Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Факультет комп’ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

Рівень вищої освіти перший (бакалаврський)

Система виявлення аномалій у трафіку

(тема)

Виконав:
здобувач 4 року навчання,
групи КІУКІ-21-6
Кирило СЛАБУНОВ
(власне ім’я, прізвище)

Спеціальність
123 «Комп’ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма
Комп’ютерна інженерія
(повна назва освітньої програми)

Керівник: ас. Ірина КРЮКОВА
(посада, власне ім’я, прізвище)

Допускається до захисту

Завідувач кафедри ЕОМ _____ Андрій КОВАЛЕНКО
(підпис) (власне ім’я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет комп’ютерної інженерії та управління

Кафедра електронних обчислювальних машин

Рівень вищої освіти перший (бакалаврський)

Спеціальність 123 «Комп’ютерна інженерія»
(код і повна назва)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Комп’ютерна інженерія
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
“____” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Слабунову Кирилу Артемовичу
(прізвище, ім’я, по батькові)

1. Тема роботи Система виявлення аномалій у трафіку

затверджена наказом по університету від “ 26 ” травня 2025 р. № 424 Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 17 червня 2025 р.

3. Вхідні дані до роботи
виявлення аномалій, мережевий трафік, розширена згортка, GRU,
увага до каналу, кібербезпека, глибоке навчання, IDS
Pytnon

4. Перелік питань, що потрібно опрацювати у роботі
Основні теоретичні дослідження
Методи штучного інтелекту для виявлення анамалій
Метод виявлення аномалій
Тестування системи
Висновки

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій 12

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання теми кваліфікаційної роботи	26.05	
2	Аналіз літератури	27.05-29.05	
3	Побудова системи	28.05-10.06	
4	Тестування системи та отримання результатів	11.06-12.06	
5	Формування пояснювальної записки	13.06-14.06	
6	Перевірка на плагіат	15.06-17.06	
7	Рецензування роботи	17.06	
8	Подача роботи в ЕК	18.06	
9	Захист роботи	24.06	

Дата видачі завдання " 26 " травня 2025 р.

Здобувач _____
(підпис)

Керівник роботи _____ ас. Ірина КРЮКОВА
(підпис) (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 57 с., 17 рис., 6 табл., 2 дод., 20 джерел.

ВИЯВЛЕННЯ АНОМАЛІЙ, МЕРЕЖЕВИЙ ТРАФІК, РОЗШИРЕНА ЗГОРТКА, GRU, УВАГА ДО КАНАЛУ, КІБЕРБЕЗПЕКА, ГЛИБОКЕ НАВЧАННЯ, IDS.

Метою кваліфікаційної роботи є створення системи виявлення аномалій у трафіку.

У ході виконання кваліфікаційної роботи досліджено проблему виявлення аномального мережевого трафіку в умовах зростаючих кіберзагроз. Запропоновано модель, що поєднує розширену згортку (DC-1D), GRU та механізм уваги до каналу для вилучення просторово-часових ознак та зменшення втрат інформації. Розроблена система дозволяє ефективно класифікувати аномальні патерни трафіку та знижує рівень хибних спрацювань у порівнянні з традиційними методами.

ABSTRACT

Bachelor’s thesis: 57 pages, 17 figures, 6 tables, 2 appendices, 20 sources.

ANOMALY DETECTION, NETWORK TRAFFIC, DILATED CONVOLUTION, GRU, CHANNEL ATTENTION, CYBERSECURITY, DEEP LEARNING, IDS.

The major goal of this thesis is to create a system for detecting anomalies in traffic.

The study addresses the problem of detecting anomalous network traffic amid growing cyber threats. A model combining dilated convolution (DC-1D), GRU, and channel attention is proposed to extract spatio-temporal features and minimize information loss. The developed system effectively classifies anomalous traffic patterns and reduces the false positive rate compared to traditional methods.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

IDS – Intrusion Detection System

DC – Dilated Convolution

DC-1D – One-Dimensional Dilated Convolution

GRU – Gated Recurrent Unit

CNN – Convolutional Neural Network

ML – Machine Learning

DL – Deep Learning

FP – False Positive

ВСТУП

Стрімкий розвиток мережевих технологій супроводжується постійним зростанням різноманітних методів кібератак, що значно підвищує ризик мережевих вторгнень у будь-який час та будь-якому місці. Подібні інциденти спричиняють не лише значні економічні збитки, але й становлять загрозу національній інформаційній безпеці, що свідчить про критичну актуальність проблеми захисту кіберпростору.

Історично перша система виявлення вторгнень (IDS) була представлена у квітні 1980 року в технічному звіті Повітряних сил США. Сучасні IDS класифікуються на два основні типи:

- метод виявлення на основі сигнатур (misuse detection) – передбачає створення бази даних шаблонів відомих загроз. Система аналізує мережеву поведінку та ідентифікує аномалії при збігу з шаблонами;
- метод виявлення аномалій– ґрунтується на аналізі відхилень від нормальної поведінки мережі.

Незважаючи на ефективність сигнатурного підходу, він має суттєві обмеження:

- нездатність виявляти нові типи атак, відсутні у базі даних;
- зростання операційних витрат на оновлення та підтримку бази шаблонів у міру масштабування мережі, що знижує ефективність методу.

На противагу цьому, виявлення аномалій пропонує проактивний підхід, спрямований на ідентифікацію невідомих загроз через аналіз статистичних відхилень. Цей метод є ключовим напрямком сучасних досліджень у галузі кібербезпеки, зокрема для аналізу мережевого трафіку з метою виявлення аномальних паттернів, пов’язаних із кібератаками. Його перевага полягає в здатності адаптуватися до динамічних змін у мережевому середовищі, що робить його перспективним інструментом для протидії складним та еволюційним загрозам.

Щоб подолати проблеми вибору ознак у традиційному машинному навчанні та підвищити точність методів глибокого навчання для виявлення аномального трафіку, ми пропонуємо новий метод. Ця модель інтегрує розширену згортку, GRU та мережу уваги каналу, ефективно поєднуючи розширені згорткові структури з GRU для вилучення як часових, так і просторових ознак для виявлення аномальних закономірностей у мережевому трафіку. Одновимірна структура розширеної згортки (DC-1D) розроблена для розширення рецептивного поля, що дозволяє комплексно вилучати ознаки трафіку, мінімізуючи при цьому втрати інформації, зазвичай спричинені операціями об'єднання. Структура DC фіксує просторові залежності в даних, тоді як GRU обробляє дані часових рядів для фіксації динамічних змін трафіку. Крім того, модуль уваги каналу призначає ваги на основі важливості ознакам у різних каналах, підвищуючи репрезентативну здатність моделі та покращуючи її здатність виявляти аномальний трафік.

1 ОСНОВНІ ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ

1.1 Огляд літературних джерел

З розвитком мережевих технологій постійно з'являються різні методи атак, що призводить до мережевих вторгнень, які можуть відбуватися будь-коли та будь-де. Ці дії можуть завдати значних економічних збитків і навіть загрожувати національній інформаційній безпеці, що підкреслює зростаючу серйозність проблем мережевої інформаційної безпеки. У 1980 року в технічному звіті ВПС США вперше була представлена система виявлення вторгнень (IDS). IDS в основному поділяється на виявлення неправильного використання [1] та виявлення аномалій. Виявлення неправильного використання передбачає створення бази даних аномальних характеристик роботи для моніторингу поведінки мережі; коли поведінка мережі відповідає базі даних, вона ідентифікується як аномальна. Однак виявлення неправильного використання має неминучі недоліки: по-перше, воно не може виявити аномальну поведінку, якої немає в базі даних; по-друге, зі зростанням масштабів мережі зростають витрати на оновлення та підтримку бази даних, що ускладнює виявлення неправильного використання. Виявлення аномалій ідентифікує атаки, виявляючи відхилення від нормальної поведінки мережі, тим самим ефективно запобігаючи мережевим вторгненням.

Виявлення аномалій є критично важливим напрямком досліджень мережевої безпеки, особливо у виявленні аномалій мережевого трафіку, де аналіз даних трафіку використовується для виявлення наявності атак [2].

З розвитком машинного навчання експерти в усьому світі застосували ці методи для виявлення аномалій трафіку. В [3] використовували методи вибору ознак у поєднанні з моделями BN, NB, J48 та DT для виявлення вторгнень. Цей підхід досяг високої швидкості виявлення, що робить його

придатним для мережевих середовищ з високим трафіком. Автори запропонували метод виявлення вторгнень, що поєднує k-NN та метод опорних векторів (SVM), який здатний навчатися та оновлювати нові дані протягом прийнятного періоду часу, а час його прогнозування не зростає швидко під час поступового навчання [4].

В роботі [5] запропонували метод вибору ознак, заснований на вдосконаленому генетичному алгоритмі (ГА), для покращення класифікації ознак.

В досліджені [6] удосконалили традиційний байєсівський метод, використовуючи кілька логарифмічних функцій. Ця оптимізація значно покращила продуктивність байєсівського методу в завданнях виявлення, забезпечуючи ефективніший обчислювальний підхід для байєсівських методів виявлення. В [7] запропонували метод виявлення, заснований на SVM та вбудовуванні наївних байєсівських ознак. У цьому підході вихідні дані були перетворені за допомогою наївного байєсівського методу, а потім класифіковані за допомогою SVM, що покращило швидкість вилучення ознак.

Однак ці моделі машинного навчання можуть неефективно адаптуватися до нових зразків аномального трафіку, і труднощі з налаштуванням можуть виникнути під час процесу навчання. Нещодавно глибокі нейронні мережі досягли широкого впровадження в різних сценаріях аналізу великих даних, що викликало сплеск досліджень у галузі глибокого навчання. Натхненні результатами досліджень у галузі комп'ютерного зору, автори [8] включили залишкові мережі до згорткових нейронних мереж (ЗНМ), досягнувши вищої точності та коефіцієнтів виявлення порівняно з GoogleNet [10] та LeNet-5 [11].

Автори [12] запропонували метод моніторингу аномального трафіку на основі ЗНМ та мереж довгої короткочасної пам'яті (LSTM), який покращив продуктивність системи, але спричинив значні витрати на навчання та тривалий час навчання. Дослідження [13] показує недоліки традиційних

методів виявлення аномалій мережі, такі як поганий вибір ознак та слабке узагальнення, запропонувавши оптимізований метод з LSTM та покращену залишкову нейронну мережу.

Подальші досягнення включають алгоритм виявлення аномалій мережевого трафіку, розроблений авторами [14] на основі згорткової рекурентної нейронної мережі, який ефективно витягує просторові та часові особливості даних мережевого трафіку. Також було інтегровано механізм часової уваги в модель CNN-BiLSTM, скоротивши час навчання та покращивши ефективність прогнозування, дозволяючи більше зосередитися на часових кроках у даних трафіку. В [15] використовували модель ResNet-BiLSTM для виявлення трафіку, досягаючи швидкої класифікації з мінімальними навчальними вибірками на основі попередньої інформації.

1.2 Аспекти метода виявлення аномалій мережевого трафіку

Однак, вищезгадані методи не враховують різну важливість ознак на різних каналах, а безперервні операції згортки та об'єднання можуть призвести до втрати інформації. Для вирішення цих проблем в роботі пропонуємо, метод виявлення аномалій мережевого трафіку, заснований на розширеній згортці та увазі каналу. Модель розроблена для ефективної та точної класифікації інформації про трафік. Запропонована архітектура системи та метод розширюють попередню роботу [16]. Основні нововведення полягають у наступному.

