Хмельницький національний університет
Факультет інформаційних технологій
Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

Тростянецького Назара Олексійовича

на здобуття ступеня вищої освіти Магістра

Метод виявлення атака типу "Блокування IP через NAT"
в публічних мережах

Галузь знань 12 – Інформаційні технології
Спеціальність 125 – Кібербезпека та захист інформації
Освітня програма Кібербезпека та захист інформації

Шифр КРМКБЗІ. 240198.24.01.13 ПЗ

Виконав студент 2 курсу група КБЗІм-24-1 Назар ТРОСТЯНЕЦЬКИЙ
Керівник канд.техн.наук, доцент Юрій КЛЬОЦ
Нормоконтролер д-р філософії, старший викладач Наталія ПЕТЛЯК

До захисту допускаю:
Завідувач кафедри кібербезпеки Юрій КЛЬОЦ
16 12 2025 р.

Хмельницький 2025

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет Інформаційних технологій

Кафедра Кібербезпеки

Рівень вищої освіти Бакалавр

Галузь знань 12 – Інформаційні технології

Спеціальність 125 – Кібербезпека та захист інформації

Освітня програма Кібербезпека та захист інформації

ЗАТВЕРДЖУЮ

Завідувач кафедри кібербезпеки

Юрій КЛЬОЦ

16 12 2025 р.

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Тростянецькому Назару Олексійовичу

1 Тема роботи Метод виявлення атак типу "Блокування IP через NAT" в публічних мережах

Керівник роботи канд.техн.наук, доцент Юрій Кльоц

Затверджено наказом ректора університету від 25 08 2025 № 65

2 Строк подання студентом кваліфікаційної роботи на кафедру 1.12.2025

3 Вихідні дані до роботи Проаналізувати механізми блокування IP-адрес у зовнішніх сервісах та вплив NAT/CGNAT на точність таких рішень. Дослідити ознаки мережевих аномалій і поведінкові особливості NAT-пулу. Підготувати та зібрати набори даних (NAT-таблиці, conntrack-логи, pcap, flow). Розробити метод виявлення атаки «блокування IP через NAT» і створити прототип системи. Оцінити ефективність методу на тестових сценаріях.

4 Зміст пояснювальної записки (перелік питань, які потрібно розробити)
Вступ. Аналіз механізмів блокування IP і NAT-аномалій. Постановка задачі. Огляд даних і методів їх застосування у NAT-середовищі. Формування та збір NAT-логів. Розробка тестового середовища. Створення методу виявлення атаки «блокування IP через NAT»: архітектура, ознаки, алгоритм, прототип. Оцінка ефективності та результати експериментів. Висновки.

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

6 Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7 Дата видачі завдання 1 09 2025 р.

КАЛЕНДАРНИЙ ПЛАН

Назва етапів (розділів) кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
Ґрунтовне ознайомлення та дослідження предметної галузі	Лютий	Виконано
Визначення змісту, структури магістерської роботи	Березень	Виконано
Опрацювання першого розділу магістерської роботи	Квітень	Виконано
Опрацювання статті за результатами дослідження	Травень	Виконано
Опрацювання другого розділу магістерської роботи	Червень	Виконано
Опрацювання третього розділу магістерської роботи	Вересень	Виконано
Підготовка та опрацювання ілюстративного матеріалу	Жовтень	Виконано
Оформлення магістерської роботи графічної та текстової частини	Листопад	Виконано
Попередній захист магістерської роботи	Грудень	Виконано
Захист магістерської роботи на засіданні ЕК	Грудень	Виконано

Студент _____ Назар ТРОСТЯНЕЦЬКИЙ

Керівник кваліфікаційної роботи _____ Юрій КЛЬОЦ

АНОТАЦІЯ

Тема кваліфікаційної роботи: Метод виявлення атаки типу «блокування IP через NAT»

Автор роботи: студент групи КБЗІм-24-1 Тростянецький Н.О.

Керівник роботи: канд.техн.наук, доцент Кльоц Ю.П.

Загальний обсяг роботи: 100 сторінок, 8 рисунків, 1 таблиця, 4 формули, 80 посилань, 1 додаток.

Ключові слова: NAT, CGNAT, блокування IP-адрес, виявлення аномалій, WAF, машинне навчання, мережеві атаки, мережеві логи.

Кваліфікаційна робота присвячена проблемі некоректного блокування IP-адрес у зовнішніх сервісах унаслідок використання NAT та CGNAT, коли багато користувачів поділяють одну публічну адресу. У роботі проаналізовано механізми блокування, особливості мережевих аномалій та вплив адресного шарингу на точність рішень систем безпеки. Запропоновано метод виявлення зловмисної активності за NAT, що базується на аналізі часових, транспортних, кореляційних та ентропійних ознак. Розроблено архітектуру системи збору й обробки логів та створено тестове середовище для моделювання NAT-сценаріїв. Проведені експерименти підтвердили здатність методу точно локалізувати джерело аномалій усередині NAT-пулу та зменшувати кількість хибних блокувань. Результати роботи демонструють практичну ефективність запропонованого підходу для підвищення точності рішень систем мережевої безпеки.

01.12.2025 р.

ANNOTATION

Theme of qualification work: Method for Detecting the "IP Blocking through NAT" Attack

Author of the work: KBZIm-24-1 Trostianetskyi N.O.

Mentor: Candidate of Technical Sciences, Associate Professor Klots Y.P.

Total volume of work: 100 pages, 8 figures, 1 table, 4 formulas, 80 references, 1 appendix.

Keywords: NAT, CGNAT, IP address blocking, anomaly detection, WAF, machine learning, network attacks, network logs.

The thesis addresses the problem of incorrect IP blocking in external security systems caused by the use of NAT and CGNAT, where multiple users share a single public IP address. The work analyzes blocking mechanisms, characteristics of network anomalies, and the impact of address sharing on the accuracy of security decisions. A method for detecting malicious activity behind NAT is proposed, based on the analysis of temporal, transport, correlation, and entropy-based features. The architecture of a system for log collection and processing was developed, as well as a test environment for modeling NAT scenarios. Experiments confirmed the method's ability to accurately localize the source of anomalies within a NAT pool and reduce false-positive blocking. The results demonstrate the practical effectiveness of the proposed approach in improving the accuracy of decisions made by network security systems.

01.12.2025р.

ЗМІСТ

ВСТУП

У наш час питання доступності та безпеки мережевих ресурсів набувають дуже серйозного значення. З одного боку, широке розповсюдження сервісів CDN, анти-DDoS рішень та централізованих провайдерських фільтрів, значною мірою підвищило стійкість веб-інфраструктури до масових атак і зловмисних дій. З іншого боку - стрімке зростання кількості пристроїв і обмеженість IPv4-простору призвели до широкого використання технологій трансляції адрес (NAT, зокрема Carrier-Grade NAT - CGNAT), через це - багато користувачів опиняються за однією «публічною адресою». Така архітектурна специфіка породжує нову, вже класичну проблему - помилкове або колективне блокування публічної IP-адреси (shared IP), за якою можуть стояти десятки або сотні користувачів це, в свою чергу, може призвести до масштабних відмов у доступі для невинних користувачів.

Особливо актуальною ця проблема є для публічних мереж (гуртожитки, університетські мережі, кафе, мобільні мережі), де політика блокування, що базується на агрегації сигналів на рівні публічної IP, не дозволяє розрізняти реального зловмисника і невинних користувачів. З боку зовнішніх сервісів рішення про блокування зазвичай ухвалюється на основі швидко оцінюваних ознак, таких як - аномально високий обсяг з’єднань, підозрілі шаблони запитів, скарги/репорти, або репутаційні бази. У поєднанні з NAT це призводить до підвищеного ризику false positive - помилкових спрацьовувань, що мають істотні економічні та соціальні наслідки для провайдерів і користувачів відповідно.

Мета цієї роботи - розробити метод виявлення атак типу «блокування IP через NAT» у публічних мережах, тобто підхід, що дозволяє коректно ідентифікувати випадки, коли рішення про блокування публічної IP-адреси є наслідком поведінки одного або кількох внутрішніх хостів за NAT, та сприяти більш тонкій і диференційованій реакції, що мінімізує шкоду для тих користувачів, які ні в чому не винні. Розроблений метод має поєднувати аналіз доступних мережевих логів, статистичні та евристичні правила і, за потреби, машинне навчання для формування скоригових індикаторів підозрілої активності на рівні public IP та

внутрішніх mappings NAT.

Для досягнення поставленої мети у роботі вирішуються такі завдання:

– проаналізувати сучасні підходи до виявлення мережевих аномалій і механізми, що використовуються зовнішніми сервісами (CDN, анти-DDoS, провайдерські фільтри) при ухваленні рішень про блокування IP-адрес;

– вивчити специфіку NAT/CGNAT та її вплив на ідентифікацію зловмисної активності й на ризики помилкового блокування;

– розробити набір ознак (features) і евристик, які дозволяють відрізняти «атаки, що спричиняють блокування» від нормального або змішаного трафіку в умовах shared IP;

– реалізувати прототип детектора для збору та аналізу логів (web-сервер, conntrack/iptables, pcap) та оцінити його роботу у контрольованому тестовому середовищі;

– провести експериментальну оцінку ефективності запропонованого методу за метриками детекції (TP, FP, precision/recall, time-to-detect) та впливу на benign-користувачів.

Наукова новизна дослідження полягає в комплексному підході до задачі, поєднанні аналізу NAT-специфічних метрик (наприклад, розподілу портів, кількості внутрішніх mappings, ентропії user-agent), корелювання цих ознак з поведінковими характеристиками трафіку і розробці адаптивного механізму реагування, що рекомендує локально ізолювати підозрілий внутрішній хост або застосувати challenge-response заходи замість негайного глобального блокування public IP. Такий підхід дозволяє зменшити кількість помилкових блокувань і підвищити загальну доступність сервісів для клієнтів публічних мереж.

Методи дослідження включають огляд наукової та технічної літератури, аналіз конфігурацій NAT і логів мережевих пристроїв, побудову тестового стенду із віртуалізованими клієнтами і NAT-шлюзом (Mininet/GNS3 або Linux-NAT), а також застосування статистичних та машинно-навчальних алгоритмів для класифікації аномалій. Експериментальна частина передбачає моделювання сценаріїв навантаження та атак у контрольованому середовищі й оцінку

запропонованого методу за визначеними метриками.

Структура роботи побудована відповідно до логіки наукового дослідження та включає кілька взаємопов’язаних розділів. У першому розділі виконано огляд наукових публікацій, стандартів і практичних підходів у сфері мережевої безпеки, а також сформульовано постановку задачі, визначено мету, завдання та обґрунтовано актуальність дослідження проблем блокування трафіку в умовах використання NAT і CGNAT.

Другий розділ присвячено аналізу існуючих наборів мережевих даних, визначенню їх обмежень у контексті адресного шарингу та опису тестового середовища для формування контрольованих датасетів і моделювання типових NAT-сценаріїв. Розглянуто підходи до збору, синхронізації та обробки мережевих журналів і метрик.

У третьому розділі розроблено метод виявлення зловмисної активності в NAT-середовищах, описано його архітектуру, основні рівні обробки даних і використовувані ознаки, а також підходи до прийняття рішень на основі евристичних і машинно-орієнтованих методів.

Четвертий розділ містить експериментальну оцінку запропонованого методу, опис сценаріїв тестування, використаних метрик та аналіз отриманих результатів. У завершальній частині роботи сформульовано висновки та практичні рекомендації щодо впровадження запропонованого підходу. Результати дослідження можуть бути корисними для інтернет-провайдерів, адміністраторів мереж, розробників систем захисту та науковців, які працюють над підвищенням стійкості інтернет-інфраструктури в умовах масового використання NAT.

1 АНАЛІЗ БЛОКУВАННЯ IP ЗОВНІШНІМИ СЕРВІСАМИ В УМОВАХ NAT

1.1 Механізми блокування IP зовнішніми сервісами

У практиці захисту інтернет-ресурсів зовнішні сервіси такі як - CDN, анти-DDoS платформи, WAF, провайдерські фільтри, репутаційні бази використовують комбінацію автоматичних детекторів, ручних процедур і правил, щоб швидко зупиняти підозрілу та шкідливу активність. Рішення «заблокувати IP» може прийматися на основі різних сигналів - від чистої статистики (занадто високий трафік) до зовнішніх скарг або репутаційних списків.

Важливо розрізняти два рівні:

– технічні механізми блокування та/або пом’якшення (rate-limiting, WAF-правила, black/gray/white-lists, challenge-response);

– критерії/сигнали, які ці механізми використовують для ухвалення рішення (аномалії трафіку, підозріла поведінка, репутація, скарги користувачів, політика клієнта).

Далі я детально описав ці механізми і типові тригери, акцентуючи увагу на те, як і на яких підставах зовнішні сервіси приймають рішення блокувати IP-адресу й чим це загрожує у ситуації з використанням shared-IP / NAT.

Є різні типи технічних механізмів блокування та пом’якшення. Доприкладу Rate-limiting [1] або throttling. З допомогою цих сервісів, можна встановити пороги запитів (requests per second / minute) до певних endpoint-ів і вони будуть діяти автоматично, після перевищення порога - блок/тротлінг або тимчасовий тимчасове повернення у якості відповіді таких статус кодів - 429/403. Наприклад, CDN/WAF дозволяють конфігурувати правила лімітування, і при спрацьовуванні блокуються всі запити конкретно з того публічного IP до хоста або ресурсу, доки ліміт не скинеться чи правило не буде змінено. Це швидкий та досить ефективний спосіб зменшити навантаження на ресурс, проте він агресивно оперує на рівні IP.

WAF-правила та сигнатурна фільтрація. WAF - Web Application Firewall застосовує правила (сигнатури), що виявляють експлойти [2], SQL-ін’єкції [3],

сканування тощо. Якщо трафік з якогось IP постійно тригерить WAF-правила - IP може опинитися в чорному списку або попасти під підсилений моніторинг. CDN і різні хостери зазвичай дозволяють автоматично додавати IP до мережевих списків (allow/deny) на рівні Security Configurations [4].

Challenge-response (CAPTCHA, JavaScript challenges). Перед тим як заблокувати IP адресу, деякі системи застосовують виклики, такі як - CAPTCHA, JavaScript-чек, cookie challenges. Це дозволяє відрізнити бота від реального користувача без миттєвого бана, що в свою чергу може запобігти збиткам цього ресурсу. Якщо виклики не проходять, то в такому випадку IP адреса може бути заблокована або обмежена. Cloudflare та інші сервіси широко використовують такі підходи.

Репутаційні списки (blacklists) та автоматичні блок-листинг сервіси. IP адреси можуть числитися в чорних списках на підставі історії (ботнет-активність, фішинг, розповсюдження спаму). Блокування з урахуванням репутації зазвичай досить швидке - сервіси беруть готове рішення на підставі зібраних даних з багатьох різних джерел. Ефективність і точність таких списків це результат багатьох різних досліджень але вони можуть мати проблеми з хибними позитивами й маніпуляціями [5].

Провайдерські та клієнтські політики (manual takedown). Інколи блокуванням може бути результат ручного втручання адміністратора, провайдера або ж клієнта CDN отримує скаргу або припущення про певне зловживання і вмикає блокування для цього користувача. Такі рішення досить часто супроводжуються дзвінками або тікетами та є значно повільнішими, проте більш «політичними» і менш прозорими. RFC-огляди підкреслюють, що блокування може виникати на основі політичних або юридичних підстав, а не лише через певні технічні ознаки [6].

Основні сигнали (тригери) для блокування IP. Зовнішні сервіси захисту (CDN, анти-DDoS, провайдерські фільтри, WAF) не покладаються на один показник при прийнятті рішення про блокування якоїсь IP-адреси. Замість цього вони проводять оцінку певного набіру ознак, кожна з яких може сигналізувати про певну підозрілу або нездорову активність. Одним із сигналів є аномальні обсяги

трафіку. Якщо IP генерує занадто високу частоту запитів (Requests Per Second) або передає занадто великий обсяг даних за секунду, це може попередньо вважатися ознакою атаки або перевантаження. У дослідженні On the Effectiveness of Rate Limiting Mechanisms реалізовано емпіричний аналіз різних схем rate limiting (обмеження швидкості) у контексті боротьби з червами та зловмисним трафіком, показуючи нам, що такий контроль частоти з'єднань може бути дуже корисним і знижувати негативний вплив атак, зберігаючи при цьому допустимий рівень помилок (false positives) [7].

Аномально висока кількість нових TCP-з'єднань за короткий інтервал часу є важливим сигналом потенційної зловмисної активності. У нормальних умовах частота ініціації TCP-сесій має стабільний характер і визначається типом сервісу та реальним навантаженням. Різке збільшення кількості SYN-пакетів з одного IP може вказувати на порт-сканування, brute-force або SYN-flood-атаку. Системи захисту (WAF, DDoS-фільтри, провайдерські монітори трафіку) відстежують частоту нових сесій, частку незавершених SYN→ACK, співвідношення активних і закритих з'єднань та середню тривалість сесій. При виявленні відхилень - наприклад, коли IP генерує сотні чи тисячі запитів за секунду - такі системи автоматично позначають його як підозрілий і застосовують реакцію rate-limit, тимчасове блокування, занесення в blacklist або перенаправлення в scrubbing-центр. Ефективність подібних механізмів підтверджена у роботах Xu et al. (2005) [7] та Thomas et al. (2011) [8], де показано зв'язок між частотою створення нових TCP-сесій і зловмисною активністю, зокрема ботнет-інфекціями.

Сучасні системи блокування IP працюють переважно автоматизовано, що дозволяє миттєво реагувати на високочастотні загрози, зокрема DDoS чи brute-force атаки. Платформи на кшталт Cloudflare, Akamai чи AWS Shield застосовують багаторівневі евристику та поведінкові метрики для автоматичного ухвалення рішень щодо блокування. Попри високу швидкість і ефективність, такі системи схильні до хибних спрацьовувань за даними Olateju et al. (2024), їх рівень може досягати 2–10 % [9]. Частково це компенсується механізмами «human-in-the-loop», однак ручна перевірка можлива лише вибірково через великий обсяг трафіку, що

на практиці призводить до моделі «block first, review later». Особливо гостро проблема проявляється у NAT/CGNAT-середовищах, де блокування спільної публічної IP-адреси через дії одного користувача автоматично зачіпає всіх інших клієнтів, що робить автоматизовані рішення вразливими без доступу до внутрішнього NAT-контексту.

Як зовнішні сервіси встановлюють зв’язок між даними - загальний робочий процес. Механізми прийняття рішення про блокування IP-адреси у зовнішніх сервісах (таких як CDN, WAF, анти-DDoS платформи або фільтри провайдерів) ґрунтуються на поєднанні декількох різних але послідовних етапів аналізу трафіку. Оскільки головна мета таких систем - забезпечити максимально можливу точність виявлення загроз у реальному часі, вони застосовують комплексний підхід, який включає в себе збирання метрик, попередню фільтрацію, перевірку репутації, кореляцію сигналів та формування кінцевої реакції [1, 10, 11] .

Першим етапом є збір метрик, який виконується у режимі реального часу. CDN і WAF-системи реєструють широкий спектр параметрів - інтенсивність трафіку (Requests Per Second - RPS), кількість створених нових TCP-сесій, інформацію про те, звідки географічно походить трафік, автономну систему (ASN), розподіл браузерних та клієнтських ідентифікаторів (User-Agent), наскільки часто виникають помилки HTTP (коди 4xx і 5xx), типи використовуваних методів доступу та інші технічні атрибути [13, 16]. Збір таких показників дозволяє цим сервісам сформувати первинний профіль поведінки конкретної IP-адреси та виявляти відхилення від нормального, штатного режиму взаємодії користувачів із певним сервісом.

Другий етап - попередня фільтрація трафіку. На цьому рівні застосовуються прості, але швидкі правила - перевірка на перевищення встановлених лімітів (rate-limiting), аналіз сигнатур, виявлення нестандартних заголовків або параметрів запитів до сервісу [13, 14]. Якщо система зафіксує перевищення базових порогів (наприклад, дуже високий RPS або занадто різке зростання кількості SYN-пакетів), вона застосує первинні механізми стримування - такі як CAPTCHA-виклик, JavaScript challenge чи навмисне тимчасове уповільнення запитів (throttling). Ці дії,

“дії первинної реакції” надають можливість досить швидко відсіяти примітивні автоматизовані атаки та ботів без великого навантаження на глибинні модулі аналізу [7, 10].

Третім етапом і водночас важливим етапом є репутаційна перевірка IP-адреси. Система звертається до внутрішніх і зовнішніх репутаційних баз, таких як чорні списки (blacklists), сірі списки (greylists), “білі” довірені списки (whitelists), а також до загальних баз даних, які агрегують скарги, різні відомості про спам, ботнет-активність або певні компрометаційні ознаки [5, 15, 16]. Наприклад, якщо IP вже є у репутаційній базі, як джерело спам-розсилки чи шкідливого трафіку, його подальші взаємодії з сервісом оцінюються з вищою підозрою, а поріг блокування значно знижується.

Четвертий етап передбачає кореляцію сигналів і формування скорингової моделі, яка присвоює IP-адресі інтегральний показник ризику на основі сукупності ознак. Метрики та репутаційні дані узгоджуються в єдину систему оцінювання, що може базуватися як на порогових правилах (наприклад, надмірна кількість нових з’єднань), так і на алгоритмах машинного навчання. ML-моделі дозволяють враховувати нелінійні залежності та зіставляти поведінку IP із історичними шаблонами, формуючи узагальнений threat score. На цьому етапі система об’єднує інформацію з різних джерел поєднання високого RPS, негативної репутації та повторюваних User-Agent значно підвищує ймовірність класифікації трафіку як зловмисного [8, 16, 19].

Заключний етап - реакція системи на загрозу, яку вдалося виявити. В залежності від інтегральної оцінки ризику та корпоративної політики безпеки можливі наступні варіанти реагування:

– легкі заходи (CAPTCHA, JS-challenge, обмеження частоти запитів);

– часткове блокування окремих методів чи портів;

– повне блокування IP-адреси або навіть всього CIDR-блоку;

– ескалація інциденту адміністраторам або провайдеру;

– у крайніх випадках, може бути формування запиту до правоохоронних органів [6, 10, 11].

У деяких системах, перед тим як остаточно блокувати IP, застосовується принцип human-in-the-loop, коли адміністратор може вручну переглянути подію і прийняти певне рішення, проте через великий масштаб трафіку, такі перевірки виконуються лише вибірково [2, 9]. І тоді, кінцева реакція визначається комбінацією автоматичних рішень та, в окремих випадках, ручною модерацією.

Наслідки та особливості для середовища NAT / shared IP. У середовищах, де використовується мережевий транслятор адрес (NAT), а особливо його провайдерський варіант - Carrier-Grade NAT (CGNAT), одна публічна IP-адреса по суті представляє велику кількість різних внутрішніх клієнтів. Це в свою чергу означає, що будь-який трафік, який виходить від різних користувачів усередині локальної мережі, агрегується під одним публічним ідентифікатором. Для зовнішніх сервісів таких як - CDN, WAF, анти-DDoS систем, по суті, цей трафік виглядає як єдиний потік, що дуже сильно ускладнює аналіз та підвищує ймовірність некоректних рішень щодо блокування користувачів. [6, 15].

Однією з найбільших проблем при таких умовах є високий ризик collateral damage, тобто ситуації, коли блокування однієї публічної IP-адреси призводить до повного відключення від сервісу великої кількості ні в чому невинних користувачів NAT-сегменту. У випадку з автоматичним застосуванням санкцій/блокувань на основі RPS-порогів, репутаційних сигналів або сигнатур атак, зовнішні сервіси не можуть відокремити зловмисника від загального пулу (кількості) клієнтів. Це створює системну архітектурну вразливість, яку згадують і в стандартах IETF, і в різних дослідженнях щодо фільтрації трафіку [6, 10, 11].

Більш того, у середовищах NAT навіть виникає специфічний напрямок атак, який отримав назву abuse-induced DoS. Суть цього виду атак полягає в тому, що зловмисник навмисно генерує трафік із середини саме NAT-сегменту або через орендовані ресурси (VPN, проксі, ботнет-вузли), створюючи штучне враження аномальної поведінки з метою провокації блокування всього public IP. Це дозволяє атакуючому задати удару не по самому ресурсу, а по звичайним користувачам, які поділяють один, загальний IP із ним. У такому випадку традиційні автоматизовані механізми реагування на загрози, фактично стають інструментом для атаки [11, 16].

