

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій  
(повна назва)

Кафедра Інформаційно-мережна інженерія  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
**Пояснювальна записка**

рівень вищої освіти другий (магістерський)  
Аналіз алгоритмів згладжування і прогнозування трафіку даних  
(тема)

Виконав:  
здобувач 2 року навчання,  
групи ІМІм-23-1  
Кротінов А.П.  
(прізвище, ініціали)

Спеціальність 172 Електронні комунікації та радіотехніка  
(код і повна назва спеціальності)

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

Освітня програма Інформаційно-мережна інженерія  
(повна назва освітньої програми)


Керівник доц. Омельченко А. В.  
(посада, прізвище, ініціали)


Допускається до захисту

Зав. кафедри \_\_\_\_\_ Безрук В.М.  
(підпис) (прізвище, ініціали)

2025 р.

Не містить відомостей, заборонених до відкритого публікування

Студент  / Кротінов А.П. /  
( підпис ) ( прізвище та ініціали )

Керівник  / Омельченко А.В. /  
( підпис ) ( прізвище та ініціали )

## Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій  
 Кафедра Інформаційно-мережної інженерії  
 Рівень вищої освіти другий (магістерський)  
 Спеціальність 172 Електронні комунікації та радіотехніка  
 (код і повна назва)  
 Тип програми освітньо-професійна  
 (освітньо-професійна або освітньо-наукова)  
 Освітня програма Інформаційно-мережна інженерія  
 (повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

« 28 » жовтня 2024 р.

**ЗАВДАННЯ**

## НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Кротінову Андрію Петровичу  
 (прізвище, ім'я, по батькові)

1. Тема роботи Аналіз алгоритмів згладжування і прогнозування трафіку даних

затверджена наказом університету від 28 жовтня 2024 р. № 1148Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 15 січня 2025 р.

3. Вихідні дані до роботи У якості моделі трафіку використати самоподібний процес.

З використанням реалізацій фрактального трафіку трафіку дослідити алгоритми експоненціального згладжування та алгоритми прогнозування на основі авторегресійної моделі.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

4. Перелік питань, що потрібно опрацювати в роботі Вступ

Огляд характеристик мережевого трафіку

Аналіз існуючих підходів до згладжування і прогнозування трафіку

Аналіз алгоритмів згладжування трафіку

Дослідження можливостей прогнозування трафіку

Висновки

\_\_\_\_\_

\_\_\_\_\_

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) \_\_\_\_\_

Мета роботи, характеристики трафіку, основні алгоритми згладжування трафіку, основні алгоритми прогнозування трафіку, результати згладжування даних, прогнозування тривалостей транзакцій, результати дослідження, висновки тощо \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

### КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Ознайомлення із завданням. Уточнення ТЗ	28.10.24	Виконано
2	Підбір літератури за темою роботи	28.10-1.11	Виконано
3	Виконання розділу 1	2.11-16.11	Виконано
4	Виконання розділу 2	16.11-1.12	Виконано
5	Виконання розділу 3	1.12-15.12	Виконано
6	Виконання розділу 4	16.12-31.12	
7	Оформлення пояснювальної записки	1.01-10.01.24	Виконано
8	Оформлення презентаційного матеріалу, ідготовка до захисту у ЕК	11.01-14.01.25	Виконано

Дата видачі завдання 28 жовтня 2024 р.

Здобувач \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис) доц. Омельченко А.В.  
(посада, прізвище, ініціали)

## РЕФЕРАТ

Пояснювальна записка 74 с., 18 рис., 4 табл., 30 джерел.

Об'єкт роботи – корпоративна мережа підприємства.

Мета роботи – аналіз алгоритмів згладжування і прогнозування трафіку.

Проведено дослідження сучасних методів згладжування і прогнозування трафіку. Описано механізми згладжування і прогнозування трафіку. Наведено алгоритм згладжування і прогнозування трафіку. Розроблено механізм згладжування і прогнозування трафіку. Здійснено верифікацію результатів.

КОРПОРАТИВНА МЕРЕЖА, ЗГЛАДЖУВАННЯ, ТРАФІК,  
ПРОГНОЗУАННЯ, ПОКАЗНИК ХЕРСТА, НЕЙРОННА МЕРЕЖА.

## THE ABSTRACT

Explanatory slip 74 p., 18 fig., 4 tab., 30 sources.

The object of the work is the corporate network of the enterprise.

The purpose of the work is the analysis of traffic smoothing and forecasting algorithms.

A study of modern methods of traffic smoothing and forecasting has been conducted. The mechanisms of traffic smoothing and forecasting have been described. The algorithm of traffic smoothing and forecasting has been presented. The mechanism of traffic smoothing and forecasting has been developed. The results have been verified.

CORPORATE NETWORK, SMOOTHING, TRAFFIC, FORECASTING, HURST INDEX, NEURAL NETWORK

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ.....	8
ВСТУП.....	9
1 ОГЛЯД ХАРАКТЕРИСТИК МЕРЕЖНОГО ТРАФІКУ .....	12
1.1 Характеристика основних видів трафіку.....	12
1.2 Модель трафіку у вигляді самоподібного процесу трафіку .....	20
2 АНАЛІЗ ІСНУЮЧИХ ПІДХОДІВ ДО ЗГЛАДЖУВАННЯ І ПРОГНОЗУВАННЯ ТРАФІКУ .....	28
2.1 Статистичні алгоритми згладжування і прогнозування трафіку .....	28
2.2 Прогнозування трафіку з використанням нейронної мережі .....	33
3 АНАЛІЗАЛГОРИТМІВ ЗГЛАДЖУВАННЯ.....	39
3.1 Формування функцій згладжування трафіку .....	39
3.2 Аналіз якості варіантів згладжування трафіку.....	44
4 ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ПРОГНОЗУВАННЯ ТРАФІКУ .....	47
4.1 Розрахунок точності найкращого лінійного прогнозу.....	47
4.2 Задача прогнозування самоподібних процесів.....	55
ВИСНОВКИ .....	60
ПЕРЕЛІК ПОСИЛАНЬ .....	62
ДОДАТОК	
А.....	66

## ПЕРЕЛІК СКОРОЧЕНЬ

ВОЛЗ	волоконно оптичні лінії зв'язку
ВП	випадковий процес
ЕОМ	електронно-обчислювальна машина
ЦКЗ	цифровий канал зв'язку
ANN	ArtificialNeuralNetwork
ARMA	AutoregressiveMovingAverage
ARP	Address Resolution Protocol
QoS	QualityofService
SSA	Static single assignment form
SVMR	Support Vector Machine Regression
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

## ВСТУП

Актуальність дослідження. Бурхливий розвиток телекомунікаційних технологій, а також зниження вартості передачі та обробки інформації призводить до постійного зростання обсягу мережного трафіку. Як відомо, подібний трафік найбільш чутливий до затримок у передачі пакетів [1]. Так, під час передачі голосових даних критичною стає затримка понад 10 мс, для потокового відео затримка має перевищувати 100 мс. Таким чином, пред'являються підвищені вимоги до мереж передачі даних, алгоритмів керування трафіком і, зокрема, методів боротьби з перевантаженнями.

Завдання управління трафіком є одним із ключових при забезпеченні якості обслуговування (QoS – Quality of Service) абонентів. З точки зору топології мережі, управління трафіком включає мережеве планування та оптимізацію. Мережеве планування – це процес визначення топології мережі та пропускної спроможності каналів зв'язку з урахуванням передбачуваного навантаження. Оптимізація передбачає управління розподілом трафіку у існуючій мережі [2].

Боротьба з перевантаженнями у комп'ютерній мережі – важлива частина завдання забезпечення QoS та управління трафіком зокрема. Ефективні алгоритми боротьби з навантаженнями дозволяють підвищити як надійність, так і корисну пропускну здатність мережі. Перевантаження в комп'ютерній мережі виникає тоді, коли обсяг інформації, що передається, наближається до максимальної пропускної здатності мережі. Протокол IP для обміну даними не вимагає попередньої установки з'єднання, таким чином, мережний пристрій не може визначити необхідний обсяг ресурсів до отримання першого повідомлення, це може призвести до того, що вхідний потік займе всі ресурси маршрутизатора. У першу чергу це пов'язано з використанням буферів кінцевої довжини в обладнанні, що комутує. При заповненні цих буферів нові пакети починають відкидатися вузлами мережі, що своєю чергою викликає повторну передачу повідомлення, і навантаження зростає лавиноподібно. Причина такої ситуації у тому, що протоколи транспортного рівня моделі OSI, наприклад TCP, не мають безпосередніх засобів визначення стану мережі, роблячи висновки з

урахуванням діалогу з кінцевими системами. Отримувач у межах протоколу TSP надсилає відправнику підтвердження прийому та обсяг даних, які він готовий прийняти. Отже, вихідні дані алгоритмів управління каналом зв'язку виходять досить грубим методом [3].

Для нормального функціонування мережі необхідно підтримувати пропускну здатність на прийнятному рівні. Ряд методик на основі зворотного зв'язку дозволяють керувати пропускну здатністю в залежності від навантаження: протитиск, пакет, що стримує, сигналізація про перевантаження. Способи забезпечення QoS без зворотного зв'язку, такі як принцип справедливого розподілу ресурсів, алгоритми керування чергами (у тому числі RED), а також алгоритми профілювання навантаження функціонують на рівні мережевого пристрою. Вище перелічені підходи до забезпечення якості обслуговування розроблялися та тестувалися на пуассонівських моделях трафіку та не враховують самоподібної структури телекомунікаційного трафіку та розподілу з важким хвостом. Використання більш адекватних моделей трафіку може сприяти розробці ефективних протоколів передачі та методик забезпечення якості обслуговування абонентів. Зокрема, це стосується методик прогнозування трафіку та стану мережі для боротьби з навантаженнями.

На разі особливо важливо як розробити адекватний алгоритм прогнозу, так і врахувати природу трафіку, протокол передачі, архітектуру мережі, ступінь завантаженості і характер навантаження. Залежно від перерахованих вище факторів статистичні та динамічні властивості трафіку відрізнятимуться.

Системи керування навантаженнями в мережі на основі прогнозованого навантаження показують кращі результати в порівнянні з системами, що працюють з мережею в реальному часі. Варто відзначити, що в більшості робіт з розробки способів управління навантаженнями в мережі, не враховуються нелінійно-динамічні властивості процесу передачі повідомлень, хаотичний характер трафіку, а також вплив трафіку на апаратні ресурси серверів та обладнання. Тим часом виявлення залежностей між обсягом трафіку і станом

обладнання може допомогти при розробці способів прогнозу перевантажень і, як наслідок, підвищення якості обслуговування абонентів мережі.

Актуальність проблеми згладжування і прогнозування трафіку за допомогою програмних методів визначила основні напрями даного дослідження.

*Мета та завдання.* Метою даного дослідження виступає аналіз алгоритмів згладжування і прогнозування трафіку. Для досягнення поставленої мети у роботі необхідно виконати низку завдань:

- дослідити сучасні методи згладжування і прогнозування трафіку;
- описати механізми згладжування і прогнозування трафіку;
- навести алгоритм згладжування і прогнозування трафіку;
- розробити механізм згладжування і прогнозування трафіку;
- здійснити верифікацію результатів.

## 1 ОГЛЯД ХАРАКТЕРИСТИК МЕРЕЖНОГО ТРАФІКУ

### 1.1 Характеристика основних видів трафіку

Основні інформаційно-технічні характеристики мережі, які визначають можливості з надання гарантованої якості послуг для користувачів і працездатності мережі в цілому, зазвичай задаються на етапі планування і уточнюються в результаті проектування та оптимізації мережі. Зазвичай до них належать такі основні характеристики [4]:

- пропускна спроможність транспортних магістралей чи базові швидкості передачі;
- обсяг вхідного та вихідного трафіку в мережевих вузлах;
- сумарний трафік у мережевих трактах та магістралях мережі;
- надійність чи коефіцієнт готовності мережі загалом.

При плануванні мережі визначають чи задають основні вимоги, які забезпечують як гарантовану якість послуг, так і можливість подальшого її нарощування та розвитку зв'язку. Загальні вимоги до сучасної мережі передбачають такі фактори:

- необхідна смуга пропускання;
- розширюваність та масштабованість мережі;
- керованість мережі;
- інтеграція різних видів трафіку;
- сумісність апаратури та комутації;
- резервування мережевих трактів та каналів;
- надійність та готовність мережі.

Таким чином, з одного боку, на основі аналізу загальних інформаційно-технічних характеристик мережі, включаючи аналіз розподілу трафіку, здійснюється планування та організація магістралей транспортної мережі. З іншого, на основі аналізу первинного навантаження мережі та визначення або завдання ієрархії обміну потоків за допомогою функцій комутації планують загальне навантаження в мережі.

Теорія масового обслуговування вивчає системи та ситуації обслуговування випадкового потоку заявок (вимог) обмеженою кількістю призначених для цього каналів. Після закінчення деякого випадкового часу обслуговування, канал готовий до роботи над наступною заявкою. Така ситуація характерна для систем комутації (телефонних станцій, АТМ-комутаторів тощо) у мережі зв'язку, ЕОМ з обмеженою кількістю процесорів. У теорії масового обслуговування основними поняттями є [5]:

- дисципліна обслуговування;
- потоки заявок (трафік);
- закони розподілу часу обслуговування одним каналом;
- ймовірності відмов у обслуговуванні;
- середній час очікування у черзі на обслуговування;
- пропускна спроможність системи обслуговування.

При вирішенні різних інженерних завдань з використанням методів теорії масового обслуговування часто вважають, що потоки описуються розподілом Пуассона. Таке припущення як спрощує аналіз, так і часто близьке до реальних процесів. Справа в тому, що пуасонівські потоки в певному сенсі є граничними для різних потоків (трафіку). Наприклад, якщо потік виходить в результаті складення досить великого числа потоків різної структури, то сумарний потік, у досить широкому діапазоні умов, буде близький до пуасонівського потоку. Така ситуація характерна для систем комутації у мережах зв'язку та інших систем масового обслуговування. У складних технічних системах, що складаються з великої кількості елементів, потік відмов буде складатися з потоків відмов окремих її елементів. Тому потік відмов технічних систем близький до пуасонівського потоку. Така ситуація характерна для опису відмов у теорії надійності технічних систем і, зокрема, надійності ВОЛЗ та мереж зв'язку.

У процесі надання послуг на мережах зв'язку розрізняють кілька видів дисципліни обслуговування окремих заявок або вимог. Зокрема, з пріоритетом та без пріоритету. До дисциплін обслуговування без пріоритету відносять

режими обслуговування звідмовами, очікуванням(без відмов), змішані режими (з обмеженням часу перебування заявок у черзі чи за довжиною черги).

Потік заявок (трафік) поділяють на стаціонарний та нестаціонарний, ординарний та неординарний. Найчастіше потік заявок вважають стаціонарним. Імовірність надходження деякої кількості заявок за кінцевий інтервал часу визначається лише часовим інтервалом і залежить від моментів початку і кінця. Найпростіший потік – це стаціонарний і ординарний потік без наслідку. Використовуючи модель найпростішого (пуасонівського) потоку, можна обчислити щільність ймовірності розподілу інтервалів між заявками на обслуговування, середній час обслуговування одним каналом імовірність відмови виклику [6].

Для оцінки якості обробки трафіку використовується інтенсивність трафіку  $A$ , що відповідає періоду максимального навантаження або періоду пікового навантаження. Метою оптимального проектування системи зв'язку є забезпечення для заданого числа каналів максимально можливого значення питомої інтенсивності трафіку  $A$ , що пропускається однією лінією пучка, або відношення  $A/V$ . Допустиме значення  $A/V$  має відповідати прийнятному значенню показника якості обробки трафіку, протягом якого прийнята ймовірність заняття всіх каналів одночасно, тобто ймовірність події, що призводить до втрати виклику. Таким чином, можливість втрати одного виклику в системі зв'язку з  $V$  каналами в періоді пікового навантаження визначає значення показника якості обробки трафіку. Зрозуміло, що чим більше  $V$ , тим менше показник якості обробки трафіку і тим вище якість обробки трафіку.