Для запобігання втраті інформації, спричиненій зниженням частоти дискретизації під час процесу згортки, було побудовано модель вилучення ознак на основі розширеної згортки, зокрема розширену згортку-1D (DC-1D). Крім того, було розроблено структуру, що містить розширені згортки зі швидкостями розширення 1, 2 та 3, що забезпечує безперервність ознак під час процесу вилучення інформації.

Одновимірна розширена структура згортки була поєднана з блоками

стробування GRU для вилучення як просторових, так і часових ознак з даних.

Для моделі було розроблено метод оцінки важливості ознакових каналів, використовуючи увагу до каналу. Це пояснюється тим, що використання лише згортки не може виділити різні проблеми на основі важливості ознак, що впливає на ефективність виявлення. Додаючи увагу до каналу до мережі, канали можна зважувати відповідно до їхньої важливості, щоб зосередитися на важливих ознаках та покращити здатність моделі виявляти аномальний трафік.

Комбінація розширених згорток з різними коефіцієнтами дозволяє моделі охоплювати як локальні, так і глобальні просторові залежності в мережевому трафіку, підвищуючи точність виявлення аномалій.

Використання уваги до каналів забезпечує адаптивне підкреслення критичних ознак у реальному часі, що значно зменшує кількість хибно-позитивних спрацювань у порівнянні з класичними підходами.

2 МЕТОДИ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ВИЯВЛЕННЯ АНАМАЛІЙ

2.1 Одновимірна згорткова нейронна мережа

Традиційні моделі згорткових нейронних мереж (ЗНМ) зазвичай складаються з п'яти частин: вхідного шару, згорткового шару, шару об'єднання, повнозв'язного шару та вихідного шару. Структура ЗНМ показана на рисунку 2.1.

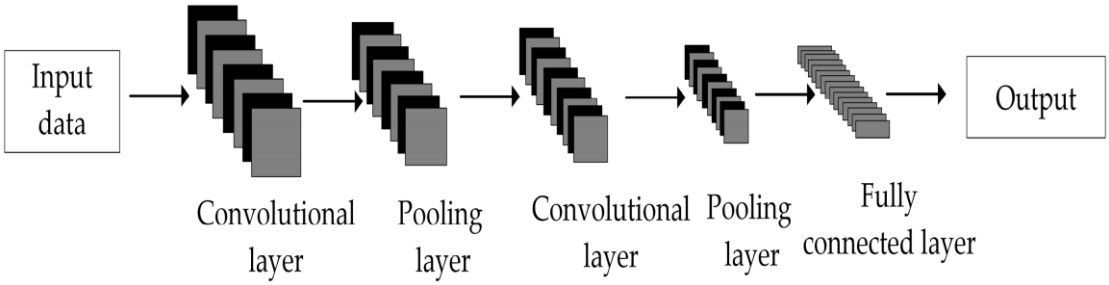


Рисунок 2.1 – Структура CNN

Основна функція згорткового шару полягає у вилученні ознак шляхом застосування згорткових ядер до вхідних даних. У згортковому шарі ваги кожного нейрона є спільними, що значно зменшує кількість параметрів. В одновимірній згортковій нейронній мережі (1D-CNN) згорткове ядро є одновимірним масивом. Вихід згорткового шару:

$$x_j^n = g\left[\sum_{i \in N}\left(x_i^{n-1} k_{ij}^n\right) + b_j^n\right], \tag{2.1}$$

де x^n_j позначає вихід j-го нейрона в n-му шарі; x^{n-1}_i представляє вихід попереднього шару ((n − 1)-го шару), який служить вхідним сигналом для n-го шару; k^n_{ij} означає згорткове ядро від i-го нейрона (n − 1)-го шару до j-го нейрона n-го шару; b^n_j – зміщення для j-го нейрона в n-му шарі; g – функція

активації, зазвичай це функція активації ReLU (випрямлена лінійна одиниця). Функція активації правила показана на рисунку 2.2.

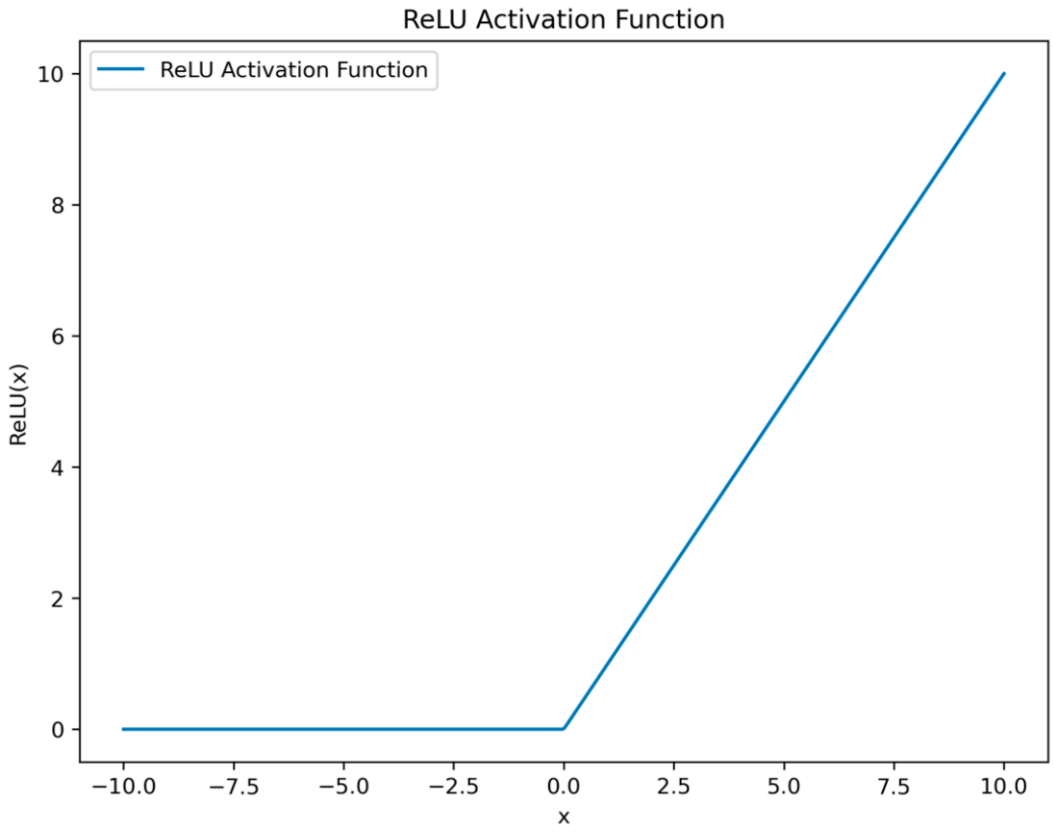


Рисунок 2.2 – Функція активації правила

Операція об'єднання даних стосується зниження частоти дискретизації, зазвичай відомого як максимальне об'єднання даних та усереднення даних. Вирази такі:

$$z_i^n = \text{maxpooling}\left(x_i^{n-1}, s_1, s_2\right), \tag{2.2}$$

$$z_i^n = \text{averapooling}\left(x_i^{n-1}, s_1, s_2\right), \tag{2.3}$$

де $z^n{}_i$ – вихідний сигнал об'єднання i-го нейрона в шарі n; s_1 та s_2 – масштаб об'єднання та розмір кроку відповідно; maxpooling() ·– максимальна функція об'єднання; averpooling() ·– це середня функція об'єднання.

2.2 GRU

У контексті вилучення часових ознак, шлюзови рекурентний блок (GRU) є важливою моделлю нейронної мережі, подібною до моделі довгої короткочасної пам'яті (LSTM), але з деякими оптимізаціями. GRU включає шлюзові структури, такі як шлюз оновлення z_t та шлюзи скидання r_t, які керують оновленням та забуванням інформації. Структура GRU показана на рисунку 2.3.

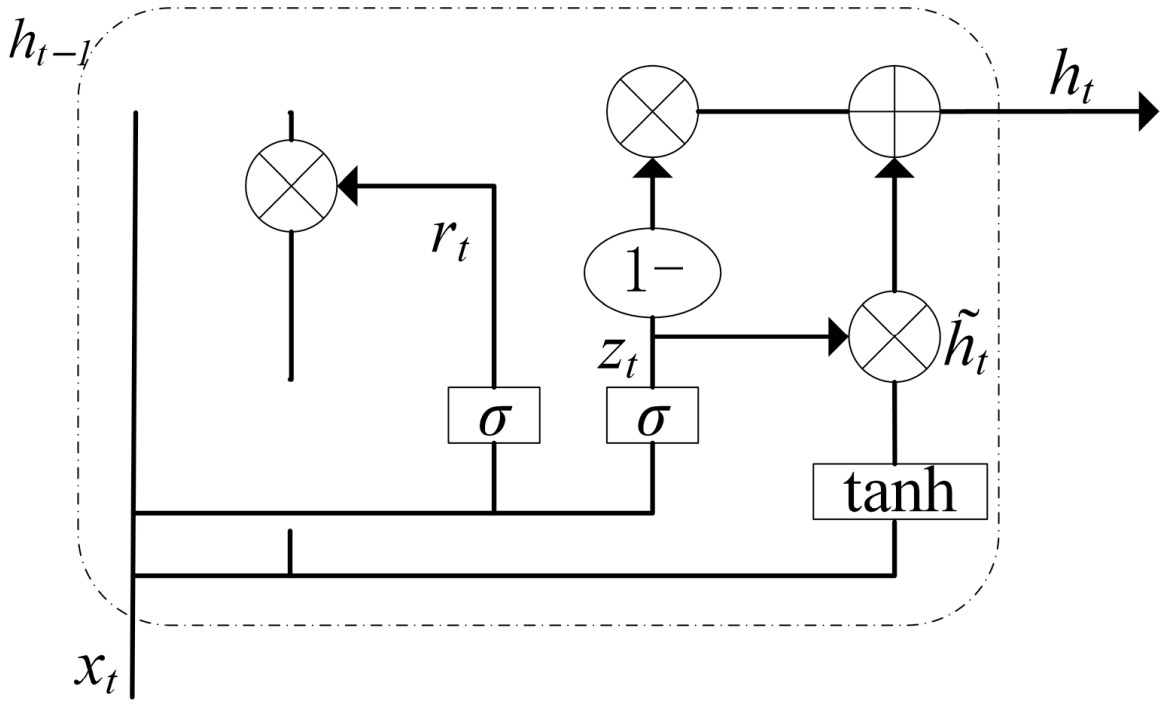


Рисунок 2.3 – Структура GRU

Порівняно з LSTM, GRU є більш лаконічними з точки зору параметрів, що допомагає зменшити складність моделі та покращити швидкість збіжності. Розумна конструкція GRU дозволяє мережі фіксувати довгострокові залежності в послідовностях, одночасно пом'якшуючи проблему зникнення градієнта. Таким чином, GRU добре справляються з багатьма завданнями, що включають послідовну обробку даних, включаючи обробку природної мови, розпізнавання мовлення та машинний переклад, серед іншого.

