Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Факультет *інфокомунікацій*
(повна назва)

Кафедра *інформаційно-мережної інженерії*
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

другий (магістерський)
(рівень вищої освіти)

(позначення документа)

Дослідження алгоритмів прогнозування інформаційного трафіку

(тема)

Виконав: студент 2 курсу, групи *ІМІм-21-1*
спеціальності *172 Телекомунікації та радіотехніка*
(код і повна назва спеціальності)
освітньо-професійна програма *Інформаційно-мережна інженерія*
(повна назва освітньої програми)
Тульчинський В.О.
(прізвище, ініціали)
Керівник *доц. Омельченко А.В.*
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ *Безрук В.М.*

(підпис) (прізвище, ініціали)

2022 р.

Не містить відомостей, заборонених до відкритого публікування

Студент _____ *Тульчинський В.О.*
(підпис) (прізвище та ініціали)

Керівник _____ *Омельченко А.В.*
(підпис) (прізвище та ініціали)

Харківський національний університет радіоелектроніки
(повна назва вищого навчального закладу)

Факультет *інфокомунікацій*

Кафедра *інформаційно-мережної інженерії*

Освітній рівень *другий (магістерський)*

Спеціальність *172 Телекомунікації та радіотехніка*
(код і назва)

Освітньо-професійна програма *Інформаційні мережі зв'язку*
(назва)

ЗАТВЕРДЖУЮ
Завідувач кафедри ІМІ
_____ *проф. Безрук В.М.*
"____" _____2022 року

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Тульчинському Віталію Олександровичу
(прізвище, ім'я, по батькові)

1. Тема роботи *Дослідження алгоритмів прогнозування інформаційного трафіку*
керівник роботи *Омельченко Анатолій Васильович, доц.* ,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
затверджені наказом ВНЗ від « 21 » *жовтня* 2022 року № *1376 Ст*

2. Строк подання студентом роботи *15 грудня 2022 р.*

3. Вихідні дані до роботи *1. З використанням реалізацій трафіку локальної мережі дослідити наступні алгоритми прогнозування: алгоритм LV; алгоритм простого експоненціального і подвійного експоненціального згладжування; алгоритми прогнозування на основі авторегресійної моделі. 2. Побудувати залежності нормованої середньоквадратичної похибки прогнозування фрактального трафіку від величини кроку прогнозування при різних значеннях параметру Херста*

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)
Вступ
1. Огляд характеристик мережного трафіку
2. Аналіз методів прогнозування трафіку
3. Дослідження точності алгоритмів прогнозування на один крок
4. Дослідження можливостей довготермінового прогнозування
Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
Слайди у форматі Power Point (назва, мета і актуальність атестаційної роботи, модель фрактального трафіку, застосування алгоритмів прогнозування трафіку, методи прогнозування, результати експериментального аналізу алгоритмів прогнозування, висновки)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів атестаційної роботи	Терміни виконання етапів роботи	Примітка
1	*Ознайомлення із завданням. Уточнення ТЗ.*	*24.10.22*	*Виконано*
2	*Підбір літератури за темою роботи.*	*25.10 – 27.10.22*	*Виконано*
3	*Виконання розділу 1*	*28.10 – 10.11.22*	*Виконано*
4	*Виконання розділу 2*	*11.11 – 20.11.22*	*Виконано*
5	*Виконання розділу 3*	*21.11 – 30.11.22*	*Виконано*
6	*Виконання розділу 4*	*01.12 – 10.12.22*	*Виконано*
7	*Оформлення пояснювальної записки*	*01.12 – 12.01.22*	*Виконано*
8	*Оформлення презентаційного матеріалу,*		
	підготовка до захисту у ЕК	*10.12 – 15.12.22*	*Виконано*

Дата видачі завдання *24 жовтня 2022 р.*

Студент _____ *Тульчинський В.О.*
(підпис) (прізвище та ініціали)

Керівник роботи _____ *Омельченко А.В.*
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка: 65 с., 19 рис., 36 джерел, 3 додатки.

Об'єкт дослідження – алгоритми прогнозування трафіку у інформаційних мережах.

Мета роботи – порівняльний аналіз алгоритмів прогнозування трафіку в інформаційних мережах.

Описана модель фрактального трафіку у вигляді дискретного у часі самоподібного процесу.

Розглянуті основні алгоритми прогнозування трафіку: 1) на основі згладжування часових рядів; 2) в рамках моделі ARIMA; 3) на основі штучної нейронної мережі.

На основі реалізації реального трафіку проведено порівняльний аналіз наступних алгоритмів прогнозування: LV-предиктору; прогнозування на основі моделей авторегресії; алгоритмів простого і подвійного експоненціального згладжування.

Виконано дослідження можливостей довготермінового прогнозування самоподібних у широкому розумінні послідовностей з різними значеннями параметру Херста.

АВТОРЕГРЕСІЙНИЙ ПРОЦЕС, ЗГЛАДЖУВАННЯ, МОДЕЛЬ ТРАФІКУ, САМОПОДІБНИЙ ПРОЦЕС, СЕРЕДНЬОКВАДРАТИЧНА ПОХИБКА, ФРАКТАЛЬНИЙ ТРАФІК

THE ABSTRACT

Explanatory note: 65 p., 19 fig., 36 sources, 3 app.

The object of research – traffic forecasting algorithms in information networks.

The purpose of the work is a comparative analysis of traffic forecasting algorithms in information networks.

The fractal model of traffic in the form of a time-discrete self-similar process is described.

The main traffic forecasting algorithms are considered: 1) based on time series smoothing; 2) within the ARIMA model; 3) based on an artificial neural network.

Based on the implementation of real traffic, a comparative analysis of the following forecasting algorithms was carried out: LV-predictor; forecasting based on autoregression models; simple and double exponential smoothing algorithms.

A study of the possibilities of long-term forecasting of self-similar sequences in the broad sense with different values of the Hurst parameter was conducted.

AUTOREGRESSIVE PROCESS, SMOOTHING, TRAFFIC MODEL, SELF SIMILAR PROCESS, MEAN SQUARE ERROR, FRACTAL TRAFFIC

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ 7
ВСТУП..... 8
1 ОГЛЯД ХАРАКТЕРИСТИК МЕРЕЖНОГО ТРАФІКУ 11
1.1 Застосування самоподібного процесу для опису мережного трафіку..... 11
1.2 Характеристик трафіку мереж з комутацією пакетів..... 14
1.2.1 Трафік Ethernet……14
1.2.2 Трафік глобальних мереж……...16
1.2.3 Відеотрафік зі змінною бітовою швидкістю……16
1.2.4 Трафік World Wide Web……..17
1.2.5 Система сигналізації SS7……17
2 Аналіз методів прогнозування трафіку 18
2.1 Застосування алгоритмів прогнозування телетрафіку..... 19
2.2 Прогнозування на основі моделі ARIMA 23
3.3 Прогноз за допомогою нейроних мереж 26
3 ДОСЛІДЖЕННЯ точності алгоритмів прогнозування на один крок 30
3.1 Показники точності прогнозування і експериментальні дані трафіку . 30
3.2 Аналіз результатів дослідження точності прогнозування..... 31
4 ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ДОВГОТЕРМІНОВОГО ПРОГНОЗУВАННЯ..... 35
4.1 Задача прогнозування самоподібних процесів з найменшою середньоквадратичною похибкою..... 35
4.2 Розрахунок точності найкращого лінійного прогнозу..... 36
ВИСНОВКИ 39
ПЕРЕЛІК ПОСИЛАНЬ 41
ДОДАТОК А. ЛІСТІНГ ПРОГРАМИ ДЛЯ ДОСЛІДЖЕННЯ АЛГОРИТМІВ ПРОГНОЗУВАННЯ ТРАФІКУ НА ОДИН КРОК..... 44
ДОДАТОК Б. ЛІСТІНГ ПРОГРАМИ ДЛЯ ДОСЛІДЖЕННЯ ПОТЕНЦІАЛЬНИХ МОЖЛИВОСТЕЙ ПРОГНОЗУВАННЯ ТРАФІКУ 50
ДОДАТОК В. СЛАЙДИ ПРЕЗЕНТАЦІЇ 55

ПЕРЕЛІК СКОРОЧЕНЬ

ANN – Artificial Neural Networks; штучна нейронна мережа;
AR – Auto-Regression; авторегресія;
ARMA – Auto-Regressive Moving Average, авторегресійне ковзне середнє;
ARIMA – Auto-Regressive Integrated Moving Average, авторегресійне проінтегроване ковзне середнє;
FARIMA – Fractional Auto-Regressive Integrated Moving Average, дрібне авторегресійне проінтегроване ковзне середнє;
IP – Internet Protocol;
ES – exponential smoothing; експоненціальне згладжування;
DES – double exponential smoothing; подвійне експоненціальне згладжування;
LRD – Long-Range Dependence; довготривала залежність;
NN – Neural Network, нейрона мережа;
QoS – Quality of Service, якість обслуговування;
RMSE – Root Mean Square Error;
TCP – Transmission Control Protocol
NMSE – Normalized Mean Square Error;
MA – Moving Average, ковзне середнє;
LV – Last Value; останнє значення;
UDP – User Datagram Protocol;
ПЗЗ – повільно згасаюча залежність;
ШЗЗ – швидко згасаюча залежність.

ВСТУП

Завдяки появі нових технологій за останнє десятиліття надзвичайно зріс обсяг трафіку Internet. Згідно з звітом Cisco, кількість користувачів Інтернету зросте з 3,9 у 2018 році до 5,3 мільярда у 2023 році [1]. Це явище спричинене зростанням кількості кінцевих пристроїв, таких як мобільні телефони, комп'ютери, а також появою нових мережевих парадигм, таких як Інтернет речей або 5G. Зростання кількості користувачів призводить до зростання мережевого трафіку.

Мережі мають справлятися із постійно зростаючим попитом на трафік і забезпечувати хорошу якість обслуговування для користувачів. У результаті ефективне використання мережевого обладнання стало вирішальним. Це завдання можна ефективно виконати, якщо мережевий трафік можна точно передбачити. Точне прогнозування трафіку застосовується в багатьох мережевих сферах, таких як енергозбереження [2], керування мережевими ресурсами [3] та бездротових сенсорних мережах [4].

Одним із застосувань передбачення трафіку є енергозбереження.

Прогнозування трафіку використовується в основних маршрутизаторах Інтернету для значного збереження енергії. Зі збільшенням вимог до трафіку та обчислювальних вимог кількість і складність процесорів, які використовуються в цих маршрутизаторах, зростають, що призводить до більшого енергоспоживання. Високе енергоспоживання обладнання та подальше зростання витрат на охолодження призводять до збільшення експлуатаційних витрат мережі. Якщо мережевий трафік можна точно передбачити, додаткові процесори в цих основних маршрутизаторах можна вимкнути під час низького трафіку для економії енергії [5–8].

По-друге, передбачення трафіку потрібне для ефективного використання ресурсів мережі. Точне прогнозування мережевого трафіку в точках доступу дозволяє ефективно розподіляти ресурси для забезпечення високої якості обслуговування [9, 10]. Програми також включають контроль перевантаження, контроль доступу, розподіл пропускної здатності мережі та виявлення аномалій. Завдяки новим програмам, таким як Youtube і Netflix, кількість відеотрафіку в мережі також зросла. Багато технік були запропоновані для прогнозування відеотрафіку [11–15]. Крім того, завдання виявлення та

запобігання зловживанням мережею стає дуже складним із зростанням обсягу трафіку та складності мереж.

Одним із засобів протидії зловживанню мережею є виявлення аномалій. Значне відхилення від нормальної поведінки трафіку може бути використано для виявлення атаки. Ефективність виявлення аномалій безпосередньо пов’язана з точністю прогнозування трафіку [16, 17].

