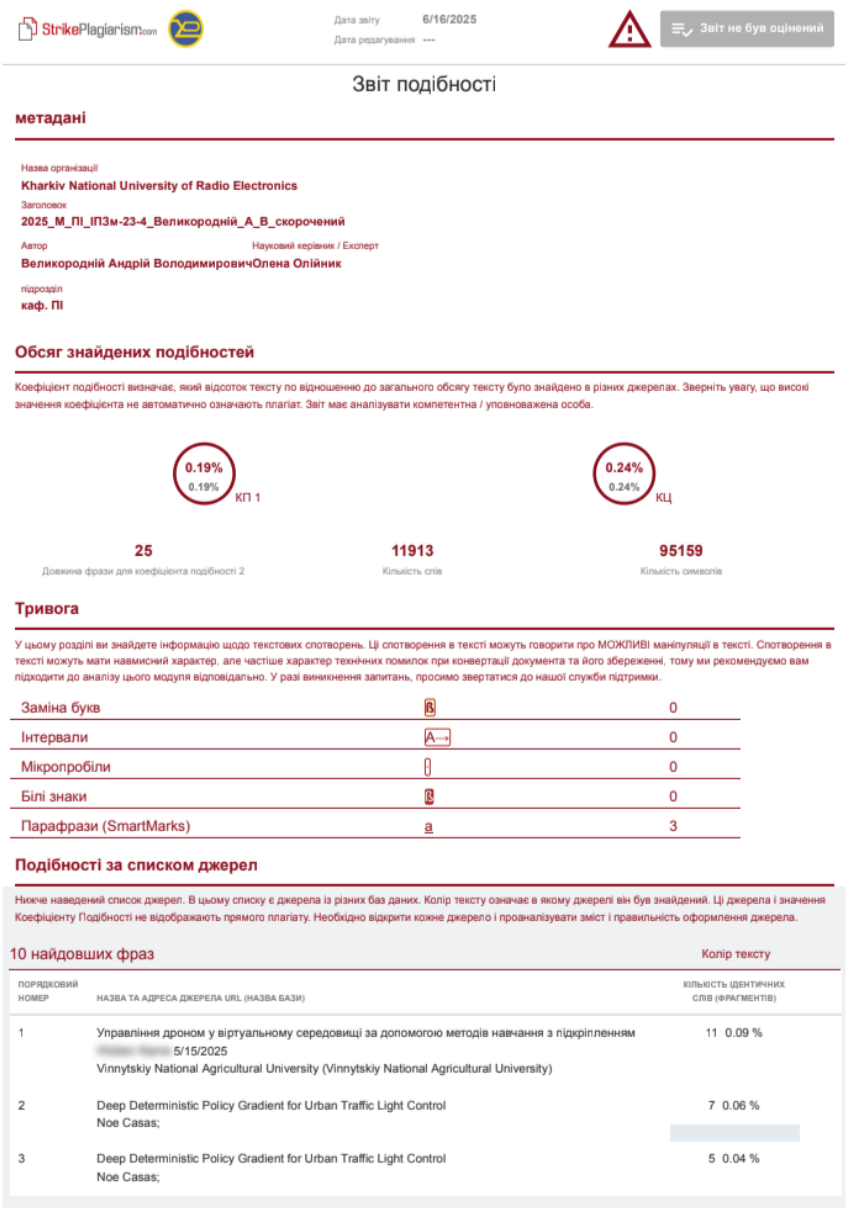
ДОДАТОК А

Звіт результатів перевірки на унікальність тексту в базі ХНУРЕ



StrikePlagiarism.com

Дата звіту 6/16/2025
Дата редагування ---

Звіт не був оцінений

Звіт подібності

метадані

Назва організації
Kharkiv National University of Radio Electronics
Заголовок
2025_М_ПІ_ІП3м-23-4_Великородній_А_В_скорочений
Автор
Великородній Андрій Володимирович
Науковий керівник / Експерт
Олена Олійник
підрозділ
каф. ПІ

Обсяг знайдених подібностей

Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.

0.19% 0.19% КП 1

0.24% 0.24% КЦ

25	11913	95159
Довжина фрази для коефіцієнта подібності 2	Кількість слів	Кількість символів

Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв	A	0
Інтервали	A⟷	0
Мікропробіли	(	0
Білі знаки	¶	0
Парафрази (SmartMarks)	a	3

Подібності за списком джерел

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз — Колір тексту

ПОРЯДКОВИЙ НОМЕР	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	Управління дроном у віртуальному середовищі за допомогою методів навчання з підкріпленням [redacted] 5/15/2025 Vinnytskiy National Agricultural University (Vinnytskiy National Agricultural University)	11 0.09 %
2	Deep Deterministic Policy Gradient for Urban Traffic Light Control Noe Casas;	7 0.06 %
3	Deep Deterministic Policy Gradient for Urban Traffic Light Control Noe Casas;	5 0.04 %

з бази даних RefBooks (0.10 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
джерело: https://arxiv.org/		
1	Deep Deterministic Policy Gradient for Urban Traffic Light Control Noe Casas;	12 (2) 0.10 %

з домашньої бази даних (0.00 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)

з програми обміну базами даних (0.09 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	Управління дроном у віртуальному середовищі за допомогою методів навчання з підкріпленням [redacted] 5/15/2025 Vinnytskiy National Agricultural University (Vinnytskiy National Agricultural University)	11 (1) 0.09 %

з Інтернету (0.00 %)

ПОРЯДКОВИЙ НОМЕР	ДЖЕРЕЛО URL	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)

Список прийнятих фрагментів (немає прийнятих фрагментів)

