Хмельницький національний університет
Факультет інформаційних технологій
Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

Кулачук Марії Романівни

на здобуття ступеня вищої освіти Бакалавра

Система виявлення DDoS-атак у мережах 5GN

Галузь знань 12 – Інформаційні технології

Спеціальність 125 – Кібербезпека

Освітня програма Кібербезпека

КРБКБ.220161.22.01.04 ПЗ

Виконала студентка 3 курсу, група КБс-22-1 _____ Марія КУЛАЧУК
Підпис, дата — Ініціали, прізвище

Керівник докт. тех. наук, професор _____ Михайло КАСЯНЧУК
Науковий ступінь, вчене звання — Підпис, дата — Ініціали, прізвище

Нормоконтролер старший викладач 02.06.25 Сергій МОСТОВИЙ
Науковий ступінь, вчене звання — Підпис, дата — Ініціали, прізвище

До захисту допускаю:

Зав. кафедри кібербезпеки _____ Юрій КЛЬОЦ
Підпис, дата — Ініціали, прізвище

02 06 2025р.

Хмельницький, 2025

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет Інформаційних технологій

Кафедра Кібербезпеки

Рівень вищої освіти Бакалавр

Галузь знань 12 – Інформаційні технології

Спеціальність 125 – Кібербезпека

Освітня програма Кібербезпека

ЗАТВЕРДЖУЮ

Завідувач кафедри кібербезпеки

Юрій КЛЬОЦ

15 02 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ
Кулачук Марії Романівні

1 Тема роботи Система виявлення DDoS-атак у мережах 5GN

Керівник роботи д.т.н, проф. кафедри кібербезпеки Михайло Миколайович Касянчук

Затверджено наказом ректора університету від 7 02 2025 № 23

2 Строк подання студентом кваліфікаційної роботи на кафедру 2.06.25

3 Вихідні дані до роботи Дослідити сучасні архітектурні особливості 5G-мереж та їхню вразливість до DDoS-атак. Розробити систему виявлення DDoS-атак із використанням глибоких нейронних мереж CNN та BLSTM. Сформувати набір даних на основі реалістичного 5G-трафіку. Провести навчання моделей та оцінку їх достовірності.

4 Зміст пояснювальної записки (перелік питань, які потрібно розробити) Огляд архітектури 5G-мереж. Опис типів DDoS-атак. Аналіз наявних підходів до виявлення атак. Постановка задачі. Розробка алгоритму виявлення атак. Формування датасетів на основі 5G-трафіку. Архітектури моделей CNN та BLSTM. Параметри навчання. Оцінка продуктивності моделей. Метрики достовірності класифікації. Порівняння результатів.

5 Перелік графічного матеріалу (із зазначенням обов'язкових креслень) Налаштування МН для прийняття рішення. Блок-схема процесу класифікації трафіку. Схема роботи LSTM та BLSTM мереж.

6 Консультанти розділів кваліфікаційної роботи

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

7 Дата видачі завдання 16 02 2025 р.

КАЛЕНДАРНИЙ ПЛАН

Назва етапів (розділів) кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
Вибір і затвердження теми кваліфікаційної роботи	07.02.2025	
Ознайомлення з предметною областю	10.02.2025	
Дослідження існуючих рішень	26.02.2025	
Постановка задачі	06.03.2025	
Визначення загальних принципів рішення задачі	18.03.2025	
Деталізація принципів рішення задачі	14.04.2025	
Розробка проєктних рішень	24.04.2025	
Апробація проєктних рішень	04.05.2025	
Оформлення пояснювальної записки згідно вимог	28.05.2025	
Оформлення графічної частини	31.05.2025	
Захист КР	10.06.2025	

Студентка _____ Марія КУЛАЧУК

Керівник кваліфікаційної роботи _____ Михайло КАСЯНЧУК

АНОТАЦІЯ

Тема кваліфікаційної роботи: «Система виявлення DDoS-атак у мережах 5GN».

Авторка роботи: Кулачук Марія Романівна.

Керівник роботи: Касянчук Михайло Миколайович.

Пояснювальна записка: 64 с., 1 додаток, 12 рис., 41 джерел.

Графічна частина: 3 презентаційних слайди.

Ключові слова: 5G, DDoS-атака, CNN, BLSTM, нейронна мережа, мережевий трафік, кібербезпека.

Завданням даної кваліфікаційної роботи є розробка системи виявлення DDoS-атак у мережах п'ятого покоління (5G) з використанням глибоких нейронних мереж CNN та BLSTM. Така система дозволяє ідентифікувати мережеві атаки на основі поведінкового аналізу трафіку й забезпечує ефективний захист критичних функцій ядра 5G-мереж. У процесі дослідження було: змодельовано мережу 5G з використанням Free5GC та UERANSIM; зібрано і сформовано датасет на основі легітимного та атакуючого трафіку; розроблено алгоритми класифікації із застосуванням CNN та BLSTM; проведено навчання моделей і тестування їх достовірності. Також виконано порівняльний аналіз результатів і здійснено оцінку ефективності системи виявлення атак у симульованому середовищі. Робота має практичне значення для підвищення безпеки мереж нового покоління в умовах зростання обсягу даних і кількості пристроїв.

25.05.25

ABSTRACT

Topic of the qualification work: «DDoS attack detection system in 5GN networks»

Author: Kulachuk Maria Romanivna

Supervisor: Kasyanchuk Mykhailo Mykolayovych

Explanatory note: 64 p., 1 appendix, 12 figures, 41 references.

LIST OF KEYWORDS: 5G, DDoS attack, CNN, BLSTM, neural network, network traffic, cybersecurity.

The objective of this qualification work is to develop a system for detecting DDoS attacks in fifth-generation (5G) networks using deep neural networks CNN and BLSTM. Such a system allows to identify network attacks based on behavioral traffic analysis and provides effective protection of critical functions of the core of 5G networks. In the course of the study, the 5G network was modeled using Free5GC and UERANSIM; a dataset was collected and formed based on legitimate and attacking traffic; classification algorithms were developed using CNN and BLSTM; models were trained and tested for their reliability. A comparative analysis of the results was also performed and the effectiveness of the attack detection system in a simulated environment was evaluated. The work is of practical importance for improving the security of next-generation networks in the face of increasing data volume and the number of devices.

25. 05. 25

ЗМІСТ

Зм.	Арк.	№ докум.	Підпис	Дата	*КРКБ.220161.22.01.04 ПЗ*		
Виконала		Кулачук М.Р.		25.05	Система виявлення DDoS-атак у мережах 5GN Пояснювальна записка	Літера	Аркуш 2 / Аркушів 64
Перевір.		Касянчук М.М.					
Н.контр.		Мостовий С.В		06.06.25		*ХНУ, КБс-22-1*	
Затвер.		Кльоц Ю.П		2.06.25			

ВСТУП

Швидкий розвиток інформаційно-комунікаційних технологій, зокрема мобільного зв'язку п'ятого покоління (5G), відкриває нові горизонти для цифрової трансформації, розвитку Інтернету речей (IoT), автономного транспорту, розумних міст і промисловості 4.0. Ця технологія забезпечує значно вищу швидкість передавання даних, менші затримки, більшу щільність підключень і надійність у порівнянні з попередніми поколіннями мобільного зв'язку. Проте, попри всі переваги, 5G-мережі стають об'єктом нових загроз і кіберінцидентів, зокрема атак типу відмови в обслуговуванні (DoS) та розподілених атак (DDoS), які є одними з найпоширеніших і найнебезпечніших інструментів порушення доступності критичних інформаційних систем. Унікальні архітектурні особливості 5G, зокрема використання мережевої віртуалізації, мікросегментації (network slicing), розподіленої обробки даних і сервісно-орієнтованого ядра, з одного боку забезпечують гнучкість і масштабованість, а з іншого — створюють нові вектори атак, які потребують комплексного підходу до захисту. Одним із викликів є захист функцій ядра мережі, таких як AMF, SMF, NRF та UPF, від перевантаження, яке може бути спровоковане цілеспрямованими DDoS-атаками. Такі атаки можуть не лише порушити функціонування певного сегменту мережі, але й вивести з ладу сервіси, що працюють у рамках виділених віртуальних зрізів, критичних для медицини, транспорту чи енергетики.

Із ростом кількості IoT-пристроїв, які часто мають обмежені ресурси захисту, підвищується ймовірність їхнього використання як частин ботнет-мереж для генерації DDoS-трафіку. Це ставить завдання не лише виявляти аномальні дії з боку зовнішніх джерел, а й проводити ідентифікацію скомпрометованих пристроїв усередині самої мережі. Актуальним стає впровадження механізмів раннього попередження, автоматичного реагування й самонавчання систем безпеки.

Традиційні методи виявлення атак, що базуються на сигнатурному або евристичному аналізі, мають обмежену ефективність у нових умовах. Зокрема, вони не здатні адаптуватися до нових типів загроз, швидко змінюваних моделей трафіку та широкого спектру аномалій. Тому актуальним стає використання

підходів, заснованих на машинному навчанні, а особливо глибокого навчання, які дозволяють не лише виявляти відомі загрози, але й розпізнавати невідомі атаки на основі поведінкового аналізу. Такі моделі можуть автоматично адаптуватися до змін у трафіку, аналізувати високовимірні дані, ідентифікувати нетипові шаблони та знижувати кількість помилкових спрацьовувань.

Сучасні дослідження демонструють перспективність застосування згорткових нейронних мереж (Convolutional Neural Networks, CNN) та двонаправлених довготривалих короткочасних мереж пам'яті (Bidirectional Long Short-Term Memory, BLSTM) для класифікації мережевого трафіку. Моделі CNN здатні виявляти просторові закономірності в потоках даних, тоді як BLSTM дозволяють враховувати тимчасові залежності між подіями, що є критично важливим у контексті динамічних мереж 5G. Поєднання цих підходів дозволяє створити потужну систему для аналізу трафіку, яка здатна швидко реагувати на загрози, навіть якщо вони ще не мають визначеної сигнатури.

Окрім технічних аспектів виявлення, важливою складовою є якість даних, які використовуються для навчання моделей. У випадку 5G-мереж стандартні набори даних не завжди відображають реальні умови, тому важливим етапом дослідження є створення та використання нових датасетів, що моделюють типові та нетипові сценарії взаємодії в інфраструктурі 5G. Врахування трафіку, що генерується як легітимними, так і шкідливими джерелами, дозволяє сформувати репрезентативну вибірку для навчання і тестування моделей, що є передумовою для їх надійності в реальному середовищі.

У цьому контексті актуальність цієї кваліфікаційної роботи полягає у розробці інтелектуальної системи виявлення DDoS-атак для 5G-мереж із використанням глибоких моделей нейронних мереж. Система повинна бути здатною до аналізу високошвидкісного мережевого трафіку в режимі, наближеному до реального часу, забезпечуючи точність, стабільність та адаптивність. Додатково розглядається можливість її масштабування та інтеграції у віртуалізовані середовища шляхом використання технологій контейнеризації (Docker) та розгортання у GPU-прискорених обчислювальних системах.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА НАЯВНИХ РІШЕНЬ

1.1 Особливості 5G-мережі

Вперше запроваджена у світі в 2019 році, 5G-мережа є черговим етапом розвитку технологій мобільного зв'язку, який став потужним наступником 4G-мереж, що наразі домінують у сфері мобільної підключеності. Ця п'ята генерація технологій покликана суттєво трансформувати інтернет-сервіси, забезпечуючи вищу швидкість і надійність з'єднання, з потенційними показниками до 20 Гбіт/с за оптимальних умов при підключенні одного користувача [1]. Завдяки збільшеній пропускній здатності 5G підтримує велику кількість одночасних підключень пристроїв без втрати продуктивності, що є важливим в епоху цифровізації [2].

Архітектура 5G є результатом еволюційного розвитку мобільного зв'язку і значно відрізняється від попередніх поколінь, таких як 3G і 4G, не лише за швидкістю передачі даних, але й за принципами побудови, масштабованістю та функціональністю. Якщо 3G був зосереджений на розширенні можливостей мобільного інтернету, а 4G забезпечив стабільну IP-основу для масового потокового мультимедіа та широкосмугового доступу, то 5G фактично відкрив нову еру — з орієнтацією на ультранадійний зв'язок, мінімальну затримку та обслуговування мільйонів пристроїв в одиниці площі.

Однією з інновацій архітектури 5G стало впровадження сервісно-орієнтованого ядра (Service-Based Architecture, SBA), де основні функції розділено на окремі модулі, які взаємодіють між собою через стандартизовані інтерфейси. Цей підхід суттєво відрізняється від монолітної структури 4G, де багато функцій були жорстко інтегровані в єдину систему, що обмежувало гнучкість і масштабованість.

Функція управління доступом і мобільністю (AMF) у 5G є центральним вузлом для ініціалізації з'єднання та підтримки мобільності абонента. Наприклад, коли пристрій намагається підключитися до мережі, AMF відповідає за автентифікацію користувача, встановлення початкового з'єднання та маршрутизацію запитів до інших мережевих функцій. Якщо користувач змінює своє фізичне розташування — наприклад, рухається з одного міста до іншого —

саме AMF забезпечує збереження сеансу та неперервність зв'язку без втрати сигналу. На відміну від 4G, де ця логіка була частково закладена в компоненти як MME (Mobility Management Entity), у 5G вона винесена окремо, що дозволяє забезпечити швидшу реакцію системи на зміни та динамічне балансування навантаження.

Інша функція — UPF (User Plane Function) — відповідає за обробку і маршрутизацію користувацького трафіку. Її можна порівняти з "автострадою" для даних, яка забезпечує передачу інформації від користувача до призначеного сервісу, наприклад, до відеострімінгового сервера чи хмарного додатку. У випадку 5G UPF здатна бути розміщеною ближче до кінцевого користувача — наприклад, у регіональному центрі обробки даних — що значно скорочує час затримки і робить можливим запуск критичних застосунків, таких як автономне водіння або хірургічні маніпуляції на відстані. У попередніх поколіннях зв'язку, зокрема в 4G, така гнучкість розгортання користувацької площини була відсутня: всі пакети проходили через центральні вузли EPC (Evolved Packet Core), що створювало вузькі місця та збільшувало затримки.

Загалом архітектура 5G створена за принципами модульності, віртуалізації та хмарноорієнтованості. Завдяки цьому вона легко адаптується до нових сценаріїв використання — від індустріального Інтернету речей до систем масового відеоспостереження. У той час як 3G була лише першим кроком до "мобільного інтернету", а 4G — відповіддю на зростаючі вимоги мультимедіа, 5G — це вже інтелектуальна мережева екосистема, яка не просто передає дані, а управляє ними в реальному часі відповідно до потреб конкретних сервісів і користувачів.

5G-мережа організована у вигляді комірок, кожна з яких має базову станцію, що передає дані через радіохвилі у визначених частотних діапазонах, підключених до глобальної мережі інтернет і телекомунікаційних систем за допомогою високопродуктивних каналів, таких як оптоволоконні кабелі або бездротові магістралі. Стандарти для таких мереж, зокрема використання програмного забезпечення New Radio (NR) і кодування Orthogonal Frequency Division Multiplexing (OFDM), розробляються в межах 3rd Generation Partnership Project (3GPP). 5G працює в різних частотних діапазонах, серед яких: низькочастотний

діапазон (600–900 МГц), що забезпечує швидкість від 5 до 250 Мбіт/с і охоплення, схоже на 4G [3]; середньочастотний діапазон (1,7–4,7 ГГц), який підходить для міських територій і забезпечує швидкість від 100 до 900 Мбіт/с [4]; високочастотний діапазон (24–47 ГГц), який надає швидкість, подібну до кабельного інтернету, але потребує щільного розміщення базових станцій через обмежений радіус дії [5]. Частоти в діапазоні 24,25–29,5 ГГц, які активно ліцензуються в усьому світі, мають ключове значення для високошвидкісних застосувань, таких як масові впровадження Інтернету речей (IoT) і міжмашинну комунікацію, що відображає їхню важливу роль у сферах, які вимагають високої пропускної здатності.

Із розвитком технологій 5G очікується, що вони покращуватимуть швидкість і ефективність бездротового зв'язку в різних галузях, суттєво розширюючи можливості та охоплення мереж нового покоління [6].

Архітектура 5G – це всебічна структура, створена для забезпечення сучасних бездротових послуг із підвищеною ефективністю, пропускною здатністю та масштабованістю. Ця архітектура також називається Service-Based Architecture (SBA) або сервісно-орієнтованою архітектурою ядра 5G. На відміну від традиційних жорстких систем, вона базується на більш гнучкому та динамічному підході для управління різноманітними додатками та запитами користувачів. Архітектура складається з трьох основних компонентів: обладнання користувача (User Equipment, UE), радіомережа доступу (Radio Access Network, RAN), функції ядра мережі (Core Network Functions).

Обладнання користувача охоплює будь-які пристрої, що дозволяють кінцевим користувачам отримувати доступ до мережі 5G. Це включає смартфони, планшети, а також різноманітні пристрої IoT. Такі пристрої є кінцевою точкою взаємодії користувача з мережею для доступу до послуг [7].

Радіомережа доступу забезпечує зв'язок між обладнанням користувача та ядром мережі. Вона складається з базових станцій, обладнаних необхідною радіотехнікою для бездротового обміну даними. Базові станції стратегічно розташовані для забезпечення оптимального покриття та продуктивності. RAN відповідає за управління радіоресурсами та підтримання якості обслуговування під

час переміщення обладнання користувача між різними комірками [8].

Ядро мережі є центром управління інтелектуальними функціями та ресурсами 5G. Воно складається з декількох компонентів:

– функція управління доступом і мобільністю (Access and Mobility Management Function, AMF) відповідає за управління підключеннями, автентифікацію користувачів, встановлення сесій та ефективне управління мобільністю. Ця функція забезпечує безперервне обслуговування під час переміщення користувача між зонами мережі [9];

– функція управління сесіями (Session Management Function, SMF) працює спільно з AMF для управління станами сесій, їхнім створенням, підтримкою та завершенням. SMF забезпечує стабільний та надійний потік даних через мережу [10];

– функція сервера автентифікації (Authentication Server Function, AUSF) грає важливу роль у забезпеченні безпеки, автентифікуючи пристрої користувачів та перевіряючи їхню ідентичність перед доступом до мережі. Це допомагає захиститися від несанкціонованого доступу та потенційних загроз [11];

– функція управління даними (Unified Data Management, UDM) централізує дані абонентів і управляє їхнім обробленням, включно з автентифікаційними даними та інформацією про підписку. Ця функція гарантує, що доступ до мережі отримують лише автентифіковані користувачі [12];

– функція управління політиками (Policy Control Function, PCF) встановлює правила використання ресурсів мережі, регулюючи трафік даних відповідно до угод про обслуговування та політик мережі, щоб забезпечити користувачам обіцяний рівень якості послуг [13];

– функція репозиторію мережі (Network Repository Function, NRF) діє як каталог, що містить записи про всі доступні функції та послуги мережі. NRF допомагає у пошуку та реєстрації функцій, забезпечуючи ефективну взаємодію між ними [14];

– функція взаємодії з мережею (Network Exposure Function, NEF) відповідає за безпечний доступ до можливостей мережі для сторонніх додатків. Це сприяє створенню екосистеми різноманітних послуг та застосунків [15];

– функція користувацької площини (User Plane Function, UPF) відповідає за високошвидкісне транспортування даних, маршрутизацію, управління трафіком та виконання політик, пов'язаних із передачею даних [16].

