Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

факультет прикладних інформаційних технологій та електроінженерії
(повна назва факультету)
кафедра радіотехнічних систем
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(назва освітнього ступеня)

на тему: **Метод та програмний засіб виявлення аномалій у часових профілях трафіку комп'ютерних мереж**

Виконав: студент II курсу, групи РАм-61
спеціальності 172 Телекомунікації та радіотехніка
(шифр і назва спеціальності)

(підпис) Воробець І.І. (прізвище та ініціали)

Керівник (підпис) Хвостівська Л.В. (прізвище та ініціали)

Нормоконтроль (підпис) Хвостівська Л.В. (прізвище та ініціали)

Завідувач кафедри (підпис) Дунець В.Л. (прізвище та ініціали)

Рецензент (підпис) Яворська Є.Б. (прізвище та ініціали)

Тернопіль
2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет прикладних інформаційних технологій та електроінженерії
(повна назва факультету)

Кафедра радіотехнічних систем
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Дунець В.Л.
(підпис) (прізвище та ініціали)
« 26 » листопада 2025 р.

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня магістр
(назва освітнього ступеня)

за спеціальністю 172 Телекомунікації та радіотехніка
(шифр і назва спеціальності)

студенту Воробець Ігор Ігорович
(прізвище, ім'я, по батькові)

1. Тема роботи Метод та програмний засіб виявлення аномалій у часових профілях трафіку комп'ютерних мереж

Керівник роботи Хвостівська Лілія Володимирівна, к.т.н., доц.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від « 26 » листопада 2025 року № 4/7-1038 .

2. Термін подання студентом завершеної роботи

3. Вихідні дані до роботи : Об'єкт дослідження: процес автоматичного виявлення аномалій у часових профілях трафіку комп'ютерних мереж. Предмет дослідження: метод та програмне забезпечення аналізу часових рядів трафіку комп'ютерних мереж з метою виявлення аномалій.

4. Зміст роботи (перелік питань, які потрібно розробити)

РОЗДІЛ 1. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА ЗАСОБІВ ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖЕВОМУ ТРАФІКУ

РОЗДІЛ 2. МАТЕМАТИЧНА МОДЕЛЬ ТА МЕТОД ВИЯВЛЕННЯ АНОМАЛІЙ У ЧАСОВИХ ПРОФІЛЯХ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

РОЗДІЛ 3. ПРОГРАМНИЙ ЗАСІБ ТА РЕЗУЛЬТАТИ ВИЯВЛЕННЯ АНОМАЛІЙ У ЧАСОВИХ ПРОФІЛЯХ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Актуальність роботи
2. Математична модель трафіку комп'ютерних мереж з елементами аномалій
3. Метод виявлення аномалій у часових профілях трафіку комп'ютерних мереж
4. Алгоритм виявлення аномалій у часових профілях трафіку комп'ютерних мереж
5. Програмний засіб виявлення аномалій у часових профілях трафіку комп'ютерних мереж
6. Результати виявлення аномалій у часових профілях трафіку комп'ютерних мереж
7. Наукова новизна
8. Наукова публікація за напрямом дослідження
9. Загальні висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці та	Окіпний І.Б., зав. каф. МТ		
безпека в надзвичайних	Стручок В.С., ст. викл. каф. ОХ		
ситуаціях			

7. Дата видачі завдання 03.09.2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	**Примітка**
1	Отримання завдання на виконання роботи	03.09.2025	
2	Аналіз отриманого завдання	13.09.2025	
3	Виконання розділу 1	23.09.2025	
4	Виконання розділу 2	10.10.2025	
5	Виконання розділу 3	25.10.2025	
6	Виконання розділу 4	15.11.2025	
7	Оформлення пояснювальної записки	20.11.2025	
8	Оформлення презентаційного матеріалу	26.11.205	
9	Перевірка кваліфікаційної роботи на антиплагіат	28.11.2025	
10	Попередній захист кваліфікаційної роботи	09.12.2025	
11	Захист кваліфікаційної роботи	23.12.2025	

Студент _____ Воробець І.І.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Хвостівська Л.В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Тема кваліфікаційної роботи: «Метод та програмний засіб виявлення аномалій у часових профілях трафіку комп’ютерних мереж» // Кваліфікаційна робота // Воробець Ігор Ігорович // ТНТУ, факультет прикладних інформаційних технологій та електроінженерії, група РАм-61 // Тернопіль, 2025 // с. – 82, рис. – 10, табл. – 2, додат. – 2, бібліогр. – 29.

Ключові слова: мережевий трафік, аномалії, часові профілі, динамічний модальний розклад (DMD), мультискейловий аналіз, робастна нормалізація, виявлення аномалій, MATLAB.

У роботі розглянуто проблему виявлення аномалій у часових профілях трафіку комп’ютерних мереж, що є актуальною через зростання обсягів передавання даних, складність мережевих структур та збільшення кількості кіберзагроз. Проведено аналіз існуючих статистичних, спектральних, фрактальних та інтелектуальних методів детекції аномалій, визначено їх обмеження щодо роботи з багатомасштабними, нестаціонарними та шумовими потоками даних.

Запропоновано математичну модель трафіку, яка описує взаємодію повільних трендів, локальних коливань, шумових збурень і аномальних відхилень. На основі цієї моделі розроблено метод виявлення аномалій, що поєднує динамічний модальний розклад (DMD), мультискейлову обробку та робастні статистичні критерії. Метод забезпечує виділення як короткочасних пікових подій, так і довготривалих структурних змін у трафіку.

Розроблено програмний засіб у MATLAB, який реалізує мультискейловий DMD-аналіз, формує енергетичні та спектральні показники, виконує робастну нормалізацію та автоматичну порогову детекцію аномалій. Проведено експериментальні дослідження на синтетичних та модельованих даних із добовою мінливістю навантаження. Отримано результати, що демонструють високу чутливість методу, стійкість до шумів та здатність виявляти аномалії різної природи.

ANNOTATION

Theme of qualification work: «Method and software tool for anomaly detection in time profiles of computer network traffic» // Vorobets Ihor // Ternopil Ivan Puluj National Technical University, Faculty of Applied Information Technologies and Electrical Engineering group RAm-61 // Ternopil, 2025 // p. – 82, fig. – 10, tab. - 2, add. – 2, bibliography -29.

Keywords: network traffic, anomalies, time profiles, Dynamic Mode Decomposition (DMD), multiscale analysis, robust normalization, anomaly detection, MATLAB.

The work focuses on developing a method and software tool for detecting anomalies in time profiles of computer network traffic. The study highlights the challenges posed by non-stationarity, multiscale dynamics and noise inherent in modern network flows. Existing statistical, spectral, fractal and machine-learning approaches are reviewed, and their limitations in analysing variable and bursty traffic are identified.

A mathematical model of traffic is proposed, representing the signal as a combination of slow trends, local oscillations, noise and anomalous events. Based on this model, a method employing multiscale Dynamic Mode Decomposition is developed. The method extracts dominant temporal modes of traffic behaviour and computes energy-based and spectral indicators of deviations. Robust normalization and adaptive thresholding ensure reliable detection under changing load conditions.

A MATLAB software tool implementing the proposed method is created, providing automated processing, visualization of modal spectra, reconstruction errors and anomaly detection results. Experimental studies confirm the effectiveness, robustness and interpretability of the approach, demonstrating its capability to detect both short-term spikes and long-term structural changes in network traffic.

ЗМІСТ

ВСТУП

Актуальність роботи

Сучасні комп’ютерні мережі складаються з великої кількості вузлів, протоколів та сервісів. Обсяги передаваної інформації швидко зростають, з’являються нові види утилізацій, шкідливих активностей та атак. В таких умовах важливо своєчасно виявляти аномалії в поведінці мережевого трафіку, особливо в часових профілях, що змінюються упродовж доби чи тижня. Неефективне виявлення може призводити до порушення якості обслуговування, зниження безпеки, втрат у продуктивності та економічних збитків.

У науковій літературі існує значна кількість методів виявлення аномалій, що охоплюють статистичні, спектральні, фрактальні підходи, а також методи машинного та глибокого навчання. Статистичні методм забезпечують простоту та інтерпретованість, проте часто виявляються чутливими до нестаціонарності трафіку. Спектральні методи та аналіз динаміки (Thottan & Ji; Lakhina, Crovella, Diot) забезпечують опис частотних властивостей трафіку, проте мають обмежену здатність локалізувати короткі аномалії. Фрактальні підходи, основані на працях Херста та подальших MFDFA-розробках, є ефективними для дослідження довготривалих залежностей, але малопридатні для оперативного аналізу. Методи машинного навчання (Cortes & Vapnik; MacQueen; Patcha & Park) демонструють високу гнучкість, однак вимагають великих навчальних вибірок та не гарантують інтерпретованості рішень.