На основі цього будують таблиці Пальма для різного числа  $V$  каналів та заданої ймовірності блокування (відмов). Таблиці Пальма дозволяють для заданого значення ймовірності блокування трафіку і заданого числа каналів  $V$  між вузлами мережі визначити максимально допустиму для даного пучка каналів інтенсивність трафіку  $A$  в Ерлангах.

Недоліком комутаційних систем із великою кількістю виходів  $V$  є підвищена чутливість до перевантажень.

Питома інтенсивність трафіку  $A/V$  (коефіцієнт використання каналів у пучку) часто використовується як основний параметр в інженерних додатках і, зокрема, для розрахунку та оптимізації трафіку в мережі зв'язку.

Сучасні цифрові мережі та системи зв'язку дозволяють передавати та комутувати різні види трафіку зі швидкостями до 10 Гбіт/с та вище. Можна виділити дві основні категорії трафіку – трафік реального часу (передача голосу, аудіо, відео тощо) та трафік передачі даних, в більшості випадків, передача якого не критична до затримок.

За характером інформації, що передається основними видами трафіку є: передача голосу; передача даних; передача відеотрафіку; передача мультимедіа (аудіо, відео, дані). Вимоги до цифрових каналів зв'язку в мережі залежно від виду і типу трафіку, що передається, істотно різняться. Голосовий трафік та відеотрафік – це трафік реального часу. Вони пред'являють жорсткі вимоги до необхідної смуги пропускання та тимчасових затримок у каналі зв'язку. Наприклад, для якісної передачі голосу, як було показано вище, потрібен цифровий канал зв'язку (ЦКЗ) типу ОЦК зі швидкістю передачі 64 кбіт/с. (без стиснення). Для високоякісної передачі музики (високоякісне аудіо) потрібна смуга частот від 16 кГц і вище, що відповідає швидкості передачі цифрових каналів зв'язку рівної 128 кбіт/с. Передача стандартного відеосигналу з високою якістю зображення потребує смуги частот до 6 МГц. При цифровому перетворенні такого сигналу з частотою дискретизації 16 мегагерц і 8-бітовому кодуванні швидкість передачі становить 1.28 Мбіт/с.

У системах відеоконференцій передаються лише зміни у діапазоні частот між двома наступними кадрами. У звичайній ситуації, коли об'єкт, що спостерігається, здійснює незначні рухи, обсяг даних, які необхідно передати, може становити тільки 1% обсягу повного кадру. Якщо зображення оновлюється нечасто і застосовуються методи стиснення відеосигналу, то для передачі сигналів відеоконференцій і отримання хорошої якості відео, потрібні ЦКЗ зі швидкістю передачі 128- 256 кбіт/с. Верхня межа швидкості передачі сигналів

відеоконференцій визначається співвідношенням ціна/якість одержуваного зображення.

Передача трафіку мультимедіа в залежності від якості та виду послуги, що надається, вимагає ЦКЗ з різною швидкістю передачі. При цьому трафік мультимедіа включає практично всі основні види трафіку, за винятком передачі даних. Вимоги до ЦКЗ щодо пропускної спроможності визначаються при плануванні та проектуванні конкретних систем. У табл. 1.1. наведено вимоги до цифрових каналів зв'язку передачі різних типів трафіку мультимедіа.

Таблиця 1.1 – Загальні характеристики трафіку різних програм

Тип трафіку	Відео зі стиснення м	Аудіо зі стиско м	Відео без стиснен ня	Аудіо без стиснен ня	Дані зі стиснення м	Дані без стиснен ня
Швидкість передачі, кбіт/с	56...35000	16...384	3000...166000	64...1536	800...1200000	155000...1200000
Коефіцієнт пульсацій	1:10	1:3	1:10	1:3	3:1000	3:1000

Трафік передачі характеризується значною неоднорідністю, вибухоподібним характером у часі і вимагає своєї передачі у різні моменти часу різної смуги пропускання.

У сучасних магістральних транспортних мережах зберігається і посилюється тенденція збільшення у загальному обсязі трафіку частка обсягу трафіку передачі. Це пов'язано з різким збільшенням обсягу послуг, які надає мережа Інтернет. Основним показником цього процесу є відношення обсягів основних видів трафіку(голос/дані) у магістральних каналах зв'язку. Чим більше це відношення у бік обсягу трафіку передачі даних, тим більшою є необхідність застосування ІР-технологій у сучасних мережах.

Основними характеристиками трафіку передачі є одиниця даних і спосіб упаковки цих одиниць. Одиницею даних може бути:біт, байт, повідомлення, блок. Дані упаковуються у файли, пакети, кадри, комірки, а можуть також передаватися без упаковки. Швидкість передачі даних вимірюється водиницях

даних за одиницю часу визначає час, необхідний для передачі одиниці даних по мережі. Зазвичай швидкість передачі вимірюють в бітах за секунду або кратних їй одиницях (Кбіт/с, Мбіт/с і т.д.). Реальний обсяг даних, що передаються по мережі, складається безпосередньо з даних (корисного навантаження) і необхідного інформаційного обрамлення, що становить «накладні витрати» на передачу. Багато технологій встановлюють обмеження мінімальний і максимальний розміри пакета. Наприклад, для технології X.25 максимальний розмір пакета становить 4096 байт, а технології FrameRelay максимальний розмір кадру становить 8096 байт.

Можна виділити такі загальні характеристики трафіку передачі даних та його обслуговування:

- показник вибухоподібного характеру трафіку;
- терпимість до затримок;
- необхідна ємність та пропускна здатність мережі.

Ці характеристики з урахуванням маршрутизації, пріоритетів, з'єднань тощо, якраз і визначають характер роботи додатків у мережі.

Показник вибухоподібного характеру (вибухоподібності) трафіку визначає частоту посилки даних користувачем у мережу. Цей показник можна визначити ставленням максимального (пікового) значення густини трафіку (швидкості передачі) до її середнього значення. Наприклад, якщо максимальна (пікова) швидкість передачі даних становить 100Мбіт/с при середній швидкості передачі 10Мбіт/с, то показник вибухоподібності трафіку складе 10.

Терпимість до затримок характеризує реакцію додатків попри всі часові затримки у мережі. Наприклад, програми, що обробляють фінансові транзакції в реальному масштабі часу, не допускають затримок. Великі затримки можуть призвести до неправильної роботи таких програм. Різні програми сильно відрізняються за допустимим часом затримки. Для програм, що працюють у реальному масштабі часу (наприклад, для відеоконференцій), час затримки не повинен перевищувати деякого граничного значення, яке є досить малим. З іншого боку, для багатьох програм допустимі значення затримки можуть

становити від декількох хвилин до декількох годин (наприклад, для електронної пошти та пересилання файлів).

Поняття ємності та пропускної спроможності мережі пов'язані між собою, але, по суті, не однакові. Місткість – це реальна кількість ресурсів, доступних користувачеві на певному шляху передачі даних. Пропускна здатність визначається загальною кількістю даних, які можуть бути передані за одиницю часу. Місткість мережі відрізняється від пропускної спроможності мережі через наявність накладних витрат, які залежать від способу використання мережі. У табл. 1.2 представлені загальні характеристики трафіку передачі для різних додатків.

Таблиця 1.2 – Загальні характеристики трафіку різних програм

Додаток / Характеристика графіка	Вибухоподібність трафіку	Терпимість до затримок	Час відповіді	Пропускна спроможність, Мбіт/с
Електронна пошта	Висока	Висока	Регламентується	0.004..0.20
Голос	Середня	Низька	Реальний час	0,004. .0.064
Передача файлів	Висока	Висока	Регламентується	0.01...600
CAD/CAM-системи	Те саме	Середня	Близький до реального	1...100
Обробка транзакцій	Те саме	Низька	Близький до реального	0,064...2,048
Обробка зображень	Те саме	Середня	Реальний час	0,256...25
Ділове відео	Низька	Низька	Реальний час	0,256... 16
Розважальне відео	Низька	Низька	Близький до реального	2.04S...50
Широкомовне відео	Низька	низька	Реальний час	0,128-128
Зв'язок локальних мереж	Висока	Висока	Те саме	10... 100
Доступ до сервера	Середня	Висока	Те саме	10...100
Високоякісне аудіо	Низька	Низька	Те саме	0.128...1

Для деяких програм потрібно гарантувати час реакції, пропускну здатність мережі та інші характеристики. Це забезпечується технологією якості обслуговування QoS (Quality of Service). Вона дозволяє використовувати розподіл за категоріями якості обслуговування та призначення пріоритетів для різних видів трафіку. Для трафіку з високим пріоритетом забезпечується гарантована якість обслуговування та кращі умови передачі в мережі, незалежно

від вимог до пропускної спроможності для трафіку менш відповідальних додатків. Якість обслуговування вибирають на основі вимог користувачів для конкретних мережевих технологій та мультисервісних мереж в цілому. Мультисервісні мережі дозволяють передавати будь-які види та типи трафіку з необхідною якістю обслуговування.

Умовно трафік передачі даних можна розділити на три категорії, що відрізняються одна від одної вимогами до часових затримок під час передачі.

Трафік реального часу. До цієї категорії відноситься трафік з аудіо- та відеоінформацією, що практично не допускає затримок. Для такого трафіку допустима затримка стала і зазвичай не перевищує 0,1 с. включаючи час на обробку у кінцевому вузлі мережі. Слід зазначити, що при стисненні інформації трафік реального часу стає дуже чутливим до помилок під час передачі.

Трафік транзакцій. Допустима часова затримка не повинна перевищувати 1 с. Великі затримки призводять до зниження продуктивності праці та дискомфорту у роботі. У ряді випадків перевищення допустимої затримки призводить до збою робочої сесії і додатків користувача потрібно її повторити.

Трафік даних. Допустима будь-яка часова затримка, аж до кількох секунд. Особливістю такого трафіку є підвищена чутливість до доступної пропускної спроможності мережі, а не часових затримок. При збільшенні пропускної здатності мережі зменшується час передачі. Програми, що вимагають передачі великих обсягів даних, зазвичай займають всю доступну смугу пропускання мережі. Рідкісним винятком є програми потокового відео. Їх істотні значення і пропускну здатність з мінімальною затримкою.

Для кожної категорії трафіку встановлюються пріоритети. Трафік із вищим пріоритетом обробляється першим. Приклад пріоритетного трафіку може бути трафік транзакцій із замовленням. Введення пріоритетів для різних видів трафіку неминуче залежить від недостатності ресурсів мережі. При цьому пріоритети можуть застосовуватися для виділення груп, прикладних програм та окремих користувачів у групах тощо.

Передача аудіо-відеоінформації чутлива до зміни часової затримки та може призводити до помітних спотворень зображень та розбірливості мови абонента.

До характеристик трафіку, що створюється різними мультимедійними послугами, належать:

- значення трафіку (миттєве, максимальне, пікове, середнє мінімальне), біт/с ;
  - коефіцієнт пачечності трафіку (пульсації);
  - середня тривалість пікового трафіку;
  - середня тривалість сеансу зв'язку;
  - формати елементів трафіку;
  - максимальний, середній, мінімальний розмір пакета;
  - інтенсивність трафіку запитів.

Слід зазначити, що мультимедійний трафік зазвичай не може бути описаний пуассонівським потоком. Для опису вказаного трафіку останнім часом застосовується теорія самоподібного трафіку.

## 1.2 Модель трафіку у вигляді самоподібного процесу трафіку

Сучасна мережева інфраструктура настільки велика, що відстежити правильність (безаномальність) руху всієї інформації в ній практично неможливо. Мережевий трафік являє собою складний динамічний процес і є суперпозицією багатьох потоків з множинними взаємозв'язками, які генеруються різними потоками [7]. При дослідженнях комп'ютерних систем і мереж застосовують як аналітичні (засновані на математичних розрахунках) моделі, так і імітаційне моделювання на основі вже готових програм емуляторів створених за допомогою універсальних мов програмування. Системи моделювання можуть бути як вузьконаправленими (спеціалізованими), так і універсальними, що дозволяють імітувати мережі найрізноманітніших типів [6]. Результат

досліджень експериментальних даних показав, що трафік сучасних комп'ютерних мереж має особливу структуру, яка проявляє ефект самоподібності. Цей ефект проявляється в тому, що статистичні характеристики трафіку як би «масштабуються» при усередненні значень взятих за різні проміжки часу. Іншими словами, під самоподібністю мається на увазі повторюваність розподілу навантаження в часі при різних масштабах (рис. 1.2) [5].

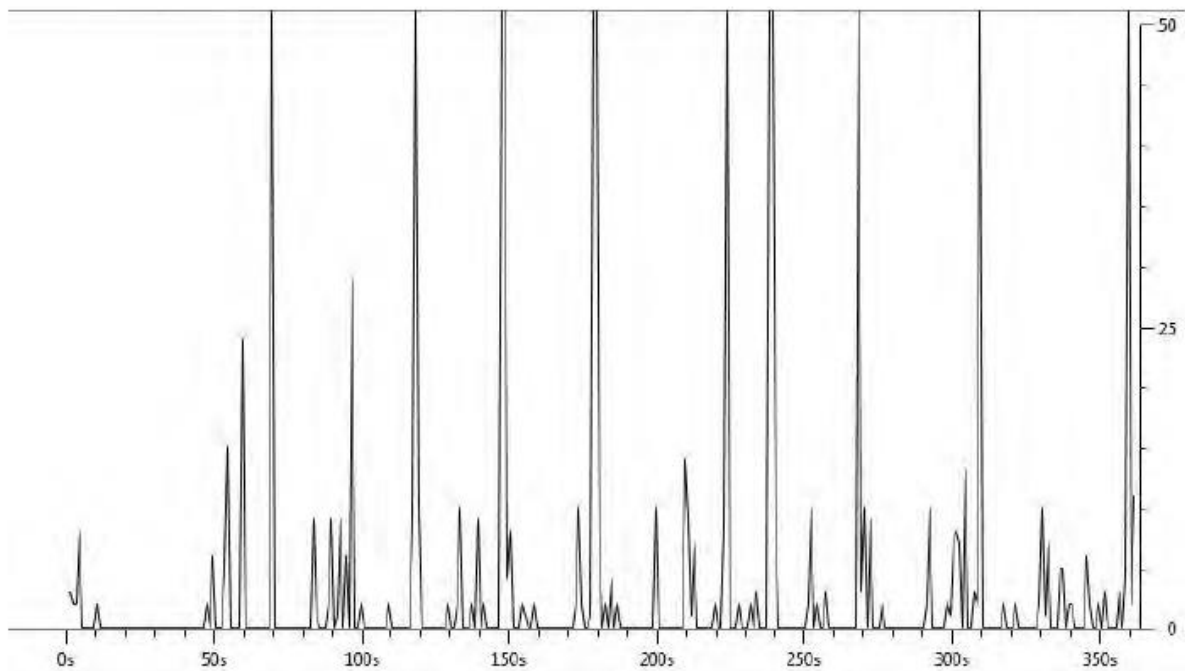


Рисунок 1.1 – Нормальне навантаження трафіку при бездротовому з'єднанні [8]

Утворення ефекту самоподібності пов'язано з властивостями TCP / IP протоколу передачі даних, який на сьогоднішній день використовується в комп'ютерних мережах. Мається на увазі особливість передачі даних у вигляді пакетів, які приходять на вузол комутації цілими пачками, а не випадковим чином. Трафік в таких мережах має явно виражений вибуховий характер, що підвищує ймовірність перевантажень в вузлах мережі, які ведуть до переповнення буферів і викликають втрати і / або затримки. Пульсації

призводять до перепадів швидкості інформаційних потоків, при яких відношення максимального значення швидкості до мінімального становить десятки раз. Дослідження самоподібності показали, що це явище значно знижує якість трафіку через мережу [5]. Однак, незважаючи на те, що прояв самоподібності надає не дуже позитивний вплив на передачу даних в мережі (він є і з ним немає способів боротися), з цього явища можна витягти вигоду [9].