Нарешті, у бездротових сенсорних мережах економія енергії є надзвичайно важливою, оскільки якість послуг, які вони надають, залежить від енергопостачання. Технологія мережевих мікросенсорів вважається однією з найбільш перспективних технологій 21 століття. Дешеві інтелектуальні пристрої з кількома вбудованими датчиками, об’єднані в мережу через бездротові з’єднання та Інтернет і розгорнуті у великій кількості, надають безпрецедентні можливості для вимірювання та контролю будинків, міст і навколишнього середовища. Крім того, мережеві мікросенсори забезпечують технологію для широкого спектру систем на оборонній арені, створюючи нові можливості для розвідки та спостереження, а також інших тактичних застосувань. Час роботи мережі можна значно збільшити, якщо вимкнути невикористовувані вузли. Для того, щоб вчасно включити ці вузли та оптимізувати енергозбереження, необхідне точне передбачення майбутніх потреб [4, 18, 19].

Статистичний аналіз вимірювань трафіку з широкого діапазону робочих пакетних мереж (наприклад, Ethernet LAN, WAN, CCSN/SS7, ISDN і VBR відео через ATM) переконливо показав наявність його фрактальних або самоподібних властивостей [21 – 23]. Цю властивість можна і слід використовувати для прогнозування трафіку.

На даний момент з’являється все більше доказів того, що LRD має фундаментальне значення для низки інженерних проблем телетрафіку, таких як вимірювання трафіку, поведінка в черзі та розмір буфера, контроль доступу та контроль заторів.

На відміну від традиційних моделей пакетного трафіку, які спричиняють експоненціальний хвіст розподілу розміру черги та зазвичай призводять до надто оптимістичних прогнозів продуктивності та неадекватного розподілу ресурсів, моделі фрактального трафіку передбачають гіперболічний або Вейбуллівський (розтягнутий експоненціальний) розподіл розміру черги та можуть таким чином впливають на контроль і управління пакетною мережею.

Це також вказує на те, що збільшення ємності каналу може бути більш ефективним для підвищення продуктивності, ніж збільшення розміру буфера.

Самоподібне явище являє собою процес, що демонструє структурні подібності в широкому діапазоні масштабів певного виміру. Іншими словами, еталонна структура повторюється в широкому діапазоні масштабів різних розмірів (геометричних, або статистичних, або часових), і статистика процесу не змінюється зі зміною. Таким чином, самоподібність може бути пов'язана з «фракталами», які являють собою об'єкти з незмінним зовнішнім виглядом у різних масштабах. Поняття фракталів включає, крім геометричного значення, статистику, а також динаміку. Це означає, що існують фрактальні процеси різних розмірів, наприклад, геометричні, статистичні та динамічні.

У випадку статистичних фракталів це щільність ймовірності, яка повторюється на кожному масштабі, як, наприклад, закон Парето. Самоподібність так же означає, що певна кореляційна структура зберігається в широкому діапазоні часових масштабів, хоча вона може мати різні форми. Для того, щоб мати можливість контролювати такі процеси, потрібно спочатку пояснити та перевірити на фізичних підставах причинно-наслідкові механізми, які можуть бути відповідальними за створення самоподібності в реалістичному мережевому середовищі. Крім того, розуміння впливу таких процесів є принаймні таким же важливим, як і розуміння їхнього фізичного походження.

Збільшення обсягу обмінюваних даних разом із новими вимогами до вищих пікових швидкостей передачі даних, підвищення надійності та зменшення затримки в мережах п'ятого покоління (5G) ставить перед дослідниками та операторами телекомунікацій нові виклики щодо підвищення продуктивності системи та, одночасно, підвищення енергоефективності мережі.

Більшість робіт розглядають прогнозування трафіку з використанням таких методів, як ARIMA. Крім того, різні версії моделі ARIMA (наприклад, SARIMA, Fractional-ARIMA) були застосовані, щоб охопити тенденції часової еволюції трафіку [24]. Зокрема, SARIMA (Seasonal ARIMA) — це техніка для вивчення сезонних і циклічних моделей, які присутні в трафіку, щоб знайти повторювану поведінку мобільних запитів користувача.

Метою даної роботи є дослідження ефективності алгоритмів прогнозування трафіку в локальних мережах.

1 ОГЛЯД ХАРАКТЕРИСТИК МЕРЕЖНОГО ТРАФІКУ

1.1 Застосування самоподібного процесу для опису фрактального трафіку

Самоподібність трафіку у комп'ютерних мережах вперше була описана в класичних статтях [21, 24]. Спочатку це явище спостерігалося у трафіку локальної мережі, а потім поширилося на трафік WWW, трафік VBR [21, 22]. Це дуже відрізнялося від попередніх поглядів, які розглядали мережевий трафік як узагальнення телефонного трафіку.

У випадку самоподібності скупченість присутня в усіх часових масштабах і не має тенденції до чистого білого шуму при агрегуванні, як у пуасонівському процесі. На інтуїтивному рівні явище самоподібності означає, що властивості або характеристики об'єкта зберігаються незалежно від масштабування в часі чи просторі. У комп’ютерних мережах нас цікавить статистична самоподібність, тобто поведінка функції автокореляції в різних масштабах часу [26].

Введемо означення самоподібних процесів, для чого використаємо наступні позначення [26]. Нехай $X = (X_1, X_2, ...)$ – стаціонарний у широкому розумінні дійсний стаціонарний випадковий процес з дискретним часом. Позначимо через $\mu = M[X_t] < \infty$ та $\sigma^2 = D[X_t] < \infty$ відповідно середнє й дисперсію процесу X, а через

$$b(k) = M[(X_{t+k} - \mu)(X_t - \mu)]; \quad r(k) = \frac{b(k)}{\sigma^2}, \qquad k \in Z_+ \quad (1.1)$$

– відповідно автокореляцію і коефіцієнт кореляції процесу X.

Нехай з процесу X отримано усереднений процес [26]

$$X_t^{(m)} = \frac{1}{m} \cdot (X_{t \cdot m - m + 1} + ... + X_{t \cdot m}), \qquad t, m \in \mathbf{N}. \quad (1.2)$$

Нехай також $r_m(k)$ і V_m – відповідно коефіцієнт кореляції й дисперсія усередненого процесу $X_t^{(m)}$.

Процес X називається самоподібним у вузькому розумінні (СВР) з параметром $H = 1 - (\beta / 2)$, $0{,}5 < H < 1$, якщо [26]

$$m^{1-H} X^{(m)} \sim X, \qquad m \in N, \tag{1.3}$$

де символ ~ позначає рівність розподілів.

Крім СВР широко використовується поняття самоподібності в рамках кореляційної теорії. Сформулюємо його.

Процес X називається строго самоподібним у широкому розумінні (ССШР) з параметром $H = 1 - (\beta / 2)$, $0 < \beta < 1$, якщо його коефіцієнт кореляції [26]

$$r(k) = g(k), \qquad k \in N, \tag{1.4}$$

де функція

$$g(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}]. \tag{1.5}$$

Причина, за якою процес X називається самоподібнім, полягає у тому, що означення (1.5) приводить до умови $r_m(k) = r(k)$, яка означає, що процес не змінює свій коефіцієнт кореляції після усереднення по блокам довжини m.

Можна показати, що функція $g(k)$ асимптотично зі зростанням k спадає за степеневим законом [26]

$$g(k) \sim \frac{1}{2}(2-k)(1-\beta)k^{-\beta}, \quad k \to \infty. \tag{1.6}$$

Зі співвідношення (1.6) випливає, що кореляційна функція самоподібних процесів спадає дуже повільно (за степеневим законом) у порівнянні, наприклад, з показниковим спаданням, властивим марківським і авторегресійним процесам.

На рис 1.1 і рис 1.2 зображений вид функції g(k) при різних значеннях параметра Херста. Ці графіки задають вид кореляційної функції ССШР процесів і побудовані згідно з виразом (1.5).

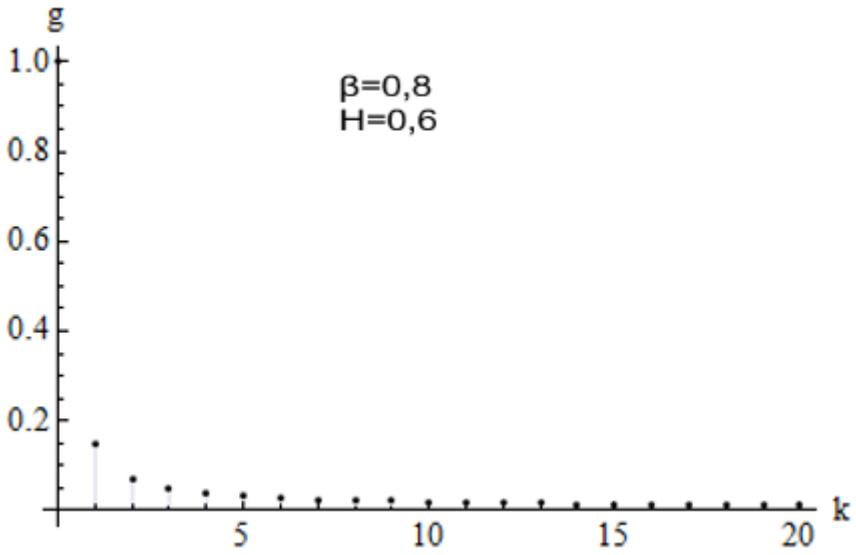

Рисунок 1.1 – Вид функції g(k) при β=0,8 (H=0,6)

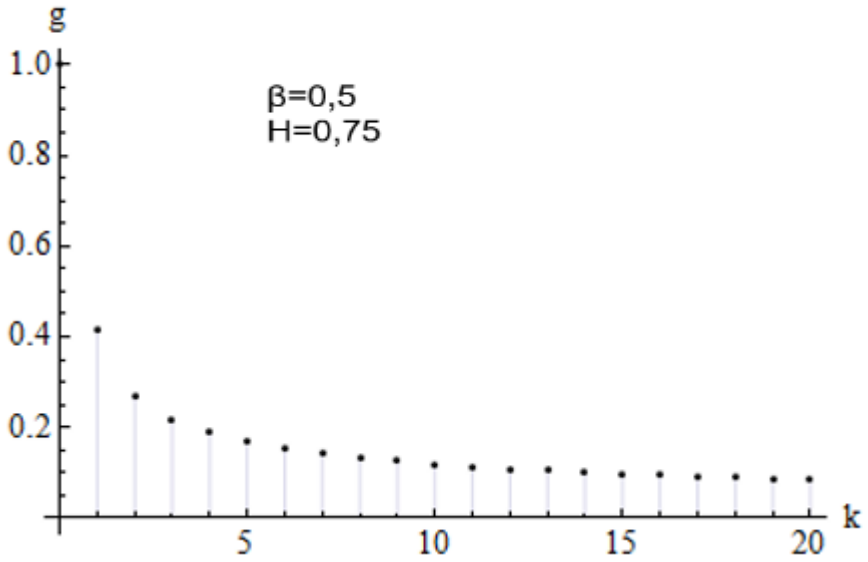

Рисунок 1.2 – Вид функції g(k) при β=0,5 (H=0,75)

Процес X називається асимптотично самоподібним у широкому розумінні (АСШР) з параметром $H = 1 - (\beta / 2)$, $0 < \beta < 1$, якщо [26]

$$\lim_{m \to \infty} r_m(k) = g(k), \qquad k \in \mathbf{N}. \tag{1.7}$$

Зміст цього визначення полягає в тому, що АСШР-процес після усереднення по блоках довжини m сходиться при $m \to \infty$ до ССШР процесу.