Сумарно існуючі методи показують обмежену здатність працювати з багатомасштабним, нестаціонарним трафіком і часто не забезпечують одночасно інтерпретованості, робастності та адаптивності. Додатковою проблемою є те, що мережеві процеси містять компоненти з різними часовими масштабами – від коротких піків до повільних трендових змін – що унеможливлює їх адекватний аналіз в межах одного часовго вікна.

Це обґрунтовує необхідність методів, здатних враховувати мультискейлову структуру сигналу, відокремлювати типову динаміку від нетипових змін,

забезпечувати робастність до шумів та зберігати інтерпретованість отриманих характеристик.

Тому актуальною є розробка методу, та розробленого на його основі прогармного засобу, які би враховували багатомасштабну структуру трафіку, дозволяли відокремлювати типову динаміку від нетипових процесів, працювали у режимі реального часу та залишалися стійким до шумів і одиничних пікових значень.

Мета дослідження: розробка методу та програмного засобу, який дозволяє автоматично виявляти аномалії у часових профілях мережевого трафіку, з урахуванням добових і тижневих циклів, зміни поведінки, статистичних характеристик, з високою точністю і з можливістю застосування у реальному середовищі.

Задачі дослідження:

1. Провести детальний аналіз і класифікацію існуючих методів виявлення аномалій у часових профілях мережевого трафіку.
2. Розробити математичну модель мережевого трафіку**,** яка враховує добові та тижневі цикли, локальні коливання, шумові складові та аномальні події, а також підтримує представлення багатоканальної структури мережі.
3. Розробити метод та алгоритм виявлення аномалій у часових профілях трафіку комп’ютерних мереж, який враховує багатомасштабну динаміку мережевих процесів і забезпечує виділення відхилень від типових закономірностей.
4. Реалізувати програмний засіб для автоматизованого аналізу та виявлення аномалій у часових профілях трафіку комп’ютерних мереж, включаючи візуалізацію динаміки, спектрів мод та виділених аномальних ділянок.
5. Провести експериментальні дослідження щодо виявлення аномалій у часових профілях трафіку комп’ютерних мереж.

Об’єкт дослідження: процес автоматичного виявлення аномалій у часових профілях трафіку комп’ютерних мереж.

Предмет дослідження: метод та програмне забезпечення аналізу часових рядів трафіку комп’ютерних мереж з метою виявлення аномалій.

Наукова новизна.

Вперше розроблено метод виявлення аномалій у часових профілях трафіку комп’ютерних мереж, у якому багатомасштабна динаміка сигналу аналізується за допомогою узагальненого динамічного модального підходу, що забезпечило підвищення чутливості до короткочасних та довготривалих відхилень і дозволило виявляти аномалії різної природи в умовах нестаціонарності трафіку.

РОЗДІЛ 1
АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА ЗАСОБІВ ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖЕВОМУ ТРАФІКУ

1.1 Особливості часових профілів трафіку комп’ютерних мереж

Часовий профіль мережевого трафіку – це відображення зміни обсягу або інтенсивності передавання даних у комп’ютерній мережі в часі. Він є результатом складної взаємодії користувачів, серверів, протоколів передавання, службових процесів і зовнішніх чинників. Вивчення часових характеристик трафіку дає змогу виявляти закономірності його зміни, прогнозувати навантаження, виявляти аномалії та підвищувати ефективність роботи мережевих систем.

Мережевий трафік формується як суперпозиція великої кількості незалежних і корельованих процесів, що відбуваються на різних часових масштабах. До таких процесів належать: короткочасні події (наприклад, запити до веб-серверів або короткі піки активності), середньотривалі процеси (завантаження великих файлів, синхронізація баз даних, оновлення систем), а також довготривалі тренди, пов’язані з добовими або тижневими циклами користувацької активності. Сукупність цих процесів створює складну часову структуру сигналу трафіку, яку не можна адекватно описати простими статистичними моделями.

Однією з характерних особливостей часових профілів трафіку є нестаціонарність. Середнє значення, дисперсія та інші статистичні характеристики сигналу змінюються з часом, що зумовлено циклічною природою людської активності, технологічними змінами в мережі або зовнішніми подіями. Наприклад, у робочі години трафік зростає через активне використання мережевих ресурсів, тоді як уночі спостерігається спад. Подібна сезонність потребує використання адаптивних методів аналізу, здатних відрізняти природні коливання від справжніх аномалій.

Другою ключовою рисою є фрактальність та самоподібність трафіку. Дослідження показують, що спектральна щільність енергії трафіку зменшується за степеневим законом, а автокореляційна функція має повільний спад, що свідчить про

довготривалу пам’ять процесу. Це означає, що короткочасні події впливають на поведінку системи протягом тривалого періоду, а динаміка трафіку характеризується нелінійними зв’язками між його елементами. Такі властивості потребують застосування методів, здатних враховувати взаємозалежність даних у різних часових масштабах – наприклад, мультискейлових або модальних підходів.

Крім того, трафік має високу мінливість (burstiness) – тобто чергування періодів спокою з різкими сплесками активності. Подібна поведінка притаманна як користувацьким запитам, так і технічним процесам (оновлення, резервне копіювання, хмарні синхронізації). Раптові піки можуть бути наслідком нормальної роботи системи, але за певних умов вони стають індикаторами аномалій, таких як перевантаження каналів, атаки типу DDoS або помилки у маршрутизації. Для їхньої коректної ідентифікації необхідно розрізняти регулярні та нетипові відхилення, що потребує використання математичних моделей з адаптивними порогами і робастними статистичними критеріями.

Ще однією особливістю є багатоканальність і просторово-часова кореляція даних. У сучасних комп’ютерних мережах трафік передається через десятки або сотні вузлів, кожен з яких має власний профіль навантаження. Проте між цими каналами існують тісні взаємозв’язки: зміна навантаження на одному вузлі може викликати реакцію в інших. Таким чином, часові профілі окремих вузлів не є незалежними, і для їхнього аналізу доцільно використовувати векторно-матричні моделі, що описують стан усієї системи у кожен момент часу. Це створює підґрунтя для використання методів багатовимірного аналізу, таких як PCA, ICA або Dynamic Mode Decomposition (DMD).

Важливо також враховувати вплив стохастичних факторів. До шумових складових трафіку належать випадкові флуктуації швидкості передавання пакетів, затримки у буферах, повторні запити та технічні збої. Вони утворюють шумову компоненту сигналу, що може маскувати аномалії. Тому методи аналізу трафіку повинні забезпечувати стійкість до шуму та флуктуацій, зберігаючи чутливість до реальних відхилень.

Таким чином, часові профілі трафіку комп’ютерних мереж характеризуються такими основними особливостями:

– нестаціонарністю **(**статистичні характеристики сигналу змінюються у часі);

– фрактальною структурою та довготривалою пам’яттю процесу;

– високою мінливістю і наявністю сплесків (burstiness);

– багатовимірністю і взаємозалежністю між каналами;

– наявністю стохастичних шумів і випадкових збурень;

– циклічністю добового та тижневого характеру навантаження.

Усі перелічені властивості ускладнюють задачу виявлення аномалій у трафіку, оскільки потребують поєднання методів, здатних одночасно враховувати як короткочасні, так і довготривалі процеси. Саме тому подальші розділи роботи присвячено побудові математичної моделі трафіку та розробці методу виявлення аномалій, який спирається на динамічний модальний аналіз і мультискейлову обробку часових даних.

1.2 Класифікація методів виявлення аномалій

Системи виявлення аномалій класифікуються за різними ознаками – способом навчання, типом використовуваних даних, підходом до моделювання та часовою природою аналізу [4-6].

Класифікація методів виявлення аномалій охоплює широкий спектр підходів, від простих статистичних до інтелектуальних моделей машинного навчання [6; 7]

На основі аналізу літератури можна виділити три головні класи (рис. 1.1):

1. Методи на основі моделювання статистики трафіку.

Застосовують оцінки середнього, дисперсії або розподілу даних.

Наприклад, Ye та Emran [8] використовували контрольні карти Шухарта для детекції змін параметрів у часових рядах.

2. Методи на основі аналітичного опису динаміки процесів**.**

Сюди належать фрактальні, спектральні, стохастичні моделі (ARIMA, ARFIMA), які відображають внутрішню структуру коливань [9; 10].

3. Методи машинного та глибокого навчання.