Властивість самоподібності асоціюється з визначенням фрактала, тобто, при зміні шкали кореляційна структура самоподібного процесу залишається незмінною. Самоподібні (фрактальні) моделі більш точно характеризують поведінку навантаженого мережевого потоку, ніж пуасонівські моделі, важливим завданням стала розробка інструментальних засобів для розуміння самоподібних процесів, і для синтезу штучного мережевого трафіку, який відображає основні характеристики цих процесів. Фрактальний аналіз трафіку дає ближчі до правди результати, ніж застосування класичних методів, заснованих на методах Пуассона, які дають невиправдано оптимістичні результати, що призводять до недооцінки навантаження, а, отже, до збоїв в роботі мережі. На даний момент фрактальність трафіку широко використовується для прогнозування поведінки комп'ютерної мережі на етапі її проектування, а так само для виявлення аномалій поведінки трафіку в існуючій мережі.

Центральними поняттями, які використовують для фрактального аналізу, є фрактальна розмірність ( $D$ ) і показник Херста ( $H$ ).

Фрактальна– розмірність (по Хаусдорфу) визначається [10]:

$$D = -\frac{\lg[N(\varepsilon)]}{\lg[\varepsilon]}, \quad (1.1)$$

де  $N(\varepsilon)$  – покривають задану множину непустих кубів розміром  $\varepsilon$

Показник ступеня Херста, показник Херста чи коефіцієнт Херста – міра, що у аналізі часових рядів. Ця величина зменшується, коли затримка між двома

однаковими парами значень у часовому ряду збільшується. Вперше це поняття використовувалося в гідрології у практичних цілях для визначення розмірів греблі на річці Ніл в умовах непередбачуваних дощів та посух, що спостерігаються протягом тривалого часу. Назва «Експонента Херста» або «Коефіцієнт Херста» дано на честь Гарольда Едвіна Херста (1880 – 1978) – провідного дослідника того часу в цій галузі. Стандартне позначення  $H$  також дано на його честь <sup>1</sup>.

Показник Херста характеризує ступінь самоподібності процесу:

- 1)  $0 < H < 0.5$  - випадковий процес, який не володіє самоподібністю, характеризується прагненням до середнього значення;
- 2)  $H = 0.5$  - повністю випадковий процес без вираженої тенденції;
- 3)  $H > 0.5$  - трендостійкий процес, який володіє тривалою пам'яттю і є самоподібним.

Фрактальна розмірність безпосередньо пов'язана з показником Херста співвідношенням:  $D = 2 - H$ . Це співвідношення справедливо, коли структура кривої, яка описує фрактальну функцію, досліджується з високою роздільною здатністю, тобто в локальній межі [6].

Алгоритм оцінки безпеки трафіку можна розділити на п'ять етапів:

- 1) збір трафіку;
- 2) статистичний аналіз;
- 3) оцінка показника Херста;
- 4) виявлення аномалій;
- 5) оцінка безпеки.

Для зменшення впливу на нормальне функціонування мережі трафік дублюється на спеціальний сервер, що займається збором трафіку. З пакетів, прийнятих від маршрутизатора, витягується інформація про тип пакету, а також чотири метрики трафіку: загальне число пакетів, число TCP пакетів, UDP пакетів, ARP пакетів в одиницю часу. Обчислюються показники Херста для

чотирьох метрик трафіку ітеративним методом оцінки в режимі реального часу [11].

Отримані значення використовуються для виявлення аномалій і поновлення моделі нормального трафіку. Поточне обчислене значення показника Херста порівнюється зі значенням з нормальної моделі поведінки трафіка. Якщо значення виходить за межі допустимого, трафік вважається аномальним. Нормальна модель трафіку будується шляхом аналізу нормальної роботи мережі протягом певного проміжку часу.

Модель включає нормальне значення показника Херста і довірчий інтервал, і може бути оновлена при виявленні аномалій. Критерієм оцінки безпеки є рівень ризику, який вираховується методом середньозважених величин, який враховує результати виявлення аномалій від чотирьох метрик трафіку. Рівень ризику надає адміністраторам поточний стан передачі даних в мережі з точки зору безпеки [6]. Для синтезу та аналізу алгоритмів виявлення атак часто використовують базу даних KDD-99, яка містить близько 5 мільйонів записів про мережеві з'єднання. При цьому представлені в цій базі 22 види атак діляться на 4 основні категорії.

Одним з популярних способів моделювання трафіку, що має задані самоподібні властивості, є використання моделі "Input M/G/ $\infty$ ". Оскільки українська термінологія в цій галузі ще не сформувалася, будемо використовувати послівний переклад і називати її «моделлю типу вхідна M/G/PP». Популярність цієї моделі обумовлена низкою її переваг. По-перше, вона відносно проста і її легко використовувати для генерації штучних трас трафіку довільної довжини, по-друге вона дає просте фізичне обґрунтування самоподібних властивостей трафіку (таке, наприклад, як наявність довготривалих залежностей і т.д.).

Числові характеристики трафіку мережі. Випадковий процес (ВП) надходження викликів або пакетів в систему обслуговування, що утворює потік заявок (трафік) характеризується законом розподілу, що встановлює зв'язок між значенням випадкової величини (кількістю заявок) і ймовірністю появи цього

значення. У більшості випадків для розрахунку параметрів QoS достатньо знати про закон розподілу лише деякі його числові характеристики. Загальними числовими характеристиками моменти розподілу різних порядків. Наприклад, для розрахунку в умовах пуассонівського розподілу достатньо математичного очікування  $M$ , а для нормального розподілу – необхідно мати значення  $M$  та дисперсії  $D$ . Іноді використовуються моменти третього та четвертого порядків – асиметрія та ексцес. Моменти вищих порядків використовуються рідко. При цьому якщо за заданими законами розподілу моменти розподілу можна визначити однозначно, то зворотне завдання можна вирішити не завжди.

Основні характеристики випадкового процесу  $M$  і  $D$ , будучи дуже важливими, водночас є вичерпними, які є іноді неінформативними для прогнозування значення випадкової величини. Можливі варіанти, коли ВП характеризуються однаковими значеннями математичного очікування та дисперсії, але внутрішня структура цих процесів є різною. Одні можуть мати плавно мінливі реалізації, а інші – яскраво виражену коливальну структуру при стрибкоподібній зміні окремих значень випадкової величини (наприклад, різке зростання кількості пакетів у мережі, що призводить до «пачечності – burstness» трафіку). Для «плавних» процесів характерна велика передбачуваність реалізацій, а для «пачечних» – дуже мала залежність ймовірності між двома випадковими величинами ВП. У такому випадку говорять, що закон розподілу, характеризує ВП, несе у собі деяку невизначеність і дозволяє з більшою чи меншою надійністю передбачити значення випадкової величини. Наприклад, при рівномірному розподілі всі значення випадкової величини рівноймовірні, а при експонентному – найменші значення мають найбільшу ймовірність.

Числовою характеристикою розподілу, яка може бути його мірою невизначеності, є ентропія закону розподілу, яка визначається для дискретного розподілу як

$$H(m) = - \sum_{i=1}^m p_i \log p_i,$$

(1.2)

де за повної групи подій  $\sum_{i=1}^m p_i = 1$ .

Ентропія залежить від значень, прийнятих випадковою величиною, лише від своїх ймовірностей. Зазвичай цікавить не абсолютне значення ентропії, а порівняння ентропій різних законів. Крім того, основа логарифму може бути взята будь-якою.

Моделювання самоподібного трафіку. В умовах відсутності точних аналітичних методів розрахунку параметрів QoS при самоподібному трафіку найкращим способом перевірки гіпотези про можливість застосування відомих методів розрахунку цих параметрів, але для інших моделей зі подібними значеннями ентропії розподілів є імітаційне моделювання. Для цього необхідно генерувати випадкові реалізації фрактального броунівського руху fBM.

Моделювання процесів фрактального характеру перебуває ще на стадії розвитку. При генеруванні самоподібного трафіку критерієм самоподібності є так звана довгострокова залежність (longrangedependent). Для виявлення цієї залежності використовують визначення функції кореляції. Для самоподібного процесу кореляційні властивості процесу, усередненого на різних часових інтервалах, залишаються незмінними. При цьому сам процес має пачковий (bursty) характер.

Пачковий характер генерованого трафіку сприяє його адекватності реального характеру трафіку в мультисервісних мережах. Тут при широкому діапазоні швидкостей передачі навантаження є різномірним, оскільки передачу потоків різних додатків і служб забезпечує одна й та сама мережа з єдиними протоколами та законами управління. Джерела певної служби характеризуються максимальною та середньою швидкостями передачі, тобто коефіцієнтом пачкування (burstness) та середньою тривалістю піку навантаження. Наприклад, пачкування для мовних служб можливе через паузи в розмові.

Найбільш відомим методом формування самоподібного потоку є метод Мандельброта. Він заснований на суперпозиції кількох (що строго чергуються) незалежних і таких, що мають однаковий розподіл ON/OFF джерел, інтервали

між ON і OFF є періодами які мають ефект Ноа (Noaheffect). Причому саме ефект Ноа у розподілі тривалостей ON/OFF періодів є базовим при моделюванні самоподібного трафіку. Ефект Ноа є синонімом синдрому нескінченної дисперсії. Математично для досягнення ефекту Ноа можна використовувати розподіл Парето, який часто називають «розподілом із довгим хвостом».

Щільність розподілу Парето задається функцією:

$$f(x) = \frac{a}{b} \left(\frac{b}{x}\right)^{a+1}, \quad (1.3)$$

де  $a$  – параметр форми;

$b$  – мода розподілу (мінімальне значення випадкової величини  $x$ ).

Причому, при  $a \leq 2$  дисперсія нескінченна (що і потрібно у якості однієї з умов самоподібності). Наявність у розподілі так званого «довгого хвоста» забезпечує властивість пачечності трафіку, оскільки у розподілі суттєво зростають ймовірності довгих інтервалів між подіями (наприклад, відсутність пакетів на інтервалі) та для «підтримання» заданого середнього значення кількості подій необхідна їх концентрація (збільшення) на інших інтервалах часу.

Параметр форми  $a$  розподілу Парето та параметр Херста  $H$  знаходяться в такій залежності [1]:

$$H = \frac{3-a}{2}. \quad (1.4)$$

У практичному моделюванні розподіл Парето (1.3) виходить шляхом переходу від рівномірного розподілу шляхом зворотної функції:

$$Z_i = \frac{b}{a\sqrt[U]{U_i}}, \quad (1.5)$$

де  $Z_i$  –  $i$ -й інтервал між подіями;

$U$  – довільне число, рівномірно розподілене на інтервалі  $[0, 1]$ .

## 2 АНАЛІЗ ІСНУЮЧИХ ПІДХОДІВ ДО ЗГЛАДЖУВАННЯ І ПРОГНОЗУВАННЯ ТРАФІКУ

### 2.1 Статистичні алгоритми згладжування і прогнозування трафіку

Задача прогнозування нових значень часового ряду за історичною вибіркою не нове [18], проте дослідження в цій галузі не втрачають актуальності і постійно виникають нові алгоритми та методики прогнозування [19]. Нещодавно дослідники почали застосовувати методи прогнозування часових рядів по відношенню до мережевого трафіку даних для управління трафіком, забезпечення QoS і боротьби з перевантаженнями [20]. На даний момент робіт у цій галузі робіт небагато, вони рідко спираються на дані щодо завантаженості реальної корпоративної мережі, а моделі прогнозування найчастіше авторегресійні, рідше – на основі нейромереж. До того ж алгоритмів прогнозування часових рядів та їх варіацій досить багато, а дослідження кожного з них є досить трудомістким завданням. Таким чином, завдання управління трафіком на основі короткострокових прогнозів стану мережі залишається дуже актуальним.

Формально задача прогнозування часового ряду  $X(t) = X(1), X(2), \dots, X(T)$  у момент часу  $T$  полягає у визначенні значень ряду  $X(t)$  у моменти часу  $T + 1, T + 2, \dots, T + T'$ , де  $T'$  – час попередження чи горизонт прогнозу. Іншими словами, визначення майбутнього стану процесу ґрунтується на відомих значеннях ряду, а прогноз ґрунтується на моделі, що відображає функціональну залежність між майбутніми та минулими значеннями ряду. У роботі [21] множина моделей прогнозування поділяється на статистичні та структурні. До статистичних моделей належать регресійні та авторегресійні, а також моделі згладжування. Моделі прогнозування на основі нейромереж, ланцюгів Маркова та класифікаційних дерев відносяться до структурних моделей.

Алгоритми згладжування. Різні алгоритми згладжування часового ряду або розрахунку середнього значення можуть використовуватися як проста і наочна модель, що прогнозує [15]. Найпростіший алгоритм ковзного середнього полягає у обчисленні наступного значення часового ряду на підставі середнього за період:

$$X_{t+1} = (X_t + X_{t-1} + X_{t-2} + \dots + X_{t-N+1})/N,$$

де  $X_{t+1}$  – прогнозоване значення низки на даний момент  $t+1$ ,

$X_{t-1}$  – значення низки на даний момент  $t-1$ ,

$N$  – період згладжування.

У даному випадку важливо підібрати відповідне значення  $N$ . До того ж логічно припустити, що  $X_t$  значення ряду має сильніше впливати на прогноз, ніж  $X_{t-N}$ . Алгоритм експоненціального згладжування враховує цей недолік через деякий коефіцієнт:

$$X'_{t+1} = \alpha X_t + (1 - \alpha)X'_t,$$

де  $0 < \alpha < 1$  – ваговий коефіцієнт,

$X_t$  – реальне значення ряду в момент  $t$ ,

$X'_t$  – прогнозне значення ряду.

Таким чином, є можливість отримати виважену ковзну середню, в якій наступне прогнозоване значення формується із прогнозованого та реального поточного значення ряду. При цьому коефіцієнт підбирається відповідно до характеру ряду – чим сильніше змінюються значення ряду, тим більше має бути коефіцієнт, що швидше пристосуватися до великих змін. І що вище шумова складова ряду, то менше має бути значення коефіцієнта, щоб згладжувати флуктуації. Для вирішення задачі підбору коефіцієнта розроблено алгоритм адаптивного згладжування.

Існує безліч варіацій алгоритму адаптивного згладжування, в найпростішому випадку коефіцієнт підбирається наступним чином:

$$\alpha_{t+2} = \left| \frac{X'_{t+1} - X_{t+1}}{X_{t+1}} \right|.$$

При цьому якщо значення коефіцієнта виходить рівним або більше 1, то приймається  $=0.99(9)$ , а якщо рівним 0, то  $=0.0(0)1$ . Розглянуті алгоритми згладжування часових рядів не враховують наявності у ряду тренду та сезонної складової.

Алгоритм експонентного згладжування з урахуванням тренду при згладжуванні враховує напрямок ряду:

$$L_t = \alpha X_t + (1 - \alpha)(L_{t-1} + T_{t-1});$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1},$$

де  $L$  – рівень тренду, а  $T$  – тренд, при цьому  $0 < \alpha < 1$  та  $0 < \beta < 1$ . Тоді прогнозоване значення ряду на  $m$  кроків:

$$X_{t+m} = L_t + (T_t m).$$

З урахуванням сезонної складової низки прогноз будується так [15]:

$$L_t = \alpha(X_t/SA_{t-c}) + (1 - \alpha)(L_{t-1} + T_{t-1});$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1};$$

$$SA_t = \gamma(X_t/L_t) + (1 - \gamma)(SA_{t-c}),$$

де  $L$  – рівень тренду,  $T$ -тренд,

$SA_t$  – сезонна складова ряду на момент часу  $t$ ,

$C$  – розмір циклу для сезонної складової,  $0 < \alpha < 1$  та  $0 < \beta < 1$ ,  $0 < \gamma < 1$ .

