розшифрує повідомлення та діятиме як клієнт для побудови зв'язку між цільовими веб-сайтами. Розташування та кількість вузлів виходу ключовий момент цих систем.

Для досягнення високої продуктивності, такої як низька затримка і висока пропускна здатність, системі необхідно збалансувати кількість споживачів і постачальників послуг. Мости та виїзні вузли Tor встановлюються волонтерами, і поки що у нього немає рішень для цієї проблеми.

1.3 Цензура в мережі інтернет

Інтернет-цензура може відбуватися в стеку TCP/IP як кінцевих точок, так і в шлях між ними. Таким чином, цензура може бути класифікована на цензуру на стороні клієнта, цензура на стороні сервера, цензура на шляху та цензура в шляху.

Цензура на стороні клієнта означає, що користувачі можуть отримати лише обмежений обсяг ресурсів через вбудованих функціях у програмах цензури, таких як мережеві фільтри.

Ці програми (наприклад, TOM-Skype, Sina-UC, LINE, Green Dam) можуть порушити нормальні зв'язки багатьма способами, наприклад, показуючи неправильні результати, коли користувач запускає деякі чутливі слова або URL-адреси в чорному списку або забороняє установку програмного забезпечення.

Загальні методи вимірювання цензури на стороні клієнта включають побудову конфіденційних списків ключових слів або зворотню інженерію програми.

У реалізації на стороні сервера правила виконуються на віддаленому сервері, який можна вибірково видаляти, приховувати або блокувати доступ до певного вмісту відповідно до правила. Подібно до виявлення цензури на стороні клієнта, дослідники намагаються створити список вмісту, як зразки тестування.

Виявили, що Weibo, китайський Twitter, видаляв найбільш чутливі повідомлення протягом дня за допомогою ретроспективного механізму на основі ключових слів. Аналогічно, Wechat, домінуючий додаток для чату в Китаї, також застосовував фільтрацію ключових слів і URL-адрес в особистому чаті, а також груповому чаті. Користувач отримуватиме попередження, коли спрацьовують чутливі слова.

Незважаючи на клієнтську та серверну цензуру, цензори також можуть взяти контроль через канал зв'язку, наприклад, керування кількома маршрутизаторами всередині мережі і ввести неправдиву інформацію про маршрутизацію, щоб ці маршрутизатори могли відкинути або переслати пакети в неправильні місця.

Цензор розгортає деякі пристрої, крім міжнародного шлюзу, для моніторингу або щоб переривання мережевого трафіку, з метою дослідження. Пристрої, такі як NIDS (мережева система виявлення вторгнень) може виконувати величезний обсяг аналізу, роблячи копії пакетів у мережевому каналі зв’язку.

Ці системи мають можливості читати всередині, а також вводити додаткову інформацію в пакети.

Дослідження зосереджено лише на поведінці супротивників, які з'явилися між клієнтом і сервером, точніше, GFW, система фільтрації національного рівня Китаю. GFW використовує різні технічні методи для порушення зв’язків між ними, цензурні регіони та регіони без цензури, включаючи фільтрацію, втручання, підробку та спостереження.

Існує дві основні категорії цензури: цензура на шляху та цензура в шляху. Загальна модель цензури показана на рисунку 1.1.

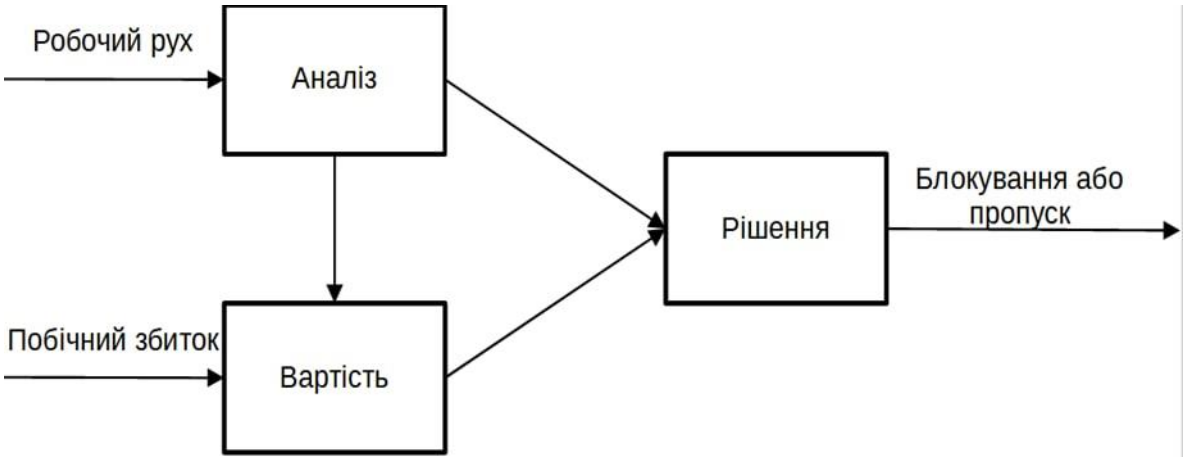


Рисунок 1.1 - Загальна модель цензури

GFW приймає мережевий трафік, такий як IP-пакети, а потім аналізує їх за набором раніше визначених правил, результатом такого аналізу буде введення функції вартості та функції рішення. Нарешті, система вирішує, чи варто рух заблокувати або прямувати до пункту призначення.

Модель найбільш загальна така, яку можна застосувати до будь-яких систем цензури, GFW є більш складним, але це важко зробити через непрозорість інфраструктури цензури. Одним із способів вирішити цю проблему є читання літератури або публікацій, написаних GFW дизайнерів або попередньо визначити набір зразків тестування, щоб визначити, який тип методів Використовує GFW.

1.4 Внутрішня цензура

Внутрішня цензура покладається на аналіз на рівні потоку[28], який базується на 3-х кортежі IP-адреси джерела, IP-адреси призначення та протоколу. Це ефективно і на багато простіше в порівнянні з аналізом на рівні потоку, оскільки він розглядає лише заголовки IP і перевіряє, чи є IP-адреса призначена законною.

Блокування IP-адреси є одним із найперших методів, розгорнутих GFW. Це відбувається на мережевому рівні.

Тривіальне рішення - підтримувати список контролю доступу (ACL) заблокованих веб-сайтів на шлюзових маршрутизаторах. Коли пакети проходять через маршрутизатор, він спочатку порівнює IP-адреси призначення з ACL, а потім виконує серію дій, тобто або пересилає пакети, або мовчки відкидає їх. Хоча це рішення просте і зрозуміле, воно не ефективне для магістральних мереж які повинні обробляти величезний обсяг мережевого трафіку, особливо коли він там це велика кількість IP-адрес в ACL, оскільки ця операція займає додатковий час щоб відповідати IP-адресі призначення. Крім того, оскільки маршрутизатори мають відносно невеликий розмір пам’яті, вони не підходять для зберігання величезної кількості IP-адрес.

Вони можуть запропонувати полегшений метод керування на основі маршрутизації розповсюдження для великомасштабної мережі, яка також називається Border Gateway Protocol (BGP) .

У цьому методі GFW вводить неправильну статичну інформацію про маршрутизацію наприклад, вручну налаштований запис маршрутизації до маршрутизатора шлюзу, потім цей маршрутизатор однорангові з усіма маршрутизаторами шлюзу з неправильною інформацією про маршрутизацію за допомогою BGP і перерозподіл Open Shortest Path First (OSPF). Отже, неправильна інформація поширюється на всі маршрутизатори всередині домену, що підлягає цензурі, що означає GFW захоплює весь трафік, намагаючись отримати доступ до заблокованих веб-сайтів в автономній системі, цей трафік буде втрачено під час транспортування або перенаправлено на аналізатори трафіку. Блокування IP має

два обмеження. По-перше, залежить від впливу блокування IP-адреси на точність інформації про маршрутизацію, і її необхідно ретельно підтримувати і оновлювати. Інше обмеження полягає в тому, що користувачі можуть легко обійти цензуру використання проксі-серверів за межами цензурного домену.

До чорного списку IP можна поєднувати блокування портів, наприклад, вибіркове закриття служб на сервері, перевіривши на фільтрацію пакетів із певними портами, такими як порт 22 для Secure Shell (SSH), 80 для протоколу передачі гіпертексту (HTTP), 443 для протоколу передачі гіпертексту Безпечний (HTTPS).

1.5 Подорожня цензура

Оскільки блокування IP-адресів можна легко вирішити за допомогою методів проксі, а підтримання точного списку ACL коштує дуже багато. Крім того, GFW не здатний до виявлення або внесення всіх IP-адрес у чорний список. Через ці недоліки, дизайнери починають шукати альтернативи для розпізнавання мережевого потоку, який намагається обійти цензуру.

Хорошим рішенням є встановлення аналізаторів трафіку мережевий і блоковий чутливий потік відповідно до вмісту в транспортному рівні і прикладний шар. Система In-Path не підходить для виконання важкого аналізу. Інакше це призведе до затримок руху та заторів. Натомість GFW починає використовувати систему on-path, яка може обробляти надзвичайно високу пропускну здатність.

GFW захоплює всі IP-пакети з мережі до своїх аналізаторів трафіку бічного каналу, а потім збирає ці пакети відповідно до порядкового номера всередині TCP. заголовки для виконання DPI на прикладному рівні. GFW має свою реалізацію стеку TCP/IP, щоб повторно синхронізувати сегменти TCP для подальшого аналізу. На даний момент це може розрізняти багато популярних протоколів, таких як HTTP і система доменних імен (DNS). Вони мають кілька механізмів аналізу трафіку на рівні потоку, на основі яких на 5-ти кортежі з IP-адресою джерела, портом джерела, IP-адресою призначення, пунктом призначення порт і протокол [29].

Класифікація на основі портів Аналіз трафіку полягає в перевірці номера порту всередині заголовка TCP (наприклад, порт 80 означає трафік HTTP). Цей метод дуже підходить для класифікації масивного мережевого трафіку шляхом простого відображення сервісу на добре відомому порту число, але точність може бути дуже низькою, щоб запобігти атаці служб або програм, системні адміністратори відкривають свою службу за іншим номером порту замість загальновідомого номера за замовчуванням, деякі програми використовують динамічне виділення портів . Мур показав це за допомогою добре відомих методів класифікації номерів портів, велика кількість мережевого трафіку невідома, а невелика кількість мережевого трафіку неправильно класифіковано [30].

Глибока перевірка пакетів – це спосіб перевірки потоків мережевого трафіку через певні контрольні пункти, щоб приймати рішення в режимі реального часу. За допомогою цієї технології GFW може прослуховувати конфіденційні ключові слова всередині пакету або підозрілого мережевого трафіку. На відміну від традиційного мережевого аналізу, який перевіряє лише структуровану інформацію в заголовках пакетів, наприклад IP, TCP, Заголовки UDP, DPI переглядає вміст пакетів, а саме програми. Поширеним способом отримання мережевих пакетів для DPI є використання дзеркального відображення портів, також відомого як span-порт, і оптичного розгалужувача.

Кількість точок на дюйм (DPI) може розрізнити купу протоколів прикладного рівня, таких як одноранговий (P2P), VoIP, потокове аудіо/відео, можна розглядати як чудовий спосіб ідентифікувати реальний мережевий трафік від інкапсуляції пакетів. Основні підходи DPI включаючи класифікацію на основі корисного навантаження та класифікацію на основі шаблонів:

- класифікація на основі корисного навантаження;
- класифікація на основі шаблонів;
- класифікація машинного навчання.

Шаблони трафіку можна ідентифікувати за корисним навантаженням всередині пакетів транспортного рівня. Цей метод може отримати високу точність, тоді як обчислювальна складність висока. Показали, що шляхом поєднання

класифікації на основі порту та класифікації на основі корисного навантаження можна збільшити правильно визначену трафік з приблизно 70% до майже 79%. Інші недоліки, що включають вторгнення в конфіденційність, і не повинні впливати на зашифрований мережевий трафік.

Класифікація на основі шаблонів, цей метод класифікує потік мережевого трафіку на основі поведінки хоста в транспортний шар.

У порівнянні з класифікацією на основі портів і корисного навантаження, він використовує лише шаблони підключення, а не порт і корисне навантаження. Караген розробив систематичну методологію для визначення прикладного рівня [31].

Потік P2P за допомогою шаблонів однорангового з'єднання, не покладаючись на навантаження пакетів. На основі цього дослідження вони пізніше придумали нову систему під назвою «BLINC» для аналізу моделі трафіку з соціального, функціонального та прикладного рівня [32]. Оцінка показує, що вони можуть класифікувати 80%-90% випадків трафік з точністю понад 95%. Перевага такого підходу в тому він може ідентифікувати мережевий трафік, коли корисне навантаження зашифровано. Крім того, це зменшує складність обчислень щодо класифікації на основі корисного навантаження. Однак цей підхід все ще перебуває на стадії експерименту через відсутність великого набору даних, наприклад, важко визначити новий протокол або протокол, винайдений самостійно.

Класифікація машинного навчання, методи згадані вище, є деякими простими правилами фільтрації, які застосовуються до фільтрації підозрілих потоків мережевого трафіку, а машинне навчання розроблено для класифікації значного обсягу трафіку в реальному часі. Він поєднує в собі кілька функцій і застосовує алгоритми навчання, включаючи навчання з наглядом, навчання без нагляду та напівкероване навчання для досягнення високої точності. Аналізатори мережі зазвичай збирають велику кількість пакетів на основі поведінки потоку мережевого трафіку та набору параметрів всередині мережі.

Пакети як функції або дискримінатори, класифікатори мережевого трафіку які можуть відрізняти один тип трафіку від іншого. Юань пропонує точну

класифікацію мережевого трафіку на основі методу SVM, а результат експерименту показує, що точність досягнута більше 97,17% [33].

Перевага класифікації мережевого трафіку на основі машинного навчання полягає в тому, що він працює як для незашифрованих, так і для зашифрованих даних, оскільки він не покладається на корисне навантаження пакетів. Мережева класифікація машинного навчання може досягти високої точності тільки тоді коли відповідні функції вибираються для великого набору даних, але обчислювальна складність досить значна для навчання точної моделі, вона не відразу працює для нових протоколів або ситуації, коли деякі біти всередині пакета змінюються в старих мережевих протоколах.

У деяких особливих випадках GFW повинен підтвердити, що сервер працює, заборонити протоколи. У цій ситуації GFW встановить з’єднання з підозрілим проксі-сервером, як звичайний клієнт. Якщо сервер відповідає правильно, і вони успішно встановлюють з'єднання, GFW візьметься блокувати дію. Дослідження, проведене доводить, що GFW використовує активне зондування для виявлення мостів Tor [34].

Підхід до прийняття рішень Після виявлення чутливих потоків у транспортному шарі та прикладному рівні GFW застосовує кілька механізмів блокування, як правило, скидання TCP-з'єднання і DNS отруєння.

Скидання з’єднання TCP також відоме як атака скидання. В основному це працює на прикладному та транспортному рівнях. GFW захоплює всі IP пакети, які намагаються пройти через міжнародний шлюз, повторно синхронізує їх відповідно до порядкового номера.

Після виявлення конфіденційного вмісту в файлі, наприклад, чутливі ключові слова з’являлися в запитах HTTP GET, GFW впроваджує пакети RST(type-1) і RST/ACK(type2) як клієнту, так і серверу, щоб примусово припинити поточне з’єднання. Ван та інші оцінили недавню поведінку GFW, після кількох тестів вони виявили, що GFW створює TCB та залежить від пакетів SYN і SYN/ACK, а також від GFW в деяких випадках переходить у стан ре синхронізації [35]. Сю та інші

знайшли що фільтрація відбувається в прикордонних шлюзах і багатьох провінційних мережах [36].

Виконання скидання TCP-з'єднання є дорогим, оскільки воно потребує запису з'єднання стану в TCB, також відомий як блок керування передачею, ця структура даних відстежує різну інформацію про кожне з'єднання, наприклад, локальне і номери віддалених портів, буфери відправника та одержувача, протоколи та поточний сегменти.

Викрадання та отруєння DNS системи доменних імен – це ієрархічна децентралізована система імен яка підтримує каталог доменних імен і перекладає їх на адреси IP-оголошень.

Методики використовуються разом з блокуванням IP, оскільки одне доменне ім'я може відповідати кільком IP-серверам, що означає, що GFW може блокувати декілька адрес одночасно, але якщо користувачі введуть правильну IP-адресу домену, вони все ще могли отримати доступ до місць призначення. DNS-сервер нормально надається постачальником послуг інтернету (ISP).

Для їх запам'ятовування використовується кеш переклади за певний період. Тому він може негайно відповідати на запити DNS поки не закінчиться термін дії кешу. Коли DNS-сервер отримав фальшивий переклад його кеш, повертає неправильну IP-адресу і доставляє трафік іншому сервер.

Сервер GFW зазвичай вводить помилкові запити на локальні DNS-сервери та також виконує виявлення вторгнення в порт 53. Як тільки GFW виявляє чутливий запит, він негайно вводить у відповідь DNS фальшиву IP-адресу.

Оскільки підроблена адреса повертається набагато раніше, ніж законна, сервер ігнорує останнє, що надійшло, і лише пересилає користувачам помилкову відповідь (рисунок 1.2).

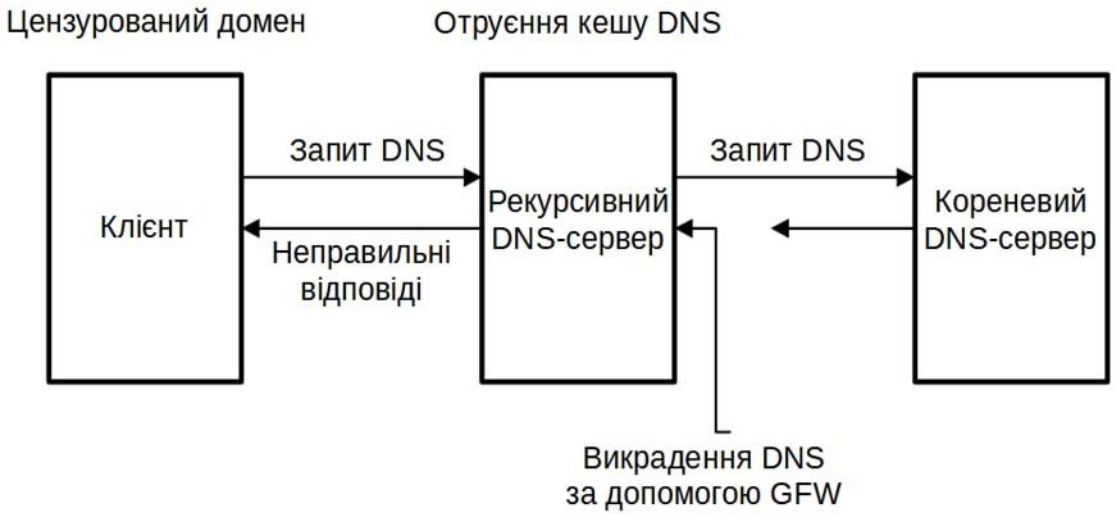


Рисунок 1.2 - Викрадання та отруєння DNS

Найкращіший підхід цензорів – це блокування пов’язаної адреси Інтернет-протоколу (IP) та номера порту призначення. Для того, щоб обійти цензуру, одним із найпопулярніших контрзаходів є надсилання запитів до проксі вузла і цей вузол надсилає заблоковану інформацію назад. Успіх систем протидії цензурі на основі проксі на практиці призвів до того, що цензори почали розгортати передові механізми глибокої перевірки пакетів (DPI) який може ідентифікувати трафік на основі інформації на прикладному рівні, а також поведінка мережевого потоку між двома кінцевими точками. На основі технології DPI фільтрація за ключовими словами на основі URL-адреси може застосовуватися до маршрутизаторів, через які проходять пакети. Таким чином, цензори можуть аналізувати та змінювати мережевий трафік.

Глибока перевірки пакетів широко використовується в інтернет-цензурі, більшість стійких систем до цензури використовують методи шифрування та обфускації щоб ухилятися від дій DPI.

Ці підходи можна розділити на чотири основні категорії описані нижче:

- шифрування;
- рандомайзер;
- мімікрія;
- тунелювання.

1.6 Шифрування корисного користування

Звичайне шифрування може бути відносно хорошим обфускатором для запобігання мережі від цензорів. Шифрування в цьому випадку означає шифрування на основі семантики, тобто шифрується тільки корисне навантаження транспортного рівня. Наприклад, заголовок протоколу керування передачею (TCP) або протоколу дейтаграм користувача (UDP) не зашифровано, поки їх корисне навантаження зашифровано.

Хоча шифрування може запобігти дані зі зчитування третьою стороною, але тільки шифрування корисного навантаження не гарантує надійних з'єднань у всьому зв'язку, оскільки більшість шифрування вказують протокол, який вони використовували для передачі даних у заголовку звичайного тексту, це дає цензору можливість легко виявити відбитки пальців, які можна використовувати з мережевих пакетів.

Узгодження таких протоколів шифрування завжди мають деякі відмінності і шаблони, які також очевидні для цензури, він зазвичай передає конкретні повідомлення такі як методи шифрування або сертифікати сервера, які можна легко використовувати як відбитки пальців або підписи цензора. Відомий приклад: GFW може виявити конфіденційний мережевий потік HTTPS і виконати атаку скидання TCP.

1.7 Рандомайзер та мімікрія

На відміну від шифрування, одним з альтернативних підходів є рандомізація корисного навантаження та застосування потокового шифру до кожного байту. Такий підхід ускладнює ідентифікацію відбитків пальців, оскільки немає характерних закономірностей для спостереження.

Вінтер запропонував тонкий рівень протоколу над транспортним рівнем, що називається ScrambleSuit, цей протокол використовується для обфускації даних

програми та всього мережевого трафіку можна не відрізнити за допомогою псевдовипадкового корисного навантаження [37].

Така ідея також використовується в Pluggable, Tor, Transport, Obfsproxy, включаючи obfs2, obfs3, і obfs4. obfs2 — це протоколи рівня обфускації всередині протоколу TCP.

Метою проектування є запобігання розпізнавання конкретного комунікаційного протоколу третьою стороною. Однак він не забезпечує аутентифікацію та цілісність даних.

Протоколи мають дві фази:

- встановлення ключів;

- обмін супер шифрованим трафіком.

У пізнішій версії obfs3 пропонує захист від пасивних глибоких пакетів інспекції. Цей метод неможливо виявити без запуску пробної атаки проти його рукостискання. Потім Tor пропонує протокол obfs4, який є комбінацією протоколу ScrambleSuit, техніки elligator2 та протоколу ntor призначений для захисту від цензорів глибокої перевірки пакетів який досліджує відбитки пальців, а також корисне навантаження пакетів [38].

Протокол не відрізняється від цензорів шляхом, побудови випадкових пакетів корисного навантаження. Хоча видимий відбиток пальця не може бути виявлений, «відбиток пальця немає» сам по собі стає функцією, що означає, що цензори можуть зловживати відмінністю між простими текстовими заголовками звичайних протоколів і рандомізацією. В результаті цього типу додатків можна легко виявити простим евристики, наприклад, перевірки довжини та тести на основі ентропії.

Системи, стійкі до цензури, засновані на мімікрії, намагаються створювати корисні пакети виглядають як допустимі пакети. Могаддам запропонував метод SkypeMorph, призначений для підключення транспорту Tor щоб замаскувати мережевий трафік такий як відео дзвінок Skype між двома кінцевими точками [39]. Пізніше Вайнберг розробив дизайн StegoTorus, який підключає транспортні вилки Tor Obfsproxy [40].

Він розділив потоки Tor між кількома з’єднаннями та вставив трафік потоки, які виглядають як HTML, JavaScript або PDF. Ван запропонував новий фреймворк для стійкого до цензури веб-перегляду під назвою CensorSpoofer, він приховує вміст запиту вище по потоку, наприклад URL-адреси в миттєвих повідомленнях та електронних листах і завантажує веб-вміст із цільових серверів, виконуючи спуфінг IP [41].