На рис. 1.3 проілюстрована швидкість спадання автокореляційної функції ССШР процесів при різних значеннях параметра H.

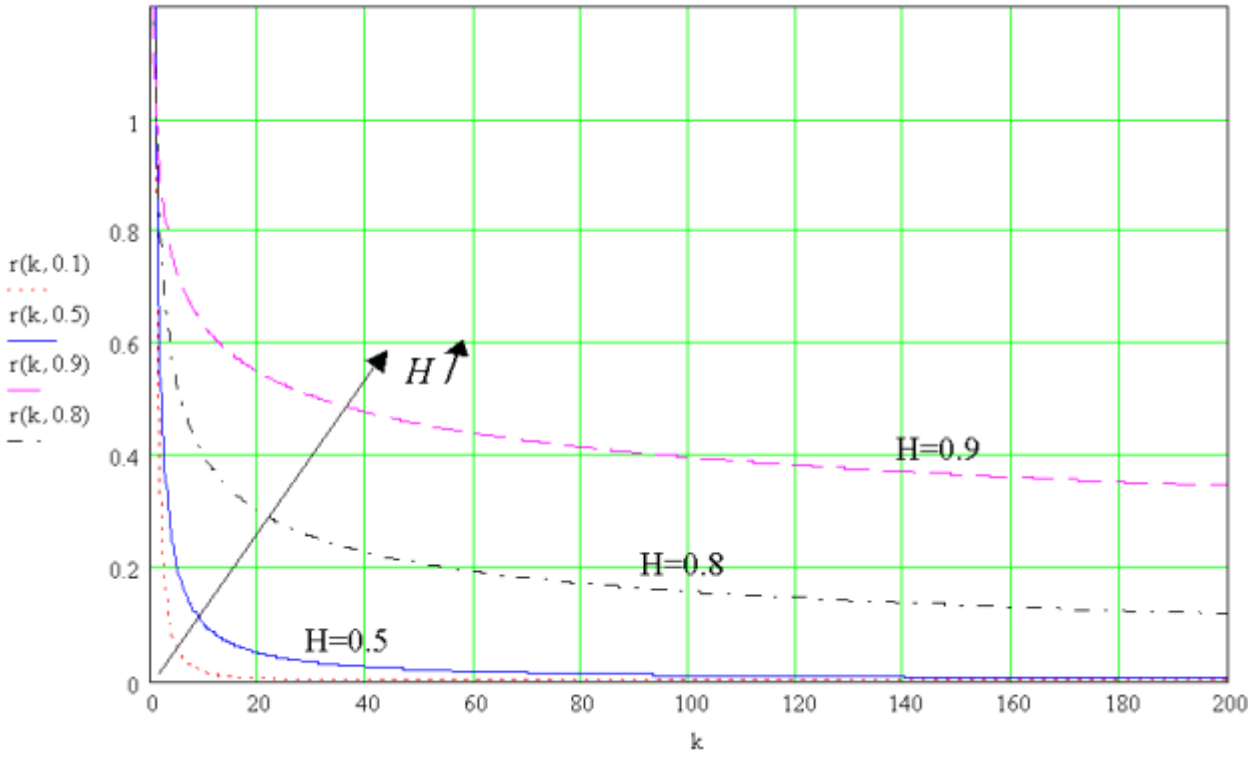

Рисунок 1.3 – Вид кореляційної функції ССШР процесів при різних H

1.2 Характеристик трафіку мереж з комутацією пакетів

Статистичний аналіз вимірювань трафіку високої роздільної здатності з широкого діапазону робочих пакетних мереж (наприклад, мережі Ether net LAN, WAN, CCSN/SS7, ISDN і відео VBR через ATM) переконливо показав наявність самоподібних властивостей як у локальній зоні, так і в трасах руху на великій території [20]. Наведемо короткий опис моделей, які придатні для ряду конкретних застосувань.

1.2.1 Трафік Ethernet

Використовуючи величезну кількість даних вимірювань трафіку Ethernet, зібраних між 1989 і 1992 роками з різних локальних мереж Ethernet (LAN), автори згаданих вище документів (W. Leland, M. Taqqu, W. Willinger і D.

Wilson) продемонстрували, що трафік Ethernet є самоподібним. Крім того, було проведено детальний агрегаційний аналіз, R/S аналіз, дисперсійно-часовий графік, періодограми, оцінки спектра тощо [21, 24].

На рис. 1.4 у якості прикладу наведено графіки коефіцієнта кореляції для досліджуваного трафіку в логарифмічному масштабі. Для порівняння на тому ж графіку наведена теоретична залежність, що плавно змінюються і відповідає значенню коефіцієнта кореляції ССШР процесу з коефіцієнтом Херста H=0,724.

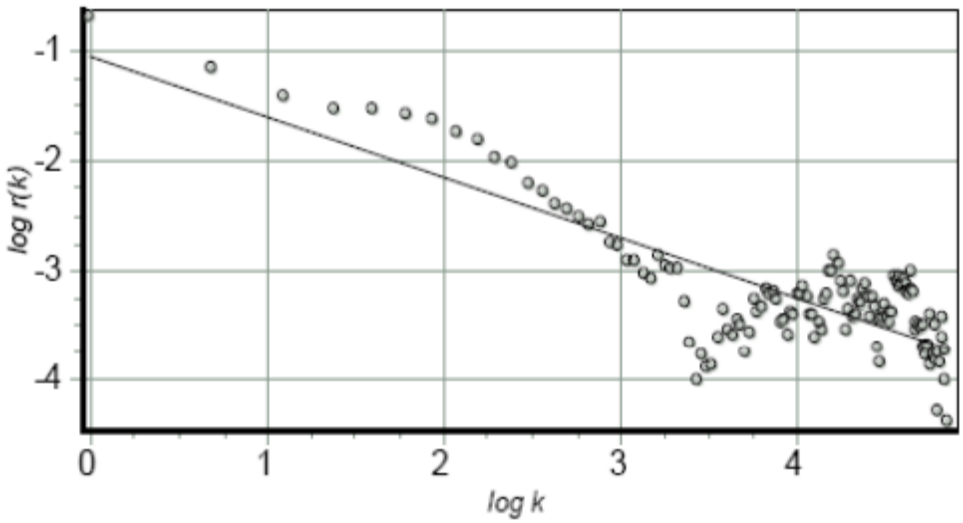

Рисунок 1.4 – Вид нормованої кореляційної функції у логарифмічному масштабі

Дослідники прийшли до наступних висновків відносно трафіку Ethernet [21, 24]:

- пуасонівська модель трафіку не відповідає дійсності;
- агрегований трафік локальних мереж є самоподібним;
- параметр Херста приймає значення 0,7<H<0,9;
- трафік погано усереднюється;
- закон великих чисел не виконується.

Згодом було опубліковано багато інших статей, які підтвердили самоподібну природу трафіку даних в Ethernet та інших мережних середовищах.

1.2.2 Трафік глобальних мереж

У 1994 році двоє дослідників з лабораторії Лоуренса Берклі і з Каліфорнійського університету опублікували статтю [23]. У ній автори повідомили про результати вимірювання та аналізу трафіку WAN (трафік TCP) для програм, таких як FTP і TELNET. Мільйони TCP-пакетів і з'єднань були зібрані з різних сайтів із відстеженням від 1 години до 30 днів.

Основі результати цих досліджень наступні:

- пуасонівська модель серйозно недооцінює стрибкоподібність TCP-трафіку в широкому діапазоні часових масштабів;
- самоподібність присутня агрегованому трафіку WAN;
- прибуття з'єднання TELNET описується законом Пуасона, але прибуття пакетів – ні;
- FTP з'єднання не виявляються пуассонівськими. Надходження сеансу FTP добре відповідають законом Пуасона , але надходження з'єднання даних мають набагато більшу швидкість надходження. Крім того, кількість байтів на пакет має важкий розподіл.

Подібно до досліджень Ethernet, документи про трафік WAN супроводжувалися багатьма іншими дослідженнями вимірювання та аналізу, які підтвердили самоподібну природу трафіку WAN.

1.2.3 Відеотрафік зі змінною бітовою швидкістю

Було опубліковано декілька досліджень, які показали самоподібну природу відеотрафіку, який передається через банкомати та Інтернет. Найважливішими публікаціями є публікації [20, 21].

Деякі з основних повідомлених результатів:

- передача відео демонструє самоподібність;
- довжина кадру відповідає розподілу Парето, принаймні в хвості розподілу;
- LRD, здається, є невід'ємною особливістю відеотрафіку VBR, незалежно від використовуваних кодеків і записаної сцени.

1.2.4 Трафік World Wide Web

Група дослідників із Бостонського університету у 1995 та 1996 роках опублікувала низку робіт, у яких вони повідомляють про дослідження вимірювання трафіку WWW, головним результатом яких було те, що цей трафік є самоподібним [22]. Було зібрано та проаналізовано понад півмільйона слідів веб-запитів із приблизно 40 робочих станцій Бостонського університету.

Деякі з основних повідомлених результатів:

- трафік, створюваний браузерами, має самоподібний характер;
- робота кожного веб-браузер описується моделлю ON-OFF, і дані добре відповідають розподілу Парето;
- файли, доступні через Інтернет, мають жорсткий розподіл розміру (бімодальний розподіл).

Зокрема, було помічено, що ступінь обмеженості розмірів файлів може безпосередньо визначати ступінь самоподібності трафіку, і на це істотно не впливають різні параметри, такі як варіації пропускної здатності вузького місця та/або буфера са швидкість, перехресний рух і розподіл часу між прибуттями. Розподіл розмірів HTTP-документів можна змоделювати сумішшю логнормальних розподілів і розподілів Парето.

1.2.5 Система сигналізації SS7

Вимірювання та аналіз трафіку SS7 проводилися в Дублінському інституті передових досліджень [27]. Автори зібрали та проаналізували понад 170 мільйонів сигнальних повідомлень із різноманітних мереж, контрольованих SS7.

Основні результати наступні:

- модель Пуассона не враховує властивості трафіку SS7;
- трафік SS7 є самоподібним.

2 АНАЛІЗ МЕТОДІВ ПРОГНОЗУВАННЯ ТРАФІКУ

2.1 Застосування алгоритмів прогнозування телетрафіку

Методи прогнозування трафіку у інформаційних мережах давно цікавлять операторів, які зацікавлені у зниженні вартості мереж. Важливим інструментом у досягненні цієї мети є прогнозування. Прогнозування дозволяє мережним операторам розрахувати потенційну вартість нової мережі/послуги для даної GoS на етапі планування та проектування, таким чином гарантуючи, що витрати зведені до мінімуму. Окрім того алгоритми прогнозування трафіку можуть використовуватися для управління мережами з метою недопущення заторів і забезпечення їх оптимальної роботи.

З розвитком інформаційних мереж з'явилися інші застосування алгоритмів прогнозування. Суттєвою проблемою при обслуговуванні телекомунікаційного трафіку у мультисервісних мережах. Трафік у цих мережах значною мірою є фрактальним, тобто має значну кореляцію на великих інтервалах часу. Ця властивість дозволяє покращити коефіцієнт використання каналу та загальної ефективності системи.

У роботах [28, 29] описаний пристрій з динамічним управлінням швидкістю вихідного потоку. Його схема представлена на рис. 2.1. Для управління потоком на її виході застосовується передбачення швидкості трафіку на вході. Такий механізм управління трафіком дозволяє ефективно його вирівняти і в той же час не допустити перевантаження буферу.

Особливо ефективною є робота подібного пристрою у таких схемах, де в одному фізичному каналі організовано декілька логічних каналів.

В літературі [28, 29] також описується застосування алгоритму прогнозування для управління інтенсивністю трафіку з використанням механізму Token Bucket. Це дозволяє підвищити ефективність використання мережних ресурсів.