1.2 Типи DDoS-атак та особливості їх реалізації

Атаки відмови в обслуговуванні (Denial of Service, DoS) та розподілені атаки відмови в обслуговуванні (DDoS) є загрозою для доступності мережевих ресурсів, значно порушуючи доступ як для організацій, так і для окремих користувачів. Такі атаки перевантажують мережеві сервіси надмірним трафіком, унеможливлюючи доступ до них для легітимних користувачів [17-19]. У межах цієї роботи моделюються такі атаки, щоб підкреслити важливість розуміння різних методів DoS і DDoS та необхідність впровадження надійного захисту для мережевої інфраструктури.

DDoS-атаки є складнішою формою DoS-атак, оскільки трафік надходить з багатьох джерел, наприклад, з ботнетів — мереж заражених пристроїв, контрольованих зловмисником (рис. 1). Це ускладнює відновлення доступності сервісу та часто призводить до значних фінансових і репутаційних збитків для атакованих організацій [20]. Серед різновидів цих атак: об'ємні атаки, які перенасичують пропускну здатність мережі; атаки на рівні протоколів, наприклад, TCP-флуд і UDP-флуд, що використовують уразливості мережевих протоколів; атаки на рівні додатків, які спрямовані на специфічні «слабкі» місця веб-додатків [21-23]. Такі атаки можуть тривати від кількох хвилин до кількох днів, ускладнюючи роботу систем адміністрування та вимагаючи оперативної реакції. Залежно від характеру атаки, навантаження може бути спрямоване як на зовнішній периметр мережі, так і безпосередньо на внутрішні вузли інфраструктури, включаючи ядро 5G. Особливо небезпечними є ситуації, коли атака спрямована на віртуалізовані компоненти, які обслуговують критичні функції, як-от AMF або UPF. Таке навантаження може призвести до збоїв у роботі всієї мережі та порушення обслуговування навіть у віддалених сегментах. У зв'язку з цим зростає

потреба в інтелектуальних системах виявлення та протидії атакам, здатних працювати в реальному часі й адаптуватися до змінних сценаріїв загроз.

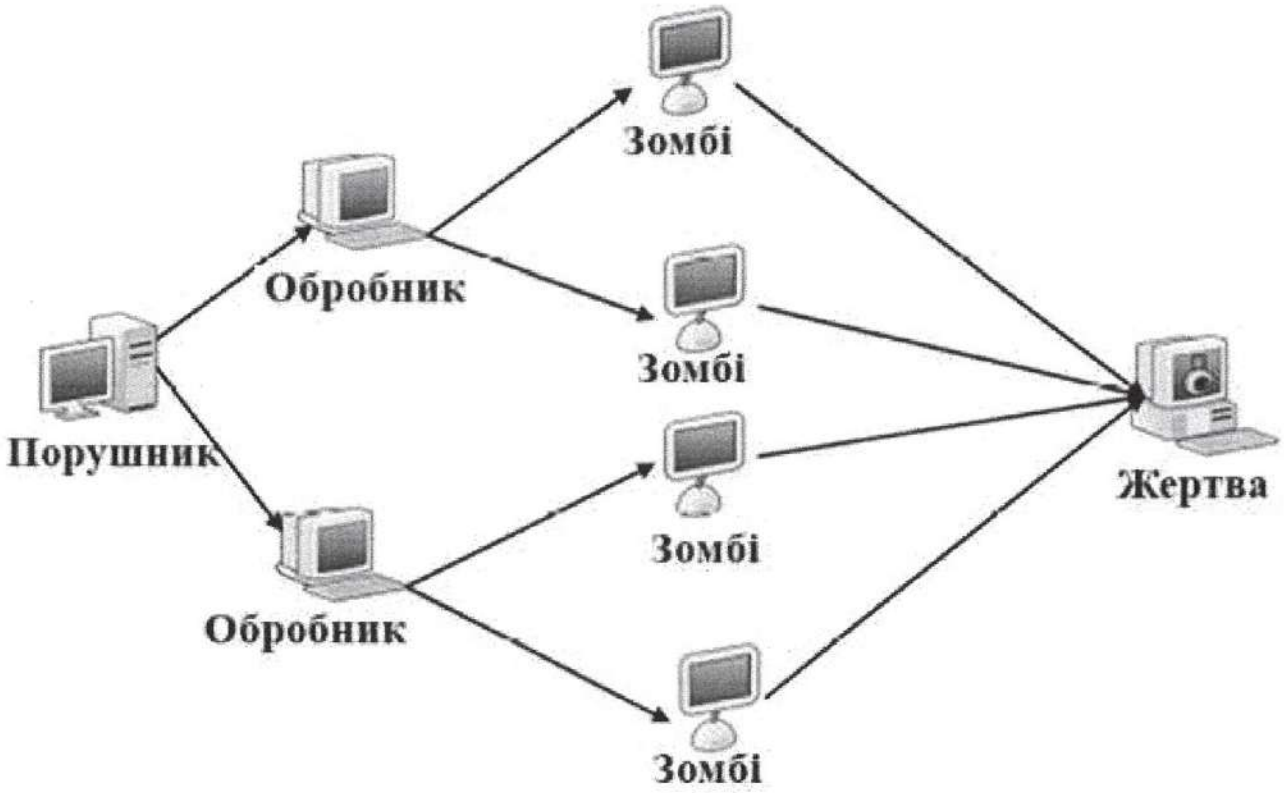


Рисунок 1 – Архітектура DDoS-атаки

TCP-флуд є різновидом DoS-атак, яка вичерпує ресурси цільового сервера шляхом надмірної кількості запитів на встановлення TCP-з'єднань (рис. 2). Це перевищує типову атаку TCP SYN-флуд, що в основному експлуатує процес трьохетапного рукостискання. У звичайному TCP-з'єднанні клієнт надсилає SYN-пакет, сервер відповідає SYN-ACK, а клієнт завершує процес, надсилаючи ACK. Проте при атаці TCP-флуд зловмисник може завершувати або частково завершувати численні з'єднання швидко і безперервно. Це змушує сервер постійно створювати нові записи у буфері черги підключень, очікуючи нормальної передачі даних. Буфер швидко заповнюється, що унеможливлює нові легітимні підключення. Унаслідок цього легітимні клієнти можуть стикатися із затримками або повною недоступністю сервісу, особливо якщо атака еволюціонує у DDoS із залученням багатьох систем. Розподілений характер таких атак ускладнює їх

нейтралізацію, оскільки трафік надходить із різних джерел, ідентифікація шкідливих запитів без впливу на нормальну роботу мережі стає складною [24].

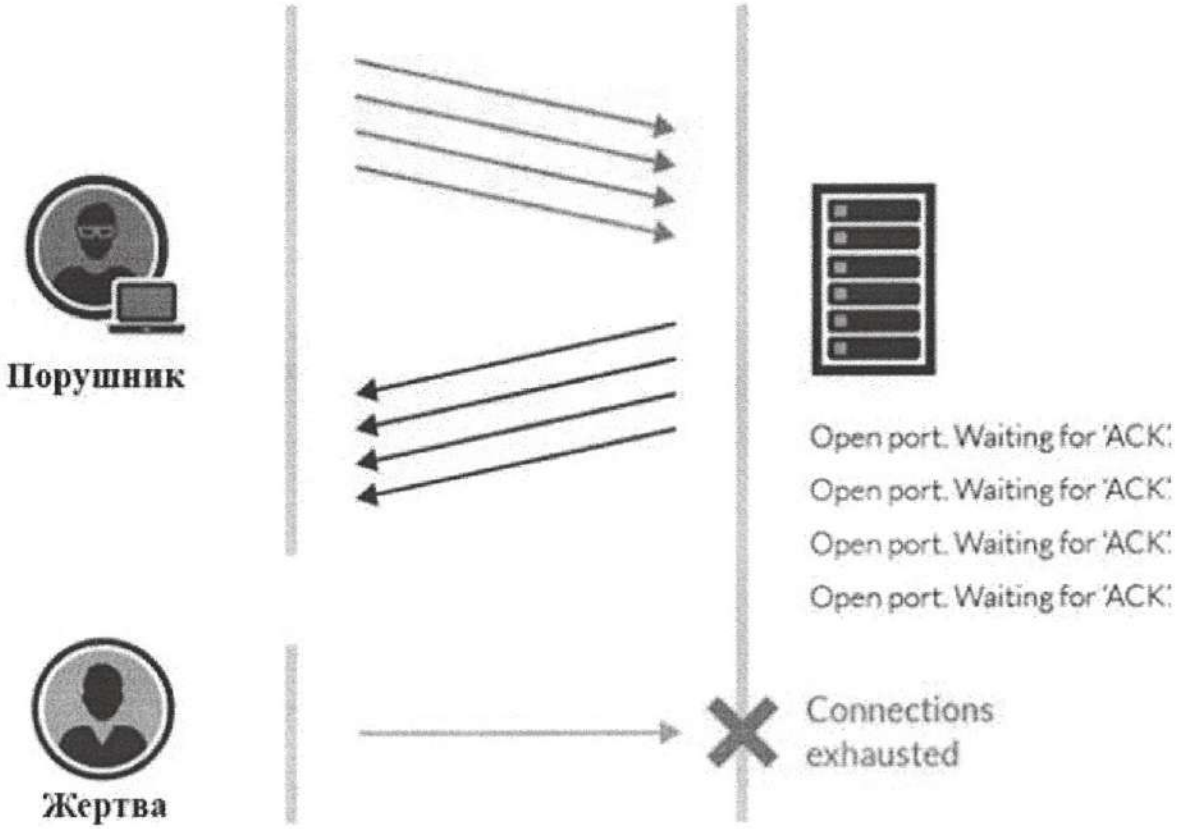


Рисунок 2 – Схема TCP SYN-флуд атаки

UDP-флуд – це тип DoS-атаки, яка перевантажує сервери великою кількістю UDP-пакетів (рис. 3). На відміну від TCP-атак, які вимагають встановлення з'єднання, UDP є протоколом без з'єднання, що дозволяє зловмиснику швидко надсилати значний обсяг пакетів на випадкові порти цільового сервера. Це змушує сервер перевіряти наявність додатка на цих портах і відповідати повідомленням ICMP «Destination Unreachable», якщо жоден додаток не знайдено. Потік вхідних пакетів і необхідність відповіді на кожен з них можуть вичерпати пропускну здатність мережі та ресурси сервера, що унеможливлює обробку легітимних запитів [25]. Простота створення UDP-пакетів робить цей тип атак особливо небезпечним і важким для відстеження, оскільки зловмисники можуть підробляти IP-адреси, приховуючи свою особу. Такі атаки можуть бути підсилені за допомогою відображаючих атак (Reflector DDoS), коли великий обсяг відповідей

надсилається на цільовий сервер через третіх сторін, що значно збільшує їхній вплив.

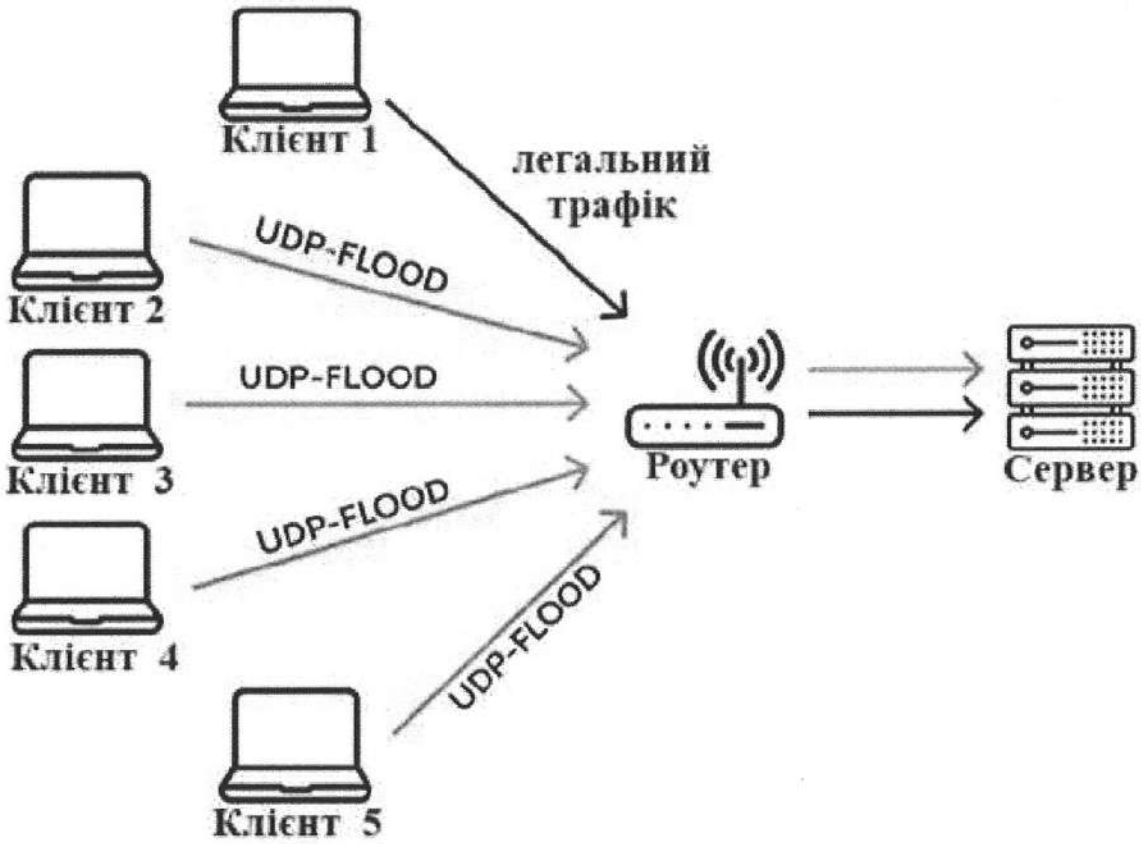


Рисунок 3 – Приклад UDP-флуд атаки

Атака на NRF (Crash NRF) спрямована на функцію Network Repository Function (NRF) у 5G-мережах. Вона використовує вразливості у free5GC — популярній програмній реалізації ядра 5G з відкритим кодом. Атака відправляє неправильно сформований HTTP GET-запит до NRF, у якому відсутні параметри requester-nf-type і target-nf-type. Вразливий NRF може завершити роботу через такий запит, спричиняючи відмову в обслуговуванні (DoS). Однак ця проблема була вирішена у версії free5GC 3.1.1 шляхом впровадження перевірки вхідних даних. Інцидент підкреслює важливість ретельної перевірки введення та регулярного оновлення систем безпеки для підтримки стабільності мережі.

Атака Fake AMF Delete спрямована на функцію управління доступом і мобільністю (Access and Mobility Management Function, AMF) у 5G-мережах, що використовують free5GC. У межах цієї атаки відправляється запит DELETE, який

видаляє AMF із репозиторію мережі за його ідентифікатором сесії, що може призвести до серйозних порушень роботи мережі, включно з відключенням важливих компонентів, які керують підключеннями та мобільністю користувачів. Ця атака демонструє необхідність суворих заходів безпеки для управління мережами 5G, аби запобігти несанкціонованим змінам і забезпечити стабільність роботи.

1.3 Аналіз наявних рішень виявлення DDoS-атак

Дослідження, представлене в статті [26], спрямоване на розробку методів виявлення DDoS-атак у радіодоступній мережі (RAN) на початкових етапах, що дозволяє протидіяти загрозам ще до того, як вони зможуть суттєво вплинути на мережеві ресурси. Основною метою роботи є створення інструментів для ідентифікації скомпрометованих користувацьких пристроїв (UE), які є частиною стільникових ботнетів. Запропонований підхід базується на аналізі статистики радіопротоколів, зокрема на вивченні атрибутів, отриманих із RAN, без необхідності аналізувати інкапсульовані пакети площини користувача, наприклад IP-пакети. Однією з особливостей методології є використання статистичних функцій, що дозволяють виявляти відмінності між шкідливим і легітимним трафіком. Наприклад, розмір протокольних даних радіолінії (RLC PDU) у вихідній лінії зв'язку демонструє зростання під час атаки, що стає важливим маркером аномалій. Автори підкреслюють, що їхній метод дозволяє ефективно виявляти шкідливий трафік на ранніх етапах, ще до того, як атака зможе перевантажити мережу.

У дослідженні [27] автори пропонують модель пом'якшення загроз для 5G-мереж (5GN-SMM), яка поєднує штучну нейронну мережу (Artificial Neural Network, ANN) із моделлю інтерпретативної структури (Interpretive Structure Modeling, ISM). Основна мета роботи – підвищити ефективність та точність виявлення та нейтралізації загроз у 5G-мережах, інтегруючи передові обчислювальні методи та систематичне моделювання. Однією з переваг

запропонованого підходу є глибокий аналіз загроз та їхніх взаємозв'язків. Розроблено гібридну методологію, яка використовує ANN для реального часу виявлення загроз та оцінки ризиків, а також ISM для аналізу взаємозалежностей між загрозами та вразливостями, створюючи структуровану модель для їхнього управління. Одна з ключових переваг моделі – це здатність ANN точно ідентифікувати потенційні загрози безпеці, тоді як ISM допомагає визначати пріоритети цих загроз і розуміти їхню взаємозалежність. Запропонована модель була протестована на основі реального кейсу, що продемонструвало її здатність краще справлятися із сучасними викликами, ніж традиційні методики. Водночас дослідження має і певні недоліки. Зокрема, модель потребує подальшого тестування на більш різноманітних сценаріях використання 5G-мереж, а також її адаптації до нових типів загроз, які можуть виникнути в майбутньому. Крім того, впровадження моделі може бути складним через необхідність високих обчислювальних ресурсів і підтримки спеціалізованої інфраструктури.

У [28] запропоновано інноваційну систему для створення моделей машинного навчання (МН), які здатні виявляти атаки в транспортних мережах 5G і наступного покоління (5GB). Основна ідея цієї системи полягає у забезпеченні безпечної та конфіденційної співпраці між вузлами «Транспорт – усе» (Vehicle-to-Everything, V2X), що є важливим для точного навчання моделей. Для цього структура поєднує федеративне навчання (Federated Learning, FL), блокчейн і смарт-контракти. Федеративне навчання дає змогу транспортним вузлам тренувати локальні МН-моделі, не передаючи необроблені дані, що зменшує ризик витоку інформації та підвищує конфіденційність. Завдяки інтеграції блокчейну й смарт-контрактів забезпечується прозорість, чесність та довіра між учасниками системи. Додатково структура включає алгоритм консенсусу з інтелектуальним механізмом винагород, який стимулює участь ефективних вузлів, що створюють високоточні локальні моделі. Це дозволяє підвищити загальну якість роботи системи та точність виявлення загроз. Серед переваг пропонованої системи можна виділити її здатність забезпечувати високий рівень конфіденційності даних завдяки федеративному навчанню, а також створення прозорої взаємодії між учасниками через блокчейн і смарт-контракти. Інтелектуальний вибір вузлів для участі в системі сприяє

формуванню точних моделей, а швидкий час досягнення консенсусу підвищує її ефективність. Однак система має й певні недоліки. Її впровадження є технічно складним, оскільки вимагає інтеграції декількох складних технологій, таких як FL, блокчейн і смарт-контракти. Крім того, розширення системи може викликати труднощі через потенційне ускладнення управління блокчейном і зростання часу консенсусу із збільшенням кількості учасників. Також використання блокчейну може призводити до підвищення енергоспоживання та обчислювальних витрат.