Колаж дизайну, який дозволяє користувачам вбудовувати повідомлення із запитами всередині створеного користувачами вмісту. Наприклад, сайти для обміну фотографіями. Хоча мімікрії додатків намагаються зробити пакети корисних даних схожими на стандартні протоколи, вони не використовують реалізацію стандартного протоколу, що означає обфускатор відрізняється від фактичних протоколів, які він намагається імітувати, і може бути легко виявлений супротивником.

Наприклад, дослідники виявили, що SkypeMorph і StegoTorus зазнає невдачі навіть проти найслабшого цензора [42]. Тоді Houmansadr показав що неспостережливість шляхом наслідування є принципово хибним і частковим підходом наслідування гірше, ніж відсутність наслідування взагалі. Вони пропонують альтернативу, використовуючи оригінальний протокол замість імітації протоколу.

1.8 Постановка задачі

Метою роботи є проведення дослідження для вимірювання поведінки GFW. Крім того, модель загрози GFW створюється після вимірювання її реальної поведінки, розробити та оцінити систематичний підхід до ухилення від цензури, мульти-проксі, одноранговий проксі-сервер на основі системи доступу до Інтернету та конфіденційності.

У порівнянні з поточними проксі-серверами або VPN-системами на базі клієнтського сервера, концепції дизайну мульти-проксування, експлуатують ресурси та групи вузлів. Воно має три основні функціональні можливості.

По-перше, вона спрямована на забезпечення основних послуг обходу, і трафіка, може бути перенаправлена через різні вузли.

По-друге, система запроваджує токен-економіку, а це означає, що всі учасники в мережі повинні платити або заробляти токени відповідно. Таким чином, мульти-проксі може збалансувати кількість споживачів та обхід провайдера. У порівнянні з безкоштовним використанням Tor мостів і вихідних вузлів, мульти-проксі може створити надійне і здорове середовище для користувачів для досягнення низької затримки і високої пропускної здатності. Система також використовує маршрутизацію цибулі для конфіденційності розгляду ініціаторів запиту.

Мульти-проксі вибирає групу вузлів кандидатів для побудови тунелів для передачі даних між творцями та цільових веб-сайтів. Питання дослідження наведені в розділі 2, щоб мати базове розуміння методів цензури, цей розділ також спрямований на надання деякого фону системи опору GFW та цензури (CRS), включаючи різні методи та архітектуру з літератури.

Розділ 3 дасть комплексне вимірювання GFW і розгляне, як цензура в Китаї працює, наприклад, феноменальний і операційний принцип, що лежить в основі, різні методи блокування, модель загрози GFW. Розділ 4 полягає в розробці та впровадженні нової системи під назвою мульти-проксування, щоб обійти цензуру.

1.9 Висновки

В даному розділі було описано технології, цензурування в інтернеті та їх наслідки. Були розглянуті сучасні методи блокування вмісту та боротьби з цензурою , яким чином можна розробити ефективний спосіб ухилення від цензури та були розглянуті рекомендації щодо обходу. Перші два запитання спрямовані на дослідження сучасних технологій блокування контенту та анти цензурних прийомів. Наступні два запитання були зосереджені на розробці та оцінці ефективних систем цензури.

2 МОДЕЛЬ ПРОЦЕСУ ВИЯВЛЕННЯ З'ЄДНАННЯ В КОМП'ЮТЕРНИХ СИСТЕМАХ

2.1 Експериментальне встановлення

Чорний список GFW з відкритим кодом, розміщений на Github вибирається для проектування а обґрунтований вимірювальний експеримент і побудова моделі загрози GFW.

Цей список містить майже всі заблоковані веб-сайти, про які повідомляють користувачі мережі, він часто оновлюється та ретельно підтримується групою волонтерів 22 лютого 2009 р.

Через непрозорість GFW файл не може повністю записати всі заблоковані доменні імена, але більшість із них.

Загалом, волонтери мають доступ до веб-сайтів вручну або за допомогою скриптів для перевірки доступності перед створенням проблем у сховищі. Проблеми ретельно перевіряються адміністраторами, щоб перевірити правильність.

Ці доменні імена в списку можуть бути заблоковані GFW за допомогою різних методів, наприклад, блокування IP-адреси або методів виявлення вторгнень, включаючи скидання TCP-з'єднання та викрадення DNS/отруєння кешу DNS.

На рисунку 2.1 показаний конвеєр вимірювань для обробки списку GFW. Хоча використовується остання версія, оновлена 30 червня 2018 року, вона може містити деяка застаріла інформація.

Наприклад, деякі сервери можуть змінити свою адресу записів на DNS-сервері, або GFW більше не блокує частину доменів у період.

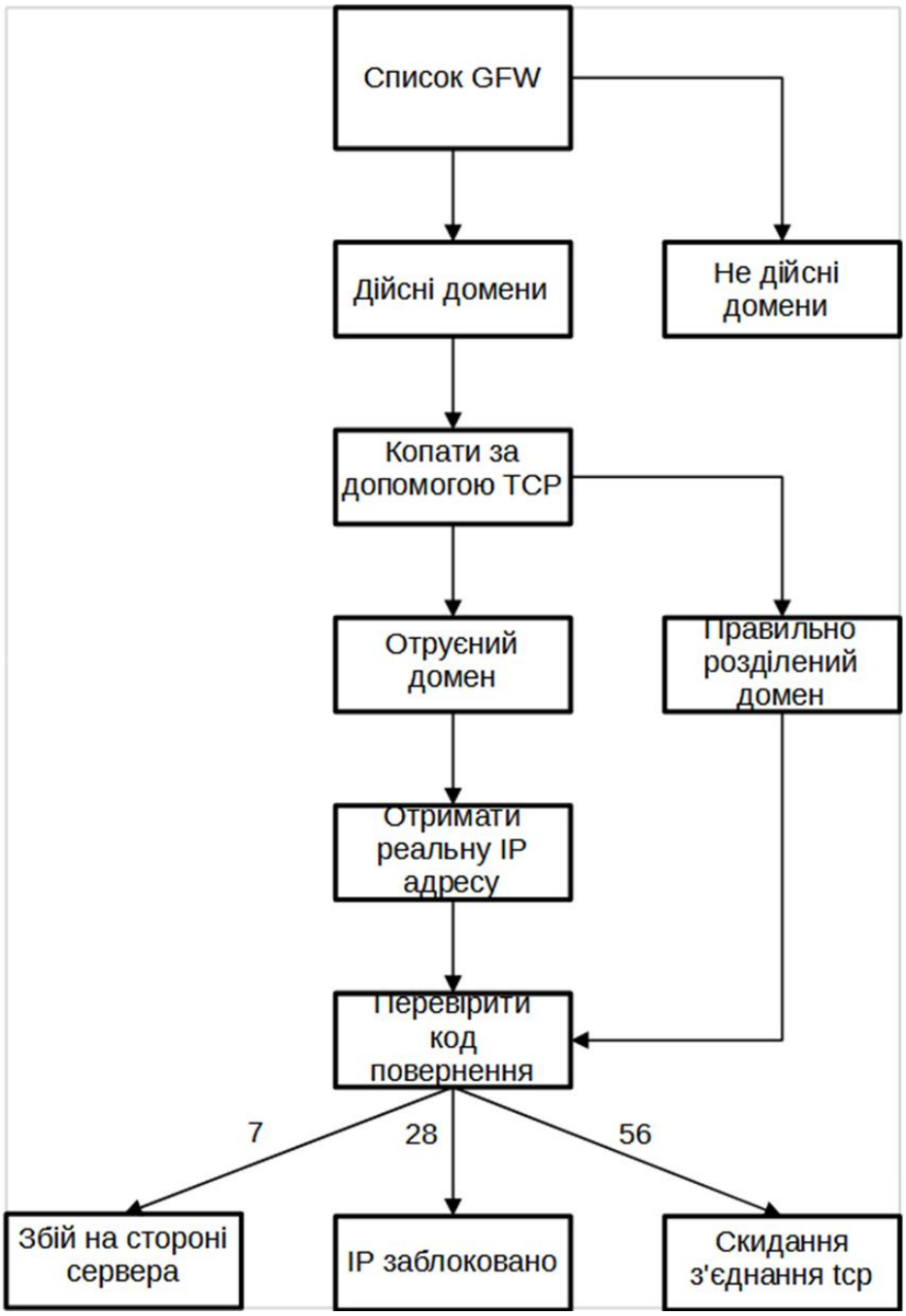


Рисунок 2.1 - Експериментальний конвеєр для класифікації різних першопричин

На першому кроці список фільтрується за URL-адресами, IP-адресами та коментарями, і всього 6170 унікальних повних доменних імен (FQDN).

На наступному кроці видаляються дійсні доменні імена. Нарешті, 4265 дійсних доменів використовуються для категоризації.

Наступним кроком є з'ясування поведінки GFW та отримання розподілу веб-сайтів, на які впливає блокування IP-адреси, скидання TCP-з’єднання та

викрадення/кеш DNS отруєння. Що стосується оцінки частки блокування IP-адреси та скидання TCP-з'єднання отруєних доменних імен, список правильних IP-адрес виводиться за межі цензурованого домену, а потім вони класифікуються за допомогою інструмента командного рядка curl.

Деякі сервери використовують віртуальний хост на основі імені, що означає, що існує кілька доменів імена на одній IP-адресі. Таким чином, поле хост будується всередині HTTP-заголовка, щоб переконатися, що клієнт може правильно підключитися до веб-сайту.

Щоб розрізнити всі ці випадки, тобто відсоток постраждалих веб-сайтів за допомогою цих методів окремо для створення запитів і аналізу трафіку використовується безліч інструментів командного рядка, включаючи curl, dig, traceroute і tcpdum, curl використовується для створення запитів HTTP/HTTPS через протокол TCP у вимірюваннях блокування IP і скидання TCP.

Програма Traceroute використовується для відправки Пакета Internet Control Message Protocol (ICMP) для відстеження шляху між клієнтом і сервером, dig використовується в експерименті DNS, щоб отримати роздільну здатність адреси доменні імена, tcpdump використовується для аналізу трафіку на стороні клієнта, щоб знайти шаблони механізми роботи GFW.

2.2 Блокування IP

Блокування IP є одним із перших методів, прийнятих GFW через його простоту і безпосередність. Він працює на мережевому рівні. Після того, як пакет проходить через інтернет магістраль і готова до відправки в мережу інтернет, система блокування на шляху відповідає IP-адресі призначення в заголовку IP із своїм чорним списком.

Якщо збігається, то пакети скидаються або фільтруються. Немає способу уникнути механізму блокування IP якщо поле адреси призначення не заповнено розблокованою IP-адресою.

Під час аналізу трафіку програма постійно надсилає ті самі пакети, оскільки немає відповіді з іншого боку.

Наприклад, хост надсилає SYN пакетів багаторазово, оскільки він не отримує відповіді в TCP-з'єднанні малюнок 2.2 показаний механізм блокування IP.

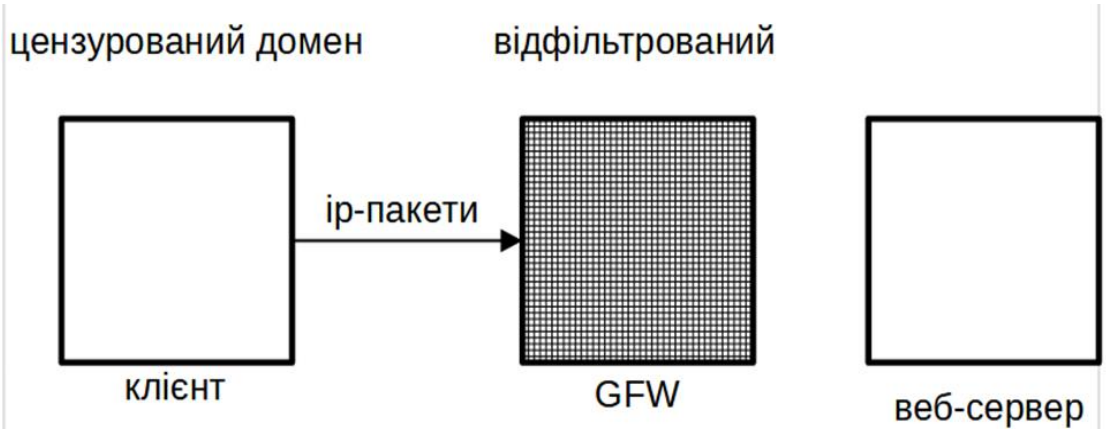


Рисунок 2.2 - Блокування IP

Причина в тому, що IP-пакети загубилися в середині. Пристрої на шляху відкидають або фільтрують і ніколи не потрапляють на сервер. Тому сервер не дасть відповіді клієнту.

Перелік показує шлях трасування між доменом без цензури та однією IP-адресою з www.google.com. Рисунки 2.3 та 2.4 показують результат трасування, запитаний із цензурованого домену.

У першому фрагменті пакети успішно досягають місця призначення, яке означає, що серверна сторона не блокує пакети ICMP, поки з другої ult, пакети припиняють пересилання після того, як досягнуть маршрутизатора 119.147.219.241, який належить China Telecom з AS4816, однієї з китайських магістральних мереж.

```
$ traceroute -n 216.58.211.100
traceroute to 216.58.211.100 (216.58.211.100), 64 hops max, 52 byte packets
1  145.94.160.2  2.483 ms  1.272 ms  7.041 ms
2  10.200.23.58  1.112 ms  1.121 ms  1.156 ms
3  10.200.246.121  1.269 ms  1.160 ms  1.147 ms
4  10.200.246.5  1.439 ms  1.153 ms  1.181 ms
5  10.200.24.5  1.385 ms  1.283 ms  1.287 ms
6  145.145.26.97  2.792 ms  4.937 ms  4.145 ms
7  145.145.166.86  3.026 ms  2.910 ms  2.788 ms
8  108.170.241.161  4.056 ms
   108.170.241.129  2.997 ms
   108.170.241.161  4.022 ms
9  108.170.237.45  3.070 ms  3.379 ms  2.808 ms
10  216.58.211.100  2.804 ms  3.064 ms  2.965 ms
```

Рисунок 2.3 – Результат тестування

```
$ traceroute -n 216.58.211.100
traceroute to 216.58.211.100 (216.58.211.100), 30 hops max, 60 byte packets
1  * * *
2  11.212.252.65  5.862 ms  6.276 ms  6.484 ms
3  11.218.131.97  5.384 ms 11.218.131.173  4.878 ms 11.218.131.149  5.013 ms
4  11.218.131.246  1.081 ms  1.116 ms 11.218.131.234  1.078 ms
5  119.38.212.110  2.434 ms 119.38.212.102  1.416 ms 119.38.212.114  2.260 ms
6  116.251.113.137  1.381 ms 42.120.242.217  2.470 ms 116.251.113.141  1.451 ms
7  183.2.180.229  1.898 ms 183.2.180.93  1.909 ms 183.2.180.217  1.574 ms
8  183.2.182.117  2.552 ms 183.2.182.125  2.157 ms 183.2.182.129  2.589 ms
9  119.147.219.241  5.912 ms 119.147.222.13  10.056 ms 119.147.220.41  12.080 ms
10  * * *
11  * * *
...
```

Рисунок 2.4 - Трасировка із цензурованого домену

Отже GFW має список заблокованих IP-адрес і діапазон IP-адрес. Він занадто великий щоб ми відповідали всім IP-адресам. Інструмент командного рядка ping призначений для перевірки доступності між двома вузлами на мережевому рівні, але враховуючи, що ping досягає свого створюючи пакети ICMP, і деякі брандмауери на стороні сервера можуть блокувати доступ до пакетів ICMP з міркувань безпеки.

Ці сценарії розроблені за допомогою інструмента командного рядка curl. Він підтримує набір протоколів прикладного рівня, наприклад, FTP, HTTP і HTTPS.

В експерименті пакети потрібно відправити доступні порти. Доступний означає, що порти відкриті та не фільтруються брандмауером на стороні сервера.

Явно, порт 80 (HTTP) дозволений більшістю серверів. тому протокол HTTP використовується для запуску GFW. завивка є зручним інструментом, оскільки він може повертати як результати, так і код статусу, наприклад, код статусу дорівнює 0, якщо з’єднання успіх.

Перевірка іншого коду стану може виявити проблеми з мережевим підключенням. Наприклад, код статусу 7 і 28 вказує, що хтось на шляху перехоплює нормальний рух. Код стану 7 означає, що curl не вдалося встановити TCP-з'єднання до господаря.

Причиною можуть бути вказані неправильні номери портів, неправильні імена хостів, або на шляху існують брандмауери. Код статусу 28 означає, що з'єднання досягає зазначений період очікування. У скрипті максимальний час для підключення до веб-сайту десять секунд. Спосіб розрізнення кількох ситуацій у коді стану 7 - це виконання сценарії кілька разів у клієнтах як всередині, так і за межами цензурного домену. Результат показує, що є 498 дійсних IP-адрес доменів, заблокованих GFW.

2.3 Скидання TCP з’єднання

TCP (Transmission Control Protocol) є надійним протоколом зв'язку в транспортний шар. Він забезпечує кілька механізмів, таких як виявлення помилок, контроль потоку, контроль перевантажень і повторна передача. Однак він не призначений для безпечної передачі даних між двома вузлами, тобто протокол TCP не забезпечує конфіденційність корисного навантаження та аутентифікація ідентифікаторів двох вузлів.

Отже, TCP з'єднання можуть бути легко перехоплені або підроблені зловмисниками. Є різні підходи до переривання TCP-з'єднання, наприклад, атака скидання TCP, заливання SYN атаки та атаки перехоплення сеансу TCP. Простим способом розірвати існуюче з’єднання між клієнтом і сервером є скидання TCP-з’єднання (рис 2.5). Дві різні ситуації які можуть викликати скидання TCP-з'єднання, розглядаються в цьому розділі.

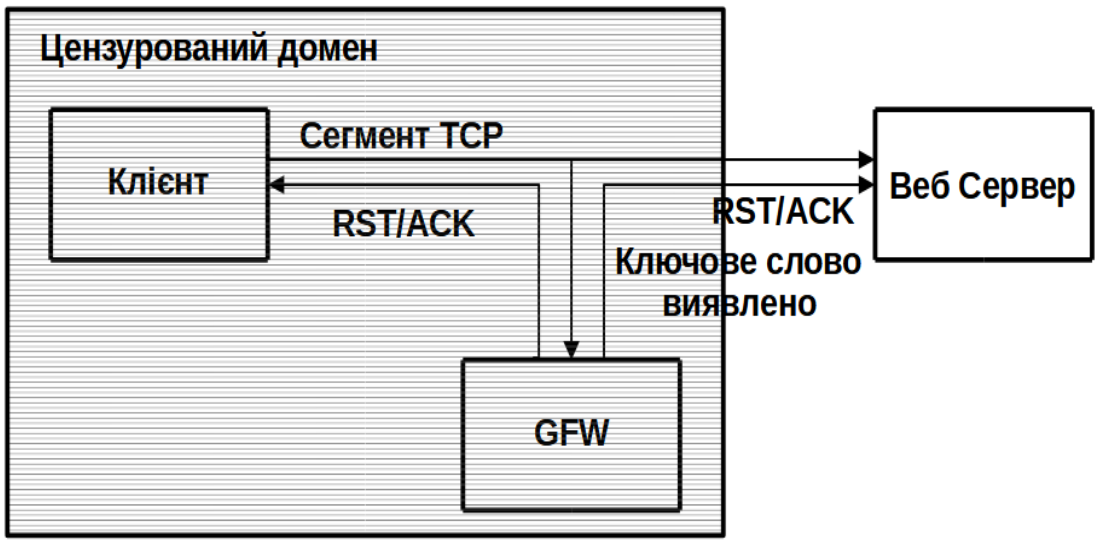


Рисунок 2.5 - Скидання TCP-з'єднання

Код статусу повернення curl 56 означає, що домен ініціює атаку скидання TCP.

Код виходу 35 з’являється , якщо є помилка підключення TLS/SSL.

Що стосується цільового веб-сайту, на нього впливає скидання TCP-з'єднання, у шаблоні використовується лише один IP вимірювальний експеримент, щоб мінімізувати вплив поля Time to Live (TTL) всередині заголовок IP.

У типовій ситуації один сервер завжди спілкується з клієнт з тим же значенням TTL.

Його можна легко виявити, якщо GFW порушує роботу з'єднання, оскільки GFW конструює пакети з випадковими значеннями TTL.

2.5 Виявлення ключових слів HTTP

Згідно з аналізом трафіку в рисунку 2.6, коли клієнт вперше підключається до заблокованого веб-сайту, з’єднання перехоплюється GFW після клієнта та сервера успішно виконано тристороннє рукостискання, і клієнт надсилає конфіденційний запит у цьому випадку mail-archive.com з'явився в полі заголовка HTTP Host.

Тоді GFW перериває з'єднання, надсилаючи один пакет RST, а потім три однакових RST/ACK пакети.

Справжнє підтвердження приходить пізніше і ігнорується клієнтом.

```
00:00:00.000000 IP (tos 0x0, ttl 64, id 38830, offset 0, flags [DF], proto TCP (6), length 60)
  192.168.1.117.49576 > 72.52.77.8.http: Flags [S], cksum 0x5788 (incorrect -> 0xd21f), seq
  ↪  2636432921, win 29200, options [mss 1460,sackOK,TS val 898126263 ecr 0,nop,wscale 7],
  ↪  length 0
00:00:00.167822 IP (tos 0x14, ttl 49, id 0, offset 0, flags [DF], proto TCP (6), length 60)
  72.52.77.8.http > 192.168.1.117.49576: Flags [S.], cksum 0xce77 (correct), seq 2021363056, ack
  ↪  2636432922, win 28960, options [mss 1460,sackOK,TS val 4038397413 ecr 898126263,nop,wscale
  ↪  7], length 0
00:00:00.167858 IP (tos 0x0, ttl 64, id 38831, offset 0, flags [DF], proto TCP (6), length 52)
  192.168.1.117.49576 > 72.52.77.8.http: Flags [.], cksum 0x5780 (incorrect -> 0x6cd7), ack 1,
  ↪  win 229, options [nop,nop,TS val 898126431 ecr 4038397413], length 0
00:00:00.167955 IP (tos 0x0, ttl 64, id 38832, offset 0, flags [DF], proto TCP (6), length 133)
  192.168.1.117.49576 > 72.52.77.8.http: Flags [P.], cksum 0x57d1 (incorrect -> 0x3c39), seq
  ↪  1:82, ack 1, win 229, options [nop,nop,TS val 898126431 ecr 4038397413], length 81: HTTP,
  ↪  length: 81
    HEAD / HTTP/1.1
    User-Agent: curl/7.29.0
    Host: mail-archive.com
    Accept: */*

00:00:00.173402 IP (tos 0x14, ttl 61, id 0, offset 0, flags [none], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.49576: Flags [R], cksum 0x5bff (correct), seq 2021363057, win
  ↪  13474, length 0
00:00:00.174417 IP (tos 0x14, ttl 97, id 6299, offset 0, flags [DF], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.49576: Flags [R.], cksum 0x19f9 (correct), seq 1, ack 82, win
  ↪  4872, length 0
00:00:00.174436 IP (tos 0x14, ttl 97, id 6299, offset 0, flags [DF], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.49576: Flags [R.], cksum 0x19f9 (correct), seq 1, ack 82, win
  ↪  4872, length 0
00:00:00.174458 IP (tos 0x14, ttl 97, id 6299, offset 0, flags [DF], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.49576: Flags [R.], cksum 0x19f9 (correct), seq 1, ack 82, win
  ↪  4872, length 0
00:00:00.335710 IP (tos 0x14, ttl 49, id 55300, offset 0, flags [DF], proto TCP (6), length 52)
  72.52.77.8.http > 192.168.1.117.49576: Flags [.], cksum 0x6be0 (correct), ack 82, win 227,
  ↪  options [nop,nop,TS val 4038397581 ecr 898126431], length 0
00:00:00.335734 IP (tos 0x14, ttl 64, id 18607, offset 0, flags [DF], proto TCP (6), length 40)
  192.168.1.117.49576 > 72.52.77.8.http: Flags [R], cksum 0x32fe (correct), seq 2636433003, win
  ↪  0, length 0
```

Рисунок 2.6 - Надсилання конфіденційного запиту

Рисунок 2.6 показує, як працює GFW, якщо клієнт знову підключається до веб-сайту через певний проміжок часу, після того як клієнт відправить пакет SYN, GFW негайно видає себе за сервер і надсилає пакет SYN/ACK, який змушує клієнта посилати послідовні пакети до GFW.