ПОРЯДКОВИЙ НОМЕР	ЗМІСТ	КІЛЬКІСТЬ ОДНАКОВИХ СЛІВ (ФРАГМЕНТІВ)

ДОДАТОК Б

Слайди презентації

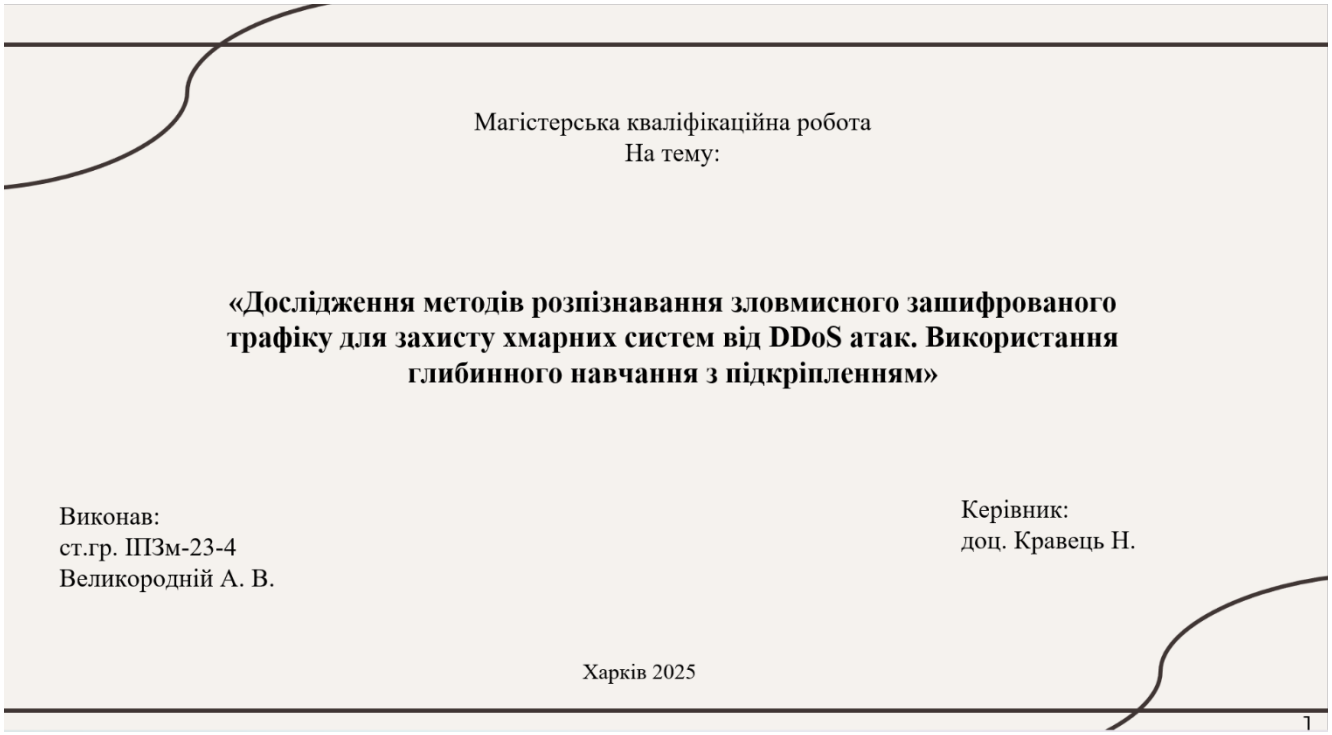


Рисунок Б.1 – Слайд 1

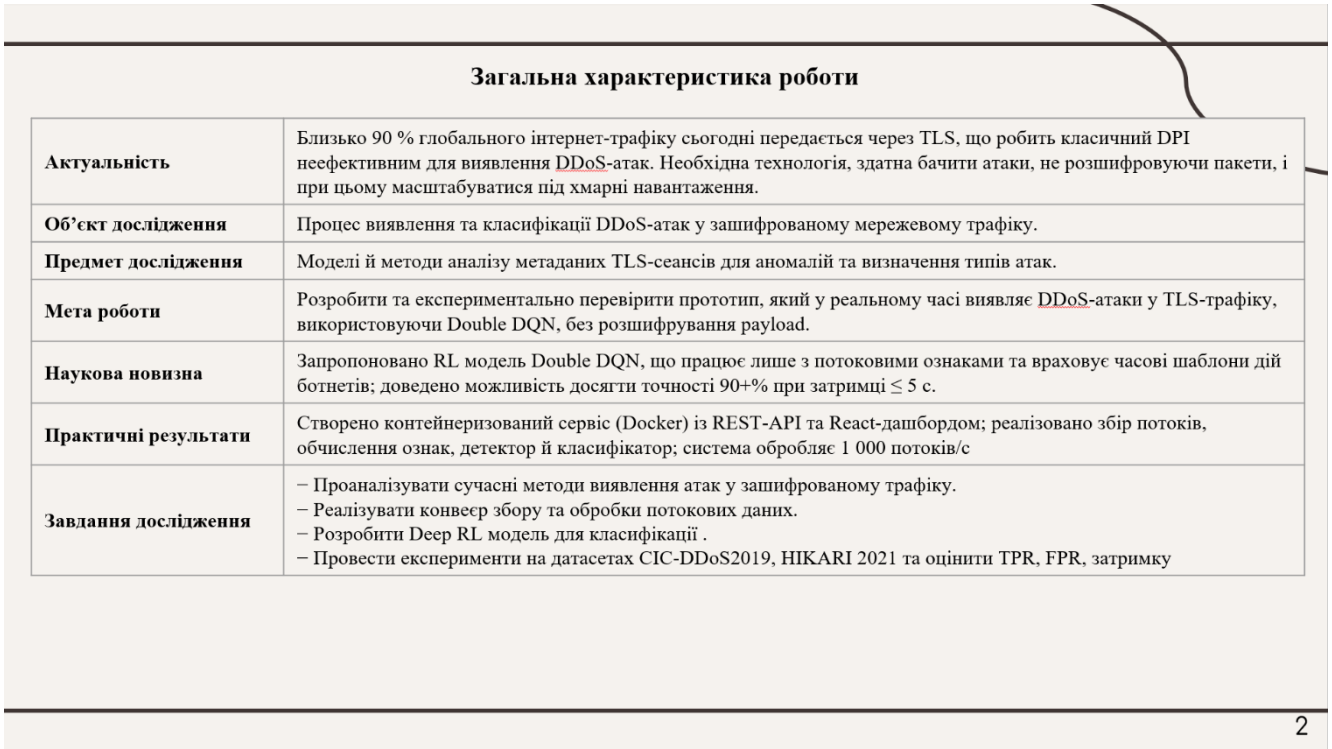


Рисунок Б.2 – Слайд 2

Ключові підходи до побудови системи виявлення DDoS-атак у зашифрованому TLS-трафіку

Метод	Переваги	Недоліки
Класичні supervised ML (Random Forest, SVM)	Швидке навчання на структурованих даних; висока інтерпретованість моделі; велика стабільність та зрілість реалізацій	Обмеженість у складних просторах станів; неможливість безпосередньо працювати з відгуками середовища; необхідність ручного інженерінгу ознак
Deep Q-Network (DQN)	Автоматичне витягування ознак; здатність працювати з високорозмірними вхідними даними; відкриття ефективності Deep RL	Завищення оцінок Q-функції (overestimation bias); нестабільність навчання; чутливість до гіперпараметрів
Double DQN	Усунення overestimation bias; більш стабільна та швидша збіжність; мінімальні зміни порівняно з DQN	Збільшені обчислювальні витрати через підтримку двох мереж; потреба додаткового тюнінгу інтервалу оновлення цільової мережі
Recurrent DQN (R2D2, DRQN)	Обробка частково спостережуваних задач (POMDP); гнучкість у роботі з послідовностями; підвищена ефективність вибірки в R2D2	Ускладнена архітектура; повільніше навчання через RNN-шари; потреба в складних стратегічних Replay-буферах

3

Рисунок Б.3 – Слайд 3

Постановка задачі

- **Мета:**
  Створити систему, що виявляє DDoS-атаки у TLS-трафіку без розшифрування пакетів і одразу пояснює, чому потік позначено як загрозу.