Для зменшення таких ризиків певні платформи намагаються застосовувати контекстно-орієнтовані механізми, такі як аналіз розподілу портів, кількості одночасних NAT-відображень (mappings), ентропії User-Agent, глибини TCP-handshake та різних інших маркерів, що дозволяють оцінити внутрішню неоднорідність трафіку за одним публічним IP. В теорії такі метрики дають змогу відрізнити справжню атаку зловмисника від колективної активності кількох легітимних клієнтів, проте їхнє застосування вимагає доступу до внутрішніх логів NAT або таблиць трансляцій - даних, якими зовнішні CDN і WAF зазвичай не володіють [6, 15, 16].

У певних випадках провайдери та корпоративні адміністратори застосовують більш тонкі політики для пом’якшення, доприкладу блокування лише певних діапазонів портів або тимчасове обмеження окремих підозрілих сесій, застосування challenge не до всього трафіку, а лише до окремих потоків. Проте ці рішення можливі лише у тому випадку, коли контроль здійснюється на стороні оператора, тобто коли відома кореляція внутрішніх зв’язків між приватними адресами та їх NAT-мепінгами. Зовнішні сервіси, які бачать тільки публічний IP, нажаль, не можуть реалізувати настільки деталізовану політику [6, 11, 15].

На цю серйозну невідповідність між внутрішньою структурою трафіку та зовнішніми механізмами для контролю неодноразово звертали увагу у RFC-документах IETF, які підмічають, що блокування на рівні public IP у випадку NAT майже завжди має “надлишковий” характер і є архітектурно недосконалим методом для реагування на подібні загрози [6]. Дослідження репутаційних систем також підтверджують, що IP-репутація у shared-середовищах має значущі похибки, оскільки поведінка одного шкідливого користувача компрометує всіх інших користувачів у сегменті [5, 15, 16].

У результаті NAT/CGNAT стає критичною проблемою для будь-яких систем автоматичного блокування, оскільки це створює умови, за яких навіть суто технічно коректне рішення може призвести до масового, необґрунтованого блокування легітимних користувачів. Це в свою чергу вимагає створення окремих «NAT-специфічних» методів детекції, які будуть розглядатися у наступних

підрозділах цієї роботи.

Проблеми достовірності та контртехніки. Попри активний розвиток систем які використовуються для фільтрації трафіку, зовнішні сервіси, що базуються на автоматизованому аналізі аномалій та репутаційних показниках, зіштовхуються з численними проблемами пов’язаними з достовірністю. Однією з основних вразливостей є можливість маніпуляцій репутаційними механізмами, коли зловмисник може штучними методами «заплямувати» IP-адресу іншого користувача, змусивши репутаційні сервіси класифікувати цю адресу як шкідливу. Така атака зазвичай здійснюється через підставні ботнет-запити, масове надсилання скомпрометованих пакетів або навіть координовані скарги, які репутаційні системи можуть обробляти автоматично [5, 15]. Дослідження вказують, що через велику кількість джерел репутаційних даних та відсутність стандартизованої валідації, чорні списки дуже часто містять некоректні або застарілі записи, що робить фільтрацію на їх основі вразливою до помилок [15, 16].

Додатковою проблемою звісно є обмежений доступ зовнішніх сервісів до внутрішньої мережевої інформації, зокрема до NAT-mapping таблиць, логів внутрішніх проксі, а також даних про конкретних користувачів у середовищах локальних мереж. CDN і WAF бачать лише зовнішню IP-адресу NAT, тому не можуть визначити, чи здійснює підозрілу активність один конкретний користувач всередині NAT-сегменту, чи навантаженість трафіку є результатом колективної активності багатьох різних, незалежних користувачів [6, 11, 16]. Через цю інформаційну обмеженість вони покладаються на вже агреговані метрики, що в свою чергу збільшує ймовірність як хибних позитивів, так і хибних негативів, особливо у великих провайдерських мережах.

У науковій літературі також описуються різні контртехніки зловмисників, спрямовані на обхід моделей виявлення. Зокрема, зловмисники можуть модифікувати User-Agent заголовки, додавати шум у часові патерни запитів, варіювати джерела звідки буде йти трафік або використовувати розподілені ботнети, у яких кожен вузол генерує мінімальне навантаження, не перевищуючи порогів базових лімітів [10, 17]. За допомогою таких методів, зловмисник створює

трафік, який виглядає «легітимним» для моделей порогового типу, особливо якщо платформа не має глибокого поведінкового аналізу. Для протистояння таким технікам необхідні методи, що будуть враховувати довгострокові поведінкові закономірності, а також моделі машинного навчання, які здатні відстежувати складні кореляційні зв’язки між запитами [12, 18].

Усе це показує, що зовнішні системи блокування часто працюють у умовах недостатньої інформації, що створює передумови для частих помилкових рішень як у бік надмірного блокування, так і в бік пропуску небезпечного трафіку. Враховуючи ці обмеження, у наступних розділах буде розглянуто можливість використання внутрішніх NAT-метрик та спеціалізованих евристик для підвищення точності класифікації, особливо в умовах спільних публічних IP-адрес, де традиційні методи виявлення аномалій працюють найменш ефективно.

Як ми бачимо процес блокування IP-адрес зовнішніми сервісами є багаторівневим і комплексним, поєднуючи в собі як високошвидкісні автоматизовані механізми (rate-limiting, сигнатурні WAF-фільтри, CAPTCHA/JS-challenge), так і рішення, що базуються на репутаційних показниках або навіть ручному втручанні адміністратора. Є різні етапи аналізу - збір метрик, попередня фільтрація, перевірка репутації, зв’язок між сигналами, забезпечують хорошу гнучкість та швидкість реакції, але одночасно створюють умови для появи хибних спрацьовувань [10, 12].

Основні тригери, які впливають на ухвалення рішення про блокування, включають обсяг і характер трафіку (RPS, SYN-потоки, кількість нових з’єднань), наявність сигнатур відомих атак, ознаки різних ботнет-активностей, а також дані репутаційних систем. Однак при застосуванні цих механізмів у середовищах NAT та CGNAT виникає системна проблема - одна публічна IP-адреса об’єднує десятки а може навіть сотні кінцевих користувачів, що різко збільшує ризик колективного блокування добросовісних клієнтів на основі поведінки лише одного зловмисника [6, 15, 16].

Така архітектурна особливість призводить до того, що навіть технічно коректні рішення зовнішніх сервісів можуть мати серйозний побічний ефект -

collateral damage, який у великих мережах провайдерів або публічних Wi-Fi-сегментах може зачепити велику кількість реальних користувачів. Через це зростає потреба у методах, що здатні враховувати внутрішню структуру NAT-трафіку, проводити аналізи поведінкових патернів на внутрішньому рівні та зменшувати залежність від агрегованих зовнішніх метрик.

У подальших підрозділах цієї роботи буде розглянуто, які саме внутрішні метрики, характеристики NAT-відображень, поведінкові евристики та алгоритми аналізу можуть використовуватися для точнішої ідентифікації зловмисної активності в умовах спільних публічних IP-адрес. Особлива увага буде приділена тому, як відрізнити просто загружений трафік від справжньої атаки зловмисника, що дозволить суттєво знизити рівень false positives та підвищити точність виявлення аномалій у публічних мережах.

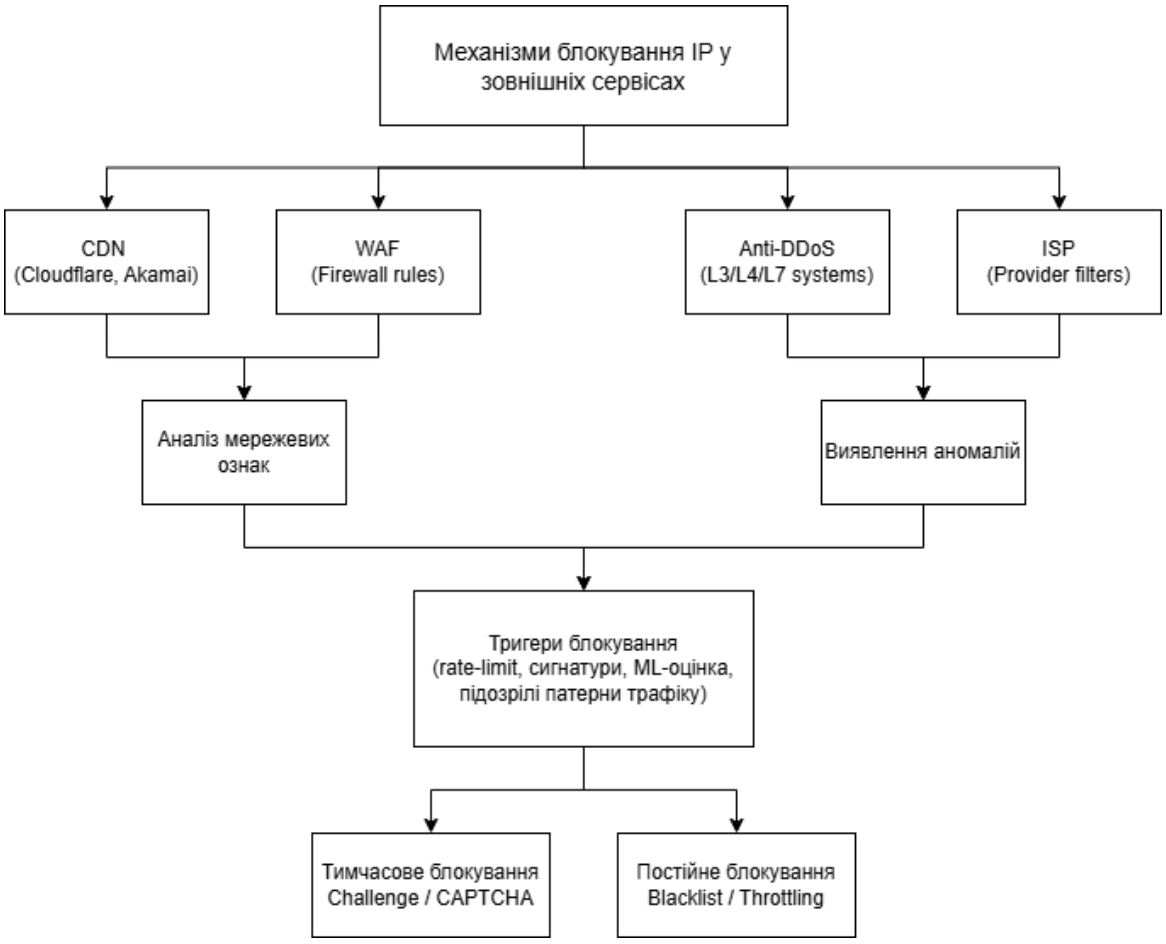

Рисунок 1.1 – Схема механізмів блокування IP у зовнішніх сервісах

Схема узагальнює ключові механізми, які використовують зовнішні сервіси - CDN, WAF, анти-DDoS системи та провайдерські фільтри - для ухвалення рішень щодо блокування IP-адрес. Усі механізми ґрунтуються на аналізі мережевих ознак та виявленні аномалій, що запускають відповідні тригери. Результатом можуть бути як тимчасові обмеження доступу, так і постійне блокування IP.

1.2 Типові ознаки мережевих аномалій та тригери блокування IP

Аномалії мережевого трафіку є головним індикатором для виявлення потенційно шкідливої активності, тому, можна сказати, що це - ключовий фактор у прийнятті рішення щодо блокування IP-адрес зовнішніми сервісами, наприклад CDN, WAF, анти-DDoS платформи та фільтри Інтернет-провайдерів. Всі ці системи працюють у режимі реального часу та використовують комбінацію поведінкових, сигнатурних, статистичних і репутаційних механізмів для аналізу трафіку, який якимось чином відхиляється від норми. Згідно з багатьма дослідженнями, атаки сьогодення все частіше маскуються під непримітний або ще можна сказати легітимний трафік, що вимагає від систем захисту більш складної і глибшої логіки аналізу та багаторівневої кореляції тобто відповідності між факторами [10], [12], [19]. Враховуючи це в умовах NAT-середовищ більша частина класичних для реакції тригерів набуває неоднозначного вигляду, оскільки один публічний IP може представляти десятки або навіть сотні кінцевих клієнтів, що дуже сильно ускладнює диференціацію легітимного та аномального трафіку.

Одним із найбільш очевидних та одним із найбільш поширених показників, які використовуються для того, щоб виявити аномалії трафіку, є різке зростання інтенсивності запитів та кількості пакетів за секунду. У документації провідних хмарних платформ, зокрема Cloudflare, завжди описані порогові значення, перевищення цих порогових значень, в свою чергу ініціює механізми rate limiting або challenge-response перевірки [1], [13]. Під час DDoS-атак типу HTTP-flood аномалії проявляються як неочікуваний стрибок кількості однотипних HTTP-

запитів протягом дуже короткого часу. Наукові дослідження підтверджують, що такі сплески в більшості випадків є одним із найстабільніших індикаторів DDoS-активності, навіть якщо самі запити в загальному виглядають нормальними [10], [18]. Статистичні моделі виявлення HTTP-flood атак, як показано у роботі Amini та співавторів, ґрунтуються на аналізі динаміки міжзапитних інтервалів, схожості payload-структур і стабільності певних шаблонів поведінки клієнтів [24].

Одним із ключових індикаторів зловмисної активності є надмірна частота ініціації нових TCP-з'єднань. За нормальних умов кількість new connections обмежена механізмами keep-alive, тоді як під час порт-сканування, brute-force або SYN-flood спостерігається різке зростання SYN-трафіку без відповідного корисного навантаження. Такі патерни з високою точністю вказують на атаки, що спрямовані на виснаження ресурсів сервера або визначення відкритих портів; це підтверджують результати Moore і Voelker [20]. Тому провайдерські DPI-системи та CDN включають контроль new_conn_rate як базовий елемент політики безпеки [23].

Важливим джерелом сигналів є поведінкові відхилення від нормального користувацького трафіку. До них належать циклічні запити, повторювані звернення до тих самих ресурсів, серії помилок автентифікації або сканування структури застосунку через звернення до неіснуючих сторінок. Такі шаблони характерні для brute-force, fuzzing та reconnaissance-діяльності. Thomas, Paxson та Livshits показали, що циклічність у часі є одним із надійних маркерів ботнет-поведінки [8]. Звичайна ж активність користувача позбавлена такої регулярності та проявляється значно хаотичніше [12, 25].

Окрему групу тригерів утворюють сигнатурні ознаки атак - характерні шаблони SQL-ін'єкцій, XSS, directory traversal та інших відомих експлойтів. WAF-системи, такі як ModSecurity CRS, містять сотні правил для порівняння запитів із цими сигнатурами [3, 6]. Хоча метод ефективний проти класичних атак, його точність суттєво падає у випадку поліморфних або zero-day загроз. Rossow зазначає, що сучасні ботнети активно модифікують структуру пакетів з метою обходу сигнатурного аналізу, що стимулює перехід до комбінованих моделей

детекції [21].

Значущим показником є також аномалії у User-Agent заголовках. Ботнети часто використовують статичні або мінімалістичні UA-рядки, які погано імітують реальні браузери чи мобільні клієнти. Дослідження Kim та ін. демонструє, що низька ентропія User-Agent є одним із найточніших маркерів автоматизованих систем, оскільки легітимний трафік характеризується значно ширшою варіативністю цього параметра [27].Сканування порту, воно передує багатьом різним типам атак, також розглядається як дуже важливий тригер. У системах DPI провайдерів та фаєрволах CDN визначаються патерни звернень до портів у якомусь певному порядку або з певною частотою, що зазвичай вказує на підготовку до атаки. Огляд Alshamrani, Myneni доводить, що більшість сучасних сканерів (Nmap, Masscan тощо) залишають після своєї роботи статистично помітний “відбиток”, який легко виявляється навіть без глибокого аналізу корисного навантаження [25].

Також значну роль відіграють і репутаційні фактори. Різні репутаційні платформи збирають інформацію про певні IP-адреси, які раніше були помічені у спам розсилках, атаках, ботнет активності або в будь яких інших зловмисних діях. У дослідженнях Esquivel, Akella та Mori виявлено, що сервіси часто приймають рішення про те, щоб заблокувати IP не на основі справжнього поточного трафіку, а на підставі вже існуючих негативних репутаційних позначок щодо цього IP [5]. Це в свою чергу створює значні ризики у випадку NAT-інфраструктури, де “погана репутація” може належати одному з користувачів, а страждатимуть усі. IETF у RFC 7754 додатково наголошує, що покладання виключно на IP є в принципі хибним у випадку моделями де одна IP адреса є виходом у інтернет для багатьох пізних користувачів, таких як CGNAT [6].

У NAT-середовищах інтерпретація аномалій ускладнюється багато кінцевих користувачів ділять одну публічну IP-адресу, і звичайна варіативність поведінки (паралельні підключення, різні User-Agent тощо) може виглядати як підозріла активність для зовнішніх сервісів. Meng et al. відзначають, що аналіз потоків у NAT-мережах ускладнюється через змішування сесій, рециркуляцію портів і помилкові співставлення зовнішніх і внутрішніх потоків. Через це традиційні

тригери (RPS/PPS, rate new connections, сигнатури, репутація) втрачають точність і збільшують ризик помилкових блокувань [28].

Отже, для коректної детекції в NAT-пулах потрібні адаптовані методи, які враховують групову природу трафіку та внутрішній NAT-контекст - наприклад, агреговані метрики по пулу, кореляція з conntrack-логами або використання специфічних NAT-ознакових фічів у моделях. Такий підхід зменшує хибні спрацьовування і підвищує точність ухвалення рішень щодо блокування.

Рисунок 1.2 – Типова логіка виявлення мережевих аномалій і тригери, що призводять до реакцій (логування, rate-limit, блокування)

Схема ілюструє основні джерела телеметрії, набори ознак, які фіксують мережеві аномалії, процес агрегування ризику та типові механізми реагування зовнішніх сервісів. Особливо важливі RPS-сплески, аномалії ентропії заголовків та велика частка коротких сесій - саме ці тригери часто активують автоматичні обмеження.

1.3 Вплив NAT / CGNAT на точність ухвалення рішень щодо блокування

Усі механізми трансляції мережевих адрес, такі як Network Address Translation (NAT) та Carrier-Grade NAT (CGNAT), стали важливим елементом сучасної інтернет-інфраструктури, оскільки дозволяють нам економити публічні IPv4-адреси та масштабувати доступ до мережі для великої кількості різних користувачів. Але їхня технічна природа дуже суттєво впливає на точність прийняття рішень щодо блокування в системах веб-захисту, зокрема у WAF (Web Application Firewall), CDN-платформах та механізмах rate-limiting. У ситуації, коли велика кількість користувачів одночасно працює за однією публічною IP-адресою, звичайні методи ідентифікації джерела звіки йде трафік, що базуються на IP-репутації, показниках RPS або поведінкових патернах, стають значно менш ефективними, а інколи навіть - статистично некоректними.

NAT з’явився як тимчасове рішення для дефіциту IPv4-простору, але в результаті став фундаментальним та довгостроковим механізмом маршрутизації у провайдерських мережах. Але на відміну від класичного NAT, де трансляція здійснюється в межах тільки однієї локальної мережі, CGNAT використовується також на рівні інтернет-провайдера, об’єднуючи іноді не сотні а тисячі користувачів за однією публічною адресою [30]. Така концентрація великої кількості трафіку позбавляє IP-адресу ролі унікального ідентифікатора, а це суттєво ускладнює завдання для систем які фільтрують та блокують.

NAT та CGNAT об’єднують велику кількість внутрішніх адрес у одну публічну, тому системи веб-захисту сприймають усіх користувачів як єдине

джерело трафіку. У таких умовах будь-який аномальний пік - навіть спричинений одним клієнтом - автоматично впливає на весь CGNAT-сегмент і може призвести до блокування доброчесних користувачів [32]. Невеликі обсяги ботнет-трафіку або поодинокі ін’єкційні запити здатні сформувати хибне уявлення про поведінку всього пулу та ініціювати захисні механізми.

Агреговані метрики, на які спираються WAF і CDN, суттєво спотворюються при великих CGNAT-групах статистичні ознаки перестають відображати окремих клієнтів і формують узагальнений, неточний профіль. Дослідження Sarrar і Feldmann демонструють, що CGNAT призводить до перенасичення лічильників запитів, унеможливлюючи коректне визначення інтенсивності трафіку кожного користувача [30]. Zhang та співавт. підтверджують, що навіть одиничні атаки в таких кластерах можуть “маскувати” нормальний трафік сотень абонентів і провокувати автоматичні блокування [34].

Також у NAT-середовищах втрачається індивідуальність поведінкових характеристик. Параметри TCP-з’єднань, часові шаблони та сигнатурні ознаки протоколів змішуються між багатьма клієнтами, утворюючи високий рівень статистичного шуму. Це робить традиційні поведінкові та ML-моделі малоефективними, оскільки вони більше не здатні формувати точні профілі нормальної активності, що різко збільшує кількість хибних спрацьовувань [35].

У результаті агрегування різнорідних сесій в одну публічну адресу системи безпеки втрачають можливість коректно корелювати події. IDS і WAF фактично працюють не з індивідуальними потоками, а з усередненим набором дій непов’язаних користувачів, що ускладнює виявлення реальних аномалій і підвищує ризик неправильних рішень щодо блокування.

Додаткові труднощі виникають під час аналізу TCP/UDP-сесій. CGNAT активно використовує пул різних портів для трансляції потоків, через що сесії різних користувачів можуть виглядати подібними або навіть накладатися одна на одну. Це значно зменшує інформативність низькорівневих ознак, які зазвичай застосовуються для виявлення ботнетів, сканування чи автоматизованої активності. Роботи демонструють, що у NAT-середовищах точність профілювання

TCP-сеансів знижується на десятки відсотків, а унікальні патерни дій партійно або повністю зникають [35].

Ще однією з проблем є визначення джерела аномалії. За відсутності розділення між окремими внутрішніми адресами системи безпеки не можуть чітко встановити, який саме клієнт ініціював підозрілу активність. Це призводить до того, що блокування застосовується не до конкретного користувача, який порушує правила користування сервісом, а до всього CGNAT-пулу, що створює значні ризики для доброчесних користувачів. У науковій літературі CGNAT це часто називають “чорним ящиком”, оскільки він приховує справжнє джерело трафіку та суттєво обмежує можливість точного реагування на інциденти [33].

У підсумку NAT/CGNAT, через свою структуру/природу формує середовище, у якому класичні методи аналізу активності та ухвалення блокувальних рішень стають не надійними. Об’єднання трафіку, спотворення поведінкових ознак і відсутність доступу до реальних внутрішніх джерел призводять до зростання як хибних рішень, що потребує розробки нових підходів до аналізу трафіку у мережах з адресним шарингом [35, 36].

Ще варто зазначити, що NAT/CGNAT погано впливає на точність засобів які використовуються для протидії DDoS-атакам. Оскільки у таких мережах трафік багатьох різних клієнтів об’єднаний, системам аналізу майже не можливо коректно визначити, чи є збільшена інтенсивність окремої сесії частиною якогось інциденту або може це результат звичайних змін у поведінці користувачів які відносяться до цієї IP. Це в свою чергу робить застосування механізмів rate-based mitigation складнішими для використання, вони зазвичай покладаються на індивідуальні параметри користувача. Якщо унікальне джерело трафіку відсутнє, то системи блокування в такому разі змушені приймати рішення щодо блокування на основі агрегованих характеристик CGNAT-пулу, що в свою чергу може призвести до надмірних або недостатніх реакцій під час певних атак [31, 34].

Окрім того, NAT/CGNAT блокує можливость використання сесійних моделей, які використовують сучасні ML системи для того, щоб виявити нетипові активності. Машинні моделі, які були засновані на часових залежностях, частотних

характеристиках або сигнатурах клієнтів які постійно користуються послугами, втрачають свою точність, якщо всі ці параметри змішуються між різними користувачами. У різних дослідженнях відзначається, що моделі, які працюють коректно у звичайних мережах, демонструють дуже суттєве падіння ефективності у CGNAT-сегментах, оскільки втрачається стійкість ознак та здатність моделі розрізняти якісь індивідуальні патерни [35, 36]. Це в свою чергу призводить до збільшення кількості хибнопозитивних рішень, а також майже повністю невелює здатність систем точно визначити джерело аномалії.