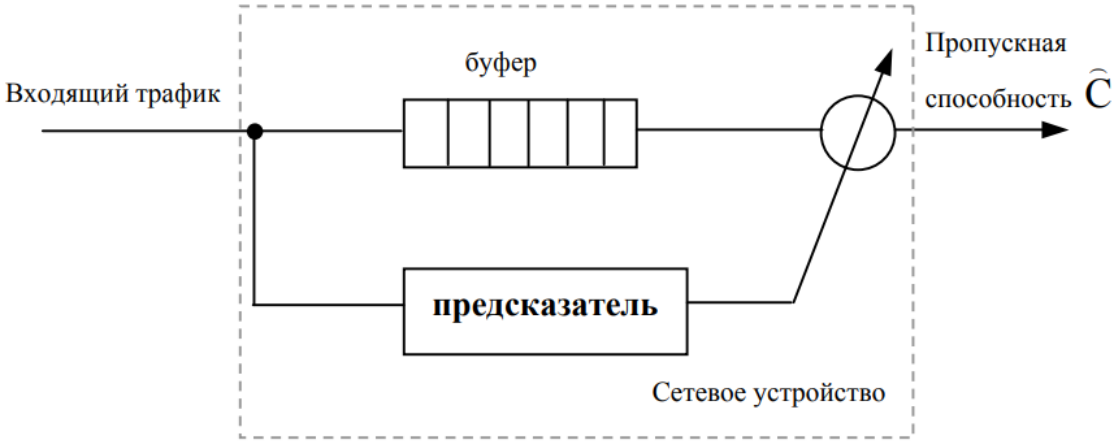

Рисунок 2.1 − Структура пристрою з динамічним управлінням швидкістю вихідних потоків на основі прогнозування

2.2 Прогнозування шляхом згладжування спостережень

Розглянемо спочатку алгоритми на основі згладжування часових рядів спостережень.

У якості найбільш простого прогнозу часто використовується метод LV. Такий прогноз називається ще наївним. Для нього всі прогнози визначаються як значення останнього спостереження. Цей метод добре працює у багатьох прикладних задачах [30].

Трохи складнішим є метод середнього значення. Тут прогнози всіх майбутніх значень дорівнюють середньому значенню вибірки даних. Подальшим його розвитком є метод WMA (віконної ковзної середньої). Тут прогнозне значення формується як вибіркове середнє n минулих спостережень [30].

Ще складнішими є методи експоненціального згладжування [31]. Вони включають просте (одинарне) експоненціальне згладжування та подвійне експоненціальне згладжування. Метод простого експоненціального згладжування використовується, якщо дані часових рядів не містять трендів і сезонних (циклічних) моделей. Навпаки, метод подвійного експоненціального згладжування використовується, якщо дані часових рядів містять моделі трендів і не містять сезонних моделей.

Шаблони даних можна розділити на чотири типи, а саме трендові шаблони, сезонні шаблони, циклічні шаблони та нерегулярні шаблони [31]. Стаціонарний і нестаціонарний типи даних не містять сезонних моделей. У цьому випадку прогнозування може бути зроблено за допомогою методу ковзного середнього та методів одинарного та багаторазового експоненціального згладжування. Якщо дані містять сезонні моделі, метод дасть низькі прогнози. Тому, щоб мінімізувати помилки в результатах прогнозування, необхідно визначити відповідний метод сезонних моделей даних. Метод згладжування Холта-Вінтерса є одним із відповідних методів прогнозування даних, що містять сезонні моделі.

У загальному випадку модель експоненціального згладжування описується виразом [31]

$$x_t = \mu_t + S_{t,p} + \beta_t t + a_t, \tag{2.1}$$

де μ_t – представляє змінний у часі середній рівень;

β_t – нахил «тренду»;

$S_{t,p}$ – означає змінну у часі циклічну (сезонну) складову с параметром сезонності p;

a_t – помилка спостереження, що вважається білим шумом;

параметр p є натуральним числом і відповідає номеру циклу.

Модель простого експоненціального згладжування слід використовувати, коли дані часових рядів не мають тренду та сезонності. Проста модель експоненціального згладжування визначається рівнянням моделі

$$x_t = \mu_t + a_t. \tag{2.2}$$

Рівняння згладжування має вигляд

$$\hat{X}_{t+1} = \alpha \cdot x_t + (1-\alpha) \cdot \hat{X}_t, \tag{2.3}$$

де $\alpha \in (0,1)$.

Параметри α в прогнозі (2.3) зазвичай вибирається з умови найменшої середньоквадратичної похибки прогнозування.

Нескладно показати, що метод експоненціального згладжування співпадає з оптимальним прогнозом на основі моделі ARIMA(0,1,1) [32].

Подвійне експоненціальне згладжування (Хольта-Вінтерса) є розвитком простого експоненціального згладжування [31]. Його слід використовувати, коли дані часових рядів мають тренд, але не мають сезонності. Модель подвійного експоненціального згладжування задана рівнянням моделі

$$x_t = \mu_t + \beta_t t + a_t. \tag{2.4}$$

У даному випадку рівняння згладжування на d кроків приймає вигляд

$$\hat{X}_{t+d} = A_t + d \cdot B_t, \tag{2.5}$$

де змінні A_t і B_t є відповідно оцінками величини рівня і тренду часового ряду в момент часу t. Вони знаходяться наступним чином:

$$A_t = \alpha_1 \cdot y_t + (1-\alpha_1) \cdot (A_{t-1} - B_{t-1}), \tag{2.6}$$

$$B_t = \alpha_2 \cdot (A_t - A_{t-1}) + (1-\alpha_2) \cdot B_{t-1}. \tag{2.7}$$

Згідно (2.5) прогноз на d-кроків вперед, формується як сума останнього оціненого рівня A_t та добутку кроку прогнозування d на оцінку тренду (нахилу) B_t. Зазвичай $A_1 = X_1$, а $B_1 = 0$.

Параметри α_1 і α_2 в прогнозі (2.5, 2.6) слід вибирати з умови найменшої середньоквадратичної похибки прогнозування.

Зазначимо, що при $\alpha_2 = 0$ алгоритм Хольта співпадає з алгоритмом експоненціального згладжування. Якщо ж $\alpha_1 = \alpha_2$ то алгоритм Хольта співпадає з відомим алгоритмом Брауна.

Модель сезонного експоненціального згладжування слід використовувати, коли дані часових рядів не мають тренду, але мають сезонність. Ця модель описується рівнянням

$$x_t = \mu_t + S_{t,p} + a_t. \tag{2.8}$$

Прогноз при сезонному експоненціальному згладжуванні здійснюється згідно з рівнянням

$$\hat{X}_{t+d} = A_t + C_{t+(d \bmod L)-L}, \tag{2.9}$$

де допоміжні змінні

$$A_t = \alpha \cdot (X_t - C_{t-L}) + (1-\alpha) \cdot A_{t-1},$$
$$C_t = \delta \cdot (y_t - A_t) + (1-\delta) \cdot C_{t-L}.$$

Тобто спочатку здійснюється прогноз на d-кроків вперед, використовуючи останній доступний оцінений стан рівня, а потім додаєте останній доступний згладжений сезонний фактор $C_{t+(d \bmod L)-L}$, який відповідає поточному часу з врахуванням циклу прогнозування.

Експоненціальне згладжування з врахуванням тренду і циклічності враховує повний варіант моделі (2.1). Цей варіант згладжування має назву методу Тейла-Вейджа. Прогноз тут здійснюється у відповідності з рівнянням [31]:

$$\hat{X}_{t+d} = A_t + d \cdot B_t + C_{t+(d \bmod L)-L}, \tag{2.10}$$

де допоміжні змінні знаходяться наступним чином:

$$A_t = \alpha_1 \cdot (X_t - C_{t-L}) + (1-\alpha_1) \cdot (A_{t-1} - B_{t-1}),$$

$$B_t = \alpha_2 \cdot (A_t - A_{t-1}) + (1-\alpha_2) \cdot B_{t-1},$$

$$C_t = \alpha_3 \cdot (y_t - A_t) + (1-\alpha_3) \cdot C_{t-L}.$$

Зазвичай приймаються наступні початкові значення допоміжних змінних $A_1 = X_1$, $B_1 = 0$, а у першому циклі нехтують циклічністю.

Як і раніше параметри $\alpha_1, \alpha_2, \alpha_3 \in (0,1)$ вибираються з умови мінімальної помилки прогнозування.

2.2 Прогнозування на основі моделі ARIMA

Клас процесів ARIMA широко використовується для прогнозування часових рядів. Модель ARIMA можна розглядати як «фільтр», який намагається відокремити сигнал від шуму, а потім сигнал екстраполюється в майбутнє для отримання прогнозів. Прогнозне рівняння ARIMA для стаціонарного часового ряду є лінійним (тобто регресійного типу) рівнянням, у якому предиктори складаються з лагів залежної змінної та/або лагів помилок прогнозу. У відповідності з моделлю ARIMA прогнозоване значення формується як константа зважена сума одного чи кількох останніх значень та зважена сума одного чи кількох останніх значень помилок прогнозування [32].

Акронім ARIMA розшифровується як Auto-Regressive Integrated Moving Average [32]. Лаги стаціонарного ряду в рівнянні прогнозування називаються «авторегресійними» членами, лаги помилок прогнозу називаються «ковзними середніми» членами, а часовий ряд, який потрібно відрізнити, щоб стати стаціонарним, називається «інтегрованим». варіант стаціонарної серії. Моделі випадкового блукання та випадкового тренду, авторегресійні моделі та моделі експоненціального згладжування є окремими випадками моделей ARIMA. Несезонна модель ARIMA класифікується як модель "ARIMA(p,d,q)", де: p – кількість членів авторегресії, d — кількість несезонних відмінностей, необхідних для стаціонарності, і q – кількість помилок прогнозу із запізненням у рівнянні прогнозу.

За означенням процес ARIMA описується рівнянням [32]

$$\Delta^d X_t = c + \sum_{i=1}^{p} \varphi_i \Delta^d X_{t-i} + \sum_{j=1}^{p} b_j \varepsilon_{t-j} + \varepsilon_t, \qquad (2.11)$$

де ε_t – стаціонарний часовий ряд;

c, φ_i, b_j – параметри моделі;

Δ^d – оператор кінцевої різниці часового ряду порядку d.

Пояснимо суть операції кінцевої різниці. Нехай $y_t = \Delta^d X_t$. Тоді при d=0 маємо $y_t = Y_t$; при d=1: $y_t = Y_t - Y_{t-1}$; при d=2: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$ і так далі. Слід звернути увагу, що друга різниця це перша

різниця першої різниці, яка є дискретним аналогом другої похідної, тобто локальне прискорення ряду.

Загальне рівняння прогнозування виглядає так:

$$\hat{X}_t = \mu + \varphi_1 X_{t-1} + ... + \varphi_p X_{t-p} - \theta_1 \varepsilon_{t-1} - ... - \theta_q \varepsilon_{t-q}. \qquad (2.12)$$

Тут параметри ковзного середнього (θ) визначені таким чином, що їхні знаки є від'ємними в рівнянні, дотримуючись угоди, введеної Боксом і Дженкінсом.

Процес ARIMA(p,0,0) називається авторегресійним процесом (AR) порядку p. Для AR-процесу прогнозне рівняння має наступний вигляд

$$\hat{X}_{t+1} = \mu + \sum_{i=1}^{p} \varphi_i \cdot X_{t+1-i}, \qquad (2.13)$$

де φ_i, $i = \overline{1,M}$ – параметри (коефіцієнти) авторегресії.

Коефіцієнти авторегресії зазвичай оцінюються до прогнозування по навчальним вибіркам даних. Одні з найпростіших оцінок АР-параметрів отримуються методом Юла-Уолкера [12].