Тепер GFW може надсилати назад пакет RST/ACK і пакет RST, щоб повідомити клієнта про те що встановлення з’єднання припинено.

Пакет SYN/ACK від реального сервера надходить після того, як клієнт закрив з’єднання. Цей результат показує, що GFW зберігає стан, оскільки він може запам'ятати стани з'єднання клієнта і сервера за певний період часу.

Час зберігається від 90-х до 95-х, а коли час минув, GFW втрачає старий стан з'єднання і знову надсилає пакет скидання після того, як клієнт і заблокований веб-сайти закінчує тристороннє рукостискання.

У цьому списку показано, що GFW надсилає один пакет RST, а потім три пакети RST/ACK для завершення з'єднання між клієнтом і сервером після виявлення ключових слів у файлі Host і URL. Підроблені пакети, які надсилає GFW, пофарбовані в червоний колір.

2.6 Виявлення ключових слів DNS

Система доменних імен (DNS), як правило, є ієрархічною децентралізованою системою, яка зіставляє доменні імена з IP-адресами. Запит на DNS-сервер є першим кроком після цього користувачі вводять доменне ім’я в адресному рядку та натискають Enter.

Протокол DNS підтримує зв'язок як за протоколами TCP, так і за протоколами UDP, але не забезпечує конфіденційність при передачі даних, оскільки корисне навантаження надсилається у вигляді простого тексту. Завдяки цій функції GFW також містить модуль для запуску скидання з’єднання TCP, орієнтоване на протокол DNS.

Два віртуальних приватних сервера (VPS) у цензурному домені та без цензури розгорнуті для перевірки цього механізму. Вибираються два DNS-сервери в тому числі Google Public DNS, який знаходиться за межами домену, підданого цензурі, і Ali DNS розташовано всередині домену, підданого цензурі. Для цього використовується інструмент захоплення пакетів tcpdump моніторинг мережевого трафіку до порту 53, а інструмент пошуку DNS dig повинен створити Запити DNS

через TCP. Запит DNS надсилається через TCP до загальнодоступного DNS 8.8.8.8 Google, а вихід показує, де з'єднання скинуто. Мережевий трафік показує це після.

При звичайному тристоронньому рукостисканні хост видає запит DNS, який містить www.google.com з іншого боку, деякі пристрої в середині видають 8.8.8.8 і негайно повертає три пакети RST/ACK, щоб підтвердити попередній дані надсилаються, а потім повідомляється, що з'єднання з хостом скинуто, як показано в рисунку 2.7.

```
00:00:00.000000 IP (tos 0x0, ttl 64, id 15284, offset 0, flags [DF], proto TCP (6), length 60)
  192.168.1.117.52458 > 8.8.8.8.domain: Flags [S], cksum 0xd25b (incorrect -> 0xfd94), seq
  ↪  2016075502, win 29200, options [mss 1460,sackOK,TS val 4237944154 ecr 0,nop,wscale 7],
  ↪  length 0
00:00:00.023463 IP (tos 0x14, ttl 39, id 10393, offset 0, flags [none], proto TCP (6), length
↪  60)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [S.], cksum 0x6bfc (correct), seq 3781683001, ack
  ↪  2016075503, win 60192, options [mss 1380,sackOK,TS val 1040060966 ecr
  ↪  4237944154,nop,wscale 8], length 0
00:00:00.023505 IP (tos 0x0, ttl 64, id 15285, offset 0, flags [DF], proto TCP (6), length 52)
  192.168.1.117.52458 > 8.8.8.8.domain: Flags [.], cksum 0xd253 (incorrect -> 0x849e), ack 1,
  ↪  win 229, options [nop,nop,TS val 4237944177 ecr 1040060966], length 0
00:00:00.023721 IP (tos 0x0, ttl 64, id 15286, offset 0, flags [DF], proto TCP (6), length 97)
  192.168.1.117.52458 > 8.8.8.8.domain: Flags [P.], cksum 0xd280 (incorrect -> 0xa472), seq
  ↪  1:46, ack 1, win 229, options [nop,nop,TS val 4237944177 ecr 1040060966], length 4510227+
  ↪  [lau] A? www.google.com. (43)
00:00:00.030778 IP (tos 0x14, ttl 157, id 19375, offset 0, flags [DF], proto TCP (6), length 40)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [R.], cksum 0xe209 (correct), seq 1, ack 46, win
  ↪  3728, length 0
00:00:00.030781 IP (tos 0x14, ttl 157, id 19375, offset 0, flags [DF], proto TCP (6), length 40)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [R.], cksum 0xe209 (correct), seq 1, ack 46, win
  ↪  3728, length 0
00:00:00.030817 IP (tos 0x14, ttl 157, id 19375, offset 0, flags [DF], proto TCP (6), length 40)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [R.], cksum 0xe209 (correct), seq 1, ack 46, win
  ↪  3728, length 0
00:00:00.047164 IP (tos 0x14, ttl 39, id 10403, offset 0, flags [none], proto TCP (6), length
↪  52)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [.], cksum 0x8453 (correct), ack 46, win 236,
  ↪  options [nop,nop,TS val 1040060989 ecr 4237944177], length 0
00:00:00.047196 IP (tos 0x14, ttl 64, id 37215, offset 0, flags [DF], proto TCP (6), length 40)
  192.168.1.117.52458 > 8.8.8.8.domain: Flags [R], cksum 0xb94c (correct), seq 2016075548, win
  ↪  0, length 0
00:00:00.047757 IP (tos 0x14, ttl 39, id 10404, offset 0, flags [none], proto TCP (6), length
↪  113)
  8.8.8.8.domain > 192.168.1.117.52458: Flags [P.], cksum 0x9a74 (correct), seq 1:62, ack 46,
  ↪  win 236, options [nop,nop,TS val 1040060990 ecr 4237944177], length 6110227 1/0/1
  ↪  www.google.com. A 172.217.27.132 (59)
00:00:00.047764 IP (tos 0x14, ttl 64, id 37216, offset 0, flags [DF], proto TCP (6), length 40)
  192.168.1.117.52458 > 8.8.8.8.domain: Flags [R], cksum 0xb94c (correct), seq 2016075548, win
  ↪  0, length 0
```

Рисунок 2.7 - Приклад захоплення пакетів через TCP

TCP-з'єднання закривається хостом це означає, що у такого процесу немає прослуховування на попередньому порту.

Справжні пакети ACK і буфери, які містять правильну адресу розв'язання надійшла пізніше з 8.8.8.8, були проігноровані операційною системою, а хост відповідає, що пакетами RST — зазначив, що це один із способів відрізнити правильний DNS-сервер від підробленого – це перевірити час життя (TTL).

У цьому випадку правильний DNS-сервер має значення TTL з 39, тоді як imposter дорівнює 157, значення TTL з 64 означає, що пакети надсилаються з хоста.

Рисунок 2.8 показує, як GFW виконує атаку скидання. Це схоже на Скидання з'єднання TCP для виявлення ключових слів HTTP.

Різниця в тому, що HTTP виявлення має стан, оскільки GFW зберігає стан підключення у своїй базі даних для період часу.

Натомість для протоколу DNS, якщо запитується те саме доменне ім'я багаторазово, GFW завжди повертатиме три пакети RST/ACK для завершення підключення.

Це відображається для DNS через TCP, GFW не записуватиме TCP стан підключення.

```
00:00:03.732147 IP (tos 0x0, ttl 64, id 43057, offset 0, flags [DF], proto TCP (6), length 60)
  192.168.1.117.48636 > 72.52.77.8.http: Flags [S], cksum 0x5788 (incorrect -> 0xb244), seq
  ↪  2201214804, win 29200, options [mss 1460,sackOK,TS val 873953125 ecr 0,nop,wscale 7],
  ↪  length 0
00:00:03.739782 IP (tos 0x14, ttl 67, id 36317, offset 0, flags [DF], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.48636: Flags [S.], cksum 0x4ef6 (correct), seq 41278637, ack
  ↪  2201214805, win 2442, length 0
00:00:03.739808 IP (tos 0x0, ttl 64, id 43058, offset 0, flags [DF], proto TCP (6), length 40)
  192.168.1.117.48636 > 72.52.77.8.http: Flags [.], cksum 0x5774 (incorrect -> 0xe670), ack 1,
  ↪  win 29200, length 0
00:00:03.739961 IP (tos 0x0, ttl 64, id 43059, offset 0, flags [DF], proto TCP (6), length 121)
  192.168.1.117.48636 > 72.52.77.8.http: Flags [P.], cksum 0x57c5 (incorrect -> 0xb5d2), seq
  ↪  1:82, ack 1, win 29200, length 81: HTTP, length: 81
        HEAD / HTTP/1.1
        User-Agent: curl/7.29.0
        Host: mail-archive.com
        Accept: */*

00:00:03.745413 IP (tos 0x14, ttl 54, id 0, offset 0, flags [none], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.48636: Flags [R], cksum 0x76c0 (correct), seq 41278638, win
  ↪  17494, length 0
00:00:03.746091 IP (tos 0x14, ttl 68, id 35307, offset 0, flags [DF], proto TCP (6), length 40)
  72.52.77.8.http > 192.168.1.117.48636: Flags [R.], cksum 0x4ef2 (correct), seq 1, ack 1, win
  ↪  2443, length 0
00:00:03.893600 IP (tos 0x14, ttl 49, id 0, offset 0, flags [DF], proto TCP (6), length 60)
  72.52.77.8.http > 192.168.1.117.48636: Flags [S.], cksum 0x0e1b (correct), seq 521495779, ack
  ↪  2201214805, win 28960, options [mss 1460,sackOK,TS val 4014223819 ecr 873953125,nop,wscale
  ↪  7], length 0
00:00:03.893643 IP (tos 0x14, ttl 64, id 45705, offset 0, flags [DF], proto TCP (6), length 40)
  192.168.1.117.48636 > 72.52.77.8.http: Flags [R], cksum 0x37b1 (correct), seq 2201214805, win
  ↪  0, length 0
```

Рисунок 2.8 Приклад захоплення пакетів скидання TCP-з'єднання за протоколом HTTP протягом певного періоду

2.7 Модуль скидання TCP з'єднання

На малюнку 2.9 показано, як працює основний пристрій скидання TCP-з'єднання [43]. Він містить кілька модулів які підключаються почергово.

Рисунок 2.9 Модель пристрою скидання TCP-з'єднання

Модуль захоплення збирає всі пакети TCP із мережевого трафіку в реальному часі. Модуль оновлення оновлює форму з'єднання, що зберігається в базі даних. Тим часом модуль прийняття рішень приймає рішення чи слід виконувати правила запобігання вторгненню та використовувати модуль блокування для створення підроблених пакетів після обчислення перед відправленням цілям.

Поведінку GFW можна застосувати до цієї моделі. GFW сидить у шлях міжнародного шлюзу і робить копію мережевих пакетів. TCP Модуль захоплення пакетів відповідає за захоплення пакетів TCP після цього GFW збирає інформацію про підключення, включаючи вихідний контроль доступу до медіа (MAC), IP адреса, TCP-порт, а також MAC-адреса призначення, IP-адреса, TCP-порт в додаток до цього, GFW також збирає порядковий номер і номер підтвердження з пакета запиту/відповіді з обох сторін, щоб створити дійсний пакет.

Один раз клієнт підключається до сервера, інформація про TCP-з’єднання буде зібрана або оновлена, а потім збережена в базі даних протягом певного періоду. Запит на цензуровані ключові слова, надіслані клієнтом, може запустити модуль блокування, який надішле команда для створення підробленого пакета відповідно до попереднього оновленого з'єднання стан, наприклад, обчислення номерів SYN і ACK, перемикання адрес джерела та адреси призначення. Підроблені пакети RST або RST/ACK надсилаються як на клієнтську, так і на серверну сторону щоб розірвати зв'язок з обох сторін.

2.8 Викрадення та отруєння кешу DNS

У цьому розділі показано експеримент щодо того, як GFW виконує перебір DNS Протокол UDP, а також отруєння DNS, що є побічним ефектом, спричиненим DNS викрадення. Викрадання DNS використовує властивості протоколу UDP.

Відомо, що протокол UDP ненадійний, відправники надсилають пакети і ні вимагають визнання. Це дозволяє зловмисникам надсилати підроблені відповіді DNS. Клієнт приймає введені відповіді, оскільки вони надходять раніше, ніж законні відповіді.

Під час надсилання запитів DNS через протокол UDP вихід dig не виконується, неможна показати будь-яке повідомлення про помилку.

Замість цього він повертає записи типу A з адресою IPv4 показано в рисунку 2.10.

Однак ця адреса недійсна, оскільки не вказує на справжній веб-сайт. Підроблені пакети, які надсилає GFW, пофарбовані в червоний колір.

```
00:00:00.000000 IP (tos 0x0, ttl 64, id 26525, offset 0, flags [none], proto UDP (17), length
↪  71)
  192.168.1.117.49964 > 8.8.8.8.domain: 14538+ [1au] A? www.google.com. (43)
00:00:00.005060 IP (tos 0x14, ttl 44, id 0, offset 0, flags [none], proto UDP (17), length 76)
  8.8.8.8.domain > 192.168.1.117.49964: 14538 1/0/0 www.google.com. A 31.13.85.16 (48)
00:00:00.008538 IP (tos 0x14, ttl 103, id 29352, offset 0, flags [DF], proto UDP (17), length
↪  76)
  8.8.8.8.domain > 192.168.1.117.49964: 14538 1/0/0 www.google.com. A 31.13.80.17 (48)
00:00:00.022273 IP (tos 0x14, ttl 39, id 29402, offset 0, flags [none], proto UDP (17), length
↪  87)
  8.8.8.8.domain > 192.168.1.117.49964: 14538 1/0/1 www.google.com. A 216.58.197.100 (59)
```

Рисунок 2.10 - Приклад захоплення пакетів через UDP

Аналіз трафіку показує, що після того, як хост надсилає запит DNS який містить заблоковані ключові слова, деякі пристрої маскуються під цільовий DNS сервер і негайно надсилають два пакети з підробленими ресурсними записами. Через кілька мілісекунд правильна відповідь ігнорується dig, оскільки вона вже використовувала першу підроблена відповідь. Рисунок 2.11 показує, як працює викрадення DNS для TCP та UDP протокол.

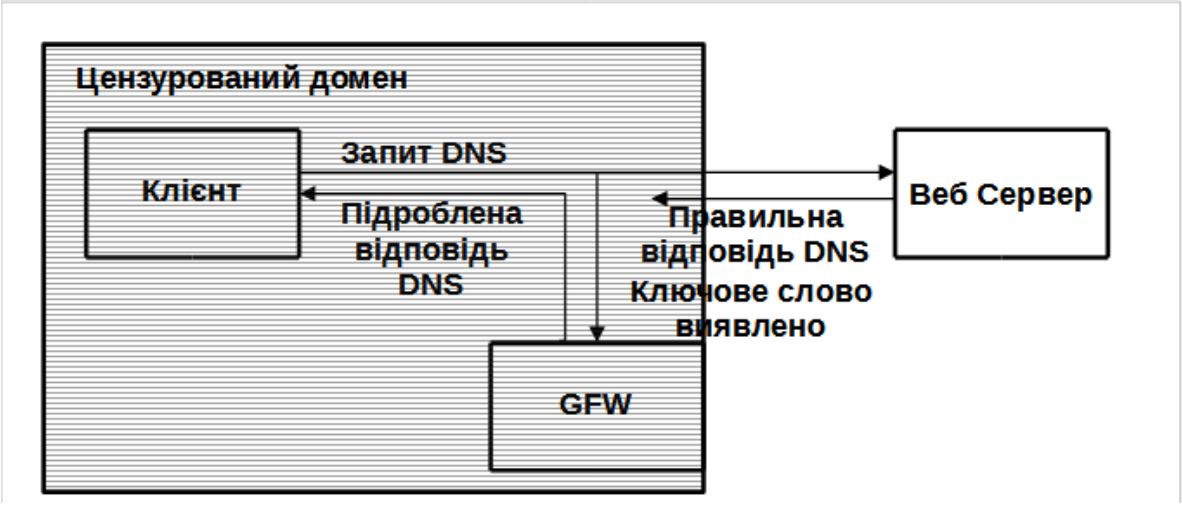


Рисунок 2.11 - Викрадання DNS

Загалом, отруєння кешу DNS є більш ефективним, ніж викрадення DNS, оскільки зловмисники атакують ціль лише один раз і дозволяють фальшивим ресурсним записам зберігатися в кеші DNS протягом одного або двох днів. Що стосується отруєння DNS, то змінено DNS-сервер на сервер Ali DNS (223.5.5.5 і

223.6.6.6) всередині домену, підданого цензурі, а потім запит DNS надсилається через TCP та UDP.

Згідно з виводом tcpdump, файл з'єднання не було перехоплено третьою стороною, але обидва ресурсні записи були неправильний. Причина в тому, що коли DNS-сервер Ali намагається виконати рекурсивну операцію Запит DNS (наприклад, надіслати запит на кореневий сервер DNS), деякі пристрої між ними перехоплювати з'єднання і створювати фальшиві пакети, що призводить до недійсних запис ресурсу, що зберігається в кеші DNS.

Оскільки GFW може читати вміст пакетів, атаку легко виконати, оскільки GFW не має вгадати 16 бітовий ідентифікатор у пакетах DNS. Він копіює ідентифікатор із запиту та використовує цільовий DNS адресу сервера як адресу джерела для підробленої відповіді, а потім надсилає її назад до користувач.

При виконанні тих самих експериментальних кроків за межами цензурного домену.

Поведінка GFW така ж, за винятком того, що всі повернені записи ресурсу є неправильними через отруєння кешу DNS. Згідно з результатами тесту, GFW використовує DPI для перевірки ключових слів всередині DNS пакетів на прикладному рівні. Якщо пакет містить ключові слова, які внесені до чорного списку GFW запустить конкретні правила.

GFW має різні правила фільтрації залежно від базового транспортного протоколу, надсилайте пакети RST/ACK для перехоплення TCP-з'єднання або фальшиву відповідь через UDP.

На основі отриманих результатів за межами цензурного домену правила фільтрації виглядають двох направленими для як вхідного, так і вихідного мережевого трафіку.

Відповідь, створена протоколом GFW через UDP, проста оскільки незалежно від того, який тип запитується, наприклад, сервер імен (NS), Start of Authority (SOA) або поштовий обмінник (MX), результатом вирішення завжди є запис типу A з підробленою IP-адресою. Ці отруєні доменні імена збираються, щоб отримати геолокацію.

Кожен із них 101 раз запитується до загальнодоступного DNS-сервера Google використовуючи протокол UDP, щоб запустити GFW і побачити шаблон, що стоїть за підроблені адреси. Зібрано 770 унікальних підроблених IP-адрес.

У таблиці 2.1 показано номер AS, компанії та наявність IP-адреси, пов'язані з цими отруєними IP-адресами.

Таблиця 2.1 – Отруєні IP-адреси

Номер AS	Власник	Країна	Приклад діапазону CIDR	Поява IP
AS32934	Facebook, Inc.	US	31.13.64.0/19 69.63.176.0/20 173.252.64.0/19 66.220.144.0/20	456(41.1%)
AS40824	WZ Communications Inc.	US	204.155.144.0/20 74.117.176.0/21 199.101.132.0/22	116(10.5%)
AS36351	SoftLayer Technologies Inc.	US	174.36.0.0/15 67.15.0.0/16 74.86.0.0/16 67.228.0.0/16 208.101.0.0/18	67(6.0%)
AS19679	Dropbox, Inc.	US	108.160.160.0/20	64(5.8%)
AS13414	Twitter Inc.	US	199.96.60.0/22 199.16.156.0/22 199.59.148.0/22	63(5.7%)
AS10753	Level 3 Parent, LLC	US	69.63.176.0/20 66.220.144.0/20	57(5.1%)

Кінець таблиці 2.1 – Отруєні IP-адреси

Номер AS	Власник	Країна	Приклад діапазону CIDR	Поява IP
AS15169	Google LLC	US	74.125.0.0/16 172.217.0.0/16 209.85.128.0/17 216.58.192.0/19	39(3.5%)
AS11172	Alestra, S. de R.L. de C.V.	MX	74.86.0.0/16 208.101.0.0/18 174.36.192.0/18 67.228.96.0/19 75.126.112.0/20	38(3.4%)
AS22773	Cox Communications Inc.	US	74.86.0.0/16 174.36.192.0/18 67.228.192.0/19	36(3.2%)
AS22561	CENTURYLINK, INC.	US	66.220.144.0/20	30(2.7%)
AS2914	NTT America, Inc.	US	128.242.0.0/16 168.143.0.0/16 124.40.32.0/19	30(2.7%)
AS19281	Quad9	US	216.58.192.0/19	13(1.2%)
AS13767	DataBank Holdings, Ltd.	US	72.233.0.0/19 72.232.168.0/22	7(0.6%)
AS22576	DataPipe, Inc.	US	72.233.0.0/19 72.232.168.0/22	7(0.6%)
AS31815	Media Temple, Inc.	US	64.13.192.0/18	7(0.6%)

Більшість фіктивних IP-адрес надходять зі списку блокування IP-адрес. Частка отруєних доменних імен наведена в таблиці 2.2

Таблиця 2.2 Частка отруєних доменних імен

Отруєні домени	1435
Правильно вирішені домени	4265
Недійсні домени	470
Всього	6170

Результати свідчать що якщо користувачі не вживають активних заходів для запобігання викрадення/отруєння DNS, ефект такий самий, як і блокування IP: пакети не пересилаються до місця призначення.

2.9 Висновки до розділу

Експерименти показують, що GFW має як пристрої на шляху, так і в шляху. Пристрої на шляху зосереджені на відкиданні або фільтрації пакетів, тоді як пристрої в шляху мають можливість спостерігати за пакетами на льоту, але не підробляють пакети безпосередньо.

Методи виявлення, включно з визначенням ключових слів, будь-який підозрілий мережевий трафік запускають попередньо визначені правила. Ці пристрої надалі створюватимуть деякі підроблені пакети для перехоплення звичайних з’єднань між клієнтом і сервером.

- GFW не спрямований на дешифрування даних між сервером і клієнтом. Метою GFW є перехоплення, припинення або блокування з'єднань.
- GFW може блокувати пакети на мережевому рівні своїми пристроями на шляху.
- GFW може контролювати стан з'єднання на прикладному рівні та переглядати пакети.

- GFW має кілька модулів виявлення вторгнень, і кожен з них має модуль правил запобігання вторгненню, наприклад, виявлення ключових слів у заголовках HTTP ініціює три пакети RST/ACK.

Результати вимірювань наведені в Таблиці 3.2 та Таблиці 3.3.

Загалом, скидання TCP-з'єднання займає значну частину в списку GFW, і існує 139 доменів, які постраждали як від отруєння DNS, так і від блокування IP, тоді як 923 домени страждають від отруєння DNS і скидання TCP-з'єднання.

Після порівняння результатів із 100 кращими глобальними сайтами Alexa, 36 із 100 веб-сайтів заблоковані IP-адресами, і більшість із них також страждають від отруєння DNS.

Ці веб-сайти, включаючи популярні пошукові системи, наприклад, Google, веб-сайти для обміну відео, наприклад, YouTube і Dailymotion, соціальні мережі, наприклад, Facebook і Twitter, а також хмарні сховища, такі як Dropbox.

На більшість непопулярних або невідомих веб-сайтів впливає лише отруєння DNS або скидання TCP-з'єднання.