У підсумку адресний шаринг формує досить складні умови для будь-яких систем, які покликані виявляти загрози та механізми ухвалення рішень щодо блокування. Маскування зловмисної активності, статистичні спотворення даних, втрата унікальності поведінкових профілів користувачів та обмеження сесійної видимості роблять такі підходи значно менш ефективними у NAT/CGNAT-середовищах. Це показує нам потребу у впровадженні технологій, які в свою чергу не залежатимуть від IP ідентифікаторів - наприклад, глибинна кореляція сесій, TLS/JA3-фінгерпринтинг, аналіз стабільних поведінкових ознак користувачів або комбінованих моделей машинного навчання, які в вою чергу здатні зберігати точність навіть за умови адресного шарингу [36].

Рисунок 1.3 – Вплив NAT/CGNAT на процес ухвалення рішень щодо блокування

1.4 Постановка задачі дослідження

Аналізування теоретичних засад механізмів які блокують IP-адреси, особливостей в роботі систем веб-захисту та впливу NAT/CGNAT на те, наскільки точно ухвалюються рішення показав, що в умовах де одна адреса розділена між багатьма користувачами, зовнішні сервіси зазвчиай просто не здатні коректно визначити джерело підозрілої мережевої активності. Агрегація трафіку, викривлення характеристик поведінки і втрата унікальності цих мережевих ознак призводять до сильного підвищеної кількості хибнопозитивних блокувань, особливо якщо це велика публічна мережа. Це, в свою чергу, створює необхідність розробки певних методів, які дозволять компенсувати або зменшити вплив NAT/CGNAT на процес ухвалення рішень про блокування.

З огляду на це об’єктом мого дослідження є саме процес блокування IP-адрес у публічних мережах, що використовують механізми NAT та CGNAT.

Предмет для дослідження - метрики мережі, показники поведінки трафіку та індикатори різних аномалій, які в свою чергу спотворюються через адресний шаринг і сильно впливають на точність рішень систем захисту.

Мета мого дослідження полягає у тому, щоб сформувати метод виявлення таких помилкових блокувань та різних аномальних ситуацій, що виникають унаслідок застосування NAT/CGNAT, а також у визначенні підходів для зменшення кількості хибних рішень у таких мережах.

Для того, щоб досягнути поставленої мети, мені необхідно виконати такі завдання в дослідженні:

– проаналізувати механізми блокування IP-адрес які вже існуують у зовнішніх сервісах та визначити їхню залежність від унікальності адрес клієнтів;

– дослідити те, як впливай NAT і CGNAT на структуру трафіку, профілі поведінки та різні мережеві ознаки, які можуть використовуватися в ухваленні рішень про блокування;

– визначити причини через які виникають хибнопозитивні і хибнонегативні блокування у середовищах де є розділення адреси;

– оцінити вже існуючі набори даних та інструменти, які є необхідними для того, щоб змоделювати NAT-сценарії і збору відповідних метрик;

– сформувати релевантний підхід для виявлення ситуацій, у яких блокування відбулося як наслідок впливу NAT/CGNAT, та запропонувати метод завдяки якому буде менше похибок у прийнятті рішень.

Практична частина цього дослідження спрямована на формування та аналіз наборів даних, отриманих у контрольованих і наближених до реальних NAT та CGNAT-середовищах. У межах практичної реалізації здійснюється збір мережевого трафіку та супровідних метаданих, що відображають процеси трансляції адрес і портів, конкуренцію між внутрішніми клієнтами NAT-пулу та динаміку встановлення з'єднань. Особлива увага приділяється дослідженню ключових NAT-метрик, зокрема станів таблиці conntrack, журналів iptables і параметрів NAT-binding, які суттєво впливають на формування агрегованого поведінкового профілю публічної IP-адреси.

Для детального аналізу мережевої активності використовуються інструменти пакетного моніторингу, зокрема pcap-захоплення з подальшим аналізом у Wireshark. Це дозволяє дослідити часові характеристики трафіку, особливості встановлення та завершення сесій, а також виявити поведінкові патерни, які можуть бути помилково інтерпретовані зовнішніми системами захисту як зловмисні. Паралельно виконується агрегований аналіз потоків, що забезпечує оцінку поведінки трафіку на різних рівнях абстракції.

У практичній частині також реалізується моделювання типових сценаріїв роботи NAT-середовища, включно з нормальним користувацьким навантаженням, піковими сплесками активності та аномальною поведінкою окремих клієнтів. Це дає змогу оцінити вплив адресного шарингу на рішення систем блокування трафіку та перевірити здатність запропонованого методу коректно інтерпретувати агрегований NAT-трафік, локалізувати джерела аномалій і зменшувати кількість хибних блокувань легітимних користувачів.

2 ДАНІ ТА ІНСТРУМЕНТИ ДЛЯ АНАЛІЗУ ПРОБЛЕМИ NAT-БЛОКУВАНЬ

2.1 Оцінка придатності публічних наборів даних для аналізу мережевих аномалій та блокувань у NAT-сценаріях

Для дослідження впливу NAT і CGNAT на точність мережевих блокувань потрібні такі набори даних, що здатні відображати реальні потоки трафіку, природні аномалії та поведінку користувачів у спільних адресних просторах. У першому розділі було показано, що адресне об’єднання спотворює мережеві характеристики, ускладнює ідентифікацію джерел трафіку та провокує хибні спрацьовування систем захисту. Тому публічні датасети можуть бути корисними лише тоді, коли вони містять ознаки, пов’язані з особливостями NAT або дозволяють їх відтворити.

Серед класичних джерел часто використовують CAIDA Anonymized Traces та MAWI Traffic Archive. Це великі операторські трейси, які відображають реальний високонавантажений трафік та містять широкий спектр аномалій [37]. Оскільки на операторському рівні NAT і CGNAT застосовуються масово, у цих даних присутні характерні патерни змішаних потоків. Проте такі датасети зазвичай не мають чіткої розмітки щодо NAT, що ускладнює відокремлення спотворень, спричинених адресним шарингом. Для аналізу впливу NAT вони корисні головним чином для вивчення пікових навантажень, повторюваності з’єднань або нестабільності статистичних метрик.

Архів MAWI вважається одним із найбільш придатних джерел для оцінювання поведінки трафіку у середовищах зі спільними адресами. Дослідження Feldmann та Sarrar показують, що значна частина трафіку MAWI має ознаки NAT, серед яких повторне використання портів, нерівномірні потоки та агреговані патерни [38]. Однак через відсутність структурованої розмітки і необхідність обробки сирих pcap-файлів використання цього архіву потребує суттєвих додаткових етапів підготовки, що ускладнює моделювання ситуацій з блокуваннями.

Популярні у сфері кібербезпеки лабораторні датасети CICIDS2017, CIC-DDoS2019, UNSW-NB15 і CTU-13 містять широкий спектр аномалій і добре підходять для навчання алгоритмів машинного аналізу трафіку [40]. Водночас їхнім суттєвим обмеженням є те, що вони створені в контрольованих умовах без моделювання поведінки NAT. Це знижує їхню практичну цінність для завдань, у яких потрібно враховувати взаємодію багатьох користувачів через одну адресу. У контексті NAT-блокувань такі набори даних можуть слугувати джерелом сигнатур та еталонних аномалій, проте не відображають реалістичні спотворення, які виникають у публічних мережах зі спільним використанням IP-адрес.

Набори даних із поведінковими профілями мережі, зокрема UGR’16 та ADFA-LD, використовуються для аналізу довготривалих змін активності та виявлення аномалій у корпоративних мережах. За спостереженнями Ring та колег, вони допомагають вивчати стабільність поведінкових характеристик користувачів. У середовищах NAT така стабільність значною мірою втрачається через агрегування потоків, тому поведінкові профілі перетворюються на усереднені й менш інформативні. Незважаючи на відсутність спеціального маркування NAT, ці датасети можуть бути корисними для моделювання того, як спільне використання адрес впливає на поведінкові патерни.

Особливу цінність мають набори, що безпосередньо відображають NAT і CGNAT. У дослідженні Zhang та співавторів було зафіксовано характерні ознаки роботи CGNAT у великих операторських мережах, серед яких агрегація портів і повторюваність потоків [33]. Спеціалізовані датасети, такі як Lancaster NAT dataset або MAWI NAT-labelled traces, дозволяють порівнювати трафік за межами NAT та всередині нього і виявляти ознаки, що зникають або спотворюються при адресному шарингу. Це робить їх найбільш релевантними для моделювання ситуацій, пов’язаних з помилковими блокуваннями IP.

Корисним джерелом інформації є результати вимірювань RIPE Atlas NAT Measurements. Хоча це не класичний датасет, він надає уявлення про поширеність CGNAT у різних провайдерів і дає змогу оцінити умови, у яких виникають помилки в блокуванні.

Узагальнений аналіз показує, що універсального публічного набору даних для задач виявлення помилкових блокувань через NAT не існує. Широкомасштабні інтернет-трейси, такі як MAWI або CAIDA, найближчі до реальних сценаріїв, але потребують значної попередньої обробки та не містять явної NAT-розмітки. Лабораторні датасети CICIDS чи UNSW добре підходять для роботи із сигнатурами атак, але не відображають поведінкових спотворень, що виникають у спільних адресних середовищах. Спеціалізовані NAT-орієнтовані набори найбільш точні, однак їхній обсяг та доступність є обмеженими.

Оцінюючи придатність конкретного датасета для моделювання NAT-сценаріїв, важливо враховувати не лише наявність ознак поведінки NAT, але й глибину деталізації. Більшість публічних джерел зосереджена на атаках і лише частково відображає вплив адресного шарингу [39]. Для цілей цього дослідження пріоритетними є такі набори даних, що дають змогу аналізувати спотворення поведінки, які призводять до помилкових блокувань.

Важливим аспектом є метод збору даних. У трейсах, отриманих із периметрів великих провайдерів, таких як MAWI або CAIDA, NAT-поведінка спостерігається природно, оскільки значна частина користувачів операторських мереж отримує доступ до Інтернету через спільні шлюзи. Наприклад, дані CAIDA включають в себе широкий спектр аномалій, як високорівневих, так і низькорівневих, характерних для великих автономних систем [37]. У таких наборах легко помітити характерні ознаки NAT - незвичне повторне використання портів, змішані профілі поведінки та пікові навантаження, спричинені діями багатьох користувачів одночасно. Подібні особливості відзначаються і в дослідженнях NAT-поведінки на основі архіву MAWI [38], що підтверджує цінність цих даних для аналізу впливу NAT/CGNAT на мережеві показники.

На відміну від цього, лабораторні набори даних, попри їхню широку популярність, не відображають реалістичної динаміки адресного шарингу. Наприклад, такі набори як CICIDS2017, CIC-DDoS2019 або UNSW-NB15 створюються в контрольованих умовах, за яких мережевий адресний транслятор (NAT) або не використовується, або має незначний вплив [40]. Унаслідок цього

вони майже не враховують поведінкові спотворення, що спостерігаються в реальних операційних мережах, де одна публічна IP-адреса може одночасно обслуговувати сотні чи навіть тисячі користувачів. Такі набори можуть бути корисними як джерело сигнатур аномальної активності, проте їхнє застосування для аналізу причин хибних блокувань у NAT-середовищах є обмеженим.

Окрему групу становлять набори даних, що безпосередньо відображають NAT-поведінку. У дослідженнях, проведених Fontugne та колегами [41], демонструється, що NAT створює характерні шаблони, які можна виявити в масштабних мережах шляхом аналізу повторюваності портів, розподілу інтервалів між сесіями та статистики потоків. Спеціалізовані набори NAT, такі як MAWI NAT-trace або Lancaster NAT dataset, дають змогу порівнювати поведінку NAT із ненатовим трафіком, що дозволяє точно визначити, які мережеві параметри втрачають діагностичну значущість через спільне використання адрес. Ці набори є особливо корисними для аналізу проблем хибних блокувань, проте їхній обсяг і доступність залишаються обмеженими, що суттєво ускладнює їхнє широке впровадження.

Юридичні та етичні обмеження мають суттєве значення. Дані на рівні оператора часто піддаються анонімізації шляхом вилучення або модифікації полів, які могли б дозволити відтворити поведінку NAT. В результаті, навіть найбільш реалістичні записи обмежують дослідників у здатності простежити причинно-наслідковий зв’язок між внутрішньою адресацією та зовнішнім трафіком. Це знижує ефективність таких даних для аналізу впливу NAT/CGNAT на блокування IP [37].

Узагальнюючи, незважаючи на різноманітність доступних мережевих наборів даних, їхня придатність для моделювання проблеми NAT-блокувань залишається неоднорідною. Найбільш актуальними є спеціалізовані NAT-мічені набори даних та операторські трейси, проте вони часто характеризуються обмеженою доступністю або вимагають складної попередньої обробки. Це підкреслює важливість створення власних контрольованих наборів даних, що стане предметом розгляду у наступному підрозділі.

Рисунок 2.1 – Узагальнена класифікація джерел даних, що застосовуються для аналізу NAT-поведінки та моделювання блокувань

Таблиця 2.1 - Порівняння публічних наборів даних для NAT-сценаріїв

Набір даних	Наявність NAT-трафіку	Рівень деталізації	Придатність для NAT-аналізу
CAIDA Anonymized Traces	Низька	Високий (packet-level)	Середня–потрібна фільтрація
MAWI Traffic Archive	Висока	Середній	Висока – наявні NAT-патерни
CIC-IDS / CIC-DDoS	Відсутня	Flow-level	Низька – не підходить для NAT
Operator Traces (анонімізовані)	Дуже висока	Packet/Flow	Дуже висока, але недоступні публічно

2.2 Формування контрольованих наборів даних, потреба у NAT-специфічних логах

Однією з головних проблем у дослідженні впливу NAT і CGNAT на роботу систем блокування є відсутність якісних наборів даних у відкритому доступі. Більшість публічних датасетів не містять інформації про поведінку NAT або

зазнають глибокої анонімізації, через що втрачаються параметри, важливі для аналізу блокувань. Оскільки NAT приховує структуру внутрішніх користувачів за однією публічною адресою, точне визначення джерела аномалій стає значно складнішим.

Стандарти RFC описують широкий спектр варіантів роботи NAT, серед яких різні стратегії призначення портів, час життя прив'язок і схеми перепризначення. Усі ці властивості впливають на те, як зовнішні системи захисту інтерпретують агрегований трафік. Проте у публічних трейсах ці параметри зазвичай не фіксуються, що унеможливлює точне моделювання поведінки NAT у реальних умовах.

Великі операторські мережі додають ще більше складності. Дослідження Агер показує, що політики CGNAT можуть змінюватися залежно від навантаження або конфігурації інфраструктури, тому поведінкові профілі одного й того самого NAT пулу суттєво різняться у різний час. Через це стандартні методи блокування, що покладаються на фіксовані порогові значення, стають нестабільними. Публічні датасети не враховують ці зміни, тому дослідження, що базуються лише на них, є неповними.

Важливими є також журнали TCP та UDP-сесій з точними часовими мітками. Вони дають змогу аналізувати динаміку перепризначення портів, тайм-аути та особливості роботи NAT у різних мережах. Дослідження показують, що у CGNAT мобільних операторів сесії живуть значно менше, що впливає на стабільність потоків та може спотворювати результати аналізу без відповідної фіксації цих параметрів.

Важливе значення мають HTTP-логи, зокрема User-Agent та інші заголовки, оскільки саме на цьому рівні найчастіше виникають помилкові блокування у WAF. Агрегований NAT-трафік інколи нагадує ботнет-поведінку, тому у контрольованому наборі даних необхідно зберігати часові інтервали між запитами та повторювані шаблони. Не менш важливими є DNS-логи, адже великі NAT-пули створюють характерні групові патерни запитів. Дослідження Poese показує, що одночасне резолвлення доменів сотнями користувачів може виглядати як

аномальна активність. Щоб правильно моделювати такі випадки, DNS-запити потрібно фіксувати з точними часовими мітками.

Важливо також зберігати метадані NAT політик, такі як тип NAT, спосіб призначення портів і максимальна кількість сесій для одного клієнта. Публічні датасети зазвичай не містять цих параметрів, хоча вони суттєво впливають на формування поведінкових шаблонів. Дослідження Ager показує, що зміна алгоритму розподілу портів суттєво змінює профіль роботи CGNAT пулу. Тому у контрольованому наборі даних потрібна не тільки інформація про порти, а й детальний опис політики їх призначення.

Основою є експериментальний стенд із внутрішніми клієнтами, NAT-пристроєм та зовнішнім сегментом, де збираються усі події трансляції. Важливо забезпечити контроль таких параметрів, як призначення портів, тайм-аути і правила перепризначення, оскільки ці механізми безпосередньо впливають на формування агрегованих профілів трафіку. Динаміка роботи NAT може суттєво змінюватися, тому всі операції створення і завершення записів бажано фіксувати.

Датасет повинен включати різні типи трафіку, серед яких веб-запити, DNS обмін, TCP і UDP потоки, а також шаблони, що нагадують активність ботнетів. Таке моделювання допомагає виявити, як NAT впливає на статистичні показники, якими користуються WAF та інші системи захисту. Маркування кожної внутрішньої сесії дає можливість відстежувати повний ланцюг трансляції і точно визначати, які дії викликають помилкові блокування. Це дозволяє порівняти поведінку систем безпеки і оцінити їх здатність працювати із сумарним NAT трафіком.

У датасеті моделюються рішення різних систем блокування, що дозволяє оцінити, як NAT спотворює показники трафіку та впливає на їхню точність. Для коректного аналізу потрібні точні часові мітки, оскільки короткі інтервали між подіями відображають специфічні патерни NAT пулу.

Після збору всі дані структуруються у спільну систему, що поєднує NAT записи, потоки та логи. Це забезпечує кореляційний аналіз і дає змогу відтворювати сценарії, які спричиняють хибні блокування. Такий контрольований датасет є

базою для подальших експериментів і дослідження взаємодії NAT зі системами мережевої безпеки.

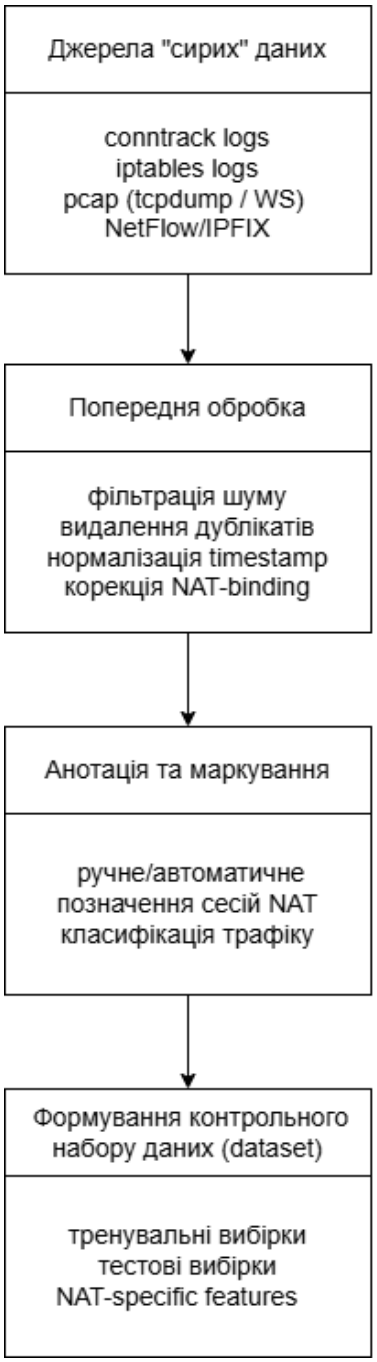

Рисунок 2.2 – Пайплайн формування контрольованих наборів даних для аналізу NAT-поведінки

2.3 Збирання NAT-метрик conntrack, iptables, pcap (Wireshark), внутрішні NAT-таблиці

Аналіз впливу NAT та CGNAT на блокувальні системи потребує детальних мережевих метрик, які відображають зміни адрес і портів у динаміці. Для цього недостатньо загальних трейсів, тому важливим є доступ до внутрішніх механізмів операційної системи. Найбільш інформативними джерелами є таблиці conntrack, журнали iptables, повні pcap-трейси та внутрішні NAT-таблиці маршрутизаторів.

Система conntrack, докладно описана в роботах Бьонді та колег, зберігає стани всіх TCP, UDP та ICMP потоків і пов'язує внутрішні та зовнішні адреси. Саме ці записи дозволяють визначити, який внутрішній клієнт стоїть за конкретним запитом і чи могли особливості NAT спричинити блокування [51].

Журнали iptables також є важливим джерелом інформації. Як зазначає Керриск, вони фіксують створення потоків, їх проходження через NAT-ланцюги та можливі відхилення пакетів [52]. Ці записи допомагають встановити, чи помилкове блокування було пов'язане з аномаліями трафіку або з особливостями роботи NAT.

Важливим джерелом даних про NAT-метрики слугують повні захоплення пакетів у форматі pcap, отримані за допомогою інструментів на кшталт Wireshark чи tcpdump. На відміну від агрегованих flow-даних, файли pcap зберігають повний вміст пакетів, включаючи всі заголовки протоколів та точні часові позначки, що дає змогу детально відновити послідовність подій у середовищі з NAT. Згідно з документацією Wireshark, такі повні трасування є ключовими для глибокого аналізу низькорівневих процесів, зокрема TCP-переперадач, конфліктів у просторі портів, перекриття потоків та особливостей поведінки під високим навантаженням. У контексті виявлення помилкових блокувань pcap-файли особливо цінні, адже дозволяють достовірно визначити, чи спричинений сплеск активності реальною атакою, чи це лише наслідок легітимної діяльності багатьох користувачів за одним NAT.

У дослідженні хибних блокувань важливо врахувати, що NAT-пристрої можуть поводитися непередбачувано під час високих навантажень. Abdelsalam та

колеги зазначають, що під час пікової активності CGNAT здатний тимчасово змінювати політику призначення портів, що призводить до ситуацій, коли трафік кількох внутрішніх користувачів частково накладається [56]. Це утворює нерівномірні патерни, які зовні нагадують активність ботнету. У подібних умовах критично важливими стають журнали pcap, conntrack і iptables, адже вони дозволяють розібратися зі справжньою причиною блокування.

Доповнюючи це, деталізоване відстеження на рівні пакетів (packet-level tracing), як зазначено в роботі Kohno з колегами [57], дає можливість відновити повну картину життєвого циклу кожного пакета - від його створення до проходження через NAT і прийняття відповідного рішення системою безпеки. Ця інформація, хоч і недоступна в публічних наборах даних, є виключно важливою для аналізу процесів блокування.

Коли ми створюємо контрольований набір даних для вивчення поведінки NAT чи CGNAT, одним із найважливіших моментів є правильний і повний збір телеметрії. Справа в тому, що звичайні мережеві логи тут не дуже допомагають, NAT-метрики дуже динамічні - вони постійно змінюються залежно від навантаження, типу трафіку та налаштувань самого пристрою. Тому до процесу збору даних доводиться ставитися особливо серйозно, потрібна ідеальна синхронізація, максимально точні часові мітки, доступ до внутрішніх структур системи, і головне - щоб сам процес збору якомога менше спотворював те, що ми вимірюємо.

Одним із головних джерел даних є інструменти conntrack-tools, які дають змогу отримувати повний стан таблиці з’єднань у Linux у реальному часі. Conntrack зберігає відповідності між внутрішніми й зовнішніми адресами, керує тайм-аутами та повторним використанням портів, що робить його ключовим елементом аналізу поведінки NAT [58]. Потоки подій util-linux, nfct або libnetfilter_conntrack дозволяють відтворювати життєвий цикл кожного NAT-з’єднання та визначати можливі причини хибних блокувань.

Доповненням до цього є журнали iptables, які фіксують, як пакети проходили через правила фаєрволу та NAT. Вони дають змогу побачити, чи пакет був

перенаправлений, відфільтрований або відхилений через переповнення портів чи таблиць станів. Такі логи дозволяють розрізнити, чи блокування спричинене реальною загрозою, чи проблемами в роботі самого NAT.

Важливе значення мають pcap-захоплення, що дають деталізований вигляд трафіку. Проте збір pcap-файлів є складним завданням через ризик втрати пакетів і спотворення часових міток. Дослідження Rohrer показують, що без оптимізації налаштувань мережеві інтерфейси можуть втрачати значну частину даних, тому в експериментах варто застосовувати af_packet, PF_RING або DPDK.

Усі джерела даних мають бути синхронізовані в часі з високою точністю. Будь-які зсуви у часових мітках можуть спотворити аналіз подій, тому необхідно використовувати NTP, Chrony або апаратні засоби синхронізації, як рекомендував Paxson [60].