На рис. 2.2 зображена схема пристрою для прогнозування авторегресійних послідовностей згідно з рівнянням (2.13).

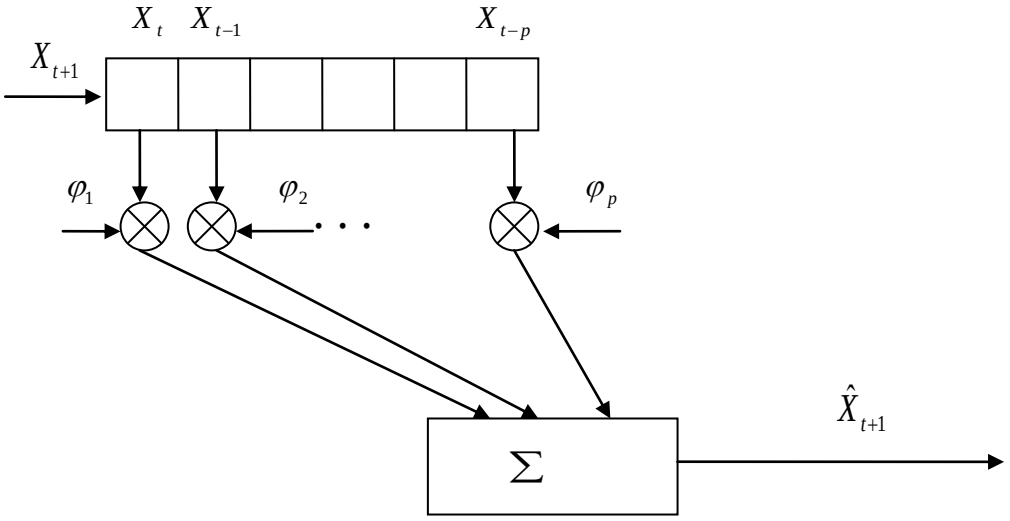

Рисунок 2.2 – Схема пристрою для прогнозування авторегресійних послідовностей

Процес ARIMA(0,0,q) має назву процесу ковзного середнього. Для нього прогнозне рівняння має такий вигляд

$$\hat{z}_{t+1} = \sum_{j=1}^{q} \theta_j a_{t+1-j} , \tag{2.14}$$

де відліки процесу $a_t = z_t - \hat{z}_t$.

Наведемо ще декілька цікавих прогнозних рівнянь, що витікають з моделі ARIMA.

Процес ARIMA(0,1,0) має назву «випадкове блукання». Цей процес є найпростішою моделлю нестаціонарного часового ряду. Його можна розглядати як граничний випадок моделі AR(1), в якій коефіцієнт авторегресії дорівнює 1. Прогнозне рівняння для цієї моделі можна записати так:

$$\hat{X}_t = \mu + \hat{X}_{t-1},$$

де постійний член є середньою зміною X від періоду до періоду (тобто довгостроковий дрейф).

Модель ARIMA(1,1,0) - диференційована авторегресійна модель першого порядку. Якщо помилки моделі випадкового блукання корельовані, проблему можна вирішити, додавши один лаг залежної змінної до рівняння передбачення, тобто за допомогою регресія першої різниці X на себе з відставанням на один період. Це дасть таке прогнозне рівняння

$$\hat{X}_t = \mu + \hat{X}_{t-1} + \varphi_1(\hat{X}_{t-1} - \hat{X}_{t-2}).$$

Модель ARIMA(0,1,1) без константи приводить до прогнозу у виді простого експоненціального згладжування. Слід враховувати, що для деяких нестаціонарних часових рядів (наприклад, тих, які демонструють шумні коливання навколо середнього, що повільно змінюється), модель випадкового блукання не працює так добре, як ковзне середнє минулих значень. Іншими словами, замість того, щоб брати останнє спостереження як прогноз наступного спостереження, краще використовувати середнє значення кількох останніх

спостережень, щоб відфільтрувати шум і точніше оцінити локальне середнє значення. Рівняння прогнозу для простої експоненціальної моделі згладжування можна записати в ряді математично еквівалентних форм, однією з яких є так звана форма «виправлення помилок», у якій попередній прогноз коригується в напрямку допущеної помилки:

$$\hat{X}_t = \hat{X}_{t-1} + \theta\varepsilon_{t-1}. \quad (2.15)$$

Оскільки $\varepsilon_{t-1} = X_{t-1} - \hat{X}_{t-1}$ за визначенням, то можна переписати так:

$$\hat{X}_t = \theta X_{t-1} + (1-\theta)\hat{X}_{t-1}.$$

Це означає, що прогноз зводиться до простого експоненціального згладжування.

Зазначимо, що прогнози на декілька кроків в рамках моделі ARIMA можуть бути побудовані з прогнозу на один крок заміною невідомих значень величин на прогнозні значення [32].

2.3 Прогноз за допомогою нейроних мереж

Нейрона мережа – це алгоритм, який намагається знайти приховані закономірності у великій кількості даних, відтворюючи те, як працює людський мозок. Такі мережі будуються з елементарних структур, які називаються нейронами.

Принцип роботи нейрону описується виразом

$$U = \sum_{i=1}^{m} w_i \cdot Z_i + b; \qquad Z = \varphi(U), \quad (2.16)$$

де $Z_1,...,Z_m$ – вхідні сигнали;

$w_1,...,w_m$ – вагові коефіцієнти;

b – деякий поріг;

$\varphi(U)$ – скалярна функція.

У якості φ(U), що називається функцією активації зазвичай використовується якась порогова функція: порогова функція із жорстким обмеженням; функція sgn, лінійна чи напівлінійна функція, або плавно

обмежена порогова функція див. рис. 2.3. У якості функції з плавним обмеженням часто використовується сигмоподібна функція типу [33]

$$\varphi(U) = \frac{1}{1 + exp(-U)}. \quad (2.17)$$

У деяких програмах використовується гіперболічний тангенс, що дає вихідні значення в діапазоні [-1,1].

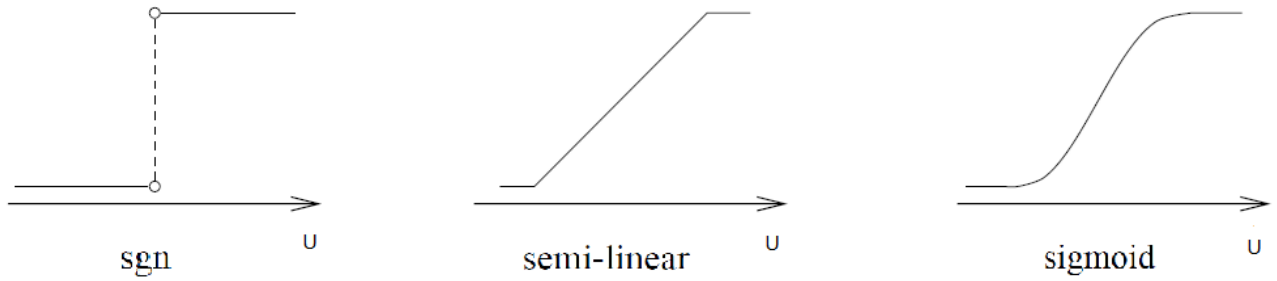

Рисунок 2.3 – Різні варіанти функції активації

Нейрону мережу можна розглядати як мережу «нейронів», які організовані в шари. Прогнози (або входи) [33].

Найпростіші мережі не містять прихованих шарів. На рис. 2.4 показана версія нейроної мережі лінійної регресії з чотирма предикторами.

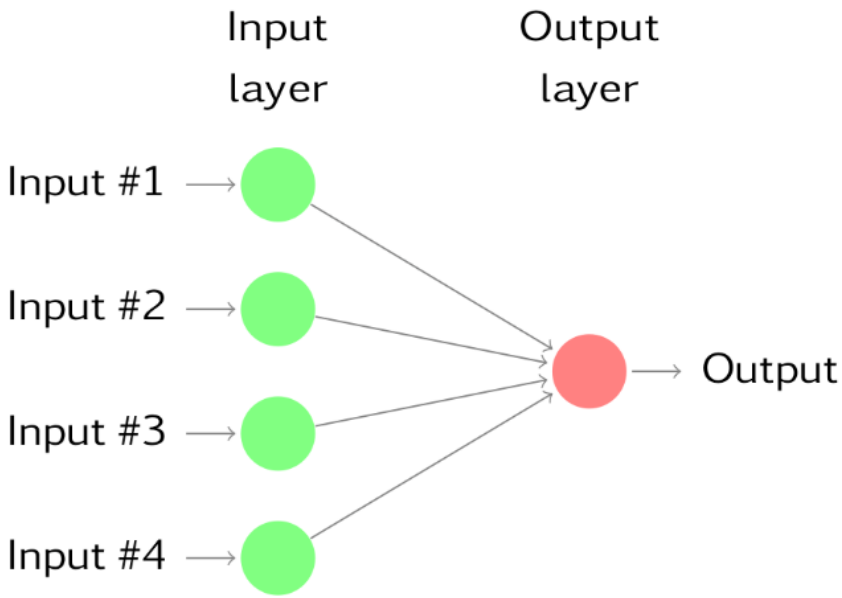

Рисунок 2.4 – Проста нейрона мережа, еквівалентна лінійній регресії

Коефіцієнти, додані до цих предикторів, називаються «вагами». Прогнози отримують шляхом лінійної комбінації вхідних даних. Вагові коефіцієнти вибираються в структурі мережі за допомогою «алгоритму навчання», який мінімізує «функцію вартості», таку як MSE [33].

Як тільки ми додаємо проміжний шар із прихованими нейронами, нейрона мережа стає нелінійною. Простий приклад показано на рис. 2.5.

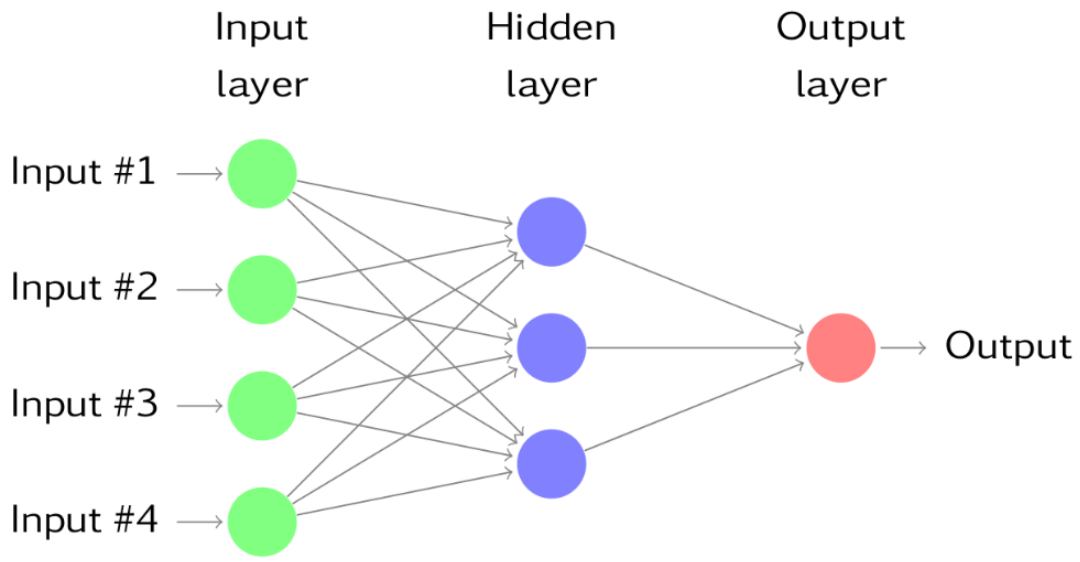

Рисунок 2.5 – Нейрона мережа з чотирма входами та одним прихованим шаром із трьома прихованими нейронами

Ця структура відома як багаторівнева мережа прямого зв’язку, де кожен рівень вузлів отримує вхідні дані від попередніх рівнів. Вхідні дані для кожного вузла об’єднуються за допомогою зваженої лінійної комбінації

$$z_j = b_j + \sum_{i=1}^{4} w_{i,j} x_i . \qquad (2.18)$$

У прихованому шарі це потім змінюється за допомогою нелінійної функції, такої як сигмоїда, щоб надати вхідні дані для наступного шару. Це, як правило, зменшує вплив екстремальних вхідних значень, таким чином роблячи мережу дещо стійкою до викидів [33].