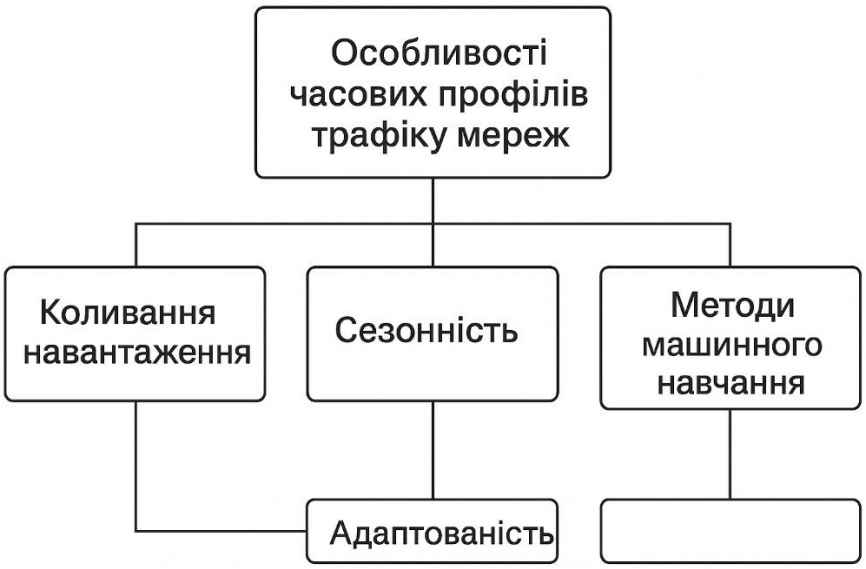


Рис.1.1. Головні класи методів виявлення аномалій

Їх розвиток зумовлений наявністю великих обсягів мережевих даних (NetFlow, CICIDS2017). За дослідженнями Ahmed et al. [11], ці методи демонструють найвищу гнучкість при обробці нестаціонарних потоків.

Узагальнено, класифікацію методів можна представити як ієрархію (рис. 1.2), де нижчі рівні забезпечують аналітичну інтерпретацію, а вищі — адаптивність і самонавчання.

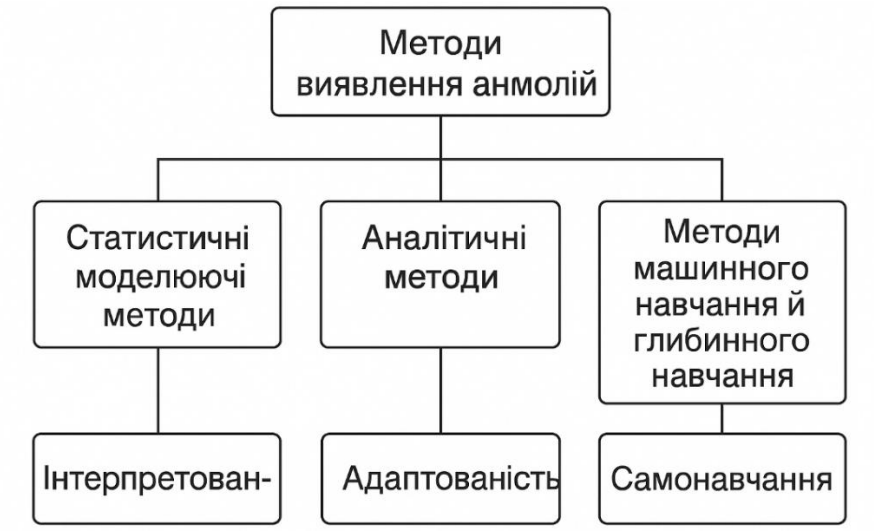


Рис.1.2. Класифікація методів

Ефективне виявлення аномалій у трафіку комп'ютерних мереж є одним із найважливіших завдань у сфері мережевого моніторингу, кібербезпеки та управління якістю обслуговування (QoS).

Під аномалією розуміють таку поведінку трафіку, що істотно відхиляється від нормального режиму функціонування мережі — наприклад, різке зростання кількості пакетів, сплеск затримок або зміну статистики потоків, викликану атаками типу DDoS, скануванням портів чи несправностями обладнання [1, 2].

Математично процес виявлення аномалій у часових профілях мережевого трафіку можна подати у вигляді перевірки гіпотези:

$$\begin{cases} H_0 : x_t \in \Omega_N, & (нормальний\ трафік) \\ H_1 : x_t \notin \Omega_N, & (аномалія) \end{cases}. \qquad (1.1)$$

де x_t – вектор характеристик потоку у момент часу t;

Ω_N – множина типових (нормальних) станів системи [3].

Для виявлення таких відхилень розроблено велику кількість методів, які за принципом побудови, типом навчання та природою аналізованих даних можна поділити на кілька основних груп (рис. 1.1): статистичні, методи машинного

навчання, спектрально-сигнальні, фрактальні (мультискейлові) та евристичні (сигнатурні) [4–6].

1.2.1 Статистичні методи аналізу мережевого трафіку

Статистичні методи є базовими інструментами для аналізу часових профілів трафіку. Вони передбачають побудову статистичної моделі нормального потоку даних і виявлення відхилень від неї.

Найпростіший випадок – параметрична модель, у якій припускається, що трафік x_t підкоряється нормальному розподілу:

$$p(x_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_t - \mu)}{2\sigma^2} \right], \tag{1.2}$$

де μ та σ – оцінки середнього значення та стандартного відхилення об’єму трафіку за певний інтервал [7].

Аномалія фіксується, якщо:

$$|x_t - \mu| > k\sigma, \tag{1.3}$$

де k – коефіцієнт (зазвичай k=3).

Такий підхід ефективний для стабільних каналів, але чутливий до нестаціонарності.

Для складніших сценаріїв використовують непараметричні моделі, зокрема ядрову оцінку щільності (KDE):

$$\hat{p}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h} \right), \tag{1.4}$$

де $K(\cdot)$ – ядро (зазвичай гаусове), h — ширина вікна, x_i — виміряні об’єми трафіку.

Якщо оцінена щільність $\hat{p}(x)$ падає нижче порогу τ, точка вважається аномальною [8].

Для часових рядів трафіку застосовуються авторегресійні моделі (AR, ARIMA):

$$x_t = \sum_{i=1}^{p} a_i x_{t-i} + \varepsilon_t , \tag{1.5}$$

де ε_t – випадкова складова.

Величина залишку $r_t=x_t-x\hat{}_t$ аналізується за статистичними критеріями, наприклад:

$$|r_t| > \theta \Rightarrow аномалія , \tag{1.6}$$

Ці методи широко застосовуються в системах IDS/IPS першого покоління, що контролюють трафік маршрутизаторів і фаєрволів [9].

1.2.2 Методи машинного навчання для виявлення мережевих аномалій

Машинне навчання забезпечує адаптивне виявлення аномалій у динамічному середовищі мережі, де статистичні моделі часто не відображають складної поведінки трафіку.

1.2.2.1 Методи з учителем (supervised learning)

Використовують розмічені набори даних, де класи «норма» та «аномалія» відомі.

Класичним прикладом є метод SVM (опорні вектори), який шукає гіперплощину, що відділяє нормальні та аномальні стани:

$$f(x) = sign(w^T x + b), \tag{1.7}$$

де w – ваговий вектор, b – зміщення [10].

Такі моделі використовуються для аналізу ознак мережевих пакетів (кількість TCP SYN, IP-адреси, розмір пакета тощо).

1.2.2.2 Методи без учителя (unsupervised learning)

Застосовуються, коли навчальні вибірки без розмітки. Найпоширеніші – кластеризаційні підходи. Наприклад, у методі k-середніх кожна точка трафіку відноситься до найближчого центру кластера μ_j:

$$d_j = \left\| x_i - \mu_j \right\|, \tag{1.8}$$

а точки, що не належать жодному щільному кластеру, розглядаються як аномалії [11].

Інші варіанти – Isolation Forest**,** DBSCAN або Autoencoder для глибокої репрезентації даних.

1.2.2.3 Напівконтрольовані (semi-supervised)

Використовують лише нормальні дані.

Наприклад, One-Class SVM шукає гіперплощину, яка відокремлює всі нормальні точки від початку координат:

$$\min_{w,\rho,\xi_i} \frac{1}{2}\|w\|^2 + \frac{1}{\nu n}\sum \xi_i - \rho, \tag{1.9}$$

де ν — частка допустимих аномалій [12].

1.2.2.4 Глибинне навчання

У мережевому аналізі часто застосовуються рекурентні (LSTM) або згорткові (CNN) нейронні мережі.

Рекурентна структура описується виразом:

$$h_t = \sigma\left(W_h h_{t-1} + W_x x_t + b\right), \tag{1.10}$$

де h_t – прихований стан, x_t – вектор ознак трафіку,

W_h, W_x – вагові матриці.

Такі моделі ефективні для потоків з часовими залежностями, наприклад, при аналізі добових циклів або змін навантаження [13].

1.2.3 Спектральні та сигнальні методи

Сигнальні підходи трактують трафік як часовий сигнал, що дозволяє виявляти приховані частотні закономірності.

Дискретне перетворення Фур'є (DFT)**:**

$$X(f) = \sum_{t=0}^{N-1} x(t) e^{-j 2\pi f t / N}, \tag{1.11}$$

дає змогу аналізувати частотний спектр трафіку й виявляти аномалії у певних діапазонах частот, наприклад, періодичні атаки або повторювані збурення [14].