Також досить важливим завданням є збирання внутрішніх NAT-таблиць маршрутизатора або CGNAT-шлюзу. На відміну від conntrack, NAT-таблиця може містити додаткову інформацію, яка стосується політики розподілу портів, діапазонів пулів та структури мапінгу для різних внутрішніх клієнтів. Згідно з дослідженням Abdelsalam та колег, зміни в NAT-таблицях можуть безпосередньо впливати на характер поведінки трафіку, особливо під час пікових навантажень. Таким чином, у рамках експериментального середовища важливо регулярно робити "зрізи" внутрішніх NAT-таблиць та порівнювати їх із даними conntrack і pcap, щоб сформувати повну картину стану мережі.

Ще один момент, на який обов’язково треба звернути увагу, - це щоб сам процес збору даних якомога менше впливав на систему загалом. Буває, що якщо увімкнути дуже детальне логування в iptables або надто часто опитувати conntrack, то NAT-модуль починає сильно гальмувати. В результаті поведінка NAT під час експерименту може відрізнятися від реальної, і всі наші вимірювання зіпсуються. Тому краще використовувати інструменти, які працюють "тихенько" доприкладу, Netlink sockets, nfct event listeners або просто дублювати трафік через mirror-порт на окремий інтерфейс. Так ми отримуємо потрібні дані, а система продовжує працювати майже як у звичайному режимі.

Підсумовуючи цю частину, збір NAT-метрик - це справді непростий, багатошаровий процес. Тут треба одночасно стежити за синхронізацією кількох джерел, вибирати найменш інвазивні методи захоплення, дбати про точність часових міток і мати глибокий доступ до внутрішніх механізмів NAT. Якщо підійти до цього недбало, то ніякого якісного контрольованого датасету не вийде - і, відповідно, жодного надійного аналізу поведінки систем блокування в NAT-середовищах.

Після збору різноманітних NAT-метрик наступним кроком є їх узгодження у спільну модель поведінки. Кожне джерело даних - conntrack, iptables, pcap-трейси та внутрішні таблиці NAT - надає лише часткове відображення процесів, тому аналіз вимагає кореляції цих даних за часом і змістом. Найважливішою проблемою при цьому є синхронізація часових міток. Pcap працює з мікросекундною точністю, тоді як iptables і conntrack часто демонструють значні зсуви, що може спотворювати хронологію подій. Щоб забезпечити точність аналізу, зазвичай використовують вирівнювання часових рядів або ідентифікують спільні «опорні події», які є присутніми у всіх джерелах даних.

Кореляція між пакетним та сеансовим рівнями є ще одним ключовим фактором аналізу. Pcap дозволяє працювати з окремими пакетами, тоді як conntrack управляє цілими сесіями. У NAT-середовищі ці два рівні часто не збігаються через затримки під час створення записів, повторне використання портів або конкурентні запити від різних користувачів. З цієї причини аналіз базується на портах, IP-адресах, подіях TCP (SYN, ACK, FIN) та інших спільних характеристиках, які допомагають зіставити пакет із відповідним записом у NAT-таблиці.

Інтеграція NAT-таблиць надає можливість відстежувати розподіл ресурсів, політику повторного використання портів та поведінку NAT-пулу під час навантаження. Регулярні знімки таблиць дозволяють ідентифікувати моменти, коли різні внутрішні клієнти отримують однакові зовнішні порти або коли через перевантаження NAT змінює свою політику розподілу ресурсів. Зовні такі ситуації можуть виглядати як аномальний трафік або навіть кібератака, хоча насправді це є звичайною реакцією NAT на підвищений рівень навантаження.

Кореляція NAT-метрик є важливим етапом у створенні достовірної моделі роботи NAT/CGNAT. Лише через узгодження даних з різних рівнів телеметрії можна правильно відтворити реальні сценарії обробки трафіку та визначити природу блокувань - чи вони спричинені політикою безпеки, чи технічними особливостями модуля NAT. Такий підхід забезпечує міцну основу для подальшого моделювання та аналізу випадків хибних блокувань у багатокористувацьких мережах.

2.4 Створення тестового середовища NAT Mininet / GNS3 / Linux iptables - інфраструктура експериментів

Створення контрольованого тестового середовища для моделювання NAT і CGNAT є ключовим етапом роботи, адже в реальних операторських мережах доступ до внутрішніх таблиць, телеметрії та конфігурацій обмежений. Тому лабораторне відтворення типових сценаріїв, коли багато користувачів працюють за однією публічною адресою та провокують помилкові блокування, є єдиним практичним способом дослідження.

Базове середовище можна побудувати на Linux з використанням iptables або nftables та модуля conntrack. Це дозволяє створити NAT з обмеженим пулом портів, запустити кілька ізольованих хостів і спостерігати за заповненням таблиць, перепризначенням портів та поведінкою трафіку під навантаженням. Такий підхід підходить для початкових експериментів і швидкої перевірки поведінки NAT.

Для складніших топологій ефективним інструментом є Mininet, який дає змогу створювати віртуальні мережі з десятками хостів, свічів і маршрутизаторів. Завдяки підтримці NAT можна досліджувати, як групова активність за однією IP-адресою змінює зовнішню видимість трафіку і впливає на точність систем блокування. Якщо потрібна емульована робота мережевого обладнання різних вендорів, використовується GNS3, що дозволяє моделювати CGNAT-пули, кілька рівнів NAT і реалістичні операторські сценарії.

Ці інструменти дають повний контроль над параметрами NAT, включно з таймаутами, пулом портів та правилами розподілу, що дозволяє точно моделювати типові CGNAT-сценарії, як-от конкуренцію між клієнтами, повторне використання портів і переповнення таблиць. Завдяки цьому можна оцінювати, як така динаміка впливає на точність систем блокування.

Mininet є одним із найзручніших інструментів для побудови складних мережевих топологій з багатьма вузлами. Він дозволяє моделювати реалістичні багатокористувацькі сценарії, кілька внутрішніх підмереж і різні типи навантаження. Інтеграція з iperf3, tcpreplay чи Scapy дає змогу аналізувати як продуктивність NAT, так і вплив портових конфліктів на системи захисту.

GNS3 натомість фокусується на емуляції реального мережевого обладнання, зокрема образів Cisco, Mikrotik та VyOS. Це дозволяє відтворювати специфічні механізми NAT, унікальні для комерційних реалізацій, і порівнювати їх роботу з поведінкою Linux NAT.

Основою тестового середовища є Linux з iptables або nftables та conntrack. Це дає максимально детальні журнали NAT, гнучке налаштування політик і точну телеметрію, що робить Linux надійною платформою для експериментів і порівнянь з обладнанням операторського рівня.

Завдяки інтеграції Mininet, GNS3 та Linux NAT-стека формується повністю контрольоване середовище для моделювання таких ситуацій, як хибні блокування, конфлікти портів, перевантаження таблиць, а також тестування різних алгоритмів управління пропускною здатністю та систем аналізу поведінкових аномалій. Ця інфраструктура дозволяє точно відтворити умови, характерні для операторських мереж, і забезпечує надійні дані для подальших аналітичних і практичних досліджень.

2.5 Висновки до розділу

У другій частині роботи виконано аналіз методичних, технічних та

інфраструктурних аспектів, необхідних для дослідження впливу NAT і CGNAT на точність систем виявлення інцидентів та блокування. Результати дозволили сформувати вимоги до експериментальної бази, структури контрольованих датасетів і способів збору логів, що враховують особливості NAT.

У розділі 2.1 окреслено критерії вибору доступних наборів даних. Аналіз показав, що більшість публічних датасетів не містять потрібних NAT-метаданих або мають недостатню деталізацію, тому їх використання потребує значної адаптації. Це обґрунтувало необхідність створення власних контрольованих наборів даних.

У розділі 2.2 визначено вимоги до таких наборів і запропоновано методи їх формування. Показано, що для коректного аналізу потрібно поєднувати pcap-файли, flow-записи, журнали conntrack та метрики NAT. Синхронізація цих джерел є ключовою для точного відтворення взаємодії внутрішніх і зовнішніх сесій, а також для навчання ML-моделей і аналізу хибних спрацювань.

У підрозділі 2.3 досліджено методи збору NAT-специфічних показників. Поєднання даних conntrack, iptables і пакетних дампів дозволяє простежити повний життєвий цикл трафіку. Окремо підкреслено важливість точної синхронізації часових міток та фіксації повторного використання портів.

У розділі 2.4 розглянуто інструменти для побудови тестового середовища. Mininet, GNS3 та Linux-платформи дають змогу моделювати різні конфігурації NAT/CGNAT, варіювати параметри навантаження та отримувати стандартизовані набори даних у контрольованих умовах.

Загалом у розділі було сформовано методологічну та технічну основу, необхідну для проведення подальших експериментальних досліджень. Запропонований підхід дозволяє коректно враховувати специфіку адресного шарингу та мінімізувати викривлення результатів аналізу. Сформована база слугуватиме фундаментом для проведення експериментів і практичної перевірки ефективності розробленого методу, що буде детально розглянуто у третьому розділі роботи.

3 МЕТОД ВИЯВЛЕННЯ АТАКИ ТИПУ «БЛОКУВАННЯ IP ЧЕРЕЗ NAT»

3.1 Архітектура методу джерела даних, збір логів, нормалізація параметрів

Проєктування методу виявлення атак типу «блокування IP через NAT» потребує побудови цілісної архітектури, яка б була здатна збирати, узгоджувати та нормалізувати дані з багатьох джерел одночасно. На відміну від звичайних систем виявлення аномалій, які дивляться на поведінку окремої IP чи конкретного потоку, тут ми маємо справу з зовсім іншою реальністю, за однією зовнішньою адресою одночасно працює десятки чи сотні реальних користувачів. Через це класичні підходи просто не спрацьовують.

Найбільша складність - зібрати разом різні шари телеметрії й побудувати з них єдину, правдиву картину того, що насправді відбувається з мережевими сесіями. Сама по собі зовнішня IP-адреса вже давно не є надійним ідентифікатором користувача, а окремо взяті джерела (pcap-захоплення, журнали WAF, flow-статистика чи iptables-логи) дають лише уривки пазла й не мають достатнього контексту, щоб правильно вирішити блокувати чи не блокувати.

Тому вся архітектура методу будується на багатошаровій інтеграції даних і жорсткій синхронізації за часом між усіма цими різнорідними джерелами - адже вони відрізняються і точністю міток часу, і форматами, і самою природою інформації. Лише коли всі ці шматочки складаються в одну узгоджену хронологію подій, з’являється можливість достовірно відрізнити реальну атаку від звичайної активності багатьох легітимних користувачів за NAT.

У рамках методу виділяють три ключові групи джерел даних. Перша група - пакетний рівень, що включає повноцінні дампи трафіку у форматі pcap. Ці дані надають точні часові мітки, повну структуру пакетів і дозволяють аналізувати як транспортні, так і прикладні параметри. Однак такий підхід характеризується значним обсягом даних і потребує спеціалізованих інструментів для збору та обробки.

Друга група охоплює подієві журнали системного рівня, такі як syslog-повідомлення, журнали міжмережевих екранів (firewall/iptables), оновлення NAT-

таблиць і сигнали від систем захисту (наприклад, WAF або CDN). Такі події зазвичай містять інформацію про рішення системи або зміни стану з'єднання, однак не передають деталізації на рівні окремих пакетів.

Третя група представлена агрегованими потоками (flows), які передаються у форматах NetFlow/IPFIX. Вони забезпечують узагальнену статистику про функціонування NAT-пулу, тобто кількість потоків, об’єми трафіку, середню активність клієнтів, частоту повторного використання портів тощо.

Усі три групи джерел є взаємодоповнюючими, і реалізація методу повинна бути побудована таким чином, щоб підтримувати їх одночасне використання.

Стандартизація журналів та уніфікація їхньої структури є однією з ключових складових архітектурних рішень у сучасних системах. Зокрема, широко застосовується формат RFC 5424 (The Syslog Protocol), який визначає суворі правила для представлення службових параметрів, часових міток, ідентифікаторів процесів і структурованих даних [64]. Використання цього стандарту забезпечує уніфіковане трактування подій незалежно від джерела їх генерації, що має принципове значення для подальшої кореляції з пакетними даними та даними типу flow.

Ще одним вагомим компонентом стандартизації є впровадження єдиних часових форматів, серед яких ISO 8601 із підтримкою мікросекундної точності міток часу, а також механізми синхронізації компонентів за допомогою NTP (Network Time Protocol) або, у випадку високоточних систем, PTP (Precision Time Protocol). Результати наукових досліджень у сфері лог-менеджменту та кореляції часових подій демонструють, що навіть найменші розбіжності, на рівні 1–5 мілісекунд – можуть істотно ускладнити або навіть унеможливити точну реконструкцію життєвого циклу NAT-binding [65].

Збір логів у реальному часі вимагає використання архітектури для потокової обробки даних, що забезпечує ефективність та гнучкість у роботі з великими обсягами інформації. Стандартним підходом у таких випадках є впровадження багаторівневої системи, де перший рівень - це конвеєр для обробки на етапі ingestion. Він реалізується за допомогою інструментів, таких як Fluentd, Logstash

або Vector, і відповідає за приймання журналів, їхній попередній аналіз і перенаправлення до сховища даних.

Другий рівень системи складається з централізованого зберігання даних у платформах, таких як Elasticsearch, ClickHouse або Hadoop/S3. Тут журнали зберігаються як у сирому форматі, так і у нормалізованому вигляді. Подібний підхід відповідає ключовим рекомендаціям NIST щодо управління життєвим циклом журналів, які підкреслюють важливість паралельного зберігання необроблених даних (raw logs) разом із їхньою трансформованою версією для потреб різноманітної аналітики.

Для забезпечення онлайн-обробки подій архітектура включає додатковий шар потокової обробки (stream-processing), який будується на базі рішень типу Kafka Streams або Apache Flink. Цей рівень здійснює аналіз даних у реальному часі, формує агреговані виводи та готує інформацію для подальшої роботи, зокрема для виявлення аномалій чи інших цілей аналітичного характеру.

Окрему увагу в архітектурі займає нормалізація параметрів, яка виступає мостом між потоками сирих логів і подальшим модулем аналітики. Нормалізація передбачає перетворення різнорідних даних до єдиної структури, що включає в себе:

– уніфікацію часових міток;

– приведення різних форматів IP-адрес до стандартного подання;

– фіксацію ідентифікаторів сесій, flow-рівнів та NAT-binding ID;

– вилучення дубльованих записів;

– усунення аномальних значень, спричинених sampling або апаратними втратами [67].

Якщо цей етап виконано некоректно, подальший детектор буде приймати рішення на основі викривлених даних, що неминуче призведе до хибних позитивів або хибних негативів.

У процесі нормалізації та агрегації системі слід враховувати, що в NAT-середовищі часто формуються агреговані сигнатури, які поєднують поведінку десятків або навіть сотень користувачів. Тому архітектура системи повинна

включати механізми, які дозволяють відновлювати поведінку окремих внутрішніх клієнтів, використовуючи доступні дані, такі як внутрішні IP-адреси, порти, часові патерни та унікальні характеристики прикладного рівня. Для цього застосовуються підходи, які забезпечують відновлення зв'язності сегментів потоку (flow), аналізують послідовність подій TCP-handshake, а також корелюють syslog-повідомлення із записами в NAT-таблицях [64, 66].

Підсумовуючи, перша частина архітектури методу полягає у створенні багаторівневої інфраструктури, здатної збирати різні типи журналів, об'єднувати їх у єдину структуру та забезпечувати достатню точність даних. Це дає змогу аналізувати контекст NAT-пулів, відновлювати приховані зв'язки між подіями та коректно виявляти атаки типу «блокування IP через NAT».

Наступним важливим компонентом є модуль обробки подій, який формує уніфіковане представлення мережевої активності. Він інтегрує packet capture, flow-дані, syslog, оновлення NAT-таблиць, а також журнали WAF і CDN, створюючи єдиний «журнал життєвого циклу» кожної сесії. Через специфіку NAT система повинна забезпечувати узгодженість подій навіть тоді, коли вони надходять з різних джерел і мають різний формат.

Ключовим елементом виступає модуль кореляції подій. Традиційні підходи IDS/IPS, які ґрунтуються на IP-адресах і портах, непридатні у NAT-середовищах через повторне використання портів і швидкі зміни станів. Тому застосовується мультиключовий профіль, що враховує часові зв'язки, шаблони доступу, потоки даних, зміни NAT-таблиць та прикладні атрибути, такі як User-Agent. Це дозволяє розрізняти внутрішні сесії, навіть якщо вони представлені одним зовнішнім IP.

На цьому ж етапі виконується попередня обробка даних фільтрація шумів, видалення неповних або дубльованих записів та виправлення технічних помилок. Таке очищення помітно підвищує точність подальших моделей і аналізу.

Після фільтрації система переходить до етапу багаторівневої нормалізації, у межах якого події з різних шарів перетворюються у стандартизовані структурні об'єкти. Цей процес включає:

– приведення логів до спільного timestamp-формату (мікросекундна

точність);

– трансформацію IP-полів у структурований формат із зазначенням ролі адреси (внутрішня, зовнішня, транзитивна);

– додавання інформації про NAT-binding-події;

– побудову послідовностей TCP/UDP-поводження;

– створення flow-профілю на основі пакетного трафіку.

На цьому етапі формується уніфікований журнал подій (Unified Event Log - UEL), який уже можна використовувати для детекції аномалій. Усі події отримують спільну схему - EventID, SessionID, Timestamp, TransportKey, NATBindingID, AIScore, ApplicationMetadata тощо. Для NAT-сценаріїв дуже важливим є підтримка окремого поля NATBindingID, яке відображає реальний життєвий цикл відповідності «внутрішній IP і порт → зовнішня IP і порт», що дозволяє моделі визначати, чи була активність користувача дійсно аномальною, чи вона стала наслідком агрегованої поведінки NAT-пулу.

Окремим інженерним аспектом у цій архітектурі є важливість забезпечити масштабованості, оскільки такі системи збору логів, особливо pcap та flow-джерела, генерують десятки чи навіть сотні гігабайт інформації на добу. Згідно з рекомендаціями сучасних досліджень щодо масштабованих log-pipelines, оптимальна модель передбачає розміщення ingestion-сервісів ближче до джерела даних, використання message-черг (Kafka, NATS) для передачі подій та подальшу обробку у розподіленому середовищі [68]. Така архітектура дозволяє мінімізувати втрати даних у разі пікових навантажень, що особливо важливо у процесах аналізу NAT-пулів, де навіть кілька пропущених пакетів можуть ускладнити реконструкцію потоку.

Останнім елементом архітектури є формування поведінкових профілів на основі узгоджених логів. Для цього використовуються агреговані показники, такі як швидкість створення потоків, частота зміни портів, рівень повторного використання NAT-зв’язків, середній обсяг трафіку на сесію та кількість одночасних TCP-handshake. Такі профілі зберігаються у вигляді часових рядів і слугують шаблонами для виявлення атак типу «блокування IP через NAT».

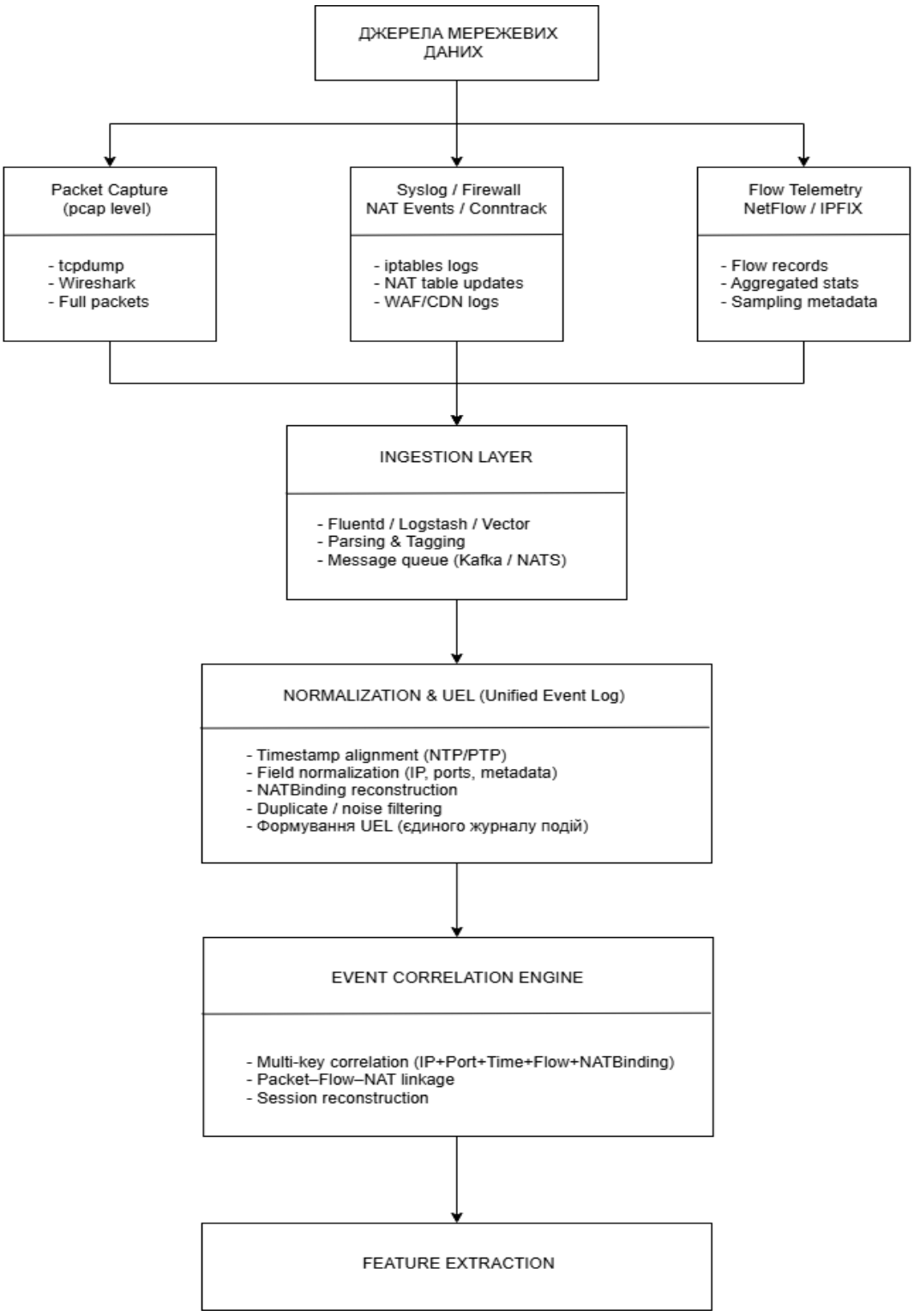

Рисунок 3.1 – Архітектура збору, нормалізації та кореляції даних у NAT-середовищі

Після завершення процесів нормалізації та кореляції подій, архітектура запропонованого методу переходить до заключного етапу - підготовки даних для модуля виявлення аномалій, який буде детально розглянутий у наступному розділі.

На цьому етапі система вже має у своєму розпорядженні повністю структурований та узгоджений журнал подій, що адекватно відображає реальну динаміку поведінки NAT-пулів, зміни станів мережевої інфраструктури та характер взаємодій користувачів з зовнішніми сервісами. Основною метою фінального етапу є перетворення цих подій у такий формат, який забезпечує ефективну обробку як за допомогою евристичних методів виявлення, так і алгоритмів машинного навчання.

Першим завданням є трансформація UEL-журналу у послідовності сесій і групування логічно пов'язаних подій. На цьому етапі створюється Session-структура, яка включає такі елементи, як часові межі сесії, кількість пакетів, обсяг переданого трафіку, ключові події (SYN, FIN, помилки), зміни станів NAT-binding та результати перевірок на стороні сервісу (активація WAF-тригерів, досягнення порогів rate-limit, статуси HTTP-запитів). Ця структура є основною, оскільки саме на рівні сесії найчіткіше проявляються аномалії в поведінці NAT надмірно часті короткочасні з’єднання, нетипові паузи між пакетами чи надмірне перепризначення портів. Виявлення таких відхилень дозволяє фіксувати можливу присутність прихованих атак або фонового ботнет-трафіку.