У мережі GRU, вентиль оновлення z_t визначає, які частини попереднього прихованого стану слід передати на поточний часовий крок на основі поточного вхідного сигналу та прихованого стану з попереднього часового кроку. Вентиль скидання r_t вирішує, яку інформацію в попередньому прихованому стані слід забути, на основі поточного вхідного сигналу та прихованого стану з попереднього часового кроку. Процес обчислення GRU виглядає наступним чином:

$$r_t = \text{sig}\left(W_r \cdot [h_{t-1}, x_t]\right), \tag{2.4}$$

$$z_t = \text{sig}\left(W_z \cdot [h_{t-1}, x_t]\right), \tag{2.5}$$

$$\widetilde{h_t} = \tanh\left(W_{\tilde{h}} \cdot [r_t \times h_{t-1}, x]\right), \tag{2.6}$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot h_t, \tag{2.7}$$

де r_t представляє логічний елемент скидання, z_t представляє логічний елемент оновлення, $\text{sig}(\cdot)$·– сигмоїдна функція активації, W_r – вагова матриця для вентиля скидання, W_z – вагова матриця для вентиля оновлення, $\widetilde{h_t}$ – це прихований стан-кандидат, а h_t – прихований стан на кроці часу t.

2.3 Механізм уваги до каналу

Механізм уваги розроблений для імітації людського зору. Під час обробки візуальної інформації людські очі вибірково ігнорують неважливі деталі та зосереджуються більше на значущій інформації, що дозволяє ефективніше розподіляти ресурси між життєво важливими завданнями. Цей

механізм може оптимізувати модель згорткової нейронної мережі (CNN) та підвищити її продуктивність. Зокрема, увага до каналу здатна отримувати інформацію про канал та встановлювати зв'язки між каналами, тим самим покращуючи здатність моделі розпізнавати ознаки.

Якщо дані, які нам потрібно обробити, мають вигляд $X \in R^{C \times H \times W}$, C – кількість каналів, а H та W – висота та ширина карти ознак, тоді формула для уваги до каналу така:

$$a = \sigma\left(F\left(GAP\left(X\right)\right)\right), \tag{2.8}$$

де a – вектор уваги, а σ – сигмоподібна функція активації. F – повнозв'язний шар, а GAP – глобальне об'єднання. Після отримання уваги каналу a, Y використовується для представлення уваги, яку слід приділити кожному каналу вхідних даних X:

$$Y = a_I \cdot X_I, \tag{2.9}$$

де a_i представляє і-й елемент у векторі уваги, а X_i представляє і-й канал у вхідних даних.

3 МЕТОД ВИЯВЛЕННЯ АНОМАЛІЙ

3.1 Загальна структура моделі

Модель складається з трьох частин: обробка даних, вилучення ознак та класифікація результатів, як показано на рисунку 3.1.

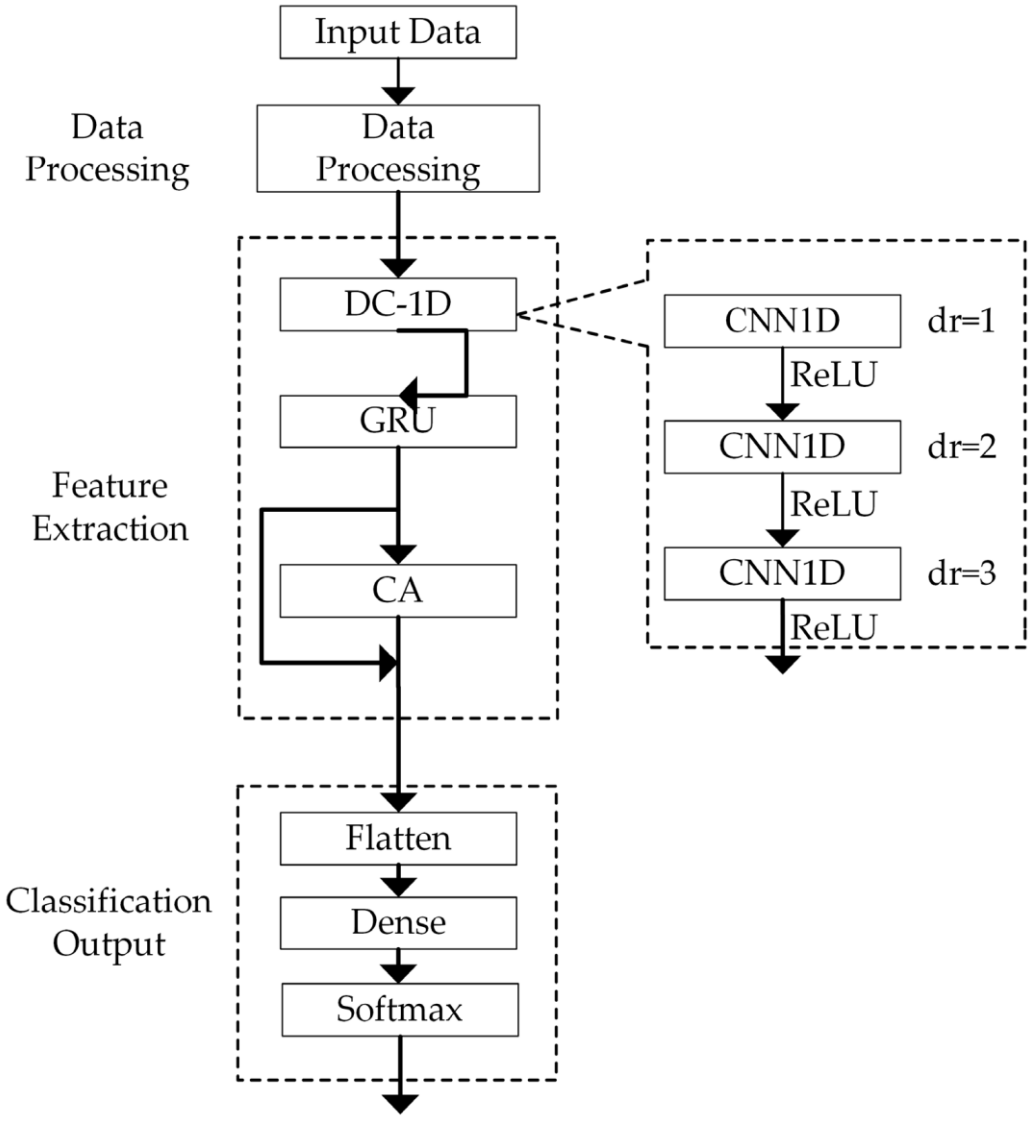


Рисунок 3.1 – Загальна структура моделі

Попередня обробка даних проводиться, оскільки різні інтервали ознак у векторі ознак трафіку мають різні розмірності. Щоб пом'якшити вплив цих

відмінностей у розмірності на результати класифікації, ознаки спочатку слід нормалізувати. Нормалізація середнього значення та дисперсії використовується для перетворення даних у нормальний розподіл із середнім значенням 0 та дисперсією 1. Формула для нормалізації така:

$$x' = \frac{x - x_m}{s}, \tag{3.1}$$

де x′ – оброблені дані, x – характеристичне значення кожної групи вхідних ознак, x_m– середнє значення вхідних ознак, а s – стандартне відхилення кожної групи вхідних ознак.

Потім, на етапі вилучення ознак, попередньо оброблені дані вводяться в DC-1D. Модуль DC-1D складається з трьох згорткових шарів зі швидкостями розширення 1, 2 та 3. У DC-1D вхідні дані послідовно проходять через три розширені згорткові шари для вилучення ознак, щоб створити карти ознак, причому кожна згортка обробляється за допомогою функції активації ReLU. Потім модуль GRU використовується для вилучення її часових характеристик. Вилучені ознаки вводяться в модуль CA, де дані зважуються за каналами через модуль CA.

На етапі виведення класифікації зважені ознаки вводяться в шар Flatten для вирівнювання, а потім обробляються шаром Dense та шаром Softmax для цілей класифікації.

3.2 Структура DC-1D

Щоб уникнути втрати інформації, спричиненої операціями об'єднання в традиційних згортках, у цій роботі замість операцій об'єднання використовуються розширені згортки. Розширені згортки можуть розширювати рецептивне поле без зміни роздільної здатності ознак, тим самим покращуючи швидкість обчислень. Якщо розмір традиційного ядра

згортки дорівнює p, а швидкість розширення дорівнює r, то розмір розширеного ядра згортки p' задається:

$$p' = (r - 1)(p - 1) + p \quad , \tag{3.2}$$

У цьому випадку рецептивні поля традиційної згортки та порожньої згортки відповідно:

$$R_{i+1} = R_i + (p - 1)s \quad , \tag{3.3}$$

$$R'_{i+1} = R'_i + (p' - 1)s \quad , \tag{3.4}$$

де R_i та R_{i+1} представляють рецептивні поля поточного та наступного шарів у традиційній згортці відповідно. Терміни R'_i та R'_{i+1} позначають рецептивні поля поточного та наступного шарів у розширеній згортці відповідно. Крім того, s – це добуток кроків від першого шару до (i+1)-го шару. Підставляючи рівняння перше у рівняння наступне, отримуємо:

$$R'_{i+1} = R'_i + [((r - 1)(p - 1) + p) - 1]s \quad , \tag{3.5}$$

З цього видно, що рецептивне поле швидко зростає під час виконання багатошарових операцій з розширеними згортками, що дозволяє отримати значно більше рецептивне поле.

Оскільки розширені згортки можуть призвести до втрати безперервності між ознаками, щоб уникнути цього та досягти багатомасштабного збору інформації, розширені згортки зі швидкостями розширення 1, 2 та 3 об'єднуються для формування структури DC-1D, як

показано на рисунку 3.2.

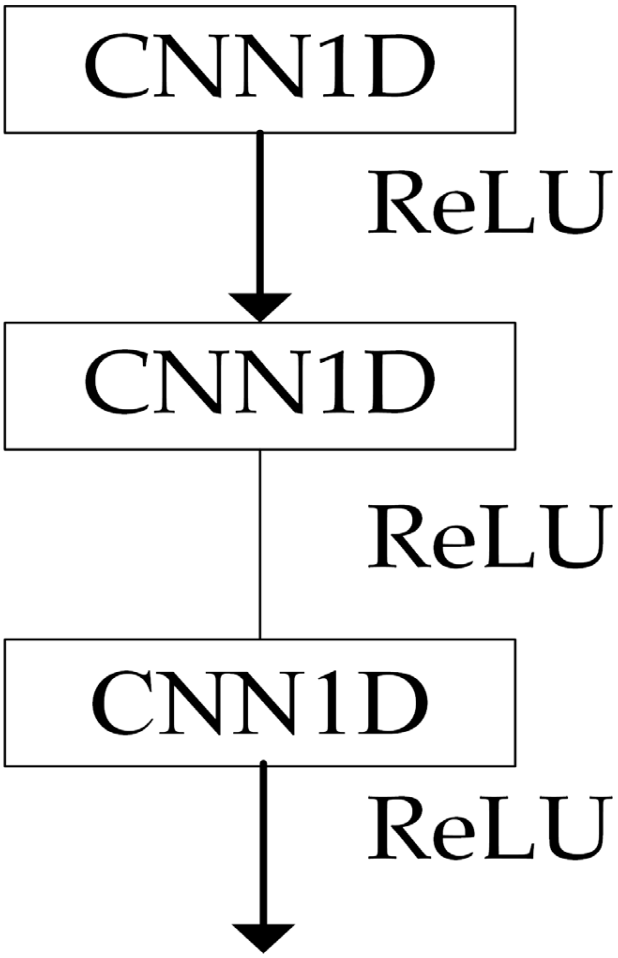


Рисунок 3.2 – Структура DC-1D

Одновимірні дані, що складаються з кожної характеристики мережевого трафіку, беруться як вхідні дані x. Вихідні дані після обробки через структуру DC-1D видаються за формулою:

$$F_c = (Conv_{r3}\,(Conv_{r2}\,(Conv_{r1}\,(x)))) \quad , \tag{3.6}$$

де $Conv_{r1}$, $Conv_{r2}$, та $Conv_{r3}$ представляють розширені згортки зі швидкостями розширення 1, 2 та 3 відповідно. Вихідна карта ознак позначається як F_c.