Вейвлет-перетворення**:**

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int x(t) \psi^* \left(\frac{t-b}{a} \right), \tag{1.12}$$

дозволяє аналізувати зміни інтенсивності трафіку одночасно у часі й частоті.

Більш сучасним інструментом є динамічний модальний розклад (Dynamic Mode Decomposition, DMD**)** [15]:

$$X_{k+1} \approx A X_k, \quad A = Y X^{+} \tag{1.13}$$

де A — оператор динаміки системи, $X^{+}X^{\wedge}+X^{+}$ — псевдообернена матриця.

DMD дозволяє виділити моди, що відповідають характерним коливанням трафіку, і фіксувати появу нових мод як аномалій (наприклад, під час змін маршрутизації або вторгнень).

1.2.4 Фрактальні та мультискейлові підходи

Мережевий трафік має властивість самоподібності та довготривалих залежностей [16].

Для кількісної оцінки застосовується експонента Херста (Hurst exponent):

$$R/S \infty T^H, \tag{1.14}$$

де $H \in [0,1]$ при $H > 0{,}5$ процес демонструє довготривалу кореляцію.

Зміни значення H свідчать про появу аномалій у структурі трафіку – наприклад, збільшення інтенсивності коротких потоків при атаці SYN flood.

Мультифрактальний аналіз дозволяє описати складну ієрархічну структуру трафіку, де малі та великі пакети мають різну масштабну поведінку [17].

У таких моделях спектр мультифрактальності f(α) використовується як індикатор різних типів аномалій — наприклад, сплесків або зміни варіабельності потоків.

1.2.5 Евристичні та сигнатурні методи

Сигнатурні методи виявлення аномалій базуються на порівнянні поточного трафіку з еталонними шаблонами атак або подій [18].

Якщо функція подібності між поточним профілем $x(t)$ і сигнатурою s_j перевищує поріг:

$$\mathrm{sim}\left(x(t), s_j\right) > \eta, \tag{1.15}$$

то подія класифікується як аномалія.

Цей підхід реалізований у таких системах, як Snort**,** Suricata або Zeek.

Перевагою є висока точність для відомих атак, проте такі системи не виявляють невідомих або змінених сценаріїв.

1.3 Обмеження існуючих підходів і постановка задачі дослідження

1.3.1 Аналіз проблематики існуючих методів

Проведений у попередніх підрозділах аналіз показав, що сучасні методи виявлення аномалій у мережевому трафіку мають суттєві обмеження, які ускладнюють їх практичне застосування в умовах динамічних та масштабних мереж.

Кожен із підходів – статистичний, фрактальний, спектральний або машинного навчання – демонструє ефективність лише в обмеженому класі задач, тоді як у реальних системах аномалії мають комплексний, багаторівневий і нестаціонарний характер [1-3].

Статистичні моделі (наприклад, ARIMA, Гаусові розподіли, KDE) спираються на припущення про стаціонарність трафіку та відому форму розподілу.

Однак у реальних мережах:

– трафік має фрактальну структуру та важкохвостий розподіл [4];

– нормальні параметри трафіку змінюються в часі;

– просте порогове визначення $|x_t - \mu| > k_\sigma$ призводить до значної кількості хибних спрацювань.

Це робить класичні статистичні системи непридатними для довготривалого безперервного моніторингу, особливо в умовах великого обсягу даних.

Фрактальні моделі (на основі Hurst-експоненти або MFDFA) ефективно виявляють довготривалі закономірності, але:

– не розрізняють локальні аномалії короткої тривалості (наприклад, короткі атаки або тимчасові перевантаження);

– потребують великих обсягів даних для стабільної оцінки показників H або спектра $f(\alpha)$;

– складні у реалізації в реальному часі через високу обчислювальну складність.

Таким чином, ці підходи мають високу аналітичну цінність, але низьку придатність для оперативного моніторингу [5].

Спектральні методи дозволяють досліджувати частотну структуру трафіку, однак:

– у випадку нестаціонарних потоків результат сильно залежить від вибору вікна або масштабу *a*;

– при значній кількості потоків (мультиплексований трафік) методи потребують багатовимірного розкладу, що ускладнює інтерпретацію;

– аналіз частотних компонент не завжди дозволяє встановити час появи аномалії.

Тому для виявлення коротких імпульсних подій або різких змін спектральні методи потребують доповнення часовими моделями [6].

Попри високу гнучкість і здатність до самоадаптації, методи машинного навчання мають такі проблеми:

– потребують великих і збалансованих навчальних вибірок, у той час як реальні аномалії трапляються рідко (проблема «class imbalance»);

– результати залежать від якості попередньої обробки ознак;

– моделі часто є «чорними ящиками», що ускладнює інтерпретацію рішень (наприклад, у глибоких нейронних мережах);

– при роботі в реальному часі виникають складнощі з оновленням моделі без повного перенавчання [7].

Крім того, більшість ML-систем орієнтовані на розпізнавання відомих типів атак, а не на виявлення нових або прихованих аномалій у часовій структурі трафіку.

В табл. 1.1 наведено існуючі методи виявлення аномалій у часових профілях трафіку комп’ютерних мереж.

Таблиця 1.1

Узагальнююча порівняльна таблиця існуючих методів

Клас методів	Основна ідея	Переваги	Недоліки	Приклади
Статистичні	Аналіз відхилень від моделі нормального трафіку	Простота, інтерпретованість	Неадаптивність до змін	ARIMA, KDE
Машинне навчання	Навчання закономірностей у даних	Адаптивність, точність	Потреба в навчальних вибірках	SVM, Autoencoder
Спектральні	Аналіз частотної структури трафіку	Виявлення прихованих коливань	Складність параметризації	FFT, DMD
Фрактальні	Аналіз масштабних властивостей	Стійкість до шумів, глибоке розуміння структури	Висока складність	Hurst, MFDFA
Евристичні	Використання шаблонів або правил	Висока точність для відомих подій	Нездатність до нових типів атак	Snort, Suricata

1.3.2 Суперечності та наукова проблема

Узагальнення результатів огляду дозволяє сформулювати основну науково-технічну суперечність між:

– необхідністю високої чутливості та адаптивності при виявленні нових аномалій у трафіку;

– та потребою зберігати інтерпретованість, стійкість і обчислювальну ефективність методів для їх практичної реалізації в системах моніторингу.

Інакше кажучи, існує розрив між аналітичними методами, що описують динаміку трафіку (статистичні, фрактальні, спектральні), і навчальними системами (ML, DL), які здатні виявляти складні закономірності, але не забезпечують прозорості та фізичної інтерпретації процесів.

Це створює передумови для розробки гібридних методів, що поєднують математичні моделі часової динаміки з можливостями інтелектуального навчання,

орієнтованих на часові профілі трафіку — тобто послідовності агрегованих статистик за інтервалами часу.

1.4 Висновки до розділу 1

У розділі проведено системний аналіз сучасних методів і засобів виявлення аномалій у часових профілях трафіку комп’ютерних мереж. Показано, що мережевий трафік є складним багатокомпонентним процесом, який характеризується нестаціонарністю, високою мінливістю, фрактальністю, наявністю шумів і просторово-часовою кореляцією між каналами. Такі властивості істотно ускладнюють задачу виявлення аномалій та потребують методів, здатних одночасно враховувати як короткі піки активності, так і довготривалі зміни режимів роботи мережі.

У ході аналізу встановлено, що статистичні методи забезпечують простоту та інтерпретованість, але виявляють високу чутливість до нестаціонарності сигналу. Методи машинного навчання демонструють гнучкість, однак потребують великих навчальних вибірок та часто не забезпечують інтерпретованості. Спектральні та сигнальні підходи дозволяють визначати приховані частотні компоненти, проте недостатньо ефективні для локалізації коротких подій. Фрактальні та мультискейлові методи добре описують довготривалі закономірності, але мають значну обчислювальну складність і не завжди придатні для роботи в реальному часі. Евристичні сигнатурні системи точні для відомих атак, але не виявляють нових або модифікованих типів аномалій.

Узагальнення результатів аналізу показало наявність суперечності між вимогами до адаптивності, чутливості та здатності працювати з багатомасштабною структурою трафіку — та необхідністю зберігати інтерпретованість і обчислювальну ефективність. Це визначило наукову проблему дослідження та обґрунтувало потребу у розробці методу, який би поєднував можливості динамічного аналізу, мультискейлового представлення сигналу та робастних статистичних критеріїв.

РОЗДІЛ 2
МАТЕМАТИЧНА МОДЕЛЬ ТА МЕТОД ВИЯВЛЕННЯ АНОМАЛІЙ У ЧАСОВИХ ПРОФІЛЯХ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

2.1 Математична модель трафіку комп'ютерних мереж з елементами аномалій

2.1.1 Структура математичної моделі

Трафік розглядається як часовий сигнал $x(t)$, що є результатом суперпозиції кількох процесів:

$$x(t) = T(t) + L(t) + N(t) + A(t), \tag{2.1}$$

де $T(t)$ – повільний тренд (добова або багатогодинна динаміка навантаження);

$L(t)$ – локальні коливання середньої тривалості (типова користувацька активність);

$N(t)$ – випадковий шум, що відображає стохастичні збурення;

$A(t)$ – аномальні відхилення, які порушують нормальний режим мережі.