У [29] автори пропонують підхід до виявлення DDoS-атак у мережах 5G, зосереджуючись на використанні глибокого перенесеного навчання (Deep Transfer Learning, DTL). Основна ідея роботи полягає в тому, щоб покращити продуктивність моделей МН у виявленні атак завдяки перенесенню знань, отриманих під час навчання на одному наборі даних, до іншого набору даних із подібними, але обмеженими обсягами анотованого трафіку. Автори використовують тестове середовище 5G, де було створено набір даних із шістьма мільйонами потоків трафіку, які включають як легітимний, так і шкідливий трафік, зібраний із різних сегментів мережі. Цей великий набір даних став основою для тренування моделей глибокого навчання, таких як двонаправлена довготривала короткочасна пам'ять (BiLSTM), згорткові нейронні мережі (CNN), ResNet і Inception. Моделі, створені на основі цього джерела даних, далі були доопрацьовані за допомогою DTL для меншого цільового набору даних, 5G-NIDD, який містить обмежену кількість анотованого трафіку, створеного у реальній 5G-мережі. Запропонований підхід демонструє значне покращення в точності та F1-мірі під час виявлення атак. Крім того, використання сучасних моделей, таких як BiLSTM і CNN, дозволяє точно розпізнавати складні патерни атак у мережах 5G, які характеризуються великою кількістю підключених пристроїв і змінною структурою мережі. Однак метод також має певні недоліки. Зокрема, моделі CNN, хоча й демонструють високу ефективність, потребують більше часу для навчання порівняно з BiLSTM, що може бути критичним у реальних сценаріях із обмеженим часом на обробку. Крім того, ефективність моделей все ще суттєво залежить від якості та репрезентативності початкового набору даних, що ставить виклик перед розробниками щодо забезпечення доступу до різноманітних і релевантних даних.

Автори [30] пропонують рішення для виявлення атак типу Distributed Denial-of-Service (DDoS) в умовах мереж 5G, зокрема в контексті використання технології "network slicing", що дозволяє розділяти одну фізичну мережу на кілька віртуальних з різними рівнями ізоляції та якості обслуговування (QoS). У запропонованому підході використовується глибинне навчання (Deep Learning), що дозволяє виявляти атаки DDoS, а потім створювати спеціальну сегментацію мережі для ізоляції та пом'якшення наслідків атак. Це рішення має низку переваг, зокрема високу ефективність у виявленні атак та мінімальний вплив на продуктивність мережі для легітимних користувачів. Однак варто зазначити, що використання обмеженої кількості фізичних ресурсів для ізоляції атак може створювати додаткові навантаження на систему і вимагати точних налаштувань для оптимальної роботи.

У [31] йдеться про розробку гібридного методу виявлення та пом'якшення DDoS-атак у мережах 5G з використанням технологій програмно-визначених мереж (SDN) і машинного навчання. Автори пропонують поєднання статистичних методів і алгоритмів класифікації, таких як RepTree, Naive Bayes, RandomTree та інші, для динамічної оцінки мережевого трафіку та визначення аномалій. Основною особливістю запропонованого рішення є застосування динамічного порогу, що адаптується до змін мережевого середовища, дозволяючи швидко та з високою точністю виявляти атаки. Рішення не вимагає додаткового обладнання, що значно знижує витрати на впровадження. Крім того, використання динамічної конфігурації SDN дозволяє ефективно керувати мережевими ресурсами, забезпечуючи захист мережі в режимі, наближеному до реального часу. Однак, застосування динамічного порогу потребує ретельного налаштування, щоб уникнути помилкових блокувань легітимного трафіку. Крім того, процес обробки даних і навчання алгоритмів може створювати додаткове навантаження на систему, особливо у великих мережах із високим трафіком.

У [32] автори пропонують підхід, заснований на федеративному навчанні, для виявлення DDoS-атак на протокол GTP у мережах 5G. Модель використовує колективний інтелект кількох пристроїв, що дозволяє ефективно та конфіденційно виявляти атаки. Крім того, вони розробили тестову архітектуру мережі 5G, яка

імітує складну публічну мережу. Це дає змогу оцінювати безпекові застосунки на основі штучного інтелекту, а також тестувати реалізацію та впровадження запропонованої моделі. Результати експериментів показують, що запропонована модель федеративного навчання, працюючи без нагляду, ефективно виявляє DDoS-атаки в мережі 5G, при цьому зберігаючи конфіденційність даних кожного пристрою. Це підкреслює перспективи використання федеративного навчання для підвищення рівня безпеки в мережах 5G і майбутніх технологіях. Основною перевагою підходу є його здатність зберігати конфіденційність даних та ефективно адаптуватися до нових загроз. Водночас недоліком може бути потреба у високих обчислювальних ресурсах і складність координації між пристроями в реальних умовах мережі.

У [33] автори акцентують увагу на проблемі стрімкого зростання обсягу трафіку у бездротових мережах, зумовленого розвитком технологій та збільшенням кількості пристроїв, які працюють у межах однієї інфраструктури. Метою роботи є розробка моделі виявлення DDoS-атак у мережах 5G, яка складається з двох фаз: вилучення ознак і виявлення атак. У моделі для виконання завдання виявлення використовуються класифікатори Long Short-Term Memory (LSTM) та Recurrent Neural Network (RNN). Для оптимізації ваг класифікатора RNN та покращення точності моделі застосовано алгоритм оптимізації Seagull на основі протилежного навчання (OLSOA).

У роботі [34] автори пропонують використання глибинного навчання для виявлення і нейтралізації таких атак. Зокрема, вони застосували два алгоритми глибинного навчання — Convolutional Neural Network (CNN) і Feed Forward Neural Network (FNN) — для аналізу спеціально створеного набору даних, що моделює роботу IoT пристроїв у мережах 5G. Для побудови інфраструктури 5G було використано середовище моделювання OMNeT++ із фреймворками INET і Simu5G. Основною перевагою цього підходу є його здатність ефективно підвищувати рівень безпеки IoT пристроїв у межах мереж 5G завдяки високій точності аналізу. Водночас, недоліками можуть бути вимоги до значних обчислювальних ресурсів для роботи моделей глибинного навчання, а також залежність від якісного набору даних для навчання алгоритмів.

Дослідження [35] спрямоване на захист мереж 5G від сигнальних DDoS-атак, які націлені на ключові функції ядра мережі (CN), такі як AMF, SMF та UPF. Запропонований метод базується на МН із використанням аналізу на основі ентропії (EBA) та статистики (SBA), що дозволяє проактивно виявляти аномалії в мережевому трафіку. Розроблена система включає збір даних із ключових інтерфейсів протоколу 5G CN, попередню обробку інформації, вибір ознак, класифікацію за допомогою МН-алгоритмів (таких як випадковий ліс, SVM і Байєсівська класифікація) та оцінку їхньої продуктивності. Використання EBA дозволяє визначати нерівномірний розподіл даних, а SBA порівнює статистичні розподіли для виявлення аномалій, що виникають під час DDoS-атак. Структура системи передбачає використання Apache Kafka і Redis для швидкої обробки потоків даних та зберігання інформації, що сприяє оперативному виявленню атак. Основними перевагами підходу є висока точність виявлення, здатність працювати в реальному часі та зменшення негативного впливу атак на функції мережі. Недоліками є залежність від якісних даних для навчання, висока обчислювальна складність та необхідність інтеграції системи в складну інфраструктуру мереж 5G.

У [36] автори пропонують підхід, який базується на використанні методів виявлення вторгнень (IDS) для аналізу трафіку, профілювання поведінки та впровадження адаптивних механізмів реагування. У дослідженні розглянуто традиційні методи IDS, зокрема підходи, засновані на сигнатурах і аномаліях, а також техніки машинного навчання. Основна увага приділяється розробці рішень, що враховують складність і специфіку мереж 5G, як-от великий обсяг даних і високі вимоги до оперативності виявлення загроз. Запропоновані стратегії включають аналіз трафіку в реальному часі, що дозволяє швидко виявляти аномалії, та адаптивні механізми, які автоматично реагують на загрози. Основною перевагою підходу є його здатність забезпечувати високу точність і оперативність виявлення DDoS-атак у складних умовах 5G, мінімізуючи негативний вплив на мережу. Проте недоліками можуть бути підвищені вимоги до обчислювальних ресурсів і залежність від складності впровадження таких систем у масштабованих інфраструктурах. Автори підкреслюють, що підвищення стійкості 5G до нових

кіберзагроз є важливим кроком для забезпечення надійності цих мереж у майбутньому.

Запропоновані методи захисту мереж 5G від DDoS-атак демонструють значний прогрес у сфері кібербезпеки, адже використання машинного та глибокого навчання, блокчейну, федеративного навчання та адаптивних механізмів моніторингу дозволяє ефективно виявляти та пом'якшувати загрози. Завдяки цим підходам забезпечується висока точність аналізу трафіку, що зменшує кількість помилкових спрацьовувань і дозволяє вчасно ідентифікувати аномалії. Використання адаптивних алгоритмів, таких як ANN-ISM або DTL, допомагає системам безпеки швидко пристосовуватися до нових загроз, що особливо важливо в динамічному середовищі 5G.

Проте впровадження таких рішень супроводжується певними труднощами. Зокрема, більшість запропонованих методів тестувалися в контрольованих середовищах, тому їхня ефективність у реальних 5G-мережах може суттєво відрізнятися. Також існує ризик зниження ефективності алгоритмів із подальшим зростанням кількості підключених пристроїв та загального обсягу трафіку.

1.4 Постановка задачі

Метою кваліфікаційної роботи розробка ефективної системи виявлення DDoS-атаки в мережах 5G, що базується на сучасних підходах машинного навчання. Для досягнення цієї мети необхідно наступне завдання:

– створити та обробити набір даних, що містить реалістичний мережевий трафік 5G із врахуванням атакуючих і легітимних сценаріїв. Для цього варто змоделювати як легітимну взаємодію користувачів із мережею, так і реалізацію різних типів атак, включаючи UDP Flood, TCP SYN Flood, HTTP Flood та інші. Зібраний трафік потрібно перетворено в CSV із виділенням формату основних характеристик, які можуть ідентифікувати атаки. Додатково провести фільтрацію та нормалізацію даних для підвищення ефективності моделей машинного навчання;

– розробити алгоритм роботи та систему виявлення DDoS-атаки в мережах 5G;

– розробити та навчити моделі машинного навчання для точного та швидкого виявлення атак. Основну увагу варто приділити використанню згорткових нейронних мереж (CNN) та двонаправленої довготривалої короткочасної пам'яті (BLSTM). Слід передбачити тестування різних архітектурних моделей, оптимізацію гіперпараметрів та порівняння їхньої продуктивності на основі метрики, таких як точність, повнота, F1-міра;

– провести оцінку ефективності запропонованої системи шляхом її тестування в реальному середовищі. Важливим етапом є перевірка працездатності розроблених алгоритмів у реальних умовах роботи 5G-мереж. Для цього потрібно буде провести тестування в симуляційному середовищі, що імітує поведінку реальної інфраструктури, включаючи змінний трафік, мобільність користувачів та вплив на атаку. План застосування рівня помилкових спрацьовувань, час реагування системи та ефективність класифікації атакуючого трафіку. Аналіз отриманих результатів дозволить зменшити переваги та обмеження розробленої системи, а також надати рекомендації щодо її подальшого вдосконалення.

2 АЛГОРИТМ РОБОТИ СИСТЕМИ ВИЯВЛЕННЯ DDOS-АТАК

2.1 Застосування машинного навчання для виявлення атак

Алгоритми МН дозволяють системам самостійно навчатися на основі шаблонів у даних та приймати обґрунтовані рішення або робити прогнози без необхідності явного програмування. Ці алгоритми варіюються від простих методів, таких як лінійна регресія, до складніших ансамблевих підходів, які підвищують точність прогнозів і адаптивність. Завдяки своїй різноманітності алгоритми МН можуть легко інтегруватися в різні сфери, покращуючи їхню надійність і ефективність у реальних сценаріях.

Глибоке навчання, спеціалізована галузь МН, використовує глибокі нейронні мережі для обробки складних даних. Ці мережі складаються з багатошарових структур вузлів або нейронів, чудово моделюють складні, нелінійні взаємозв'язки у великих масивах даних, що дозволяє глибокому навчанню досягати безпрецедентної точності, особливо в таких галузях, як обробка природної мови та комп'ютерний зір.

Ефективне управління даними є значущим фактором у роботі з інформацією. Воно включає збір, перевірку, очищення та трансформацію даних, забезпечуючи їхню точність, повноту та аналітичну придатність. Дані етапи важливі для якості даних, які надходять до моделей МН, сприяючи точнішим аналізам і прогнозам.

Підготовка даних включає наступні етапи:

– видалення даних включає видалення викидів або виправлення помилок;

– нормалізація включає масштабування вхідних характеристик для забезпечення рівного внеску кожної з них;

– створення нових змінних на основі наявних даних для підвищення продуктивності моделі.

Вказані процеси є необхідними для оптимізації роботи алгоритмів і досягнення успіху у впровадженні МН.

Процес створення моделей МН починається з вибору відповідної моделі залежно від природи завдання, складності даних і бажаного результату. Цей вибір може варіюватися від простих дерев рішень до складних нейронних мереж. Після

вибору моделі є оптимізація гіперпараметрів. Методи, як-от grid search (перебір параметрів), випадковий пошук або байєсівська оптимізація, використовуються для пошуку оптимальних налаштувань, які підвищують ефективність моделі. Моделі тренуються на визначених наборах даних, а їх продуктивність перевіряється на окремих тестових наборах, щоб уникнути перенавчання та забезпечити здатність до узагальнення на нових, невідомих даних.

Застосування машинного та глибокого навчання охоплює широкий спектр галузей:

– інтеграція МН із багатомасштабним моделюванням відкрила нові можливості для аналізу складних систем у різних наукових сферах, дозволяючи передбачати поведінку на різних рівнях;

– у сфері дистанційного зондування глибоке навчання революціонізувало аналіз великих обсягів геопросторових даних, покращуючи моніторинг та управління довкіллям;

– у медицині ці технології значно вдосконалили діагностику, персоналізовану медицину та управління пацієнтами, підвищуючи як результати лікування, так і ефективність роботи.

Техніки МН поділяються на три основні категорії:

– кероване навчання (Supervised Learning, навчання з вчителем) передбачає тренування моделі на основі позначених даних, де кожна точка даних має відповідну мітку. Модель навчається встановлювати відповідність між вхідними та вихідними даними, що дозволяє робити прогнози для нових, невідомих даних;

– некероване навчання (Unsupervised Learning, навчання без вчителя) працює з непозначеними даними. Метою є моделювання внутрішньої структури чи розподілу даних, щоб краще їх зрозуміти;

– навчання з підкріпленням (Reinforcement Learning), де агент навчається взаємодіяти з середовищем шляхом виконання дій і отримання винагород чи штрафів за ці дії.

Моделі класифікації зосереджуються на прогнозуванні категоріальних міток. Вони призначені для завдань, де результат є дискретною категорією. Наприклад, модель класифікації може визначати, чи є електронний лист спамом, або

розпізнавати цифру на рукописному зображенні. Основні техніки класифікації включають:

– Random Forest (випадковий ліс) - метод ансамблевого навчання, який будує велику кількість дерев рішень під час тренування і виводить клас, що є найбільш частим (для класифікації), або середнє значення (для регресії) передбачень окремих дерев. Random Forest допомагає запобігти перенавчанню, властивому окремим деревам рішень [37];

– Support Vector Machine (SVM, метод опорних векторів) - клас алгоритмів для задач класифікації та регресії. У задачах класифікації SVM знаходить гіперплощину в багатовимірному просторі, яка максимально розділяє класи даних. Мета — знайти площину з максимальною відстанню між точками обох класів, що забезпечує точну класифікацію [38];

– XGBoost (екстремальне градієнтне підсилення) - реалізація градієнтного підсилення дерев рішень, оптимізована для швидкості та продуктивності. XGBoost підходить як для класифікації, так і для регресії, здатний працювати з пропущеними даними, що робить його надійним інструментом для складних наборів даних [39].

Ансамблеві моделі комбінують кілька базових моделей для створення оптимальної моделі. Вони покликані використовувати сильні сторони окремих моделей, одночасно зменшуючи їхні недоліки. До поширених ансамблевих методів належать Bagging, Boosting та Stacking. Bagging тренує кілька моделей на різних підмножинах даних і усереднює їхні прогнози, щоб зменшити варіативність і уникнути перенавчання. Boosting покроково виправляє помилки попередніх моделей, що покращує загальну продуктивність. Stacking комбінує різні моделі за допомогою метамоделі (наприклад, лінійної регресії), щоб отримати кращі результати. Однак ця техніка вимагає ретельного налаштування, щоб уникнути перенавчання.

Дослідження показують, що ансамблеві методи суттєво підвищують точність і надійність у різних сферах, зокрема машинному навчанні, прогнозуванні погоди та аналізі медичних даних. Хоча такі моделі пропонують високу продуктивність, вони також потребують значних обчислювальних ресурсів.

У нейронній мережі різні типи шарів виконують важливі функції обробки даних і витягування ознак:

– Dense Layer повністю пов'язує кожен нейрон одного шару з усіма нейронами наступного, важливий для вивчення складних шаблонів шляхом налаштування ваг під час тренування;

– Dropout Layer запобігає перенавчанню, випадково відключаючи частину нейронів у процесі тренування, що покращує здатність моделі до узагальнення;

– Embedding Layer перетворює категоріальні дані в щільні вектори, що дозволяє мережі ефективніше вивчати зв'язки між категоріями;

– GlobalMaxPooling1D зменшує розміри просторових даних, обираючи максимальне значення в часовому вимірі, використовується в задачах класифікації тексту для вилучення найважливіших ознак;

– SpatialDropout1D це техніка регуляризації для однорозмірних характеристик, зокрема послідовностей, яка покращує узагальнювальні можливості мережі.

Комбінація цих шарів, а також функцій активації та специфічних архітектур, забезпечує здатність мереж навчатися, робити прогнози та ефективно працювати із новими даними.

Архітектури нейронних мереж охоплюють широкий спектр моделей, кожна з яких розроблена для вирішення певних завдань і оптимізована під специфічні типи даних. Зокрема, двонаправлена довготривала короткочасна пам'ять (Bidirectional Long Short-Term Memory, BiLSTM) є типом рекурентної нейронної мережі, яка здатна вивчати залежності у вхідних послідовностях у двох напрямках — як у прямому, так і в зворотному. Це робить BiLSTM корисною для завдань, які вимагають урахування контексту, таких як обробка природної мови, розпізнавання мовлення або аналіз часових рядів. З іншого боку, згорткові нейронні мережі (Convolutional Neural Networks, CNN) найбільше використовуються для обробки візуальних даних, таких як зображення та відео. Завдяки механізмам згорток і пулінгу CNN ефективно витягують просторові ознаки, що робить їх також корисними в прогнозуванні мережевого трафіку. Окрему увагу привертає модель двонаправленого подання кодувальника з трансформерів (Bidirectional Encoder

Representations from Transformers, BERT), яка базується на механізмі трансформерів і здатна аналізувати контекст слів у двох напрямках, що значно покращує її здатність розуміти природну мову. Модель BERT та її варіанти, такі як CyBERT, показали високу ефективність у різних завданнях, включаючи забезпечення кібербезпеки, класифікацію мережевого трафіку та виявлення загроз. Завдяки механізму уваги, вбудованому в трансформери, ця архітектура активно застосовується для виявлення аномалій, аналізу мережевих вторгнень та вирішення інших складних задач.

Важливим аспектом розвитку нейронних мереж є поява архітектур глибокого навчання, які інтегрують переваги CNN, глибоких нейронних мереж (Deep Neural Networks, DNN) і рекурентних нейронних мереж (Recurrent Neural Networks, RNN). Наприклад, LSTM і BiLSTM виявилися ефективними в завданнях передбачення категорій Загального регламенту захисту даних (General Data Protection Regulation, GDPR) у політиках конфіденційності або розпізнавання іменованих сутностей. Прогрес у некерованому навчанні словникових вбудовувань, а також вдосконалення алгоритмів тренування значно знизили потребу в ручному проектуванні ознак, що сприяло появі більш гнучких і точних моделей. Це свідчить про те, що розвиток нейронних мереж продовжує відкривати нові можливості для аналізу даних і вирішення складних завдань.