Тоді прогнозоване значення ряду на  $m$  кроків уперед:

$$X_{t+m} = (L_t + (T_t m)) S A_{t-C+m}.$$

Незважаючи на спроби не просто усереднити ряд ковзним вікном, а врахувати його сезонну, трендову та циклічну складову, прогнози на основі моделей, що згладжують, відфільтровують досить багато інформації про реальний процес. Разом із шумами усувається інформація, яка могла б бути використана під час прогнозування ряду. Розглянуті далі регресійні та авторегресійні моделі відрізняються більшою точністю, враховують більше факторів, що впливають на розподіл ряду.

Регресійні та авторегресійні алгоритми. У статистичній обробці даних регресійний аналіз застосовується для визначення залежності між вихідною залежною змінною та зовнішніми незалежними змінними, так званими регресорами. Найпростіша регресійна модель називається лінійною і передбачає наявність єдиного зовнішнього фактора, що впливає на процес із лінійним зв'язком:

$$X_t = \beta_0 + \beta_1 Y_t + \varepsilon_t,$$

де  $Y_i$  – незалежна змінна,  $\beta_0$  та  $\beta_1$  – коефіцієнти регресії,  $\varepsilon_t$  – помилка моделі. У разі множинної регресійної моделі зовнішніх факторів, що впливають на значення ряду – кілька:

$$X_t = \beta_0 + \beta_1 Y^1_t + \beta_2 Y^2_t + \dots + \beta_n Y^n_t + \varepsilon_t.$$

Залежність між вихідним процесом та регресором не обов'язково має бути лінійною, можливий випадок при якому залежність описується довільною функцією.

Найчастіше при дослідженні процесу складно виділити зовнішні фактори, що впливають, при цьому регресійний аналіз ґрунтується тільки на самих значеннях ряду, такі моделі називаються авторегресійними. Одна з найвідоміших авторегресійних моделей – ARMA (AutoregressiveMovingAverage) – є комбінацією авторегресійної моделі та ковзного середнього. Авторегресійна (AR(p)) модель може бути виражена наступним чином:

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t,$$

де  $X_t$  – поточне значення низки,

$\varepsilon_t$  – випадкова помилка,

$\varphi_i (i = 1, 2, \dots, p)$  – вагові коефіцієнти,

$c$  – константа.

Коефіцієнти  $\varphi_i (i = 1, 2, \dots, p)$  в рівняннях Юла-Уокера визначають взаємозв'язки між значеннями часу в часових рядах. Ці рівняння можна виразити у матричному вигляді і використовуються вони для оцінки параметрів моделей ARMA/ARIMA, а також базуються на обчислених автокореляціях  $\rho_k$  (для  $k = 1, p$ ):

$$R[i] = \sum_{j=1}^p \Phi[j] R[i-j] = 0, i = 1 \div p,$$

$$R[0] - \sum_{j=1}^p \Phi[j] R[j] = D_a,$$

де  $R[j] = E\{x[t]x[t-j]\}$  – значення функції кореляції випадкового процесу,

$D_a = \sigma_a^2$  – дисперсія помилок передбачення моделі AP,

$R[0]$  – дисперсія випадкового процесу  $x[t]$ .

У записі  $AR(p)$ ,  $p$  – це порядок моделі. Трохи видозмінений запис представленої раніше моделі ковзного середнього  $MA(q)$ :

$$X_t = \mu + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t,$$

де  $\mu$  – середнє значення вибірки,

$\theta_j (j = 1, 2, \dots, q)$  – параметри моделі та  $q$  порядок моделі.

Тоді  $ARMA(p, q)$  модель:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j},$$

$ARMA$  модель працює з припущенням, що ряд стаціонарний. Операція усунення тренда у разі може призвести до стаціонарності низки чи навіть усунути сам тренд. Якщо в якості вихідних даних використовується не сам часовий ряд, а різниця його компонентів, то в такому випадку модель називається  $ARIMA(p, d, q)$  – Autoregression integrated moving average, де  $d$  – порядок різниці між значеннями ряду. На основі  $ARMA$  моделі розроблено багато інших прогнозуючих моделей.  $FARIMA$  – одна з них, коли операція детрендингу проводиться з дробовим порядком [16].  $SARIMA$  – модель, в якій враховано сезонну складову ряду [87].

## 2.2 Прогнозування трафіку з використанням нейронної мережі

Нейронні мережі користуються великою популярністю в задачі прогнозування часових рядів через можливість імітації процесів із множинним зовнішнім впливом. Ефективні у завданнях класифікації об'єктів та

розпізнавання образів. Принципово структура нейрона ANN схожа за будовою на людський нейрон і представлена рис. 2.1.

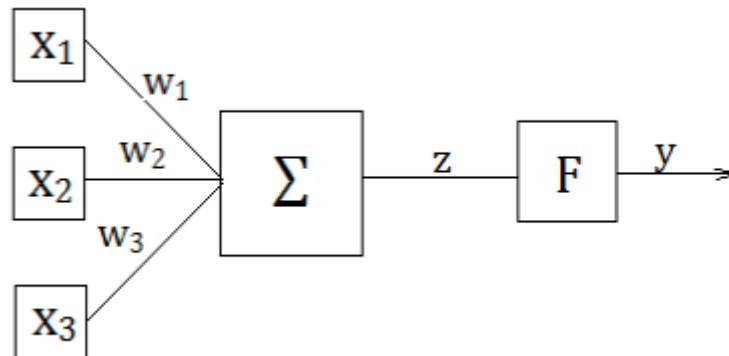


Рисунок 2.1 – Принципова схема нейрона

На вхід штучного нейрона надходить вектор вхідних сигналів  $X_1, X_2, X_3, \dots, X_n$ , котрим вираховується середнє значення  $z$ , що у активаційної функції  $F$  до розрахунку вихідного значення  $Y$ . Активаційна функція може бути функцією одиничного стрибка, лінійної функцією, тощо, наприклад:

$$F(z) = \frac{1}{1 + e^{-z}}.$$

Приклад ANN, що складається з 3 вхідних та 2 прихованих нейронів, представлений на рисунку 2.2. Вхідний шар служить для розподілу вхідного вектора значень і не містить нейронів. Зазначимо, що нейрони в повному обсязі обов'язково мають бути пов'язані між собою, тобто для деяких зв'язків  $w$  може бути нульовим. Також нейронна мережа може мати зворотний зв'язок для переналаштування ваг. Вагові коефіцієнти зв'язків призначаються у процесі навчання, і навіть можуть перераховуватися залежно від результатів роботи мережі.

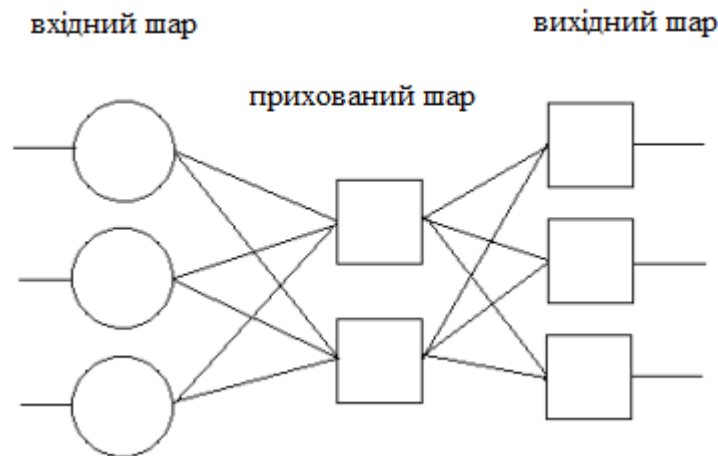


Рисунок 2.2 – Структурна схема нейронної мережі

Відповідно до [18] можна виділити основні особливості ANN (Artificial Neural Network):

ANN можуть моделювати нелінійні залежності між вхідними та вихідними даними.

ANN навчається на вхідних даних, тобто дані визначають модель між вхідними та вихідними даними.

ANN можуть узагальнювати вхідні дані, що призводить до зниження ефективності роботи мережі за зміни характеру даних.

На відміну від статистичних моделей, ANN вносить припущення про розподіл вхідних даних.

У ході використання ANN виникає низка труднощів, які згідно [19] виглядають наступним чином:

Складність підбору оптимальної комбінації параметрів мережі, таких як швидкість навчання, кількість прихованих шарів, прихованих нейронів на кожному шарі.

Складність вибору алгоритму навчання та суворі вимоги до навчальної вибірки.

Навчена нейронна мережа служить деякою «чорною скринькою», тобто досить складно виділити сукупність правил, за якими ANN приймає рішення.

Дуже нетривіальним є завдання вибору відповідної моделі прогнозування, а також добору параметрів для лінійної статистичної моделі або налаштування

нелінійної моделі. Застосування генетичних алгоритмів може допомогти у вирішенні такого завдання.

Перші роботи у цьому напрямі описували використання генетичних алгоритмів у доборі коефіцієнтів для регресійних [10] і авто регресійних моделей [11]. При цьому популяція формувалася ковзним вікном по ряду даних, а коефіцієнти регресії служили вектором генів. Після подальших схрещувань та мутацій формувалася нова популяція, пристосованість якої визначалася як точність прогнозування низки.

У більш сучасних роботах генетичні алгоритми використовуються для налаштування штучних нейронних мереж, тобто вибору архітектури мережі: числа вхідних нейронів, числа вихідних нейронів, ваги синапсів (у тому числі нульові, якщо необхідно усунути зв'язок між деякими нейронами).

Алгоритм пошуку оптимальної ANN для деякого числового ряду наступний:

Випадково генерується вихідна популяція (набір хромосом) нейронних мереж.

Розраховується фенотип (архітектура ANN) та значення оптимізаційної функції для кожної особи популяції.

Застосовуються методи генетичних алгоритмів, такі як елітизм, відбір вдалих особин, кросовер та мутація, для формування нової популяції.

Попередні два кроки повторюються кілька разів до настання ефекту перенавчання мережі.

Методи локальної апроксимації застосовуються для прогнозування хаотичних і квазіперіодичних рядів, тобто процесів, у яких може бути відсутня глобальна лінійна складова [13]. Тобто прогнозування ґрунтується на локальній підпоследовності ряду, причому модель все ще може залишатися лінійною в рамках деякої локальної вибірки. Вибірка ж формується не за часовою близькістю значень низки, а, по близькості у просторі затримок.

Основна особливість методу у тому, що довжина прогнозу визначається можливостями моделі, а динамічними властивостями самого ряду. При цьому

методом локальної апроксимації оцінюється динаміка низки, а аналіз основних закономірностей здійснюється з допомогою інших методів, наприклад сингулярного спектрального аналізу (SSA).

Алгоритм побудови прогнозу на один крок уперед приблизно наступний [14]:

Побудова матриця затримок та вибір локального уявлення.

Визначення числа сусідів.

Оцінка параметрів обраної моделі та побудова прогнозу у припущенні, що  $X$  змінюється за тим самим законом і з тими самими параметрами, що його сусіди.

Прогнозування значень ряду методом локальної апроксимації може бути виконано одним із наступних способів:

Ітеративний. При цьому коефіцієнти впливу відхилень у рамках кожної з матриць затримок та середнє значення прогнозів від кожного із сусідів розраховуються один раз, а спрогнозований вектор використовується як новий стартовий.

Ітеративний з перерахунком. Відмінність від попереднього полягає у перерахуванні всіх параметрів після отримання прогнозу та вибору нових сусідів.

Прямий. Полягає в тому, що після обчислення прогнозного значення, воно не додається до вихідних даних у подальших розрахунках.

Методи локальної апроксимації мають багато спільного з широко відомими авторегресійними методами прогнозування рядів.

Розглянуті алгоритми прогнозування процесів є загальновідомими і в завданнях дослідження часових рядів застосовуються досить давно. Однак для прогнозування реального процесу, в силу його характеру, деякі моделі підходять краще за інших. При цьому підбір відповідної моделі, що описує досліджуваний процес, є нетривіальним завданням, рішення якого необхідно базувати на точних статистичних та динамічних даних про природу самого процесу.

Згідно з відомою методологією Бокса-Дженкінса [20], аналіз та прогноз часових рядів можна розділити на три етапи:

Ідентифікація моделі, що найбільш точно описує процес:

- перевірка низки на стаціонарність;
- аналіз повної та приватної АКФ.

Проведення оцінки моделі та перевірка її адекватності:

- оцінка параметрів моделі, розрахованих першому етапі.
- перевірка моделі на відповідність вихідним даним.

Розрахунок прогнозних значень низки.

## 3 АНАЛІЗАЛГОРИТМІВ ЗГЛАДЖУВАННЯ

### 3.1 Формування функцій згладжування трафіку

Аналіз та прогнозування трафіку є важливими складовими управління сучасними телекомунікаційними системами та мережами. На основі подібних прогнозів своєчасно виділяються та резервуються ресурси обладнання, необхідні для підтримки належної якості надання послуг. Проблемам точного прогнозування трафіку у провідних широкосмугових мережах та стільникових мережах присвячено принаймні два напрями досліджень. Одне з них засноване на припасуванні моделей (наприклад, on-off модель, модель ARIMA [4], модель FARIMA [5], модель мобільності [6], модель мережевого трафіку [7] та альфа-стабільна модель [8]) для вивчення характеристик трафіку, таких як просторова та тимчасова релевантність або самоподібність [9], та отримання майбутнього трафіку за допомогою відповідних методів прогнозування. Інший напрямок досліджень заснований на сучасних методах обробки сигналів (наприклад, метод аналізу основних компонентів, метод фільтрації Калмана [10] або метод стиснення [11]) для реєстрації еволюції трафіку.

Останнім часом було проведено велику роботу з дослідження динамічних характеристик бездротового трафіку, зокрема з урахуванням їх нестационарності і сезонності, збільшення точності прогнозів [12–19]. Прогнозування стільникового трафіку в цих роботах можна розглядати як завдання аналізу часових рядів, продуктивність якої залежить від її лінійних статистичних моделей, таких як авторегресійна інтегрована ковзна середня (ARIMA) та альфа-стабільна модель.

Метод експоненційного згладжування (Exponential Smoothing Method) був запропонований QT Tran та ін для прогнозування трафіку стільникової мережі; N. Sapankevych and R. Sankar застосували лінійну регресію (Linear Regression LR), а J. Wang зі співавторами використовували регресію машини опорних векторів

(SupportVectorMachineRegression - SVMR) для прогнозування трафіку стільникової мережі.

Дослідження функцій згладжування трафіку проведемо на наборі даних зібраних з мобільних пристроїв, встановлених у кооперативній інтелектуальній транспортній системі (C-ITS) Швеції. Датасет був сформований з даних, що пересилаються автобусами на точки збору даних IoT, і включає 44 844 рядків запису з 40 стовпцями категорій. Серед 40 категорій, багато з яких не значущі для даного дослідження, слід виділити такі важливі категорії, як:

- index – унікальний номер транзакції;
- time stamp – часова мітка початку транзакції (дата та час);
- transaction time – час транзакції (у секундах);
- protocol – протокол, що використовується (UDP, TCP або HTTPS).

Для швидкої візуальної оцінки відсутніх даних побудовано діаграму (рис. 3.1), на ній білим кольором відзначені відсутні дані.