Параметри в (2.18) «навчаються» з даних [33]. Значення ваг часто обмежені, щоб запобігти тому, щоб вони стали занадто великими. Спочатку ваги приймають випадкові значення, які потім оновлюються з використанням спостережених даних. Отже, є елемент випадковості в прогнозах, створених

нейроною мережею. Тому мережа зазвичай навчається кілька разів з використанням різних випадкових початкових точок, а результати усереднюються. Кількість прихованих шарів і кількість вузлів у кожному прихованому шарі необхідно вказати заздалегідь.

3 ДОСЛІДЖЕННЯ ТОЧНОСТІ АЛГОРИТМІВ ПРОГНОЗУВАННЯ НА ОДИН КРОК

3.1 Показники точності прогнозування і експериментальні дані трафіку

Точність алгоримів прогнозування зазвичай характеризується показником нормованої середньоквадратичної похибки прогнозування NMSE (Normalized Mean Square Error) [30]:

$$Q_{NMSE} = \sqrt{\frac{\sum_{k=1}^{K} (\hat{X}_k - X_k)^2}{\sum_{k=1}^{K} X_k{}^2}}, \tag{3.1}$$

де X_k – послідовність спостережень;

$\hat{X}_k$ – прогнозні значення X_k;

K – кількість відліків.

Показник NMSE може приймати невід`ємні значення: 0 – ідеальний прогноз; 1 – тривіальний прогноз. Чим менше значення показника NMSE, тим кращим є прогноз.

У даній роботі будемо використовувані наступний показник якості прогнозування

$$Q_{NMSE} = \frac{\sum_{k=1}^{K} (\hat{X}_k - X_k)^2}{\sum_{k=1}^{K} X_k{}^2}. \tag{3.2}$$

У даній роботі для дослідження точності алгоритмів прогнозування використана реалізація трафіку, яка була отримана авторами роботи [24]. Вона названа авторами BC-Oct89Ext.TL і являє собою залежність від часу розмірів мережних кадрів (Ethernet-кадрів. Дані представлені двома колонками чисел: перший стовпець представляє собою час в секундах, другий – довжину поля даних Ethernet-кадру в байтах [24] .

Оцінка кореляційної функції трафіку з рівнем агрегації ΔT=0.1 с наведена на рис. 3.1. Нижче на рис. 3.2 наведена оцінка кореляційної функції того ж трафіку з рівнем агрегації ΔT=0.5 с.

Наявність кореляції сусідніх значень відліків трафіку дає підстави до його прогнозування. Повільний спад кореляційної функції повинне забезпечити можливість прогнозування трафіку навіть на великих проміжках часу.

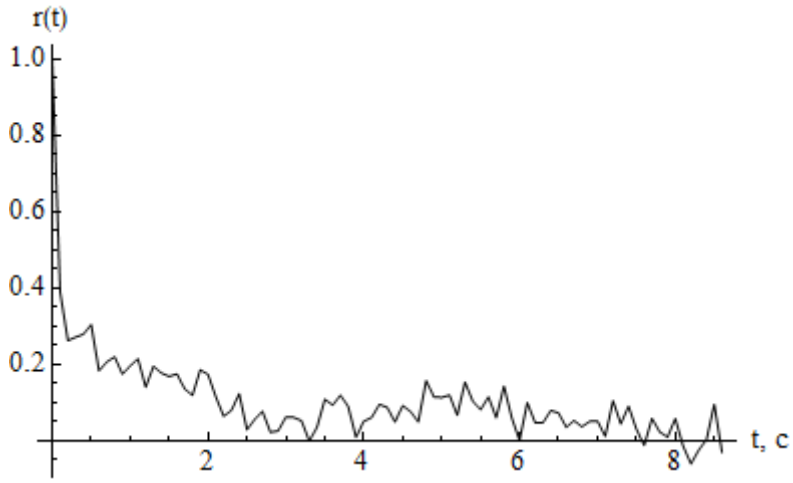

Рисунок 3.1 – Оцінка кореляційної функції трафіку з ΔT=0.1

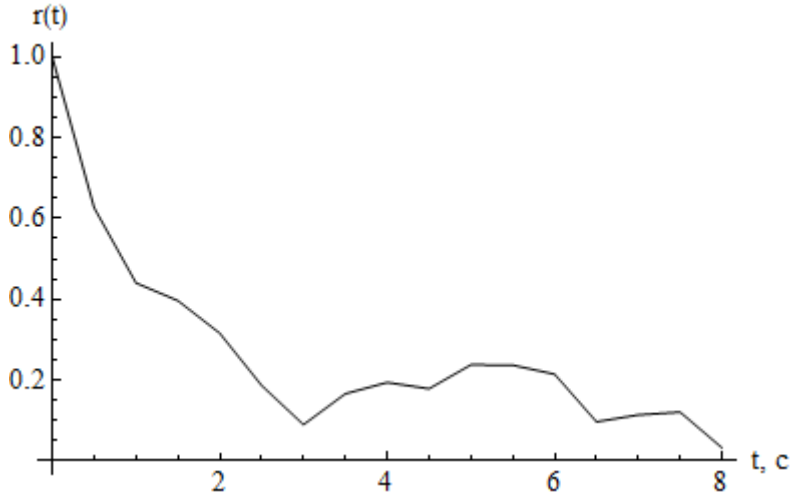

Рисунок 3.2 – Оцінка кореляційної функції трафіку з ΔT=0.5

3.2 Аналіз результатів дослідження точності прогнозування

Дослідження точності алгоритмів прогнозування трафіку проводилися в середовищі програмування Mathematica.

Метою першого етапу досліджень було визначення оптимального кроку агрегування трафіку. Для цьому досліджувалася залежність показника точності прогнозування Q_{NMSE} від величини кроку агрегування трафіку Δ. Результати виведено подані у вигляді графіків на рис. 3.3. Тут по коротким пунктиром показана залежність для алгоритму останнього значення LV; довгим пунктиром – для авторегресійного прогнозу порядку 1; суцільною кривою – для авторегресійного прогнозу порядку 2.

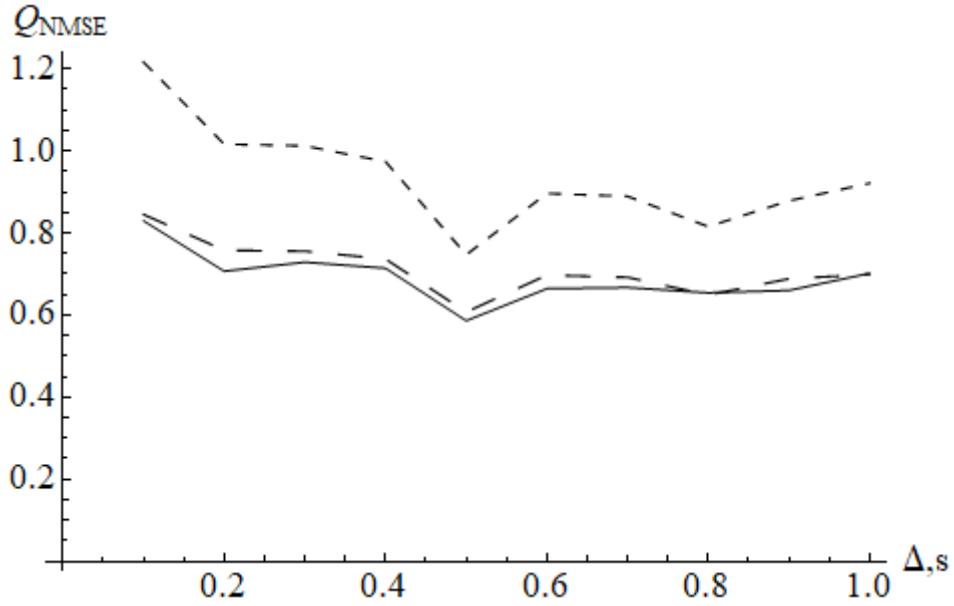

Рисунок 3.3 – Залежності показника точності Q_{NMSE} від кроку агрегування Δ

Із графіків, зображених на рис. 3.3, видно що авторегресійна модель забезпечує кращий прогноз, ніж метод LV, а найкращі значення прогнозу досягаються при кроці агрегування $\Delta = 0{,}5\ c$. Такий результат можна пояснити, що зі збільшенням кроку агрегування трафік дещо згладжується, проте при цьому зростає крок прогнозування. Значення $\Delta = 0{,}5\ c$ у даному випадку є компромісним.

На другому етапі досліджувалася залежність точності прогнозування трафіку на один крок від величини коефіцієнтів згладжування для алгоритмів простого і подвійного експоненціального згладжування. Результати дослідження подані на рис. 3.4. Тут по осі абсцис відкладені значення коефіцієнту α для алгоритму простого згладжування α1 для подвійного

експоненціального згладжування, а по осі ординат –значення показника точності Q_{NMSE}. Суцільній кривій тут відповідає алгоритм простого експоненціального згладжування. Короткому і довгому пунктиру відповідає алгоритм подвійного згладжування зі значенням параметру α2=0.1 і α2=0.2 відповідно.

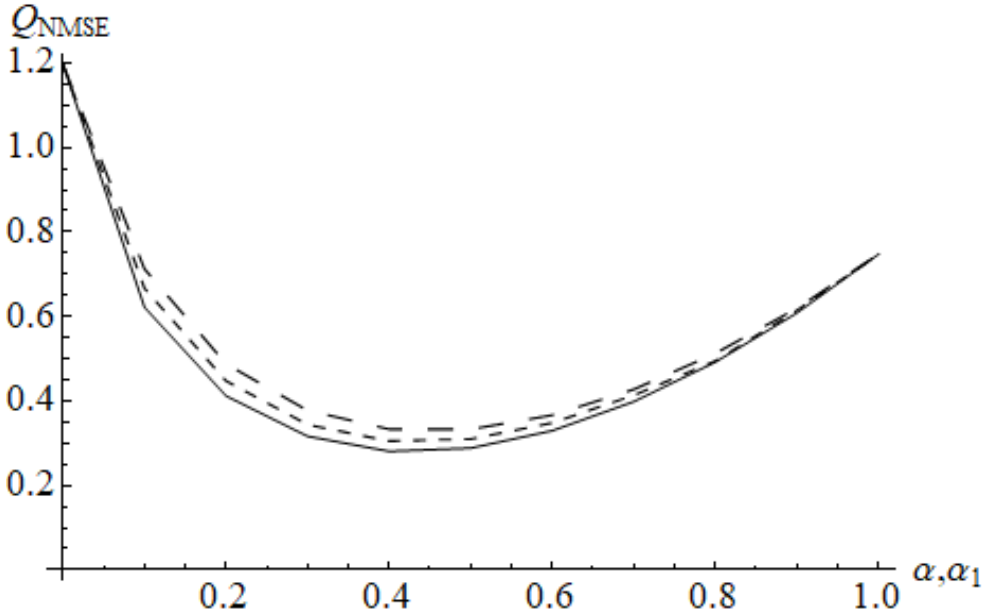

Рисунок 3.4 – Залежності показника точності Q_{NMSE} від кроку агрегування Δ

Аналіз графіків, наведених на рис. 3.4 дозволяє зробити висновок, що найкращий результат прогнозування даного трафіку забезпечується алгоритмом експоненціального згладжування. При цьому найкраще значення параметру згладжування α=0,4.