3.3 DC-GRU

Робочий процес моделі DC-GRU починається з попередньої обробки набору даних. Попередньо оброблені дані потім вводяться в модель CNN для вилучення просторових ознак. Згодом ці просторові ознаки подаються в GRU для вилучення часових ознак. Детальний процес реалізації проілюстровано на на рисунку 3.3, а архітектура моделі CNN-GRU зображена на рисунку 3.4.

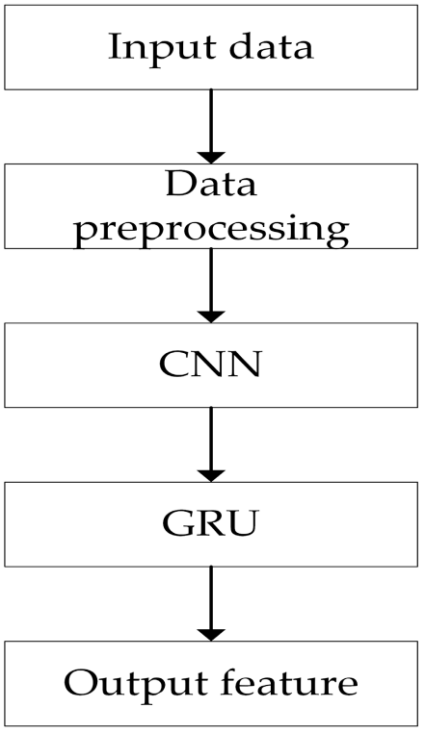


Рисунок 3.3 – Структурна блок-схема

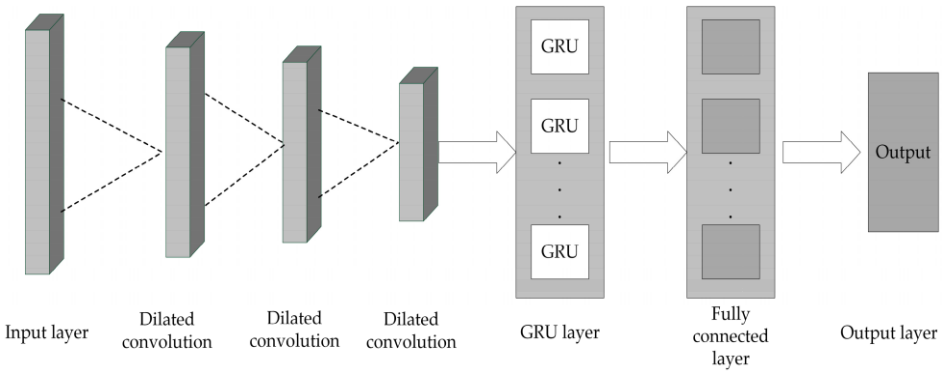


Рисунок 3.4 – Модель DC-GRU

3.4 Модуль CA

Через те, що різні дані трафіку можуть демонструвати однакові ознаки, важливість цих ознак для класифікації трафіку різна. Існуючі алгоритми виявлення аномалій трафіку не враховують ступінь, до якої ознаки трафіку в різних каналах впливають на результати класифікації. Для вирішення цієї проблеми використовується модуль CA (channel attention - увага до каналу), як показано на рисунку 3.5.

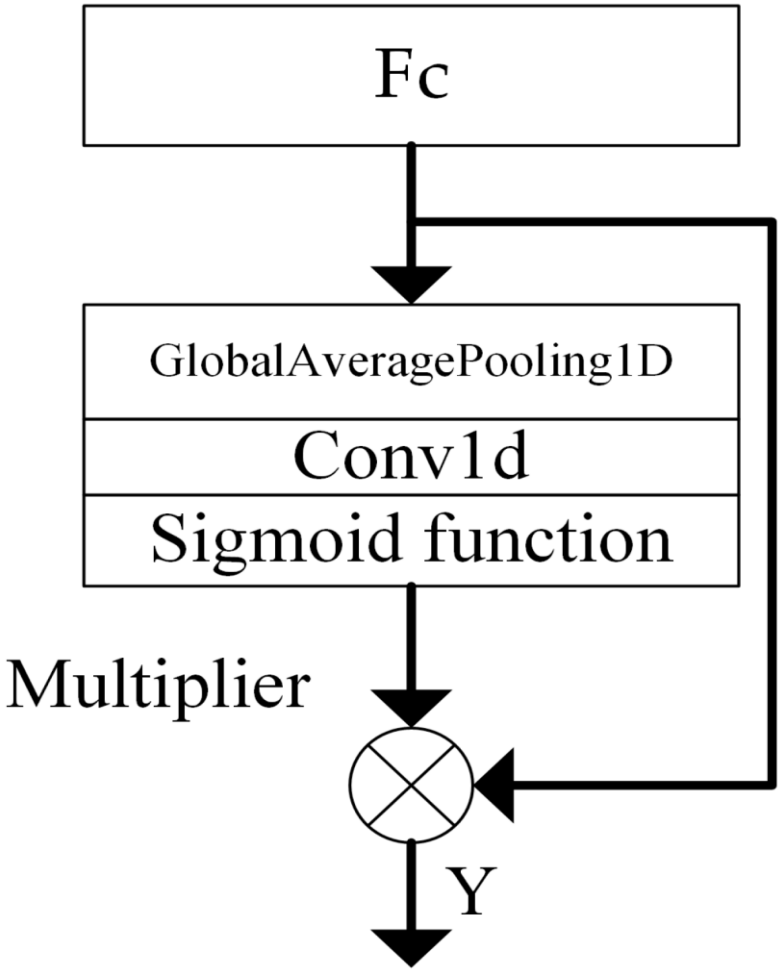


Рисунок 3.5 – Модуль CA

Модуль CA може адаптивно визначати ядро згортки k за допомогою одновимірної згортки, тим самим уникаючи необхідності ручного налаштування ядра згортки. Адаптивний розрахунок ядра згортки k виглядає

наступним чином:

$$k = \left| \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right|_{odd}, \quad (3.7)$$

де C – розмірність каналу, γ = 2, b = 1, де непарне число позначає непарне число.

4 ТЕСТУВАННЯ СИСТЕМИ

4.1. Експериментальне середовище

Експерименти в цій роботі проводилися в операційній системі Windows з використанням фреймворку глибокого навчання TensorFlow. Конкретне апаратне та програмне середовище детально описано в таблиці 4.1.

Таблиця 4.1 – Експериментальне середовищ

Експериментальне середовище	Конфігурація обладнання
Операційна система	Windows 11 64-біт
Процесор	Intel i7-12700H
Графічний процесор	NVIDIA GeForce RTX3070Ti
Мови програмування	Python 3.12
Прикладне програмне забезпечення	Pycharm 2020.1

4.2 Набори даних та показники оцінювання

Для навчання та валідації було використано набір даних CIC-IDS-2017. Цей набір даних містить дані про трафік з 3 липня 2017 року по 7 липня 2017 року, включаючи звичайний трафік та 14 типів атак. В експерименті враховувався звичайний трафік (BENIGN) та дев'ять типів атак. Співвідношення навчального набору до тестового набору становило 7:3.

Набір даних CIC-IDS-2017 містить загалом 79 ідентифікаторів із 78 ознаками, а останній ідентифікатор використовувався для позначення того, чи відображають дані нормальну чи аномальну поведінку.

Матрицю плутанини було використано як показник оцінки, як показано в таблиці 4.2. У матриці похибки TP (істинно позитивний результат) – це кількість випадків аномального дорожнього руху, правильно

ідентифікованих як аномальні; TN (істинно негативний результат) – це кількість випадків нормального дорожнього руху, правильно ідентифікованих як нормальні; FP (хибнопозитивний результат) – це кількість випадків аномального дорожнього руху, неправильно ідентифікованих як нормальні; та FN (хибнонегативний результат) – це кількість випадків нормального дорожнього руху, неправильно ідентифікованих як аномальні.

Таблиця 4.2 – Матриця похибки

Матриця похибки		Прогнозоване значення	
		Позитивний клас	Негативний клас
Справжнє значення	Позитивний клас	TP	FN
	Негативний клас	FP	TN

Для оцінки ефективності моделі використовувалися такі показники, як точність, прецизійність, повнота та F1-оцінка. Формули мають наступний вигляд:

$$\text{Accuracy} = \frac{\text{TP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{4.1}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{4.2}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{4.3}$$

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (4.4)$$

4.3 Попередня обробка даних

Під час обробки даних мережевого трафіку присутні нечислові ознаки, які не підходять для безпосереднього введення в моделі виявлення аномалій. Щоб вирішити цю проблему, було використано гаряче кодування для перетворення цих нечислових ознак на числові, що робить їх сумісними з моделлю. Крім того, можуть існувати значні відмінності в масштабах різних ознак. Такі розбіжності можуть впливати на ефективність класифікації, що вимагає нормалізації значень ознак. Нормалізація гарантує, що всі ознаки знаходяться в порівнянній шкалі, тим самим забезпечуючи більш справедливі порівняння в межах моделі.

Наприклад, у загальнодоступному наборі даних KDDCUP99 ознаки «service», «flag» та «protocol_ty» не є числовими, тому для перетворення на двійкові числові значення потрібно кодувати за допомогою одноразового кодування. Коли ознака «protocol_type», яка включає три атрибути — TCP, UDP та ICMP – кодується за допомогою одноразового кодування, результати становлять 100, 010 та 001 відповідно. Ознака «service» містить 70 атрибутів і після одноразового кодування перетворюється на 70-вимірний двійковий вектор. Ознака «flag» містить 11 атрибутів, а її одноразове кодування дає 11-вимірний двійковий вектор.

Більше того, значні відмінності в діапазонах значень ознак можуть призвести до повільної конвергенції моделі виявлення аномалій, збільшуючи експериментальний час і зрештою впливаючи на продуктивність виявлення. Тому нормалізація ознак є необхідною. Наприклад, у наборі даних CIC-IDS-2017 ознака «Загальна довжина пересилання пакетів» коливається від [0 181

012], тоді як ознака «Максимальний потік IAT» коливається від [-1 120 000 000], що вказує на суттєві розбіжності. Щоб пом'якшити проблеми, що виникають через ці відмінності, виконується нормалізація для уніфікації діапазону кожної ознаки в інтервалі [0, 1]. Цей процес стабілізує модель і прискорює конвергенцію, тим самим заощаджуючи експериментальний час і підвищуючи продуктивність виявлення. У цьому дослідженні було використано нормалізацію середнього значення та дисперсії, як визначено рівнянням (10). Обчислюючи середнє значення та стандартне відхилення кожної ознаки, значення були перетворені відповідно до стандартизованого розподілу. Отже, незалежно від їх початкового діапазону, значення ознак були ефективно зіставлені зі спільним числовим інтервалом, забезпечуючи більш надійну основу для навчання та прогнозування моделі.