Результат показує, що обхід GFW шляхом безпосередньої атаки на його пристрій на шляху є менш ефективним, наприклад, десинхронізація стану з'єднання в середині бази даних, оскільки більшість популярних веб-сайтів заблоковано на мережевому рівні.

Надійним способом є використання IP-адрес, які не можуть бути заблоковані GFW, для пересилання трафіку, наприклад, за допомогою віртуальної приватної мережі (VPN) для шифрування всього корисного IP-адреси або використання проксі-сервера за межами домену, підданого цензурі.

Відповідно до вимірювань під час скидання TCP-з'єднання, GFW використовуватиме DPI, щоб визначити, чи є ключові слова всередині пакетів, наприклад, поле Host у заголовках HTTP або поле Question у пакетах запитів DNS, або чи є шаблон для виявлення, наприклад, обмін сертифікатами в протоколі рукостискання TLS.

Таблиця 2.3 – Кількість доменів, на які впливає кожен механізм блокування

Отруєні домени		Правильно вирішені домени	
Блокування IP	Скидання з'єднання TCP	Блокування IP	Скидання з'єднання TCP
139	923	498	2219

Ці правила виявлення вторгнень вказують на те, що корисне навантаження має бути безглуздим для виявлення з GFW, одним із контрзаходів може бути шифрування або рандомізація, про що згадувалося в попередньому розділі.

Крім того, фаза рукостискання та фаза передачі даних не можуть містити чітких ознак.

Тунелювання даних через інші допустимі протоколи може бути способом вирішення цієї проблеми.

3 МЕТОД ПЕРЕАДРЕСАЦІЇ ТА ТОКЕНОВА ЕКОНОМІКА

3.1 Цілі дизайну

В останні роки були розроблені різні методи для обходу цензури, наприклад, тунель VPN і SSH. Shadowsocks, як зашифрований відкритий код проксі програмне забезпечення широко використовується в материковому Китаї для обходу інтернет цензури. З тих пір цей проект отримав 26753 зірки та 16626 розвилок на Github 2012.

Програмне забезпечення було портовано на різні платформи та операційні системи.

Ідея Shadowsocks полягає в тому, щоб розділити socks-proxy на дві частини: клієнт всередині домену, підданого цензурі, і сервер за межами цензурного домену. Різні для захисту даних на шляху між клієнтом підтримуються методи шифрування і сервер, наприклад, AES, RC4-MD5, Salsa20, Chacha20 тощо.

Тому Shadowsocks працює ефективно для окремих користувачів завдяки його легкому дизайну та гарній обфускації.

Звичайний варіант використання полягає в тому, що користувачі використовують свої клієнти Shadowsocks для підключення власних клієнтів Сервер Shadowsocks поза цензурним доменом.

Однак у системи є і деякі недоліки. Shadowsocks базується на архітектурі Client Server у своїй поточній реалізації, що означає, що вона підтримує лише один або кілька клієнтів, які використовують один сервер.

Це може призвести до серверної сторони єдиної точки відмови.

Незважаючи на те, що метод шифрування достатньо хороший, щоб обмежити протокол із технологією DPI GFW, GFW все ще має деякі можливості для інтеграції передових методів для виявлення мережевого трафіку, наприклад, машинне навчання та тест на ентропію [44-45].

Після виявлення спеціального мережевого трафіку сервер Shadowsocks може бути легко виявлений або заблокований GFW.

Користувачі не мають іншого вибору, крім налаштування своїх нових серверів Shadowsocks. Пропускна здатність може бути обмежена в цій архітектурі, оскільки зазвичай це може обмежувати один користувач використовувати пропускну здатність лише одного сервера з урахуванням вартості сервера.

Проект Tor побудований на розподіленій і анонімній мережі, розроблений новий pluggable transport meek у 2014 році [46].

Це підключення використовує домен fronting техніку як шар обфускації, щоб уникнути цензуру.

Заборонений мережевий трафік передається до пересилача, який називається сервером meek у розблокованій мережі CDN. Meek тепер є єдиним способом для клієнтів tor отримати доступ до заблокованих веб-сайтів з Китаю.

На рисунку 3.1 показана схема Tor, створена за допомогою meek. Перший стрибок – це міст з лагідним сервером під назвою «cymrubridge02». Цей міст є загальним для всіх браузери Tor, які використовують лагідний клієнт.

На рисунку 3.2 показано номер його клієнтів коливання за останній рік, і воно зросло з 5 000 до 10 000.

IP-адреси за сервером meek “meek.azureedge.net” наведено результати в таблиці 3.1 отримуються шляхом запиту доменного імені в глобальному масштабі.

Таблиця 3.1 IP-адреси сервера

Азія	117.18.232.200
Європа	152.199.19.160
США	72.21.81.200

Хоча це так сервер здатний обробляти до 10 000 клієнтів одночасно, сервер все одно може бути вузьким місцем виконання.

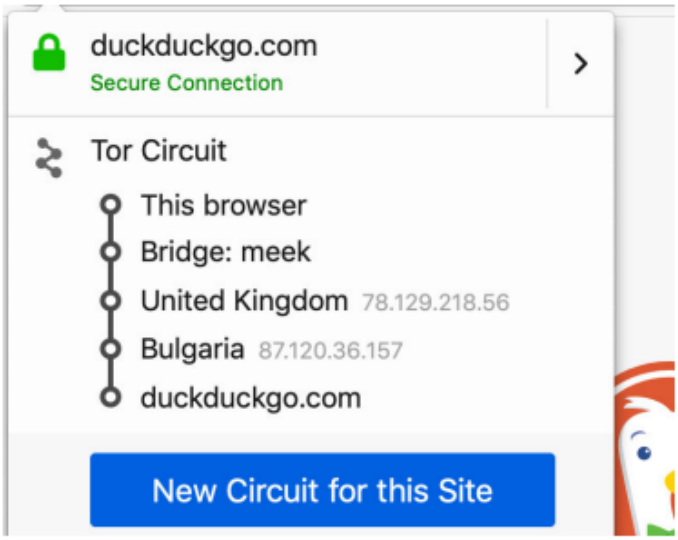


Рисунок 3.1 - Схема Tor із підключеною транспортною системою

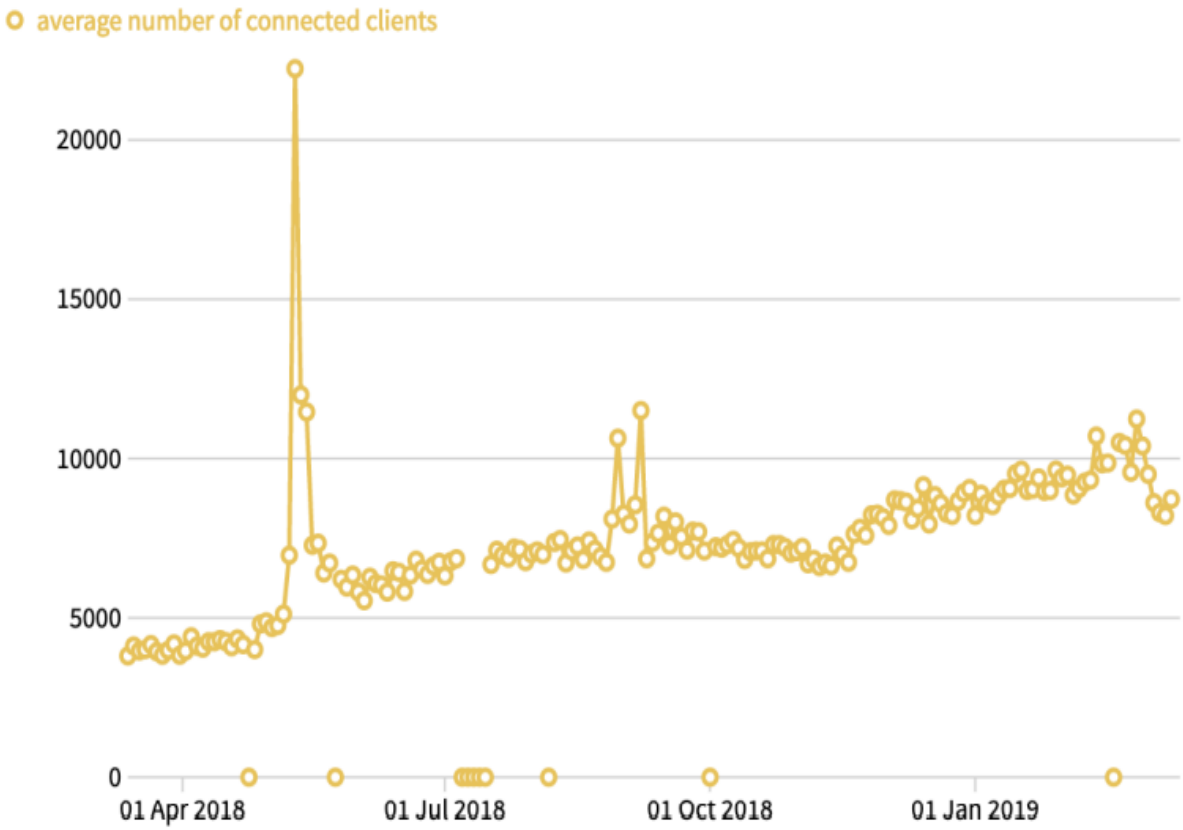


Рисунок 3.2 - Середня кількість підключених клієнтів

Мульти-проксі покладається на концепція однорангової мережі та забезпечує можливість користувачам всередині цензурованого домену використовувати

декілька проксі-серверів одночасно, а також пропонує маркер економіка як метод стимулювання збалансування споживачів і постачальників послуг.

3.2 Архітектура системи

На рисунку 3.3 описані основні компоненти та стек протоколів мульти-проксі система, що включає протокол SOCKS5 на межі, економіка маркерів заснована на протоколі Trustchain і технології Intel SGX, цибульна маршрутизація та однорангове виявлення протокол. Усі функції побудовані на основі однорангової системи.

Рисунок 3.3 - Архітектура системи обходу

На рисунку 3.4 показано основний шлях маршрутизації трафіку системи мульти-проксі.

У системі існують два типи проксі-вузлів, переважно клієнтські проксі-вузол і проксі-вузол сервера.

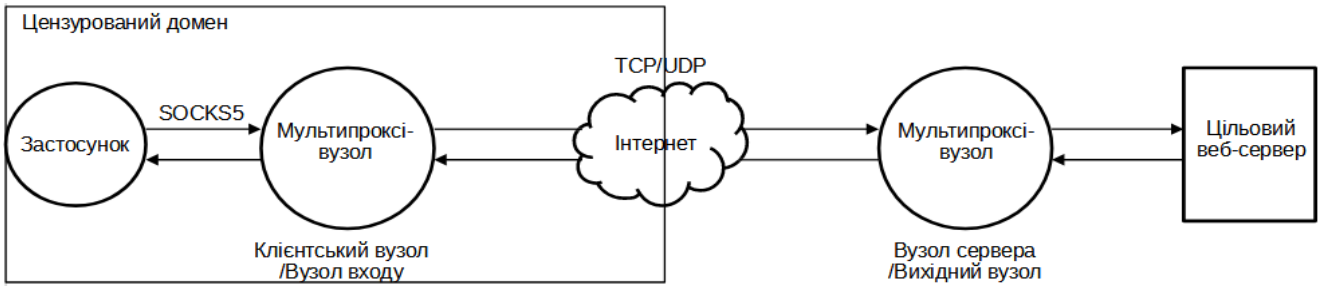


Рисунок 3.4 Шлях маршрутизації трафіку

Проксі-сервер клієнта може бути автором запиту, а також може діяти як вузол ретрансляції, який використовується для пересилання у мережі трафіку, коли схеми вбудовані в цибульну маршрутизацію.

Як система обходу, основна функціональність — це переадресація трафіку, тому клієнтський вузол повинен пересилати свій трафік до вузлів, які надають відповідну послугу обходу.

Для цього потрібен протокол пересилання мережевих даних між кількома додатками та системою мульти-проксі на одному хості.

Для передачі даних між вузлами повідомлення мають бути тунельовані та зашифровані, щоб уникнути використання різних методів аналізу мережевого трафіку GFW.

Також проксі-вузол сервера відомий як вихідний вузол, є критичним компонентом, оскільки він діє не лише як сервер отримувати дані від клієнтського проксі-сервера, але також виступає в якості клієнта для побудови з'єднання між цільовими веб-серверами для надання послуги обходу.

Клієнтський вузол знайде принаймні один вихідний вузол, щоб побудувати схеми та отримати доступ цільові веб-сервери.

Основні компоненти системи наведені в таблиці 3.2.

Таблиця 3.2 Компоненти системи мульти-проксі

Функціональність	Ціль	Вимоги
Переадресація трафіку	Надати базову послугу обходу	Протокол SOCKS5 Шифрування даних
Токенова економіка	Зробити систему надійною	Протокол Trustchain Intel SGX, SCONE
Повідомлення з кількома переходами	Захист конфіденційності для авторів запитів	Створення схем Тунелювання даних

Друга функціональність розроблена для кращого використання кількох вузлів. Для створення системи необхідно збалансувати кількість постачальників послуг і споживачів доступний і міцний.

Мульти-проксі також забезпечує анонімну маршрутизацію цибулі зв'язок через мережу.

Мульти-проксі також використовує цибульну маршрутизацію до захищати особистість і конфіденційність авторів запиту.

Повідомлення інкапсулюються в кожному вузлі, через який він проходить, таким чином, що проміжні вузли не можуть знати джерело та призначення повідомлень.

Цей механізм вимагає протоколу обміну ключами криптографії між ланцюгами.

3.3 Переадресація трафіку

Основною функцією мульти-проксі є переадресація трафіку.

Це досягається шляхом додавання проксі-шару поверх традиційної архітектури клієнт-сервер. Рівень проксі містить два основних компоненти: клієнтський вузол і вузол сервера.

Клієнтський вузол використовується для отримання даних з оригінальних програм, наприклад, браузерів.

Цей клієнтський вузол можна розглядати як розширення програми, оскільки він може змінювати зашифровані дані з вихідних програм, або змінити адресу призначення одержувача.

Щоб переслати заявку даних на мульти-проксі, потрібен відповідний протокол передачі даних. SOCKS5 використовується як протокол передачі даних між додатками та клієнтським вузлом.

SOCKS — це діючий протокол обміну повідомленнями між прикладним і транспортним рівнем.

Він розроблений для прозорого і безпечного проходу через брандмауер.

В останній версії п’ять, які визначені в RFC 1928[47] підтримує аутентифікацію, IPv6 і передачу даних як по TCP, так і по UDP. SOCK5 містить три фази: узгодження, передачу адреси та передачу даних, як показано на рисунку 3.5.

Коли протокол зініціюється, клієнт SOCKS5 і сервер SOCKS5 обмінюються різними методами, які будуть використовуватися на етапі переговорів, потім клієнт надсилає адресу призначення серверу.

Сервер згодом повідомляє клієнта про успішне з’єднання, відповідаючи а обов'язкове повідомлення.

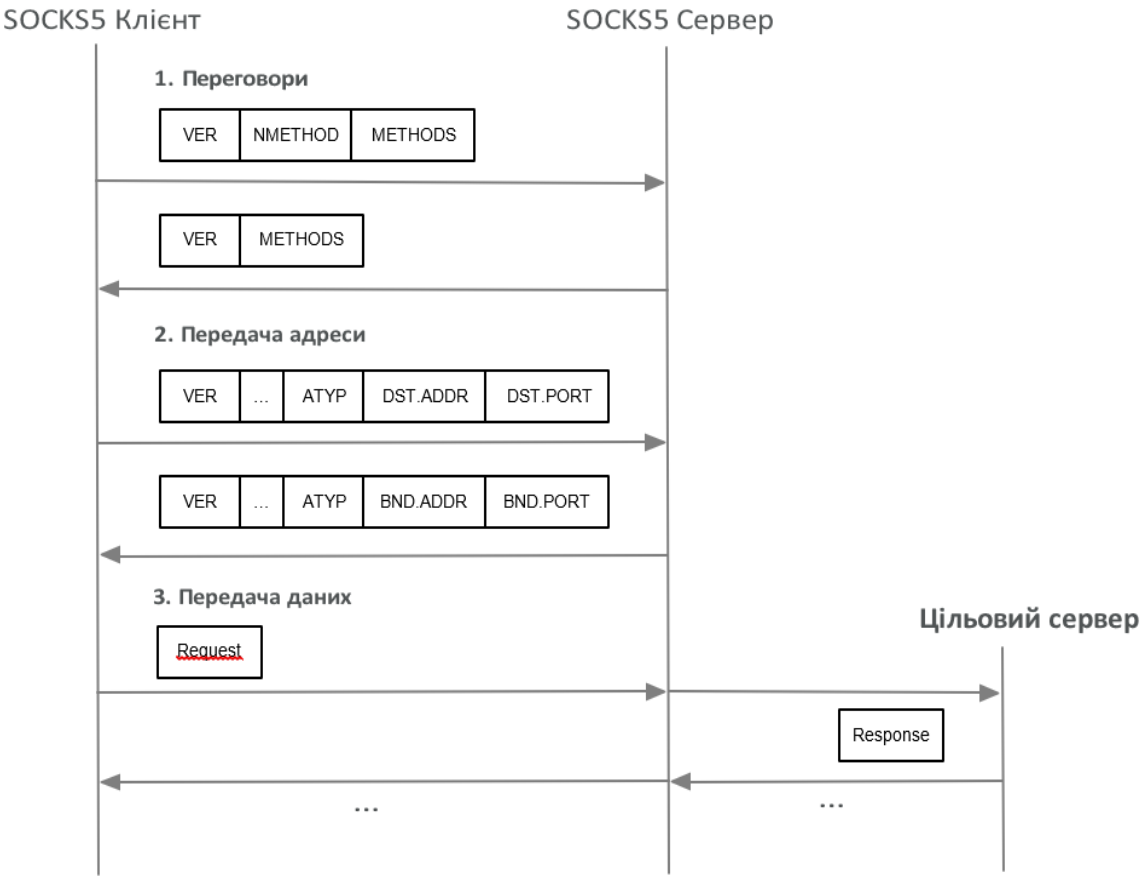


Рисунок 3.5 - Потік роботи протоколу SOCKS5

Клієнт і сервер починають передавати дані на третьому етапі якщо з'єднання було створено успішно. Причини, чому мульти-проксі використовує протокол SOCKS5 між додатками а клієнтський проксі завдяки своїй гнучкості та розширюваності.

Оскільки SOCKS5 ні підтримує лише протокол транспортного рівня, наприклад TCP та UDP, але також може пересилати кілька протоколів прикладного рівня, такі як HTTP і HTTPS. У системі мульти-проксі вузол сервера розташований у нецензурному місці домену та чекає даних від клієнтського вузла через захищений канал для відправлення у безкоштовний Інтернет.

Метою є дешифрування, реконструкція або зміна даних перед пересиланням на вказану адресу призначення, наприклад на веб-сервер заблокований системою цензури. Після того, як вузол сервера отримає дані від віддаленого веб-сайту, він перенаправить трафік назад на клієнтський вузол.

Рівень проксі працює належним чином, лише якщо GFW не знає адреси серверного вузла, оскільки сервер можна занести в чорний список. Щоб зробити систему більш масштабованою, тобто клієнтський вузол може використовувати кілька серверних вузлів для пересилання свого трафіку.

Таким чином, навіть якщо деякі сервери вилучено GFW, який може відбувається в традиційній установці Shadowsocks, клієнтський вузол все ще має інші можливості.

Оскільки проксі-сервери Shadowsocks зазвичай розташовуються в публічних хмарах, цензура не може заблокувати цілі домени, оскільки це призведе до величезної побічної шкоди. Об’єднавши всі проксі SOCKS в єдину спільну інфраструктуру, система стає більш стійкою до блокування окремих серверів, ніж у а одностороння установка. Крім того, кілька серверних вузлів можуть покращити пропускну здатність.

З цих причин мульти-проксі побудовано на основі однорангової мережі. Вузли однорангової мережі відрізняються від традиційних серверів, вони мають фіксовані та відомі IP-адреси. Подібно до людських стосунків реальному світі, хтось приєднується до мережі, представивши учасника, який є вже є частиною мережі. Принаймні цього вимагає спосіб однорангового приєднання підключається до вузла, який уже є в цій мережі.

На практиці статус вузлів динамічно змінюється, так що вузол, що знову приєднався, не може відразу знайти однорангового. З цієї причини деякі вузли завантаження, налаштовані для надання початкової конфігурації однорангової мережі, наприклад IP адреси та ідентифікатори послуг.

Як показано на рисунку 3.6, для однорангових, які цього не роблять знаючи будь-яких інших однорангових партнерів у мережі, вони спочатку підключаються до надійного вузла завантажувального ременя та зареєструють свої адреси, а також ідентифікатори служб, які використовуються щоб уточнити, які послуги вони надають.

Вузол початкового завантаження надсилає свої конфігурації назад до обох однорангових пристроїв. Вузлам дозволено виявляти та спілкуватися з ними один до одного, лише якщо вони мають однакові ідентифікатори служби.

Після вступів, кожен вузол мережі підтримує свій список відомих однорангових. Вони посилають ping і періодично надсилати повідомлення pong, щоб відстежувати статус інших вузлів, наприклад, оновлення нові вузли або видалити автономні вузли.

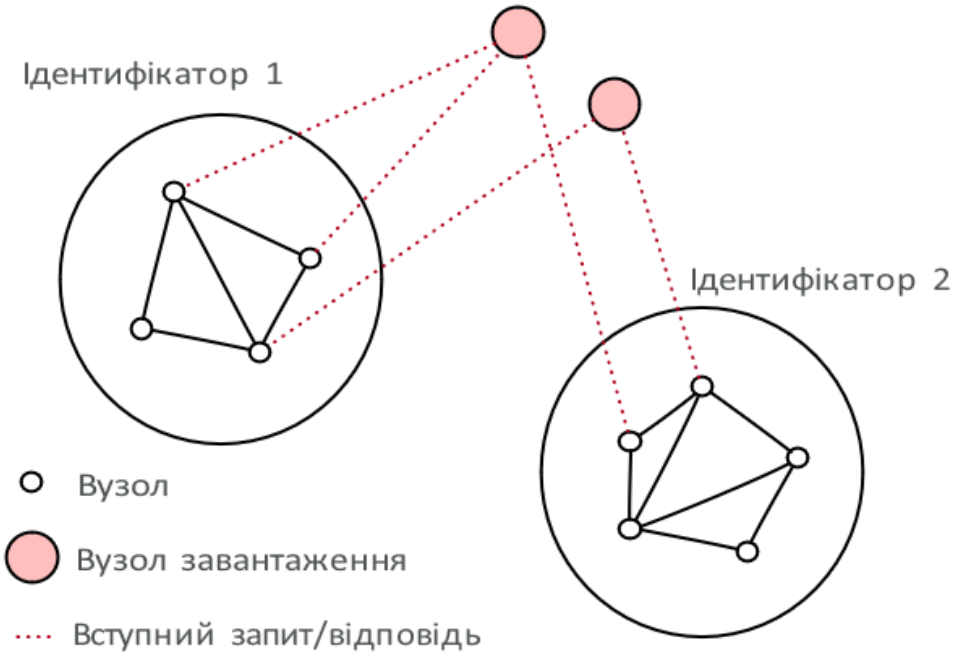


Рисунок 3.6 – Завантаження однорангової системи

3.4 Токенова економіка

Нерегульована система може працювати на практиці або досягти стабільного стану лише за умови інтересу між сторонами, які надають послуги, та тими, хто споживає послуги, збалансовані. У цьому випадку мульти-проксі працює надійно, лише якщо люди в домені без цензури готові поділитися своїми ресурсами та в нести їх у звичайний обхід.

У мульти-проксування головне завдання – збалансувати кількість серверів і клієнтів. Зараз багато людей вже запускають свої сервери Shadowsocks у публічних доступних хмарах.

Це означає новий механізм стимулювання людей налаштовувати свій сервер екземплярів у систему мульти-проксі, коли вони хочуть використовувати її, щоб уникнути проблеми фрирайдерів [48].