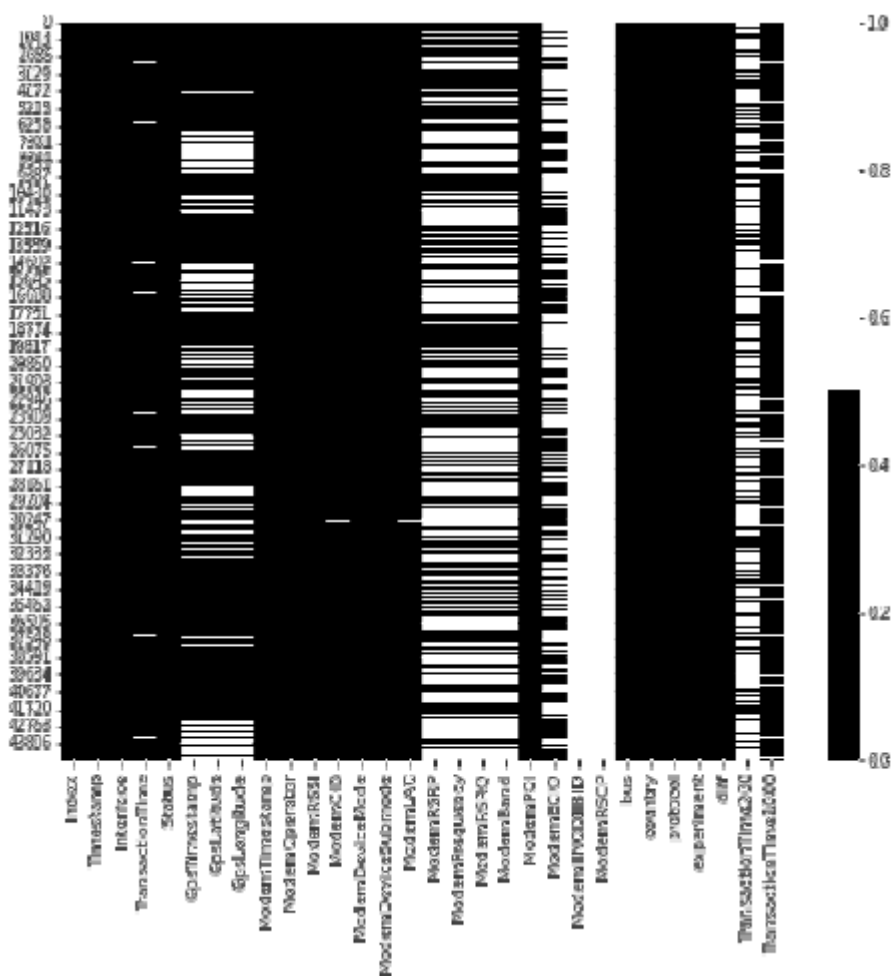


Рисунок 3.1 – Діаграма, що ілюструє відсутні дані (відображаються білим кольором)

Для стовпців, що цікавлять нас, відсоток відсутніх даних становить: Index – 0 %, Timestamp – 0 %, TransactionTime – 3 %, protocol – 0 %.

Дані TransactionTime, що відсутні, були доповнені медіанним значенням (рівним 0.305502 с), розрахованим по стовпцю TransactionTime.

При аналізі даних важливо визначити дані, які дуже відрізняються від більшості (outliers – «викиди»). В даному випадку найцікавішими є дані про час транзакції: чи існують «викиди», і чи є ці дані аномальними. Побудуємо стовпчасту діаграму, що ілюструє дані про час транзакцій TransactionTime (рис. 3.2).

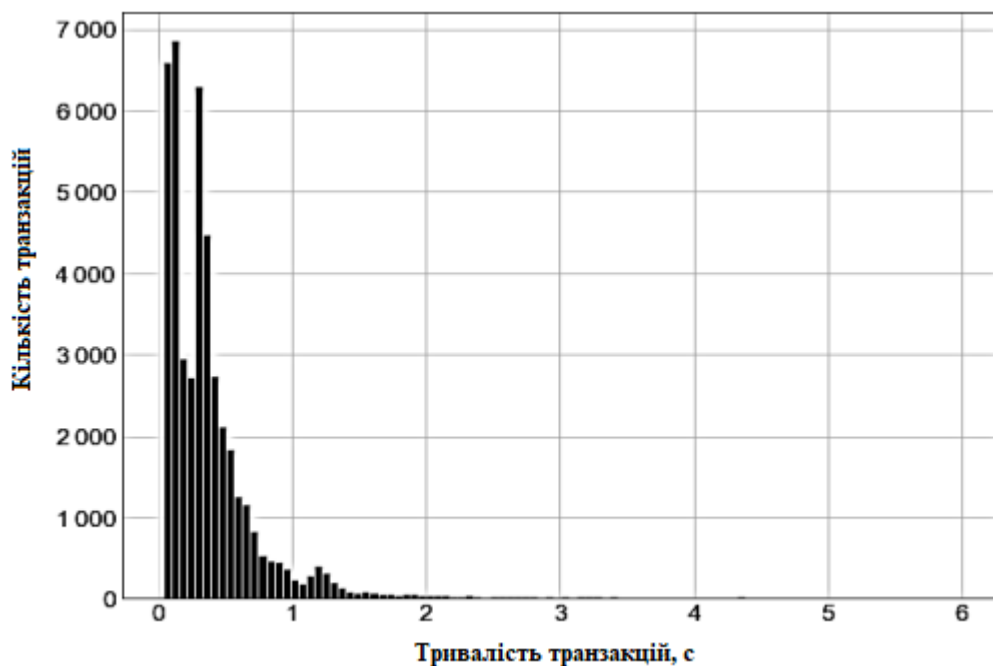


Рисунок 3.2 – «Викиди» даних часу транзакцій (хвіст діаграми)

Наочніше аномальні дані представлені на діаграмі boxplot (рис. 3.3).

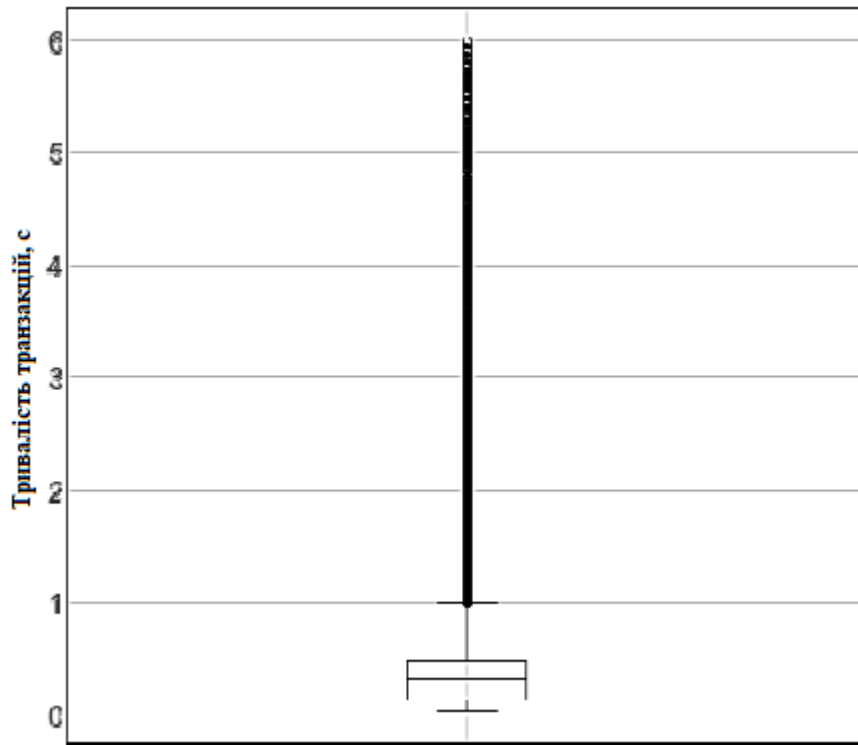


Рисунок 3.3 – Діаграма boxplot

Діаграма "ящик з вусами" (boxplot) є графічним методом візуалізації розподілу даних. В даному випадку ця діаграма наочно ілюструє, що більша кількість транзакцій має тривалість від десятих часток секунди до півсекунди, а транзакції, що мають тривалість приблизно більше 1 с, вже є відносно рідкісними в даних. Boxplot (рис. 3.3) складається з «коробки», «вусиків» та крапок. Коробка показує інтерквартильний розмах розподілу, тобто, відповідно, 25 % ( $Q1$ ) та 75 % ( $Q3$ ) перцентілі. Характеристика всередині коробки означає медіану розподілу. Вуса відображають весь розкид точок, крім викидів, тобто мінімальні та максимальні значення, які потрапляють у проміжок ( $Q1 - 1,5 \times IQR, Q3 + 1,5 \times IQR$ ), де  $IQR = Q3 - Q1$  – інтерквартильний розмах, а  $\times$  - знак множення. Точками на графіці позначаються викиди – ті значення, які вписуються у проміжок значень, заданий вусами графіка (на рис. 3.3 це помітно для транзакцій з тривалістю від 1 до 6 с). Проаналізуємо описову статистику за наданими даними (табл. 3.1).

Таблиця 3.1 – Описова статистика датасету для TransactionTime

Показник	Значення
Середнє значення	0,413150
Стандартне відхилення	0,509285
Мінімальне значення	0,033924
Квартиль 25%	0,129970
Квартиль 50%	0,305502
Квартиль 75%	0,480927
Максимальне значення	5,982710

Приміром, для ознаки TransactionTime 5,98 – максимальне значення, тоді як 75 % кварталів дорівнює лише 0,48. Проте значення транзакції в 5,98 с не є викидом – це цілком реалістичний час деяких транзакцій. Тим не менш, можливо, що при навчанні моделі ці дані доцільно відкинути.

Хоча збір даних і проводився три місяці, він проводився нерегулярно, і тому подальший аналіз можливий лише окремих серій даних, відповідних дням збору інформації. Наприклад, на 13 грудня динаміка зміни тривалості транзакцій протягом дня з урахуванням протоколу, що використовувався, виглядає наступним чином (рис. 3.4).

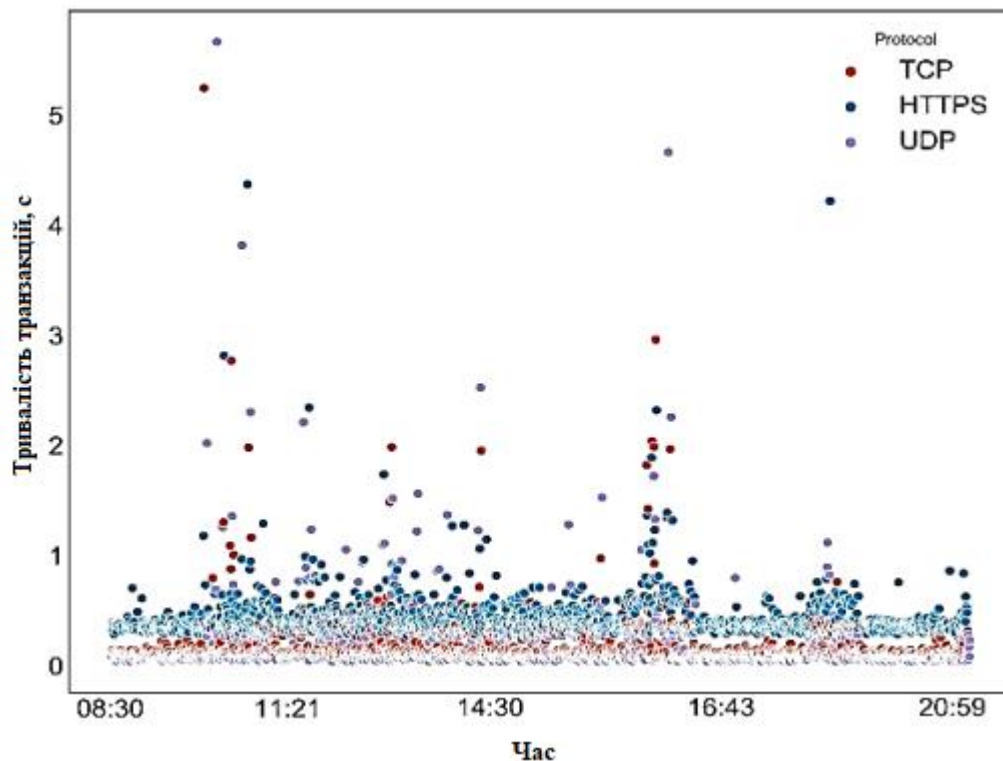


Рисунок 3.4 – Динаміка зміни тривалості транзакцій протягом дня з урахуванням протоколу, що використовувався

На графіку, зображеному на рис. 3.4, простежуються тенденції, що пов'язують використовуваний протокол із часом транзакції – щільні скупчення точок. Помітні й викиди – окремі точки, які далеко від скупчень.

### 3.2 Аналіз якості варіантів згладжування трафіку

Для аналізу та короткострокового прогнозування обмежимося однією годиною трафіку, зібраного в один із днів, при цьому вважатимемо, що протокол передачі неважливий. В якості даних продовжуємо використання попереднього набору даних зібраних з мобільних пристроїв, встановлених у кооперативній інтелектуальній транспортній системі (C-ITS) Швеції. Датасет був сформований з даних, що пересилаються автобусами на точки збору даних IoT, і включає 44 844 рядків запису з 40 стовпцями категорій.

Особливість даних трафіку полягає в тому, що отримані дані нерегулярні – в одні часові рамки збирається велика кількість даних, в інші транзакції відсутні і записів немає. Таким чином, з потоку цих даних утворюється нерегулярний часовий ряд. Для роботи з такими рядами вони мають бути приведені до регулярних рядів шляхом агрегації даних на часових ділянках, які були взяті з певною частотою. В даному випадку була обрана частота часового ряду в 10 с, а як агрегатор – середній час транзакції за цей час. Якщо за цей час не було жодної транзакції, середній час дорівнював нулю (при прогнозуванні майбутніх значень ряду іноді доцільно вважати його дуже малим, але не нульовим, або інтерполювати його за сусідніми значеннями). Аналізуватимемо середні значення часу транзакції у десятисекундний проміжок часу. Для згладжування даних застосовується багато методів. Одним із популярних методів є згладжування за допомогою ковзного середнього. На рис. 3.5 та 3.6

представлені графіки згладжування для вікон у 30 та 60 с (шкала часу представлена у форматі «години: хвилини»).