4.4 Налаштування параметрів моделі

Щоб вирішити потенційні проблеми вибуху градієнтів або зникнення градієнтів у глибоких згорткових нейронних мережах (ЗНМ), було проведено експерименти на ЗНМ з різною кількістю шарів (від одного до п'яти) для порівняння їхньої точності. Усі інші параметри в експериментах залишалися постійними, окрім кількості згорткових шарів. Результати експериментів показано на рисунку 4.1.

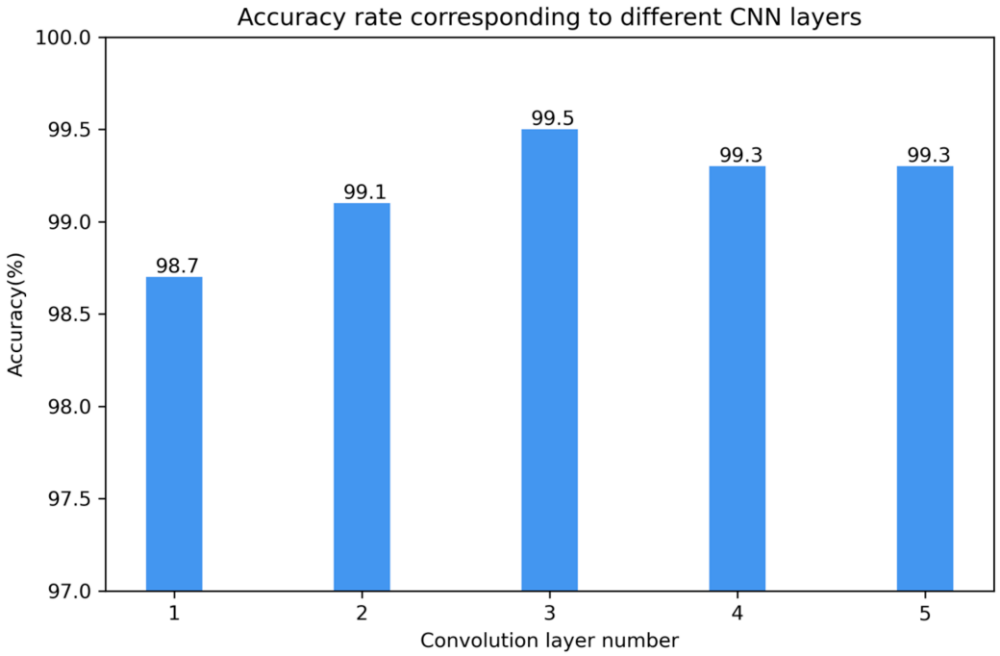


Рисунок 4.1 – Точність, що відповідає різній кількості згорткових шарів

З результатів на рисунку можна зробити висновок, що продуктивність була найгіршою, коли кількість згорткових шарів становила один, а найкраща – коли їх було три, з точністю 99,5%. Отже, ця модель використовує мережеву структуру з трьома згортковими шарами.

Хоча розширена згортка може розширити рецептивне поле та захопити більше контекстної інформації без збільшення кількості параметрів, поширена проблема, відома як «ефект сітки», може виникнути, коли швидкість розширення розширеної згортки велика. Цей ефект виникає, коли згорткове ядро охоплює лише частину вхідної карти ознак, що призводить до розривів у вихідній карті ознак з точки зору просторового розташування. Цей розрив може призвести до втрати інформації або фрагментації, що впливає на продуктивність моделі. Щоб вирішити цю проблему, ми використали гібридну розширену згортку [19] (ГРЗ) та визначили швидкості розширення для розширеної згортки на основі тришарової згортки. ГРЗ має три характеристики. По-перше, картина швидкості розширення має зигзагоподібну конфігурацію. По-друге, швидкості розширення послідовно розташованих ГРЗ не повинні мати спільного коефіцієнта, більшого за 1. Нарешті, ГРЗ задовольняє формулу:

$$M_i = \max\left[M_{i+1} - 2r_i, M_{i+1} - 2\left(M_{i+1} - r_i\right), r_i\right], \quad (4.5)$$

де r_i представляє швидкість розширення і-го шару. M_i представляє найбільшу швидкість розширення і-го шару.

Швидкості розширення були встановлені таким чином, щоб задовольняти рівняння для швидкостей розширення 1, 2 та 3, а також 1, 2 та 5. Експерименти проводилися, встановлюючи швидкості розширення на 1, 2 та 3, а також 1, 2 та 5, зберігаючи при цьому інші умови постійними. Результати експериментів показано на рисунку 4.2. З рисунку 4.2 видно, що коли швидкості розширення були встановлені на 1, 2 та 3, точність досягла 99,5%, що значно краще, ніж коли швидкості розширення становили 1, 2 та 5.

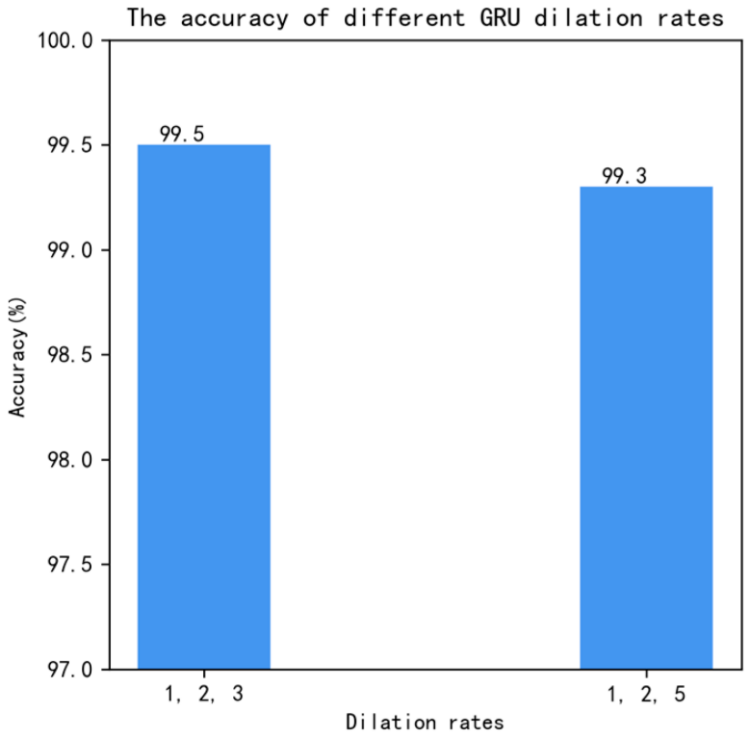


Рисунок 4.2 – Швидкість дилатації

У моделях глибокого навчання кількість шарів у нейронній мережі може суттєво впливати на кінцеву продуктивність класифікації, що особливо помітно при використанні мереж GRU. Тому, щоб визначити оптимальну

кількість шарів, усі інші параметри в експериментах залишалися постійними, і коригувалася лише кількість шарів GRU. Різні конфігурації шарів GRU були поєднані зі структурою DC-1D (одновимірна розширена згортка), зокрема, включаючи GRU з 1, 2, 3 та 4 шарами. Ця комбінація дозволила дослідити здатність GRU обробляти складні ознаки та зберігати інформацію про послідовності, спостерігаючи за її впливом на загальну продуктивність моделі. Кожна конфігурація шарів пройшла однаковий процес навчання та перевірки, щоб забезпечити порівнянність результатів. Результати експериментів показано на рисунку 4.3.

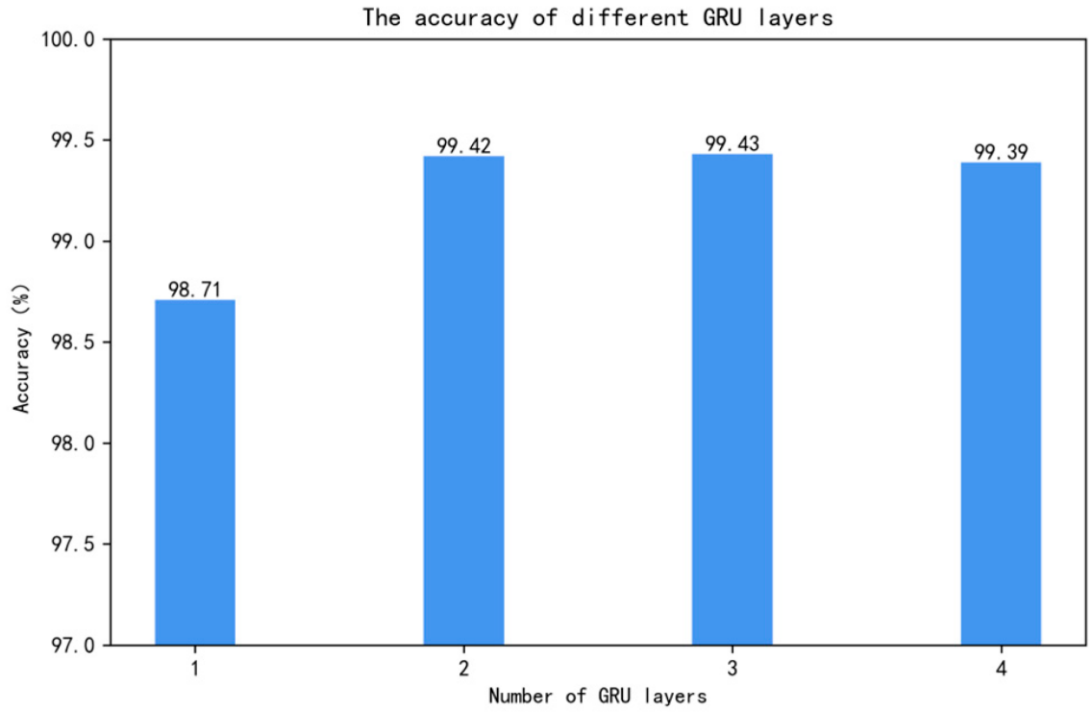


Рисунок 4.3 – Точність, що відповідає різній кількості шарів GRU

Враховуючи, що обчислювальні витрати та ефективність моделей глибокого навчання є вирішальними факторами в розробці та виборі моделі, незначна різниця в точності між використанням двошарового та тришарового GRU стала критичним моментом прийняття рішення. Хоча тришаровий GRU забезпечував дещо вищу точність, це також призводило до збільшення обчислювальних витрат та складності. Тому, після балансування точності та обчислювальних витрат, рішення про використання двошарового GRU як

остаточної моделі було прийнято на основі всебічного врахування продуктивності та ефективності. Цей вибір не тільки забезпечив високу точність, але й забезпечив ефективність та практичність моделі.

Оскільки вибір нейронів GRU також впливає на продуктивність моделі, були проведені експерименти для визначення оптимальної кількості нейронів. Двошаровий GRU був встановлений на 16, 32, 64 та 128 нейронів, експериментальні результати для яких показані на рисунку 4.4. Результати показують, що найвища точність була досягнута, коли кількість нейронів була встановлена на рівні 64-64.