Проблема фрирайдерів є значною загрозою для рівноправних систем таких систем, як BitTorrent в які однорангові користувачі лише завантажують файли, не завантажуючи нічого, що є несправедливим для однорангових які сприяють розвитку мережі і можуть мати негативний вплив на систему.

Рішення для оптимізації швидкості завантаження в BitTorrent полягає в тому, щоб використовувати принцип «ти за око», який бере початок з високоефективної стратегії теорії ігор.

У цій стратегії BitTorrent задушить однолітків, які не співпрацюють, які не беруть участь у завантаженні файлів і розподілені своїх обмежених слотів для завантаження іншим, більш кооперативним одноранговим.

Проблема безкоштовних користувачів у системі мульти-проксі означає, що однорангові користувачі використовують послуги обходу, що надаються вузлами сервера, не вносячи внесок у відповідь, надаючи пропускну здатність свого проксі-сервера для інших вузлів.

Це може призвести до незбалансованої кількості проксі-серверів і клієнтів, а саме кількість клієнтів значно перевищує кількість серверів.

Щоб уникнути цього явища, мульти-проксі вводить токени як спосіб створити стимули та збалансувати надання ресурсів із споживанням. Точніше, в системі мульти-проксі клієнтський вузол повинен платити токенів перед використанням служби обходу, наданої вузлом сервера. Тим часом вузол сервера може заробляти токени, надаючи свою послугу обходу.

У відкритому середовищі, як-от Інтернет, деякі вузли можуть егоїстично не використовувати систему для свого використання. У цьому розділі описані загрози та виклики, які виникають у мульти-проксі та як він має протистояти.

Є три основні загрози, якими буде існувати система зустрічаються: єдина точка збою, фальсифікація маркерів і вторгнення в дані конфіденційність. Перші дві загрози можна узагальнити як де і як зберігати токени.

Тому мульти-проксі зосереджується на вирішенні двох проблем економіки токенів, а саме, як захистити токени та як захистити дані та конфіденційну інформацію, що передаються між ними.

Однією із слабких сторін централізованих серверів транзакцій є єдина точка збою.

Єдина точка відмови (SPOF) може очолити всю систему економіки маркерів припинити роботу, якщо централізовані сервери транзакцій зламані або виходять з ладу, і проблема довіри означає, що централізовані сервери транзакцій мають права на втручання або переглянути записи, і ніхто в мережі не зможе перевірити їх правильність.

Навіть за допомогою децентралізованої системи обліку такі є загрози, які існують в системі.

Перша загроза полягає в тому, що шкідливий вузол сервера не надає еквівалентні послуги після отримання кількості токенів.

У більшості У випадках серверний вузол може отримати токен без надання еквівалентної послуги.

Наприклад, шкідливий сервер може бути розташований всередині домену, підданого цензурі, для єдина мета майнінг токенів.

Однак він не надає справжню послугу обходу, оскільки не має можливості передавати пакети не цензури в Інтернет.

Цю шкідливу поведінку можна виявити за допомогою аутентифікації виклик-відповідь, як показано на рисунку 3.7.

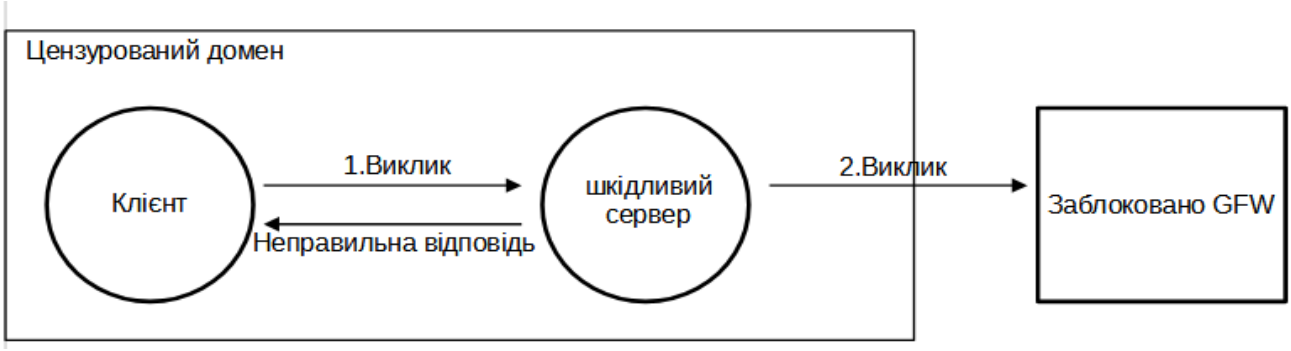


Рисунок 3.7 - Механізм відповіді на виклик

Якщо вузол сервера вирішує бути обхідним сервер, він повинен завершити автентифікацію виклик-відповідь, до якої звикли визначити, чи правильно вузол сервера виконує свою роль, у цьому випадку відправити назад вміст деяких зазначених заблокованих веб-сайтів.

Шкідливий вузол сервера також може обманювати в майнінгу, щоб отримати більше токенів, наприклад, у моделі майнінгу на основі часу зловмисний сервер може налаштувати систему годинник швидше.

У моделі пропускної здатності або мережевого трафіку серверу це легко зробити підробити пропускну здатність або пакети, які він передає.

Інша загроза полягає в тому, що серверний вузол може маніпулювати даними, які надходять з клієнтський вузол.

По-перше, знаючи майже інформацію про клієнтський вузол, наприклад, IP адресу, номер порту та повідомлення, вузол сервера може виконувати аналіз трафіку так що вузол клієнта може втратити конфіденційність.

На додаток до цього, сервер також може запустити атаку «Людина посередині» (MITM) кількома способами, наприклад, зупинити пересилання або скиньте повідомлення та змініть повідомлення в якомусь незашифрованому випадку.

3.5 Повідомлення з кількома переміщеннями

Одним із способів захисту конфіденційності є обмеження знань про вузли, тобто жоден вузол не знає повну інформацію про мережу. Повідомлення з кількома переходами маршрутизація є контрзаходом для захисту конфіденційності, оскільки вона запобігає ідентифікації автори запиту.

Наскрізне анонімне спілкування для захисту конфіденційності використовує анонімне спілкування.

Цього можна досягти шляхом побудови ланцюгів передачі даних, тобто вузлів, розташованих по-різних шляхах.

Маршрутизація цибулі — це спосіб послідовного обгортання та шифрування повідомлень на рівнях через мережу, яка має кілька проміжних вузлів.

Віна може добре захистити конфіденційність автора, оскільки повідомлення шифруються, надсилаються між ними і жоден проміжний вузол всередині схеми не може визначити джерело та кінцеве призначення повідомлення, крім вихідного вузла.

В системі мульти-проксування, повідомлення захищені механізмом лукової маршрутизації.

Автори запиту можуть вказати, скільки стрибків вони хочуть створити. Тоді воно починає будувати схеми між списком вузлів, це можна досягти за допомогою асиметричного ключа криптографії, наприклад, алгоритм обміну ключами Діффі-Хеллмана, для узгодження різних сесійних ключів через мережу.

Рисунок 3.8 показує схему з 2 стрибками від вузла клієнта до вузла виходу (вузла сервера)..

Вузли сервера оголошуються як вихідні вузли і стають останніми переходами для пересилання даних на цільовий веб-сервер. Після того, як ланцюг буде побудовано, дані будуть передані всередину схеми.

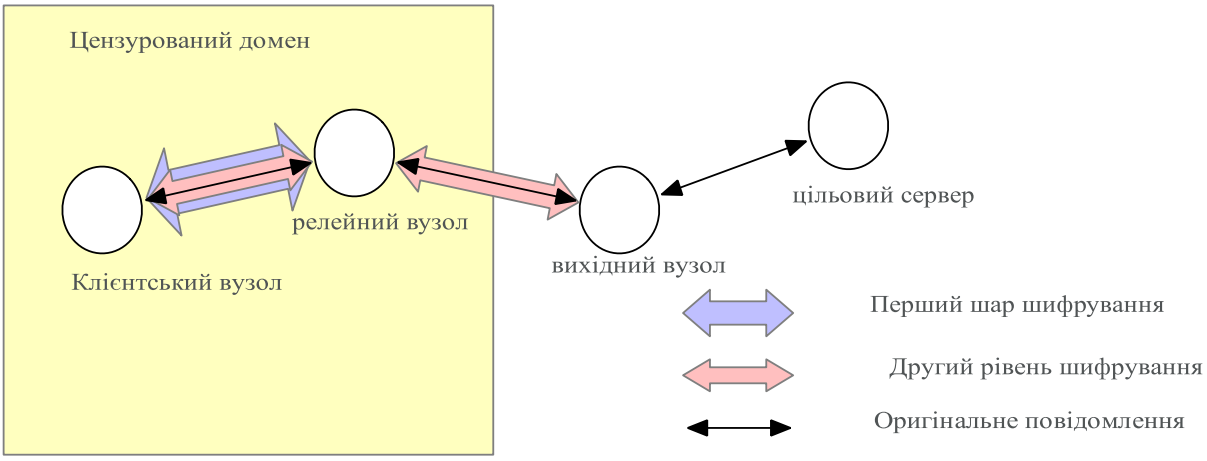


Рисунок 3.8 - Схема маршрутизації цибулі з 2 стрибками

3.6 Висновки

У цьому розділі представлено процес побудови ланцюга для спільноти тунелів. Автор запиту спочатку вказує скільки стрибків він хоче створити, тоді він запитує наступний стрибок за допомогою повідомлення «створити», як тільки hop 1 отримує повідомлення, він генерує спільний ключ сеансу та відправляє назад «створене» повідомлення автору.

Потім автор надсилає повідомлення «розширити» для переходу 1, а перехід 1 пересилає повідомлення «створити» до переходу 2.

Коли hop 2 отримає повідомлення та згенерує новий ключ сеансу. Потім надсилає повернути «створене» повідомлення для переходу 1. Перехід 1 надсилає «розширене» повідомлення назад ініціатору.

Процес повторюється, коли він досягає вихідного вузла. Кожен вузол в ланцюг знає лише поточні стрибки та мінус один, перш ніж відправляти його до наступного стрибка, тому кожен проміжний перехід не може розрізнити джерело повідомлень і кінцевий призначення, оскільки всі вони отримують однакові повідомлення протягом періоду будівництва ланцюга.

4 ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ РОБОТИ

4.1 Аналіз системи на недоліки

Для таких платформ, як Facebook і Shopify, потрібна єдина кінцева точка для початку рукостискання аутентифікації або базова URL-адреса для вбудовування/відкриття вашої програми.

Ця вимога робить виснажливим процес створення кількох середовищ, оскільки вони зазвичай вимагають ручного налаштування кількох програм на цих платформах. Мульти-проксі забезпечує безпроблемне рішення як для користувачів, так і для розробників.

Використовуючи мульти-проксі, користувачі можуть легко вибрати, до якого середовища вони хочуть отримати доступ, а розробники повинні налаштувати одну програму на цих платформах. Мульти-проксі також дозволяє розробникам створювати динамічні середовища, оскільки їм не потрібно створювати додаткову програму для кожного середовища.

Мультипроксі покладається на файл cookie браузера, щоб визначити, яке середовище активне на даний момент, і відповідно маршрутизувати запити. Аутентифікаційне рукостискання та запуск програми відбувається на рівні браузера, на цьому етапі немає прямого зв’язку між платформою програми та програмою.

Деякі платформи, такі як Shopify, надсилають події назад у програми за допомогою Webhooks, у цьому випадку у вас є два варіанти:

Встановити конкретну кінцеву точку webhook для кожного середовища (якщо можливо);

Налаштувати мульти-проксі для трансляції цих вебхуків у всі середовища.

Мульти-проксі має вбудований веб-інтерфейс, щоб користувачі могли переглядати програми та вибирати активне середовище кожної програми.

Веб-інтерфейс — це просто гарний інтерфейс для створення файлів cookie, які використовуються мульти-проксі для маршрутизації запитів і встановлення належного значення.

Список додатків зображений на рисунку 4.1.

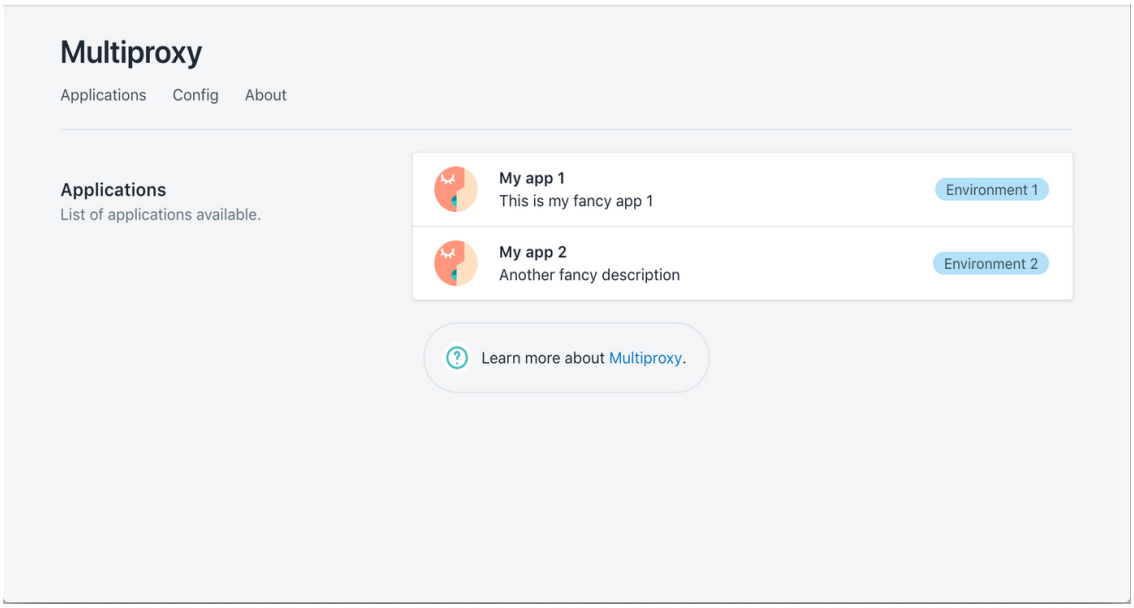


Рисунок 4.1 - Список додатків

Оскільки ви можете налаштувати веб-інтерфейс в іншому домені, веб-інтерфейс буде спілкуватися з кінцевою точкою програми за допомогою XHR + Cross-Origin Request Sharing (CORS).

Кожного разу, коли користувач натискає, щоб вибрати середовище, веб-інтерфейс надсилає XHR до кінцевої точки програми мульти-проксі із заданим ключем середовища, цей запит відповідно оновлює файли cookie та мульти-проксі.

Список середовищ зображений на рисунку 4.2

Розроблене середовище дозволяє користувача переглядати програми та вибирати активне середовище кожної програми для безпечної передачі інформації в мережі інтернет.

4.2 Продуктивність мережі

Ефективність мережі можна розглядати як найважливішу частину системи оцінки. Розглянемо такі фактори як затримка та пропускна здатність.

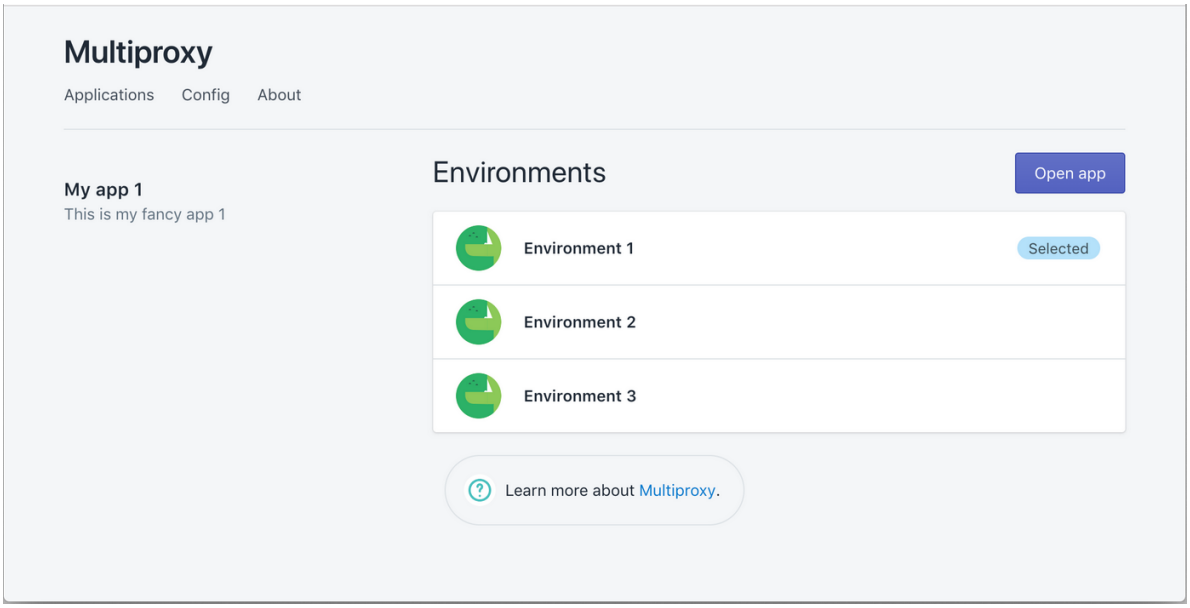


Рисунок 4.2 - Середовища роботи розробленої програми

Затримка мережі, яка також відома як затримка від кінця до кінця, відноситься до загального часу, який проходить пакет від джерела до місця призначення. На затримку мережі може впливати кілька причин, викликаних будь-яким об’єктом передачі між початковою та кінцевою точкою. Наприклад, LAN, WAN або шлях від ISP до хоста.

Одним з інтуїтивно зрозумілих способів вимірювання затримки є обчислення часу передачі, віднімаючи час початку від часу надходження в різні вузли. Однак, оскільки синхронізувати годинники між різними пристроями є нетривіальною проблемою, час Roundtrip (RTT) зазвичай використовується для вимірювання затримки між двома вузлами.

Затримка в обидва боки означає, що порівняння за весь час проводяться з одного і того ж вузла.

Оскільки затримка мережі часто змінюється, одним із способів отримати точну затримку є повторення вимірювання пропускної здатності протягом дня через регулярні інтервали часу та отримання максимального, мінімального, середнього та стандартного відхилення.

Пропускна здатність мережі означає кількість успішно доставлених даних по каналу зв'язку за одиницю часу. Він подається у бітах на секунду (bps). Пропускну

здатність можна оцінити шляхом обчислення ділення загального числа байтів, які один вузол отримав через RTT. Однак фактична пропускна здатність зазвичай нижча за теоретичну пропускну здатність середовища через різні додаткові фактори, такі як одночасне використання мережі, перевантаження та керування потоком.

4.3 Продуктивність системи

Хоча сьогодні більшість комп'ютерів мають потужні процесори і великий об'єм пам'яті, продуктивність системи все ще є важливим фактором для роботи користувача. Використання ЦП та пам'яті оцінюється під час використання мульти-проксування.

Використання ЦП вказує пропорції циклів, призначених для виконання конкретної програми. Експеримент вимірює використання ЦП, тобто час, поділений на час виконання процесу, і виражене у відсотках.

Використання пам'яті означає, скільки фізичної пам'яті використала програма під час свого виконання. При оцінці використовується фізична пам'ять, яку займає процес, замість співвідношення між розміром, встановленим резидентом, до фізичної пам'яті машини, оскільки відношення буде близьким до нуля, якщо машина має великий обсяг пам'яті.

4.4. Методології та експериментальні кроки

Розташування сервера вважається основним фактором впливу мульти-проксі. Щоб отримати як ідеальну продуктивність, так і практичну роботу. Мульти-проксі розгортається у двох різних ситуаціях.

Перший випадок — це надмірне забезпечення, яке забезпечує достатню потужність для передачі мережевого трафіку. Таким чином, передбачається, що мережевий трафік має низьку затримку та відносно високу пропускну здатність з низькою швидкістю втрати пакетів. Ідеальну продуктивність можна виміряти в

першому випадку, тоді як у другому випадку є більше факторів невизначеності, наприклад, скільки ресурсів або буферів можуть утримувати проміжні маршрутизатори.

Вибір місця перелічено таким чином: Усі вузли розгорнуті в одному кластері DAS5. Вузли розгорнуті серед різних екземплярів хмари в глобальному масштабі.

Розподілений суперкомп'ютер ASCI 5 (DAS5) — це шести кластерна розподілена система, яка забезпечує спільну обчислювальну інфраструктуру для паралельних і розподілених завдань.

Головний вузол кожного кластера керує завданнями за допомогою системи пакетної черги SLURM. Кожен вузол для експериментів має два восьми ядерних процесора Intel Xeon E5 (2,40 ГГц) з технологією Intel Hyper-Threading, яка дозволяє одноядерним процесорам вести себе як два логічні процесори, тому кожен вузол має загалом 32 логічні процесори. Об'єм пам'яті становить 62 ГБ. Для мережевих конфігурацій використовує мережу InfiniBand, яка забезпечує до 48 Гбіт/с для передачі даних.

Оскільки деякі користувацькі або зареєстровані порти фільтруються брандмауером, мульти-проксі, що працює в різних головних вузлах, не може безпосередньо підключитися, тому експеримент з вузлами, розташованими в різних кластерах, не може бути завершено.

Що стосується вибору глобальних екземплярів сервісів, то клієнтом є Alibaba, розташований у Південному Китаї.

Внутрішні переходи-кандидати та вузли виходу – це всі екземпляри Google Cloud, які розташовані окремо в Нідерландах та Південній Каролініі, США. Ці хмарні екземпляри використовують високопродуктивний мережевий рівень преміум-класу, а топологія мережі показана на рисунку 4.1.

Клієнт у Китаї спочатку підключається до одного екземпляра хмари, розташованого в Нідерландах, цей екземпляр діє як клієнт і пересилає дані на інший екземпляр у Нідерландах і, нарешті, досягти вихідного вузла в США.

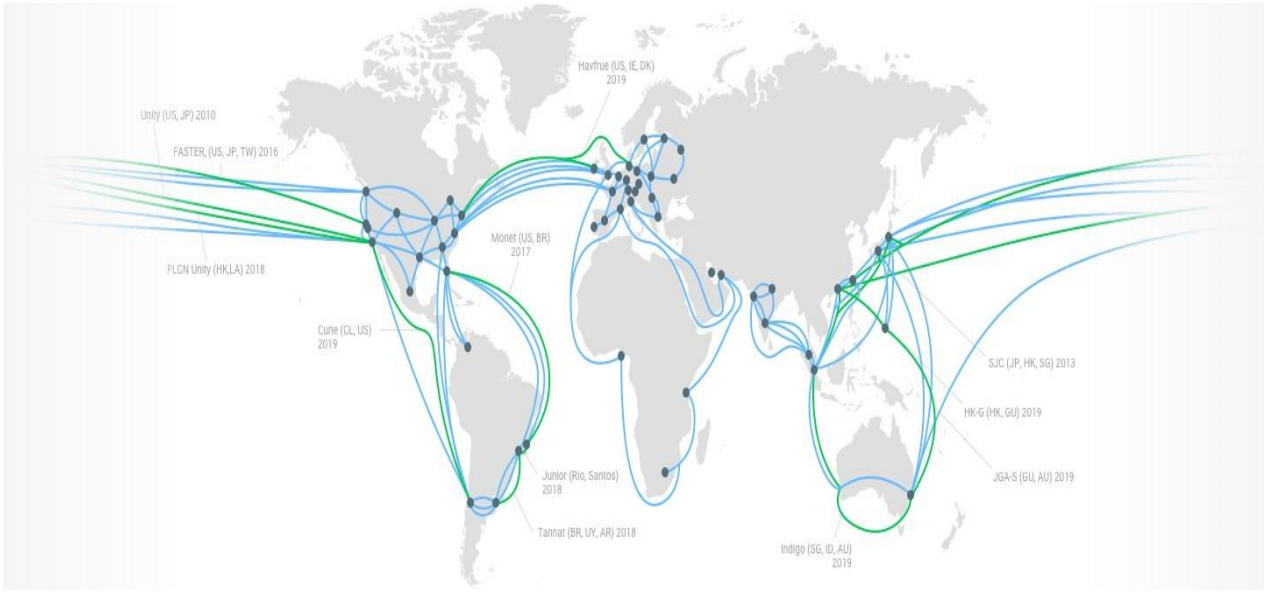


Рисунок 4.3 - Топологія мережі преміум-класу екземплярів Google Cloud

Порівняні дослідження проводяться для вимірювання практичної продуктивності мульти-проксі. Його порівнюють з іншими репрезентативними системами обходу цензури. Ці системи поділяються на три категорії відповідно до їхніх протоколів та архітектури. Для кожної категорії вибираються найбільш популярні та представницькі системи.