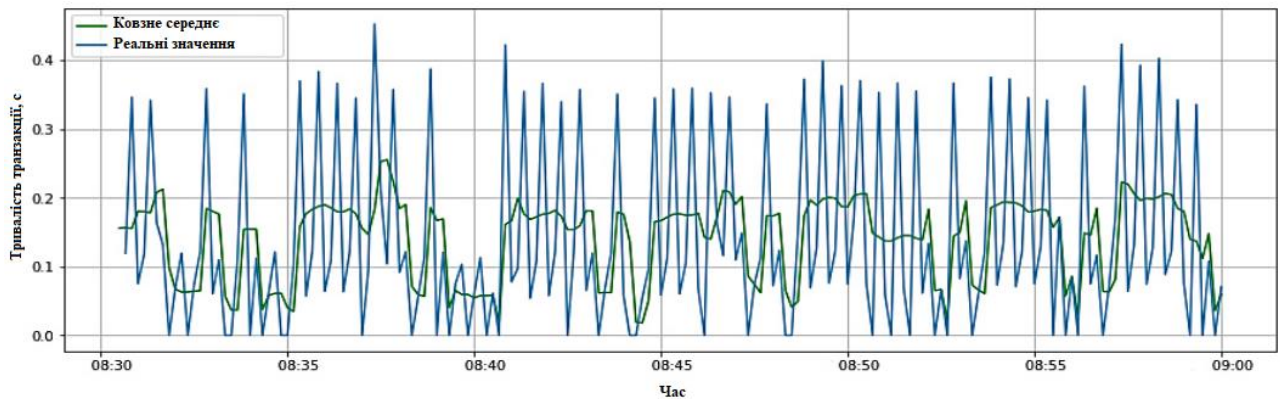


Рисунок 3.5 – Згладжування даних за допомогою ковзного середнього (розмір вікна 30 с)

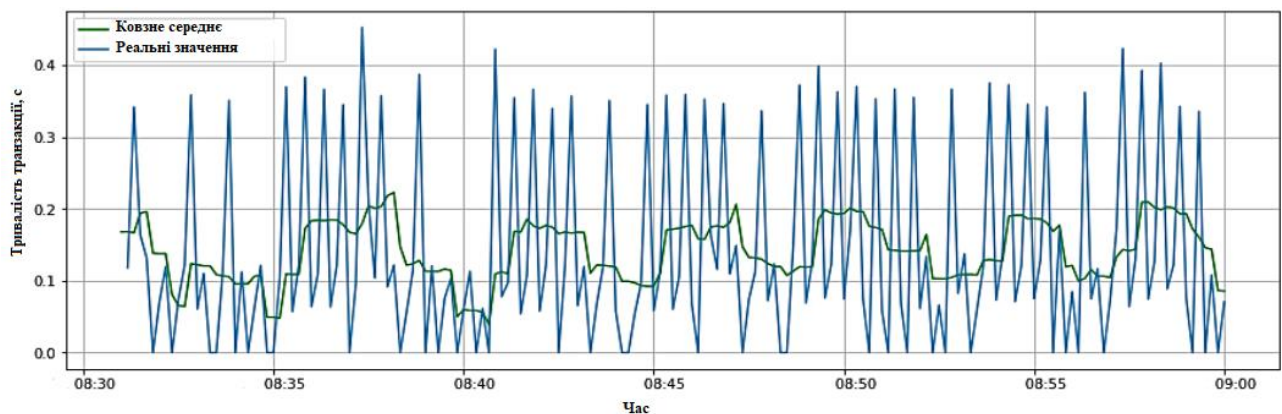


Рисунок 3.6 – Згладжування даних за допомогою ковзного середнього (розмір вікна 60 с)

Іншим методом згладжування даних є експонентне згладжування. Для цього методу основним параметром згладжування є коефіцієнт  $\alpha$  (може набувати значення від 0 до 1). Він визначає вагу останнього спостереження під час обчислення прогнозу. Вибір значення  $\alpha$  впливає на швидкість реакції на зміни даних і ступінь згладжування шумів. Маленькі значення  $\alpha$  роблять метод більш стійким до шумів, але менш чутливим до змін тренду та сезонності. Великі значення  $\alpha$  роблять метод більш чутливим до останніх даних, але і більш

схильним до шумів. Для різних значень графік згладжування представлений на рис. 3.7 (шкала часу представлена у десятках секунд).

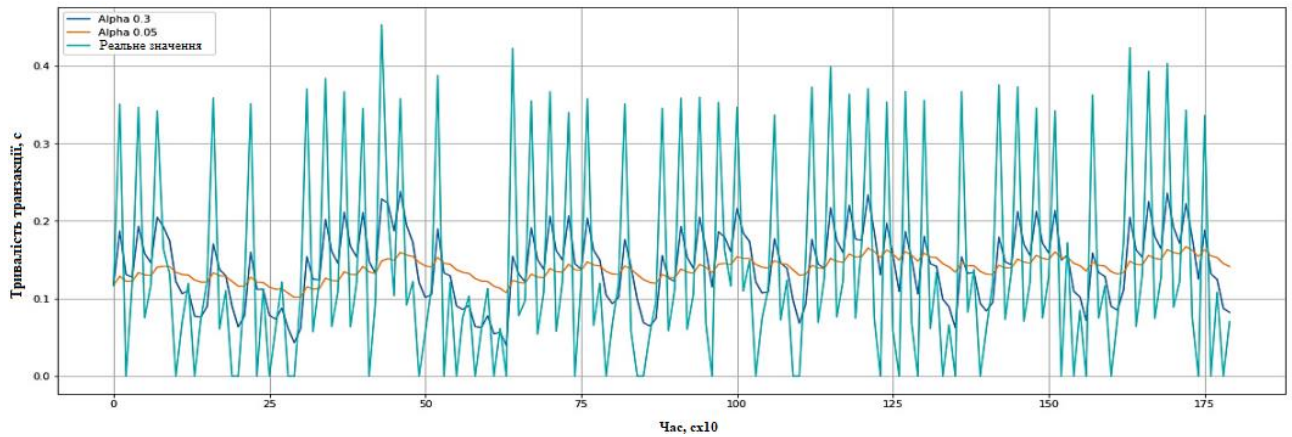


Рисунок 3.7 – Експонентне згладжування даних

Прогнозування параметрів трафіку сучасних телекомунікаційних систем має велике значення для їхнього розвитку та надання якісних послуг абонентам. На основі подібних прогнозів відбувається виділення ресурсів системи: провідних та радіоканалів, смуги частот, проводиться автоматична реконфігурація обладнання у разі збоїв або при впровадженні нової апаратури зі зв'язку.

## 4 ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ПРОГНОЗУВАННЯ ТРАФІКУ

### 4.1 Розрахунок точності найкращого лінійного прогнозу

Експонентне згладжування та моделі ARIMA є двома найбільш широко використовуваними підходами до прогнозування і є потужним інструментом при роботі з часовими рядами [20]. У той час як моделі експоненціального згладжування засновані на описі тренду та сезонності даних, моделі ARIMA спрямовані на опис автокореляцій даних часового ряду.

Модель ARIMA складається з автокореляційної частини (модель AR), яка представлена формулою

$$y_t = \varphi_0 + \varphi_1 \cdot y_{t-1} + \varphi_2 \cdot y_{t-2} + \varphi_p \cdot y_{t-p} + \varepsilon_t,$$

де  $y_t$  – значення часового ряду в момент часу  $t$ ;

$\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$  – коефіцієнти автокореляції, які визначають зв'язок між поточним значенням та попередніми значеннями ряду;

$y_{t-1}, y_{t-2}, y_{t-p}$  – значення часового ряду в моменти часу  $t - 1, t - 2, \dots, t - p$  відповідно;

$\varepsilon_t$  – випадковий шум, що описує вплив змінних, не врахованих в моделі.

Модель із ковзним середнім (модель MA) порядку  $q$  задається рівнянням

$$y_t = \varepsilon_t - \theta_1 \cdot \varepsilon_{t-1} - \theta_2 \cdot \varepsilon_{t-2} - \dots - \theta_q \cdot \varepsilon_{t-q},$$

де  $\varepsilon_t$  – поточна помилка (випадковий шум);

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$  – значення залишківі часових періодів тому (попередні помилки);

$\theta_1, \theta_2, \dots, \theta_q$  – коефіцієнти ковзного середнього, які визначають впливи попередніх помилок на поточне значення часового ряду.

Моделі ковзного середнього прогнозують значення функції на основі лінійної комбінації обмеженого числа залишків  $q$ , авторегресійні моделі прогнозують значення, використовуючи лінійну функцію апроксимації обмеженого числа  $p$  минулих значень.

Модель ARIMA (модель Бокса – Дженкінса) поєднує ці моделі:

$$y_t = c + \varphi_1 \cdot y_{t-1} + \dots + \varphi_p \cdot y_{t-p} + \theta_1 \cdot \varepsilon_{t-1} + \dots + \theta_q \cdot \varepsilon_{t-q} + \varepsilon_t,$$

де  $c$  – константа.

Модель ARIMA ( $p, q, d$ ) має три параметри:

$p$  – порядок авторегресії. Він дозволяє відповісти на запитання, чи буде черговий елемент ряду близький до значення  $X$ , якщо до нього були близькі  $p$  попередніх значень;

$q$  – порядок ковзного середнього. Дозволяє встановити похибку моделі як лінійну комбінацію значень помилок, що спостерігалися раніше;

$d$  – порядок інтегрування. Він показує, наскільки елемент ряду близький за значенням до  $k$  попередніх значень, якщо різниця між ними мінімальна.

Багато моделей є окремими випадками [19] моделі ARIMA:

білий шум – модель ARIMA (0, 0, 0) без константи;

випадкове блукання – модель ARIMA (0, 1, 0) без константи;

випадкове блукання з дрейфом – модель ARIMA (0, 1, 0) з константою;

авторегресія – модель ARIMA ( $p$ , 0, 0);

ковзне середнє – модель ARIMA (0, 0,  $q$ ).

До моменту навчання моделі необхідно переконатися, що часовий ряд є стаціонарним. Стаціонарний часовий ряд – це ряд, статистичні властивості якого залежить від часу спостереження за цим рядом. Стаціонарність низки визначається критерієм Дікі - Фуллера [21].

В даному випадку він дорівнює 0,001655 і ряд можна вважати стаціонарним.

Для автоматичного підбору [22] параметрів моделі використовувалася функція `auto_arima` із бібліотеки `pmdarima`. Найкращими параметрами моделі стали параметри (3, 0, 0), і представлена модель має авторегресійний характер. На рис. 4.1 представлені графік навчання та короткостроковий прогноз інтенсивності трафіку незалежно від використовуваного протоколу передачі.

Для аналізу даних та прогнозу інтенсивності трафіку дані були поділені за окремими протоколами: TCP, UDP, HTTPS. Для аналізу та прогнозу використовувалися такі моделі: наївна сезонна, експоненціальне згладжування, лінійна регресія, ARIMA, Theta та N-BEATS. Як інструмент була використана бібліотека Darts (для мови програмування Python). Вона підтримує різні підходи до прогнозування часових рядів: від класичних статистичних моделей, таких як ARIMA та експоненціальне згладжування, до методів, що базуються на машинному та глибокому навчанні. Крім цього бібліотека Darts включає функції аналізу статистичних властивостей часових рядів, а також оцінки точності моделей прогнозування.

Наївна сезонна модель, як випливає з назви, досить проста: вона повторює кілька останніх значень часового ряду. При успішному доборі цих значень модель, незважаючи на простоту, може давати хороші результати.

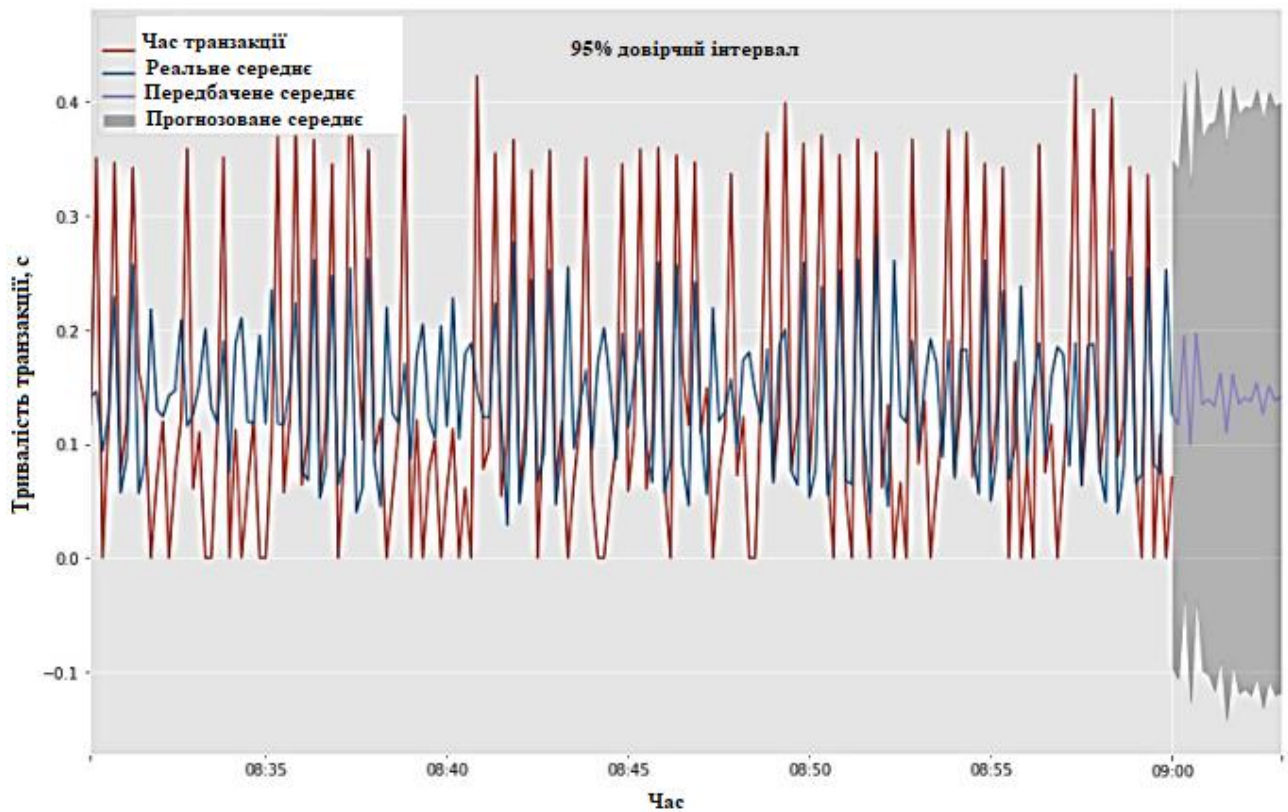


Рисунок 4.1 – Графік навчання моделі ARIMA та короткостроковий прогноз тривалості транзакції (*передачі даних*)

Раніше вже згадуване експоненціальне згладжування – одна з найпопулярніших моделей як для згладжування даних, так і для прогнозування часових рядів. Цей метод був розроблений Брауном і Холтом, і в його основі лежить розрахунок експоненціальних ковзних середніх рядів, що згладжуються.

Лінійна регресія також є однією з моделей, що часто використовуються. Її перевагами є відносна простота і, як наслідок, хороша швидкість обробки даних та прогнозування. Найчастіше це є визначальною характеристикою використання цієї моделі.

Тета-метод, розроблений Асимакопулосом і Ніколопулосом, знаходить дедалі більше застосування при прогнозуванні часових рядів. Детальний опис методу представлено у роботі [24], модифікація обговорюється у роботі [25].

Метод N-BEATS [27] відноситься до методів прогнозування з використанням нейронних мереж. Архітектура нейронної мережі N-BEATS

використовує не рекурентні, а повнозв'язні шари і найбільш ефективно застосовується для одновимірних часових рядів (рис. 4.2).