- **Проблема:**
  Шифрування приховує вміст пакетів, тож класичний DPI та сигнатури вже не працюють; статистичні ліміти дають забагато хибних тривог, а більшість існуючих рішень не враховують часові патерни атак.
- **Що потрібно:**
  Розробити Double DQN модель RL модель, що класифікує набори мережевих потоків;
- **Для чого:**
  Щоб центри безпеки могли швидко зупиняти розподілені атаки, отримувати пояснювані алерти та зменшити витрати на ручний аналіз хибних спрацювань.

4

Рисунок Б.4 – Слайд 4

Процес виявлення шкідливого трафіку

і інформацію про набір піПісля надходження набору мережевих-потоків система виділяє статистичні ознаки та нормалізує їх, щоб мережа отримала однорідні дані.

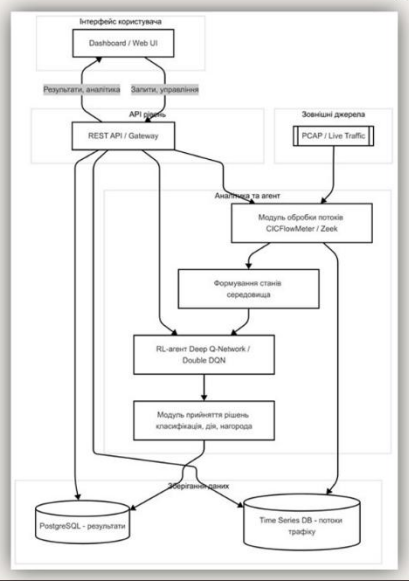
Double DQN обирає дію (позначити трафік як нормальний чи зловмисний)

При класифікації трафіку як зловмисний, користувач отримує сповіщення підозрілих потоків

5

Рисунок Б.5 – Слайд 5

Архітектура системи



- Завантажуємо CSV CICFlowMeter, формуємо ознаки та послідовності, нормалізуємо і масштабуємо вхідні дані.
- Формуємо вектори станів середовища для RL-агента. RL-агент Deep Q-Network / Double DQN приймає рішення щодо класифікації потоків і нарахування винагороди.
- Результати класифікації зберігаємо в PostgreSQL, а часоряди потоків – в Time Series DB.
- REST API / Gateway обслуговує запити, надаючи дані React-дашборду та зовнішнім SOC-сервісам.
  Усі модулі виконуються в Docker-контейнерах