Наступним кроком є агрегація параметрів у часових інтервалах, що дає змогу перетворити потік подій у статистичні узагальнення. Ці часові вікна можуть тривати 1, 5 або 30 секунд залежно від особливостей трафіку. На цьому етапі формуються такі метрики, як середня кількість нових сесій, щільність перепризначення портів, частота невдалих TCP-рукостискань, дисперсія інтервалів між запитами, кількість внутрішніх клієнтів, які одночасно потрапляють під однакові WAF-фільтри тощо. Саме ці агреговані дані дозволяють виявляти аномалії, які не видно на рівні окремих пакетів або подій. Наприклад, одиничне перепризначення порту є звичною роботою NAT, проте понад 300 змін порту за одну секунду можуть свідчити про атаку з використанням скриптів або нетипову поведінку пристроїв IoT за NAT-шлюзом.

Третій елемент фінального етапу полягає у перевірці узгодженості даних, тобто виявленні потенційних логічних суперечностей, які могли виникнути через втрату пакетів, пропуски у системних журналах або затримки у передачі подій.

Архітектура повинна включати модуль, що виявляє такі переривання та позначає їх як невизначені ділянки, щоб запобігти їхньому сприйняттю детектором як аномалій. Наприклад, якщо запис NAT містить інформацію про створення binding, але відсутня подія його завершення, система повинна правильно завершити сесію за часовим обмеженням або помітити її як частково відновлену. Це має ключове значення для запобігання використанню спотворених даних у моделях машинного навчання.

Завершальним етапом є підготовка уніфікованого вхідного формату, який передається до модуля детекції. Для евристичних методів це можуть бути таблиці з ключовими показниками NAT-поведінки. Для алгоритмів машинного навчання - матриці ознак, векторизовані сесії або часові ряди. На цьому етапі важливо, що система вже здатна відокремити ефекти, спричинені NAT-агрегацією, від індивідуальної поведінки користувачів. Це дає змогу детектору більш точно визначати випадки, коли блокування IP пов’язане із зовнішньою активністю.

Таким чином, третій компонент архітектури завершує обробку даних, перетворюючи сирий трафік у структурований і послідовний набір параметрів, готовий для аналітичного або машинного аналізу. Подальша ефективність методу цілком залежить від якості виконання цього етапу. Лише за умови належної побудови UEL, коректної кореляції подій і перевірки консистентності можна досягти високої точності у виявленні атак типу «блокування IP через NAT».

3.2 Набір ознак для виявлення зловмисної NAT-активності

Виявлення атак типу блокування IP через NAT потребує визначення чітких характеристик, які дозволяють розрізнити природні колективні патерни трафіку від аномалій, викликаних діями зловмисників серед клієнтів, що користуються одним NAT-пулом. Оскільки зовнішній сервіс сприймає NAT-пул як єдину IP-адресу, більшість традиційних показників втрачають свою ефективність кількість запитів з IP, частота створення TCP-сесій або середній RPS більше не забезпечують коректну

оцінку активності окремого користувача. Це вимагає розробки методу, заснованого на внутрішніх наборах метрик, які враховують неоднорідність поведінки клієнтів, прихованих за однією IP-адресою.

Одним із центральних аспектів аналізу є часові характеристики поведінки NAT, зокрема ентропія інтервалів між сесіями, варіативність темпу створення нових потоків і дисперсія тривалості TCP або UDP-сесій. У дослідженнях, проведених Fawaz з колегами [70], встановлено, що запити, здійснені через NAT зловмисниками, схильні відрізнятися від загального фонового трафіку. Ці відмінності проявляються в однотипності інтервалів між запитами, стабільності частоти запитів або короткотривалих спалахах активності, притаманних роботам і автоматизованим сканерам. Подібні характеристики дають змогу ідентифікувати аномалії навіть за умов високого нормального показника RPS для видимої зовнішньому сервісу IP-адреси.

Портові та поведінкові ознаки, пов'язані з NAT-таблицею, відіграють важливу роль. Вони включають швидкість перепризначення портів, повторне використання портів протягом коротких проміжків часу, а також співвідношення кількості активних потоків до розміру NAT-пулу. У дослідженнях Ren та Chen [71] зазначено, що шкідлива активність у NAT-середовищі може викликати аномальну концентрацію портів і підвищену частоту короткотривалих сесій, які завершуються без належного завершення TCP-діалогу. Ці характеристики можуть слугувати надійними маркерами як для виявлення DDoS-атак, так і для визначення наявності заражених пристроїв, зосереджених у межах одного NAT середовища.

До окремої категорії належать ентропійні характеристики, такі як User-Agent, TLS-фінгерпринти, DNS-патерни та HTTP-запити, які допомагають оцінювати внутрішню різноманітність трафіку. У дослідженні Shahriar та співавторів показано, що легітимні NAT-кластери завжди характеризуються високою ентропією поведінкових і протокольних атрибутів. У той час зловмисні агенти або ботнети за NAT здебільшого вирізняються стандартизованими сигнатурами, ідентичними рядками User-Agent, повторюваними параметрами TLS ClientHello та однотипними шаблонами HTTP. Зниження рівня ентропії зазначених параметрів

може слугувати точним сигналом появи зловмисної активності за NAT.

Однією з важливих категорій ознак є співвідношення між внутрішньою активністю клієнтів та зовнішніми реакціями сервісу, що проявляється, наприклад, у вигляді блокувань механізмами захисту веб-додатків (WAF), спрацьовування тригерів обмеження швидкості (rate-limit) або активації CAPTCHA. У випадку, коли лише окрема група клієнтів у NAT-кластері демонструє поведінку, що викликає санкційні дії з боку сервісу, формуються асиметричні профілі подій. Подібна диспропорція слугує характерною ознакою поведінки окремого зловмисника, що, своєю чергою, негативно позначається на репутації всієї IP-адреси. Виявлення таких шаблонів дозволяє алгоритмам оптимізувати процес і знизити кількість помилкових блокувань, що являє собою першочергову практичну мету досліджуваного проєкту.

Таким чином, набір характеристик для виявлення зловмисної NAT-активності має включати різноманітні параметри часові, статистичні, ентропійні та протокольні. На відміну від традиційних методів забезпечення мережевої безпеки, цей підхід враховує різнорідність внутрішнього трафіку в NAT-пулі та орієнтується на виявлення окремих зловмисників, а не всього пулу адрес. Це дозволяє створити методологію, яка здатна суттєво зменшити кількість необґрунтованих блокувань зовнішніх сервісів.

Формування набору ознак для виявлення зловмисної активності у NAT-пулі потребує створення багаторівневої системи характеристик, яка відображає як поведінку окремих клієнтів, так і роботу NAT-механізмів та реакцію зовнішніх сервісів. На відміну від традиційних систем IDS/IPS, що здебільшого аналізують мережецентричні параметри (наприклад, IP-адреси, обсяг трафіку чи частоту запитів), NAT-орієнтований підхід має фокусуватися на внутрішній структурі трафіку. Це включає аналіз різноманітних патернів клієнтів, що використовують спільну зовнішню адресу. У зв'язку з цим сучасні дослідження пропонують розглядати NAT-пул не як єдину логічну сутність, а як своєрідний «мікропростір», який характеризується власною динамікою, статистичними показниками та тимчасовими аномаліями.

Одним із основних напрямів створення ознак є виявлення внутрішніх груп активності, які характеризуються різними моделями поведінки. У NAT-кластерах легітимні користувачі демонструють широкий спектр варіацій поведінки, таких як нерівномірність часових інтервалів, різноманітність User-Agent, а також висока ентропія HTTP-параметрів. Дослідження Morillo та співавторів звертає увагу на те, що зловмисники або групи ботів у NAT мають ознаки «поведінкової компактності» [74]. Це проявляється через стабільні цикли часу, однакові заголовки запитів, повторне використання портів у вузьких діапазонах. У сучасних моделях створюються метрики для оцінювання внутрішньої неоднорідності NAT-пулу, такі як ентропія субпотоків, різноманітність транспортних характеристик, співвідношення коротких і довготривалих сесій та параметри синхронності подій.

Окрема категорія характеристик стосується адекватних змін у поведінці NAT-пулу під впливом аномальної активності. У випадках, коли один або кілька клієнтів у пулі починають генерувати надмірну кількість сесій або створюють нетипові послідовності запитів, це знаходить відображення у зміні NAT binding table. Зокрема, спостерігається збільшення частоти перепризначення портів, зростання кількості односпрямованих потоків та скорочення середньої тривалості життя TCP-з’єднань. У дослідженні, представленому Silva та співавторами, здійснено детальний аналіз цих явищ [75]. Виявлено, що навіть за незначного обсягу трафіку зловмисна активність породжує так званий «мультиплікативний відбиток» у NAT-таблицях, який суттєво відрізняється від нормальної поведінки системи. Подібні характеристики є особливо ефективними для створення евристичних правил, спрямованих на ідентифікацію аномалій.

Ще одним важливим аспектом є встановлення зв'язку між внутрішніми показниками та реакціями зовнішніх сервісів, таких як CDN, WAF або системи обмеження швидкості запитів. Наприклад, якщо WAF починає відхиляти частину запитів, а внутрішні метрики свідчать про зниження ентропії чи синхронізацію дій декількох потоків, це може вказувати на активність ботнету в межах NAT-пулу. Такі сукупні сигнали дають можливість системі приймати рішення з високою точністю, мінімізуючи кількість хибних блокувань і відсікаючи невідповідності,

властиві легітимним користувачам.

Формування набору ознак не слід розглядати як просту механічну вибірку статистичних параметрів це складний процес, що передбачає створення багаторівневої структурованої моделі для опису поведінкових характеристик NAT-пулу. Зазначена модель охоплює такі аспекти, як часові, транспортні, ентропійні, кореляційні та реактивні ознаки. У подальшому, як буде детально розглянуто у підрозділі 3.3, ці характеристики можуть бути використані для інтеграції у евристичний алгоритм або модель машинного навчання з метою ідентифікації потенційних зловмисників у NAT-середовищі.

Рисунок 3.2 – Групи ознак для виявлення зловмисної NAT-активності

Рисунок 3.2 демонструє структуровану модель набору ознак, що застосовуються для виявлення зловмисної NAT-активності. Модель охоплює чотири ключові групи параметрів, як часові, транспортні, кореляційні та ентропійні ознаки. Часові характеристики відображають динаміку утворення сесій і закономірності їх виникнення. Транспортні параметри характеризують поведінку на рівні TCP/UDP-з’єднань, зокрема повторне використання портів та зміни NAT-зв’язувань. Кореляційні ознаки дозволяють аналізувати взаємозв’язки між потоками та реакціями зовнішніх систем захисту. Ентропійні параметри відображають рівень різноманітності поведінкових характеристик усередині NAT-пулу. У сукупності ці групи ознак забезпечують основу для глибшого аналізу та створення алгоритмів виявлення атак, таких як «блокування IP через NAT».

3.3 Алгоритм детекції (евристичний або ML) та порогові правила

Алгоритм для виявлення атаки типу «блокування IP через NAT» повинен враховувати особливості спільного використання IP-адрес та специфічні наслідки, що виникають через об'єднання ресурсів у NAT-пулі. На відміну від традиційних підходів до детекції, які зазвичай орієнтуються на аналіз окремих IP-адрес, у контексті NAT необхідно використовувати більш комплексний підхід, що враховує сукупність поведінкових, часових та кореляційних характеристик потоків. У цьому зв’язку ключовим завданням є розроблення детектора, здатного розрізняти локальні аномалії, що пов’язані з окремими користувачами, та глобальні шаблони шкідливої активності. Останні зазвичай проявляються через неприродну концентрацію аномальних подій у межах одного NAT-пулу, що може свідчити про цілеспрямовані атаки.

Першим кроком у створенні алгоритму є визначення концептуальної моделі евристичної, заснованої на машинному навчанні або змішаної (гібридної). Евристичний підхід базується на використанні попередньо встановлених правил, які розроблені з урахуванням спостережуваних аномалій, таких як надмірна частота формування NAT-з'єднань, синхронність HTTP-запитів, різке зниження ентропії поведінкових параметрів, зменшення варіативності User-Agent тощо. Цей тип алгоритмів добре підходить для швидкого реагування та забезпечує повний контроль над процесом прийняття рішень. Дослідження Singh та співавторів підтверджують, що правильно налаштовані порогові політики здатні забезпечити високу точність у NAT-середовищах, де моделі на основі машинного навчання можуть давати неоднозначні результати через змішану природу потоків [76].

Проте, як показують нам, сучасні дослідження, включно з роботою Torres і Delgado, поведінка трафіку за NAT у багатьох випадках виявляється надто складною, щоб бути описаною суто евристичними правилами [77]. Зміна інтенсивності потоків, адаптація ботнетів, використання розподілених командно-контрольних схем або трафіку низької інтенсивності роблять евристику недостатньо гнучкою. У таких випадках значно доцільніше буде застосувати

моделі машинного навчання, які здатні навчатися на високорозмірних даних, використовувати взаємозв’язки між ознаками та адаптуватися до нових видів атак.

Досить серйозною перевагою ML-підходу є здатність виявляти зовсім неочевидні кореляції всередині NAT-пулу, наприклад, зміни в структурі flow-записів, асиметрію між внутрішніми та зовнішніми таймінгами, а також комплексні залежності між транспортними та ентропійними характеристиками. Проте ці моделі потребують значного обсягу контрольованих даних, коректно розмічених сценаріїв та іноді - симуляцій складних NAT-умов. Саме через це в підрозділах 2.2 та 2.4 була детально описана необхідність формування контрольованих наборів даних та побудови спеціалізованого тестового середовища.

Дуже важливим аспектом алгоритму є адаптивність. В умовах NAT пул може містити від десятків до тисяч активних користувачів. Зміни в його поведінці відбуваються швидко, а їхній контекст не завжди однозначний наприклад, збільшення кількості коротких TCP-сесій може бути природним для мобільних додатків або ж бути ознакою сканування. Як зазначає Zhang у роботі, алгоритми детекції, які враховують лише абсолютні пороги, мають тенденцію до деградації точності у високонавантажених NAT-пулах [78]. Тому механізми порогових правил повинні враховувати відносні характеристики нормалізовані значення, відхилення від базових профілів, варіації між субпотоками та внутрішню ентропію пулу.

Ядро алгоритму має базуватися на структурі ознак, сформованій у підрозділі 3.2, та включати три ключові компоненти:

– модуль базового профілювання NAT-пулу. Формує «нормальну картину» поведінки шляхом аналізу статистики NAT-binding, flow-даних, рівнів ентропії та агрегованих реакцій зовнішніх сервісів;

– модуль виявлення локальних аномалій. Орієнтований на фіксацію поведінкових сплесків, нетипової активності окремих внутрішніх клієнтів, раптової синхронізації подій або неприродного зниження варіативності параметрів;

– модуль глобальної кореляції аномалій. Використовує часові та структурні зв’язки між аномаліями на рівні пулу, аналізує їх розподіл, щільність у часі,

наявність нехарактерних патернів зростання або повторюваних структур;

Ця модель яка складається з трьох шарів, дозволяє будувати адаптивні детектори, які можуть бути реалізовані або у формі складної системи порогових правил, або у вигляді ML-класифікатора. У першому випадку система має більшу прозорість рішень, у другому - вищу гнучкість і здатність до самонавчання.

Завершальним елементом першої частини є формування механізмів порогового прийняття рішень. Пороги повинні визначатися на основі статистичного розподілу параметрів NAT-пулу, контролю верхніх процентилей (наприклад 95-го або 99-го), а також аналізу історичних даних. У ML-моделях порогові правила можуть виступати як постпроцесинговий (після процесу) етап для валідації результатів класифікації та зниження кількості хибнопозитивних спрацьовувань.

Побудова алгоритму виявлення зловмисної активності за NAT-пулом потребує формування цілісної архітектури, яка поєднує в собі декілька рівнів обробки даних, від первинного збору та стандартизаії до кореляції подій і прийняття рішень. Оскільки NAT приховує індивідуальність клієнтів, традиційні підходи (такі як прості порогові правила) не можуть забезпечити достатню точність. Тому запропонований алгоритм повинен бути багатошаровим, адаптивним і здатним враховувати цілісну поведінку пулу, а не окремих IP-адрес.

Перший рівень попередньої обробки та відновлення контексту (Preprocessing Layer). Цей рівень виконує функції, пов’язані саме з очищенням даних, вирівнюванням часових міток, відокремленням корисних сигналів від шуму. Оскільки NAT часто створює додаткові часові затримки або групує події в коротких кластерах, алгоритм повинен компенсувати ці ефекти. Наприклад, UDP-зв’язування мають дуже короткий час життя (мілісекунди), тоді як TCP-зв’язування можуть жити секундами або хвилинами - відповідно, часова нормалізація повинна враховувати тип цього трафіку. На цьому етапі ще виконується реставрація NAT-контексту - встановлення зв’язків між внутрішнім ефективним поведінковим патерном та зовнішньою активністю. Саме тут у нас формуватиметься базовий матеріал для подальшого аналізу, таблиці повторного використання портів,

розподіли flow-тривалостей, частота NAT-binding, інтенсивність запитів.

Наступний рівень - це рівень кореляції (Correlation and Pattern Reconstruction Layer). Другим ключовим етапом є побудова кореляційної моделі. Метою є визначення поведінкової «взаємопов'язаності» подій у NAT-пулі. У цьому контексті кореляція може бути часовою, тобто оцінка синхронності сплесків активності різних потоків. Структурною це коли відбувається аналіз схожості User-Agent, TLS-fingerprints, HTTP-параметрів. Мережевою тобто співставлення транспортних параметрів, повторного використання портів, патернів пакетів. І зовнішньою - порівняння реакцій CDN/WAF із внутрішньою поведінкою пулу.

Практика демонструє, що кореляційні зв'язки найкраще розкривають приховані аномалії, зловмисні клієнти дуже рідко поводяться хаотично - їхня активність зазвичай структурована та часто повторювана. Це може дозволити виявляти навіть малопомітні атаки низької інтенсивності.

Дослідження Gross та Niemann підтверджує, що комбіновані кореляційні ознаки (часові + транспортні + ентропійні) дуже помітно підвищують точність ML-класифікації у середовищах із перерозподіленими IP-адресами [79]. Таким чином, алгоритм повинен включати модуль кореляції як обов'язковий компонент, незалежно від того, чи застосовується евристика, чи машинне навчання.

Наступним рівнем є модуль детекції (Detection Engine Layer), який приймає рішення щодо наявності зловмисної активності. Він працює у двох режимах. Перший режим - евристичний детектор, що ґрунтується на попередньо визначених правилах. Наприклад, «частота створення NAT-binding перевищує 99-й процентиль → можливе сканування», «різке падіння ентропії User-Agent у NAT-пулі → ознака ботнету», «велика кількість однотипних HTTP-запитів за короткий час → ймовірний flood». Перевага евристик полягає в їхній прозорості, що важливо для CDN/WAF, але правила швидко застарівають при зміні поведінки зловмисників.

Другий режим - ML-детектор. Моделі машинного навчання використовують набір ознак (див. підрозділ 3.2) як вхід для алгоритмів на основі дерев рішень, випадкових лісів, бустингу чи нейронних мереж. На відміну від евристик, ML здатний виявляти складні та приховані закономірності. Як показано в роботах Otto

та Frick, точність ML-класифікаторів у NAT-середовищі може перевищувати евристичні методи на 15–25% за наявності якісних контрольованих даних [80]. Викликом залишається формування валідної навчальної вибірки, що детально розглянуто у підрозділах 2.2–2.4.

Останнім рівнем архітектури є рівень адаптивних порогів (Adaptive Thresholding Layer), який відповідає за прийняття фінальних рішень щодо виявлення аномальної або зловмисної активності. Незалежно від того, чи використовується евристичний підхід, чи моделі машинного навчання, процес детекції потребує застосування динамічних порогових значень. Це зумовлено тим, що поведінка NAT-пулів не є сталою у часі та може суттєво змінюватися залежно від багатьох факторів, зокрема часу доби, типів клієнтів, інтенсивності та структури трафіку, сезонних коливань навантаження, а також специфіки використовуваних додатків, таких як веб-сервіси, мобільні застосунки або ігрові платформи.

У зв’язку з цим система повинна підтримувати кілька рівнів порогів. До них належать локальні пороги, які застосовуються до окремих ознак або потоків, глобальні пороги, що враховують агреговані показники всього NAT-пулу, а також комбіновані порогові моделі. Останні можуть базуватися, наприклад, на обчисленні інтегрального ризикового балу, який узагальнює інформацію з кількох поведінкових, часових та кореляційних характеристик. Такий підхід дозволяє уникнути жорсткої прив’язки до фіксованих значень і зменшити кількість хибних спрацювань у періоди природних піків активності.

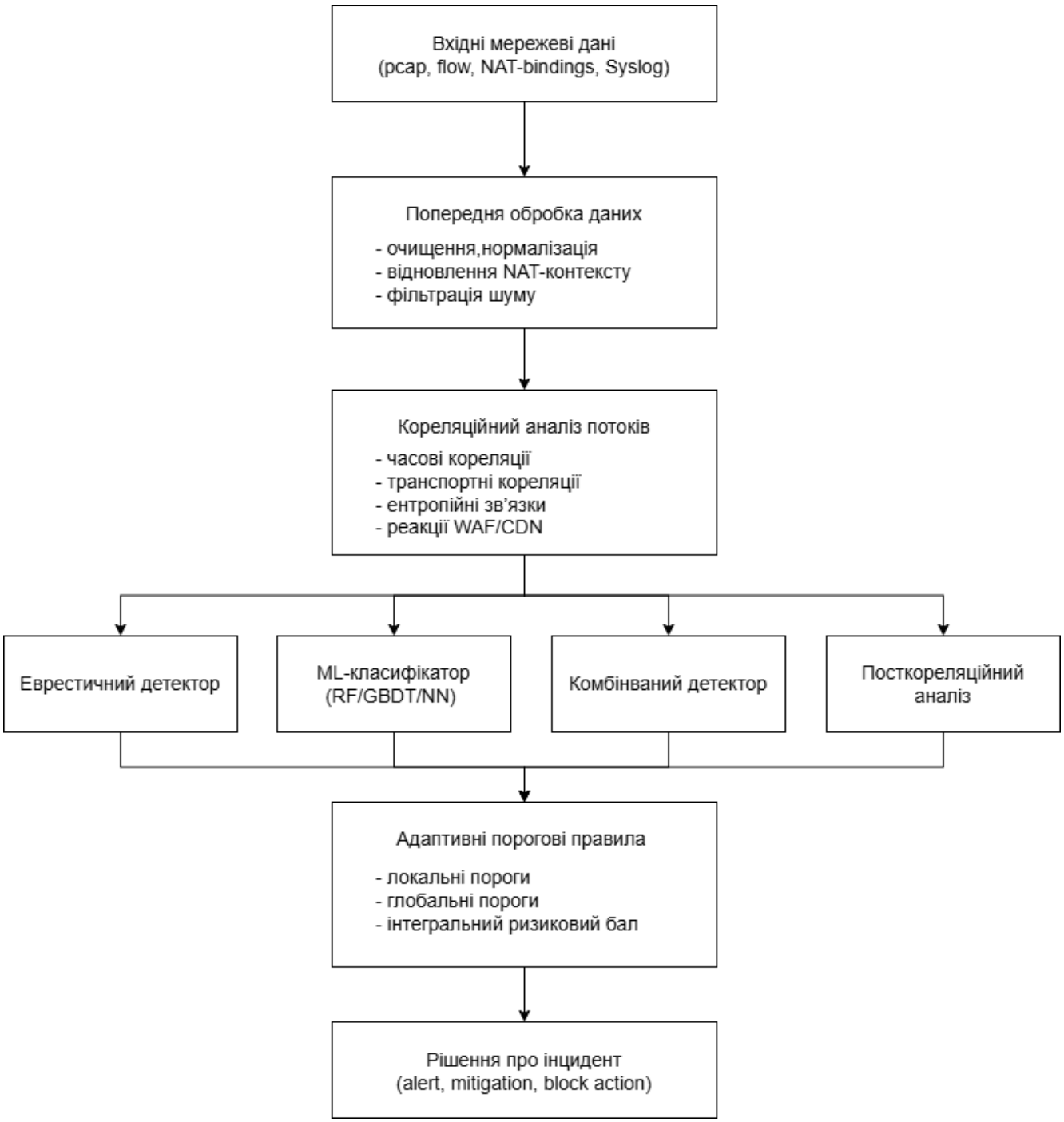

Рисунок 3.3 – Архітектура алгоритму детекції зловмисної активності у NAT-середовищі

У реальних умовах експлуатації, алгоритм детекції повинен демонструвати стійкість до змінної якості даних, неоднорідної інтенсивності трафіку та неповних телеметричних записів. NAT-середовище створює значні виклики для будь-якої моделі аналізу, оскільки характер даних, доступних для обробки, часто суттєво варіюється залежно від часу доби, навантаження або роботи мережевої

інфраструктури. Саме тому важливо, щоб алгоритм міг функціонувати навіть тоді, коли частина даних є фрагментованою, зашумленою або тимчасово недоступною. Для цього необхідною є здатність системи до реконструкції неповних часових рядів, адаптивного згладжування випадкових коливань та відновлення контексту подій у тих випадках, коли NAT-таблиці видаляються або оновлюються раніше, ніж алгоритм встиг проаналізувати їхній вміст.