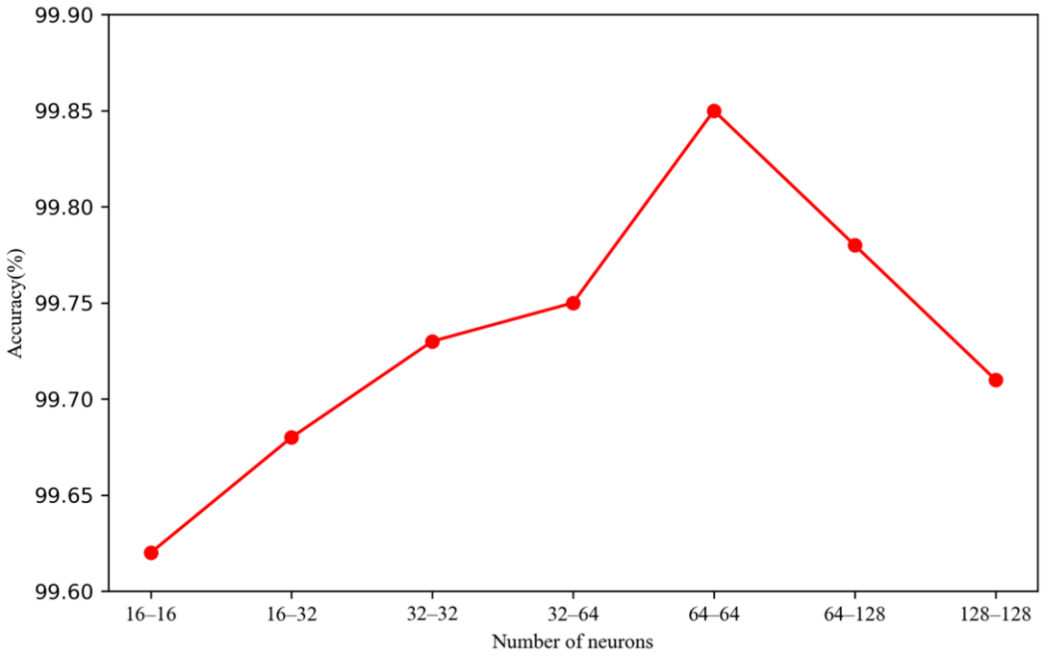


Рисунок 4.4 – Кількість нейронів

За допомогою вищезгаданого експерименту з проектування параметрів ми розробили структуру DC-GRU. Загальна структура складається з трьох одновимірних розширених згорток зі швидкостями розширення 1, 2 та 3. Загалом, для досягнення кращої точності дослідники схильні проектувати мережеві моделі з більшою глибиною для отримання ознак вищого рівня. Однак, чим глибші шари, тим більше параметрів задіяно. Теоретично, чим більше параметрів, тим краща апроксимаційна здатність моделі, але це також

збільшує ризик перенавчання. На основі численних експериментів розмір ядра згортки та кількість ядер згортки в цьому дослідженні були встановлені на рівні 3 та 64 відповідно.

У розділі GRU під час експериментів було виявлено, що відповідне збільшення глибини мережі може покращити прогностичні можливості моделі. У цьому дослідженні було використано два шари GRU з вихідною розмірністю 64.

4.5 Перевірка ефекту моделі

Під час експерименту модель сходилася, коли епоху було встановлено на 10, і на цьому етапі точність тестового набору досягла 99,5%. У таблиці 4.3 наведено значення точності, повноти та F1-оцінки для різних типів атак у наборі даних CIC-IDS-2017.

Таблиця 4.3 – Значення точності, повноти та F1 для набору даних CIC-IDS-2017

Тип потоку	Precision	Recall	F1
BENIGN	99.8%	99.5%	99.7%
Bot	93.4%	62.7%	75.1%
DDoS	99.9%	99.7%	99.8%
DoS GoldenEye	99.6%	99.3%	99.5%
DoS Hulk	98.5%	99.8%	99.1%
DoS Slowhttptest	97.7%	98.7%	98.2%
DoS slowloris	99.1%	98.6%	98.8%
FTP-Patator	99.8%	99.5%	99.7%
PortScan	96.5%	99.1%	97.8%
SSH-Patator	97.1%	98.0%	97.5%

Як видно з таблиці 4.3, точність DCGCANet для 10 типів мережевої

поведінки була вище 90%, а для типів ботів з меншою кількістю даних вона також перевищувала 92%. Коефіцієнт відтворення та оцінка F1 також залишалися на високих рівнях, що достатньо демонструє ефективність DCGCANet у класифікації трафіку.

Для більш інтуїтивної демонстрації продуктивності моделі було створено теплову карту, яка порівнює точність, повноту та показник F1 для різних типів атак, як показано на рисунку 4.5.

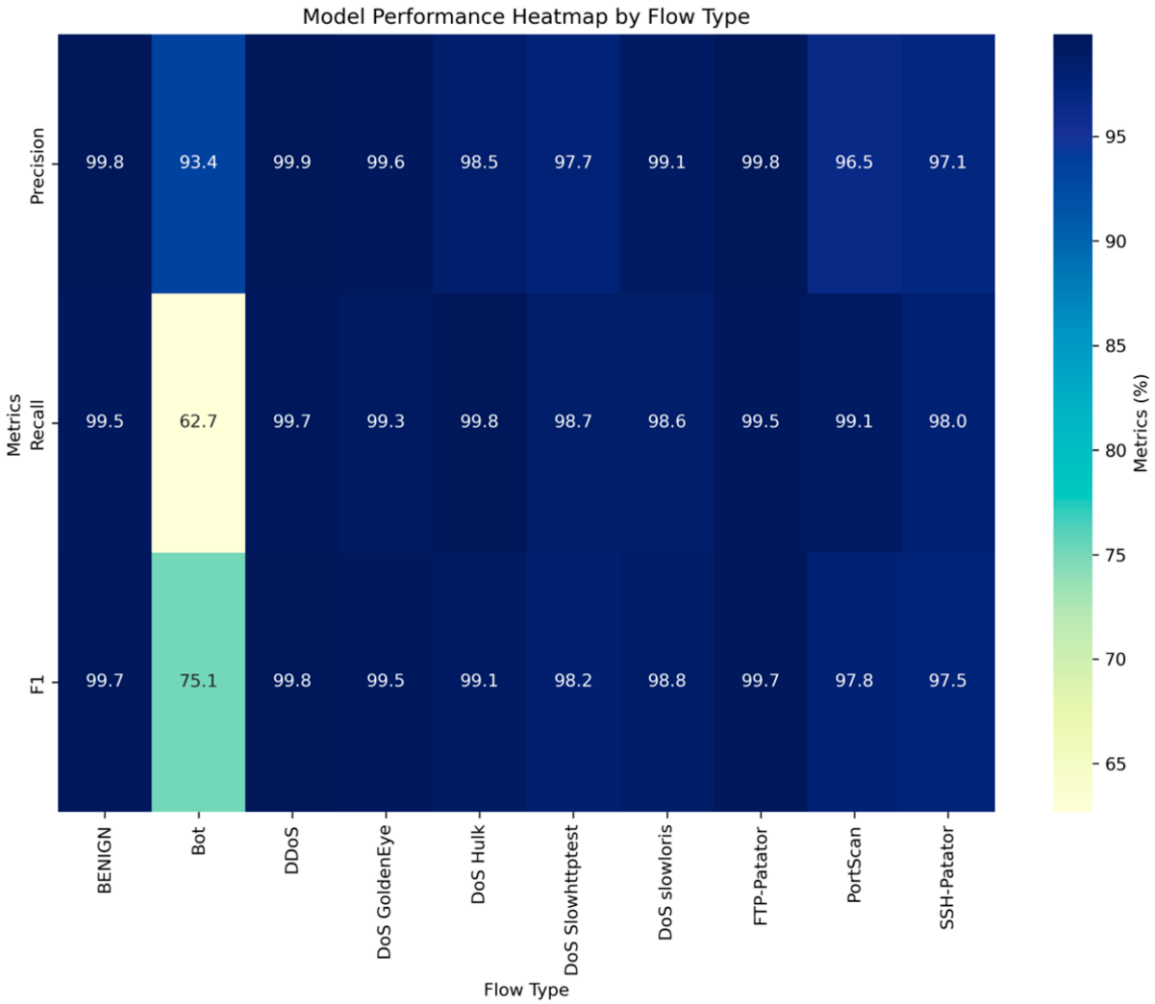


Рисунок 4.5 – Кольорова карта продуктивності моделі за типом потоку

Для подальшої перевірки ефективності моделі DCGCANet та її компонентів, експерименти з абляції були проведені в тих самих експериментальних умовах. Результати експериментів показано на рисунках 4.6-4.9. У цьому експерименті різні модулі або моделі були окремо

відключені, щоб перевірити вплив та ефективність запропонованої моделі шляхом порівняння. На рисунку 4.6 показано результати матриці плутанини для моделі CNN-1D, на рисунку 4.7 представлені результати для моделі DC-GRU, на рисунку 4.8 показано результати для комбінованих модулів DC та CA, а на рисунку 4.9 проілюстровано експериментальні результати для моделі DCGCANet.