6

Рисунок Б.6 – Слайд 6

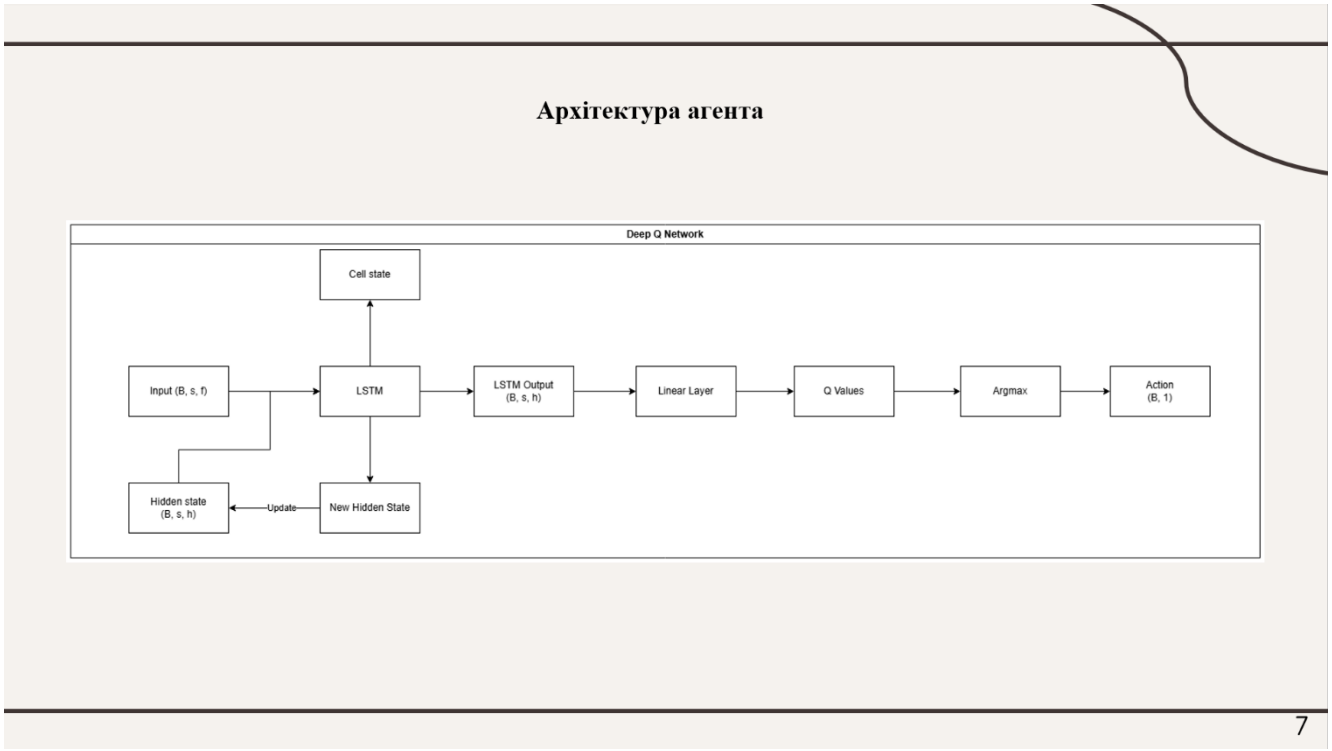


Рисунок Б.7 – Слайд 7

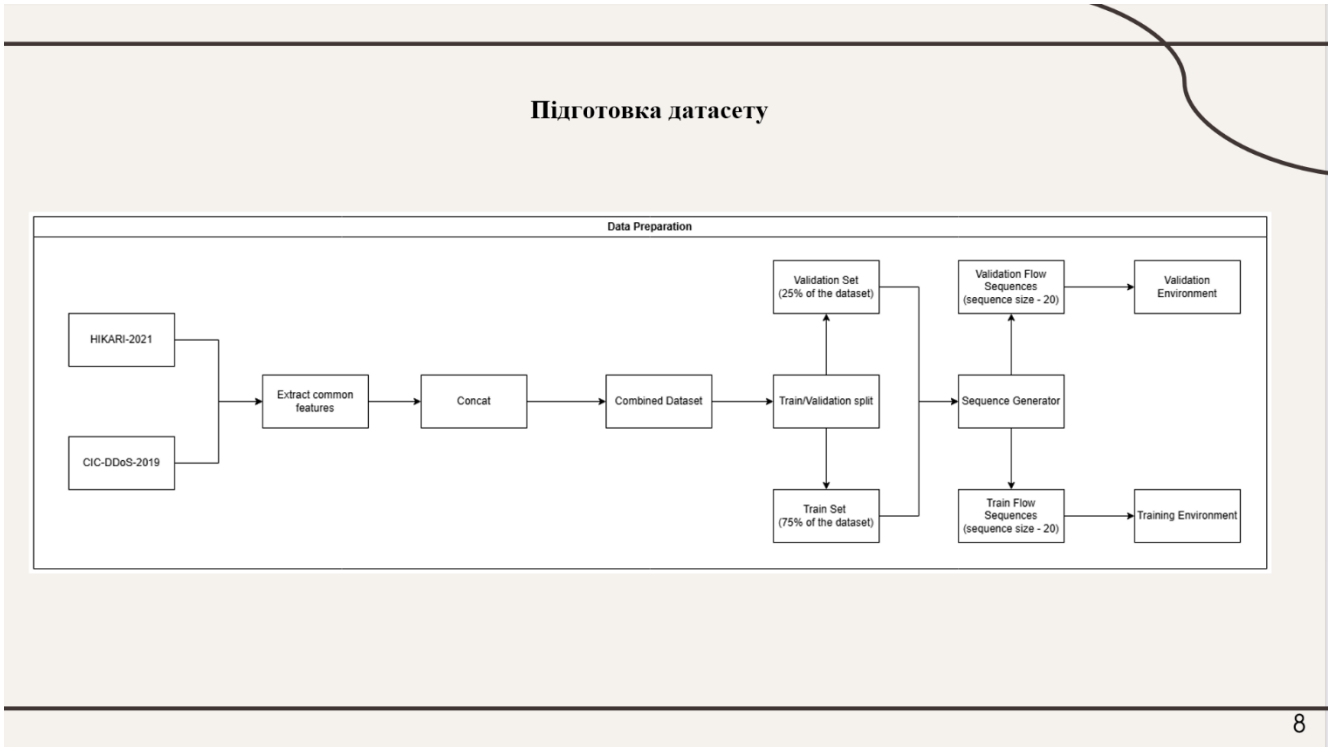


Рисунок Б.8 – Слайд 8

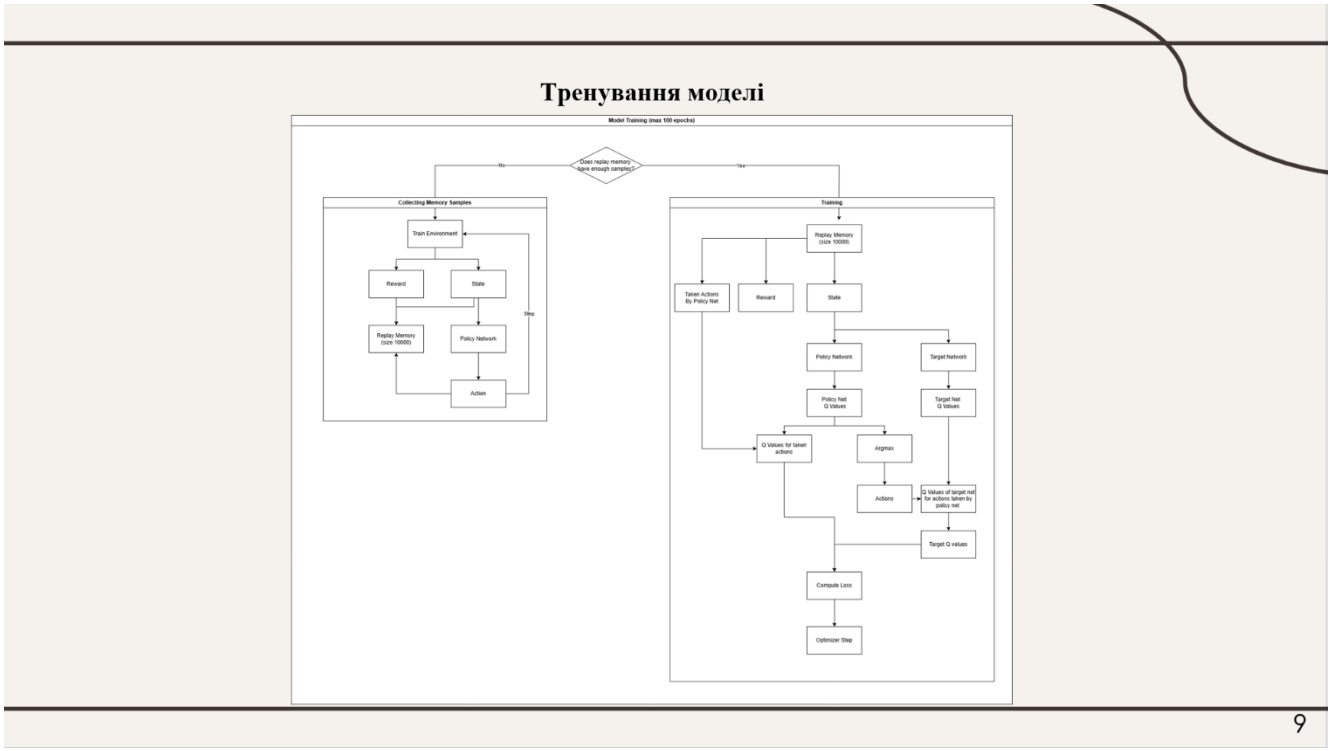


Рисунок Б.9 – Слайд 9

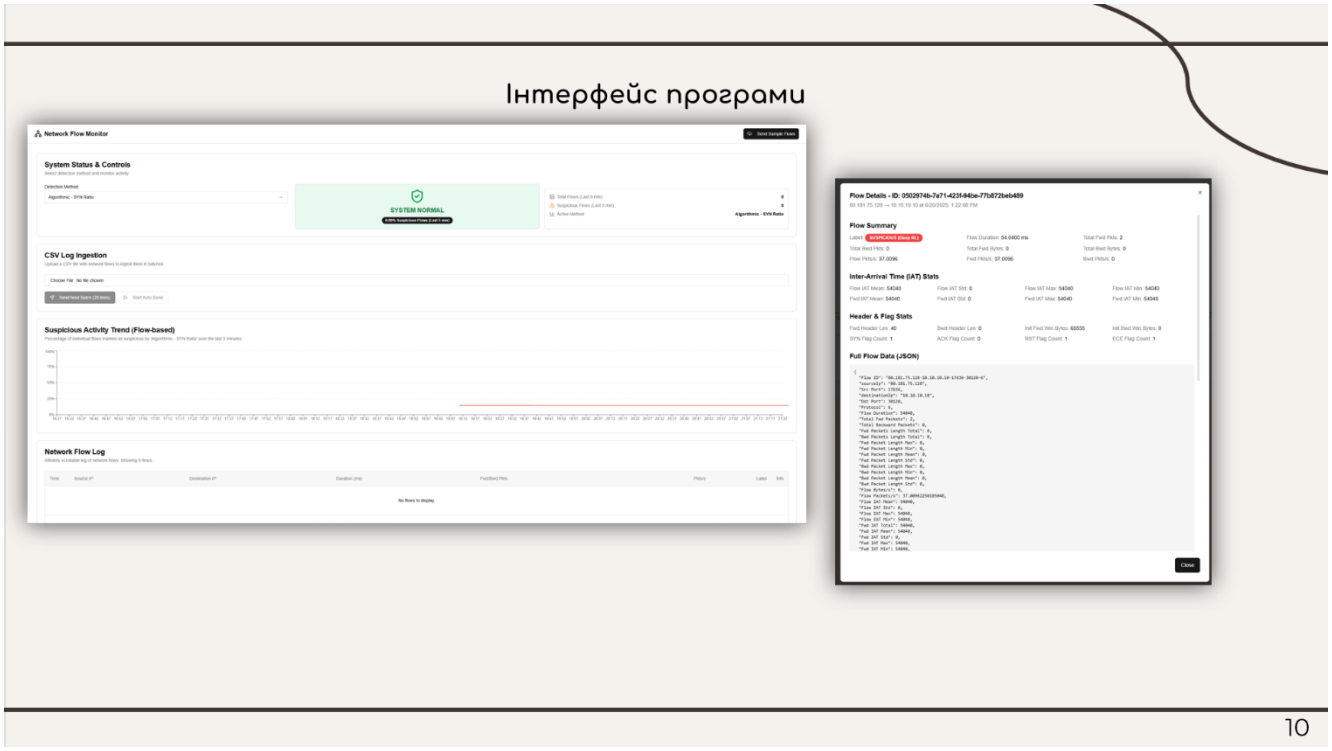
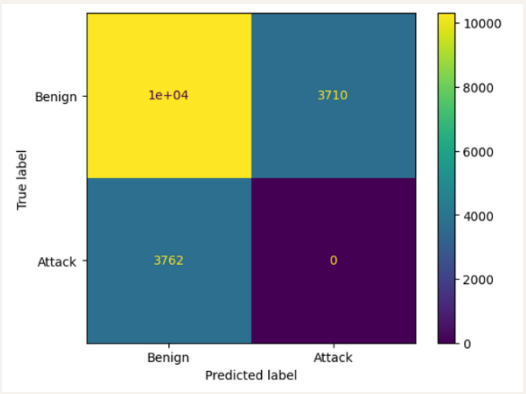


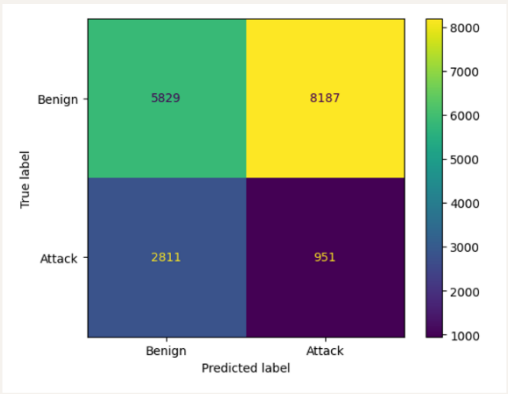
Рисунок Б.10 – Слайд 10