Операції з моделями МН є складним і багатогранним процесом, що включає кілька етапів, серед яких компіляція, навчання та оцінювання моделей. Ці етапи тісно пов'язані між собою та формують основу для створення ефективних систем МН, які здатні вирішувати реальні завдання.

На етапі компіляції моделі відбувається її початкове налаштування відповідно до специфіки завдання. Це включає вибір таких важливих компонентів, як функції втрат, оптимізатори та метрики. Функції втрат відіграють критичну роль у навчанні, оскільки вони визначають, наскільки добре модель прогнозує результати порівняно з реальними даними. Наприклад, середньоквадратична похибка часто використовується для задач регресії, а крос-ентропія підходить для класифікаційних задач, дозволяючи моделі зосереджуватися на мінімізації помилок. Оптимізатори, такі як стохастичний градієнтний спуск, Adam або

RMSprop, керують процесом коригування параметрів моделі, щоб вона поступово покращувала свої прогнози. Метрики, у свою чергу, забезпечують кількісну оцінку продуктивності моделі, вимірюючи такі показники, як точність, повнота або F1-міра, що дає можливість об'єктивно аналізувати її результати.

Навчання моделі є наступним етапом. У цьому процесі модель ітераціями обробляє тренувальні дані, коригуючи свої параметри для мінімізації втрат. Цей процес зазвичай триває кілька епох, протягом яких модель поступово вдосконалює свої передбачення та зменшує помилки. Щоб уникнути перенавчання та налаштувати гіперпараметри, навчання перевіряється на валідаційному наборі даних, який не використовується безпосередньо під час тренування. Це дозволяє збалансувати складність моделі та її здатність до узагальнення, забезпечуючи адекватну продуктивність у реальних умовах.

Завершальним етапом є оцінювання моделі, що дозволяє визначити, наскільки добре вона впорається із завданням у реальних умовах. Для цього модель тестується на наборі даних, який раніше не використовувався в навчанні чи валідації. Порівнюючи прогнози моделі з реальними результатами, визначаються її здатність до узагальнення та готовність до практичного використання. Для оцінювання використовуються метрики, визначені під час компіляції, зокрема точність, площа під кривою ROC, F1-міра та інші, які дають повне уявлення про сильні сторони й обмеження моделі.

Таким чином, ці три етапи — компіляція, навчання та оцінювання — утворюють єдиний, злагоджений процес, що спрямований на створення моделей, здатних не лише добре навчатися, а й демонструвати стабільну продуктивність у реальних задачах. Вони не лише забезпечують якісне налаштування моделей, але й дозволяють краще зрозуміти їхню ефективність, межі застосування та потенціал для подальшого вдосконалення.

2.2 Створення набору даних для навчання CNN та BLSTM моделей

Існує кілька загальнодоступних наборів даних, призначених для аналізу DDoS-атак, зокрема CICDoS, CICIDS2017, CSECICIDS2018 та CIC-DDoS2019. Однак жоден із них не був створений спеціально для зрізів мережі 5G. Тому було реалізовано інфраструктуру за допомогою симуляторів Free5GC та UERANSIM, що дало змогу створити нові набори даних, орієнтовані саме на дослідження мережевих даних 5G.

У процесі створення наборів даних для аналізу безпеки мереж 5G було використано локальну інфраструктуру на основі симуляторів Free5GC та UERANSIM, що дозволило згенерувати реалістичний мережевий трафік як у безпечних умовах, так і під час атак DoS/DDoS. Дослідження включало кілька етапів, у ході яких було реалізовано різні сценарії взаємодії користувацького обладнання із серверами через функції користувацької площини (UPF). Основною метою було створення масштабованих, реалістичних та різноманітних наборів даних, які відображали б поведінку мережевого трафіку в умовах реального використання та під час цілеспрямованих атак.

Перший етап дослідження передбачав створення початкового набору даних, де мережевий трафік містив як звичайну діяльність користувачів, так і DoS/DDoS-атаки. Для цього генерувалися запити ICMP, що дозволяли оцінити з'єднання та час відгуку між пристроями. Також здійснювався обмін даними між клієнтськими пристроями й серверами, включаючи ACK-сканування, TCP SYN-атаки, UDP-сканування, аналіз брандмауерів та часових міток, SYN-сканування, а також FIN, PUSH і URG-сканування. Крім того, оцінювалися показники продуктивності електронної пошти, для чого було розгорнуто поштовий сервер, що дозволило дослідити механізми передачі та обробки електронних повідомлень у середовищі 5G. Для моделювання атак використовувався потужний інструмент hping3, за допомогою якого виконувалися атаки з використанням UDP-флудингу, TCP PUSH, TCP FIN і TCP RST. Ці атаки спрямовувалися на поштовий сервер, що дало змогу оцінити його стійкість до навантажень та ефективність обробки атакованого трафіку. У результаті експериментів було отримано детальні дані про поведінку мережевого трафіку як у нормальних умовах, так і під час атаки.

Наступний етап передбачав створення вдосконаленого набору даних, що

базувався на тій самій інфраструктурі, але з розширеними можливостями. Було додано додаткову базову станцію gNodeB, а кількість користувацьких пристроїв збільшилася до восьми, що дозволило оцінити вплив розширеної топології на продуктивність мережі. Під час експерименту атаки DoS/DDoS здійснювалися через різні UPF, а для імітації безпечного трафіку застосовувався спеціально розроблений автоматизований скрипт на основі headless-браузера. Він виконував перегляд вебсайтів, імітуючи реальну активність користувачів, що дозволило отримати реалістичні шаблони мережевого трафіку. Завдяки headless-браузеру здійснювався автоматичний перегляд 500 популярних вебресурсів, що дало змогу отримати репрезентативний зразок вебконтенту та мережевих комунікацій.

Третій етап експерименту стосувався створення наступного набору даних, який відрізнявся від попередніх тим, що базувався не на віртуальних машинах, а на контейнерній архітектурі. Використання контейнеризації дозволило більш точно імітувати функціональність мережевих компонентів зрізу 5G. У середовищі Docker створювалися ізольовані контейнери, що відповідали елементам мереж. Завдяки цьому було змодельовано складну інфраструктуру, яка забезпечувала автентичні умови для збору даних про мережевий трафік. Під час цього етапу знову використовувався автоматизований headless-браузер, що імітував вебперегляд реальними користувачами. Крім перегляду вебсайтів, до сценаріїв безпечного трафіку було додано перегляд відео на YouTube, пряму трансляцію, роботу з файлами (завантаження, копіювання, вставку та видалення), SSH-з'єднання та ICMP-пінгування. Щодо атакуючого трафіку, то до вже відомих методів (UDP-флудинг, TCP SYN-атака, TCP PUSH, TCP FIN, TCP RST) додалися TCP ACK, TCP URG, TCP XMAS і TCP YMAS, що зробило експеримент ще більш комплексним та наближеним до реальних загроз у мережах 5G.

Для моніторингу та запису трафіку використовувався аналізатор мережевих протоколів Wireshark, який дозволив захоплювати та зберігати пакети у форматі .pcap. Це дало можливість детально досліджувати отримані дані та робити висновки щодо ефективності роботи мережі як у стандартних умовах, так і під час атак.

У результаті проведених експериментів вдалося створити три повноцінні

набори даних, які містять інформацію про реальний трафік у мережах 5G. Вони охоплюють широкий спектр сценаріїв використання, включаючи звичайну активність користувачів, перегляд вебсайтів, роботу з файлами, перегляд відео та атаки DoS/DDoS. Ці набори даних стали цінним ресурсом для подальших досліджень у сфері безпеки 5G, дозволяючи аналізувати різні методи атак, розробляти нові механізми захисту та оцінювати ефективність існуючих систем виявлення загроз. Захоплені файли .pcap були перетворені у формат .csv, що містив 84 характеристики мережевого трафіку. Для ідентифікації DDoS-атак було відібрано основні характеристики.

Для візуалізації трафіку використовувався метод Radviz, який дозволяє аналізувати та відображати зв'язки між різними параметрами, розміщуючи їх на осях, що виходять із центру. Такий підхід допоміг виявити закономірності та особливості трафіку, що дало змогу краще зрозуміти характер DDoS-атак у мережевих даних.

Серед характеристик, виділених інструментом CICFlowMeter, були:

– тривалість потоку (Flow Duration);

– загальна кількість пакетів у прямому напрямку (Total Fwd Packet);

– загальна кількість пакетів у зворотному напрямку (Total Bwd Packet);

– загальний розмір пакетів у прямому напрямку (Total Length of Fwd Packet);

– загальний розмір пакетів у зворотному напрямку (Total Length of Bwd Packet);

– мінімальне, максимальне, середнє та стандартне відхилення розміру пакетів у прямому та зворотному напрямках.

З 84 характеристик, отриманих шляхом перетворення файлів .pcap у формат .csv за допомогою CICFlowMeter, було відібрано вісім найбільш значущих. У результаті залишилися наступні характеристики: тривалість потоку (Flow Duration), стандартне відхилення довжини пакетів у прямому напрямку (Fwd Packet Length Std), кількість ACK-пакетів (ACK Flag Count), мережевий протокол (Protocol), загальна кількість пакетів у прямому напрямку (Total Fwd Packet) та мінімальний розмір сегмента у прямому напрямку (MinSegSizeForward). Подальший аналіз показав, що значення параметра MinSegSizeForward у всіх .csv-

файлах було рівним «0», що зробило цю характеристику неінформативною. Через це вона також була виключена зі списку.

Для перевірки найбільш значущих характеристик було застосовано метод головних компонент (Principal Component Analysis, PCA) для другого набору даних. Використовуючи бібліотеку sklearn у Python, було обчислено 10 характеристик з найвищою дисперсією. Найбільш вагомими характеристиками виявилися: Protocol – 18.45%, Flow Duration – 14.12%, Total Forward Packets – 11.92%, Total Backward Packets – 9.54%, Total Length of Forward Packets – 4.65%, Total Length of Backward Packets – 3.63%.

Спочатку для моделі було відібрано шість характеристик на основі PCA, а також раніше вибрану Fwd Packet Length Std. Однак додавання інших характеристик зменшувало точність моделі при валідації та тестуванні. У результаті було прийнято рішення залишити лише сім найбільш ефективних характеристик. Щоб оцінити ефективність характеристик, було проведено порівняльний аналіз із відомими наборами даних.

CIC-DDoS2019 – широко використовуваний набір даних для аналізу DDoS-атак, що включає UDP Flood, TCP SYN Flood, HTTP Flood, ICMP Flood, SYN-ACK Flood та інші методи атаки. Цей набір дозволяє дослідникам аналізувати закономірності атак та розробляти стратегії їх виявлення й запобігання.

5G-ND – набір даних, отриманий у тестовій 5G-мережі. Містить інформацію про DoS-атаки (ICMP Flood, UDP Flood, SYN Flood, HTTP Flood, Slowrate DoS) і сканування портів (SYN Scan, TCP Connect Scan, UDP Scan).

Виділені характеристики атаки й безпечного трафіку були візуалізовані за допомогою box plot, що дозволило оцінити розподіл значень у різних наборах даних:

– тривалість потоку (Flow Duration) аналізувалася протягом 1000 секунд, причому однаковий розподіл спостерігався як у трафіку атак, так і у безпечному трафіку;

– загальна кількість переданих пакетів у прямому (Total Forward Packets) та зворотному (Total Backward Packets) напрямках була представлена у байтах з відображенням даних з викидами та без них;

– сумарна кількість переданих пакетів (Total Forward and Backward Packets) продемонструвала загальний обсяг трафіку;

– середня довжина пакета (Total Packet Length Mean) вказувала на характерні розміри переданих даних.

Аналіз box plot показав, що згенеровані набори даних мають схожі розподіли з іншими існуючими наборами. Це ускладнює використання моделей, що базуються на правилах, для виявлення атак, оскільки трафік, що атакує, і безпечний трафік мають значні накладання. Виявлення та точна класифікація атак у таких умовах вимагає застосування більш складних методів машинного навчання.

Таблиця 2.1 демонструє кількість трафіку в експериментальних наборах даних.

Таблиця 2.1 - Кількість трафіку в експериментальних наборах даних

Набір даних	Зловмисний трафік	Безпечний трафік	Загальний трафік
Dataset-1	20,000	50,000	70,000
Dataset-2	5,000,000	1,000,000	6,000,000
Dataset-3	5,000,000	1,000,000	6,000,000
CIC-DDoS2019	1,000,000	1,000,000	2,000,000
5G-NIDD	300,000	15,000	315,000
5GAD-2022	Не враховувався	84,000	84,000

Значна схожість розподілу характеристик у зловмисному та безпечному трафіку свідчить про складність класифікації атак на основі стандартних методів. Це підкреслює необхідність застосування вдосконалених підходів, зокрема глибокого навчання, для ефективного виявлення нових типів атак у мережах 5G.

2.3 Навчання CNN та BLSTM моделей

На рисунку 4 продемонстровано алгоритм навчання CNN та BLSTM моделей.

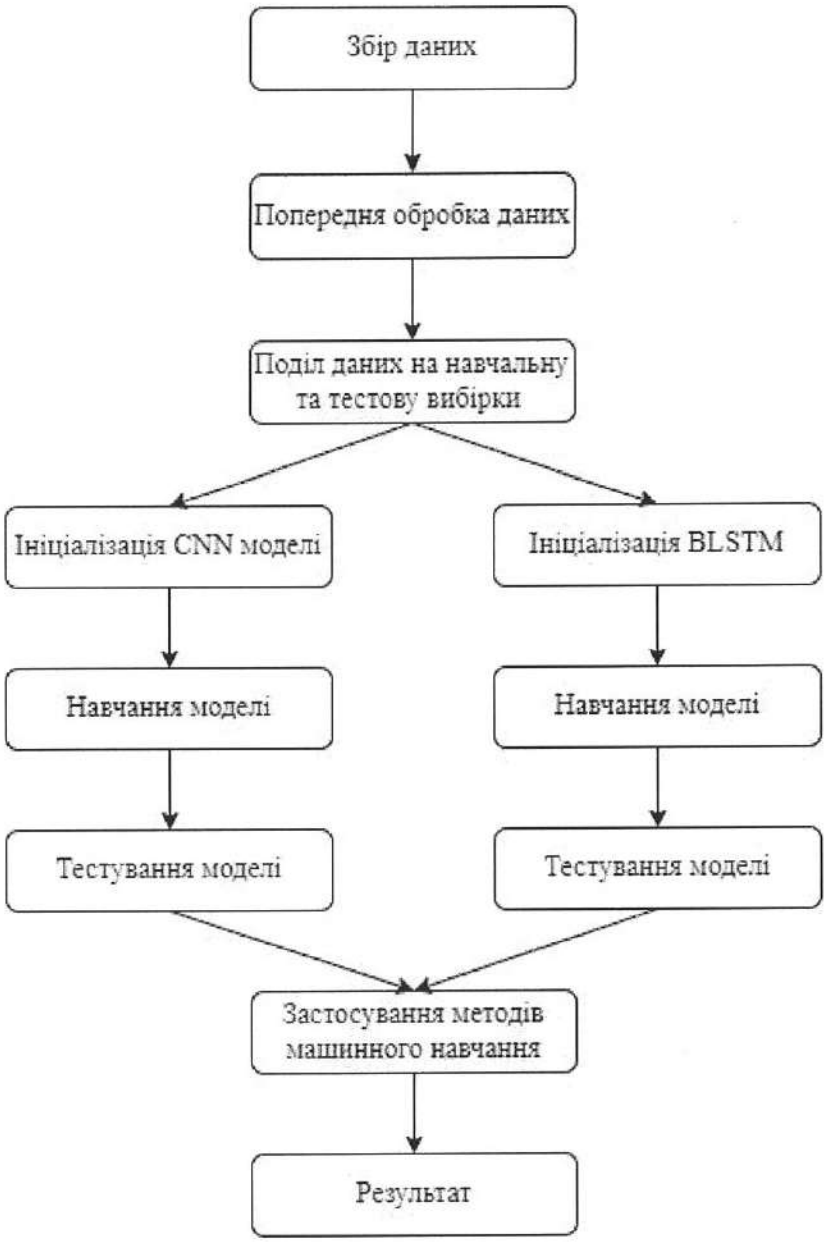


Рисунок 4 – Етапи налаштування МН для прийняття рішення

На початковому етапі було зібрано набір даних, що містить інформацію про мережевий трафік для виявлення зловмисної активності. Процес формування набору даних описано у параграфі 2.2.

На етапі попередньої обробки даних виконувалося попереднє опрацювання вихідного набору даних, щоб зробити його придатним для використання в алгоритмах МН. Для навчання та оцінки моделей МН дані було поділено на навчальну та тестову вибірки. Було використано співвідношення 80/20, де 80% даних застосовувалися для навчання моделей, а 20% – для тестування. Було обрано два алгоритми машинного навчання для порівняльного аналізу: CNN та BLSTM. Кожен із цих алгоритмів має свої переваги залежно від специфіки набору даних. На етапі навчання та тестування моделей кожна модель навчалася на тренувальній

вибірці та перевірялася на тестовій вибірці. Це дозволило моделям навчитися розпізнавати закономірності в даних та робити узагальнені прогнози на нових зразках трафіку. Останнім кроком є використання протестованих моделей для прийняття рішення при класифікації трафіку.

Загалом модель бінарної класифікації DoS/DDoS складається з чотирьох основних етапів: введення вхідних даних (мережевих потоків), вибору характеристик, тренування та класифікації трафіку як безпечного або зловмисного. Кожен вхідний елемент містить один мережевий потік. Для тренування та валідації моделей CNN та BLSTM було відібрано сім характеристик. Після навчання модель проходила тестування на різних зразках трафіку для визначення атак. Загальна схема процесу класифікації представлена на рисунку 5.

Перший етап передбачав збір мережевих потоків, які є послідовністю пакетів, що передаються між мережевими вузлами. Ці потоки містять інформацію про особливості взаємодії та комунікаційні шаблони трафіку. Щоб створити ефективну модель класифікації DoS/DDoS, було застосовано методи відбору ознак для виявлення найбільш значущих характеристик. Було відібрано сім параметрів, що дозволяють найкраще розрізняти безпечний та атакувальний трафік. Відібрані характеристики використовувалися для навчання та валідації моделі класифікації, яка включала CNN та BLSTM. CNN дозволяє аналізувати просторові залежності між характеристиками трафіку, тоді як BLSTM допомагає виявити тимчасові зв'язки. Обидві моделі навчалися на розмічених даних, де кожен мережевий потік позначався як безпечний або зловмисний.