Робота алгоритму експоненціального згладжування зі значенням α=0,4 проілюстрована на рис. 3.5. Тут пунктиром показані значення трафіку з $\Delta = 0{,}5\ c$, а суцільною кривою прогнозовані значення.

Із результатів дослідження, отриманих у даному розділі можна зробити такі висновки:

1) найкращий прогноз заданого трафіка забезпечує просте експоненціальне згладжування. Такий прогноз еквівалентний, як раніше зазначалося, використанню моделі часового ряду ARIMA(0,1,1).

2) перехід до подвійного згладжування не дає покращення точності прогнозування, що можна пояснити відсутністю у заданому трафіку явно вираженого тренду;

3) використання авторегресійних моделей порядку 1 та 2, які належать класам ARIMA(1,0,0) та ARIMA(1,0,0) відповідно, значно поступається у точності прогнозування простому експоненціальному згладжуванню, що є оптимальним для процесу ARIMA(0,1,1).

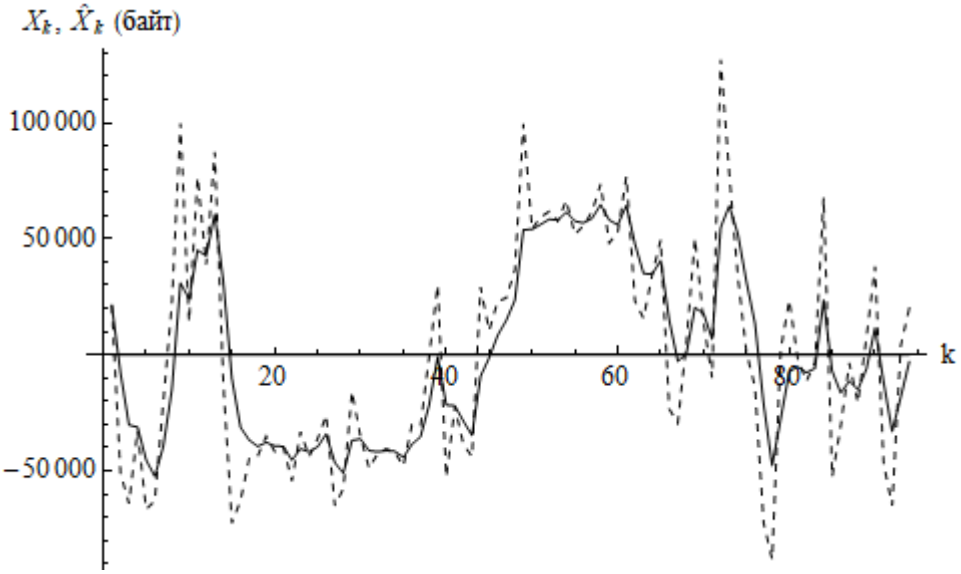

Рисунок 3.5 – Трафік і його прогноз

4 ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ДОВГОТЕРМІНОВОГО ПРОГНОЗУВАННЯ

У розд. 3 було досліджено ефективність алгоритмів прогнозування на один крок вперед. Дослідимо алгоритми прогнозування на багато кроків вперед. Потенціальна можливість такого прогнозу визначається повільнім згасанням його кореляційної функції.

4.1 Задача прогнозування самоподібних процесів з найменшою середньоквадратичною похибкою

У постановці задачі будемо вважати, що спостерігається випадкова послідовність спостережень

$$z_k = \vec{x}_k \vec{a} + \xi_k,\ k = \overline{0,t-1}\ \ , \tag{4.1}$$

де $\vec{a} = (a_0,...,a_p)$ – вектор коефіцієнтів регресії розмірності $p+1$;

$\vec{x}_k = (1,k,...,k^p),\ k = \overline{0,t-1}$ – вектор-строчка значень регресорів;

$\xi_k,\ k = \overline{0,t-1}$ – послідовність випадкових величин, що утворюють вектор $\vec{\xi} = (\xi_0,...,\xi_{t-1})$ з кореляційною матрицею $M[\vec{\xi}\vec{\xi}^{tr}] = R$;

p – порядок моделі.

Необхідно знайти оптимальний у середньоквадратичному розумінні прогноз послідовності $\hat{z}_{t+h}$ для відліку z_{t+h}, $h \geq 0$.

Якщо значення вектора параметрів $\vec{a}$ відомо то оптимальний лінійний прогноз має наступний вид [34, 35]

$$\hat{z}_{t+h} = \vec{r}_c^{\,tr} R^{-1} (\vec{z} - X\vec{a}) + \vec{x}_{t+h}\vec{a}, \tag{4.2}$$

де

$$X = \begin{bmatrix} x_{0,0} & x_{0,1} & x_{0,2} & \cdots & x_{0,p-1} \\ x_{1,0} & x_{1,1} & x_{1,2} & \cdots & x_{1,p-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{t-1,0} & x_{t-1,1} & x_{t-1,2} & \cdots & x_{t-1,p-1} \end{bmatrix} - \tag{4.3}$$

регресійна матриця розміром $t \times p$;

$R = M[\vec{\xi}\vec{\xi}^{tr}]$ – кореляційна матриця розміру $t \times t$;

$\vec{r}_c = M[\xi_{t+h}\vec{\xi}]$ – вектор взаємних кореляцій величини ξ_{t+h} с відліками вектора $\vec{\xi}$.

Зауважимо що вираз (4.2) допускає подання у іншому виді

$$\hat{z}_{t+h} = \vec{r}_c^{\,tr} R^{-1}\vec{z} + (\vec{x}_{t+h} - \vec{r}_c^{\,tr} R^{-1} X)\vec{a}. \qquad (4.4)$$

Відомо [34, 35], що оцінка лінійного прогнозу (4.2) є незміщеною, а його дисперсія

$$D_0 = (\hat{z}_{t+h} - z_{t+h})^2 = \sigma^2 - \vec{r}_c^{\,tr} R^{-1}\vec{r}_c, \qquad (4.5)$$

де σ^2 – дисперсії процесу ξ_k.

Для випадку, коли вектор параметрів тренду $\vec{a}$ є невідомим, в [36] обґрунтовано наступний алгоритм адаптивного прогнозу

$$\hat{z}_{t+h} = \vec{r}_c^{\,tr} R^{-1}\vec{z} + (\vec{x}_{t+h} - \vec{r}_c^{\,tr} R^{-1} X)\hat{\vec{a}}, \qquad (4.6)$$

де $\hat{\vec{a}}$ оцінка МНК невідомого параметра $\vec{a}$.

Прогноз (4.6) є також лінійним та незміщеним, а його дисперсія визначається наступним чином [36]

$$\begin{aligned} D_0 = {} & \sigma^2 - \vec{r}_c^{\,tr} R^{-1}\vec{r}_c + \vec{x}_{t+h} V \vec{x}_{t+h}^{\,tr} - \\ & - 2\vec{x}_{t+h} V X^{\,tr} R^{-1}\vec{r}_c + \vec{r}_c^{\,tr} R^{-1} X V X^{\,tr} R^{-1}\vec{r}_c \end{aligned} \qquad (4.7)$$

4.2 Розрахунок точності найкращого лінійного прогнозу

Опишемо результати аналітичних розрахунків без тренду. На рис. 4.1 подані аналітичні залежності нормованої середньоквадратичної похибки прогнозування Q_{NMSE} від величини кроку прогнозування h, побудовані з

використанням виразу (4.5) і врахуванням виду кореляційної функції самоподібного процесу (1.5). Тут різні криві відповідають різним значенням параметра Херста H: 0,55; 0,65; 0,75; 0,85; 0,95, причому з ростом значення параметру Херста довжина пунктиру зростає, а суцільна крива відповідає значенню H=0,95.

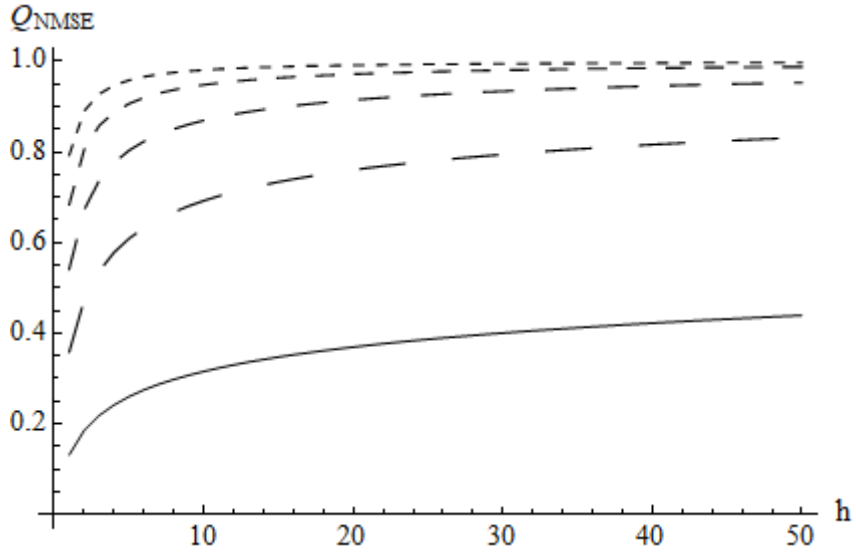

Рисунок 4.1 – Залежності нормованої середньоквадратичної похибки прогнозування від кроку прогнозування при різних значеннях параметру Херста

З аналізу залежностей приведених на рис. 4.1 можна зробити висновок, що ключове значення на ефективність прогнозування самоподібного процесу має значення параметру Херста. При H<0.55 величина Q_{NMSE} самоподібного процесу перевищує 0,8. Це означає, що довготермінове прогнозування у цьому випадку малоефективне, оскільки вибіркове середнє значення трафіку є не набагато гіршим прогнозом (тут забезпечується Q_{NMSE} =1).

На рис. 4.2 подані аналітичні залежності Q_{NMSE} від величини кроку прогнозування h для випадку послідовності з кореляційною функцією реального трафіку з рівнем агрегації ΔT=0.5, що показана на рис. 3.2. Як уже зазначалось у цьому випадку значення параметру Херста H=0,72. Результати розрахунків для реального трафіку, що відображені на рис. 4.2 досить точно узгоджуються з результатами розрахунків для самоподібного процесу з H=0,75, що відображені на рис. 4.1.