Таке декомпозиційне подання забезпечує модульність: кожна складова може бути проаналізована та параметризована окремо, що спрощує подальший аналіз методом DMD.

2.1.2 Локальні коливання

Локальні коливання описують помірно швидкі зміни трафіку, спричинені колективною активністю користувачів, оновленням даних, запитами до серверів тощо.

Вони мають квазісинусоїдальний характер із різними частотами й фазами.

Математично це подається у вигляді гармонічного ряду:

$$L(t) = \sum_{i=1}^{N} a_i \sin(2\pi f_i t + \varphi_i), \tag{2.2}$$

де a_i – амплітуда коливань iii-ї частоти;

f_i – частота зміни активності;

φ_i – випадкова фаза, що визначає часовий зсув;

N – кількість частотних компонент.

Ця частина формує середньочасову динаміку, наприклад, періодичні підвищення навантаження у робочі години чи короткі піки при обміні великими даними.

2.1.3 Повільні тренди

Повільна складова $T(t)$ описує детерміновану зміну середнього рівня трафіку протягом доби.

Типова мережа демонструє циклічну поведінку — низьке навантаження вночі та зростання у вечірній час.

Це можна відтворити періодичною функцією:

$$T(t) = A_0\left[1 + \alpha \sin\left(\frac{2\pi t}{T_d} - \varphi_0\right)\right], \tag{2.3}$$

де A_0 – середній рівень трафіку;

α – відносна амплітуда добових коливань;

T_d=12год – період (половина добового циклу, що моделює денну/вечірню зміну активності);

φ_0 – фазовий зсув.

Ця функція створює основну добову форму сигналу, до якої додаються швидші або випадкові коливання.

2.1.4 Випадковий шум

Шумова складова $N(t)$ відображає випадкові стохастичні процеси, що впливають на обсяг переданих даних.

Це можуть бути: коливання затримок пакетів, фонові запити систем, перезапити або технічні спотворення.

Вона моделюється як білий або кольоровий шум із нульовим середнім:

$$N(t) = \sigma_n \xi(t), \tag{2.4}$$

де σ_n – дисперсія шуму (визначає інтенсивність випадкових відхилень);

$\xi(t) \sim N(0,1)$ – стандартний гаусовий процес.

Іноді використовують корельований шум (наприклад, флікер-шум або процес Вінера), щоб краще відтворити тимчасову залежність мережевих флуктуацій.

2.1.5 Аномалії

Аномальна складова $A(t)$ відображає раптові відхилення, які порушують нормальний режим.

Це можуть бути атаки типу DDoS, перевантаження каналів, збої в обладнанні або піки при передаванні великих даних.

Такі події мають коротку тривалість і значну амплітуду, тому добре описуються через суму гаусових імпульсів:

$$A(t) = \sum_{k=1}^{K} b_k \exp\left(-\frac{(t - t_k)^2}{2\sigma_k^2} \right), \tag{2.5}$$

де K – кількість аномалій;

b_k – амплітуда (інтенсивність) k-го сплеску;

t_k – час виникнення події;

σ_k – ширина (тривалість) імпульсу.

У реальних мережах такі імпульси можуть мати нерівномірну амплітуду та з’являтись у непередбачувані моменти, що робить їх зручною основою для тестування алгоритмів виявлення аномалій.

2.1.6 Повна модель каналу

З урахуванням усіх складових (2.1)-(2.5) отримаємо повну математичну модель трафіку одного каналу:

$$x(t) = A_0\left[1+\alpha\sin\left(\frac{2\pi t}{T_d}-\varphi_0\right)\right]+\sum_{i=1}^{N}a_i\sin(2\pi f_i t+\varphi_i)+\sigma_n\xi(t)+\sum_{k=1}^{K}b_k\exp\left(-\frac{(t-t_k)}{2\sigma_k^2}\right), \quad (2.6)$$

Ця формула забезпечує комплексне представлення трафіку, що містить:

- регулярну зміну навантаження (тренд $T(t)$);
- псевдоперіодичні коливання (локальна динаміка $L(t)$);
- шумові збурення ($N(t)$);
- аномальні події ($A(t)$).

2.1.7 Мультиканальна структура

У реальних комп’ютерних мережах передача даних здійснюється по багатьох паралельних каналах або вузлах.

Тому повна модель може бути узагальнена у векторну форму:

$$X(t)=\begin{bmatrix}x_1(t)\\x_2(t)\\\dots\\x_M(t)\end{bmatrix}=T(t)+L(t)+N(t)+A(t), \quad (2.7)$$

де M – кількість каналів або вузлів, а кожна складова є вектором відповідних процесів.

Канали можуть бути як незалежними, так і корельованими між собою (наприклад, при спільному використанні мережевих ресурсів).

Таке представлення дозволяє застосовувати методи багатовимірного аналізу, включно з DMD, PCA, ICA або SVD, для виявлення узагальнених закономірностей у системі.

2.1.8 Просторово-часова інтерпретація

Векторна модель (2.7) дозволяє розглядати трафік не лише як часовий процес, а й як просторово-часову динамічну систему, де стан мережі у момент часу t описується множиною спостережень у різних вузлах.

Тоді зміни можна подати у вигляді матричного рівняння:

$$X_{t+1} = A_s X_t + E_t, \tag{2.8}$$

де A_s – матриця еволюції системи, що характеризує взаємозв'язки між каналами;

E_t – залишковий шум або аномалії.

Саме цей вираз використовується надалі для побудови DMD-розкладу і формування динамічної моделі зміни стану мережі.

2.1.9 Узагальнення та параметри моделі

Для параметризації моделі використовують такі групи параметрів, які наведено в табл.2.1.

Таблиця 2.1

Параметри математичної моделі

Позначення	**Фізичний зміст**	**Типове значення**
1	2	3
A_0	Середній рівень навантаження	1–10 МБ/с
α	Амплітуда добових змін	0.3–0.6

Продовж. табл. 2.1

1	2	3
T_d	Період тренду	12 год
a_i, f_i	Амплітуди та частоти локальних коливань	залежно від системи
σ_n	Інтенсивність шуму	0.05–0.2
b_k,σ_k	Параметри аномалій	випадкові
M	Кількість каналів	5–20

Запропонована математична модель трафіку комп'ютерних мереж (2.6)-(2.8) описує комплексну поведінку системи в часі та просторі, враховуючи регулярні, стохастичні й аномальні процеси.

2.2 Метод та програмний засіб виявлення аномалій у часових профілях трафіку комп'ютерних мереж

2.2.1 Вступ та постановка задачі

У сучасних комп'ютерних мережах трафік формується як складний часовий процес, що відображає взаємодію тисяч користувачів, служб і додатків. Такий процес характеризується високою розмірністю, нелінійністю та наявністю багатьох прихованих динамічних компонентів – регулярних циклів, трендів, шумових складових і випадкових відхилень.

Для ефективного аналізу подібних систем необхідні методи, які здатні виділяти домінантні режими динаміки без попереднього знання моделі процесу.

Одним із таких методів є динамічний модальний розклад (Dynamic Mode Decomposition, DMD), який дозволяє аналізувати складні часові ряди, подаючи їх у вигляді суперпозиції базових коливальних і експоненційних мод.

Перевага DMD полягає в тому, що він не потребує апріорних рівнянь динаміки, а використовує лише експериментальні дані.

У контексті комп’ютерних мереж цей метод дає змогу розділити трафік на типові режими поведінки (добові коливання, регулярні пікові періоди, фонова активність) і аномальні компоненти, які не описуються цими режимами.

2.2.2 Теоретичні основи динамічного модального розкладу

Динамічний модальний розклад (Dynamic Mode Decomposition, DMD) – це сучасний метод аналізу багатовимірних часових процесів, який дозволяє розкласти складну динаміку трафіку комп’ютерної мережі на набір незалежних модальних компонентів (мод), кожна з яких характеризується певною частотою, амплітудою та швидкістю зміни.

На відміну від класичних статистичних методів (таких як PCA чи ARIMA), DMD не просто шукає головні напрямки змін, а виявляє внутрішні часові закономірності, тобто динамічні моди, що описують еволюцію системи у часі.

Метод DMD походить із спектрального аналізу динамічних систем. Ідея полягає у тому, що спостережувані часові ряди можна представити як лінійну комбінацію мод, які розвиваються у часі з певними власними частотами та темпами зростання.