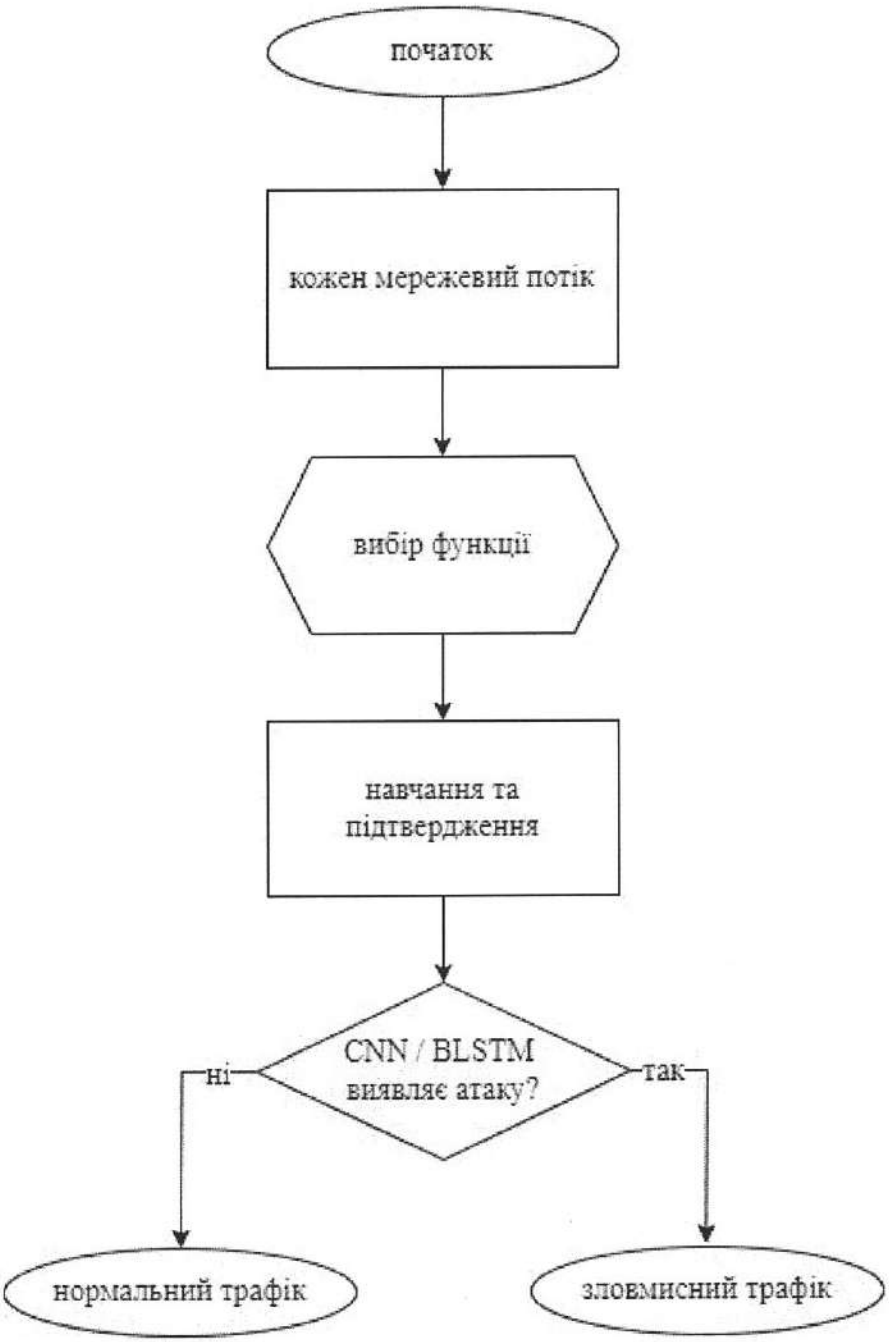


Рисунок 5 - Блок-схема процесу класифікації трафіку

Після тренування моделі CNN та BLSTM проходили тестування на різних вибірках трафіку, де вони класифікували потоки як атаки або звичайний трафік, використовуючи отримані закономірності. Це дозволяло виявляти потенційні DoS/DDoS-атаки в режимі реального часу.

Моделі на основі LSTM набули широкого застосування у виявленні атак, оскільки вони здатні аналізувати послідовні шаблони в даних та тимчасові залежності. Структуру комірки представлено на рисунку 6, а схема роботи LSTM мережі продемонстрована на рисунку 7.

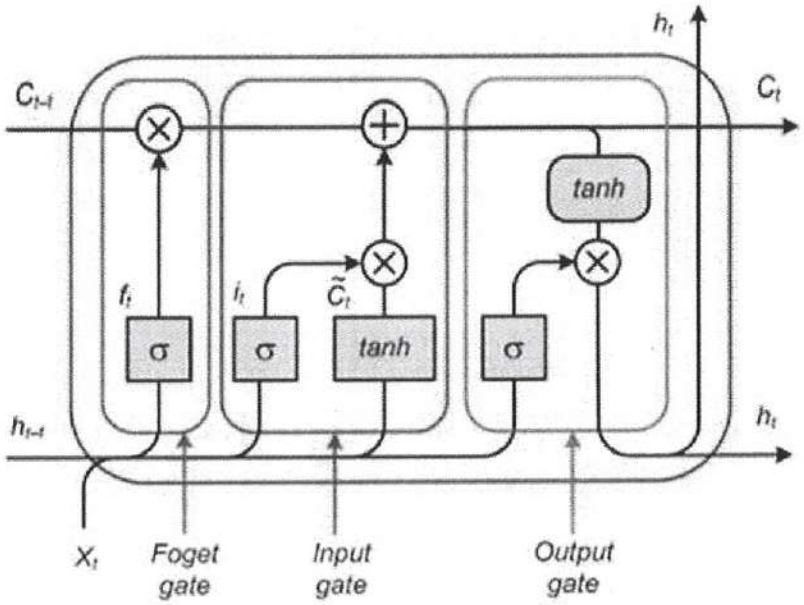


Рисунок 6 – Структура комірки LSTM

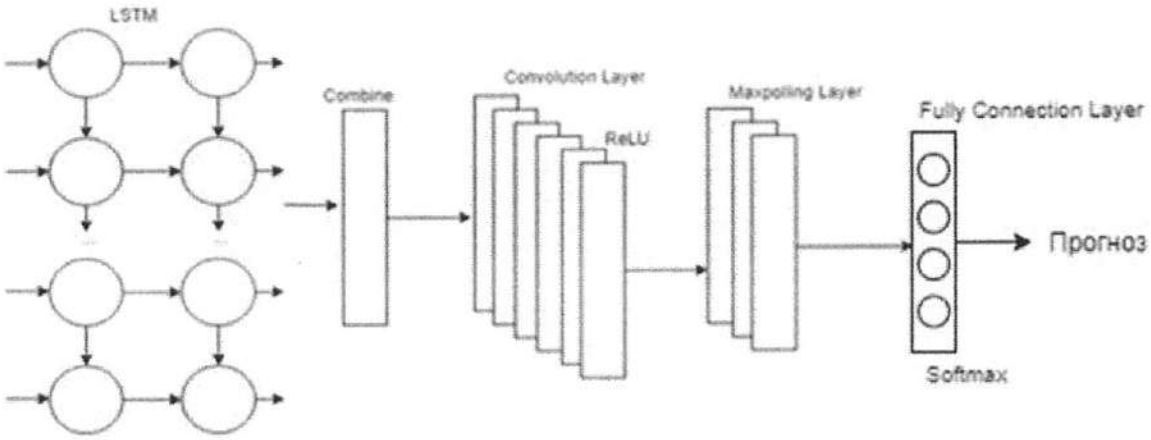


Рисунок 7 – Схема роботи LSTM мережі

Архітектура BLSTM покращує цю здатність, обробляючи вхідну послідовність у двох напрямках – вперед і назад, що забезпечує глибше розуміння контексту. Зі свого боку, CNN виконує класифікацію трафіку шляхом застосування згорткових фільтрів, які виділяють ключові ознаки, та використання шарів підвибірки (pooling), що зменшують розмірність вхідних даних. Виділені характеристики передаються у повнозв'язні шари, де відбувається фінальна класифікація трафіку на основі виявлених закономірностей. Схема роботи BLSTM мережі продемонстрована на рисунку 8.

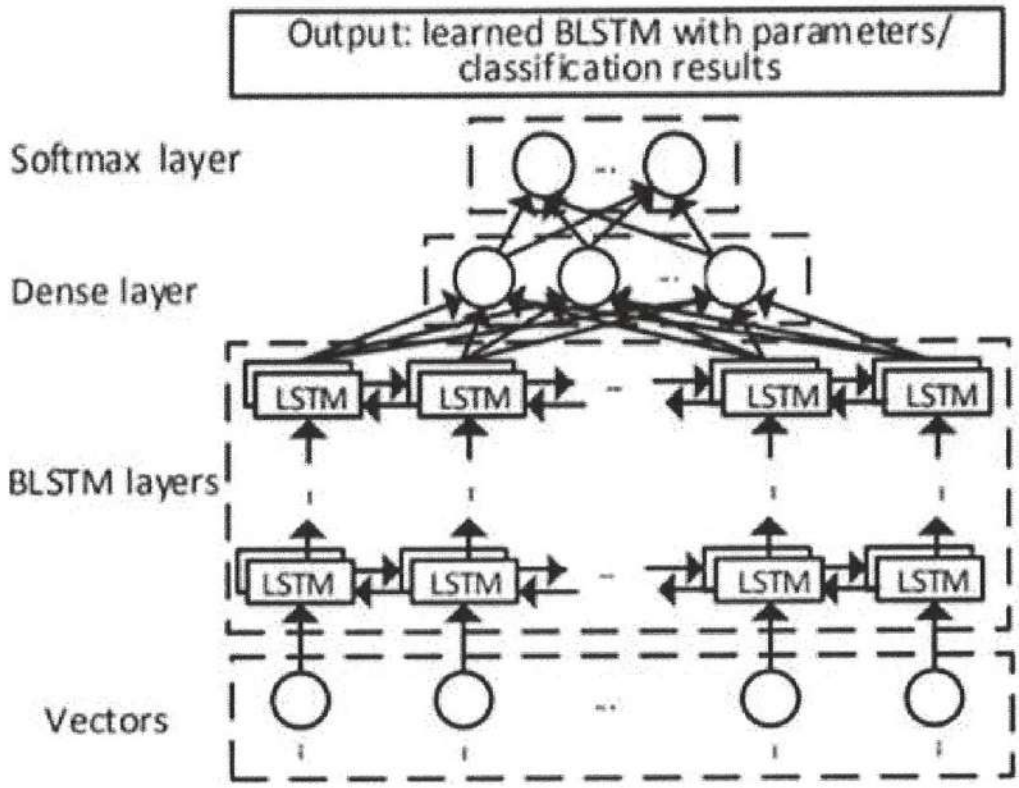


Рисунок 8 - Схема роботи BLSTM мережі

Для початкової оцінки було використано набір даних, що містив 1 мільйон рядків із зразками як безпечного, так і зловмисного трафіку. Дані охоплювали широкий спектр мережевих активностей, включаючи звичайний користувацький трафік, а також різні типи DoS- та DDoS-атак. Для кожного тестування набір даних ділився у співвідношенні 80/20: 80% використовувалося для навчання, а 20% – для тестування. Такий підхід гарантував збалансоване представлення як безпечного, так і зловмисного трафіку.

BLSTM навчалася за допомогою алгоритму Backpropagation Through Time (BPTT), який дозволяє моделі вивчати часові залежності в даних. У якості активаційної функції використовувалася tanh, що дозволяє ефективно обробляти складні патерни, приводячи вхідні значення до діапазону [-1, 1]. Тренувальний процес передбачав проходження моделі через набір даних протягом певної кількості епох із коригуванням вагових коефіцієнтів на основі похибки.

CNN складалася з чотирьох шарів, зокрема двох згорткових шарів (1D Conv), у той час як BLSTM містила чотири шари, один із яких був двонаправленим LSTM-шаром, а решта – повнозв'язні шари. В обох моделях у фінальному шарі

використовувалася активаційна функція sigmoid, а в якості оптимізатора – Adam. Навчання проходило протягом 60 епох із механізмом ранньої зупинки (early stopping) для запобігання перенавчання.

У цій роботі основним середовищем для тестування та початкового навчання нейронних мереж обрано платформу Google Colab. Це хмарний сервіс, створений компанією Google, який забезпечує зручний інтерфейс для написання, запуску та налагодження коду на мові програмування Python без необхідності встановлення додаткового програмного забезпечення на локальний комп'ютер. Його популярність серед дослідників і практиків у галузі штучного інтелекту зумовлена не лише простотою використання, а й можливістю безкоштовного доступу до графічних (GPU) та тензорних (TPU) процесорів. Саме ці ресурси дозволяють значно пришвидшити обчислення, що особливо важливо для задач машинного навчання, пов'язаних із великими обсягами даних та глибокими нейронними мережами. Ще однією перевагою Colab є вбудована підтримка всіх ключових бібліотек, зокрема TensorFlow, Keras, PyTorch, OpenCV, NumPy, Pandas та Scikit-learn, що дає змогу миттєво приступити до роботи, не витрачаючи час на налаштування середовища.

Проте для побудови повноцінної, масштабованої системи виявлення DDoS-атак важливо враховувати сценарії, у яких хмарні сервіси можуть бути недоступними або обмеженими. Саме тому в рамках дослідження додатково розглянуто варіанти локального та серверного розгортання системи. Такий підхід дозволив оцінити продуктивність і вимоги до апаратного забезпечення в реальних умовах.

На локальному рівні система була протестована на персональному комп'ютері з процесором Intel Core i7-11700, який має вісім фізичних ядер, і 32 гігабайтами оперативної пам'яті. За цих умов повний цикл навчання моделі на основі згорткової нейронної мережі (CNN) займав приблизно чотири години, тоді як модель BLSTM потребувала близько шести годин. Такий обсяг часу обумовлений відсутністю апаратного прискорення, характерного для спеціалізованих графічних процесорів, що є критично важливим фактором при роботі з великими масивами мережевих даних.

З метою пришвидшення обчислень та підвищення ефективності тренування моделей система також була розгорнута на віддаленому сервері з використанням відеокарти NVIDIA RTX 3060, що має 12 гігабайт відеопам'яті. У цьому середовищі час навчання CNN скоротився до сорока п'яти хвилин, а BLSTM — до сімдесяти хвилин. Це стало можливим завдяки використанню графічного прискорення та підтримці більших обсягів даних у процесі обробки. Зокрема, модель могла працювати з розмірами батчів до 128, що значно підвищувало швидкість і стабільність навчання без ризику перевантаження пам'яті. Подібні результати підтверджують, що в умовах реального впровадження систем виявлення атак у мережах нового покоління використання GPU є доцільним і навіть необхідним.

Окрім апаратного аспекту, важливу роль відіграє й правильна організація середовища. Для забезпечення портативності, стабільності та повторюваності експериментів реалізація системи здійснювалася у вигляді контейнеризованого застосунку, що працює на основі Docker. Такий підхід дав змогу повністю ізолювати робоче середовище, включаючи бібліотеки, залежності та середовище виконання, від зовнішніх чинників. У поєднанні з TensorFlow GPU це дозволило забезпечити масштабованість рішення — як на локальних вузлах, так і в корпоративних середовищах або кластерних обчислювальних системах.

Таким чином, побудова системи передбачала гнучку архітектуру, яка допускає використання як хмарних сервісів для швидкого прототипування, так і повноцінного локального або серверного розгортання з урахуванням конкретних технічних умов. Це робить запропонований підхід універсальним і практично застосовним у різних типах інформаційних систем.Щоб забезпечити баланс у даних, із зловмисного трафіку було випадково відібрано 1 мільйон рядків, що відповідало кількості безпечного трафіку. Кожен рядок отримав мітку «атака» або «безпечний» відповідно до оригінальних .csv-файлів.

Для BLSTM, яка працює з тривимірними (3D) даними, двовимірні (2D) дані були перетворені у формат 3D шляхом додавання додаткового виміру. Для CNN аналогічно виконувалося переформатування вхідних даних у 3D-формат для відповідності вимогам моделі.

Для досягнення високої точності виявлення DDoS-атак у рамках дослідження було проведено процес оптимізації гіперпараметрів для моделей нейронних мереж типу CNN і BLSTM. У цьому процесі застосовувалися три концептуальні підходи. Перший — це метод повного перебору комбінацій параметрів, відомий як Grid Search, який дозволяє визначити найефективнішу конфігурацію на основі вичерпного аналізу всіх можливих варіантів. Другий підхід, Random Search, передбачає випадковий вибір конфігурацій із заздалегідь визначеного простору параметрів, що значно скорочує обчислювальні витрати. Третій — Bayesian Optimization, який базується на побудові ймовірнісної моделі функції втрат для прогнозування ефективності нових налаштувань. Хоча цей підхід не був реалізований практично в рамках системи, його доцільність була розглянута на теоретичному рівні.

У ході оптимізації розглядалися такі характеристики моделі, як кількість шарів, обсяг нейронів у кожному з них, розмір батчу, швидкість навчання, кількість епох, тип функції активації та вибір оптимізатора серед поширених варіантів, зокрема Adam, RMSProp і SGD. Для моделі CNN було створено набір варіантів гіперпараметрів, де кількість фільтрів варіювалася між 32, 64 і 128, а ядра мали розміри (3,3) і (5,5). Навчання здійснювалось зі швидкістю 0.001, 0.0005 або 0.0001 у поєднанні з оптимізаторами типу Adam і RMSProp. Розміри батчів становили 32 або 64, а кількість епох — 20, 30 або 50. За результатами тестування понад п'ятдесяти різних конфігурацій, було обрано ті, які забезпечили найвищу F1-міру на валідаційному наборі даних.

Щодо моделі BLSTM, то процес налаштування охоплював визначення оптимальної кількості блоків LSTM, розміру прихованого стану, а також впровадження регуляризації у вигляді Dropout і SpatialDropout1D. Також розглядалися різні варіанти об'єднання шарів, які впливали на здатність моделі узагальнювати часові залежності у вхідних даних.

Моделі CNN та BLSTM пройшли навчання, валідацію та тестування на створеному наборі даних.

2.4 Висновки до розділу

Проведене дослідження алгоритмів виявлення DDoS-атак із використанням методів машинного навчання підтвердило їхню ефективність у мережах 5G. Було детально розглянуто процес створення наборів даних та навчання моделей CNN і BLSTM, що дозволило оцінити їхні можливості у класифікації трафіку та виявленні аномалій. Застосування машинного навчання забезпечує значне підвищення точності ідентифікації шкідливого трафіку порівняно з традиційними методами. CNN виявилася ефективною у розпізнаванні атак за статичними параметрами, тоді як BLSTM краще аналізує часові залежності в мережевих потоках. Важливим аспектом дослідження стало створення та адаптація якісних наборів даних, що містять як реальні сценарії атак, так і безпечний трафік, що є критично важливим для навчання моделей. Проаналізовано процес створення та обробки наборів даних для виявлення DDoS-атак у мережах 5G підтвердив важливість використання реалістичних симуляцій трафіку. Реалізація тестового середовища на основі Free5GC та UERANSIM дозволила отримати якісний набір даних, що враховують як звичайну активність користувачів, так і вплив шкідливих атак. Важливим аспектом дослідження стало порівняння отриманих наборів із загальнодоступними джерелами, що дало змогу виявити ключові характеристики трафіку, які мають найбільший вплив на точність класифікації атак. Попередня обробка даних, включаючи нормалізацію та відбір значущих характеристик, відіграла важливу роль у підвищенні ефективності навчання. Розподіл даних на тренувальні та тестові вибірки у співвідношенні 80/20 забезпечив збалансоване представлення безпечного та зловмисного трафіку, що сприяло узагальненню моделей та підвищенню їхньої стійкості до нових зразків даних. Результати тестування CNN та BLSTM показали, що кожна з моделей має свої переваги. CNN ефективніше аналізує просторові залежності між характеристиками трафіку, тоді як BLSTM краще справляється із виявленням часових патернів, що критично важливо для аналізу послідовних мережевих потоків. Поєднання цих підходів може сприяти створенню більш надійних систем виявлення атак у мережах 5G.