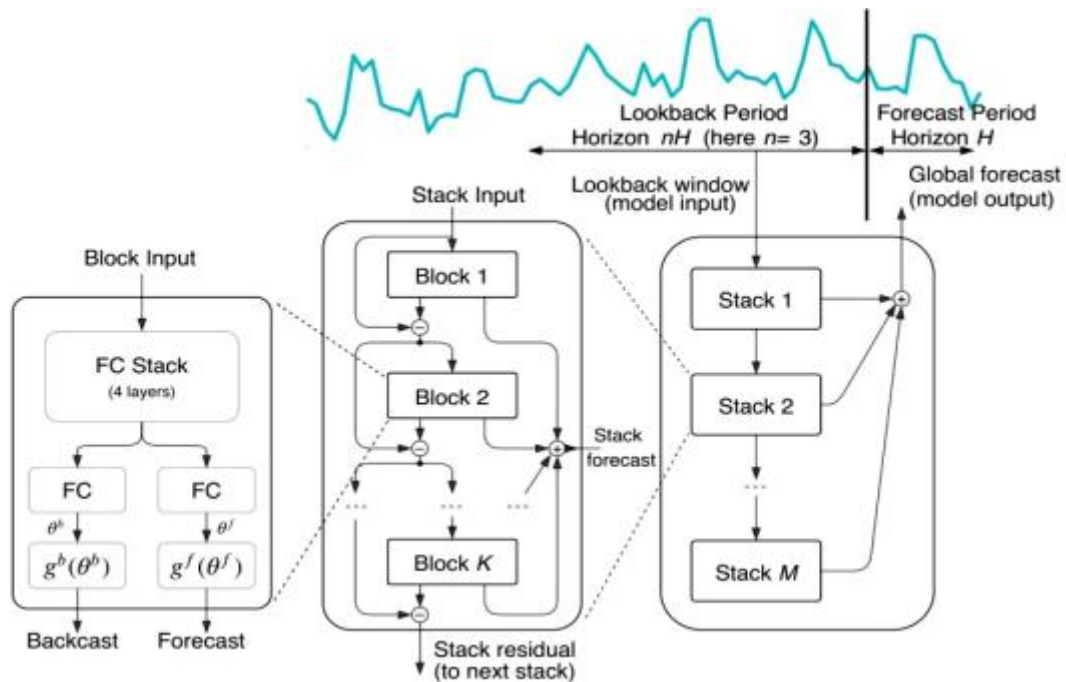


Рисунок 4.2 – Архітектура нейронної мережі N-BEATS

Архітектура N-BEATS і двох основних блоків: блоків «прогнозування» (Forecast) і блоків «оновлення» (Backcast). Блок прогнозування відповідає за прогнозування майбутніх значень часового ряду. Кожен такий блок складається з кількох повнозв'язкових шарів (FS – fullyconnectedlayers) та шарів нормалізації. Блок оновлення, навпаки, використовується для оновлення передбачення, ґрунтуючись на попередніх значеннях часового ряду. Він також складається з повно зв'язних шарів та шарів нормалізації. Блоки оновлення служать для корекції помилок передбачень, що вносяться в процесі прогнозування. У мережі N-BEATS часовий ряд подається на вхід і мережа пропускає його через кілька повнозв'язкових шарів. При цьому вихід даних із останнього шару поділяється на дві частини: частина Backcast та частина Forecast. Цю структуру називають блоком (Block) мережі. Вихід Backcast віднімається з вхідного вектора перед наступним блоком, а вихід Forecast додається до прогнозу вектора від попереднього блоку. Така структура мережі дозволяє частині Backcast

відфільтрувати лише необхідну інформацію із вхідних даних, а частини Forecast скоригувати прогноз та зменшити помилку. Послідовність блоків називається стеком (Stack). Виходи з усіх стеків підсумовуються, і таким чином виробляється глобальний прогноз.

Як показники якості навчання моделей використовувалися такі метрики [23]:

- середня абсолютна помилка (MAE);
- корінь із середньої квадратичної помилки (RMSE);
- середня абсолютна процентна помилка (MAPE);
- симетрична MAPE (SMAPE).

Для протоколу UDP найбільш вдалою моделлю стала модель ARIMA. Прогнозування тривалості транзакцій UDP за допомогою методу ARIMA представлено на рис. 4.3.

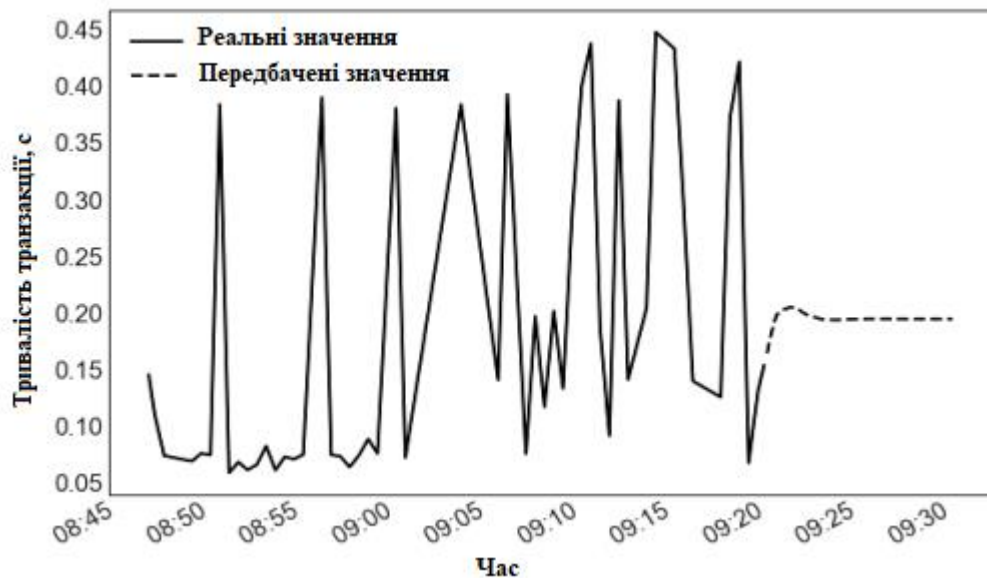


Рисунок 4.3 – Прогнозування тривалості транзакцій UDP за допомогою методу ARIMA

Для протоколу CP вибирається лінійна регресія та моделі Theta (рис. 4.4, 4.5).

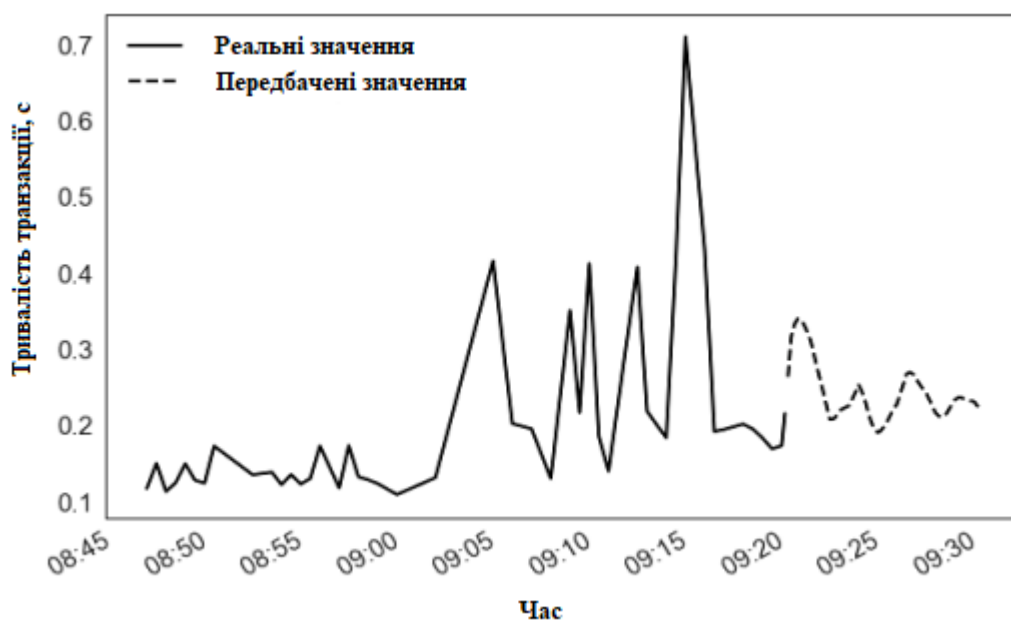


Рисунок 4.4 – Прогнозування тривалості транзакцій ТСП лінійною регресією

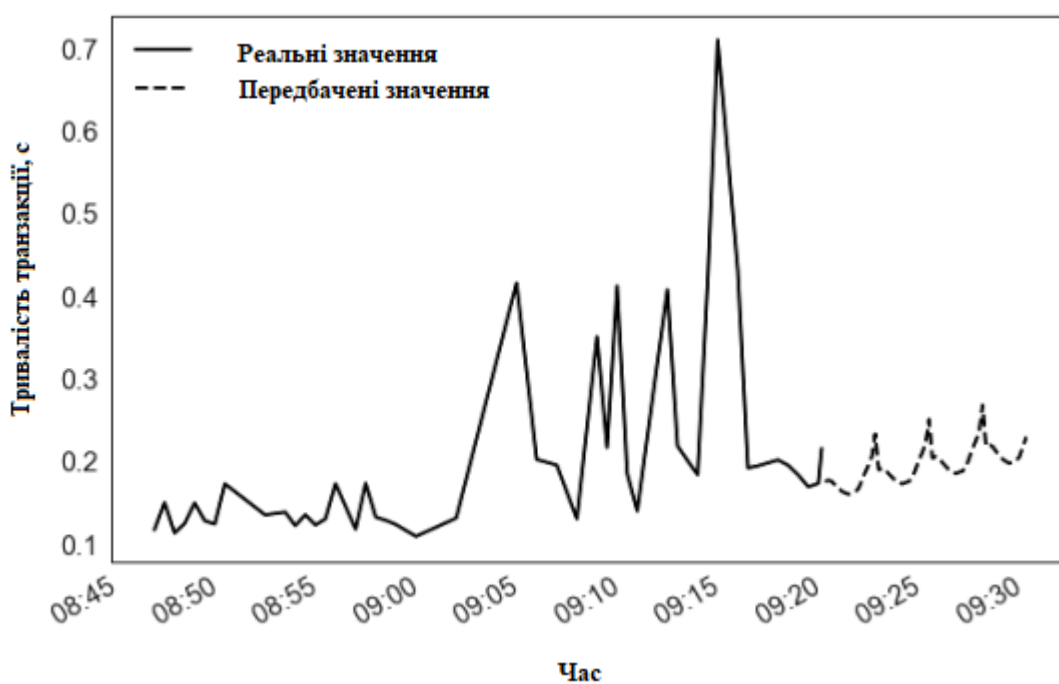


Рисунок 4.5 – Прогнозування тривалості транзакцій ТСП моделлю Theta

Для протоколу HTTPS – лінійна регресія, ARIMA та N-BEATS (рис. 4.6).

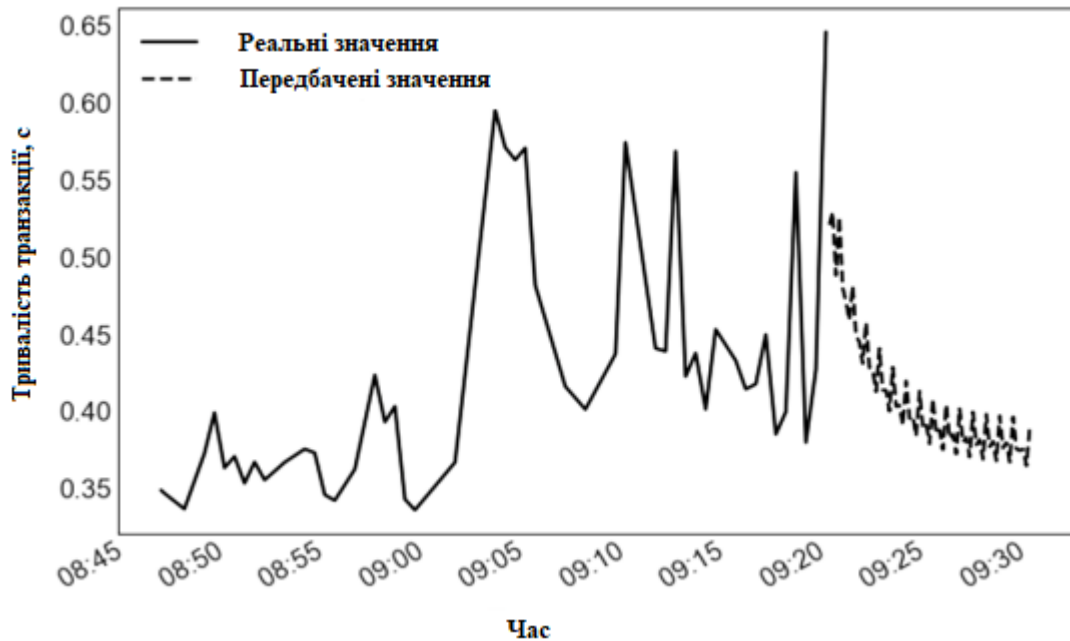


Рисунок 4.6 – Прогнозування тривалості транзакцій HTTPS моделлю N-BEATS

Порівняльний аналіз метрик якості навчання представлений у табл. 4.1.

Таблиця 4.1 – Значення метрик якості навчання моделей

Модель	Метрика			
	MAE	RMSE	MAPE	SMAPE
Протокол UDP				
Наївна сезонна	0,0910	0,1242	56,4889	44,5716
Експонентне згладжування	0,0607	0,0834	32,1989	41,5392
Лінійна регресія	0,0512	0,0599	30,5981	27,4312
ARIMA	0,0466	0,0550	27,9993	25,1173
Theta	0,2794	1,0059	185,8544	63,7469
N-BEATS	0,0585	0,0663	35,4984	30,8297
Протокол TCP				
Наївна сезонна	0,0647	0,1011	23,0194	24,0684
Експонентне згладжування	1,2578	1,4767	604,3389	127,4042
Лінійна регресія	0,0344	0,0482	13,2520	13,1049
ARIMA	0,0742	0,0801	32,0026	28,6844
Theta	0,0582	0,1003	18,5042	22,9805
N-BEATS	0,0645	0,0957	22,9030	25,3388
Протокол HTTPS				
Наївна сезонна	0,0845	0,1035	16,7633	17,5413
Експонентне згладжування	2,4104	2,7390	527,2670	130,8127

Лінійна регресія	0,0546	0,0679	11,3512	11,4193
ARIMA	0,0658	0,0779	14,8398	13,4941
Theta	0,0968	0,1134	21,6294	18,7798
N-BEATS	0,0777	0,0911	15,3963	16,9991

При короткостроковому прогнозуванні важливим чинником є час, який витрачається обробку даних: що менше часу і обчислювальних потужностей потрібно для роботи моделі, тим краще. З цієї позиції відносно прості моделі, наприклад, лінійна регресія, метод ковзного середнього або наївна модель, більш кращі в порівнянні зі складними моделями (ARIMA, N-BEATS), якщо забезпечують прийнятну якість прогнозу.

Порівняльний аналіз різних моделей навчання дозволяє зробити такі висновки: для протоколу UDP для часового ряду, що досліджується, найбільш підходящою моделлю стала модель ARIMA, для протоколу TCP – лінійна регресія і модель Theta, для протоколу HTTPS – лінійна регресія, ARIMA і N-BEATS. Якість навчання моделей та якість прогнозування сильно залежить від вихідного часового ряду, якості та повноти зібраних даних. Найчастіше відносно прості методи, такі як наївна сезонна модель або лінійна регресія, дають прийнятну якість короткострокового прогнозу, маючи при цьому значно більшу швидкодію, ніж складніші та точніші моделі.

#### 4.2 Задача прогнозування самоподібних процесів

Насправді випадкові процеси зберігають властивість самоподібності лише до певної межі. Ця межа чи міра статистичної стійкості процесу при багаторазовому масштабуванні визначається так званим параметром Херста  $H$  чи параметром самоподібності. Дамо визначення самоподібного процесу. Нехай даний стаціонарний випадковий процес дискретного аргументу (часу)  $X = (X_1, X_2, \dots, X_t)$ . Позначимо через  $r(k)$  коефіцієнт кореляції. Позначимо через  $X^{(m)} = (X_1^{(m)}, X_2^{(m)}, \dots, X_t^{(m)})$  усереднений по блоках довжини  $m$  процес  $X$ , компоненти якого визначаються рівністю:

$$X_t^{(m)} = \frac{1}{m} (X_{tm-m+1} + \dots + X_{tm}), m, t \in N.$$

Надалі називатимемо такий ряд агрегованим. Відповідно  $r_m(k)$  коефіцієнт кореляції процесу  $X^{(m)}$ . Процес  $X$  є строго самоподібним у широкому значенні слова, з коефіцієнтом Херста  $H$ , якщо виконується умова:

$$r_m(k) = r(k), m \in \{2, 3, \dots\}.$$

Тобто самоподібний процес не змінює свій коефіцієнт кореляції після усереднення блоків. Тобто процес  $X^{(m)}$  не відрізнятиметься від процесу  $X$ , як мінімум щодо статистичних показників другого порядку.