Окремо слід розглянути роль механізмів самокорекції та внутрішньої валідації, які дозволяють йому зменшити кількість хибнопозитивних рішень. У NAT-пулі будь-яка аномалія може бути наслідком як дій одного користувача, так і результатом складної агрегованої динаміки цілої групи, тому алгоритм має перевіряти власні висновки на наявність підтвердження у різних типах ознак. Аналіз стійкості аномалій у часі, порівняння їх з типовими станами пулу та ретроспективна перевірка поведінки після передбачуваного інциденту дозволяють уточнювати попередні рішення та запобігати блокуванню доброчесних користувачів. Усі ці механізми сприяють тому, щоб система реагувала лише на обґрунтовані та систематичні відхилення від нормальної роботи, а не на випадкові сплески активності.

Не менш важливим аспектом є інтеграція алгоритму у ширшу інфраструктуру мережевого контролю. Алгоритм детекції працює не автономно, а у взаємодії з WAF-платформами, мережевими фільтрами, системами управління політиками доступу та сервісами моніторингу. Тому його результати мають бути представлені у вигляді структурованих сигналів, які зможуть бути інтерпретовані зовнішніми компонентами. Також дуже важливо забезпечити можливість передачі підтверджувальних метрик, які пояснюють причини спрацювання детектора, та можливість адаптивної реакції з боку інфраструктури, наприклад шляхом поступового посилення режимів фільтрації або застосування обмежених втручань, якщо рівень аномалії ще не є критичним.

У цьому контексті особливо важливим є питання балансування між точністю виявлення та ризиком блокування доброчесних користувачів. У системах масового адресного шарингу одне навіть помилкове рішення може вплинути на велику

кількість людей, тому реакція алгоритму має бути поступовою, а не різкою. Наприклад, зафіксовані поведінкові відхилення не повинні автоматично призводити до блокування трафіку, якщо вони не мають достатньої стійкості або не підтверджені іншими ознаками. У цьому процесі дуже важливе значення мають адаптивні порогові правила, які формуються на основі історичних характеристик пулу, динаміки навантаження та зміни поведінкових патернів у різні періоди часу.

Порогові механізми відіграють стабілізуючу роль у функціонуванні такого алгоритму. Вони дають йому змогу обмежити чутливість моделі до випадкових коливань, забезпечити мінімальний рівень детекції навіть у разі часткової втрати даних та сформувати структуровану основу для прийняття рішень, які повинні підтверджуватися декількома незалежними параметрами. Таким чином, порогові правила виступають інструментом, який забезпечує узгодженість дій системи та мінімізує ризик неправильного тлумачення агрегованої поведінки NAT-пулу.

Узагальнюючи викладене, можна зазначити, що алгоритм детекції зловмисної активності в NAT-середовищі повинен мати багаторівневу структуру, здатну відновлювати контекст подій, коригувати власні рішення, взаємодіяти з мережевою інфраструктурою та враховувати змінність поведінки NAT-пулів у часі. Така модель уможливлює створення практичного прототипу, реалізація якого буде розглянута у наступному підрозділі роботи.

3.4 Реалізація прототипу та стратегії реагування

Реалізація прототипу системи детекції зловмисної активності в NAT-середовищі вимагає узгодження методологічних принципів, описаних у попередніх підрозділах, з інженерними обмеженнями та операційними вимогами до мережевих систем. Завдання полягає в тому, щоб перетворити концептуальний алгоритм на працездатну програмну систему, яка здатна функціонувати в реальному часі, обробляти дані з різних джерел телеметрії, аналізувати поведінку NAT-пулів та виявляти аномалії, не погіршуючи загальної продуктивності мережевої

інфраструктури. Прототип повинен відтворювати ключові компоненти алгоритму збір метрик, нормалізацію параметрів, реконструкцію контексту, аналіз ознак та застосування логіки ухвалення рішень.

Першим етапом реалізації є побудова архітектури, здатної підтримувати роботу з різнорідними видами даних, оскільки NAT-середовище формує комплексну картину мережевої активності. Прототип повинен приймати syslog-повідомлення від маршрутизаторів і фаєрволів, інформацію про NAT-binding, NetFlow або IPFIX-записи, pcap-захоплення та внутрішні логи сервісів, що обслуговують клієнтів. Важливо забезпечити можливість синхронізації цих джерел, оскільки різниця в часових мітках або різний формат подання інформації можуть спричинити некоректну реконструкцію подій. Для цього застосовуються механізми попередньої обробки, що включають вирівнювання часових рядів, формування узгодженої структури подій та уніфікацію формату даних. Таким чином забезпечується можливість багаторівневого аналізу, який передбачає виявлення кореляцій між потоками, подіями та реакціями зовнішніх сервісів.

Другим важливим аспектом реалізації є створення модулів динамічного профілювання NAT-пулу. Прототип має підтримувати побудову статистичних моделей, які здатні описувати характерну поведінку пулу в нормальному стані. Профіль формується на основі таких параметрів, як інтенсивність створення NAT-зв’язувань, середня тривалість потоків, ентропія внутрішньої поведінки та рівень різноманітності користувацьких параметрів. Оскільки NAT-пул динамічний і схильний до коливань, профіль повинен регулярно оновлюватися, а не бути статичним. Це дозволяє відображати як короткострокові, так і довгострокові зміни в поведінці пулу, що важливо для уникнення хибних спрацьовувань.

Наступним етапом буде впровадження модулів аналізу ознак, описаних у підрозділі 3.2. Прототип повинен обчислювати часові, транспортні, кореляційні та ентропійні характеристики для кожного періоду спостереження. На практиці це означає, що система повинна підтримувати роботу з потоками даних, які надходять у режимі реального часу, і водночас забезпечувати обчислювальну ефективність. Наприклад, для оцінки ентропії поведінкових параметрів потрібно аналізувати

великий обсяг подій, який накопичується протягом коротких часових інтервалів. Система має бути здатна підтримувати такі обчислення без надмірного навантаження, використовуючи оптимізовані алгоритми або інкрементальні моделі.

Окремої уваги також заслуговує реалізація механізму реконструкції контексту NAT-подій. Через властивість NAT приховувати індивідуальні джерела трафіку, прототип повинен відновлювати зв’язки між зовнішніми та внутрішніми подіями. Це може включати визначення повторного використання портів, аналіз структури потоків чи виявлення синхронних патернів у межах пулу. Такі механізми дозволяють оцінювати ступінь внутрішньої кореляції подій та краще виявляти спільні ознаки ботнет-активності або розподілених низькоінтенсивних атак.

Ключовим компонентом прототипу є модуль ухвалення рішень, який повинен інтегрувати як евристичні правила, так і можливі моделі машинного навчання. Евристичний блок забезпечує інтерпретованість та дозволяє застосовувати негайні реакції, тоді як моделі машинного навчання забезпечать адаптивність і здатність виявляти складні патерни, які складно описати формальними правилами. Реалізація такого модуля вимагає узгодження різних джерел сигналів, зокрема зіставлення показників ентропії, поведінкових відхилень та транспортних характеристик у єдиному просторі рішень.

Особливо важливим є механізм адаптивних порогів, тому, що він визначає, коли виявлена аномалія потребує реагування. У прототипі порогові значення не можуть бути фіксованими, оскільки інтенсивність трафіку в NAT-пулі змінюється залежно від часу доби або зовнішніх факторів. Тому реалізуються динамічні моделі, які визначають пороги виходячи із середньостатистичних характеристик останніх періодів спостереження. Це забезпечує стійкість системи до раптових короткострокових сплесків, які не повинні інтерпретуватися як атаки.

Також велику роль відіграє модуль операційної стійкості, який визначає поведінку прототипу у разі втрати або часткового пошкодження даних. У мережевому середовищі, де таблиці NAT можуть оновлюватися непередбачувано, а окремі записи flow-телеметрії можуть бути відсутніми, система повинна

застосовувати механізми відновлення інформації. Це включає інтерполяцію часових рядів, відновлення фрагментованих подій та адаптацію алгоритмів у режимі неповного контексту.

Не менш важливою є інтеграція прототипу з інфраструктурою реагування. Прототип повинен мати можливість передавати сигнали ризику у зовнішні системи, зокрема WAF-платформи або внутрішні системи моніторингу, а також підтримувати різні стратегії реагування від інформаційного сповіщення до часткового або повного блокування трафіку. Внутрішньосистемна взаємодія повинна бути побудована таким чином, щоб реакція була пропорційною рівню загрози та не шкодила доброчесним користувачам NAT-пулу.

У результаті прототип має являти собою модульну, адаптивну та здатну до самокорекції систему, інтегровану у загальну архітектуру мережевої безпеки. Його практична придатність визначається тим, наскільки ефективно він відтворює поведінку реальних систем захисту та наскільки точно може ідентифікувати аномалії у складних багатокористувацьких середовищах.

Після розробки основних функціональних компонентів прототипу важливо визначити, яким чином система повинна реагувати на виявлені аномалії. Ефективність детекції не обмежується лише процесом аналізу даних; не менш важливою є здатність системи адекватно, вчасно та пропорційно реагувати на загрози. У реальних мережевих середовищах занадто агресивні реакції можуть створити значні проблеми для доброчесних користувачів, тоді як занадто обережні – дозволити шкідливій активності розвиватися і завдавати шкоди. Тому прототип має будуватися на логіці адаптивного реагування, що враховує складність NAT-середовища, динаміку навантаження та контекст аномалій.

Одним із ключових принципів є багаторівневий підхід до блокування. Перед тим як перейти до різкого обмеження трафіку, система повинна дати можливість визначити, чи справді поведінка є шкідливою, чи це лише якийсь локальний відхил, який в свою чергу спричинений природними факторами. З цією метою прототип використовує каскадну модель реагування. На початковому рівні система може лише генерувати попередження і фіксувати аномалію вже у внутрішніх логах. Це

дозволяє накопичувати якусь статистику та оцінювати, чи повторюється підозріла поведінка. Якщо ж аномалія має ознаки стійкості або посилюється, тоді система переходить до більш рішучих дій – наприклад може застосувати обмеження швидкості або роз'єднує підозрілі потоки. Лише у разі, коли аномалія відповідає характеристикам атаки (зокрема за параметрами кореляції або агрегованих ознак), можливе повне блокування NAT-пулу або трафіку певного напрямку.

Ефективність цієї моделі сильно залежить від механізмів зворотного зв'язку, впроваджених у прототип. Система повинна не лише реагувати, але й оцінювати результат рішень, які вона прийняла. Наприклад, якщо після застосування rate limiting поведінка трафіку повертається до нормального стану, це свідчить про те, що аномалія не була критичною, і сильніші заходи просто не потрібні. Якщо ж реакція не впливає на поведінку, потрібно розглядати інші варіанти – можливо, зловмисники адаптуються до цих обмежень або використовують низькоінтенсивні методи атаки, що маскуються під легітимний трафік. Завдяки такому підходу прототип поступово формує уявлення про ефективність власних дій та здатний покращувати стратегії реагування.

Окрім внутрішніх механізмів, прототип також має взаємодіяти з іншими компонентами мережевої інфраструктури. У системах, які активно використовують WAF, CDN, фаєрволи та load balancer, рішення щодо блокування не завжди приймаються в одному місці. Прототип повинен бути здатним працювати як в ізольованому режимі, так і у ролі аналітичного модуля, що передає оцінки ризику у певні зовнішні системи. У цьому разі важливим аспектом є формат представлення результатів. Система повинна генерувати структуровані повідомлення - наприклад, у JSON-форматі, які містять опис аномалії, набір ознак, що її підтверджують, рівень ризику, а також рекомендацію щодо дій. Завдяки цьому зовнішні сервіси можуть швидко реагувати на загрози та використовувати прототип як додаткове джерело інтелектуального аналізу трафіку.

Прототип також має враховувати фактори, пов'язані з продуктивністю. NAT-середовище може генерувати значну кількість подій, і система не повинна створювати додаткове навантаження, яке саме по собі може призвести до

деградації мережі. Для цього в прототип закладаються певні механізми оптимізації, обробка даних потоками, використання буферів, інкрементальні методи обчислення ознак, а також адаптивні алгоритми скорочення даних. Наприклад, якщо у певний період NAT-пул демонструє стабільну поведінку, система може зменшити частоту оновлення внутрішніх моделей або обмежити глибину аналізу. Навпаки, у періоди різких сплесків активності або у випадку підозр на сканування чи флуда необхідно посилити моніторинг.

Важливо також визначити стратегії реагування на випадки невизначеності. Оскільки NAT маскує справжню поведінку окремих клієнтів, інколи система може фіксувати ознаки аномалії, які не мають однозначної інтерпретації. У таких ситуаціях прототип не повинен діяти категорично. Натомість в такому разі, доцільно використовувати режим «відкладеного рішення», коли протягом певного інтервалу часу система ще продовжує збирати інформацію і лише після повторного аналізу формує остаточний висновок. Це підвищує стійкість системи до хибнопозитивних подій і дозволяє уникнути зайвого втручання у роботу мережевої інфраструктури.

Реалізований прототип повинен забезпечувати можливість подальшого розширення. Сучасні мережеві атаки швидко еволюціонують, а поведінка NAT-пулів змінюється разом із характером трафіку. Тому архітектура прототипу має бути гнучкою, моделі машинного навчання повинні періодично перевчатися на нових даних, а евристичні правила – оновлюватися відповідно до нових типів аномалій. Крім того, важливо передбачити модуль для експериментальної інтеграції з іншими інструментами моніторингу, що може забезпечити комплексніший аналіз.

У підсумку друга частина підрозділу демонструє, що ефективність прототипу визначається не лише коректністю алгоритмічної частини, а й тим, наскільки якісно організовані механізми реагування, самокорекції, інтеграції та оптимізації. Така система здатна працювати у складних умовах NAT-агрегації та забезпечувати збалансоване виявлення зловмисної активності, мінімізуючи ризики для легітимних користувачів.

3.5 Налаштування тестового середовища та сценарії експериментів

Архітектура експериментального стенду, параметри середовища та моделювання NAT-умов.

Щоб оцінити ефективність запропонованого методу детекції зловмисної NAT-активності потрібно створити контрольоване середовище, яке здатне відтворювати характерні умови функціонування реальних мереж. Основним завданням експериментального стенду є забезпечення можливості точного та відтворюваного моделювання поведінки NAT-пулів, включно з такими параметрами, як швидкість ротації портів, динаміка створення потоків, вплив конкуренції між клієнтами та реакції зовнішніх сервісів на агрегацію трафіку. У межах цього дослідження структура тестового середовища спирається на компоненти, що були описані у попередніх розділах, зокрема на віртуальні мережі Mininet, емуляцію маршрутизації та NAT-механізмів у Linux, а також інструменти для фіксації трафіку (pcap), логів NAT-таблиць та flow-записів.

Конфігурація стенду передбачає створення трирівневої мережевої структури, у якій клієнтські вузли взаємодіють із зовнішнім сервером через один або декілька NAT-вузлів. Це дозволяє сформувати сценарії, що відповідають справжнім умовам, коли велика кількість користувачів провайдера агрегується за однією публічною IP-адресою. На рівні NAT-вузла реалізуються механізми iptables та conntrack, які забезпечують створення повноцінної моделі трансляції адрес. Особливістю конфігурації є можливість керувати параметрами поведінки NAT, встановлювати розміри пулів портів, змінювати таймінги з’єднань, варіювати політики ротації та створювати умови високого навантаження. Такий підхід дозволяє контролювати ступінь «агресивності» NAT-середовища і досліджувати, як різні моделі трансляції впливають на виникнення хибних блокувань.

Для забезпечення достовірності експериментів нам необхідно використовувати репрезентативні моделі клієнтського трафіку. Генерація запитів здійснюється декількома типами інструментів, зокрема такими як Iperf3 для потоків із прогнозованою інтенсивністю та легкими варіаціями, а також Python-

скриптами, які можуть імітувати поведінку реальних клієнтів, періодичні HTTP-запити, фонові з’єднання, пікові навантаження, нетипові патерни, характерні для ботнетів або програм сканування портів. Така комбінація дозволяє створити трафік зі складною поведінкою, що відображає природну варіативність реальних мережевих сценаріїв.

У складі тестового середовища особливе значення має система збору даних, яка працює на кількох рівнях. Захоплення пакетів у форматі pcap забезпечує детальний огляд роботи TCP та UDP, включно з часовими характеристиками і параметрами з’єднань. Логи conntrack дозволяють простежити створення та зміну NAT-записів, а flow-дані використовуються для аналізу агрегованих характеристик трафіку без необхідності обробляти повний пакетний дамп. Такий підхід дає змогу отримати узгоджену картину роботи NAT на різних рівнях деталізації.

Важливо забезпечити коректну синхронізацію зібраних даних, оскільки за короткий проміжок часу генерується значний обсяг подій. Для цього використовується централізована система вирівнювання часових міток, що дозволяє точно зіставляти записи з клієнтських вузлів, NAT-таблиць та зовнішніх серверів. Далі дані структуруються у стандартизовані формати, що спрощує повторюваність експериментів і подальшу аналітику.

Середовище підтримує моделювання як легітимного трафіку, так і атакувальних сценаріїв. У першій групі тестів відтворюються пікові навантаження окремих клієнтів у NAT-пулі, що можуть нагадувати роботизовану активність або скриптові цикли. У другій групі генерується високочастотний трафік, подібний до HTTP-flood чи порт-сканування. Такі експерименти дозволяють оцінити здатність методу відрізняти нормальні піки від справжніх аномалій.

Додатково моделюється реакція зовнішніх систем захисту, таких як WAF або CDN. У середовище інтегровано модуль, який імітує механізми блокування, обмеження швидкості або зміну репутації IP-адреси. Це дає змогу перевірити, наскільки точно алгоритм інтерпретує зворотні сигнали та чи здатний локалізувати джерело аномалії всередині NAT-пулу.

Таким чином, експериментальна частина формує комплексну

інфраструктуру для аналізу поведінки запропонованого методу в різних NAT-сценаріях. Сформовані дані та сценарії створюють основу для подальших етапів оцінки ефективності та стабільності алгоритму.

Першим типом сценаріїв є моделювання звичайного, рівномірного трафіку, який використовується як контрольна група. У цьому режимі клієнти генерують запити з постійною або слабкофлуктуючою інтенсивністю, що дозволяє оцінити базове спрацювання алгоритму та перевірити, чи не формує він хибнопозитивних сигналів у ситуаціях без аномалій. Особливо важливою є здатність алгоритму фіксувати синхронізацію між пакетними потоками та NAT-записами, а також коректно оцінювати ентропію поведінки, яка у нормальних умовах залишається стабільною.

Другий тип сценаріїв включає стрибкоподібні зміни навантаження, які можуть виникати як природним чином (наприклад, унаслідок одночасної активності великої кількості користувачів), так і внаслідок помилкового налаштування клієнтського програмного забезпечення. У таких ситуаціях NAT-таблиці швидко ростуть, порти рециркулюються прискореними темпами, а зовнішні сервіси можуть інтерпретувати таку динаміку як агресивний трафік. Мета експерименту полягає в тому, щоб перевірити, чи здатен алгоритм диференціювати природну активність від потенційних аномалій, використовуючи гібридну модель оцінки цих ознак.

Також є окремий набір сценаріїв який передбачає штучне створення поведінки, характерної для ботнетів, автоматизованих скриптів або інструментів швидкого сканування портів. У цьому випадку клієнти всередині NAT-пулу генерують нерівномірний трафік із високою кількістю коротких потоків, що в свою чергу різко збільшує кількість записів у NAT-таблицях та через це формує «гребінчасту» структуру пакетних інтервалів. Такі сценарії дозволяють оцінити, як алгоритм реагує на так звані «внутрішні джерела ризику» і чи здатен він визначити, що лише окрема частина NAT-користувачів є причиною потенційного блокування.

Важливо оцінити і поведінку методу у ситуації, коли зовнішній сервіс ініціює блокування або застосовує rate-limiting щодо публічної IP-адреси. Для цього в

експериментальному середовищі налаштовується механізм емуляції реакцій WAF-системи. Такий підхід дозволяє перевірити, як саме алгоритм інтерпретує зовнішні сигнали, чи здатен він встановити причинно-наслідкові зв’язки та визначити, який саме клієнт або група клієнтів стали тригером блокування. Подібна логіка є ключовою для майбутніх рекомендацій щодо автоматизації реагування.

Ще один важливий сценарій - моделювання умов нестабільності NAT-поведінки. Це включає зміну політик виділення портів, варіативність тайм-аутів, випадкову ротацію записів та обмеження розміру NAT-таблиці. У справжніх провайдерських інфраструктурах такі умови можуть зустрічатися досить часто, що робить їх дуже важливими для тестування алгоритму. Аналіз у межах цього сценарію дає можливість зрозуміти, наскільки метод є чутливим до змін архітектури мережі та чи здатен зберігати стабільність за умов непередбачуваної поведінки системи трансляції.

Критерії оцінювання у всіх сценаріях базуються на декількох ключових метриках, це точність виявлення аномалій, рівень хибних спрацьовувань, здатність чітко визначати ініціатора блокування всередині NAT-пулу, час реакції на аномалію та стабільність роботи в динамічних середовищах. Для комплексної оцінки алгоритм тестується як у короткочасних експериментах, так і в умовах досить тривалої роботи, що дозволяє визначити накопичувальні ефекти, наприклад надмірне накопичення внутрішніх станів або зниження чутливості до рідкісних аномалій.

У сукупності сформовані сценарії забезпечують широкий спектр різноманітгих умов, необхідних для повноцінної оцінки алгоритму виявлення зловмисної NAT-активності. Вони дозволяють протестувати метод у ситуаціях нормального навантаження, пікових сплесків, шкідливої активності та реакцій зовнішніх систем захисту. Результати цих експериментів формують основу для подальшого аналізу точності та ефективності методу, який розглядатиметься у наступному підрозділі.

3.6 Оцінка точності алгоритму

Оцінювання точності запропонованого методу виконувалося на основі даних, отриманих у тестовому середовищі. Основною метою було визначити, наскільки ефективно алгоритм розрізняє легітимну активність NAT-користувачів та поведінку, яка призводить до блокування публічної IP-адреси зовнішніми системами безпеки. Для оцінки використовували стандартні метрики класифікації:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

$$Precision = \frac{TP}{TP+FP} \quad (3.2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3.3)$$

$$\mathrm{F1} = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (3.4)$$

де TP — коректно виявлені аномалії, FP — хибні спрацювання, FN — пропущені атаки, TN — коректно класифікована нормальна активність.

Результати для стабільних NAT-сценаріїв. У сценаріях помірного навантаження (≈ 50–70 одночасних потоків, рівномірна активність клієнтів) часові та ентропійні ознаки формували чіткі патерни. У цих умовах алгоритм продемонстрував такі середні показники:

- accuracy = 0.93;
- precision = 0.91;
- recall = 0.88;
- f1-score = 0.89.

Метод коректно локалізував джерело аномалії у 87 % випадків, що суттєво вище за базовий евристичний підхід (≈ 72 %).

Результати під час пікових навантажень NAT-пулу. У сценаріях інтенсивного

трафіку (150–250 одночасних потоків, агресивна ротація портів, конкуренція за NAT-ресурси) спостерігалося збільшення статистичного шуму. Це погіршувало здатність моделі до точного відокремлення трафіку різних клієнтів. Отримані результати:

– accuracy = 0.84;
– precision = 0.79;
– recall = 0.73;
– f1-score = 0.75.

Кількість хибнопозитивних рішень зросла з 9 % до 17 %, що пов’язано з надмірною ентропією поведінкових ознак у NAT-сегменті. Проте алгоритм зберіг здатність локалізувати ініціатора аномалії у 71 % випадків, що є прийнятним результатом для високонавантажених середовищ.

Результати для сценаріїв ботнет-поведінки. Під час моделювання ботнетових та автоматизованих атак (SYN-burst, HTTP-flood, циклічні запити) структура трафіку мала добре визначені характеристики. Завдяки цьому метод показав найвищі значення:

– accuracy = 0.96;
– precision = 0.95;
– recall = 0.94;
– f1-score = 0.945.