3 ОЦІНКА ДОСТОВІРНОСТІ СИСТЕМИ

3.1 Метрики оцінювання достовірності

Продуктивність моделей МН оцінюється за допомогою різноманітних метрик, які можуть застосовуватись на етапах навчання, валідації та тестування моделі. Вибір конкретних метрик залежить від типу вирішуваної задачі. Наприклад, для задач бінарної класифікації підходять такі показники, як матриця плутанини (confusion matrix), акуратність (accuracy), точність (precision), повнота (recall) та F1-міра. Ідеальний класифікатор має значення цих показників, близькі до 1. Далі наведено опис кожного з них та формули розрахунку:

– правильні класифікації позитивного класу (True Positive, TP) відображають ситуацію, коли і модель, і істинна відповідь відносять зразок до позитивного класу;

– правильні класифікації негативного класу (True Negative, TN) вказують на те, що зразок належить до негативного класу;

– помилкові позитиви (False Positive, FP) демонструють ситуацію, коли модель віднесла зразок до позитивного класу, хоча насправді він належить до негативного;

– помилкові негативи (False Negative, FN) вказують на випадки, коли модель неправильно класифікувала позитивні приклади як негативні.

Акуратність (accuracy) показує загальну частку правильно класифікованих прикладів. Вона обчислюється як сума TP та TN, поділена на загальну кількість зразків. Проте ця метрика є недостатньо інформативною при великій нерівновазі класів, коли більшість трафіку є легітимним, і навіть проста модель, що завжди обирає "безпечний" клас, може демонструвати високу точність:

$$Accuracy = \frac{TP+T}{TP+FP+FN+T} \tag{3.1}$$

Точність (precision) вимірює, яку частину з передбачених як зловмисні зразків модель класифікувала правильно. Вона є особливо важливою в ситуаціях, коли критично зменшити кількість хибних спрацьовувань (FP), наприклад, щоб

уникнути блокування легітимного трафіку:

$$Precision = \frac{TP}{TP+FP} \tag{3.2}$$

Повнота (recall), також відома як чутливість, показує, яку частку дійсно зловмисного трафіку вдалося виявити. Цей показник важливий у контексті виявлення загроз, де пропуск атаки (FN) може призвести до значних наслідків:

$$Recall = \frac{TP}{TP+FN} \tag{3.3}$$

Помилка вимірює здатність системи виявлення зловмисників правильно ідентифікувати незловмисні об'єкти або події як незловмисні:

$$Specificity = \frac{FP+FN}{TP+FP+TN+FN} \tag{3.4}$$

F1-міра є гармонійним середнім між точністю та повнотою і дозволяє досягти компромісу між цими двома характеристиками. Особливо корисною вона є у випадках, коли важливо забезпечити баланс між здатністю виявляти всі загрози й уникненням надмірної кількості помилкових сигналів:

$$Fscore = \frac{Recall+Precision}{2} \tag{3.5}$$

Використання лише однієї метрики для оцінки продуктивності моделі може призвести до викривленого уявлення про її реальну ефективність. Тому в даній роботі застосовується комплексний підхід до оцінювання, що включає всі зазначені вище метрики. Це дозволяє надати повну картину роботи системи в різних ситуаціях та обґрунтувати вибір конкретної архітектури моделі в умовах реального трафіку, характерного для мереж п'ятого покоління.

Крім того, важливою особливістю в контексті 5G є динаміка трафіку, його фрагментованість через мережеві зрізи та висока щільність підключень, що

ускладнює задачу класифікації. Саме тому запропоновані метрики оцінки застосовуються не лише до агрегованих результатів, а й до різних підмножин даних, що дозволяє оцінити стабільність та узагальнювальну здатність моделі.

3.2 Оцінювання достовірності роботи пропонованої системи

Для підтвердження ефективності та надійності запропонованої системи було проведено детальне оцінювання роботи двох моделей глибокого навчання — CNN (Convolutional Neural Network) та BLSTM (Bidirectional Long Short-Term Memory). Порівняльний аналіз цих моделей дозволяє виявити їхні сильні та слабкі сторони при роботі з різними наборами даних, що включають як традиційні, так і сучасні типи DDoS-атак. Матриці плутанини для CNN та BLSTM, що дозволяють оцінити точність класифікації, наведено в таблиці 3.1. Вони відображають кількість правильних та помилкових класифікацій для кожного класу трафіку. Кількаразове виконання експериментів на кожному з датасетів дозволило зменшити вплив випадкових коливань і зробити результати більш достовірними.

Таблиця 3.1 містить детальні значення чотирьох основних метрик класифікації — акура́тності, F-міри, точності та повноти — для кожного набору даних. Також наведені відповідні стандартні відхилення, що дозволяє оцінити стабільність кожної моделі. Аналіз охоплює чотири різні датасети: Dataset-2, Dataset-3, CICDDoS 2019 та 5G NIDD, які відрізняються складністю структури трафіку, кількістю класів та обсягом даних.

За показником акуратності модель CNN стабільно демонструє кращі результати у порівнянні з BLSTM. Наприклад, для Dataset-2 CNN досягла 98,09% при стандартному відхиленні 1,39, тоді як BLSTM показала 86,76% із відхиленням лише 0,11. Це свідчить не лише про перевагу CNN у точності класифікації, а й про її стійкість до варіативності у вхідних даних. У Dataset-3 різниця ще помітніша — CNN має практично ідеальну акуратність на рівні 99,96%, у той час як BLSTM — лише 93,87%. У CICDDoS 2019 модель CNN зберігає лідерство — 98,62% проти 97,77% у BLSTM. Найбільша різниця спостерігається для складного набору 5G

NIDD, де CNN досягає 94,67%, а BLSTM лише 89,88%.

Таблиця 3.1 – Показники достовірності

Параметр / Значення	Набір даних	CNN	Стандартне відхилення для CNN	BLSTM	Стандартне відхилення для BLSTM
Акуратність	Dataset-2	98.09	1.39	86.76	0.11
	Dataset-3	99.96	0.01	93.87	8.26
	CICDDoS 2019	98.62	0.97	97.77	0.15
	5G NIDD	94.67	0.88	89.88	1.83
F-міра	Dataset-2	98.09	1.38	87.53	1.66
	Dataset-3	99.96	0.01	94.66	7.17
	CICDDoS 2019	98.60	0.99	97.73	0.16
	5G NIDD	94.87	0.51	90.54	1.73
Точність	Dataset-2	98.71	0.74	99.50	0.23
	Dataset-3	99.92	0.01	99.83	0.04
	CICDDoS 2019	98.17	1.99	95.97	0.11
	5G NIDD	95.79	2.13	96.72	2.3
Повнота	Dataset-2	97.49	2.29	83.48	9.22
	Dataset-3	99.99	0.01	90.71	12.72
	CICDDoS 2019	99.05	0.56	99.54	0.22
	5G NIDD	94.08	2.92	85.11	1.47

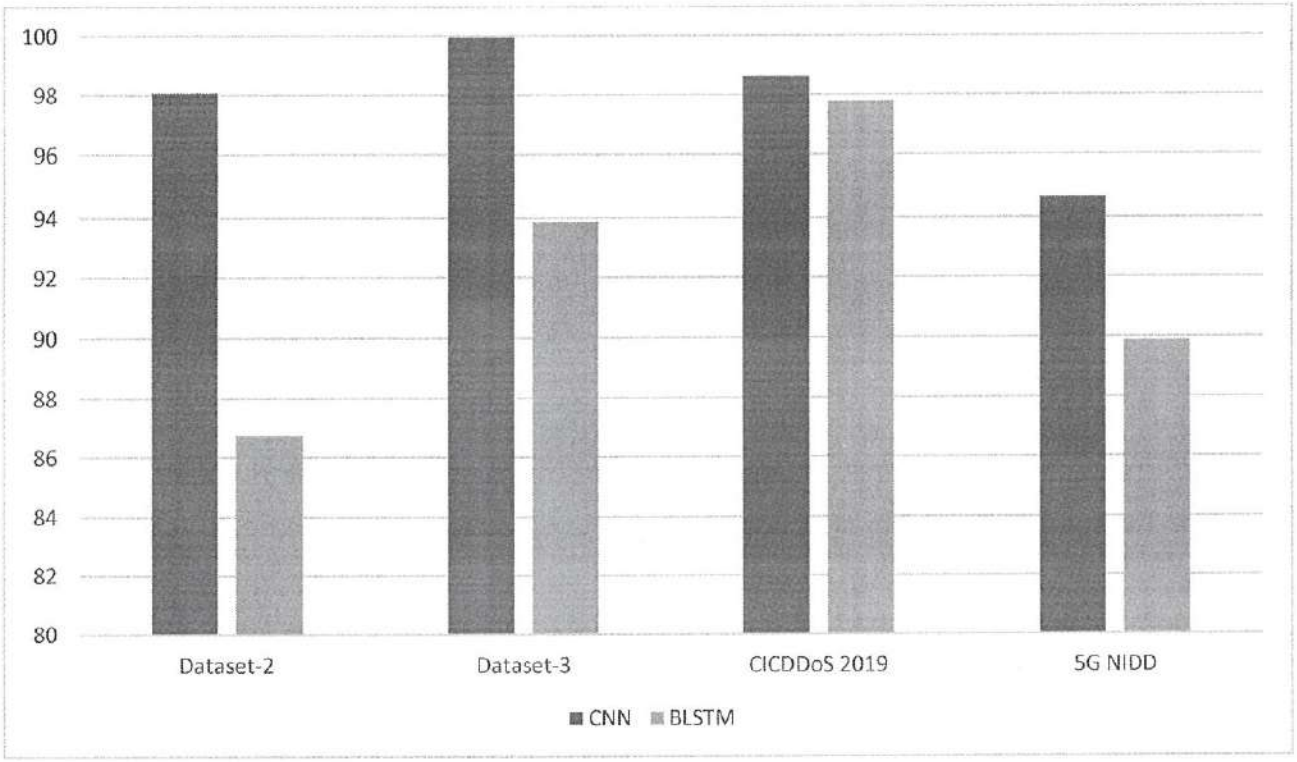


Рисунок 9 - Порівняння акуратності CNN і BLSTM для різних наборів даних

F-міра враховує як точність (precision), так і повноту (recall), і є більш комплексним показником якості класифікації. Для Dataset-2 CNN досягла 98,09% зі стандартним відхиленням 1,38, а BLSTM — 87,53%, що підтверджує значну перевагу. У Dataset-3 спостерігається аналогічна картина — 99,96% проти 94,66%. У CICDDoS 2019 модель CNN знову показує вищу F-міру — 98,60% у порівнянні з 97,73% у BLSTM. Особливо важливими є результати для набору 5G NIDD — найскладнішого з усіх, адже він містить нові типи трафіку, характерні для мереж 5G. Тут F-міра для CNN становить 94,87%, у той час як BLSTM показує дещо нижче значення — 90,54%.

Аналіз точності також демонструє перевагу CNN. Наприклад, для Dataset-3 точність CNN становить 99,92%, що є майже ідеальним результатом, тоді як BLSTM — 99,83%. Для CICDDoS 2019 точність CNN — 98,17%, а BLSTM — 95,97%. Незважаючи на те, що точність BLSTM подекуди досить висока, CNN демонструє меншу варіативність результатів та загалом стабільніші показники на всіх наборах.

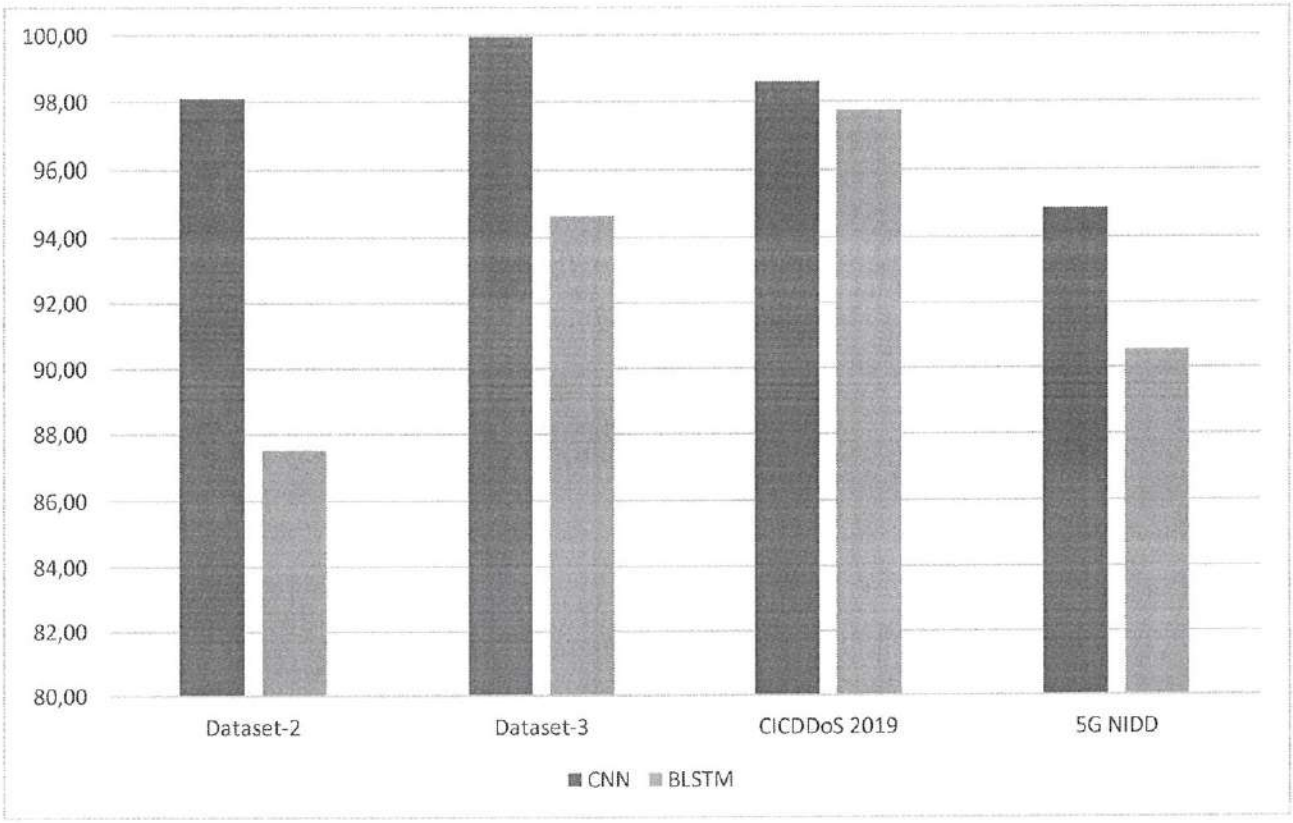


Рисунок10 - Порівняння F-міри CNN і BLSTM для різних наборів даних

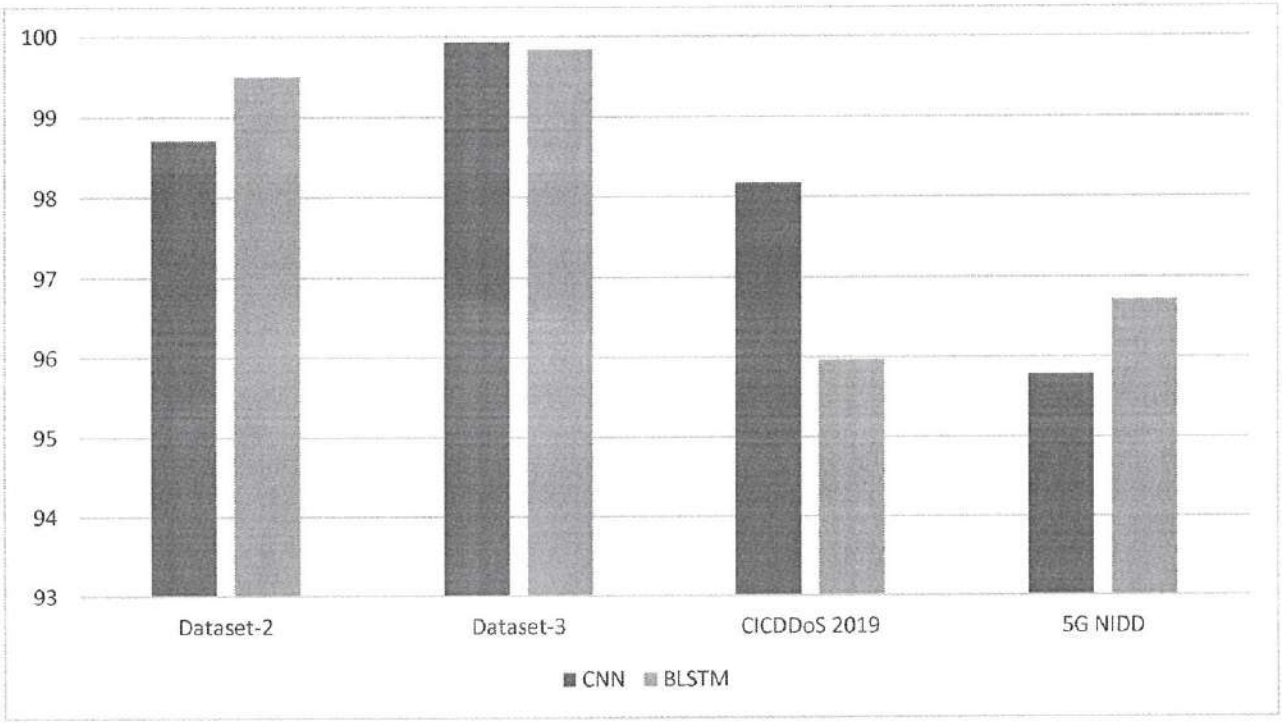


Рисунок 11- Порівняння точності CNN і BLSTM для різних наборів даних

Повнота є критично важливою метрикою у виявленні DoS/DDoS-атак, оскільки вона показує, наскільки добре система виявляє всі потенційні загрози. У Dataset-2 CNN досягає 97,49%, тоді як BLSTM — лише 83,48%. У Dataset-3 повнота

для CNN становить 99,99%, тоді як у BLSTM — 90,71%. Найвищу повноту серед усіх спостережуваних випадків демонструє CNN на CICDDoS 2019 — 99,05%, що говорить про здатність моделі виявляти майже всі справжні атаки. BLSTM, хоч і демонструє досить високий результат (99,54%), все ж поступається стабільністю. Для 5G NIDD різниця становить близько 9% — CNN досягає 94,08%, BLSTM — 85,11%.