Результат експериментальної перевірки

Матриця помилок SYN Ratio



Матриця помилок IAT Regularity

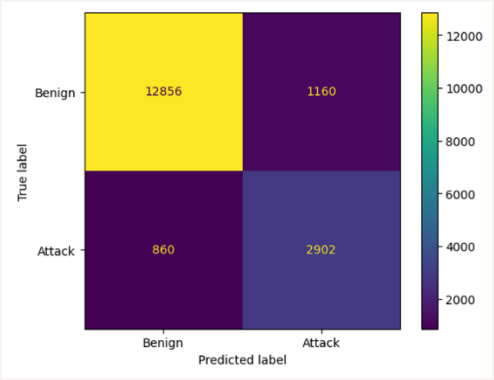


11

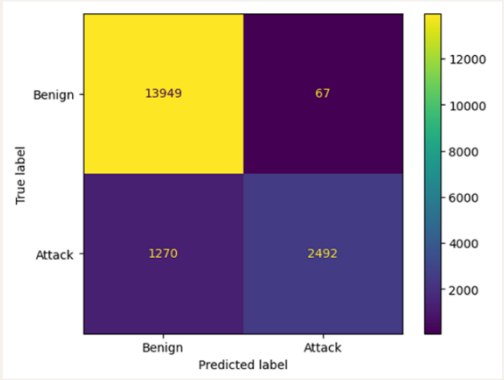
Рисунок Б.11 – Слайд 11

Результат експериментальної перевірки

Матриця помилок Unidirectionality



Матриця помилок Double Deep Q Network

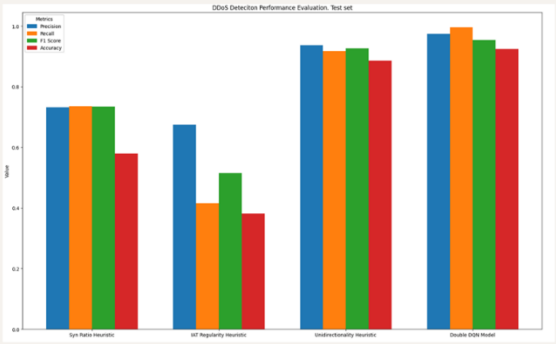


12

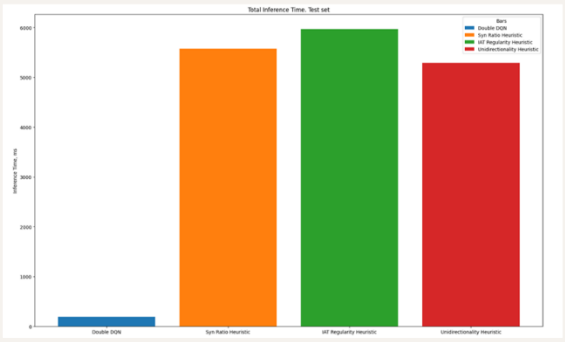
Рисунок Б.12 – Слайд 12

Результат експериментальної перевірки

Порівняння метрик усіх підходів



Порівняння швидкодії всіх підходів



13

Рисунок Б.13 – Слайд 13

Підсумки

Проаналізовано загрози DDoS у зашифрованому TLS-трафіку й доведено, що класичний DPI та порогові фільтри вже недостатні.