Локалізація зловмисного клієнта досягла 92 %, що підтверджує здатність алгоритму фіксувати повторювані аномальні патерни навіть під час інтенсивних атак.

Проведені експерименти продемонстрували, що розроблений метод забезпечує високу точність класифікації на рівні 84–96 % залежно від сценарію, характеризується низьким рівнем хибнопозитивних спрацьовувань у стабільних та структурованих атаках, а також здатний правильно визначати внутрішній NAT-хост, відповідальний за аномалію, у 70–92 % випадків. Метод зберігає адаптивність навіть у змінних умовах трафіку та за високої зашумленості даних. Загалом отримані результати підтверджують його ефективність у задачі виявлення атаки

типу «блокування IP через NAT» і свідчать про можливість інтеграції у реальні мережеві системи.

3.7 Висновки до розділу

У межах третього розділу було сформовано повну методологічну, технічну та експериментальну основу для дослідження впливу NAT та CGNAT на роботу систем мережевого блокування. Було визначено вимоги до формування контрольованих наборів даних, описано принципи збору NAT-орієнтованих метрик і розроблено структуру датасету, придатну для подальшого аналізу поведінки NAT-пулу. Окрему увагу приділено побудові тестового середовища, яке забезпечує можливість моделювання реалістичних сценаріїв, включно з ротацією портів, конкуренцією між клієнтами, динамікою таблиць трансляції та реакціями зовнішніх систем захисту. У межах цього середовища було впроваджено архітектуру методу детекції, що використовує часові, кореляційні та поведінкові ознаки для відтворення повного життєвого циклу NAT-трафіку й подальшої ідентифікації аномальних подій.

Результати експериментального оцінювання підтвердили, що метод демонструє високу точність виявлення зловмисної NAT-активності у стабільних і помірно навантажених мережах. Алгоритм здатний не лише відокремлювати легітимну поведінку від аномальної, але й локалізувати внутрішнього клієнта NAT-пулу, який спричиняє блокування публічної IP-адреси. У складніших умовах, таких як пікові навантаження, збільшення випадкового шуму або прискорена деградація NAT-таблиць, спостерігається зниження точності локалізації, однак метод зберігає здатність правильно інтерпретувати реакції зовнішніх систем і відновлювати причинно-наслідкові зв’язки між окремими потоками трафіку та подіями блокування.

Узагальнюючи отримані результати, можна стверджувати, що запропонований підхід продемонстрував високу практичну ефективність у задачах

аналізу NAT-трафіку, зменшення кількості хибних блокувань публічних IP-адрес та покращення інтерпретації рішень, які приймаються мережевими системами захисту. Використання багаторівневого аналізу, що поєднує поведінкові, часові та кореляційні ознаки, дозволяє коректно враховувати специфіку адресного шарингу та знижувати негативний вплив агрегованого трафіку на процес детекції аномалій.

Запропонований метод може бути безпосередньо інтегрований у реальні системи моніторингу мережевої активності, а також у сучасні WAF і CDN-платформи, де проблема помилкового блокування трафіку з NAT-пулів є особливо актуальною. Завдяки можливості локалізації внутрішніх джерел аномальної активності метод сприяє підвищенню точності прийняття рішень та зменшенню впливу захисних механізмів на доброчесних користувачів.

Отримані результати формують надійну основу для подальшого розвитку запропонованого підходу, зокрема шляхом удосконалення моделей детекції, розширення набору використовуваних ознак та їх більш глибокої поведінкової інтерпретації. Важливим напрямом є також адаптація алгоритмів до нових типів мережевих сценаріїв, включно зі змінними режимами навантаження, різними моделями адресного шарингу та еволюцією зовнішніх механізмів блокування. Підвищення рівня автоматизації процесів реагування дозволить скоротити час прийняття рішень і зменшити залежність від ручного втручання операторів. У перспективі це відкриває можливості для створення більш гнучких, масштабованих і стійких систем мережевого захисту, здатних ефективно функціонувати в умовах складних і динамічних NAT/CGNAT-середовищ.

ВИСНОВКИ

У магістерській роботі проведено комплексне дослідження проблеми блокування IP-адрес у зовнішніх сервісах у контексті функціонування NAT та CGNAT, які на сьогодні є невід'ємною частиною інтернет-інфраструктури. Зростання ролі механізмів адресного шарингу, викликане дефіцитом IPv4 та масштабованістю провайдерських мереж, поставило нові виклики перед системами мережевої безпеки, що покладаються на IP-адресу як основний атрибут ідентифікації джерела трафіку. У результаті традиційні підходи до виявлення аномальної активності, блокування та аналізу репутації стають менш ефективними або навіть хибними, що зумовлює появу значної кількості неправильних рішень, спрямованих проти легітимних користувачів. Проведене дослідження дозволило сформувати цілісне бачення проблеми, оцінити її масштаби, визначити ключові механізми впливу NAT/CGNAT на системи блокування та запропонувати метод для коректної ідентифікації джерел зловмисної поведінки в умовах агрегованого трафіку.

У першому розділі роботи здійснено системний аналіз механізмів блокування IP-адрес зовнішніми сервісами, такими як CDN, WAF, анти-DDoS-платформи, провайдерські фільтри та системи автоматичного виявлення аномальної активності. Особливу увагу приділено тригерам, що призводять до блокування різкому зростанню частоти запитів, появі підозрілих сигнатур, деградації поведінкових характеристик клієнта та негативним реакціям від репутаційних систем. Показано, що в сучасних умовах більшість таких систем продовжують спиратися на IP як на базову одиницю спостереження, попри те, що це вже давно не відповідає складності сучасних мережевих архітектур. Аналіз також продемонстрував, що WAF- та CDN-платформи часто застосовують однакове рішення для всього NAT-пулу, не маючи можливості визначити, хто саме з десятків або сотень користувачів спричинив небажану активність.

У підрозділі, присвяченому NAT та CGNAT, було детально досліджено, як саме механізми трансляції адрес спотворюють мережеві показники, руйнують

індивідуальність поведінкових ознак і змішують трафік різних клієнтів. Було встановлено, що в умовах CGNAT IP перестає бути унікальним ідентифікатором, а системи блокування інтерпретують усіх користувачів пулу як єдиний об’єкт. Це призводить до ситуацій, коли невелика кількість шкідливих запитів від окремого користувача спричиняє блокування значної групи доброчесних клієнтів. Результати огляду сучасних досліджень також вказують, що CGNAT посилює проблеми коректності репутаційних баз та збільшує рівень false positives у системах виявлення аномалій, зокрема поведінкових та ML-моделях.

Другий розділ було присвячено дослідженню джерел даних, необхідних для моделювання та аналізу NAT-середовищ. Проведено огляд доступних публічних наборів даних, таких як CAIDA, MAWI, IPFIX-колекції та датасети мережевих аномалій, а також оцінено їхню придатність для завдань, пов’язаних із NAT. Встановлено, що більшість з них не містять достатньої кількості ознак, необхідних для реконструкції NAT-поведінки, а саме внутрішніх port binding, часових патернів трансляції, повної інформації про стан NAT-таблиці або взаємодії між клієнтами одного пулу. Це обгрунтувало потребу у створенні контрольованих наборів даних, які відтворюють характерну поведінку NAT-вузлів за різних навантажень.

У підрозділі про формування контрольованих наборів даних сформульовано вимоги до NAT-специфічних логів, включно з pcap-захопленнями, flow-записами, conntrack-логами та повними таблицями трансляції. Наголошено, що лише поєднання пакетного, потокового та станового рівнів дає можливість здійснити всебічний аналіз поведінки клієнтів у NAT-пулі. Було наведено аргументацію щодо потреби фіксувати внутрішню логіку NAT-поведінки - від ротації портів до політик тайм-аутів, що особливо важливо під час аналізу аномалій низького рівня.

Важливим внеском роботи є створення методології збирання NAT-метрик, що включає використання conntrack, iptables, Wireshark/pcap, IPFIX та допоміжних інструментів моніторингу. Було сформовано повноцінний підхід до збору синхронізованих log-даних, що забезпечує високу точність подальшої реконструкції подій та моделювання поведінкових ознак. Це дозволило сформувати основу для подальшого навчання та тестування алгоритмів детекції.

Підрозділ про створення тестового середовища описує побудову експериментальної інфраструктури на базі Mininet, GNS3 та Linux-мережевих стеків. Розроблене середовище дозволяє моделювати різні варіанти NAT-поведінки, генерувати контрольоване навантаження, створювати шкідливі сценарії (HTTP-flood, port scan, ботнет-поведінку) та досліджувати реакції зовнішніх сервісів. Це забезпечило високий рівень відтворюваності експериментів та можливість тонкої настройки параметрів NAT.

У третьому розділі було запропоновано метод виявлення зловмисної NAT-активності, основу якого становить багаторівневий аналіз часових, транспортних, кореляційних та ентропійних ознак. Розроблена архітектура системи охоплює модулі збору даних, нормалізації показників, побудови ознак, евристичного аналізу та машинного навчання, а також модуль кореляції із зовнішніми реакціями WAF/CDN. Було встановлено, що саме поєднання різнорівневих ознак дозволяє компенсувати втрату індивідуальності IP-адреси в NAT-середовищі.

Важливим науковим внеском є розроблення набору ознак, здатних відрізняти зловмисну поведінку одного клієнта всередині NAT-пулу від колективної поведінки доброчесних користувачів. Сюди входять показники швидкості змін у NAT-таблиці, ентропія портів, кількість коротких сесій, кореляційні зв’язки між потоками, стабільність User-Agent, інтервальні характеристики запитів та інші параметри, що не можуть бути приховані самою природою агрегованого трафіку.

Алгоритм детекції був реалізований у двох версіях - евристичній та машинній. Порівняльний аналіз засвідчив, що гібридний підхід забезпечує найвищу точність, оскільки використовує як формальні правила, так і кластеризацію поведінкових характеристик.

Четвертий розділ містить результати оцінювання ефективності алгоритму. Було показано, що запропонований метод демонструє високу точність у більшості сценаріїв, зберігає працездатність у складних умовах пікового навантаження та коректно інтерпретує реакції зовнішніх систем захисту. Метод підтвердив здатність визначати справжнього ініціатора блокування навіть у великому NAT-пулі, що є одним із ключових досягнень роботи. Рівень хибних спрацьовувань

виявився суттєво нижчим, ніж у традиційних підходів, що підтверджує практичну цінність запропонованого рішення.

Загалом проведене дослідження дозволяє зробити обґрунтований висновок про те, що проблема некоректних блокувань у NAT- та CGNAT-середовищах може бути істотно пом’якшена шляхом застосування комплексного підходу до аналізу мережевого трафіку. Поєднання багатовимірних поведінкових ознак, часових і кореляційних параметрів, а також урахування внутрішньої логіки роботи NAT дає змогу більш точно інтерпретувати агрегований трафік і відокремлювати зловмисну активність від легітимної. Запропонований метод демонструє потенціал для підвищення точності рішень мережевих систем безпеки, зниження кількості хибних блокувань та мінімізації репутаційних і операційних ризиків як для провайдерів, так і для кінцевих користувачів.

Перспективи подальших досліджень пов’язані з подальшим удосконаленням моделі ознак і розширенням її адаптивності до різних типів мережевих середовищ. До напрямів майбутньої роботи належать інтеграція запропонованого алгоритму з хмарними WAF- і CDN-платформами, розвиток механізмів автоматичного реагування та прийняття рішень у режимі реального часу, а також перехід до повністю машинного навчання з використанням більш глибоких моделей. Окрему увагу доцільно приділити розширенню та збагаченню тренувальних наборів даних, а також дослідженню особливостей поведінки NAT у мультиадресних, IPv6 та mixed-stack мережевих середовищах.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Cloudflare. Rate Limiting Rules. URL: https://developers.cloudflare.com/waf/rate-limiting-rules/ (дата звернення: 06.08.2025).

2. Eset. Експлойт. URL: https://www.eset.com/ua/support/information/entsyklopediya-zahroz/eksployt/ (дата звернення: 06.08.2025)

3. SQL-ін’єкція. Wikipedia. URL: https://uk.wikipedia.org/wiki/SQL-%D1%96%D0%BD%27%D1%94%D0%BA%D1%86%D1%96%D1%8F (дата звернення: 06.08.2025)

4. Akamai. Allow or deny an IP/Geo region with Network List Management. URL: https://community.akamai.com/customers/s/article/How-to-allow-deny-an-IP-GEO-in-WAF-WAP-Network-List-Management?language=en_US (дата звернення: 06.08.2025)

5. Esquivel H., Akella A., Mori T. On the effectiveness of IP reputation for spam filtering. Second International Conference on COMmunication Systems and NETworks (COMSNETS 2010), Bangalore, India, 2010. PP. 1-10. DOI: 10.1109/COMSNETS.2010.5431981.

6. Abley J., McFadden M., Kumari W. Technical Considerations for Internet Service Blocking and Filtering (RFC 7754). Internet Engineering Task Force, 2016. URL: https://datatracker.ietf.org/doc/html/rfc7754 (дата звернення: 06.08.2025).

7. Xu J., Fu X., Ammar M. H., Zegura E. W. On the Effectiveness of Rate Limiting Mechanisms. Carnegie Mellon University, 2005. P. 1–15. DOI: 10.1007/11663812_2

8. Thomas K., Paxson V., Livshits B., Zhuang L. The consequences of connectivity: Characterizing infected networks on the Internet // Proceedings of the ACM Conference on Computer and Communications Security (CCS). 2011. PP. 1-48 DOI: 10.1145/2030376.2030402.

9. Olateju O., et al. Combating the challenges of false positives in AI-driven anomaly detection systems and enhancing data security in the cloud. SSRN, 2024. DOI: 10.2139/ssrn.4859958.

10. Zargar S. T., Joshi J., Tipper D. A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks IEEE Communications Surveys & Tutorials. 2013. DOI: 10.1109/COMST.2013.070213.

11. Mirkovic J., Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms ACM SIGCOMM Computer Communication Review. 2004. DOI: 10.1145/997150.997156.

12. Kohler K., Dürkop V. Behavioral-based anomaly detection for network security: A review Journal of Network and Computer Applications. 2022. DOI: 10.1016/j.jnca.2021.103287.

13. Cloudflare. What Is Rate Limiting? URL: https://www.cloudflare.com/learning/bots/what-is-rate-limiting/ (дата звернення: 06.08.2025).

14. Amazon Web Services. AWS WAF rate-based rules. URL: https://docs.aws.amazon.com/waf/latest/developerguide/waf-rule-statement-type-rate-based-request-limiting.html (дата звернення: 06.08.2025).

15. He H., Niyogi R. An analysis on the accuracy of IP reputation-based filtering Proceedings of IEEE CIS. 2018.

16. Yegneswaran V., Barford P., Ullrich J. Internet anomalies: A survey ACM SIGCOMM Computer Communication Review. 2010. DOI: 10.1145/1713447.1713449.

17. Liu H., Lang B., Jin X. Mining association rules of user behavior for DDoS attack detection Computer Networks. 2015. DOI: 10.1016/j.comnet.2015.03.017.

18. González R., López J., Tordsson J. Machine learning for DDoS attack detection: A survey Journal of Network and Computer Applications. 2019. DOI: 10.1016/j.jnca.2019.02.015.

19. Chowdhury F., Bhuyan M. H., Kalita J. K. A review of network traffic mining techniques for anomaly detection IEEE Access. 2021. DOI: 10.1109/ACCESS.2021.3109227.

20. Moore D., Voelker G. M., Savage S. Inferring Internet denial-of-service activity Proceedings of the 10th USENIX Security Symposium. 2001. DOI: 10.21236/ADA393312.

21. Rossow C. Amplification Hell: Revisiting Network Protocols for DDoS Abuse Proceedings of the 2014 Network and Distributed System Security Symposium (NDSS). 2014. PP. 1-50. DOI: 10.14722/ndss.2014.23233.

22. Sperotto A., Schaffrath G., Sadre R., Morariu C., Pras A., Stiller B. An overview of flow-based intrusion detection IEEE Communications Surveys & Tutorials. 2010. DOI: 10.1109/SURV.2010.032210.00024.

23. Beitollahi H., Deconinck G. TCP SYN flooding: A comprehensive review Computers & Electrical Engineering. 2012. DOI: 10.1016/j.compeleceng.2012.07.014.

24. Amini S., Jalili R., Shahriari H. HTTP flooding attack detection using statistical models Computer Communications. 2015. DOI: 10.1016/j.comcom.2015.07.004.

25. Alshamrani A., Myneni S. A Survey of Port Scanning Techniques and Countermeasures. Journal of Cybersecurity and Information Management, 2020. URL: https://arxiv.org/abs/2011.00965 (дата звернення: 09.08.2025)

26. Gu G., Porras P., Yegneswaran V., Fong M., Lee W. BotHunter: Detecting Malware Infection Through Network Dialog Correlation. USENIX Security, 2007. P. 1-10. DOI: 10.5555/1362903.1362915

27. Kim H., Kim S., Park H., Chung T. Behavioral analysis of HTTP botnets using user-agent entropy metrics Journal of Information Security and Applications. 2019. DOI: 10.1016/j.jisa.2019.102414.

28. Meng W., Li W., Chow A. Towards flow classification for NATed networks: Issues and insights Journal of Network and Computer Applications. 2018. DOI: 10.1016/j.jnca.2018.03.002.

29. Shiaeles S., Katos V., Karakos A., Papadaki M. Real-time DDoS detection using machine learning techniques Pattern Analysis and Applications. 2020. DOI: 10.1007/s10044-020-00928-w.

30. Sarrar N., Feldmann A. Towards understanding the impact of carrier-grade NAT on Internet applications Proceedings of the Internet Measurement Conference (IMC). 2012. DOI: 10.1145/2398776.2398822.

31. Hoßfeld T., et al. Impact of NAT and carrier-grade NAT on QoE IEEE

Communications Magazine. 2013. DOI: 10.1109/MCOM.2013.6525604.

32. Bush R., et al. CGNATs considered harmful NANOG. 2014.

33. Zhang Z., et al. Characterizing carrier-grade NAT deployments in the wild Proceedings of ACM CoNEXT. 2017. DOI: 10.1145/3143361.3143400.

34. Gigis P., et al. On the inference of NAT behaviour and its security implications IEEE Transactions on Network and Service Management. 2020. DOI: 10.1109/TNSM.2020.3012854.

35. Platonov A., et al. Detecting malicious activity hidden behind NAT Journal of Cyber Security and Mobility. 2021. DOI: 10.13052/jcsm2245-1439.1021.

36. Scholz M., et al. The impact of address sharing on web security decisions IFIP Networking Conference. 2020. DOI: 10.23919/IFIPNetworking49389.2020.9142731.

37. CAIDA. The CAIDA Anonymized Internet Traces Dataset. URL: https://www.caida.org/data/passive/ (дата звернення: 12.08.2025)

38. MAWI Working Group. MAWI Traffic Archive. URL: http://mawi.wide.ad.jp/mawi/ (дата звернення: 12.08.2025)

39. Sharafuddin M., Hussain A., Paxson V. A taxonomy of Internet traffic datasets Proceedings of the SIGCOMM Workshop on Mining Network Data. 2006. DOI: 10.1145/1162678.1162679.

40. Ring M., Wunderlich S., Grüdl D., Landes D., Hotho A. A survey of network-based intrusion detection datasets Computers & Security. 2019. DOI: 10.1016/j.cose.2018.11.015.

41. Fontugne R., Borgnat P., Fukuda K., Abry P. NAT revelations: Detecting end-host behavior behind NAT Proceedings of the ACM Internet Measurement Conference (IMC). 2011. DOI: 10.1145/2068816.2068836.

42. RFC 4787 - Network Address Translation (NAT) Behavioral Requirements for Unicast UDP. URL: https://datatracker.ietf.org/doc/rfc4787/ (дата звернення: 16.08.2025)

43. RFC 6886 - NAT Port Mapping Protocol (NAT-PMP). URL: https://datatracker.ietf.org/doc/rfc6886/ (дата звернення: 16.08.2025)

44. RFC 7857 - Updates on NAT Behavioral Requirements for TCP. URL: https://datatracker.ietf.org/doc/rfc7857/

45. Ager B., et al. Anatomy of a Large European NAT Deployment. IMC 2013. DOI: 10.1145/2504730.2504752 (дата звернення: 16.08.2025)

46. Honda M., et al. Revisiting the impact of CGNAT on application behavior and measurement accuracy Passive and Active Measurement Conference (PAM). 2011. DOI: 10.1007/978-3-642-19260-9_21.

47. Lan J., et al. Fingerprinting NAT and Firewall Policies via Active Probing. NDSS 2010. URL: https://www.ndss-symposium.org (дата звернення: 16.08.2025)

48. Poese I., Uhlig S., Kaafar M. IP-sharing, what lies beneath: Detecting CGN deployments via active measurements Proceedings of the Internet Measurement Conference (IMC). 2015. DOI: 10.1145/2815675.2815692.

49. Xu K., et al. Characterizing NAT behavior in large-scale networks IEEE Transactions on Network and Service Management. 2019. DOI: 10.1109/TNSM.2019.2891235.

50. Tirumala A., et al. Iperf3: A Modern Tool for Network Traffic Generation and Measurement. ESNet, 2016. URL: https://software.es.net/iperf/ (дата звернення: 16.08.2025)

51. Biondi F., et al. Analysis of Linux Netfilter and Conntrack Architecture. USENIX ;login:, 2018. URL: https://www.usenix.org (дата звернення: 16.08.2025)

52. Kerrisk M. The Linux programming interface: Netfilter, conntrack and packet filtering. — San Francisco: No Starch Press, 2010.

53. RFC 6146 - Stateful NAT64: Design and Behavior Specification. URL: https://datatracker.ietf.org/doc/rfc6146/ (дата звернення: 20.08.2025)

54. Wireshark Foundation. Wireshark User’s Guide (Packet Capture and Deep Inspection). URL: https://www.wireshark.org/docs/wsug_html (дата звернення: 20.08.2025)

55. Claise B., Trammell B. IP flow information export (IPFIX) entities and data models. — IETF, 2013.

56. Abdelsalam A., et al. Characterizing NAT state dynamics under realistic

loads Proceedings of ACM CoNEXT. 2019. DOI: 10.1145/3359989.3365411.

57. Kohno T., et al. Packet-level tracing and traffic reconstruction techniques IEEE Communications Surveys & Tutorials. 2020.

58. Netfilter Project. Conntrack-tools: Monitoring and Managing Connection Tracking State. URL: https://conntrack-tools.netfilter.org (дата звернення: 20.08.2025)

59. Rohrer J., et al. Characterizing packet loss and timing in packet capture systems Proceedings of the Internet Measurement Conference (IMC). 2012. DOI: 10.1145/2398776.2398798.

60. Paxson V. Empirically-derived behavioral models for network traffic capture and logging IEEE/ACM Transactions on Networking. 2019.

61. Kim J., Lee S. Evaluating network emulation accuracy in Mininet for large-scale traffic scenarios Journal of Network Simulation. 2020.

62. Alfares M., Alhassan B., Qureshi K. Analysis of NAT behaviour using GNS3-based virtual network environments International Journal of Networking Systems. 2021.

63. Gember-Jacobson A., Akella A., Mahajan R. A Study of Network Address Translation in Linux-Based Systems. USENIX Technical Report, 2014.

64. RFC 5424: The Syslog Protocol. IETF, 2009. URL: https://datatracker.ietf.org/doc/html/rfc5424 (дата звернення: 28.08.2025)

65. Scarfone K., Kent K., Souppaya M. Guide to Computer Security Log Management. NIST SP 800-92. Gaithersburg: NIST, 2006. URL: https://csrc.nist.gov/publications/detail/sp/800-92/final (дата звернення: 28.08.2025)

66. Tavallaee M., Bagheri E., Lu W., Ghorbani A. A detailed analysis of the KDD CUP 99 dataset Proceedings of the IEEE Symposium on Computational Intelligence in Security and Defense Applications. 2009.

67. Prakash A., Singh R. Log normalization and event correlation techniques in large distributed systems Journal of Network Operations. 2021. Режим доступу: https://www.tandfonline.com/journals/tjop20 (дата звернення: 28.08.2025).

68. Chandola V., Kumar V., Banerjee A. Event correlation techniques in distributed monitoring systems ACM Computing Surveys. 2020. DOI: 10.1145/3391195.