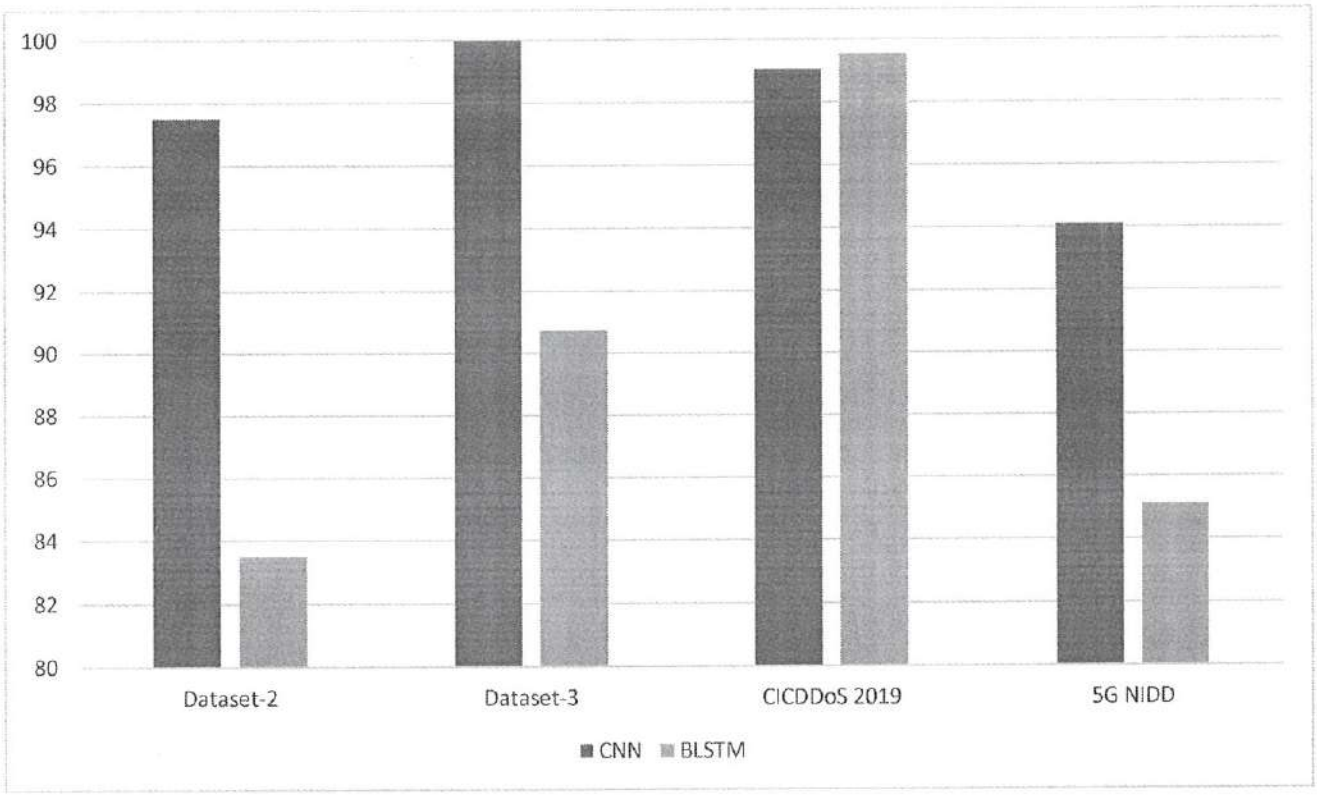


Рисунок 12 - Порівняння повноти CNN і BLSTM для різних наборів даних

Проведений аналіз дозволяє зробити чіткий висновок: модель CNN значно переважає BLSTM за всіма ключовими показниками ефективності класифікації мережевого трафіку. Це стосується як акуратності, так і F-міри, точності та повноти. CNN не лише досягає вищих абсолютних значень, але й демонструє менші стандартні відхилення, що свідчить про її стабільність і надійність у реальних умовах. Крім того, результати, отримані за допомогою матриць плутанини, підтверджують здатність CNN точно відрізняти шкідливий трафік від легітимного, що зменшує ймовірність як хибнопозитивних, так і хибнонегативних результатів. Це робить CNN особливо придатною для використання у високонавантажених

середовищах, таких як дата-центри, хмарні сервіси та оператори зв'язку, де необхідна швидка і точна ідентифікація кіберзагроз. Загалом, запропонована система з використанням CNN може бути рекомендована як ефективне рішення для побудови систем виявлення DoS/DDoS-атак нового покоління, здатне адаптуватися до змін у структурі трафіку та загроз, що постійно еволюціонують..

3.3 Висновки до розділу

У контексті забезпечення кібербезпеки сучасних телекомунікаційних мереж, особливо з урахуванням складності трафіку у мережах п'ятого покоління (5G), ключовим завданням залишається своєчасне виявлення аномальної активності. У цьому дослідженні проведено комплексне порівняння двох моделей глибокого навчання — згорткової нейронної мережі (CNN) та бінаправленої довготривалої пам'яті BLSTM — з метою визначення їх ефективності у задачах класифікації мережевого трафіку та виявлення DoS/DDoS-атак.

Результати моделювання свідчать про перевагу CNN над BLSTM за усіма основними метриками, зокрема точністю, F1-мірою, recall та precision. CNN стабільно демонструє вищу здатність до узагальнення на різноманітних наборах даних, включаючи Dataset-3 та спеціалізований набір трафіку 5G NIDD. У той час як BLSTM показує потенціал у моделюванні часових залежностей, її продуктивність виявляється більш вразливою до змін у структурі трафіку та менш стійкою до високої варіативності, характерної для сучасних мережевих середовищ.

Для поглибленого аналізу ефективності CNN у ситуаціях незбалансованого трафіку було проведено два експерименти на основі Dataset-3. У першому досліджувалося зменшення обсягу зловмисного трафіку щодо безпечного: від 20% до 100% з кроком у 20%. Наприклад, якщо безпечного трафіку було 1 мільйон записів, то до навчальної вибірки включалося від 200 тис. до 1 млн записів атак. У другому експерименті варіювався обсяг безпечного трафіку щодо зловмисного, дотримуючись того ж діапазону. Обидва експерименти показали стабільну перевагу CNN: навіть при значній диспропорції класів вона зберігала високі

значення метрик, що підтверджує її здатність ефективно функціонувати в реальних умовах, де баланс класів не гарантується.

Порівняльне тестування на чотирьох різних датасетах, включно з трафіком 5G, дозволило об'єктивно оцінити здатність моделей до масштабування та адаптації. CNN виявилася більш універсальною, демонструючи не лише високу середню ефективність, але й низьку варіативність результатів при зміні розподілу вхідних даних. Це свідчить про її придатність для розгортання в умовах динамічного трафіку, де характер атак може швидко змінюватися.

Особливо важливою є здатність CNN до обробки великого обсягу вхідних даних у режимі реального часу. Це критично важливо для інфраструктури інтернет-провайдерів та великих підприємств, які потребують високої продуктивності систем виявлення загроз. Крім того, CNN забезпечує високу повноту (recall), що означає мінімальні ризики пропуску реальних атак, а також високу точність (precision), що знижує кількість помилкових спрацювань і, відповідно, навантаження на персонал, який обробляє сповіщення. Попри здатність BLSTM моделювати часові характеристики трафіку, її обчислювальна складність та нестабільність у змінних середовищах, як показало тестування, обмежують практичне застосування в реальному часі. Модель вимагає більше ресурсів для досягнення порівнянної точності, особливо у випадку зі складними патернами трафіку, характерними для 5G.

Загалом, результати підтверджують, що CNN є доцільною базовою архітектурою для побудови системи виявлення аномалій у мережевому трафіку, орієнтованої на актуальні умови високошвидкісних мобільних мереж. Її переваги — стабільність, висока ефективність, адаптивність до великого обсягу даних та здатність функціонувати у режимі реального часу — роблять її перспективною основою для подальшого розвитку систем захисту.

ВИСНОВКИ

У процесі виконання кваліфікаційної роботи було проведено всебічне дослідження проблеми виявлення DoS/DDoS-атак у мережевому трафіку п'ятого покоління з використанням методів глибокого навчання. Актуальність цієї теми зумовлена динамічним впровадженням 5G-мереж, збільшенням кількості пристроїв, обсягів переданих даних і зростаючою складністю кіберзагроз. Нові архітектурні підходи, властиві п'ятому поколінню мобільного зв'язку, такі як мережеве розділення, хмарна віртуалізація, локалізоване обслуговування (edge computing) та висока щільність користувачів, створюють нові вектори для атак, зокрема з боку розподілених систем відмови в обслуговуванні. Традиційні методи виявлення атак виявляються недостатньо ефективними в таких умовах, що вимагає розробки адаптивних систем, здатних до самонавчання та реагування в реальному часі.

З метою створення реалістичного середовища для дослідження було змодельовано повноцінну інфраструктуру 5G з використанням програмних рішень Free5GC та UERANSIM. Це дозволило згенерувати автентичний трафік, який включає сигнальний трафік, трафік користувачів та різні типи аномальної активності. Нормальний трафік генерувався під час перегляду відео, завантаження файлів і використання браузера, тоді як зловмисний трафік охоплював різновиди атак, серед яких TCP SYN Flood, UDP Flood, Smurf, Land та інші. Усі потоки було зафіксовано за допомогою Wireshark, оброблено й перетворено у формат, придатний для машинного аналізу, з виділенням 84 ключових характеристик кожного з'єднання.

Для аналізу трафіку було реалізовано дві архітектури глибокого навчання — згорткову нейронну мережу (CNN) та бінаправлену довготривалу пам'ять (BLSTM). Навчання моделей проводилося на трьох різних датасетах: Dataset-3, CIC-DDoS2019 та 5G-NIDD, які охоплюють як класичні сценарії атак, так і специфіку трафіку мереж п'ятого покоління. У процесі налаштування моделей використовувалися сучасні прийоми, зокрема регуляризація, Dropout, Batch Normalization та оптимізація гіперпараметрів. Критерії оцінювання ефективності

включали точність, повноту, F1-міру та акуратність. Також досліджувалася стабільність моделей шляхом аналізу стандартних відхилень, що дозволило виявити ступінь коливань результатів за змінних умов.

Експерименти показали, що CNN впевнено переважає BLSTM за всіма ключовими метриками. У найпростіших випадках, наприклад у Dataset-3, точність CNN сягала 99.96%, що практично дорівнює ідеальній класифікації. У складніших умовах, таких як CIC-DDoS2019 та 5G-NIDD, CNN також демонструвала значну перевагу, зберігаючи високі показники ефективності та низькі відхилення. Особливо показовими були результати для трафіку мереж 5G, де CNN зберігала точність на рівні 94–95% і при цьому демонструвала кращу стійкість до змін структури даних. BLSTM, хоча й здатна враховувати часову динаміку, виявилася менш ефективною у ситуаціях із великою кількістю параметрів та непередбачуваними шаблонами атак.

Окрему увагу було приділено впливу дисбалансу класів у навчальних вибірках. Було змодельовано сценарії, в яких співвідношення шкідливого до безпечного трафіку змінювалося в діапазоні від 20% до 100%. Навіть у випадках значної переваги одного класу CNN зберігала високу продуктивність, що доводить її придатність до застосування у реальних умовах, де легітимний трафік значно переважає зловмисний. Це дозволяє уникнути типових проблем хибної класифікації, властивих для багатьох інших моделей, які потребують жорсткої балансованості даних.

Практичне значення дослідження полягає у тому, що модель CNN не лише показала високу точність, а й виявила здатність до роботи в режимі реального часу, що є критично важливим для операторів зв'язку, дата-центрів і хмарних сервісів. Її можна інтегрувати у компоненти 5G-інфраструктури, наприклад у функцію користувацької площини UPF, що забезпечить миттєву перевірку та фільтрацію трафіку на локальному рівні, без потреби передавання на центральні вузли. У порівнянні з BLSTM CNN є менш ресурсоємною, забезпечує швидшу обробку запитів, краще масштабується та дозволяє легше оновлювати модель у міру появи нових типів атак.

Отримані результати створюють передумови для подальшого розвитку

систем виявлення атак на основі глибокого навчання. Подальші дослідження можуть бути спрямовані на інтеграцію CNN з адаптивними механізмами навчання, що дозволить системі реагувати на нові атаки без повного перенавчання. Цілеспрямованим напрямом може бути також впровадження методів пояснюваного штучного інтелекту, які дадуть змогу адміністраторам мереж краще розуміти, на основі яких ознак було прийняте рішення про класифікацію. У перспективі можлива розробка гібридних систем, де CNN поєднуватиметься з іншими алгоритмами (наприклад, графовими нейронними мережами), що дозволить здійснювати ще глибший аналіз аномалій.

Таким чином, виконане дослідження доводить, що CNN є ефективним, стабільним та перспективним інструментом для виявлення DoS/DDoS-атак у мережевому трафіку п'ятого покоління. Результати моделювання, експериментальної перевірки та оцінювання дозволяють рекомендувати запропоновану архітектуру до впровадження у реальні системи кіберзахисту та використовувати як основу для подальших досліджень і розробок у сфері інтелектуального аналізу трафіку.

ПЕРЕЛІК ДЖЕРЕЛ

1. What are the characteristics of 5G. URL: https://www.realme.com/in/support/kw/doc/2055960 (дата звернення: 18.02.2025).

2. What is 5G? URL: https://www.ibm.com/think/topics/5g (дата звернення: 18.02.2025)

3. 5G NR frequency band. URL: https://www.sqimway.com/nr_band.php (дата звернення: 18.02.2025)

4. 5G Frequency Bands & Spectrum Allocations. URL: https://www.cablefree.net/wirelesstechnology/4glte/5g-frequency-bands-lte/ (дата звернення: 19.02.2025)

5. 5G Frequency Bands. URL: https://portal.powertec.com.au/technical-library/wireless/wireless-technologies/cellular/5g-nr/5g-frequency-bands (дата звернення: 20.02.2025)

6. Patel B., Yarlagadda V. K., Dhameliya N., Mullangi K., Vennapusa S. C. R. Advancements in 5G Technology: Enhancing Connectivity and Performance in Communication Engineering. Engineering International. 2022. Vol. 10, No. 2, pp. 117-130

7. Attar H., Issa H., Ababneh J., et al. 5G System Overview for Ongoing Smart Applications: Structure, Requirements, and Specifications. Computational Intelligence and Neuroscience. 2022. DOI: 10.1155/2022/2476841

8. Dangi R., Lalwani P., Choudhary G., You I., Pau G. Study and Investigation on 5G Technology: A Systematic Review. Sensors. 2022. Vol. 22, No. 1. DOI: 10.3390/s22010026.

9. Bhayre H., Hirannaiah R. M. Secure Authentication Mechanism for AMF in 5G Nb-IoT devices. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). Kamand, India. 2024. pp. 1-7. DOI: 10.1109/ICCCNT61001.2024.10725154.

10. Park S., Kwon S., Park Y., Kim D., You I. Session Management for Security Systems in 5G Standalone Network. IEEE Access. 2022. Vol. 10, pp. 73421-73436. DOI: 10.1109/ACCESS.2022.3187053

11. Chen Z., Ao J., Luo W., Cheng Z., Liu Y., Sheng K., Chen L. A dual-factor access authentication scheme for IoT terminal in 5G environments with network slice selection. Journal of Information Security and Applications. 2022. Vol. 68. DOI: 10.1016/j.jisa.2022.103247.

12. Huang X., Tsiatsis V., Palanigounder A., Su L., Yang B. 5G Authentication and Key Management for Applications. IEEE Communications Standards Magazine. 2021. Vol. 5, No. 2, pp. 142-148. DOI: 10.1109/MCOMSTD.001.2000024.

13. Shahen Shah A. F. M. A Survey From 1G to 5G Including the Advent of 6G: Architectures, Multiple Access Techniques, and Emerging Technologies. 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC). Las Vegas, NV, USA. 2022. pp. 1117-1123. DOI: 10.1109/CCWC54503.2022.9720781.

14. Tang Q., Ermis O., Nguyen C. D., Oliveira A. D., Hirtzig A. A Systematic Analysis of 5G Networks With Focus on 5G Core Security. IEEE Access. 2022. Vol. 10, pp. 18298-18319. DOI: 10.1109/ACCESS.2022.3151000.

15. Muzaffar R., Ahmed M., Sisinni E., Sauter T., Bernhard H.P. 5G Deployment Models and Configuration Choices for Industrial Cyber-Physical Systems – A State of Art Overview. IEEE Transactions on Industrial Cyber-Physical Systems. 2023. Vol. 1, pp. 236-256. DOI: 10.1109/TICPS.2023.3311394

16. Leyva-Pupo I., Cervelló-Pastor C. Efficient solutions to the placement and chaining problem of User Plane Functions in 5G networks. Journal of Network and Computer Applications. 2022. Vol. 197. DOI: 10.1016/j.jnca.2021.103269.

17. DDoS атаки і захист від них. URL: https://www.ukraine.com.ua/blog/hosting_ukraine/ddos-ataki-i-zashchita-ot-nih.html (дата звернення: 23.02.2025)

18. DDoS Attack Meaning. URL: https://www.fortinet.com/resources/cyberglossary/ddos-attack (дата звернення: 23.02.2025).

19. Understanding Denial-of-Service Attacks. URL: https://www.cisa.gov/news-events/news/understanding-denial-service-attacks (дата звернення: 26.02.2025).

20. Найбільші DDoS атаки, методи їх проведення та наслідки. URL:

https://foxminded.ua/naibilshi-ddos-ataky/ (дата звернення: 26.02.2025).

21. DDoS-атаки: Типи атак та рівні моделі OSI. URL: https://alexhost.com/uk/faq/ddos-ataky-typy-atak-ta-rivni-modeli-osi/ (дата звернення: 27.02.2025).

22. Complete Guide to the Types of DDoS Attacks. URL: https://www.esecurityplanet.com/networks/types-of-ddos-attacks/ (дата звернення: 27.02.2025).

23. Different types of DDoS attacks: how to protect your clients. URL: https://www.connectwise.com/blog/cybersecurity/types-of-ddos-attacks (дата звернення: 27.02.2025).

24. Vo H., Kozlowski R., Odebode I., Hodson D. D., Grimaila M. R. Simulation of SYN Flood Attack and Counter-Attack Methods Using Average Connection Times. In: Arabnia H. R., et al. Advances in Parallel & Distributed Processing, and Applications. Transactions on Computational Science and Computational Intelligence. Springer, Cham. 2021. DOI: 10.1007/978-3-030-69984-0_5

25. UDP Flood Attack: Definition & Prevention. URL: https://www.indusface.com/learning/udp-flood-attack/ (дата звернення: 7.03.2025).

26. Abdelrazek L., Fuladi R., Kövér J., Karaçay L., Gülen U. Detecting IP DDoS Attacks Using 3GPP Radio Protocols. IEEE Access. 2024. Vol. 12, pp. 24776-24790. DOI: 10.1109/ACCESS.2024.3365425.

27. Khan R. A., Khan H. U., Alwageed H. S., Alhashimi H. A., Keshta I. 5G Networks Security Mitigation Model: An ANN-ISM Hybrid Approach. IEEE Open Journal of the Communications Society. 2025. DOI: 10.1109/OJCOMS.2025.3529717.

28. Boualouache A., Brik B., Senouci S.-M., Engel T. On-Demand Security Framework for 5GB Vehicular Networks. IEEE Internet of Things Magazine. 2023. Vol. 6, No. 2, pp. 26-31. DOI: 10.1109/IOTM.001.2200233.

29. Farzaneh B., Shahriar N., Muktadir A. H. A., Towhid M. S., Khosravani M. S. DTL-5G: Deep transfer learning-based DDoS attack in 5G and beyond networks. Computer Communications. 2024. Vol. 228. DOI: 10.1016/j.comcom.2024.107927.

30. Bousalem B., Silva V. F., Langar R., Cherrier S. DDoS Attacks Detection and Mitigation in 5G and Beyond Networks: A Deep Learning-based Approach.

GLOBECOM 2022 - IEEE Global Communications Conference. Rio de Janeiro, Brazil. 2022. pp. 1259-1264. DOI: 10.1109/GLOBECOM48099.2022.10001562.

31. Aryal B., Abbas R., Collings I. B. SDN Enabled DDoS Attack Detection and Mitigation for 5G Networks. Journal of Communications. 2021. Vol. 16, No. 7, pp. 267-275. DOI: 10.12720/jcm.16.7.267-275

32. Sheikhi S., Kostakos P. DDoS attack detection using unsupervised federated learning for 5G networks and beyond. 2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit). Gothenburg, Sweden. 2023. pp. 442-447. DOI: 10.1109/EuCNC/6GSummit58263.2023.10188245.

33. Dahiya D. DDoS Attacks Detection in 5G Networks: Hybrid Model with Statistical and Higher-Order Statistical Features. Cybernetics and Systems. 2022. Vol. 54, No. 6, pp. 888–913. DOI: 10.1080/01969722.2022.2122002

34. Alzhrani R. M., Alliheedi M. A. 5G Networks and IoT Devices: Mitigating DDoS Attacks with Deep Learning Techniques. Computer Science. 2023. DOI: 10.48550/arXiv.2311.06938.