Перше - це рішення проксі, яке працює на транспортному рівні, включаючи Shadowsocks і V2Ray. Іншим є рішення віртуальної приватної мережі (VPN), які працюють на мережевому рівні, включаючи OpenVPN і OpenConnect. Останнє — це рішення I2P, і його типове застосування — Tor.

Різні CRS тестуються за допомогою різних методів шифрування та доступу. Shadowsocks налаштований як метод шифрування AES-256-CFB для захисту даних, що передаються між проксі-клієнтом і сервером, V2Ray використовує само визначений протокол зв'язку, який називається VMess.

Для рішень VPN і OpenVPN, і OpenConnect використовують шифрування AES-256-GCM і метод стиснення даних LZ4 на основі безпеки транспортного рівня. Усі сервери CRS, включаючи проксі-сервери та сервери VPN, розгорнуті в екземплярі США і діють як вихідний вузол, за винятком Tor, оскільки Tor використовує свої налаштування маршрутизації цибулі.

Для мульти-проксі, екземпляр для Китаю з екземпляром США, створеним як установка з одним переходом. Налаштування з двома стрибками складається з екземплярів Китаю, Нідерландів та США, а налаштування з трьома переходами, включаючи один екземпляр у Китаї, два екземпляри в Нідерландах та один екземпляр у США.

4.5 Методи блокування контенту

Існує чотири категорії Інтернет-цензури, цензури на стороні клієнта, цензури на стороні сервера, внутрішньої цензури та цензури на шляху. У цій дисертації вибирається одна з відомих загальнодержавних систем моніторингу та нагляду за цензурою,

Великий китайський брандмауер, як приклад, тому вона в основному зосереджена на цензурі на шляху та на шляху. GFW використовує декілька модулів для перевірки та фільтрації небажаного мережевого трафіку.

Блокування не виконує дії на фізичних рівнях, оскільки це основний рівень, що лежить в основі вищих шарів і забезпечує вимогу фізичного підключення. Прохідна цензура працює на мережевому рівні, який забезпечує засоби передачі мережевих пакетів від джерела до призначення. Блокування IP-адреси є основним засобом перехоплення трафіку.

Він порівнює пункт призначення в пакеті зі списком контролю доступу всередині міжнародних шлюзів, і за допомогою протоколу BGP небажаний мережевий трафік перенаправляється на деякі випадкові адреси або відкидається. Блокування IP — це простий і прямий спосіб блокування мережевого трафіку, оскільки маршрутизатор бере безпосередню участь у процесі маршрутизації.

Натомість цензура на шляху безпосередньо не бере участь у маршруті між двома кінцевими точками, вона копіює весь мережевий трафік, а потім відстежує, крім шлюзів, і виконує атаки «людина посередині».

Він працює на транспортному рівні, який забезпечує зв’язок між хостом між програмами.

Основними протоколами транспортного рівня є протоколи TCP і UDP, GFW все ще використовує виявлення ключових слів для вищого прикладного рівня, щоб вирішити, чи слід блокувати трафік.

Наприклад, виявлення ключових слів HTTP/DNS. Скидання з'єднання TCP використовується у фразі прийняття рішення, тобто GFW перехоплює з'єднання, надсилаючи RST-пакети на обидві кінцеві точки після тристороннього рукостискання TCP.

Цей вид атаки «людина посередині» зберігає стан, і GFW може запам'ятати стан з'єднань протягом 90–95 секунд, що означає, що якщо клієнт хоче відновити з'єднання між сервером, запити на рукостискання негайно підслуховуються та блокується GFW, а серверна сторона навіть не знає про існування клієнта. Таким чином, блокування стає ефективнішим, оскільки воно посилає лише пакети RST на сторону клієнта.

GFW в основному використовує два методи для протоколу DNS через UDP: захоплення DNS та отруєння DNS, а отруєння DNS є побічним ефектом першого.

Викрадання DNS відбувається, коли клієнт надсилає запити на DNS-сервери в домені без цензури.

Він працює шляхом прослуховування посередині, створюючи фіктивний пакет відповіді DNS з неправильною роздільною здатністю.

Оскільки GFW має менше стрибків маршрутизації, ця неправильна адреса буде прийнята клієнтом, а правильна відповідь, яка надходить пізніше, відкидається стеком TCP/IP на стороні клієнта.

Отруєння DNS має ширший негативний вплив у порівнянні з викраденням DNS.

Це трапляється, коли внутрішній DNS-сервер надсилає запити рекурсивно, а GFW виконує викрадення DNS і вводить неправильні відповіді в свій кеш.

Ці неправильні відповіді можуть залишатися на кілька хвилин або днів. Що стосується DNS через TCP, то GFW використовує скидання з'єднання TCP.

4.6 Сучасні методи боротьби з цензурою

Найпоширенішими системами обходу цензури є архітектура на основі клієнт-сервер, включаючи рішення VPN та рішення проксі.

VPN працює на мережевому рівні. Він працює шляхом встановлення наскрізних віртуальних з'єднань за допомогою виділених схем, а IP-пакети шифруються через цілі з'єднання.

Проксі працює на транспортному рівні. З точки зору безпеки проксі-сервер слабкіший, ніж VPN, але він більш гнучкий, оскільки обробляє лише потоки та дата грами на вищому рівні і не потребує шифрування всього IP-пакета. Він може розрізняти різний мережевий трафік. Створюючи чорний або білий список, він може швидко вирішити, який тип мережевого трафіку слід зашифрувати чи ні.

Проксі-сервер і VPN, вживають заходів для захисту передачі повідомлень на льоту. VPN більше зосереджується на шифруванні, такому як повідомлення автентифікації або обмін сертифікатами. Різні протоколи VPN мають різні реалізації. Вони можуть пропускати певні функції мережевого потоку через передачі.

Хоча шифрування може легко протистояти підслуховувачам у середині, це не працює для GFW, оскільки GFW спрямований на блокування мережевого трафіку, але не на його розшифрування. Це причина, чому деякі протоколи VPN нестабільні або заблоковані GFW.

Проксі використовує методи обфускації, щоб уникнути моніторингу GFW, його мережевий трафік більше схожий на звичайний мережевий потік між двома кінцевими точками.

Як правило, проксі-сервер не шифрує заголовок протоколів мережевого і транспортного рівня, а лише тунелює дані обфускації всередині тіла повідомлення протоколу, вони використовують кілька методів, наприклад різні алгоритми, для шифрування або рандомізації тунельних повідомлень.

Деякі незвичайні протоколи все ще можна розрізнити, наприклад, протокол SSH.

У результаті більшість проксі-серверів використовують звичайні протоколи HTTP або HTTPS для тунелювання вихідних запитів.

Ці рішення на основі клієнт-сервер прості в розгортанні та використанні, і зазвичай вони мають низьку затримку під час особистого використання, але їх недолік очевидний: як тільки це спостерігається GFW, екземпляр хмари негайно блокується, що призведе до деяких втрат для користувача, наприклад, служба недоступна, і їм потрібно знову розгорнути проксі-сервер із самого початку.

Інші програми використовують розподілену анонімну мережу, наприклад Tor. Ця мережа отримує кращий захист конфіденційності, оскільки вузли в мережі мають лише частину повідомлень ідентифікації, а проміжний вузол знає лише адреси своїх попередніх і наступних вузлів.

Хоча останній перехід декапсулює повідомлення, він не знає власника запиту.

Розподілена мережа використовує більше ресурсів, ніж стандартна клієнт-серверна архітектура. Сервери, якими керують волонтери, створюють розподілену мережу.

Як тільки GFW блокує ланцюг, програма може переключитися на інший канал без повторного розгортання, і це дозволяє уникнути збою однієї точки проксі-сервера.

Ще одна перевага полягає в тому, що Tor безкоштовний, а також його легко використовувати.

Користувачі можуть використовувати мережу Tor, завантаживши браузер і скориставшись підключеним транспортом у домені, підданому цензурі.

4.7 Висновки

Метою проектування є поєднання переваг клієнт-серверної та розподіленої архітектури та подолання недоліків клієнт-серверної архітектури, таких як єдина точка відмови та брак ресурсів.

Перш за все, оскільки більшість популярних веб-сайтів блокуються GFW, системі потрібен базовий модуль переадресації трафіку, який забезпечує базову послугу обходу.

Архітектура схожа на CRS на основі клієнт-сервер. Він містить проксі-сервер на стороні клієнта для отримання та обробки повідомлень від програм-авторів і пересилання цих повідомлень на проксі-сервер віддаленої сторони.

З огляду на протоколи пересилання трафіку, мульти-проксування використовує SOCKS версії 5 як протокол між додатками та клієнтським проксі, оскільки він є більш гнучким і масштабованим, ніж інші протоколи, такі як проксі HTTP/HTTPS.

Оскільки мульти-проксі також забезпечує конфіденційність для ініціаторів запитів, він може додавати кілька шарів/передавачів/переходів на шляху між клієнтським проксі-сервером і проксі-сервером.

Щоб забезпечити збалансовану кількість вузлів, вона має економіку маркерів, яка використовує протокол ланцюга довіри як розподілений реєстр, а Intel SGX запобігає шахрайству вузлів.

Структура оцінки описана в розділі 4, і вона оцінює продуктивність мережі та системи мульти-проксі в DAS5, що є ідеальним середовищем. Результат можна вважати найкращою продуктивністю, яку може досягти мульти-проксі.

Результати показують, що мульти-проксі здатний уникати GFW, а також показує розриви продуктивності між різними CRS.

Оскільки GFW спрямований на сортування підозрілого мережевого трафіку для обходу, а не на дешифрування цих повідомлень між ними.

Тому деякі нові або створені самостійно протоколи, навіть якщо вони не зашифровані, працюють для обходу, поки GFW не помічає їх.

Проте обфускація є кращим рішенням, оскільки приховування або тунелювання мережевого трафіку для обходу в стандартні протоколи, які не можна блокувати, наприклад, передавання домену, може збільшити труднощі регулювання.

ВИСНОВКИ

У роботі за результатами виконаних теоретичних та практичних досліджень розроблено середовище на основі мульти-проксування. В сучасні межежі є багато небезпек для користувача та фільтрів контенту.

Обхід фільтрів не є причиною того, чому були винайдені проксі. Основна перевага їх використання полягає в тому, що вони роблять вас більш анонімними.

Конфіденційність – це дуже важливе завдання не тільки для людей, які не хочуть мати проблем з правоохоронними органами, а й для звичайного користувача.

Незахищені канали зв'язку можуть пошкодити персональні дані користувача, до яких можуть отримати доступ як провайдер, так і зловмисник.

В роботі я дійшов висновку що, одним з найбільш надійних засобів анонімізації є проксі-сервера.

У користувачів в інтернеті є свій, унікальний IP адрес, який занесений в базу даних провайдера. Тому щоб залишатися в мережі анонімно, людина може придбати новий IP адрес, того регіону в якому проживає або будь-якого іншого. Тому головна відмінність проксі від VPN сервісів, полягає в тому, що проксі-сервер не шифрує вхідні та вихідні дані.

Отримати всі перераховані вище переваги, можна і не використовуючи додаткове ПО на комп'ютері. Набравши в адресному рядку браузера слово «Анонімайзер», користувачі побачать посилання на ресурси, що пропонують послуги анонімізації в інтернеті.

Найчастіше, інтерфейс даного сервісу є а рядок адреси, куди користувач вводить адресу цільового ресурсу.

Але на жаль, використовувати подібні сервіси вкрай небезпечно. Вводячи логін і пароль від своєї соціальної мережі, невідомо кому в руки він може потрапити.

Використання анонімайзерів підійде скоріше не для забезпечення безпеки особистих даних, а для доступу до заблокованих сайтів.

Підводячи підсумки, слід зазначити, що всі перераховані вище послуги не можуть гарантувати стовідсоткову безпеку даних. Навіть Tor браузер має свої вади і неграмотне використання, може видати людини.

Наприклад, створена сторінка в соціальній мережі на своєму IP адресу, стає непридатною для використання її через Tor браузер.

У першому розділі було досліджено вплив інтернет цензури на користувачів та виявлена її небезпека.

Було наведено сучасні і безпечні обхідні системи які допоможуть користувачам безпечніше користуватися мережею інтернет. Також були досліджені механізми аналізу та блокування для кращого розуміння роботи цензури.

У другому розділі експериментально проведені досліди роботи інтернет цензури та її наслідки для користувачів.

Інтернет цензура характеризується в контролі та припиненні публікацій або доступу до інформації в мережі Інтернет. Своєю появою інтернет-цензура зобов'язана відсутності будь-яких національних кордонів в мережі Інтернет.

Загальну проблематику інтернет-цензури можна визначити таким чином: інформація, що порушує закони держави (режим чинного уряду) та заблокована на внутрішніх ресурсах, може бути опублікована на веб серверах в інших країнах.

У третьому розділі дослідили маршрутизацію трафіку та її дизайн.

Що дозволило краще розуміти архітектуру системи мульти-проксі та її переваги над іншими системами забезпечення безпечної передачі інформації в мережі. Були наведені компоненти системи мульти-проксі які дозволяють надавати послуги обходу та роблять систему надійною.

Була досліджена схема маршрутизації цибулі.

Технологія анонімного обміну інформацією через комп'ютерну мережу. Повідомлення неодноразово шифруються і потім відсилаються через кілька мережевих вузлів, званих цибулевими маршрутизаторами.

Кожен маршрутизатор видаляє шар шифрування, щоб відкрити трасувальні інструкції, і відіслати повідомлення на наступний маршрутизатор, де все

повториться. Таким чином проміжні вузли не знають джерело, пункт призначення і зміст повідомлення.

У четвертому розділі була досліджена ефективність на практиці в розробленій програмі яка надає функції мульти-проксування.

Було доведено що кешування проксі-сервера прискорює запити на обслуговування, витягуючи вміст, збережений з попереднього запиту, зробленого тим же клієнтом або навіть іншими клієнтами.

Проксі-кешування зберігають локальні копії часто запитуваних ресурсів, дозволяючи великим організаціям значно знизити використання пропускної здатності та витрати на вихідний канал, одночасно значно підвищуючи продуктивність. Більшість провайдерів і великих компаній мають кешуючий проксі.

Кешування проксі були першим типом проксі-сервера. Веб-проксі зазвичай використовуються для кешування веб-сторінок із веб-сервера.

Погано реалізовані проксі-сервери кешування можуть викликати проблеми, такі як неможливість використовувати аутентифікацію користувача.

Анонімний проксі-сервер зазвичай намагається анонімізувати веб-серфінг. Анонімайзери можна розділити на кілька різновидів.

Сервер призначення (сервер, який в кінцевому підсумку задовольняє веб-запит) отримує запити від анонімізуючого проксі-сервера і, таким чином, не отримує інформацію про адресу кінцевого користувача.

Однак запити не є анонімними для анонімного проксі-сервера, тому між проксі-сервером і користувачем існує певний рівень довіри.

Багато проксі-серверів фінансуються через постійне рекламне посилання на користувача.

Набула подальшого розвитку інформаційна технологія відкритого проксі-сервера - це проксі-сервер пересилання, до якого може отримати доступ будь-який користувач Інтернету.

У 2008 році експерт з мережевої безпеки Гордон Ліон підрахував, що в Інтернеті керують «сотнями тисяч» відкритих проксі.

Анонімний проксі-сервер – цей сервер розкриває свою особистість як проксі-сервер, але не розкриває вихідну IP-адресу клієнта.

Хоча цей тип сервера можна легко виявити, він може бути корисним для деяких користувачів, оскільки приховує вихідну IP-адресу.

Прозорий проксі-сервер – цей сервер не тільки ідентифікує себе як проксі-сервер, але за допомогою полів заголовка HTTP, таких як X-Forwarded-For, також можна отримати вихідну IP-адресу.

Основною перевагою використання цього типу серверів є його здатність кешувати веб-сайт для швидшого пошуку.

Впровадження результатів роботи дозволили зрозуміти небезпеки та переваги використання мульти-проксування в мережі інтернет. Розроблено програмне забезпечення яке надає функції мульти-проксування.

За темою дипломної роботи опублікована одна стаття на фахову наукову конференцію «Сучасні інформаційні технології 2022», що проходила у 19-20 травня 2022. В державному університеті «Одеська Політехніка» (додаток А).

1) Шевченко Р.С. Забезпечення конфіденційності передачі інформації в мережі Інтернет // Міжнародна науково-практична конференція «Сучасні інформаційні технології 2022» ISM–2021 (Одеса, 19-20 травня 2022).

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1) Pogue D. Don't Worry about Who's watching. Scientific American. 2011. Vol. 1. Pp. 304.

2) Yee, C. K. Zolkipli, M. F. Review on Confidentiality, Integrity and Availability in Information Security. *Journal of ICT in Education*. 2021. Vol. 8, P.34-42.

3) Safety for Families and Educators. URL: https://www.web.archive.org.

4) Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*. Pp. 5802–5805.

5) Matt S. Privacy, Privacy, Where for Art Thou Privacy? 2010.

6) Internet Safety for Families and Educators. URL: http://web.archive.org. (13.05.2019).

7) Grimmelmann J. Saving Facebook. *Iowa Law Review*. 2009. Pp. 1137–1206.

8) Mediati, N. The Most Dangerous Places on the Web. *PC World*. 2018. Vol.28(11), Pp. 72–80.

9) Youn, S. *Determinants of Online Privacy Concern and Its Influence on Privacy Protection Behaviors Among Young*. 2013. Vol.1 Pp.27.

10) Larose, R., Rifon, N. J. Promoting i-Safety: Effects of Privacy Warnings and Privacy Seals on Risk Assessment and Online Privacy Behavior. *Journal of Consumer Affairs*. 2017.

11) Yee, C. Zolkipli, M. F. Review on Confidentiality, Integrity and Availability in Information Security. *Journal of ICT in Education*. 2021. Vol.8. P.34-42.

12) Valcke M., De Wever, B. Van Keer, H. Schellens, T. Long-term study of safe Internet use of young children. *Computers & Education* 2016. Vol. 57 (1): Pp. 1292–1305.

13) Maras, Marie-Helen. Internet of Things: security and privacy implications. *International Data Privacy Law* 2014. Vol.5 Pp. 99–104.

14) Larose, R. ChoI, H. Privacy Issues in Internet Surveys. Social Science *Computer Review* Vol.17. Pp. 421–434.

15) Cyphers, B. Gebhart, G. Behind the One-Way Mirror: A Deep Dive to the Technology of Corporate Surveillance. *Electronic Frontier Foundation*. 2015.

16) Kint, Cozen O'Connor-Brian. *What Is A "Reasonable Link" Under CCPA?* 2020. Pp.

17) O'Connor, C. What Is A "Reasonable Link" Under CCPA? URL: http://www.lexology.com

18) Coleman, J. CCPA Clarity in California. ACA International. 2020. Vol.1. Pp.24-26.

19) Garg, A., Mittal N. A security and confidentiality survey in wireless internet of things (iot). *In Internet of Things and Big Data Applications*. 2020. P. 65-88.

20) Grimmelmann, James. Saving Facebook. Iowa Law Review. 2019. P. 1137–1206.

21) Durieux, T.; Hamadi, Y.; Monperrus, M. Fully Automated HTML and JavaScript Rewriting for Constructing a Self-Healing Web Proxy. 2018 IEEE 29th *International Symposium on Software Reliability Engineering* (ISSRE). pp. 1–12.

22) «Trust and Privacy Online: Why Americans Want to Rewrite the Rules». *Pew Internet & American Life Project*. Released Aug. 20, 2000.

23) Kosinski, Michal; Stillwell, D.; Graepel, T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences 2013*. Vol.110. P. 5802–5805.

24) Chanal, P. M., Kakkasageri, M. S. Preserving Data Confidentiality in Internet of Things. *SN Computer Science*, 2021.Vol. 2. P. 1-12.

25) Grimmelmann, James. Saving Facebook. *Iowa Law Review*. 2019. P. 1137–1206.

26) Roya E., David F., Philipp W., Nick F., Nicholas W., and Vern P.. Examining how the great firewall discovers hidden circumvention servers. *In Proceedings of the Internet Measurement Conference*, pages .2015. Pp.445–458.

27) Michael G., Paul F., and David M. Anonymous connections and onion routing. *IEEE Journal on Selected areas in Communications*. Vol.16 Pp. 482–494, 1998.

28) Zhang H., X. Li. A Methodology for Analyzing Backbone Network Traffic at Stream-Level. *Int. Conf. Commun. Technol. Proceedings*, ICCT, 1, 2003.

29) Chadi B., Patrick T., Gianluca I., Christophe D., and Philippe O. Modeling Internet backbone traffic at the flow level. IEEE *Trans. Signal Process*., 2013 Vol.51 Pp.2111–2124.

30) Andrew W. Toward the accurate identification of network applications. *Passiv. Act. Netw. Meas*., 2015. Pp. 41–54.

31) Thomas K., Andre Br., Michalis F., and Kc Claffy. Transport layer identification of P2P traffic. IMC ’04 Proc. 4th ACM SIGCOMM Conf. Internet Meas., 2004. Pp. 121–134.

32) Ruixi Y., Zhu Li, Xiaohong Guan, and Li Xu. An svm-based machine learning method for accurate internet traffic classification. *Information Systems Frontiers*, 2018. Vol. 12(2). Pp. 149–156.

33) Roya Ensafi, David Fifield, Philipp Winter, Nick Feamster, Nicholas Weaver, and Vern Paxson. Examining how the great firewall discovers hidden circumvention servers. *In Proceedings of the 2015 Internet Measurement Conference*, pages 2017. Pp. 445–458. ACM.

34) Zhongjie W., Yue C., Zhiyun Q., Chengyu S., and Srikanth V. *Your state is not mine. Proc. Internet Meas*. Conf. - IMC ’17, 2017. Pp. 114–127.

35) Xueyang X., Morley M., and J. Alex H. Internet censorship in China: Where does the filtering occur? Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 6579 2011 LNCS:133–142.

36) Philipp W., Tobias P., and Juergen F. ScrambleSuit. Proc. 12th ACM Work. Work. Priv. Electron. Soc. - WPES ’13, 2013. Pp. 213–224.

37) Brandon W. Dust: A Blocking-Resistant Internet Transport Protocol. *Defcon*, 2013.

38) Mohajeri M. and B Li. SkypeMorph: *Protocol obfuscation for Tor bridges*. Proc. ..., 2012 Pp. 97–108.

39) Zachary Weinberg, Jeffrey Wang, Vinod Yegneswaran, Linda Briesemeister, Steven Cheung, Frank Wang, and Dan Boneh. StegoTorus : A Camouflage Proxy for the

Tor Anonymity System. Proc. 2012 ACM Conf. Comput. Commun. Secur., 2012. Pp. 109–120.

40) Qiyan W., Giang T K Nguyen, and Borisov N. CensorSpoofer : Asymmetric Communication using IP Spoofing for Censorship-Resistant Web Browsing. Proc. 2012 ACM Conf. Comput. Commun. Secur., 2012., Pp. 121–132.

41) Amir H., Chad Br., and Vitaly S.. The parrot is dead: *Observing unobservable network communications*. Proc. - IEEE Symp. Secur. Priv., 2013. Pp. 65–79.

42) Hui L., Xi Y., and Xinpeng Li. A method and device for tcp connection reset. 2013.