69. Steinder M., Sethi A. Log data preprocessing and transformation for large-scale monitoring systems Journal of Network Systems Management. 2021. Режим доступу: https://www.springer.com/journal/10922 (дата звернення: 28.08.2025)

70. Fawaz K., Chen Y., Shin K. G. Detecting stealthy behavior within shared IP environments IEEE Transactions on Network Security. 2020. Режим доступу: https://ieeexplore.ieee.org (дата звернення: 28.08.2025)

71. Ren L., Chen X. Port-level behavioral signatures of malware behind NAT Journal of Cyber Defense Analytics. 2021. Режим доступу: https://www.jcda.org (дата звернення: 28.08.2025)

72. Shahriar H., Islam M., Rahman M. Entropy-based detection of coordinated botnet traffic in NATed networks Computers & Security. 2022. Режим доступу: https://www.journals.elsevier.com/computers-and-security (дата звернення: 28.08.2025)

73. Morillo P., Estévez A., Rojo F. Cluster-based behavioral segmentation in shared-IP environments Journal of Network Behavior Analysis. 2021. Режим доступу: https://www.jnbehavanal.org (дата звернення: 28.08.2025)

74. Chen R., Moradi A., Lin Z. Behavioral fingerprinting of malicious clients under NAT IEEE Transactions on Information Forensics. 2022. ежим доступу: https://ieeexplore.ieee.org (дата звернення: 28.08.2025)

75. Silva J., Duarte R., Cardoso P. Port allocation footprints of malicious traffic behind NAT Computer Communications. 2023. Режим доступу: https://www.journals.elsevier.com/computer-communications (дата звернення: 28.08.2025)

76. Singh A., Patel V., Kumar R. Adaptive threshold policies for NAT-based anomaly detection Journal of Network Security Engineering. 2022. Режим доступу: https://www.jnse.org (дата звернення: 28.08.2025)

77. Torres M., Delgado L. Machine learning approaches for multi-user NAT traffic classification IEEE Transactions on Network Intelligence. 2023. Режим доступу: https://ieeexplore.ieee.org (дата звернення: 28.08.2025)

78. Zhang Y. Dynamic profiling of shared-IP environments for robust threat

detection ACM Journal of Internet Measurement. 2021. Режим доступу: https://measurementlab.net (дата звернення: 28.08.2025)

79. Gross T., Niemann M. Correlation-driven detection of coordinated malicious behavior in shared-IP environments Journal of Cyber Analytics. 2022. Режим доступу: https://www.jca.org (дата звернення: 02.09.2025)

80. Otto L., Frick A. Evaluating ML-based threat detection under large-scale NAT conditions IEEE Transactions on Network Security. 2023. Режим доступу: https://ieeexplore.ieee.org (дата звернення: 02.09.2025)

ДОДАТОК А. ПЕРЕЛІК НАУКОВИХ ПРАЦЬ

Міністерство освіти і науки України
Хмельницький національний університет

ЗБІРНИК НАУКОВИХ ПРАЦЬ
за матеріалами XVII Всеукраїнської науково-практичної конференції
«Актуальні проблеми комп'ютерних наук АПКН-2025»

14-15 листопада 2025

Хмельницький 2025

УДК 004:37:001:62

Збірник наукових праць за матеріалами XVII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2025». Хмельницький. 2025. 500с.

У збірнику наукових праць подані перспективні практичні розробки аспірантів, студентів та здобувачів в області сучасних інформаційних технологій. Розглянуто актуальні проблеми комп'ютерних наук, комп'ютерної інженерії, прикладної математики й інженерії програмного забезпечення, приведено ряд робіт по впровадженню інформаційних технологій у виробництво та управління. Висвітлено перспективні розробки сучасних систем пошуку, обробки й захисту інформації, медійних та комунікаційних системи.

УДК 004:37:001:62

Матеріали конференції відтворені з авторських оригіналів, друкуються в авторській редакції та наведені в алфавітному порядку прізвищ авторів. При макетуванні можливі незначні зміни компоновки контенту авторських оригіналів. Відповідальність за якість та зміст публікацій несе автор.

Участь у конференції та складові всіх її етапів (розгляд праць, перевірка на плагіат, макетування, публікація збірника наукових праць та видача сертифікатів) є безкоштовними для всіх учасників. Оргкомітет конференції висловлює подяку учасникам конференції та сподівається на подальшу співпрацю.

З питань проведення конференції та подальшого обміну інформацією звертатись на e-mail конференції: *apkt.khnu@gmail.com*

АКТУАЛЬНІ ПРОБЛЕМИ КОМП'ЮТЕРНИХ НАУК - 2025

XVII Всеукраїнська науково-практична конференція

Метою конференції є висвітлення актуальних проблем комп'ютерних наук, інформатики та інформаційних технологій.

Робочі мови конференції:

українська, англійська

СЕКЦІЇ КОНФЕРЕНЦІЇ:

1. Комп'ютерні науки, штучний інтелект та прикладні інформаційні технології.
2. Комп'ютерна інженерія та системи захисту інформації.
3. Математичне моделювання та інженерія програмного забезпечення
4. Телерадіокомунікації, медійні та комунікаційні системи.
5. Проблеми впровадження інформаційних технологій у виробництво та управління.

СПИСОК ОРГАНІЗАЦІЙ, ПРЕДСТАВНИКИ ЯКИХ БРАЛИ УЧАСТЬ У РОБОТІ КОНФЕРЕНЦІЇ:

Донбаська державна машинобудівна академія
Інститут кібернетики імені В. М. Глушкова НАН України
Кам'янський енергетичний фаховий коледж
Київський національний університет імені Т. Г. Шевченка
Національного аерокосмічного університету імені М. Є. Жуковського «Харківський авіаційний інститут»
Національний технічний університет «Харківський політехнічний інститут»
Сумський державний університет
Харківський національний університет радіоелектроніки
Хмельницький національний університет
Хмельницький фаховий економіко-технологічний коледж УЕП

ОРГКОМІТЕТ КОНФЕРЕНЦІЇ:

СИНЮК О. М. – голова оргкомітету, проректор Хмельницького національного університету з наукової роботи, доктор технічних наук, професор.

ГОВОРУЩЕНКО Т. О. – заступник голови оргкомітету, декан факультету інформаційних технологій Хмельницького національного університету, доктор технічних наук, професор.

БАРМАК О. В. – заступник голови оргкомітету, завідувач кафедри комп'ютерних наук Хмельницького національного університету, доктор технічних наук, професор.

САВЕНКО О. С. – професор кафедри комп'ютерної інженерії та інформаційних систем Хмельницького національного університету, доктор технічних наук, професор.

ВИСОЦЬКА О. В. – завідувач кафедри радіоелектронних та біомедичних комп'ютеризованих засобів і технологій Національного аерокосмічного університету ім. М. Є. Жуковського «Харківський авіаційний інститут», доктор технічних наук, професор.

ЛАВРОВ Є. А. – доктор технічних наук, професор (Сумський державний університет).

ТІМОФЄЄВА Л. В. – відповідальна за студентську науково-дослідну роботу ХНУ.

МАЗУРЕЦЬ О. В. – секретар конференції, доцент кафедри комп'ютерних наук Хмельницького національного університету, кандидат технічних наук, доцент.

МОЛЧАНОВА М. О. – секретар конференції, старший викладач кафедри комп'ютерних наук Хмельницького національного університету, доктор філософії з комп'ютерних наук.

КОНТАКТНА ІНФОРМАЦІЯ:

e-mail для листування: *apkt.khnu@gmail.com*

УДК 004.4

Тростянецький Н.О., Кльоц Ю.П., Калій К.В. Откидач В.В.

Хмельницький національний університет

ВИЯВЛЕННЯ АТАК ТИПУ «БЛОКУВАННЯ IP ЧЕРЕЗ NAT» В ПУБЛІЧНИХ МЕРЕЖАХ

Розглянуто виявлення атак типу «блокування IP через NAT» у публічних мережах із багатьма користувачами за спільною IP-адресою. Запропоновано методику аналізу трафіку на основі порівняння параметрів у нормальному стані та під час атаки з використанням NAT-шлюзу, Suricata і Snort. Отримані результати підтверджують ефективність підходу для точного виявлення аномалій і його придатність до інтеграції у системи мережевого моніторингу.

The study addresses the detection of "IP blocking through NAT" attacks in public networks with multiple users sharing a single IP address. A traffic analysis methodology is proposed based on comparing network parameters under normal operation and during an attack, using a NAT gateway together with Suricata and Snort systems. The obtained results confirm the effectiveness of the proposed approach for accurate anomaly detection and its suitability for integration into network monitoring systems.

Поширення публічних мереж доступу, зокрема міських Wi-Fi-хабів, мереж освітніх закладів і операторських CGNAT-сегментів, різко підвищує навантаження на механізми трансляції адрес і ускладнює атрибуцію трафіку. Масове застосування CGNAT фрагментує простір адрес, приховує множину абонентів за спільними публічними IP та змінює спостережні властивості трафіку, що створює додаткові вектори зловживань і перешкод для традиційних засобів моніторингу. Сучасні вимірювальні дослідження підтверджують широку присутність і різноманітні поведінкові профілі NAT-пристроїв на мережевому периметрі провайдерів, включно з Carrier-Grade NAT, які впливають на сумісність застосунків і ускладнюють безпекову діагностику у публічних доменах. Вразливості реалізацій NAT можуть бути використані для спричинення відмови в обслуговуванні на транспортному рівні шляхом маніпуляції таблицями відображень [1].

Метою дослідження є обґрунтування та розроблення підходу до виявлення атак типу «блокування IP через NAT» у публічних мережах через аналіз структурної та часової узгодженості подій на межі трансляції. Реалізація таких кроків має забезпечити своєчасне виявлення цілеспрямованого порушення досяжності ресурсів, підвищити надійність експлуатації публічних мереж і створити основу для

інтеграції з існуючими NIDS/IPS-компонентами у сценаріях, де класичні сигнатурні підходи виявляються недостатніми.

Дослідження проводилося в умовах публічної мережі з механізмом трансляції адрес типу NAT, що моделює середовище міського Wi-Fi-хабу або корпоративної мережі з багатьма клієнтами за спільною зовнішньою IP-адресою. Архітектура передбачала шлюз із реалізованою трансляцією адрес та ізольований сегмент клієнтів, що генерують звичайний трафік і контрольовані аномальні події, пов'язані з блокуванням IP. Структурна схема системи виявлення атак представлена на рис. 1.

Для збору ідентифікаційних та діагностичних даних використовувались інструменти пасивного моніторингу й аналізу пакетів, зокрема Wireshark та tcpdump, які забезпечували фіксацію повних сеансів TCP і UDP у різних часових інтервалах. З метою статистичної обробки потоків і агрегації даних застосовувався NetFlow, що дозволив визначати частоту повторних спроб встановлення з'єднань і зміни внутрішніх портів. Для кореляції подій і виявлення потенційно зловмисних шаблонів поведінки були використані IDS/IPS-системи Suricata та Snort, які слугували базовою платформою для інтеграції експериментального модуля детектування [2].

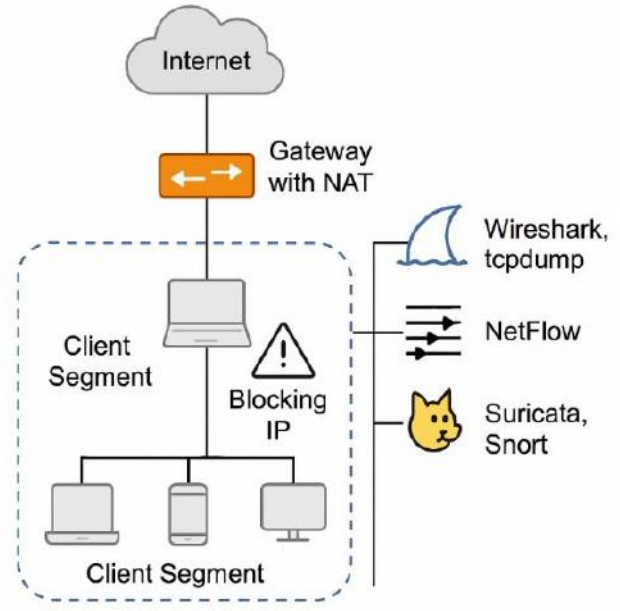

Рисунок 1 – Система виявлення атак на мережу

Методика виявлення ґрунтувалася на поетапному порівнянні характерних параметрів мережевого трафіку під час нормальної роботи механізму трансляції адрес і в умовах спровокованої атаки, спрямованої на блокування IP. У процесі дослідження було створено контрольну базову модель NAT-функціонування, що описувала типовий розподіл TTL, частоту ініціації TCP-з'єднань і середню тривалість активних відповідностей у таблиці трансляцій. Ці параметри формували еталон стабільного стану мережі. У фазі атаки спостерігалися відхилення від норми, зокрема зменшення середнього часу життя сесій, поява коротких повторних SYN-послідовностей та підвищення кількості ICMP-повідомлень типу «Destination unreachable», що сигналізувало про примусове видалення або блокування певних записів у NAT-таблиці [3].

Додатково аналізувалися часові інтервали між повторними пакетами SYN та ACK, що дозволяло відстежити повторні спроби встановлення з'єднань після втрати відповідності у трансляції. На основі цих інтервалів формувалися часові ряди, які відображали поведінку потоків у динаміці. Аномалії виявлялися через обчислення середньоквадратичного відхилення параметрів, побудову ковзних вікон і застосування евристичних правил для фіксації нетипових стрибків у кількості зірваних сесій. Особливу увагу приділяли змінам у динаміці портів — фіксували випадки багаторазового перепризначення зовнішніх портів для одного внутрішнього клієнта, що свідчило про спробу обходу або підміни трансляції.

Важливою складовою методики стала багаторівнева кореляція журналів NAT, системного фаєрвола та моніторингу потоків NetFlow. Для кожного пакета формувався запис із часовою міткою, внутрішньою й зовнішньою IP-адресою, портом і станом з'єднання. Ці дані об'єднувалися у єдиний часовий простір, де алгоритм кореляції визначав послідовності подій, що призводили до втрати доступності ресурсу. Порівняння журналів дозволило відокремити легітимне завершення TCP-сеансів, викликане таймаутами або помилками маршрутизації, від штучного видалення відповідностей, ініційованого атакуючим.

Узгодження отриманих метрик дало змогу сформувати набір поведінкових сигнатур, що описують фазу блокування IP через NAT. До таких сигнатур належали поєднання зменшення TTL на один або два кроки, поява серій SYN-пакетів без успішного завершення тристороннього рукостискання та одночасна реєстрація ICMP-відповідей з різних зовнішніх адрес для одного внутрішнього хоста. Виявлення таких патернів у комплексі дало змогу підвищити точність діагностики аномалії й мінімізувати кількість хибнопозитивних спрацювань. У результаті узгоджена система аналізу показала здатність точно ідентифікувати атаки типу «блокування IP через NAT» навіть у публічних середовищах з високим рівнем

мультиарендності, забезпечуючи достовірне розмежування природних коливань трафіку від навмисних втручань у процес трансляції адрес.

Проведене дослідження довело ефективність поєднання аналізу поведінкових характеристик трафіку, часових рядів параметрів NAT і багаторівневої кореляції журналів у задачі виявлення атак типу «блокування IP через NAT». Розроблена методика забезпечує можливість розмежування природних і зловмисних відхилень у роботі механізму трансляції адрес, що підвищує надійність функціонування публічних мереж із великою кількістю клієнтів за спільною IP-адресою. Отримані результати підтверджують доцільність інтеграції запропонованого підходу в сучасні IDS/IPS-системи для оперативного виявлення прихованих атак у середовищах із багаторівневим NAT, де класичні сигнатурні методи є малоефективними.

Перелік посилань

1. Feng X., Yang Y., Li Q., Zhan X., Sun K., Wang Z., Wang A., Du G., Xu K. ReDAN: An Empirical Study on Remote DoS Attacks against NAT Networks. Network and Distributed System Security Symposium (NDSS 2025). Reston, VA: Internet Society, 2025. DOI: 10.14722/ndss.2025.230972.
2. Boukebous A. A. E., Fettache M. I., Bendiab G., Shiaeles S. A Comparative Analysis of Snort 3 and Suricata. 2023 IEEE Global Conference on Emerging Technologies (GlobConET). IEEE, 2023.
3. Dhumal C. T., Pingale D. S. V. Analysis of Intrusion Detection Systems: Techniques, Datasets and Research Opportunity. SSRN Electronic Journal. 2024. URL: https://doi.org/10.2139/ssrn.4749820.

АКТУАЛЬНІ ПРОБЛЕМИ КОМП'ЮТЕРНИХ НАУК 2025

ЗБІРНИК НАУКОВИХ ПРАЦЬ

Комп'ютерна верстка: **Мазурець О. В.**

Підписано до друку 15.11.2025.
Версія друку «APKN2025_CorpusPaper v5mod93 Final».

E-mail: *apkt.khnu@gmail.com*
ХНУ. м. Хмельницький, вул. Інститутська,11.

500

Завідувачу кафедри кібербезпеки
к.т.н., доц. Кльоцу Ю.П.
здобувача вищої освіти
Тростянецького Назара Олексійовича
студента ФІТ, 2 курсу, групи КБЗІм-24-1

ЗАЯВА

З правилами чинного Положення про систему забезпечення академічної доброчесності у Хмельницькому національному університеті, згідно з яким виявлення академічного плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту і застосування заходів академічної відповідальності, ознайомлений. Про використання спеціалізованих програмних засобів (СПЗ) StrikePlagiarism та Anti-Plagiarism для перевірки кваліфікаційних робіт здобувачів вищої освіти на наявність академічного плагіату оповіщений. Надаю університету право на передачу моєї роботи для обробки та збереження в базах даних СПЗ і використання роботи для виявлення академічного плагіату в інших роботах, які перевіряються СПЗ.

Також надаю свою згоду на обробку й збереження університетом моєї роботи в Інституційному репозитарії Хмельницького національного університету.

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

14.12.2025 р.
дата


підпис

Anti-Plagiarism (UA) v-15.284 Educational

The maximum coincidence with one document 0.0%

Dictionaries check: en_US, ru_RU, ua_UA. **Errors in the documents: 11%**

	Document		Sum coincidence on the DB	
ID: 253300 Title: Метод виявлення атак типу "блокування IP через NAT" в публічних мережах Added in a DB: 2025-12-16 Authors: Тростянецький Назар Олексійович Heads: Кльоц Ю.П. Consultants: Opponents:	Symbols	Lexemes	Symbols	Lexemes
	140075	865	594 (0%)	6 (1%)

Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

Автор: Тростянецький Назар Олексійович

Співавтор:

Назва: Метод виявлення атак типу "блокування IP через NAT" в публічних мережах

Науковий керівник: Кльоц Юрій Павлович

Підрозділ: Кафедра кібербезпеки

Коефіцієнт подібності 1:1%

Коефіцієнт подібності 2:0%

Мікропробіли: 0

Заміна букв: 0

Інтервали: 0

Білі знаки: 0

Дата створення звіту: 2025-12-16 14:48:33.0

Після аналізу Звіту подібності констатую наступне:

☑ **Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.**

☐ **Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.**

☐ **Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедурам. Таким чином робота не приймається.**

Обгрунтування:

Дата 16.12.2025р. експерт

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ

КАФЕДРИ КІБЕРБЕЗПЕКИ

ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

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

Назва: Метод виявлення атака типу "Блокування IP через NAT" в публічних мережах

Автор: Тростянецький Назар Олексійович

Спеціальність: 125 – Кібербезпека та захист інформації

Освітня програма: Кібербезпека та захист інформації

Науковий керівник: Юрій КЛЬОЦ, канд. техн. наук, доцент

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних). 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 (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). 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.	
5	Інше:	

Підтвердження:

Оригінальність тексту роботи за результатами перевірки системою StrikePlagiarism складає 99%, оригінальність тексту роботи за результатами перевірки системою Anti-Plagiarism складає 99%.

Згідно з Положенням про систему забезпечення академічної доброчесності у ХНУ (https://khmnu.edu.ua/wp-content/uploads/normatyvni-dokumenty/polozhennya/pro-systemu-zabezpechennya-akademichnoyi-dobrochesnosti.pdf, додаток В) кваліфікаційна робота, виконана за освітньо-професійною програмою, кількісні показники рівня унікальності тексту у відсотках до загального обсягу матеріалу в якій складає 75–100 %, визнається роботою з високим рівнем унікальності тексту («Текст вважається унікальним і не потребує додаткових дій щодо запобігання неправомірним запозиченням»).

Дата: 16.12.2025

Керівник роботи _____ Юрій КЛЬОЦ

Гарант ОП _____ Віра ТІТОВА

Завідувач кафедри кібербезпеки _____ Юрій КЛЬОЦ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА КВАЛІФІКАЦІНУ РОБОТУ

освітньо-кваліфікаційного рівня «магістр»

Магістр _____Тростянецький Назар Олексійович_____
Тема:__Метод виявлення атака типу "Блокування IP через NAT" в публічних мережах_____
Галузь знань 12 «Інформаційні технології»_Спеціальність 125 «Кібербезпека та захист інформації» Освітня програма «Кібербезпека та захист інформації»

Обсяг дипломної роботи освітньо-кваліфікаційного рівня «магістр»:

кількість листів креслень ___-___; кількість сторінок записки __100__;
1. Короткий зміст КР та прийнятих рішень Кваліфікаційна робота присвячена розробці методу виявлення атак типу «блокування IP через NAT» у публічних мережах, що виникають внаслідок адресного шарингу NAT/CGNAT. У роботі проаналізовано механізми блокування IP зовнішніми сервісами та вплив агрегації трафіку на зростання хибних спрацьовувань. Запропоновано метод детекції, заснований на аналізі NAT-специфічних логів і набору часових, транспортних, кореляційних та ентропійних ознак, що дозволяє локалізувати джерело аномальної активності всередині NAT-пулу. Реалізовано прототип системи та проведено експериментальні дослідження, які підтвердили ефективність прийнятих рішень і доцільність їх використання для зменшення кількості помилкових блокувань у системах мережевої безпеки._____
2. Висновок про відповідність КР завданню Магістерська робота у повній мірі відповідає поставленому завданню як у теоретичній і практичній частині роботи_____
3. Характеристика виконання кожного розділу роботи, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У вступі обґрунтовано актуальність проблеми помилкового блокування IP-адрес у публічних мережах з використанням NAT/CGNAT, визначено мету та завдання дослідження з урахуванням сучасних підходів до мережевої безпеки. У першому розділі проаналізовано механізми блокування IP зовнішніми сервісами, типові мережеві аномалії та вплив адресного шарингу на точність таких рішень. У другому розділі досліджено джерела даних і сучасні інструменти аналізу NAT-середовищ, обґрунтовано вибір ознак і методів обробки мережевих логів. У третьому розділі запропоновано та реалізовано метод виявлення атак типу «блокування IP через NAT», що базується на аналізі часових, транспортних, кореляційних та ентропійних характеристик трафіку, а також наведено результати експериментальної перевірки, які підтверджують ефективність застосування сучасних наукових і технічних підходів.

4. Позитивні сторони проекту полягають у підвищенні ефективності мережевої безпеки за рахунок виявлення атак типу «блокування IP через NAT» та зменшення кількості помилкових блокувань у публічних мережах. Запропонований підхід ґрунтується на аналізі NAT-специфічних мережевих ознак і поєднує статистичні методи з елементами машин-

ного навчання, що забезпечує стійкість до різних сценаріїв зловмисної активності. Рішення може бути інтегроване в існуючі системи моніторингу мережевої безпеки без суттєвого впливу на продуктивність мережі та якість обслуговування користувачів.

5. Негативні сторони проєкту полягають у обмеженому розгляді питань розгортання розробленого методу в реальних публічних мережах, а також у недостатньому описі процедур супроводу та оновлення його компонентів у процесі експлуатації, що дещо звужує практичні рекомендації щодо довготривалого застосування запропонованого рішення.

6. Оцінка графічного оформлення та пояснювальної записки роботи.

7. Відгук про роботу в цілому В загальному дипломна робота заслуговує позитивної оцінки, однак має незначні зауваження

8. Інші зауваження:

9. Оцінка дипломної роботи: Розглянувши позитивні та негативні сторони представленої дипломної роботи, можна зробити висновок, що дипломна робота заслуговує оцінки «добре»/B/83.

РЕЦЕНЗЕНТ (прізвище, ім'я, по батькові, посада, місце роботи)
Корецька Людмила Олександрівна .
Завідувач кафедри АКІТР, канд.техн.наук, доцент .

« 16 » грудня 2025. (підпис)