Запропоновано RL модель Double DQN для точного визначення типу атаки без розшифрування пакетів.

На тестах CIC-DDoS2019+HIKARI2021 отримано 90%+ точності та затримку аналізу < 5 с при 1 000 потоків/с.

14

Рисунок Б.14 – Слайд 14

Слайд Б.15 – Слайд 15

Слайд Б.16 – Слайд 16

ДОДАТОК В

Апробація результатів роботи

МАТЕМАТИЧНІ МОДЕЛІ ГЛИБИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ ТА ЗАПОБІГАННЯ DDOS-АТАК У ЗАШИФРОВАНОМУ ТРАФІКУ ХМАРНИХ СИСТЕМ

Kravets Natalia
Candidate of Technical Sciences, Associate Professor
Velykorodnii Andrii
second year graduated student (master's degree)
Department of Software Engineering
Kharkiv National University of Radio Electronics, Ukraine

Вступ

У сучасних хмарних середовищах зашифрований трафік становить домінуючу частку переданих даних, що унеможливлює традиційний аналіз вмісту пакетів при виявленні DDoS-атак. Теоретичний підхід на основі глибинного навчання з підкріпленням (Deep Reinforcement Learning, DRL) спирається на математичні моделі ухвалення рішень у невизначених умовах і дає змогу формалізувати завдання захисту як Марковський процес прийняття рішень (MDP) чи навіть як частково спостережуваний MDP (POMDP). Такий формалізм закладає основу для дослідження властивостей алгоритмів, їх збіжності і меж розв'язуваності проблем безпеки в умовах шифрування.

Теоретичні основи DRL

MDP визначається п'ятіркою (S, A, P, R, γ), де

– S – простір станів;

– A – множина дій;

– P(S`|s, a) – ймовірність переходу;

– R(s, a) – функція винагороди;

– $\gamma \in [0, 1)$ - коефіцієнт дисконтингу (Puterman, 1994). У застосуванні до зашифрованого трафіку стан sts_tst формується лише з метаданих (часових інтервалів, розмірів пакетів, частоти з'єднань), отже агент стикається з POMDP, оскільки реальний мережевий стан невидимий (Kaelbling, Littman, & Cassandra, 1998).

Основою алгоритмів DRL є рівняння Беллмана для функції вартості Q:

$$Q^*(s, a) = \sum_{s' \sim P(\cdot|s,a)} [R(s, a) + \gamma \max_{\alpha'} Q^*(s', a')]$$

У табличному випадку Q-навчання доведено збігається до Q∗ (Hu et al., 2024), але при апроксимації нейронними мережами загальні теореми збіжності відсутні, що вимагає подальшого формального аналізу.

Політикальні та актор-критик методи

Теореми про збіжність методів Policy Gradient гарантують досягнення локального максимуму за умови лінійної апроксимації (Yang, Liang, Wen, & Gao, 2021). Актор-критик алгоритми поєднують оцінювач станів (критик) із безпосереднім оновленням політики (актор), але їх стабільність залежить від налаштування темпів навчання та властивостей апроксимаційного простору (Yang et al., 2021).

Похибка апроксимації та sample complexity

Кумулятивна похибка апроксимації нейромережею (Bellman error) визначає верхню межу відхилення від оптимальних політик. Дослідження Ferriyan, Wibisono, та Muchtar (2021) показали, що для задачі класифікації DDoS у зашифрованому трафіку потребуються великі навчальні вибірки, що підтверджує експоненційне зростання sample complexity в просторі ознак.

Безпека та стійкість агентів

Safe RL вводить обмеження на множину допустимих політик для зменшення ризику надмірних помилок під час навчання, а Robust RL оптимізує політику за «найгіршим» сценарієм невизначеностей у моделі переходів (Yungaicela-Naula, Vagras-Rosales, Perez-Dias, & Carrera, 2022). Формальні гарантії у POMDP із зашифрованим трафіком залишаються предметом подальших досліджень.

Особливості зашифрованого середовища

Часткова спостережуваність

Агент отримує лише метадані, що підвищує ентропію спостережень і порушує маркову властивість, ускладнюючи побудову достатніх ознак для апроксимації оптимальної політики (Vargas-Rosales, Lopez-Ortiz, & Mendez-Hernandez, 2021).

Розсіювання та шум у даних

Метадані для легітимного та атакувального трафіку часто перекриваються, що ускладнює класифікацію. Створений набір HIKARI-2021 засвідчує

різноманітність та шум у зашифрованому трафіку для задачі детекції DDoS (Ferriyan et al., 2021).

Непостійність середовища

Еволюція тактик DdoS призводить до non-stationary середовища; алгоритми повинні враховувати концептуальний дрейф і адаптивно змінювати стратегію навчання (Yungaicela-Naula et al., 2022).

Атаки на навчальний процес

Data poisoning у буфері досвіду може спотворити функцію винагороди й модель переходів. Сертифіковані захисти від таких атак потребують інтеграції robust statistics із DRL (Kheddar, Messai, Himeur, & Awad, 2023).