43) Ziye D., Zihan L., Zhouguo C., and Yubin G. The random forest based detection of shadowsock's traffic. *In Intelligent Human-Machine Systems and Cybernetics* (IHMSC), 2017 9th International Conference on. 2017. Vol 2, Pp. 75–78.

44) Qingfeng T., Jingqiao S., and Bingxing F. *Towards measuring unobservability in anonymous communication systems*. 2015.

45) David F., Chang L., Rod H., Percy Wegmann, and Vern Paxson. *Blockingresistant communication through domain fronting*. Proc. Priv. Enhancing Technol., 2015. Pp. 46–64.

46) Marcus L., Matt G., Y Lee, Ron K., David K., and L Jones. *Socks protocol version 5. Technical report*, 2016.

47) Eytan Adar and Bernardo A Huberman. Free riding on gnutella. *First monday*, 2014.

48) Bram C. Incentives build robustness in bittorrent. *In Workshop on Economics of Peer-to-Peer systems*, 2013. Vol 6, Pp.68–72.

ДОДАТОК А
(обов'язковий)

ФРАГМЕНТ ПРОГРАМНОГО КОДУ

```
import (
	"encoding/json"
	"io/ioutil"
	"log"
	"net/http"
	"net/url"
	"os"
	"strings"

	"github.com/esnunes/multiproxy/pkg/admin"
	"github.com/esnunes/multiproxy/pkg/broadcast"
	"github.com/esnunes/multiproxy/pkg/cors"
	"github.com/esnunes/multiproxy/pkg/envs"
	"github.com/esnunes/multiproxy/pkg/unicast"
)

// Environment ...
type Environment struct {
	Name     string `json:"name"`
	Key      string `json:"key"`
	Upstream string `json:"upstream"`
}

// App ...
type App struct {
	ID           int           `json:"id"`
	Name         string        `json:"name"`
	Description  string        `json:"description"`
	Broadcast    []string      `json:"broadcast"`
	Addr         string        `json:"addr"`
	Environments []Environment `json:"envs"`
}

// Config ...
type Config struct {
	Admin  string `json:"admin"`
	Cookie string `json:"cookie"`
	Apps   []App  `json:"apps"`
}

// LoadConfigFromFile ...
func LoadConfigFromFile(p string) (*Config, string, error) {
	var c Config

	rawC, err := ioutil.ReadFile(p)
	if err != nil {
		return nil, "", err
	}

	if err := json.Unmarshal(rawC, &c); err != nil {
		return nil, "", err
	}

	return &c, string(rawC), nil
}
```

```
// PatternFromAddr returns a ServeMux compatible pattern based on the given URL.
func PatternFromAddr(a string) string {
	u, _ := url.Parse(a)

	p := u.Path
	if !strings.HasSuffix(p, "/") {
		p = p + "/"
	}

	return u.Hostname() + p
}

// OriginFromAddr returns a Origin Header compatible value based on the given URL.
func OriginFromAddr(a string) string {
	u, _ := url.Parse(a)

	if u.Scheme == "" || u.Host == "" {
		return ""
	}

	return u.Scheme + "://" + u.Host
}

// ParseEnvironments ...
func ParseEnvironments(envs []Environment) ([]*url.URL, map[string]*url.URL, error) {
	addrs := make([]*url.URL, len(envs))
	rules := map[string]*url.URL{}

	for i, env := range envs {
		u, err := url.Parse(env.Upstream)
		if err != nil {
			return nil, nil, err
		}

		addrs[i] = u
		rules[env.Key] = u
	}

	return addrs, rules, nil
}

func main() {
	log.Printf("multiproxy: Multicast HTTP Reverse Proxy")

	if len(os.Args) < 2 {
		log.Fatal("Usage: multiproxy ./path/to/config.json")
	}

	c, rawC, err := LoadConfigFromFile(os.Args[1])
	if err != nil {
		log.Fatalf("Failed to load config file [%v]: %v", os.Args[1], err)
	}

	mux := http.NewServeMux()

	// admin
	mux.Handle(PatternFromAddr(c.Admin), admin.NewHandler(admin.Options{
		Debug:  false,
		Config: rawC,
	}))

	ch := cors.Cors{Origin: OriginFromAddr(c.Admin)}
```

```
	for _, app := range c.Apps {
		upstreams, rules, err := ParseEnvironments(app.Environments)
		if err != nil {
			log.Fatalf("Failed to parse environments [%v]: %v", app.Environments, err)
		}

		// broadcast
		bh := &broadcast.Handler{
			Addrs: upstreams,
		}
		for _, endp := range app.Broadcast {
			mux.Handle(PatternFromAddr(app.Addr)+endp, bh)
		}

		eh := &envs.Handler{Cookie: c.Cookie}
		mux.HandleFunc(PatternFromAddr(app.Addr)+"_multiproxy", ch.Handler(eh))

		// unicast
		mux.Handle(PatternFromAddr(app.Addr), &unicast.Handler{
			Selector: eh,
			Rules:    rules,
		})
	}

	log.Print("Listening at: 0.0.0.0:8080")
	http.ListenAndServe(":8080", mux)
}
```

ДОДАТОК Б

КОПІЯ СТАТІ НА МІЖНАРОДНУ НАУКОВУ КОНФЕРЕНЦІЮ

Шевченко Р.С. Забезпечення конфіденційності передачі інформації в мережі Інтернет // Міжнародна науково-практична конференція «Сучасні інформаційні технології 2022» ISM–2021 (Одеса, 19-20 травня 2022).

```
package main
```

УДК 004.738.5.057.4

ЗАБЕЗПЕЧЕННЯ КОНФІДЕНЦІЙНОСТІ ПЕРЕДАЧІ ІНФОРМАЦІЇ В МЕРЕЖІ ІНТЕРНЕТ

Шевченко Р.С.
Хмельницький Національний Університет, УКРАЇНА

На сьогодні процеси конфіденційності є складним та трудомісткими. Безпека передачі інформації може бути покращена при розумінні правил конфіденційності в мережі Інтернет. Саме користувач може збільшити надійність передачі інформації в Інтернеті та зберегти свої дані від зловмисників.

Вступ. Інтернет та цифрова конфіденційність розглядаються інакше, ніж традиційні очікування конфіденційності. Конфіденційність в Інтернеті в першу чергу стосується захисту інформації користувачів. Професор права Джеррі Кан пояснює, що термін «конфіденційність» означає простір, рішення та інформацію [1]. З точки зору простору, люди очікують, що в їхні фізичні простори (наприклад, будинки, автомобілі) вторгнення неможливі.

Мета роботи. Огляд проблем забезпечення конфіденційності інформації в мережі Інтернет. Для досягнення поставленої мети слід здійснити аналіз задач для забезпечення конфіденційності передачі інформації в мережі.

Основна частина роботи. Людям, які не надто піклуються про конфіденційність в Інтернеті, не потрібно досягати повної анонімності. Користувачі Інтернету можуть захистити свою конфіденційність шляхом контрольованого розкриття особистої інформації. Розкриття *IP*-адрес, профілювання, що не ідентифікує особу, та подібна інформація може стати прийнятним компромісом для зручності, яку в іншому випадку користувачі можуть втратити, використовуючи обхідні шляхи, необхідні для приховування таких деталей. З іншого боку, деякі люди прагнуть набагато більшої конфіденційності. У цьому випадку вони можуть намагатися досягти анонімності в Інтернеті, щоб не дати жодним третім сторонам можливості пов'язувати діяльність в Інтернеті з особистою інформацією користувача. Щоб зберегти конфіденційність своєї інформації, люди повинні бути обережними з тим, що вони викладають і переглядають в Інтернеті. Під час заповнення форм і купівлі товарів відстежується інформація, і оскільки вона не була приватною, деякі компанії розсилають Інтернет-користувачам спам і рекламу подібних продуктів.

Існує також кілька урядових організацій, які до певної міри захищають конфіденційність та анонімність особи в Інтернеті. У роботі [2] було звернено увагу на ряд вказівок, які допомагають окремому користувачеві Інтернету уникнути можливої крадіжки особистих даних та інших кібератак. Рекомендується, серед іншого, запобігати або обмежувати використання номерів соціального страхування в Інтернеті, бути обережними до електронних листів, включаючи спам, пам'ятати про особисту фінансову інформацію, створювати надійні паролі та керувати ними, а також використовувати безпечні способи перегляду веб сторінок.

Розміщення інформації в Інтернеті може бути шкідливим або піддавати людей до зловмисних атак. Деяка інформація, розміщена в Інтернеті, зберігається десятиліттями, залежно від умов надання послуг та політики конфіденційності окремих послуг, які пропонуються в Інтернеті. Це може включати коментарі, написані в блогах, зображеннях і веб-сайтах, таких як *Facebook* і *Twitter*, тощо. Деякі роботодавці можуть досліджувати потенційного працівника, шукаючи в Інтернеті подробиці їхньої поведінки в Інтернеті, що, можливо, вплине на результат успіху кандидата [3].

Компанії відстежують поведінку користувачів на веб-сайтах, щоб потім використовувати цю інформацію, наприклад, надсилаючи рекламу на основі історії перегляду веб-сторінок. Існує багато способів, за допомогою яких люди можуть розголошувати свою особисту інформацію, наприклад, використовуючи соціальні мережі та надсилаючи інформацію про банківські рахунки та кредитні картки на різні веб сайти.

Крім того, безпосереднє спостереження за поведінкою користувачів, наприклад, за їхніми пошуковими запитами або вмістом профілю Facebook чи інших соціальних мереж, може автоматично оброблятися, щоб зробити висновок про приватні деталі про особу, такі як політичні та релігійні погляди, раса, вживання психотропних речовин, інтелект, і особистість [4]. Користувачі, які стурбовані конфіденційністю в Інтернеті, часто посилаються на низку ризиків конфіденційності (подій, які можуть поставити під загрозу конфіденційність), з якими можна зіткнутися під час онлайн-дій [5]. Вони варіюються від збору статистичних даних про користувачів до більш зловмисних дій, таких як поширення шпигунського програмного забезпечення та експлуатація різних форм помилок (помилки програмного забезпечення).

Кілька веб-сайтів соціальних мереж намагаються захистити особисту інформацію своїх користувачів, а також надають попередження через угоду про конфіденційність та умови користування їхнім веб сайтом [3].

Наприклад у *Facebook* налаштування конфіденційності доступні для всіх зареєстрованих користувачів: вони можуть заблокувати певним особам доступ до свого профілю, вибрати своїх «друзів», а також обмежити доступ до своїх фотографій, відео та іншої інформації з свого профілю. Налаштування конфіденційності також доступні на інших веб-сайтах соціальних мереж, таких як *Google Plus і Twitter*. Користувач може застосувати такі налаштування при наданні особистої інформації в Інтернеті. *Electronic Frontier Foundation* створив набір посібників, щоб користувачам було легше використовувати ці налаштування конфіденційності [6]. Крім того, *Facebook* запустив програму *Beacon*, в рамках якої записи про в користувачів були оприлюднені для друзів. Багато людей були розлючені цим порушенням конфіденційності, і було відкрито судову справу *Lane v. Facebook, Inc* [7].

Діти та підлітки часто використовують Інтернет (включаючи соціальні медіа) у спосіб, який загрожує їх конфіденційності: це викликає дедалі більше занепокоєння серед батьків. Молодь також може не усвідомлювати, що вся їхня інформація може переглядатися та відстежуватися під час відвідування певного сайту, і що вони мають захищати власну конфіденційність [8, 9].

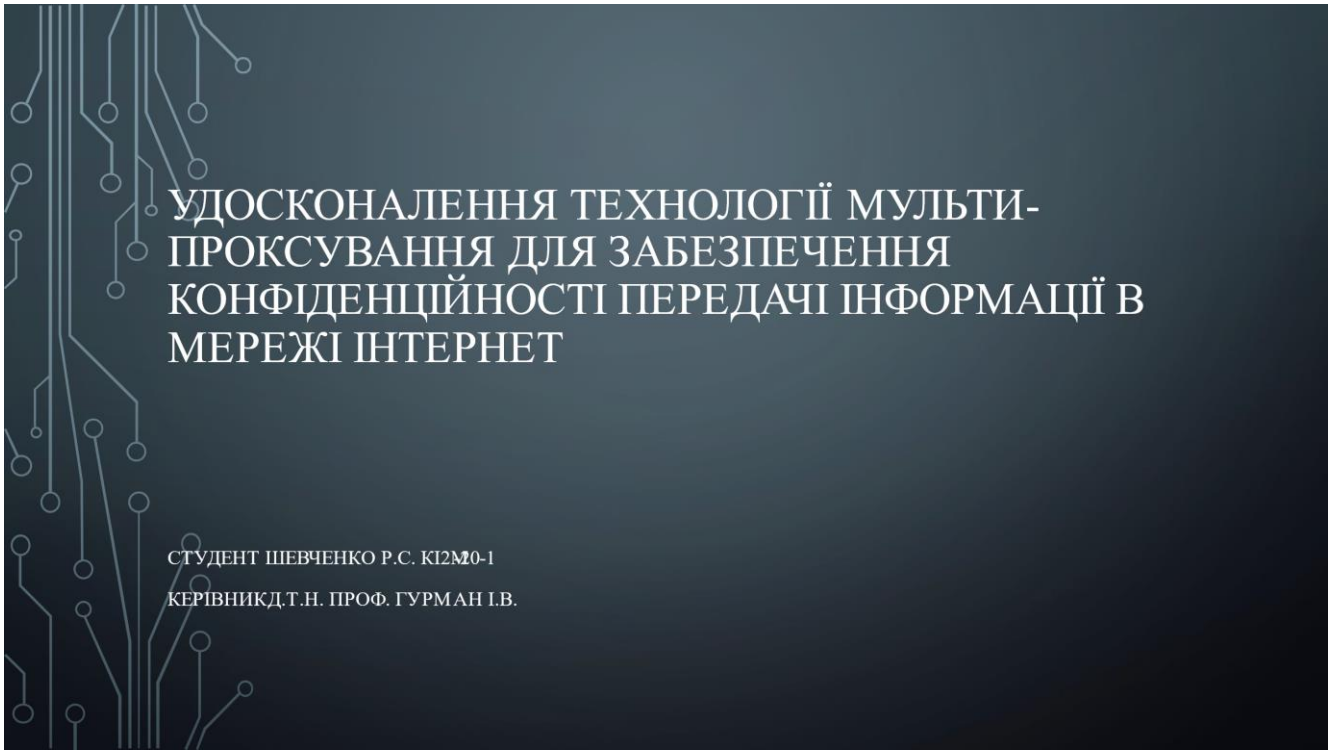
Висновки. Було проведено огляд проблем забезпечення конфіденційності передачі інформації в мережі Інтернет та інформаційних технологій для забезпечення безпеки користувачів. На сьогоднішній день ця проблема не є вирішеною, тому існує необхідність у розробленні нових підходів для забезпечення конфіденційності користувачів в мережі Інтернет, зокрема захисту при передачі приватних даних.

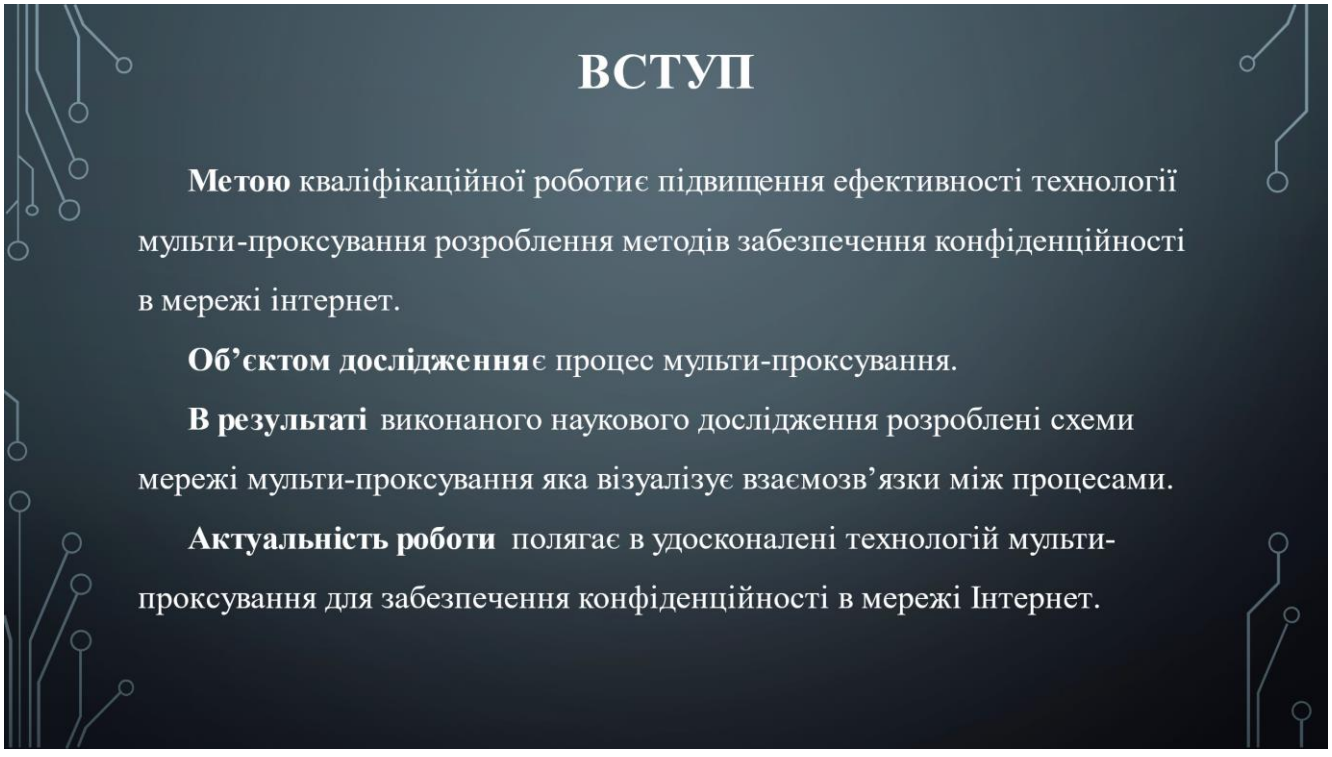
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Yee, C. K., & Zolkipli, M. F. Review on Confidentiality, Integrity and Availability in Information Security. Journal of ICT in Education. 2021. Vol. 8, P.34-42.
2. Yee, C. K., Zolkipli, M. F. Review on Confidentiality, Integrity and Availability in Information Security. Journal of ICT in Education. 2021. Vol.8. P. 34-42.
3. Pogue David. Don't Worry about Who's watching . Scientific American. 2011. P. 304.
4. Garg, A., Mittal, N. A security and confidentiality survey in wireless internet of things (iot). In Internet of Things and Big Data Applications. 2020. P. 65-88.
5. Internet Safety for Families and Educators. URL: http://web.archive.org. (13.05.2019).
6. Kosinski, Michal; Stillwell, D.; Graepel, T. Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences 2013. Vol.110. P. 5802–5805.
7. Chanal, P. M., Kakkasageri, M. S. Preserving Data Confidentiality in Internet of Things. SN Computer Science, 2021.Vol. 2. P. 1-12.
8. Grimmelmann, James. Saving Facebook. Iowa Law Review. 2019. P. 1137–1206.
9. Yee, C. K., Zolkipli, M. F. Review on Confidentiality, Integrity and Availability in Information Security. Journal of ICT in Education. 2021. Vol.8. P.34-42.

ДОДАТОК В

ПРЕЗЕНТАЦІЯ ДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ





НАУКОВА НОВИЗНА ОТРИМАНИХ РЕЗУЛЬТАТІВ:

Наукова новизна полягає в :

1) в розроблені моделі блоку мульти-проксування, фрагменту проксі та вузла, представлені у формалізованому та схематичному вигляді, а також моделі процесу мульти-проксування;

2) розроблено метод оцінювання достатності забезпечення конфіденційності даних;

АНАЛІЗ ВІДОМИХ МЕТОДІВ ТА РІШЕНЬ

Конфіденційність в Інтернеті в першу чергу стосується захисту інформації користувачів.

Людям, які не надто піклуються про конфіденційність в Інтернеті, не потрібно досягати повної анонімності. Користувачі Інтернету можуть захистити свою конфіденційність шляхом контрольованого розкриття особистої інформації.

Розкриття IP-адрес, профілювання, що не ідентифікує особу, та подібна інформація може стати прийнятним компромісом для зручності, яку в іншому випадку користувачі можуть втратити, використовуючи обхідні шляхи, необхідні для суворого придушення таких деталей. З іншого боку, деякі люди прагнуть набагато більшої конфіденційності.

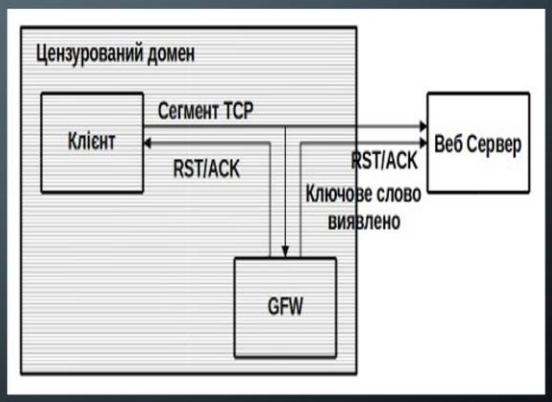
У цьому випадку вони можуть намагатися досягти анонімності в Інтернеті щоб не дати жодним третім сторонам можливості пов'язувати діяльність в Інтернеті з особистою інформацією користувача. Щоб зберегти конфіденційність своєї інформації, люди повинні бути обережними з тим, що вони викладають і переглядають в Інтернеті.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Скидання TCP з'єднання

TCP (Transmission Control Protocol) є надійним протоколом зв'язку в транспортний шар. Він забезпечує кілька механізмів, таких як виявлення помилок, контроль потоку, контроль перевантажень і повторна передача. Однак він не призначений для безпечної передачі даних між двома вузлами, тобто протокол TCP не забезпечує конфіденційність корисного навантаження та аутентифікація ідентифікаторів двох вузлів.

Отже, TCP з'єднання можуть бути легко перехоплені або підроблені зловмисниками. Є різні підходи до переривання TCP-з'єднання, наприклад, атака скидання TCP, заливання SYN атаки та атаки перехоплення сеансу TCP. Простим способом розірвати існуюче з'єднання між клієнтом і сервером є скидання TCP-з'єднання

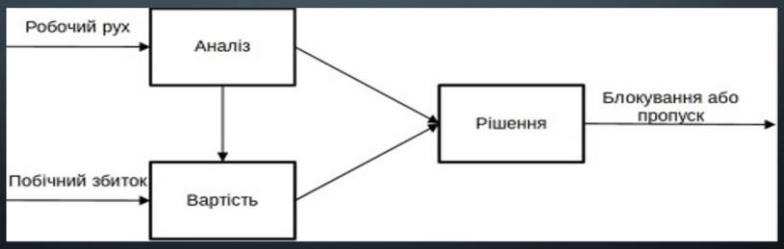


АНАЛІЗ ВІДОМИХ МЕТОДІВ ТА РІШЕНЬ

Інтернет-цензура може відбуватися в стеку TCP/IP як кінцевих точок, так і в шлях між ними. Таким чином, цензура може бути класифікована на цензуру на стороні клієнта, цензура на стороні сервера, цензура на шляху та цензура в шляху.

Цензура на стороні клієнта означає, що користувачі можуть отримати лише обмежений обсяг ресурсів через вбудованих функціях у програмах цензури, таких як мережеві фільтри.

Загальна модель цензури



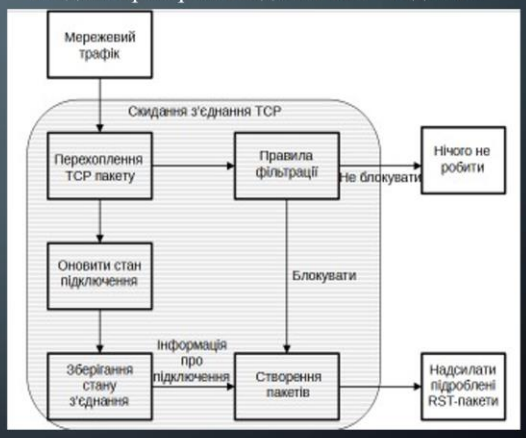
МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Модуль захоплення збирає всі пакети TCP із мережевого трафіку в реальному часі.