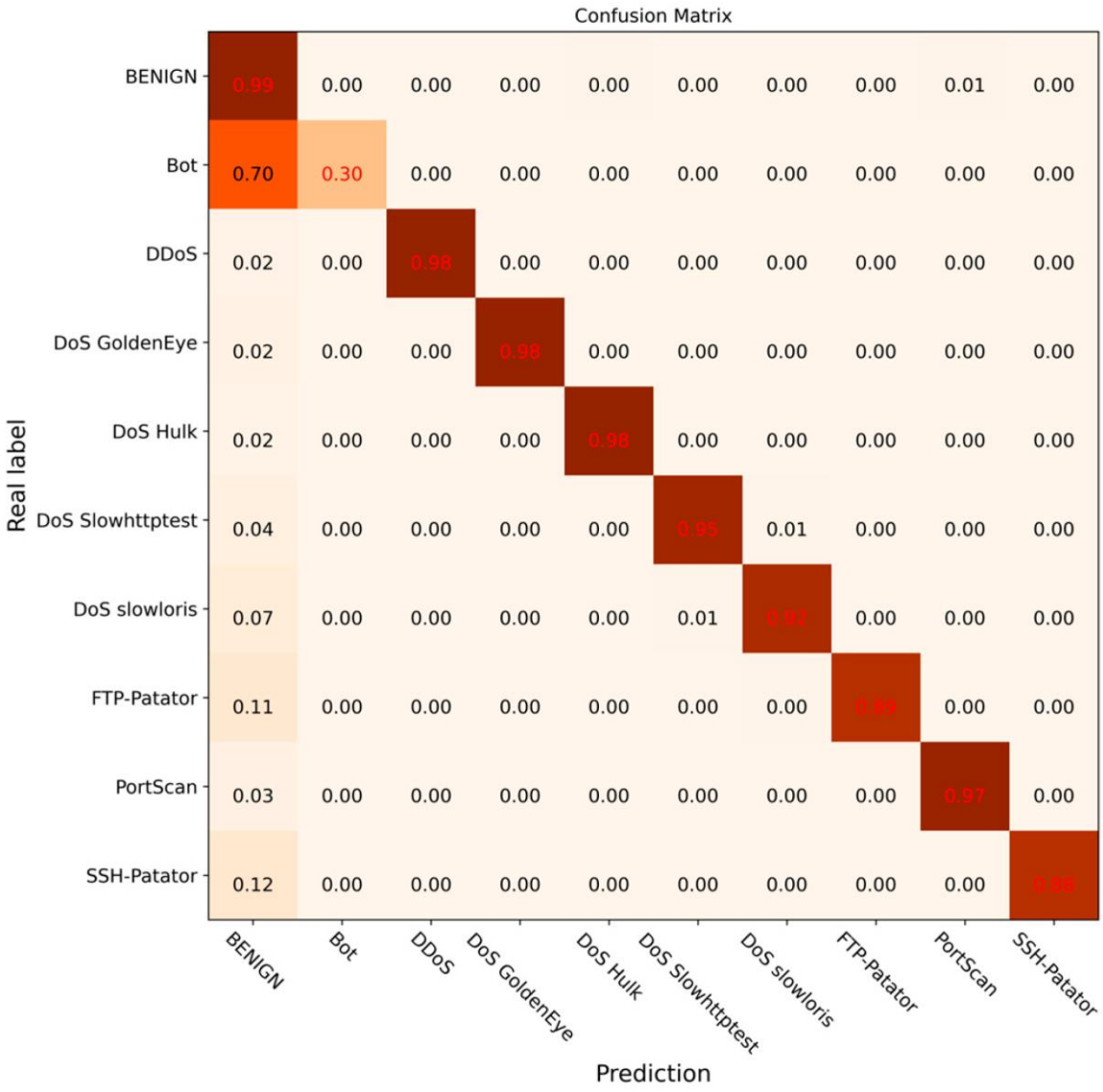


Рисунок 4.6 – Матриця похибки результатів класифікації для CNN-1D

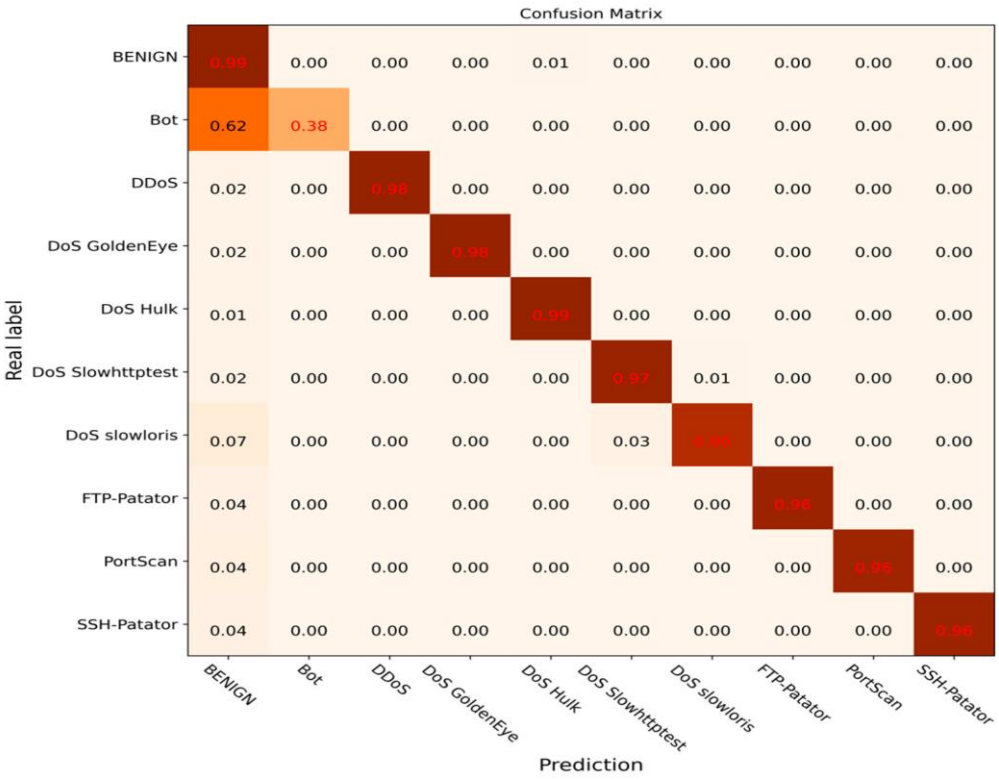
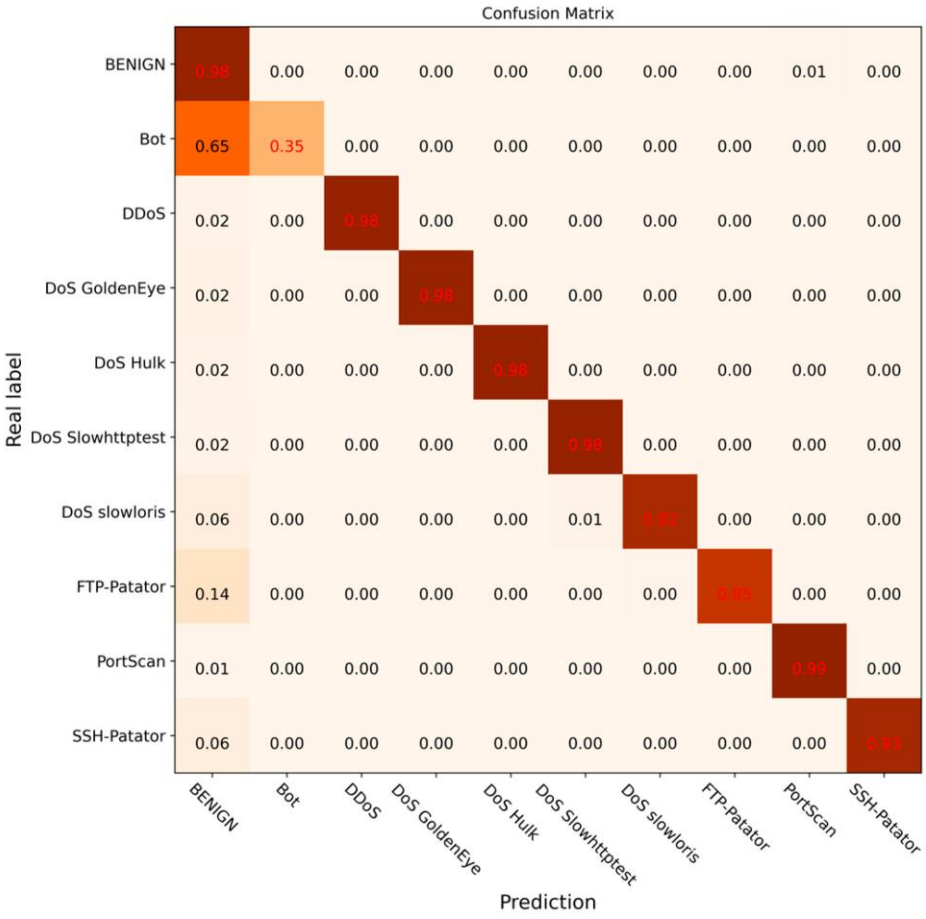


Рисунок 4.7 – Матриця похибки результатів класифікації DC-GRU



Рисувнок 4.8 – Матриця похибки результатів класифікації модуля CA

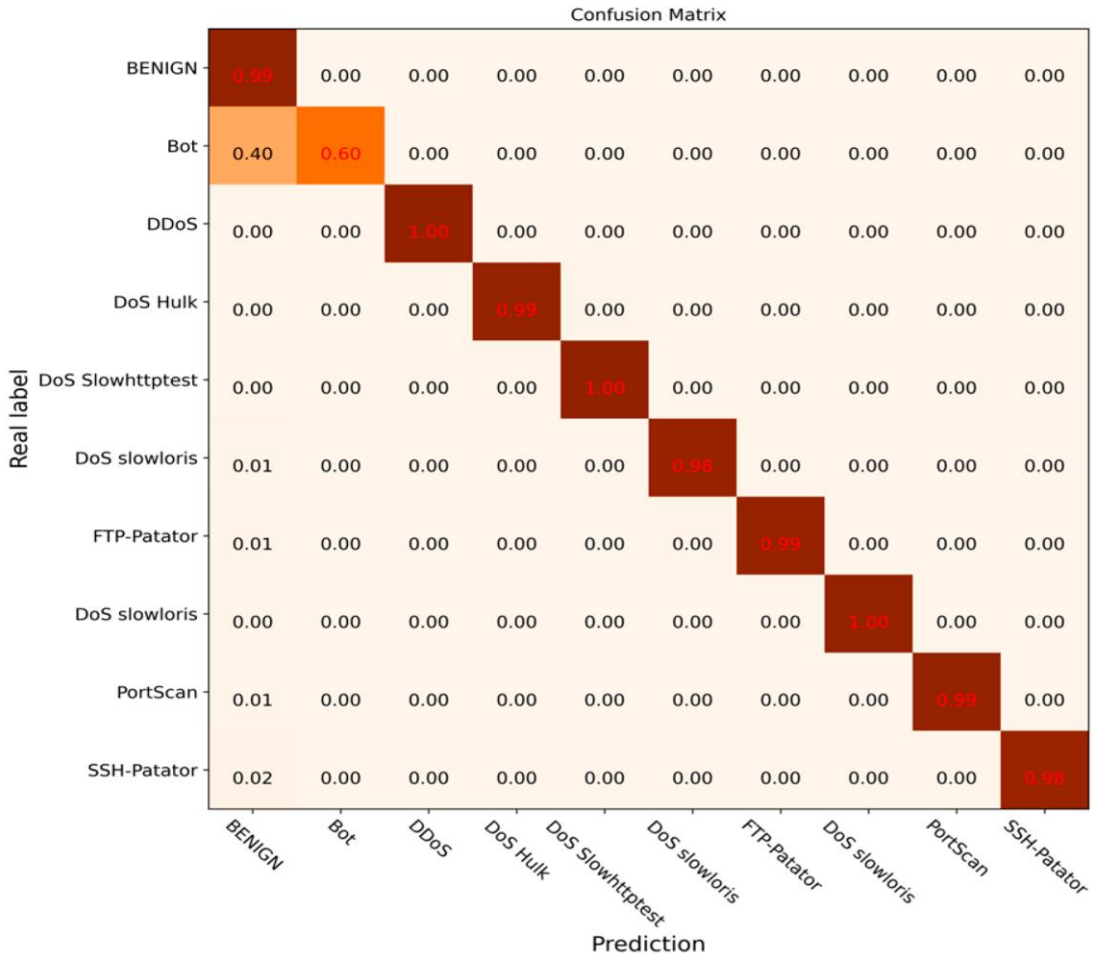


Рисунок 4.9 – Матриця похибки результатів класифікації DCGCANet

З результатів експериментів з абляції чітко видно, що як модель DC-GRU, так і модуль CA (channel attention - увага до каналу) значно покращили розпізнавання аномального трафіку. Зокрема, було виявлено, що модель DC-GRU з її унікальною розширеною структурою згортки ефективніше фіксує ключові ознаки в даних, тим самим забезпечуючи точніше представлення ознак для виявлення аномального трафіку. З іншого боку, модуль CA, зважуючи ознаки різних каналів, посилював фокус моделі на важливих ознаках, що ще більше підвищувало точність виявлення аномального трафіку. Що ще важливіше, коли ці два модулі були об'єднані, їхня взаємодоповнюваність призвела до більшого покращення загальної точності моделі.

Щоб краще проілюструвати здатність нашої моделі DCGCANet виявляти аномальний трафік, було проведено порівняння між моделлю

DCGCANet та кількома алгоритмами — KNN, багатошаровим персептроном (MLP), випадковими лісами (RF), CNN, AFM-ICNN-1D та CNN-GRU [20] – з використанням набору даних CIC-IDS-2017 за тих самих експериментальних умов. Результати цих порівнянь представлені в таблиці 4.4.