Якщо позначити через $y_t \in R_m$ вектор стану системи у момент часу t, то в локальному часовому вікні можна припустити, що існує лінійний оператор A, який описує еволюцію станів:

$$y_{t+1} = Ay_t\,, \tag{2.9}$$

У матричній формі для всього вікна довжини w:

$$X_2 = AX_1, \tag{2.10}$$

де

$$X_1 = [y_1, y_2, ..., y_{w-1}], \tag{2.11}$$

$$X_2 = [y_2, y_3, ..., y_w], \tag{2.12}$$

Оператор A можна знайти методом найменших квадратів:

$$A = X_2 X_1^+, \tag{2.13}$$

де X_1^+ – псевдообернена матриця Мура-Пенроуза.

Проте безпосереднє обчислення AAA є обчислювально складним і нестійким до шуму. Тому використовують сингулярний розклад (SVD) матриці X_1:

$$X_1 = U\Sigma V^T. \tag{2.14}$$

Далі матрицю зводять до зменшеного розміру, залишаючи лише перші r сингулярних значень, які несуть основну енергію сигналу:

$$\overline{A} = U_r^T X_2 V_r \Sigma_r^{-1}. \tag{2.15}$$

Власні значення λ_i матриці A~ визначають динаміку системи у вигляді комплексних експонент. Якщо $\lambda_i = \rho_i e^{j\theta_i}$ показує темп зростання або затухання, а θ_i – частоту коливання відповідної моди.

Таким чином, DMD дозволяє розділити динаміку сигналу на набір базових компонент, кожна з яких має чітку часову структуру.

2.2.3 Формування ознак для виявлення аномалій

Для аналізу часових профілів трафіку у кожному ковзному вікні обчислюються дві ключові характеристики:

1. Похибка реконструкції (reconstruction error) – показує, наскільки добре даний відрізок трафіку описується моделлю DMD:

$$E_{recom} = \frac{\left\| X_w - \hat{X}_w \right\|_F}{\sqrt{mw}}. \quad (2.16)$$

де Xw — матриця оригінальних даних у вікні, $\hat{X}_w$ – реконструкція, отримана з DMD, а $a\|\cdot\|_F$ – норма Фробеніуса.

Збільшення E_{recon} свідчить, що поточна динаміка відрізняється від звичайних закономірностей.

2. Зміна спектру власних значень (eigenvalue shift) – характеризує еволюцію динамічних мод у часі. Для двох сусідніх вікон обчислюється:

$$\Delta_\lambda = \sum_{k=1}^{K} \left\| \lambda_k^{(t)} \right| - \left| \lambda_k^{(t-1)} \right\|, \quad (2.17)$$

де порівнюються найбільш енергетично значущі моди. Якщо власні значення суттєво змінюються, це означає, що у трафіку з’явилися нові режими – наприклад, різкий підйом або тривале зростання.

Отже, для кожного моменту часу ми отримуємо два числових показники, які відображають короткочасну та структурну нестабільність у системі.

2.2.4 Мультискейловий аналіз

Реальний трафік комп’ютерних мереж є багатокомпонентним процесом, у якому одночасно взаємодіють процеси різних часових масштабів:

- короткочасні сплески (від секунд до хвилин),
- середньотривалі зміни (годинні коливання),
- довготривалі тренди (добові та тижневі варіації).

Кожен із цих масштабів несе власну динамічну інформацію — від коротких аномалій до глобальних змін у поведінці користувачів.

Тому для побудови ефективної системи моніторингу мережевого трафіку необхідно враховувати мультискейлову природу даних.

Мультискейловий аналіз у межах запропонованого методу базується на розділенні часових процесів на компоненти різної частотності та окремому застосуванні до них динамічного модального розкладу (DMD).

Це дозволяє виділити аномалії не лише у загальному потоці, а й у його окремих масштабних рівнях – наприклад, короткі локальні відхилення чи повільні структурні зміни.

2.2.4.1 Необхідність мультискейлового підходу

Реальний мережевий трафік має багаторівневу часову структуру, що включає:

– короткі події тривалістю кілька секунд (миттєві сплески трафіку, короткі атаки);

– середньотривалі коливання (зміни протягом хвилин або десятків хвилин, локальні перевантаження);

– повільні тренди (годинні або добові зміни активності користувачів).

Через це мережевий трафік не може бути повноцінно описаний одним часовим масштабом або одним набором параметрів моделі.

Для коректного відображення багатошарової структури сигналу застосовується мультискейловий динамічний аналіз, який дозволяє виділяти закономірності у кількох часових діапазонах одночасно.

2.2.4.2 Концепція мультискейлового підходу

Щоб охопити всі часові масштаби, динамічний модальний розклад (DMD) виконується у кількох часових вікнах різної довжини:

$$w_1, w_2, ..., w_S\,, \tag{2.18}$$

де w_1 – найкоротше вікно (для швидких процесів), а w_S – найдовше (для повільних трендів).

Наприклад, короткі вікна (кілька секунд) фіксують швидкі зміни трафіку, а довші (десятки хвилин або години) — описують повільні добові коливання.

Для кожного вікна сигнал подається у вигляді набору спостережень:

$$X_S = [x(t), x(t+\Delta t), ..., x(t+(n_S-1)\Delta t)], \quad (2.19)$$

де $s=1,\ldots,S$ – номер часової шкали;

Реальний мережевий трафік $x(t)$ є складним часовим процесом із багатошаровою структурою: короткі піки, середньотривалі зміни та довгі тренди. Щоб виявити аномалії, потрібно визначити закономірності розвитку станів системи у часі. DMD дозволяє побудувати лінійну апроксимацію еволюції сигналу у вигляді:

$$Y_S \approx A_s X_x, \quad (2.20)$$

де X_s – матриця станів у часовому вікні s, а Y_s – ті ж стани, зсунуті на один часовий крок.

Матриця A_s відображає локальні закономірності переходу мережевого стану від одного моменту до наступного.

Після обчислення A_s виконуються спектральний розклад та виділення мод:

$$A_s \phi_{i,s} = \lambda_{i,s} \phi_{i,s}, \quad (2.21)$$

де $\lambda_{i,s}$ — власні значення (характеризують темп зміни та частоту моди),

$\phi_{i,s}$ — власні вектори (описують просторову структуру моди).

У контексті мережевого трафіку:

– швидкі моди ($|\lambda_{i,s}|$ близько до 1, високі частоти) відповідають короткочасним пікам або атакам;

– повільні моди ($|\lambda_{i,s}|$ значно менше 1 або низькі частоти) відображають довготривалі тренди добового навантаження.

Таким чином, спектральні характеристики ($\lambda_{i,s}\phi_{i,s}$) дають розділення динаміки трафіку за часовими масштабами.

Переваги підходу:

– Локальна апроксимація: дозволяє лінійно моделювати сигнал навіть при глобальній нелінійності.

– Мультискейловість: окреме A_s для кожного вікна w_s дає змогу виділяти швидкі та повільні процеси.

– Інтерпретованість: власні значення і моди мають фізичне тлумачення для різних типів аномалій (піки, перевантаження, тренди).

– Основа для аномалій: зміни $\Delta\lambda^{(S)}(t)$ і залишки реконструкції $E_{recom}^{(S)}(t)$ використовуються для детекції відхилень.

DMD застосовується до кожного вікна s, щоб знайти локальну матрицю еволюції A_s, яка адекватно описує зміну станів мережевого трафіку на цьому часовому масштабі.

Власні значення $\lambda_{i,s}$ і власні вектори $\phi_{i,s}$ відображають структуру та швидкість динаміки трафіку, що дозволяє розділяти процеси за часовими масштабами і використовувати їх для виявлення аномалій.

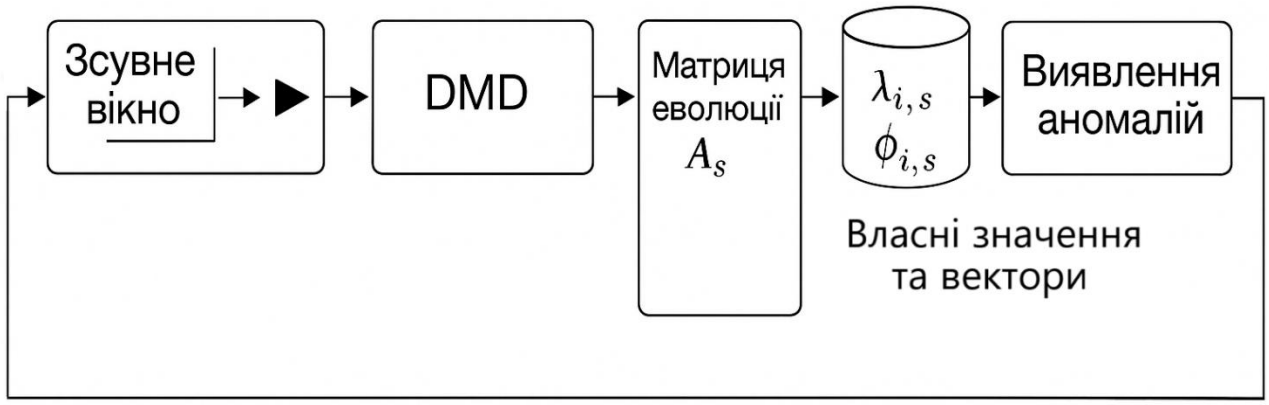


Рис. 2.1. Структура мультискейлового підходу

2.2.4.3 Енергетичний показник реконструкції

Для кожного масштабу s сигнал реконструюється за найбільш значущими DMD-модами:

$$x_s(t) = \sum_{i=1}^{r_s} b_{i,s}\phi_{i,s}e^{\omega_{i,s}t}\ , \tag{2.22}$$

де $\omega_{i,s} = \ln(\lambda_{i,s})/\Delta t$ – безперервна форма власних частот.