Так само існують асимптотично самоподібні процеси, головна властивість цих процесів полягає в тому, що при  $m \rightarrow \infty$ , процес сходиться до строго самоподібного:

$$r_m(k) \rightarrow r(k), m \rightarrow \infty.$$

Значення коефіцієнта  $H=0,5$  вказують на відсутність довгострокової залежності. Кореляція між подіями відсутня. Ряд є випадковим, а не фрактальним. Що ближче значення  $H$  до 1, то вище ступінь стійкості довгострокової залежності. При  $0 < H \leq 0,5$  часовий ряд є трендонестійким. Він мінливіший, ніж випадковий ряд, оскільки складається з частих реверсів. При  $0 < H \leq 1$  ряд має тривалу пам'ять. Тенденція зміни може бути спрогнозована.

Сучасний мережевий трафік, що включає крім звичайного трафіку передачі даних, трафік VoIP і відео-трафік, так само володіє властивістю самоподібності. При цьому як уже говорилося вище, для таких процесів характерні автокореляційна функція (АКФ), що слабо зменшується, часті сплески і спади активності, циклічна складова. На Рис.4.7 наведено реалізацію

реального добового трафіку через інтернет-провайдера, що проходить через канал зв'язку. На Рис.4.8 наведено графік АКФ отриманого часового ряду.

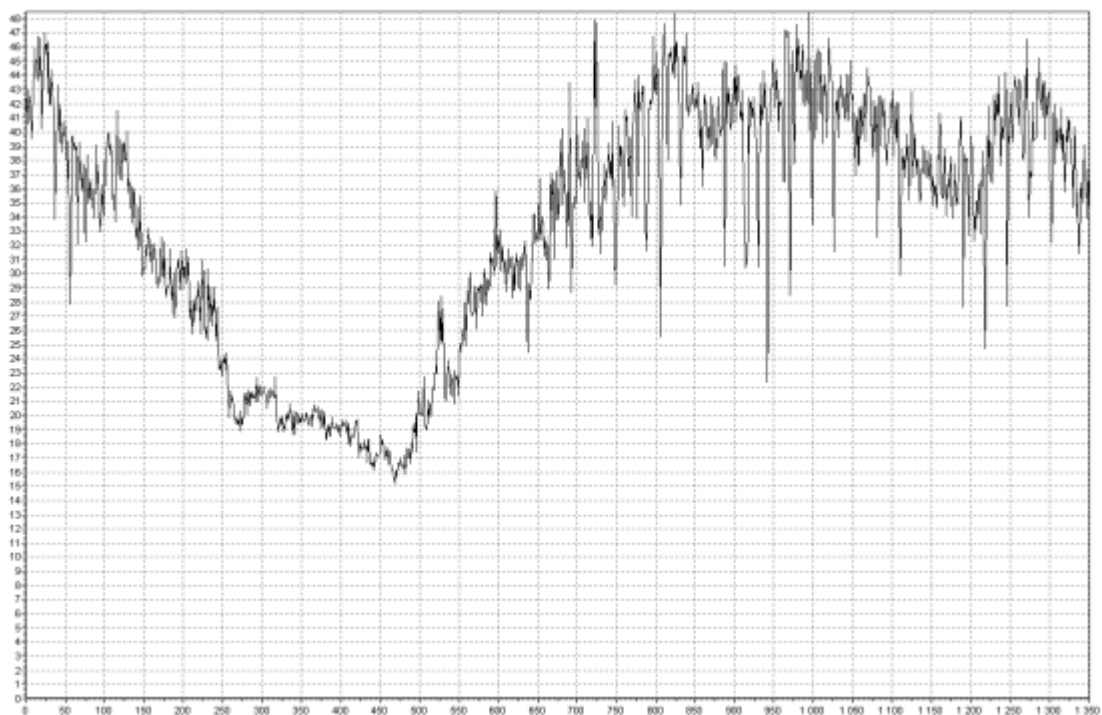


Рисунок 4.7 – Реалізація добового трафіку в Мбіт/сек



Рисунок 4.8 – АКФ реалізації добового трафіку

Незважаючи на досить тривале дослідження самоподібних процесів у телекомунікаціях, залишається низка невирішених завдань:

Фактично відсутня база, яка б замінила класичну теорію масового обслуговування;

Немає загально визнаної моделі самоподібного трафіку;

Немає єдиного методу прогнозування поведінки мережного трафіку.

Завдання прогнозування поведінки мережного трафіку має особливу актуальність на сьогоднішній день. Підбір правильно моделі для оцінювання часового ряду та його прогнозування дозволить вирішити низку важливих завдань, а саме:

Отримувати прогнози щодо доступності смуги пропускання та відповідно модифікувати стратегії запобігання перевантаженням на рівні транспортного протоколу TCP;

Забезпечити оптимальні часові характеристики, наприклад, часову затримку;

Відстежити тенденцію завантаження певних вузлів та зробити прогноз на тривалий час (година, доба, тиждень).

Слід зазначити, що саме властивість довгої пам'яті мережевих процесів дає передумови їх прогнозування і виправдовує застосування AR (авторегресійних) моделей виду:

$$x_t = \varphi_0 + \sum_{k=1}^p \varphi_k x_{t-k} + \varepsilon_t,$$

де  $\varphi_k$  – константи,  $\varepsilon_t$  – білий шум.

Для моделювання процесів з тривалою пам'яттю широкого поширення набули авторегресійні моделі ARIMA (інтегральний процес ковзного середнього) та ARFIMA (дробовий процес ковзного середнього). Також останнім

часом широкої популярності набуває вільний від моделі метод сингулярно-спектрального аналізу (метод Гусениця).

## ВИСНОВКИ

Результати даної роботи вказують на наявність фрактальних властивостей трафіку комп'ютерних мереж.

Для оцінки якості обробки трафіку використовується інтенсивність трафіку  $A$ , що відповідає періоду максимального навантаження або періоду пікового навантаження. Метою оптимального проектування системи зв'язку є забезпечення для заданого числа каналів  $V$  максимально можливого значення питомої інтенсивності трафіку  $A$ , що пропускається однією лінією пучка, або відношення  $A/V$ . Допустиме значення  $A/V$  має відповідати прийнятному значенню показника якості обробки трафіку, протягом якого прийнята ймовірність заняття всіх каналів одночасно, тобто ймовірність події, що призводить до втрати виклику. Таким чином, можливість втрати одного виклику в системі зв'язку з  $V$  каналами в періоди пікового навантаження визначає значення показника якості обробки трафіку. Зрозуміло, що чим більше  $V$ , тим менше показник якості обробки трафіку і тим вище якість обробки трафіку.

Прогнозування параметрів трафіку сучасних телекомунікаційних систем має велике значення для їхнього розвитку та надання якісних послуг абонентам. На основі подібних прогнозів відбувається виділення ресурсів системи: провідних та радіоканалів, смуги частот, проводиться автоматична реконфігурація обладнання у разі збоїв або при впровадженні нової апаратури зі зв'язку.

При короткостроковому прогнозуванні важливим чинником є час, який витрачається обробку даних: що менше часу і обчислювальних потужностей потрібно для роботи моделі, тим краще. З цієї позиції відносно прості моделі, наприклад, лінійна регресія, метод ковзного середнього або наївна модель, більш кращі в порівнянні зі складними моделями (ARIMA, N-BEATS), якщо забезпечують прийнятну якість прогнозу.

Порівняльний аналіз різних моделей навчання дозволяє зробити такі висновки: для протоколу UDP для часового ряду, що досліджується, найбільш підходящою моделлю стала модель ARIMA, для протоколу TCP – лінійна

регресія і модель Theta, для протоколу HTTPS – лінійна регресія, ARIMA і N-BEATS. Якість навчання моделей та якість прогнозування сильно залежить від вихідного часового ряду, якості та повноти зібраних даних. Найчастіше відносно прості методи, такі як наївна сезонна модель або лінійна регресія, дають прийнятну якість короткострокового прогнозу, маючи при цьому значно більшу швидкодію, ніж складніші та точніші моделі.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Цвіркун Л.І. Розробка програмного забезпечення комп'ютерних систем. Програмування: навч. посіб. [Електронний ресурс] / Л.І. Цвіркун, А.А. Євстігнєєва, Я.В. Панферова ; під заг. ред. проф. Л.І. Цвіркуна ; М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». – 568 с.
2. Цвіркун, Л.І. Розробка програмного забезпечення комп'ютерних систем. Програмування: навч. посібник / Л.І. Цвіркун, А.А. Євстігнєєва, Я.В. Панферова, під заг. ред. Л.І. Цвіркуна. – 3-є вид., випр. – Д.: Національний гірничий університет, 2016. – 223 с. – ISBN 978-966-350-595-4.
3. Цвіркун Л.І. Глобальні комп'ютерні мережі. Програмування мовою РНР: навч. посібник / Л.І. Цвіркун, Р.В. Липовий, під заг. ред. Л.І. Цвіркуна. – Д.: Національний гірничий університет, 2013. – 239 с. – ISBN 978-966-350-417-9.
4. Computer Emergency Response Team of Ukraine [Електронний ресурс]. - Режим доступу: [http://cert.gov.ua/?page\\_id=207](http://cert.gov.ua/?page_id=207). (24.11.24)
5. Hassan, Mohamed&SyedAriffin, SharifahHafizah&SyedYusof, Sharifah & N.Effiyana, Ghazali & Kanona, Mohammed & Rava, Mohammad. (2022). Smoothing-aided long-short term memory neural network-based LTE network traffic forecasting. International Journal of Electrical and Computer Engineering (IJECE). 12. 6859. 10.11591/ijece.v12i6.pp6859-6868.
6. Vinogradov M., IvanchenkoYe., Gnatyuk V. Method for efficiency assessment of cyber incidents processing by CSIRT // Ukrainian Scientific Journal of Information Security, 2017, vol. 23, issue 1, p. 56-62.
7. Gizun A. Approaches to Improve the Activity of Computer Incident Response Teams / A. Gizun, V. Gnatyuk, N. Balyk, P. Falat // Proceedings of the 2015 IEEE 8th International Conference on “Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications” (IDAACS'2015), Warsaw, Poland, September 24-26, 2015: Vol. 1. – Pp. 442-447.
8. Chan, Kit Yan & Dillon, Tharam & Singh, Jaipal & Chang, Elizabeth. (2012). Neural-Network-Based Models for Short-Term Traffic Flow Forecasting Using a Hybrid Exponential Smoothing and Levenberg–Marquardt Algorithm. IEEE

Transactions on Intelligent Transportation Systems - TITS. 13. 644-654. 10.1109/TITS.2011.2174051.

9. Gizun A.I. Baseparameters of forecasting and identification of computer attacks in information and communication systems / A.I. Gizun, S.I. Topcheev, M.O. Ryabyy // Proceedings the sixth world congress «Aviation in the XXI-stcentury». «Safety in Aviation and Space Technologies». – Vol. 1. – K.: NAU, 2014. – P. 1.11.40-1.11.44.

10. Barba, Lida & Rodriguez, Nivaldo & Montt, Cecilia. (2014). Smoothing Strategies Combined with ARIMA and Neural Networks to Improve the Forecasting of Traffic Accidents. The Scientific World Journal. 2014. 152375. 10.1155/2014/152375.

11. Cui, Zhengyan & Zhang, Junjun & Noh, Giseop & Park, Hyun. (2023). ADSTGCN: A Dynamic Adaptive Deeper Spatio-Temporal Graph Convolutional Network for Multi-Step Traffic Forecasting. Sensors. 23. 6950. 10.3390/s23156950.

12. TrầnQuang, Thanh&Hao, Li&Trinh, Quang. (2019). Cellular network traffic prediction using exponential smoothing methods. Journal of Information and Communication Technology. 18. 1-18. 10.32890/jict2019.18.1.1.

13. Валецька Т. М. Комп'ютерні мережі. Апаратні засоби. – К.: Центр навчальної літератури, 2004. 208 с.

14. Зайченко Ю.П. Комп'ютерні мережі: Навчальний посібник. – К.: Слово, 2003. – 286 с.

15. Кожем'яко В. П., Тимченко Л. І., Яровий А. А. Паралельно-ієрархічні мережі як структурно-функціональний базис для побудови спеціалізованих моделей образного комп'ютера. – Вінниця: Універсум, 2005. – 162 с.

16. Andrie, Rosa & Noprianto, Noprianto & Ilmu, Muhammad & Arai, Kohei. (2022). Traffic Density Prediction using IoT-based Double Exponential Smoothing. Knowledge Engineering and Data Science. 5. 168. 10.17977/um018v5i22022p168-178.

17. Нові комп'ютерні засоби, обчислювальні машини та мережі: Збірник наукових праць. Т.1. – К.: НАНУ, Ін-т кіберн. ім. В.М. Глушкова, Наук. рада НАН з проблеми «Кібернетика», 2001. – 170 с.
18. Saha, Sajal & Sayed, Moinul & Haque, Anwar. (2023). Empirical Mode Decomposition and Stationary Wavelet Transformation in Internet Traffic Prediction. 1-6. 10.1109/INFOCOMWKSHP57453.2023.10226162.
19. Фіногєєв А.Г. Мережеві технології. Частина 1: Навчальний посібник. – Харків: Вид-во ПГУ, 2013. – 107 с.
20. Юринець В. Є., Юринець Р. В. Комп'ютерні мережі. Інтернет. – Львів: Видавничий центр ЛНУ ім. І. Франка, 2006. – 322 с.
21. Ni, Qingjian&Peng, Wenqiang&Zhu, Yuntian&Ye, Ruotian. (2023). Graph drop out self-learning hierarchical graph convolutionnet work for traffic prediction. Engineering Applications of ArtificialIntelligence. 123. 106460. 10.1016/j.engappai.2023.106460.
22. Liu, Zibo&Shojaee, Parshin&Reddy, Chandan. (2023). Graph-based Multi-ODE Neural Networks for Spatio-Temporal Traffic Forecasting. 10.48550/arxiv.2305.18687.
23. Aldhayni, TheyaznTheyazn&Joshi, Manish. (2018). Clustering to Enhance Network Traffic Forecasting. 10.1007/978-981-10-3932-4\_37.
24. Harrou, Fouzi&Zeroual, Abdelhafid&Hittawe, Mohamad&Sun, Ying. (2022). Recurrent and convolutional neural networks for traffic management. 10.1016/B978-0-12-823432-7.00011-2.
25. Millán, Ginno & Lefranc, Gaston & Osorio, R. & Victor, Lomas-Barrie. (2021). Time Series Analysis of Computer Network Traffic in a Dedicated Link Aggregation. 10.36227/techrxiv.14575941.
26. Li, Ming & Wang, Anqi. (2020). Fractal teletraffic delay bounds in computer networks. Physica A: Statistical Mechanics and its Applications. 557. 124903. 10.1016/j.physa.2020.124903.
27. Al Azzeh, Jamil & Hadidi, Mazin & Odarchenko, Roman & Gnatyuk, Sergiy & Shevchuk, Z. & HU, ZHENGBING. (2017). Analysis of Self-Similar Traffic

Models in Computer Networks. International Review on Modelling and Simulations (IREMOS). 10. 328. 10.15866/iremos.v10i5.12009.

28. Millán, Ginno. (2021). On the Time Series Length for an Accurate Fractal Analysis in Network Systems. 10.48550/arXiv.2105.02920.

29. Sadiku, Matthew & Musa, Sarhan. (2013). Self-Similarity of Network Traffic. 10.1007/978-3-319-01646-7\_10.

30. Ding, Caichang & Chen, Yiqin & Liu, Zhiyuan & Alshehri, Ahmed & Liu, Tianyin. (2021). Fractal Characteristics of Network Traffic and Its Correlation with Network Security. Fractals. 30. 10.1142/S0218348X22400679.