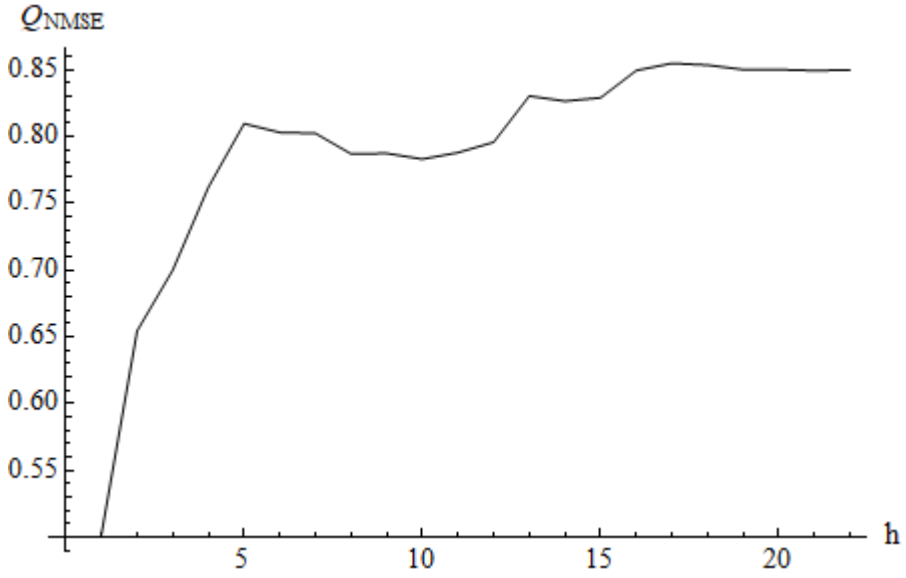

Рисунок 4.2 – Залежності нормованої середньоквадратичної похибки прогнозування від величини кроку прогнозування при різних значеннях параметру Херста для реального трафіку

Зазначимо, що отримані у розділі 3 оцінки точності прогнозу реального трафіку на один крок з використанням моделі процесу AR зі значенням Q_{NMSE} =0,6 цілком відповідають теоретичним значенням, показаним на рис 4.1, де Q_{NMSE}=0,5. Проте простий алгоритм подвійного експоненціального згладжування Хольта забезпечив для тих же умов значно кращу точність прогнозу на один крок з Q_{NMSE}=0,3. Такий висновок можна пояснити тим, що модель стаціонарного процесу, що використовувалася нами при описі трафіку не є адекватною.

ВИСНОВКИ

Результати даної роботи підтверджують наявність фрактальних властивостей трафіку комп`ютерних мереж. Цю властивість можна ефективно використовувати для прогнозування трафіку, особливо довготермінового.

У роботі з використанням реального трафіку виконано порівняльний аналіз декількох алгоритмів прогнозування: LV-предиктору; прогнозування на основі моделей ARIMA(1,0,0) та ARIMA(1,0,0); алгоритму простого експоненціального згладжування; алгоритму подвійного експоненціального згладжування. Для порівняння алгоритмів було використано показник середньоквадратичної помилки прогнозування. В результаті досліджень було встановлено наступне.

Найкращий прогноз заданого трафіку забезпечує просте експоненціальне згладжування. Такий прогноз еквівалентний, як раніше зазначалося, використанню моделі часового ряду ARIMA(0,1,1).

Перехід до подвійного згладжування не дає покращення точності прогнозування, що можна пояснити відсутністю у заданому трафіку явно вираженого тренду.

Використання авторегресійних моделей порядку 1 та 2, які належать класам ARIMA(1,0,0) та ARIMA(1,0,0) відповідно, значно поступається у точності прогнозування простому експоненціальному згладжуванню, яке є оптимальним бля процесу ARIMA(0,1,1).

Було досліджено також ефективність алгоритмів довготермінового прогнозування на багато кроків вперед.

Були побудовані залежності показника точності оптимального прогнозування ССШР процесів від кроку прогнозування при різних значеннях параметру Херста. Встановлено, що ключовий вплив на ефективність прогнозування самоподібного процесу має значення параметру Херста. При $H<0.55$ величина Q_{NMSE} самоподібного процесу перевищує 0,8.

Були побудовані також залежності точності прогнозування від кроку прогнозування при різних значеннях параметру Херста для реального трафіку. В результаті було встановлено, що оцінки точності прогнозу реального трафіку на один крок з використанням моделі процесу авторегресії зі значенням Q_{NMSE} =0,6 цілком відповідають теоретичним значенням. Проте простий алгоритм

подвійного експоненціального згладжування Хольта забезпечив для тих же умов значно кращу точність прогнозу на один крок з Q_{NMSE} =0,3. Такий висновок можна пояснити тим, що модель стаціонарного процесу, що використовується при описі трафіку не є адекватною.

ПЕРЕЛІК ПОСИЛАНЬ

1. CISCO: Cisco Annual Internet Report (2018–2023), CISCO, 2020.

2. R. Li, Z. Zhao, X. Zhou, and H. Zhang, “Energy savings scheme in radio access networks via compressive sensingbased traffic load prediction,” Transactions on Emerging Telecommunications Technologies, vol. 25, no. 4, pp. 468–478, 2012.

3. U. Paul, M. Buddhikot, and S. R. Das, “Opportunistic traffic scheduling in cellular networks,” in Proceedings of IEEE International Symposium on Dynamic Spectrum Access Networks, pp. 339–348, Bellevue, WA, USA, October 2012.

4. Z. Li, J. Bi, and S. Chen, “Traffic prediction-based fast routing algorithm for wireless multimedia sensor networks,” International Journal of Distributed Sensor Networks, vol. 9, no. 5, Article ID 176293, 2013.

5. R. Li, Z. Zhao, J. Zheng, C. Mei, Y. Cai, and H. Zhang, “The learning and prediction of application-level traffic data in cellular networks,” IEEE Transactions on Wireless Communications, vol. 16, no. 6, pp. 3899–3912, 2017.

6. M. F. Iqbal, Workload-Aware Network Processorse. University of Texas at Austin, Austin, TX, USA, 2013.

7. H. Kawase, Y. Mori, H. Hasegawa, and K.-c. Sato, “Dynamic router performance control utilizing support vector machines for energy consumption reduction,” IEEE Transactions onNetwork and Service Management, vol. 13, no. 4, pp. 860–870, 2016.

8. M. F. Iqbal and L. K. John, “Efficient traffic aware power management for multicore communications processors,” in Proceedings of IEEE/ACM Conference on Architecture for Network and Communication Systems, Austin, TX, USA, July 2012.

9. A. Svigelj, R. Sernec, and K. Alic, “Network traffic modeling for load prediction: a user-centric approach,” IEEE Network, vol. 29, no. 4, pp. 88–96, 2015.

10. R. Sivakumar, E. A. Kumar, and G. Sivaradje, “Prediction of traffic load in wireless networks using time series model,” in Proceedings of International Conference on Process Automation, Control and Computing (PACC), pp. 1–6, Coimbatore, India, July 2011.

11. S.-J. Sang-Jo Yoo, “Efficient traffic prediction scheme for realtime VBR MPEG video transmission over high-speed networks,” IEEE Transactions on Broadcasting, vol. 48, no. 1, pp. 10–18, 2002.

12. W. Xu and A. G. Qureshi, “Adaptive linear prediction of MPEG video traffic,” in Proceedings of Fifth International Symposium on Signal Processing and its Applications, vol. 1, pp. 67–70, Brisbane, Australia, August,1999.

13. A. Abdennour, “Short-term MPEG-4 video traffic prediction using ANFIS,” International Journal of Network Management, vol. 15, no. 6, pp. 377–392, 2005.

14. A. D. Doulamis, N. D. Doulamis, and S. D. Kollias, “Recursive non linear models for on line traffic prediction of VBR MPEG coded video sources,” in Proceedings of IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, vol. 6, pp. 114–119, Como, Italy, July, 2000.

15. A. S. Soares, T. P. Oliveira, and J. S. Barbar, “Computer network traffic prediction: a comparison between traditional and deep learning neural networks,” International Journal of Big Data Intelligence, vol. 3, no. 1, p. 28, 2016.

16. J. Jiang and S. Papavassiliou, “Enhancing network traffic prediction and detection via statistical network traffic separation and combination strategies,” Computer Communications, vol. 29, no. 10, pp. 1627–1638, 2006.

17. P. Romirer-Maierhofer, M. Schiavone, and A. D’Alconzo, “Device specific traffic characterization for root cause analysis in cellular networks,” in Proceedings of International Workshop on Traffic Monitoring and Analysis, Barcelona, Spain, April 2015.

18. C.-J. Yi, Y. Geng, H. dai, liang Liu, and L. Ning, “Switching algorithm with prediction strategy for maximizing lifetime in wireless sensor network,” International Journal of Distributed Sensor Networks, vol. 11, no. 11, Article ID 592093, 2015.

19. Y. Cai and L. Yu, “Sensor network traffic load prediction with markov random field theory,” in Proceedings of IEEE International Conference on Computer Science and Network Technology, Harbin, China, December 2015.

20. S. Basu, A. Mukherjee, and S. Klivansky, “Time series models for internet traffic,” in Proceedings of Fifteenth Annual Joint Conference of the IEEE Computer and Communications Societies Conference on Computer Communications - Volume 2,INFOCOM’96, pp. 611–620, Washington, DC, USA, March 1996.

21. M. F. Zhani, H. Elbiaze, and F. Kamoun, “Analysis of prediction performance of training based models using real network traffic,” International Journal of Computer Applications in Technology, vol. 37, no. 1, 2010.

20. Adrian Popescu. Traffic Self-Similarity, Department of Telecommunications and Signal Processing. Blekinge Institute of Technology Karlskrona / Swede. – 8 p.

21. Beran J., Sherman R., Taqqu M.S. and Willinger, W., Long-Range Dependence in Variable-Bit Rate Video Traffic, IEEE Transactions on Communications, Vol. 43, No. 2/3/4, February/March/April 1995.

22. Crovella M.E. and Bestavros A., Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes, Proceedings of ACM SIGMETRICS’96, 1996.

23. Paxson V. and Floyd S., Wide Area Traffic: The Failure of Poisson Modeling, IEEE/ACM Transactions on Networking, Vol. 3, No. 3, June 1995. [20] Peitgen H-O., Jur ¨ gens H. and Saupe D., Chaos and Fractals New Frontiers of Science, SpringerVerlag, 1992.

24. Leland W.E., Taqqu M.S., Willinger W., and Wilson D.V. On the self-similar nature of ethernet traffic // IEEE/ACM Transactions of Networking, 2(1), 1994. P. 1-15.

25. Causey J.W. and Kim H.S., Comparison of Admission Control Schemes in ATM Networks, International Journal of Communication Systems, Vol. 8, 1995.

26. Цыбаков Б.С. Модель телетрафика на основе самоподобного случайного процесса // Радиотехника. – 1999. – № 5. – С. 24-31.

27. Duffy D.E., McIntosh A.A., Rosenstein M. And Willinger, W., Statistical Analysis of CCSN/SS7 Traffic Data from Working Subnetworks, IEEE Journal on Selected Areas in Communications, Vol. 12, No. 3, 1994.

28. Петров В.В. Статистический анализ сетевого трафика. МЭИ, ИРЭ, Москва, Красноказарменная 13, 2003. 47 с.

29. Петров В.В. Структура телетрафика и алгоритмы обеспечения качества обслуживания при влиянии эффекта. Диссертация на соискание ученой степени к. т. н. – М., 2004. 199 с.

30. Muhammad Faisal Iqbal , Muhammad Zahid, Durdana Habib and Lizy Kurian John. Efficient Prediction of Network Traffic for Real-Time Applications. Journal of Computer Networks and Communications Volume 2019, Article ID 4067135, 11 p.

31 Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: Учеб. пособие. - М.: Финансы и статистика, 2003.- 416 с: ил

32. Бокс Дж., Дженкинс Г. Анализ временных рядов. Вып.1. Прогноз и управление. – М.: Мир, 1974.

33. Гафаров Ф.М Г12 Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018.

34. Колмогоров А.Н. Интерполирование и экстраполирование стационарных случайных последовательностей. Изв. АН СССР, сер. матем., 1941, т.5, с. 3-14.

35. Wiener N. Extrapolation, interpolation and smoothing of stationary time series. N. Y., 1949, 456 p.

36. Омельченко А. В., Роздымаха Е. А., Федоров А. В. Прогнозирование временных рядов со степенными регрессорами при коррелированном шуме наблюдения. Наукоемкие технологии в инфокоммуникациях: обработка и защита информации: коллективная монография / под ред. В.М. Безрука, В.В. Баранника. – Х.: Компания СМИТ, 2013. с .262 – 272.