Відхилення між реальним сигналом і відновленим DMD-представленням визначає енергетичну помилку реконструкції:

$$E_{recom}^{(s)}(t) = \|x(t) - \hat{x}_s(t)\|. \tag{2.23}$$

Ця величина показує, наскільки поточна поведінка мережі відповідає типовій динаміці, виявленій у масштабі *s*.

Зростання $E_{recom}^{(s)}(t)$ є індикатором можливої аномалії.

2.2.4.4 Спектральний показник зміни динаміки

Окрім енергетичного відхилення, додаткову інформацію про нестандартну поведінку дає зміна спектра власних значень:

$$\Delta\lambda^{(s)}(t) = \|\lambda_{i,s}(t) - \lambda_{i,s}(t-\Delta t)\|. \tag{2.24}$$

Якщо частоти або темпи загасання мод змінюються різко, це вказує на порушення стабільності динамічної системи — наприклад, при раптовому зростанні навантаження або зміні маршрутизації в мережі.

Таким чином, $\Delta\lambda^{(s)}(t)$ відображає структурну зміну динаміки трафіку, а не просто відхилення амплітуд.

2.2.4.5 Формування мультискейлового ансамблевого показника

Щоб узагальнити інформацію з усіх часових масштабів, обидва показники (енергетичний та спектральний) нормалізуються за медіаною і MAD (робастним розкидом):

$$\tilde{E}^{(s)}(t) = \frac{E_{recom}^{(s)}(t) - median\left(E_{recom}^{(s)}\right)}{1.4826 \cdot MAD\left(E_{recom}^{(s)}\right)}, \quad (2.25)$$

$$\tilde{\Delta}^{(s)}(t) = \frac{\Delta\lambda^{(s)}(t) - median\left(\Delta\lambda^{(s)}\right)}{1.4826 \cdot MAD\left(\Delta\lambda^{(s)}\right)}. \quad (2.26)$$

Далі формується єдиний мультискейловий ансамблевий показник:

$$S(t) = \frac{1}{S}\sum_{s=1}^{S}\left[\left(\tilde{E}^{(s)}(t)\right)^2 + \left(\tilde{\Delta}^{(s)}(t)\right)^2\right]. \quad (2.27)$$

де $\tilde{E}$ і $\tilde{\Delta}$ ~ – нормалізовані значення, про що далі.

Він інтегрує інформацію про розбіжність між реальною динамікою трафіку та моделлю DMD у кожному часовому масштабі, а також про зміну спектральних характеристик у часі.

Такий підхід дозволяє методові бути чутливим одночасно до коротких і довготривалих відхилень у поведінці трафіку.

2.2.4.6 Інтерпретація у контексті мережевого трафіку

У контексті комп’ютерних мереж мультискейловий ансамблевий показник $S(t)$:

– реагує на короткі імпульсні події (різкі піки, перевантаження каналів) через збільшення $\tilde{E}^{(s)}(t)$ у коротких вікнах;

– відображає повільні зміни режиму (добові тренди, зміни навантаження) через варіацію $\tilde{\Delta}^{(s)}(t)$ у довгих вікнах;

– залишається робастним до шуму завдяки усередненню по масштабах.

Таким чином, $S(t)$ відображає цілісну картину поведінки мережі та дозволяє відокремити справжні аномалії від звичайних коливань.

Мультискейловий аналіз у межах динамічного модального розкладу (DMD) дозволяє охопити повний спектр часових процесів у трафіку комп’ютерних мереж.

Використання кількох часових вікон $w_1,w_2,...,w_S$ забезпечує можливість виявлення як швидких локальних аномалій, так і повільних змін тренду.

Двокомпонентний ансамблевий показник, що поєднує енергетичну та спектральну інформацію, формує стійку основу для подальшої робастної нормалізації та порогової детекції.

2.2.5 Робастна нормалізація та визначення порогу

У задачі виявлення аномалій у часових профілях трафіку комп'ютерних мереж важливу роль відіграє коректна нормалізація даних і вибір порогового значення для прийняття рішень.

Оскільки реальний мережевий трафік характеризується значною варіативністю, добовими коливаннями, шумами та окремими піковими сплесками, звичайні методи статистичної нормалізації (на основі середнього значення та стандартного відхилення) часто виявляються неефективними.

Вони надто чутливі до викидів, що призводить до помилкових спрацьовувань або, навпаки, до «пропуску» аномалій.

Щоб уникнути цього, у роботі використано робастний підхід до нормалізації, який базується на використанні медіани як оцінки центральної тенденції та медіанного абсолютного розмаху (MAD) як оцінки типових коливань.

Таке поєднання дає змогу усунути вплив одиничних великих відхилень у даних і стабілізувати масштаб усіх часових профілів незалежно від рівня навантаження мережі.

2.2.5.1 Необхідність нормалізації у часових профілях трафіку

Значення трафіку комп'ютерної мережі можуть відрізнятись на кілька порядків у різні періоди доби. Наприклад, у денні години трафік може бути стабільним, а у вечірні — зростати в кілька разів через пікову користувацьку активність.

Такі відмінності у масштабі призводять до того, що однакові за формою аномалії можуть мати різні амплітуди – і тому бути некоректно ідентифікованими.

Щоб уникнути цієї проблеми, необхідно виконати нормалізацію сигналу, тобто привести усі відрізки даних до спільного масштабу.

При цьому стандартні способи (наприклад, Z-нормалізація на основі середнього та стандартного відхилення) погано працюють у присутності викидів.

Тому застосовується робастна нормалізація, яка базується на медіані та MAD.

2.2.5.2 Робастна нормалізація даних

Нехай x(t) – часовий профіль трафіку для певного каналу мережі. Робастна нормалізація виконується за формулою:

$$x_{norm}(t) = \frac{x(t) - median(x)}{1.4826 \cdot MAD(x)}, \tag{2.28}$$

де $median(x)$ – медіана вибірки, що визначає центральне значення;

$MAD(x) = median(|x(t) - median(x)|)$ – медіанний абсолютний розмах, який характеризує типову варіацію;

коефіцієнт 1.48261 використовується для приведення MAD до масштабу стандартного відхилення при нормальному розподілі.

Завдяки такому перетворенню усі часові ряди отримують уніфікований масштаб, але при цьому зберігають форму аномальних ділянок.

Це перетворення робить усі часові профілі безрозмірними, центрованими навколо нуля та з типовим масштабом одиниці, але при цьому не спотворює форму аномалій.

Таким чином, навіть при значних викидах чи нерівномірності розподілу даних, робастна нормалізація забезпечує стабільність та порівнянність часових рядів.

2.2.5.3 Робастна оцінка залишкової енергії

Після нормалізації для кожного моменту часу обчислюється залишкова енергія $r(t)$, що характеризує відхилення між реальним профілем трафіку та його реконструкцією за допомогою динамічного модального розкладу (DMD):

$$r(t) = \left\| x_{norm}(t) - \hat{x}_{norm}(t) \right\|_2, \tag{2.29}$$

де $x_{norm}(t)$ – нормалізований реальний вектор трафіку;

$\hat{x}_{norm}(t)$ – його реконструкція за допомогою DMD.

Цей показник відображає ступінь відхилення поточного стану мережі від очікуваного.

Якщо система працює у нормальному режимі, це відхилення залишається малим; у випадку появи аномалії r(t) різко зростає.

Для багатоканальної мережі розраховується ансамблевий показник залишкової енергії:

$$S(t) = \frac{1}{m} \sum_{i=1}^{m} r_i(t), \tag{2.30}$$

де m – кількість каналів або сегментів мережі. Цей показник відображає узагальнений стан мережі в момент часу t.

2.2.5.4 Робастне визначення порогу детекції

Оскільки залишкова енергія є випадковою величиною, її порівняння з фіксованим порогом не завжди коректне.