Модуль оновлення оновлює форму з'єднання, що зберігається в базі даних.

Тим часом модуль прийняття рішень приймає рішення чи слід виконувати правила запобігання вторгненню та використовувати модуль блокування для створення підроблених пакетів після обчислення перед відправленням цілям.

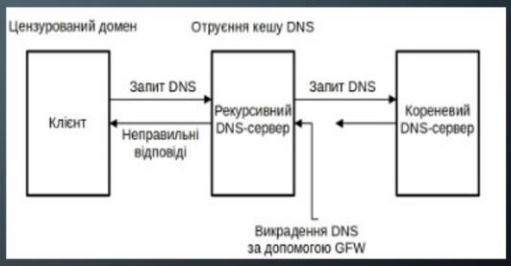
Модель пристрою скидання TCP -з'єднання



МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Для того, щоб обійти цензуру, одним із найпопулярніших контрзаходів є надсилання запитів до проксі вузла і цей вузол надсилає заблоковану інформацію назад. Успіх систем протидії цензурі на основі проксі на практиці призвів до того, що цензори почали розгортати передові механізми глибокої перевірки пакетів (DPI) який може ідентифікувати трафік на основі інформації на прикладному рівні, а також поведінка мережевого потоку між двома кінцевими точками. На основі технології DPI фільтрація за ключовими словами на основі URL-адреси може застосовуватися до маршрутизаторів, через які проходять пакети. Таким чином, цензори можуть аналізувати та змінювати мережевий трафік

Викрадання та отруєння DNS



МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Маршрутизація трафіку

Проксі-сервер клієнта може бути автором запиту, а також може діяти як вузол ретрансляції, який використовується для пересилання у мережі трафіку, коли схеми вбудовані в цибульну маршрутизацію .

Як система обходу, основна функціональність — це переадресація трафіку, тому клієнтський вузол повинен пересилати свій трафік до вузлів, які надають відповідну послугу обходу.

Для цього потрібен протокол пересилання мережевих даних між кількома додатками та системою MultiProxy на одному хості.

Для передачі даних між вузлами повідомлення мають бути тунельовані та зашифровані, щоб уникнути використання різних методів аналізу мережевого трафіку.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

КОМПОНЕНТИ СИСТЕМИ МУЛЬТИ-ПРОКСІ

Функціональність	Ціль	Вимоги
Переадресація трафіку	Надати базову послугу обходу	Протокол SOCKS5 Шифрування даних
Токенова економіка	Зробити систему надійною	Протокол Trustchain Intel SGX, SCONE
Повідомлення з кількома переходами	Захист конфіденційності для авторів запитів	Створення схем Тунелювання даних

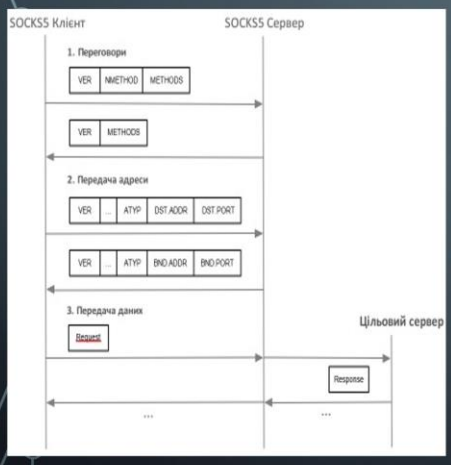
Мульти-проксі забезпечує анонімну маршрутизацію цибулі зв'язок через мережу.

Мульти-проксі використовує цибульну маршрутизацію до захищати особистість і конфіденційність авторів запиту.

Повідомлення інкапсулюються в кожному вузлі, через який він проходить, таким чином, що проміжні вузли не можуть знати джерело та призначення повідомлень .

Цей механізм вимагає протоколу обміну ключами криптографії між ланцюгами.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ



Це досягається шляхом додавання проксі-шару поверх традиційної архітектури клієнт-сервер. Рівень проксі містить два основних компоненти: клієнтський вузол і вузол сервера.

Клієнтський вузол використовується для отримання даних з оригінальних програм, наприклад, браузерів.

Цей клієнтський вузол можна розглядати як розширення програми, оскільки він може змінювати зашифровані дані з вихідних програм, або змінити адресу призначення одержувача.

Щоб переслати заявку даних на мульти-проксі, потрібен відповідний протокол передачі даних. SOCKS5 використовується як протокол передачі даних між додатками та клієнтським вузлом.

SOCKS — це діючий протокол обміну повідомленнями між прикладним і транспортним рівнем.

Він розроблений для прозорого і безпечного проходу через брандмауер.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Причини, чому мульти-проксі використовує протокол SOCKS 5 між додатками а клієнтський проксі завдяки своїй гнучкості та розширюваності.

Оскільки SOCKS 5 ні підтримує лише протокол транспортного рівня, наприклад TCP та UDP, але також може пересилати кілька протоколів прикладного рівня, такі як HTTP і HTTPS. У системі мульти-проксі вузол сервера розташований у нецензурному місці домену та чекає даних від клієнтського вузла через захищений канал для відправлення у безкоштовний Інтернет.

Метою є дешифрування, реконструкція або зміна даних перед пересиланням на вказану адресу призначення, наприклад на веб-сервер заблокований системою цензури. Після того, як вузол сервера отримає дані від віддаленого веб-сайту, він перенаправить трафік назад на клієнтський вузол.

Рівень проксі працює належним чином, оскільки сервер можна занести в чорний список. Щоб зробити систему більш масштабованою, тобто клієнтський вузол може використовувати кілька серверних вузлів для пересилання свого трафіку.

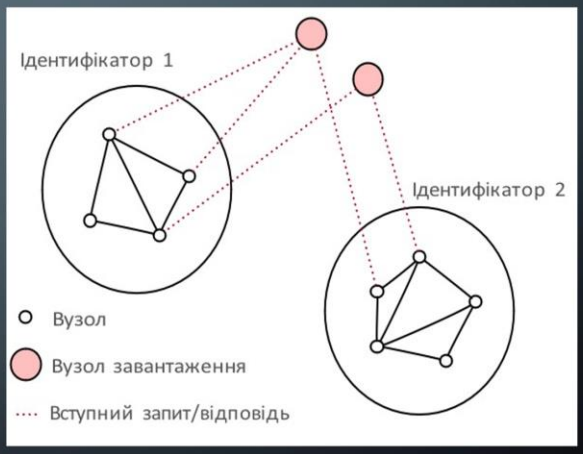
Оскільки проксі-сервери Shadowsocks зазвичай розташовуються в публічних хмарах, цензура не може заблокувати цілі домени, оскільки це призведе до величезної побічної шкоди. Об’єднавши всі проксі SOCKS в єдину спільну інфраструктуру, система стає більш стійкою до блокування окремих серверів, ніж у а одностороння установка. Крім того, кілька серверних вузлів можуть покращити пропускну здатність.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

З цих причин мульти-проксі побудовано на основі однорангової мережі. Вузли однорангової мережі відрізняються від традиційних серверів, вони мають фіксовані та відомі IP-адреси. Подібно до людських стосунків реальному світі, хтось приєднується до мережі, представивши учасника, який є вже є частиною мережі. Принаймні цього вимагає спосіб однорангового приєднання підключається до вузла, який уже є в цій мережі.

На практиці статус вузлів динамічно змінюється, так що вузол, що знову приєднався, не може відразу знайти однорангового. З цієї причини деякі вузли завантаження, налаштовані для надання початкової конфігурації одноранговоїмережі, наприклад IP адреси та ідентифікатори послуг.

Завантаження однорангової системи



МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Токенова економіка

Нерегульована система може працювати на практиці або досягти стабільного стану лише за умови інтересу між сторонами, які надають послуги, та тими, хто споживає послуги, збалансовані. У цьому випадку мульти-проксі працює надійно, лише якщо люди в домені без цензури готові поділитися своїми ресурсами та в нести їх у звичайний обхід.

У мульти-проксування головне завдання – збалансувати кількість серверів і клієнтів . Зараз багато людей вже запускають свої сервери Shadowsocks у публічних доступних хмарах.

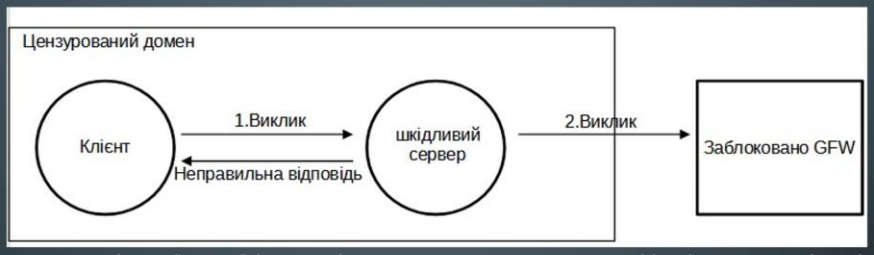
Це означає новий механізм стимулювання людей налаштовувати свій сервер у систему мульти-проксі, коли вони хочуть використовувати її, щоб уникнути проблеми фрирайдерів

Проблема фрирайдерів є значною загрозою для рівноправних систем таких систем, як BitTorrent в які однорангові користувачі лише завантажують файли, не завантажуючи нічого, що є несправедливим для однорангових які сприяють розвитку мережі і можуть мати негативний вплив на систему.

Рішення для оптимізації швидкості завантаження в BitTorrent полягає в тому, щоб використовувати принцип «ти за око», який бере початок з високоефективної стратегії теорії ігор.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

Механізм відповіді на виклик



Якщо вузол сервера вирішує бути обхідним, він повинен завершити автентифікацію виклик-відповідь, до якої звикли визначити, чи правильно вузол сервера виконує свою роль, у цьому випадку відправити назад вміст деяких зазначених заблокованих веб-сайтів.

Шкідливий вузол сервера також може обманювати в майнінгу, щоб отримати більше токенів, наприклад, у моделі майнінгу на основі часу зловмисний сервер може налаштувати систему годинник швидше.

У моделі пропускної здатності або мережевого трафіку серверу це легко зробити підробити пропускну здатність або пакети, які він передає.

Інша загроза полягає в тому, що серверний вузол може маніпулювати даними, які надходять у клієнтському вузлі.

МОДЕЛЮВАННЯ МЕТОДУ МУЛЬТИ-ПРОКСУВАННЯ

ПОВІДОМЛЕННЯ З КІЛЬКОМА ПЕРЕМІЩЕННЯМИ

Одним із способів захисту конфіденційності є обмеження знань про вузли, тобто жоден вузол не знає повну інформацію про мережу. Повідомлення з кількома переходами маршрутизація є контрзаходом для захисту конфіденційності, оскільки вона запобігає ідентифікації автори запиту.

Наскрізне анонімне спілкування для захисту конфіденційності використовує анонімне спілкування.

Цього можна досягти шляхом побудови ланцюгів передачі даних, тобто вузлів, розташованих по-різних шляхах.

Маршрутизація цибулі — це спосіб послідовного обгортання та шифрування повідомлень на рівнях через мережу, яка має кілька проміжних вузлів.

Вона може добре захистити конфіденційність автора, оскільки повідомлення шифруються, надсилаються між ними і жоден проміжний вузол всередині схеми не може визначити джерело та кінцеве призначення повідомлення, крім вихідного вузла.

Вузли сервера оголошуються як вихідні вузли і стають останніми переходами для пересилання даних на цільовий веб-сервер. Після того, як ланцюг буде побудовано, дані будуть передані всередину схеми.

ВИСНОВКИ

У роботі за результатами виконаних теоретичних та практичних досліджень розроблено середовище на основі мульти-проксування. В сучасні межежіє багато небезпек для користувача та фільтрів контенту.

Обхід фільтрів не є причиною того, чому були винайдені проксі. Основна перевага їх використання полягає в тому, що вони роблять вас більш анонімними.

Конфіденційність – це дуже важливе завдання не тільки для людей, які не хочуть мати проблем з правоохоронними органами, а й для звичайного користувача.

Незахищені канали зв'язку можуть пошкодити персональні дані користувача, до яких можуть отримати доступ як провайдер, так і зловмисник.

В роботі я дійшов висновку що, одним з найбільш надійних засобів анонімізації є проксі-сервер.

ВИСНОВКИ

У користувачів в інтернеті є свій, унікальний IP адрес, який занесений в базу даних провайдера. Тому щоб залишатися в мережі анонімно, людина може придбати новий IP адрес, того регіону в якому проживає або будь-якого іншого. Тому головна відмінність проксі від VPN сервісів, полягає в тому, що проксі-сервер не шифрує вхідні та вихідні дані.

Отримати всі перераховані вище переваги, можна і не використовуючи додаткове ПО на комп'ютері. Набравши в адресному рядку браузера слово «Анонімайзер», користувачі побачать посилання на ресурси, що пропонують послуги анонімізації в інтернеті.

Підводячи підсумки, слід зазначити, що всі перераховані вище послуги не можуть гарантувати стовідсоткову безпеку даних. Навіть Tor браузер має свої вади і неграмотне використання, може видати людини.

ВИСНОВКИ

У першому розділі було досліджено вплив інтернет цензури на користувачів та виявлена її небезпека

Було наведено сучасні безпечні обхідні системи які допоможуть користувачам безпечніше користуватися мережею інтернет. Також були досліджені механізми аналізу та блокування для кращого розуміння роботи цензури

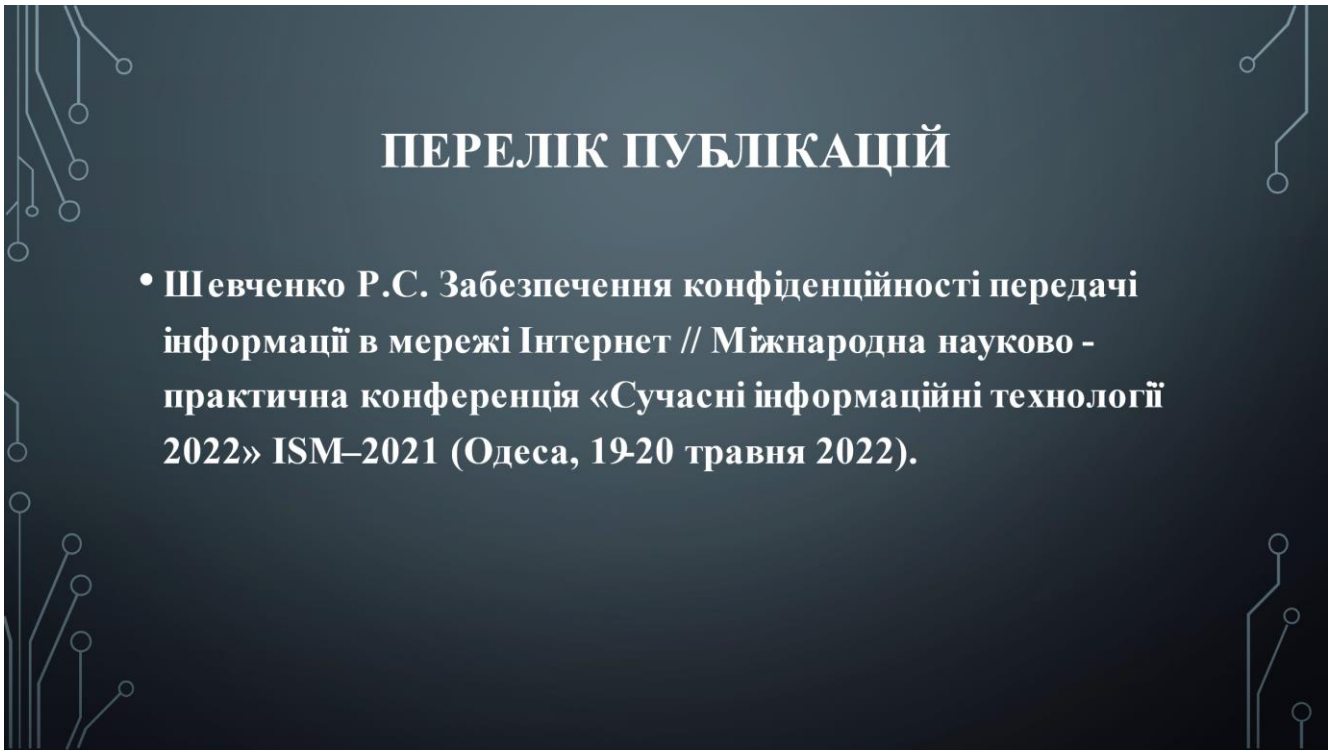
У другому розділі експериментально проведені досліди роботи інтернет цензури та її наслідки для користувачів Інтернет цензура характеризується як в контролі та припиненні публікацій або доступу до інформації в мережі Інтернет. Своєю появою інтернет-цензура зобов'язана відсутності будь-яких національних кордонів в мережі Інтернет.

Загальну проблематику інтернет-цензури можна визначити таким чином: інформація, що порушує закони держави (режим чинного уряду) та заблокована на внутрішніх ресурсах, може бути опублікована на веб серверах в інших країнах.

ВИСНОВКИ

У третьому розділі дослідили маршрутизацію трафіку та її дизайн. Що дозволило краще розуміти архітектуру системи мульти-проксі та її переваги над іншими системами забезпечення безпечної передачі інформації в мережі Були наведені компоненти системи мульти-проксі які дозволяють надавати послуги обходу та роблять систему надійною. Була досліджена схема маршрутизації цибулі. Технологія анонімного обміну інформацією через комп'ютерну мережу. Повідомлення неодноразово шифруються і потім відсилаються через кілька мережевих вузлів, званих цибулевими маршрутизаторами. Кожен маршрутизатор видаляє шар шифрування, щоб відкрити трасувальні інструкції, і відіслати повідомлення на наступний маршрутизатор де все повториться. Таким чином проміжні вузли не знають джерело, пункт призначення і зміст повідомлення

У четвертому розділі була досліджена ефективність на практиці в розробленій програмі яка надає функції мульти-проксування. Було доведено що кешування проксі-сервера прискорює запити на обслуговування, витягуючи вміст, збережений з попереднього запиту, зробленого тим же клієнтом або навіть іншими клієнтами.


ПЕРЕЛІК ПУБЛІКАЦІЙ
• Шевченко Р.С. Забезпечення конфіденційності передачі інформації в мережі Інтернет // Міжнародна науково - практична конференція «Сучасні інформаційні технології 2022» ISM–2021 (Одеса, 19-20 травня 2022).

UNICHECK

Ім'я користувача:
Кафедра КІ

ID перевірки:
1011225215

Дата перевірки:
17.05.2022 19:23:50 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
17.05.2022 19:27:18 EEST

ID користувача:
100005591

Назва документа: **Shevchenko Удосконалення технології мульти-проксування для забезпечення конфіденційн...**

Кількість сторінок: **82** Кількість слів: **15332** Кількість символів: **118047** Розмір файлу: **1.67 MB** ID файлу: **1011116709**

14.2%
Схожість

Найбільша схожість: **12.3%** з Інтернет-джерелом (https://uk.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D1%84%D1%

13.2% Джерела з Інтернету	18	Сторінка 84
1.01% Джерела з Бібліотеки	48	Сторінка 84

0.1% Цитат

Цитати	1	Сторінка 85

Не знайдено жодних посилань

0%
Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи	2

Anti-Plagiarism v-15.257

Максимальное совпадение с одним документом 0.0%

Словари проверки: en_US, ru_RU, ua_UA. **Ошибок в документах: 10%**

ID: 103583 Название: Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет Добавлено в БД: 2022-05-17 Авторы: Шевченко Р.С. Руководители: Гурман І.В. Консультанты: Опоненты:	Документ		Суммарное совпадение по Базе Данных	
	Символы	Лексемы	Символы	Лексемы
	105639	795	677 (1%)	9 (1%)

Источник плагиата

ID	Описание	Наличие плагиата в документе	
		Символы	Лексемы

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА ДИПЛОМНУ РОБОТУ

Дипломник: Шевченко Роман Сергійович

Тема: Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет

Спеціальність: 123 «Комп'ютерна інженерія»

Обсяг дипломної роботи:

Кількість листів креслень ___; кількість сторінок записки 85

1. Короткий зміст роботи та прийнятих рішень В дипломній роботі розроблено моделі блоку мульти-проксування, фрагменту проксі та вузла, представлені у схематичному вигляді, а також моделі процесу мульти-проксування. На основі представленої моделі удосконалено метод переадресації, який грунтується на токеновій економіці. На базі удосконаленого методу розроблені схеми мережі мульти-проксування, що візуалізує взаємозв'язки між процесами.

2. Висновок про відповідність роботи дипломному завданню Дипломна робота відповідає виданому завданню як в теоретичній, так і в практичній частині.

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: Розділ 1 – проведено аналіз видів та основних характеристик цензури в мережі Інтернет. Розділ 2 – удосконалено модель процесу виявлення з'єднання в комп'ютерних системах. Розділ 3 – удосконалено метод переадресації, який грунтується на токеновій економіці. Розділ 4 – представлено результати роботи розробленого методу та досліджено його ефективність. В загальному усі розділи відповідають завданню.

4. Позитивні сторони роботи: Розроблено метод підвищення анонімності користувача в мережі Інтернет, що надає додаткові можливості анонімізації.

5. Негативні сторони роботи: Надмірна кількість теоретичного матеріалу. В роботі відсутня формалізація моделі процесу виявлення з'єднання в комп'ютерних системах.

6. Оцінка графічного оформлення та пояснювальної записки роботи Пояснювальна записка відповідає нормам оформлення та виконана на задовільному рівні.

7. Відгук про роботу в цілому: В загальному дипломна робота заслуговує задовільної оцінки. Дипломна робота присвячена вирішенню актуальної задачі удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет.

8. Інші зауваження:

9. Оцінка дипломної роботи: Розглянувши позитивні та негативні сторони представленої дипломної роботи, можна зробити висновок, що вона заслуговує на оцінку «задовільно».

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи) Мартинюк Валерій Володимирович, д.т.н., професор, завідувач кафедрою Автоматизації, комп'ютерно-інтегрованих технологій і телекомунікацій

"18" травня 2022 р. _____(підпис)

Завідувачу кафедри КІІС
д-р.техн.наук, проф. Говорущенко Т. О.

Шевченка Романа Сергійовича
ПІБ здобувача вищої освіти

ФПКТС, 2 курсу, групи КІ2м-19-1

ЗАЯВА

З правилами чинного Положення «Про дотримання академічної доброчесності в Хмельницькому національному університеті» від 26.09.2020 (зі змінами від 26.11.2020), згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на плагіатоповіщений (а) та надаю свою згоду на обробку та збереження університетом моєї роботи в інституційному репозитарії університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-технічних засобів (Unicheck та Anti-Plagiarism) та використання роботи для виявлення плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення плагіату в текстах робіт.

Робота для перевірки університетом надається в друкованому та електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

16.05.2022р.
дата

підпис

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ
КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОМАЦІЙНИХ СИСТЕМ
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Удосконалення технології мульти-проксування для забезпечення конфіденційності передачі інформації в мережі Інтернет

Автор: Шевченко Роман Сергійович

Спеціальність: 123 – Комп'ютерна інженерія

Освітня програма: освітньо-наукова

Науковий керівник: Гурман Іван Васильович, к.т.н, доцент

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту.	відповідає
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

1) запозичення розміщені в розділах аналізу існуючих аналогів та прототипів, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
2) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ ідентичності/схожості, складає 14,2% і адресується до 66 першоджерел, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Максимальний збіг з одним джерелом складає 12,3% (https://uk.wikipedia.org/wiki) та стосується розділів, у яких проводиться опис предметної області у напрямку дослідження.

Керівник роботи _____ І. В. Гурман

Гарант ОП _____ О. С. Савенко

Завідувач кафедри КІІС _____ Т. О. Говорущенко