Таблиця 4.4 – Експериментальні результати різних моделей

Метод	Acc	Pre	Recall	F1	Час
KNN	96.3%	96.1%	96.3%	96.2%	201.68 с
ID3	98.3%	98.4%	98.5%	98.1%	192.38 с
MLP	76.6%	77.2%	83.3%	76.8%	215.91 с
RF	96.7%	97.2%	88.3%	92.9%	206.31 с
CNN	95.3%	97.6%	97.7%	97.3%	200.19 с
AFM-ICNN-1D	96.3%	98.3%	98.4%	98.3%	230.81 с
CNN-GRU	96.1%	96.2%	96.6%	96.2%	252.19 с
DCGCANet	99.6%	99.6%	99.6%	99.6%	273.13 с

Як показано в таблиці порівняння, що оцінює точність кількох моделей, модель, представлена в цій роботі, продемонструвала значну перевагу, незважаючи на необхідність більшого часу виявлення. Зокрема, наша модель досягла точності 99,6%, що демонструє суттєві покращення порівняно з традиційними та найсучаснішими методами, такими як KNN, ID3, MLP, RF, CNN та нещодавно розроблена модель AFM-ICNN-1D.

Як показано в таблиці 4.4, традиційні алгоритми машинного навчання, такі як kNN, RF та ID3, також досягають задовільних результатів виявлення. Однак, для зменшення розмірності та підвищення точності, у методах машинного навчання необхідний вибір ознак. Метод, запропонований у цьому дослідженні, продемонстрував чудову продуктивність у сфері вибору ознак.

Такі методи, як CNN, AFM-ICNN-1D та CNN-GRU, страждають від певної втрати інформації через операції об'єднання. Натомість цей метод

замінює об'єднання розширеною згорткою та враховує увагу до каналу. Це дозволяє моделі призначати різні ваги ознакам у різних каналах залежно від їхньої важливості, тим самим підвищуючи здатність моделі фіксувати критичну інформацію та покращуючи загальну здатність представлення. Однак через високу складність моделі було витрачено більше часу на покращення ефективності виявлення моделі.

Покращення точності на 3,3% порівняно з KNN свідчить про високу здатність обробляти складні набори даних, тоді як збільшення на 1,3% порівняно з ID3 демонструє високу ефективність у прийнятті рішень щодо класифікації. Порівняно з MLP спостерігалося покращення точності на 23%. Крім того, наша модель продемонструвала покращення на 2,9% та 4,3% порівняно з RF та CNN відповідно, що доводить її чудову продуктивність у виявленні аномалій мережевого трафіку. Крім того, спостерігаються значні покращення точності, повноти та F1-оцінка, що підтверджує повноту та надійність запропонованої моделі.

Щоб перевірити здатність DCGCANet до узагальнення, ми використали чотири підмножини P1-P4 CIC-IDS2017 як тестовий набір, і DCGCANet порівняли з моделлю LSTM на цих наборах даних. Результати представлені в таблиці 4.5.

Таблиця 4.5 – Значення точності, повноти та F1 для різних підмножин набору даних CIC-IDS-2017

DataSets	Pre		Recall		F1	
	LSTM	DCGCANet	LSTM	DCGCANet	LSTM	DCGCANet
P1	73.6%	99.3%	75.5%	99.4%	76.6%	99.4%
P2	73.1%	99.4%	73.2%	99.5%	73.1%	99.3%
P3	76.1%	99.6%	75.7%	99.7%	76.1%	99.7%
P4	68.3%	99.6%	70.3%	99.6%	68.3%	99.6%

З результатів порівняння, представлених у таблиці 4.5, можна

побачити, що значення точності, повноти та F1 запропонованої DCGCANet на чотирьох підмножинах CIC-IDS2017 перевищували 99%, і ці експериментальні результати значно кращі, ніж результати моделі LSTM. Отже, можна зробити висновок, що DCGCANet має чудову здатність до узагальнення. Це пояснюється акцентом моделі DCGCANet на важливих характеристиках каналу шляхом зважування каналу, що покращує репрезентативні можливості моделі та призводить до покращеної здатності до узагальнення.

Для подальшої перевірки здатності моделі до узагальнення, для валідації було використано набір даних KDD Cup99, результати яких наведено в таблиці 4.6. Набір даних KDD Cup99 містить нормальну поведінку та чотири типи аномальної поведінки: звичайний трафік (Normal), атака відмови в обслуговуванні (DOS), атака зондування (Probe), атака віддаленого входу (R2L) та атака користувача до root (U2R).

Таблиця 4.6 – Вплив запропонованого методу на виявлення даних NSL-KDD

Attack Categories	Acc	Pre	Recall	F1
DoS	99.1%	99.2%	98.3%	97.9%
Normal	95.3%	95.6%	92.7%	93.3%
R2L	97.3%	97.3%	98.4%	98.0%
Probe	93.1%	93.2%	96.6%	95.2%
U2R	97.6%	97.6%	97.5%	97.6%

Як показано в таблиці 4.6, точність та коефіцієнти виявлення для всіх показників п'яти типів атак у наборі даних KDD Cup99 перевищували 90% за допомогою запропонованої моделі. Модель продемонструвала відмінну ефективність виявлення як на наборах даних CIC-IDS-2017, так і на KDD Cup99, що свідчить про високу здатність до узагальнення.

ВИСНОВКИ

Оскільки методи глибокого навчання здатні витягувати корисні ознаки та фіксувати взаємозв'язки та закономірності в ознаках, складність, пов'язана з традиційною інженерією ознак, зменшується. Отже, під час передатестаційної практики пропонується модель виявлення аномалій мережевого трафіку, заснована на розширеній згортці та увазі каналу. Безперервна багатошарова одновимірна структура розширеної згортки була розроблена для ефективного підвищення точності, коефіцієнта повторюваності та F1 виявлення аномального трафіку. Крім того, модуль CA обробляє ознаки відповідно до їхньої різної важливості, підвищуючи точність розпізнавання порівняно з традиційними алгоритмами виявлення вторгнень. У порівнянні з традиційними методами машинного навчання та існуючими методами глибокого навчання, модель демонструє чудову продуктивність виявлення та вищу точність у виявленні різних типів методів атак.

Щоб вирішити проблему втрати інформації, спричиненої операціями об'єднання в традиційні згорткові нейронні мережі, та нехтуванням важливості різних функцій, ми пропонуємо вдосконалену модель глибокого навчання. Ця модель інтегрує одновимірну розширену згортку та механізми канальної уваги для ефективнішого захоплення та обробки даних мережевого трафіку.

Спочатку була побудована модель, заснована на одновимірній розширеній згортці. На відміну від традиційних операцій згортки, розширена згортка розширює рецептивне поле без необхідності об'єднання, тим самим захоплюючи ширшу контекстуальну інформацію та уникаючи її втрати. Цей підхід особливо підходить для обробки даних зі складними ознаками або довгостроковими залежностями, такими як дані мережевого трафіку.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Gao, S.G.; Li, S.M.; Li, Q.Y.; Yu, J.W. Convolutional neural network web anomaly traffic detection method based on DAPA. Inf. Technol. Netw. Secur. 2020, 39, 8–12.

2. Hang, M.; Chen, W.; Zhang, R.J. Abnormal flow detection based on improved one-dimensional convolutional neural network. Comput. Appl. 2021, 41, 433.

3. Stiawan, D.; Heryanto, A.; Bardadi, A.; Rini, D.P.; Subroto, I.M.I.; Idris, M.Y.B.; Abdullah, A.H.; Kerim, B.; Budiarto, R. An approach for optimizing ensemble intrusion detection systems. IEEE Access 2020, 9, 6930–6947.

4. Xu, B.; Chen, S.; Zhang, H.; Wu, T. Incremental k-NN SVM method in intrusion detection. In Proceedings of the 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 24–26 November 2017; pp. 712–717.

5. Halim, Z.; Yousaf, M.N.; Waqas, M.; Sulaiman, M.; Abbas, G.; Hussain, M.; Ahmad, I.; Hanif, M.J.C. An effective genetic algorithm-based feature selection method for intrusion detection systems. Comput. Secur. 2021, 110, 102448.

6. Singh, S. Poly logarithmic naive bayes intrusion detection system using linear stable PCA feature extraction. Wirel. Pers. Commun. 2022, 125, 3117–3132.

7. Gu, J.; Lu, S. An effective intrusion detection approach using SVM with naïve Bayes feature embedding. Comput. Secur. 2021, 103, 102158.

8. Li, Y.; Zhang, B. An Intrusion detection Algorithm based on Deep CNN. Comput. Appl. Softw. 2020, 37, 324–328.

9. Khan, R.U.; Zhang, X.; Kumar, R. Analysis of ResNet and GoogleNet models for malware detection. J. Comput. Virol. Hacking Tech. 2019, 15, 29–37.

10. Wang, Z.; Wang, Z.; Yi, F.; Zeng, C. Attack traffic detection based on LetNet-5 and GRU hierarchical deep neural network. In Proceedings of the

International Conference on Wireless Algorithms, Systems, and Applications, Nanjing, China, 25–27 June 2021; pp. 327–334.

11. Chen, X.Y. Network intrusion detection based on convolutional neural networks with LSTM. Inf. Technol. Netw. Secur. 2021, 40, 42–46.

12. Ma, W.; Zhang, Y.; Guo, J. Abnormal traffic detection method based on LSTM and improved residual neural network optimization. J. Commun. 2021, 42, 23–40.

13. Xu, H.; Ma, Z.; Yi, H.; Zhang, L. Network Traffic Anomaly Detection Technology Based on Convolutional Recurrent Neural Network. Netinfo Secur. 2022, 21, 54–62.

14. Sun, H.; Chen, M.; Weng, J.; Liu, Z.; Geng, G. Anomaly detection for in-vehicle network using CNN-LSTM with attention mechanism. IEEE Trans. Veh. Technol. 2021, 70, 10880–10893.

15. Jia, Z.; Yao, Y.; Wang, Q.; Wang, X.; Liu, B.; Jiang, Z. Trojan traffic detection based on meta-learning. In Proceedings of the Computational Science–ICCS 2021: 21st International Conference, Krakow, Poland, 16–18 June 2021; Proceedings, Part II 21, 2021. pp. 167–180.

16. О.С. Ляшенко, І.А. Великодний, В.Г. Знайдюк, О.Д. Журило. Модель та методи виявлення широкомасштабної атаки в середовищі IoT / Системи управління, навігації та зв'язку. Том 1 № 75 (2024), С.127-132

17. Li, C.; Qiu, Z.; Cao, X.; Chen, Z.; Gao, H.; Hua, Z. Hybrid dilated convolution with multi-scale residual fusion network for hyperspectral image classification. Micromachines 2021, 12, 545.

18. Ji, C.P.; Yu, H.F.; Dai, W. Network Traffic Anomaly Detection based on Dilated Convolution and Channel Attention. In Proceedings of the 4th IFSA Winter Conference on Automation, Robotics & Communications for Industry 4.0/5.0, (ARCI' 2024), Innsbruck, Austria, 7–9 February 2024; pp. 90–97

19. Журило, О. і Ляшенко, О. (2024) «Архітектура та системи безпеки IoT на основі туманних обчислень», СУЧАСНИЙ СТАН НАУКОВИХ ДОСЛІДЖЕНЬ ТА ТЕХНОЛОГІЙ В ПРОМИСЛОВОСТІ, (1(27), с. 54–66.

doi: 10.30837/ITSSI.2024.27.054.

20. Zhang, X.J.; Wang, H.B. Research on intrusion detection algorithm based on CNN-GRU. J. Tianjin Univ. Technol. 2022, 38, 37–42.