35. Park S., Cho B., Kim D., You I. Machine Learning Based Signaling DDoS Detection System for 5G Stand Alone Core Network. Applied Sciences. 2022. Vol. 12, Article 12456. DOI: 10.3390/app122312456

36. Maiwada U. D., Danyaro K. U., Sarlan A. B., et al. Enhancing DDoS Detection in 5G Systems through Advanced Intrusion Detection Techniques. Zhongguo Kuangye Daxue Xuebao. 2024. Vol. 29, No. 3, pp. 130-140.

37. Random Forest Classification. URL: https://www.datacamp.com/tutorial/random-forests-classifier-python (дата звернення: 7.03.2025).

38. Support Vector Machines (SVM). URL: https://medium.com/@RobuRishabh/support-vector-machines-svm-27cd45b74fbb (дата звернення: 7.03.2025).

39. XGBoost Documentation. URL: https://xgboost.readthedocs.io/en/stable/ (дата звернення: 8.03.2025).

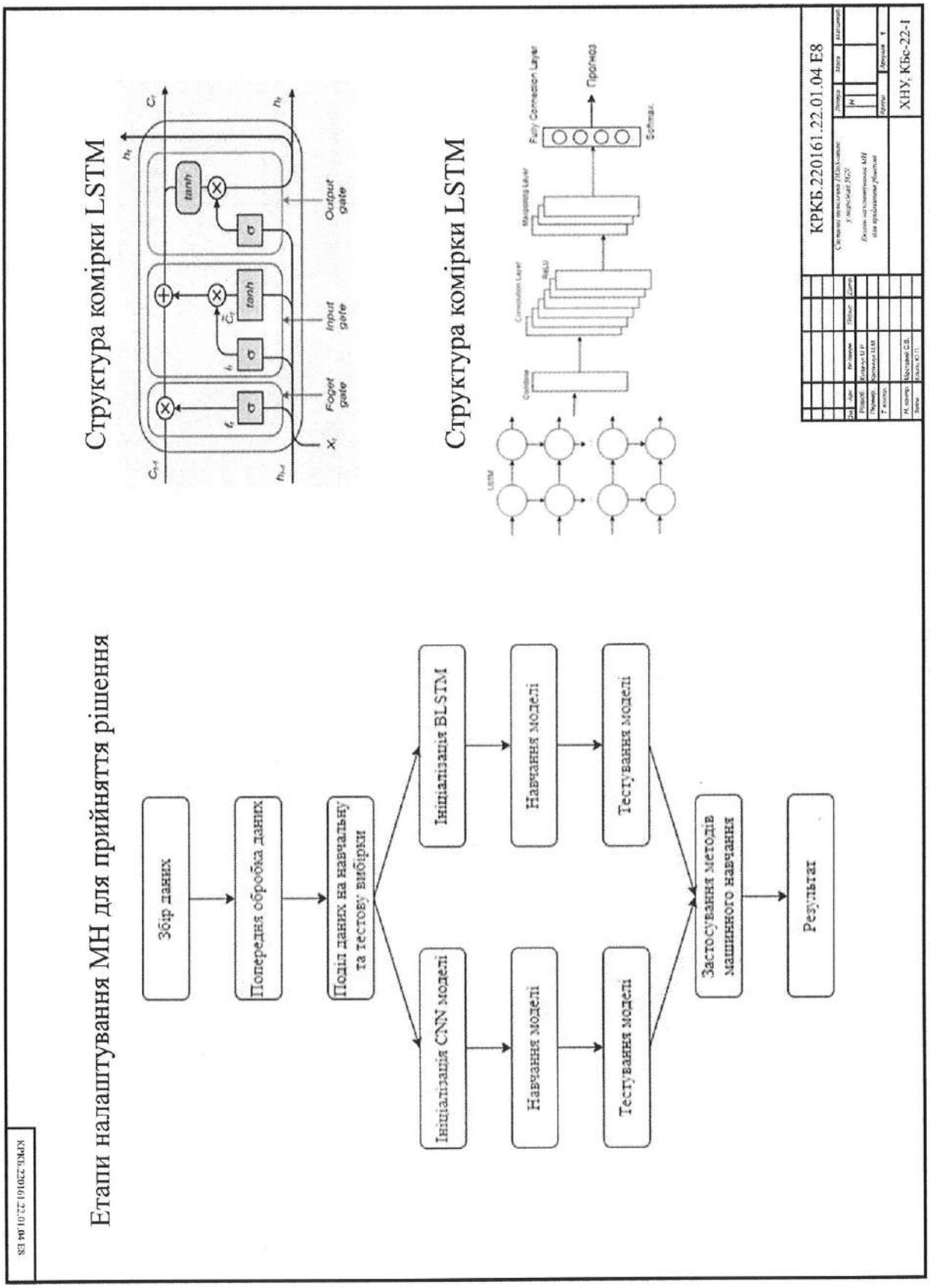
40. СОУ 207.01:2017. Текстові документи. Загальні вимоги. Хмельницький: ХНУ, 2017. 46 с. URL:

https://msn.khnu.km.ua/pluginfile.php/466522/mod_resource/content/1/132_C%20T%20A%20H%20Д%20A%20P%20T%20чист%20.pdf

41. ДСТУ 8302:2015. Бібліографічне посилання. Загальні положення та правила складання. [Чинний від 2016-07-1]. Київ, 2016. 20 с. (Державна наукова установа —Книжкова палата України імені Івана Федорова).

ДОДАТОК А

Копії графічної частини



Блок-схема процесу класифікації трафіку

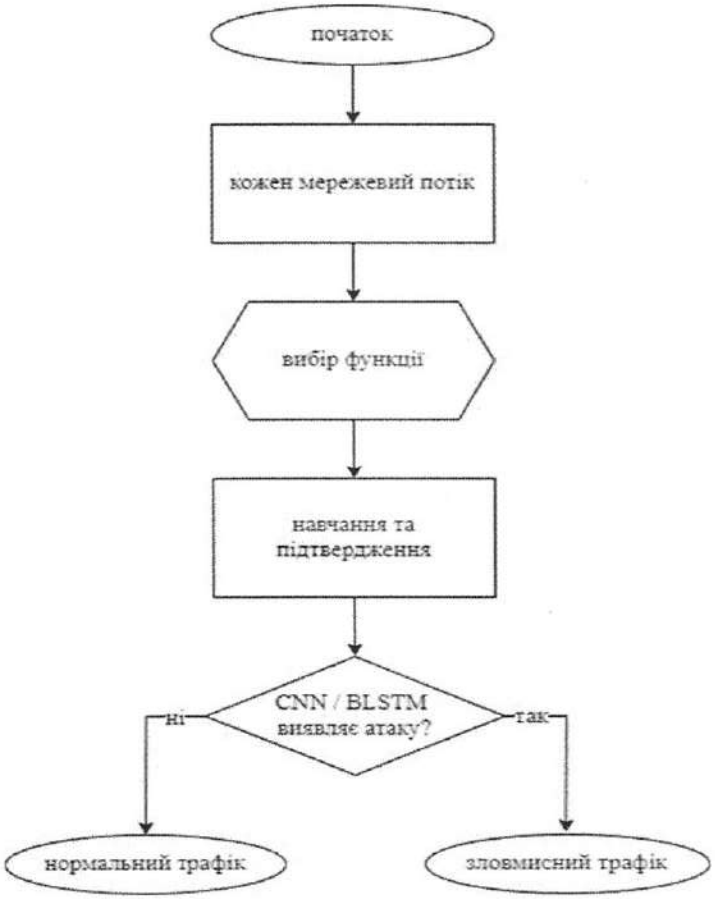
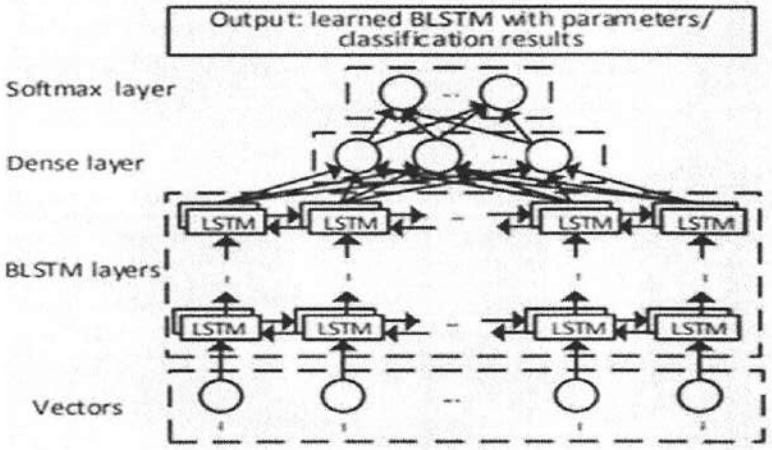
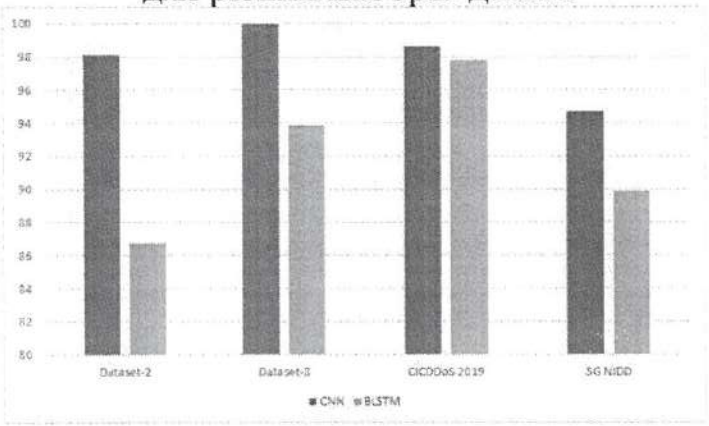


Схема роботи BLSTM мережі

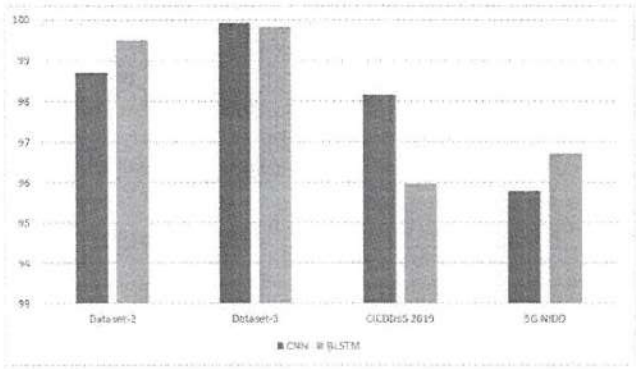


					КРКБ.220161.22.01.04 Е8		
Зм.	Арк.	№ докум.	Підпис	Дата	Системна класифікація DDoS-атак у мережах SGN	Літера	Маса	Масштаб
Розроб.		Куличук М.Р.			Блок-схема процесу класифікації трафіку			
Перевір.		Касянчук М.М.				Аркуш	Аркушів 1	
Т. контр.								
Н. контр.		Мостовий С.В.				ХНУ, КБс-22-1		
Затв.		Клюєв Ю.П.						

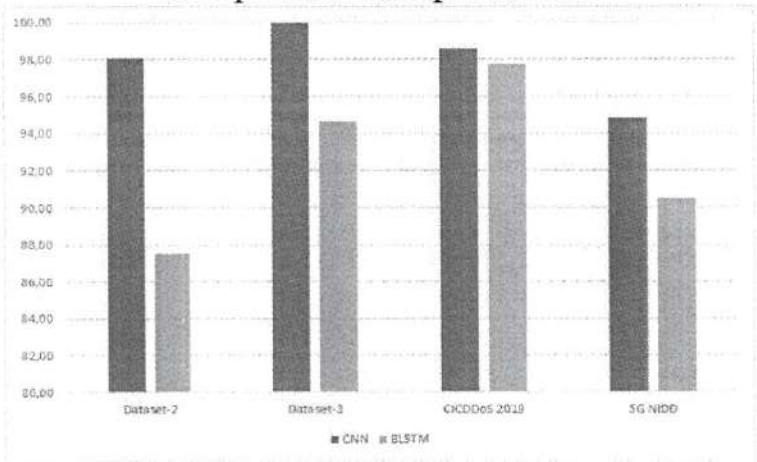
Порівняння акуратності CNN і BLSTM для різних наборів даних



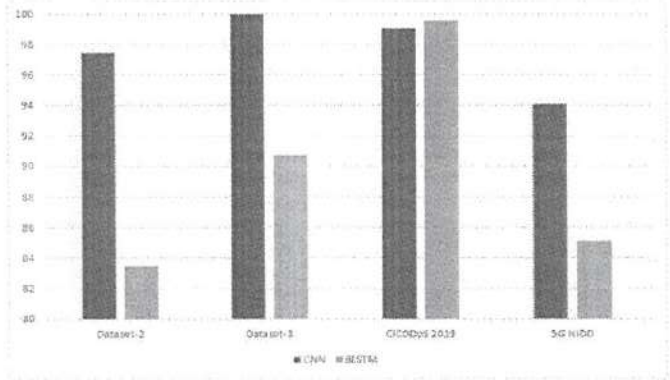
Порівняння точності CNN і BLSTM для різних наборів даних



Порівняння F-міри CNN і BLSTM для різних наборів даних



Порівняння повноти CNN і BLSTM для різних наборів даних



					КРКБ.220161.22.01.04 Е8		
Зм.	Арк.	№ докум.	Підпис	Дата	Системи виявлення DDoS-атак у мережах 5GN. Порівняння метрик CNN і BLSTM для різних наборів даних	Літера	Маса	Масштаб
Розроб.		Кулинич М.Р.				М		
Перевір.		Касинець М.М.				Аркуш	Аркушів 1	
Т. контр.								
Н. контр.		Мостецький С.В.				ХНУ, КБс-22-1		
Затв.		Кльоц Ю.П.						

Завідувачу кафедри кібербезпеки
к.т.н., доц. Кльоцу Ю.П.
Кулачук Марії Романівни
Студентки ФІТ, 3 курсу, групи КБс-22-1

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 дисциплінарної та академічної відповідальності, ознайомлена. Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на наявність академічного плагіату оповіщена та надаю свою згоду на обробку й збереження університетом моєї роботи в інституційному репозитарії Хмельницького національного університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-обчислювального комплексу StrikePlagiarism та/або програмнотехнічного засобу Anti-Plagiarism) і використання роботи для виявлення академічного плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення текстових збігів в роботах.

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

25.05.25
дата

підпис

Anti-Plagiarism (UA) v-15.281 Educational

The maximum coincidence with one document 0.0%

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

ID: 242396 Title: Система виявлення DDoS-атак у мережах 5GN Added in a DB: 2025-05-29 Authors: Кулачук Марія Романівна Heads: Касянчук М.М. Consultants: Opponents:	Document		Sum coincidence on the DB	
	Symbols	Lexemes	Symbols	Lexemes
	85066	593	133 (0%)	3 (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: Кулачук Марія Романівна

Співавтор:

Назва: Система виявлення DDoS-атак у мережах 5GN

Науковий керівник:

Підрозділ: Кафедра кібербезпеки

Коефіцієнт подібності 1:1.8%

Коефіцієнт подібності 2:0%

Мікропробіли: 0

Заміна букв: 0

Інтервали: 0

Білі знаки: 0

Дата створення звіту: 2025-05-29 12:56:53.0

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:

30.05.2025р.

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: Система виявлення DDoS-атак у мережах 5GN

Автор: Кулачук Марія Романівна

Спеціальність: 125 – Кібербезпека

Освітня програма:Кібербезпека

Науковий керівник:Михайло КАСЯНЧУК, докт. техн. наук, професор

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:

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

1) запозичення розміщені в розділах аналізу існуючих методів та технологій, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
2) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
3) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ ідентичності/схожості, складає 1.8%, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи _____ Михайло КАСЯНЧУК

Завідувач кафедри Кб _____ Юрій КЛЬОЦ

Гарант ОП _____ Віктор ЧЕШУН

Дата:

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

РЕЦЕНЗІЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

освітнього ступеня «бакалавр»

Студент Кулачук Марія Романівна

Тема Система виявлення DDoS-атак у мережах 5GN

Спеціальність 125 – Кібербезпека

Обсяг кваліфікаційної роботи освітньо-кваліфікаційного рівня «бакалавр»:

кількість листів креслень 3; кількість сторінок записки 64.

1. Короткий зміст роботи та прийнятих рішень У кваліфікаційній роботі була розроблена система виявлення DDoS-атак у мережах п'ятого покоління (5GN) з використанням глибоких нейронних мереж CNN та BLSTM. У межах дослідження змодельовано мережу 5G за допомогою Free5GC і UERANSIM, сформовано набір даних із легітимного та атакуючого трафіку, розроблено алгоритми класифікації, здійснено навчання моделей та їх тестування. Проведено оцінку ефективності системи в симульованому середовищі, що імітує реальні умови функціонування мережі. Розроблена система дозволяє ідентифікувати DDoS-атаки на основі поведінкового аналізу мережевого трафіку.

2. Висновок про відповідність кваліфікаційної роботи завданню У роботі повністю виконано поставлені завдання, визначені темою та завданням на кваліфікаційну роботу, як у теоретичній, так і в практичній частинах. Розроблено та апробовано ефективну систему виявлення DDoS-атак для мереж 5G з використанням сучасних технологій глибокого навчання.

3. Характеристика виконання кожного розділу роботи, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У вступі обґрунтовано актуальність теми, сформульовано мету та завдання, описано об'єкт, предмет та методи дослідження. У першому розділі подано глибокий аналіз архітектури 5G-мереж, класифікацію DDoS-атак і огляд сучасних методів виявлення атак. У другому розділі розроблено алгоритм виявлення атак, створено датасети на основі 5G-трафіку, описано процес навчання моделей CNN та BLSTM. У третьому розділі проведено оцінку достовірності та ефективності розробленої системи. Робота базується на актуальних наукових підходах, включаючи згорткові та рекурентні нейронні мережі, а також використовує передові симуляційні середовища.

4. Позитивні сторони Робота має високу практичну значущість, оскільки спрямована на розробку інтелектуальної системи виявлення DDoS-атак у мережах 5G — технології, яка активно впроваджується в сучасних інформаційно-телекомунікаційних системах. Запропоновані рішення дозволяють підвищити рівень безпеки критичної інфраструктури, своєчасно виявляти загрози та зменшити ймовірність збоїв і втрат унаслідок кіберінцидентів.

5. Негативні сторони роботи Система потребує значних обчислювальних ресурсів для роботи в режимі реального часу, зокрема при обробці високошвидкісного трафіку та навчанні глибоких моделей. Це може обмежити її використання в середовищах із недостатньою інфраструктурною підтримкою або в умовах енергетичних обмежень.

6. Оцінка графічного оформлення та пояснювальної записки роботи Графічне оформлення кваліфікаційної роботи відповідає темі роботи та виконане з дотриманням стандартів. В цілому, графічне оформлення є якісним, а пояснювальна записка відповідає нормам оформлення.

7. Відгук про роботу в цілому Кваліфікаційна робота заслуговує позитивної оцінки, оскільки весь матеріал роботи є структурованим, чітким та послідовним. Усі розділи роботи мають логічну послідовність, що сприяє зрозумінню викладеного матеріалу в рамках теми роботи. Графічний матеріал допомагає наочно продемонструвати доцільність та ефективність прийнятих рішень для досягнення мети.

8. Інші зауваження

9. Оцінка кваліфікаційної роботи Враховуючи всі позитивні та негативні сторони представленої кваліфікаційної роботи, можна зробити висновок, що вона заслуговує оцінки «задовільно».

РЕЦЕНЗЕНТ (прізвище, ім'я, по батькові, посада, місце роботи)
Бойко Юлій Миколайович,
доктор технічних наук, професор, професор кафедри телекомунікацій, медійних та інтелектуальних технологій

« 04 » червня 2025