Висновки

Методи DRL для виявлення та запобігання DdoS-атак у зашифрованому хмарному трафіку мають значний потенціал завдяки здатності адаптуватися до складних умов та часткової спостережуваності. Проте формальні гарантії збіжності при функціональній апроксимації залишаються невирішеними й вимагають обмежених класів апроксиматорів з доведеними властивостями. Необхідно дослідити trade-off між обсягом даних для навчання та точністю політик, а також розробити механізми safe/robust RL для контролю ризиків хибних класифікацій. Окрім того, критично важливо враховувати непостійність середовища й атаки на навчальні дані, формалізуючи adaptive та distributed DRL-алгоритми для масштабованих хмарних систем.

Список використаних джерел

1. Ferriyan, H., Wibisono, A., & Muchtar, A. (2021). HIKARI-2021 dataset: Benchmark for encrypted traffic DDoS detection. IEEE Access, 9, 7868–7886. https://doi.org/10.3390/app11177868
2. Hu, J., Yang, X., Hu, J., & Peng, Y. (2024). A Q-learning algorithm for Markov decision processes with continuous state spaces. Systems & Control Letters, 170, 105782. https://doi.org/10.1016/j.sysconle.2024.105782
3. Kheddar, H., Messai, N., Himeur, Y., & Awad, A. (2023). Deep transfer learning for intrusion detection in industrial control networks: A comprehensive review. Journal of Network and Computer Applications, 230, 103760. https://doi.org/10.1016/j.jnca.2023.103760
4. Vargas-Rosales, C., Lopez-Ortiz, G., & Mendez-Hernandez, H. (2021). SDN-based architecture for transport and application layer DDoS attack detection by using machine and deep learning. IEEE Access, 9, 3101650. https://doi.org/10.1109/ACCESS.2021.3101650
5. Yang, J., Liang, G., Wen, G., & Gao, T. (2021). A deep-learning- and reinforcement-learning-based system for encrypted network malicious traffic detection. Electronics Letters, 57(21), 906–908. https://doi.org/10.1049/ell2.12125
6. Yungaicela-Naula, M., Vagras-Rosales, C., Perez-Dias, J., & Carrera, D. (2022). A flexible SDN-based framework for slow-rate DDoS attack mitigation by using deep reinforcement learning. Journal of Network and Computer Applications, 216, 103444. https://doi.org/10.1016/j.jnca.2022.103444

ДОДАТОК Г

Експертний висновок результатів перевірки кваліфікаційної роботи на відповідність оформлення вимогам ДСТУ 3008: 2015

1

Експертний висновок результатів перевірки кваліфікаційної роботи

студент (посада) | програмної інженерії (кафедра) | ІПЗм-23-4 (група)

Андрій ВЕЛИКОРОДНІЙ
(прізвище, ім'я, по батькові)

Зауваження

Пункт ДСТУ 3008-2015	Зміст пункту	Сторінка кваліфікаційної роботи
1	2	3
	7.1 Загальні положення	
7.1.11	Рекомендовано на сторінках звіту використовувати береги такої ширини: верхній і нижній — не менше ніж 20 мм, лівий — не менше ніж 25 мм, правий — не менше ніж 10 мм.	По тексту
	7.3 Нумерація сторінок звіту	
	7.5 Рисунки	
	7.6 Таблиці	
7.6.9	Якщо рядки або колонки таблиці виходять за межі формату сторінки, таблицю поділяють на частини, розміщуючи одну частину під іншою або поруч, чи переносять частину таблиці на наступну сторінку. У кожній частині таблиці повторюють її головку та боковик. У разі поділу таблиці на частини дозволено її головку чи боковик заміняти відповідно номерами колонок або рядків, нумеруючи їх арабськими цифрами в першій частині таблиці. Слово «Таблиця » подають лише один раз над першою частиною таблиці. Над іншими частинами таблиці з абзацного відступу друкують «Продовження таблиці » або «Кінець таблиці ____ » без повторення її назви.	38
	7.7 Переліки	
	7.8 Примітки	
	7.9 Виноски	
	7.10 Формули та рівняння	
7.10.6	Пояснення познак, які входять до формули чи рівняння, треба подавати безпосередньо під формулою або рівнянням у тій послідовності, у якій їх наведено у формулі або рівнянні. Пояснення познак треба подавати без абзацного відступу з нового рядка, починаючи зі слова «де» без двокрапки. Познаки, яким встановлюють визначення чи пояснення, рекомендовано ви-рівнювати у вертикальному напрямку.	40
	7.11 Посилання	
	7.13 Список авторів	
	7.14 Скорочення та умовні познаки	
	7.15 Додатки	

Методичні вказівки до виконання кваліфікаційної роботи магістра… ЗАТВЕРДЖЕНО кафедрою ПІ протокол № 12 від 03.02.2025 р 3.2 Оформлення пояснювальної записки згідно з ДСТУ 3008:2015 Звіти у сфері науки і техніки. Структура та правила оформлювання. **Шаблон**:ЗАТВЕРДЖЕНО кафедрою ПІ протокол № 12 від 03.02.2025 р.	Увага! встановлені фіксовані береги: лівий – 25 мм., правий – 10 мм, верхній і нижній – 20 мм.	По тексту

Експерт _____ Вадим НЕЧВОЛОД
(підпис) (прізвище, ініціали)

Робота з перевірки оформлення пояснювальної записки кваліфікаційної роботи на нормоконтроль виконана у програмі Word Microsoft 365. Версія 2504 (збірка 18730.20220)
19.06.2025