Для забезпечення стійкості до шумів і поступових змін навантаження використано робастний механізм порогової детекції, у якому порогове значення визначається як:

$$\theta = median(S) + k \cdot 1.4826 \cdot MAD(S), \tag{2.31}$$

де $median(S)$ — центральна тенденція (типовий рівень залишкової енергії),

$MAD(\underline{S})$ — оцінка типових коливань без впливу викидів,

k – коефіцієнт чутливості, що визначає компроміс між кількістю хибних спрацьовувань і пропущених аномалій.

Рішення про наявність аномалії приймається за правилом:

$$\begin{cases} S(t) > \theta, & (аномалний\ стан); \\ S(t) \le \theta, & (нормальний\ стан). \end{cases} \quad (2.32)$$

Якщо ансамблевий показник $S(t)$ перевищує цей поріг, система інтерпретує поточний стан мережі як аномальний:

Такий підхід дозволяє автоматично адаптувати поріг до поточного стану мережі, не потребуючи зовнішнього налаштування.

Завдяки використанню медіани та MAD поріг не зміщується навіть у присутності сильних, але короткочасних викидів, що забезпечує робастність і стабільність процесу детекції.

Крім того, у динамічних середовищах можливо реалізувати ковзне оновлення порогу на основі останніх T спостережень, що дає змогу методу адаптуватися до змінного середнього навантаження мережі, не втрачаючи здатності реагувати на локальні аномалії.

2.2.5.5 Теоретичне обґрунтування робастного порогу

Вибір саме медіани і MAD обґрунтовується тим, що вони мають високий «breakdown point» – до 50%, тобто залишаються стабільними навіть якщо половина даних спотворена шумом або викидами.

Для порівняння:

- середнє зміщується навіть при 1-2% великих викидів;
- стандартне відхилення зростає пропорційно квадрату амплітуди викиду;
- медіана і MAD залишаються практично незмінними.

Тому така порогова модель є статистично стійкою і забезпечує надійне виявлення навіть у нестаціонарних середовищах.

2.2.5.6 Динамічне оновлення порогу

У процесі роботи системи поріг може адаптивно оновлюватися на основі ковзного вікна даних тривалістю T:

$$\theta_t = median(S_{t-T:t}) + k \cdot 1.4826 \cdot MAD(S_{t-T:t}). \quad (2.33)$$

Така адаптація дозволяє системі реагувати на повільні зміни добових чи тижневих трендів трафіку, не знижуючи чутливість до локальних аномалій.

2.2.5.7 Візуальна інтерпретація порогу

На графіку ансамблевого показника $S(t)$ порогова лінія θ слугує межею між «нормою» і «аномалією»:

– Точки, що лежать нижче θ, відповідають типовим коливанням трафіку.

– Точки, що перевищують θ, — фіксують нетипову поведінку мережі, що потребує аналізу.

У MATLAB це може бути реалізовано як побудова графіка $S(t)$ із горизонтальною лінією порогу та маркерами аномалій.

2.2.5.8 Переваги робастного підходу

Розроблений метод EDSD забезпечує виявлення аномалій у часових профілях трафіку шляхом поєднання локального спектрального аналізу (DMD) з мультискейловим ансамблевим підходом і робастною статистичною нормалізацією.

Базові переваги:

– Стійкість до шумів і пікових викидів: поріг не зміщується навіть за наявності одиничних сильних аномалій.

– Автоматичне масштабування: завдяки нормалізації, система працює однаково для вузлів із різним середнім рівнем трафіку.

– Адаптивність: може оновлюватись у реальному часі, реагуючи на зміну добового навантаження.

– Математична обґрунтованість: використання MAD і медіани відповідає принципам робастної статистики, рекомендованої для реальних технічних систем.

Робастна нормалізація та визначення порогу є ключовими елементами системи виявлення аномалій у часових профілях трафіку комп’ютерних мереж.

Вони забезпечують: інваріантність до масштабу сигналів; нечутливість до викидів; стабільність прийняття рішень при зміні режиму мережі.

У поєднанні з динамічним модальним розкладом (DMD) цей підхід формує стійку та адаптивну основу для інтелектуального моніторингу мережевого трафіку у реальному часі.

2.3 Імітація даних та добова мінливість трафіку

Для необхідності верифікації методу було розроблено імітаційну модель мережевого трафіку з урахуванням добової циклічності.

Базовий рівень трафіку задавався як гармонічний сигнал із випадковими збуреннями, який підсилюється у вечірні години та спадає вночі:

$$x(t) = A(1 + \alpha \sin(2\pi f t + \phi))\left(0.6 + 0.8\sin\left(\frac{2\pi t}{T_{day}} - \frac{\pi}{2}\right)\right) + \xi(t), \quad (2.34)$$

де T_{day} – тривалість умовного «дня», $\xi(t)$ – шум.

Додатково вводилися контрольовані аномалії:

– короткий сплеск (імітація DDoS);

– середньотривале зростання (flash crowd);

– повільна зміна частоти (зміна режиму мережі).

Таке моделювання дасть змогу перевірити, чи здатен метод відрізняти природну сезонність від справжніх відхилень.

2.4 Алгоритм виявлення аномалій у часових профілях трафіку комп’ютерних мереж

На рис.2.2 наведено алгоритм виявлення аномалій у часових профілях трафіку комп’ютерних мереж

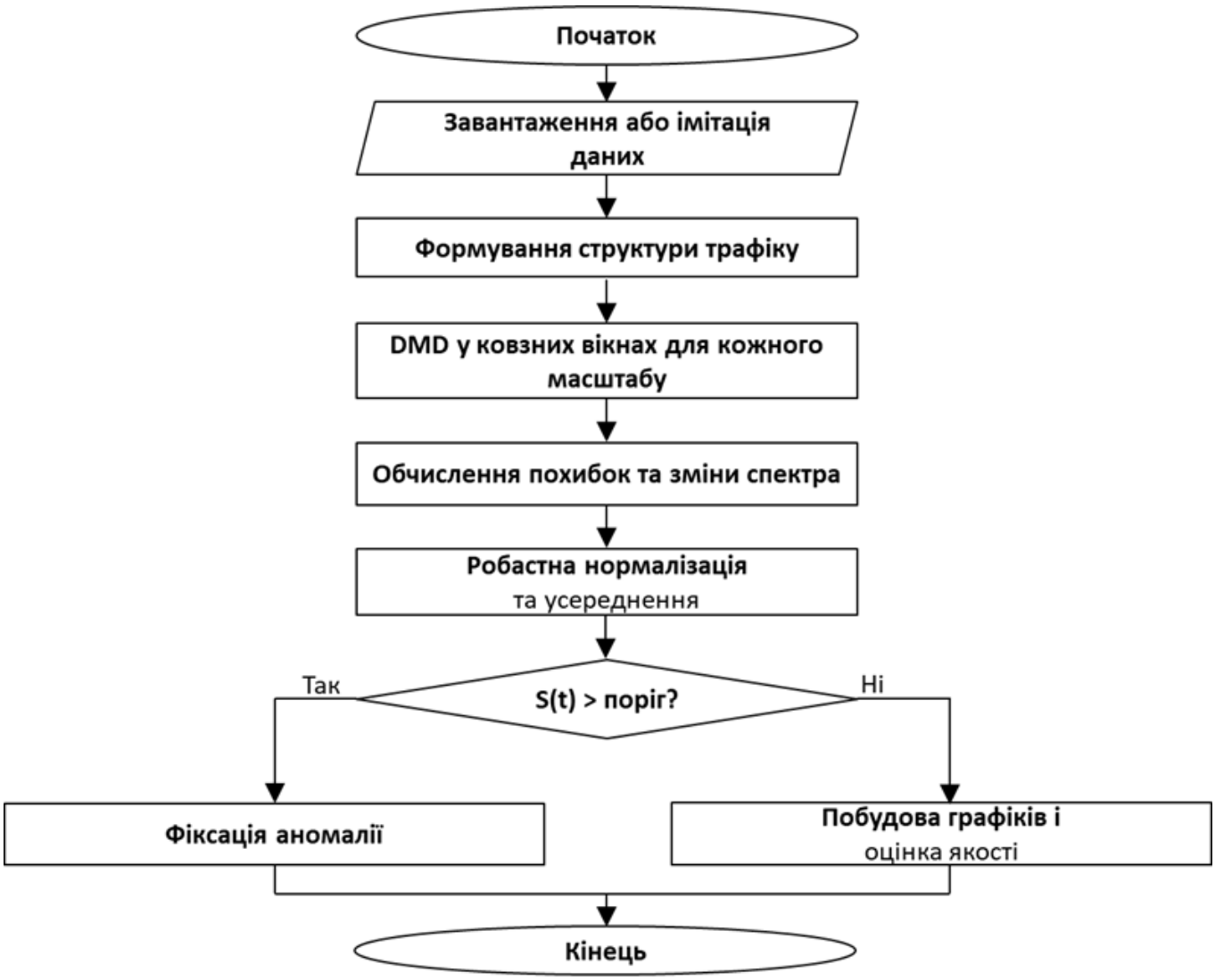


Рис.2.2. Алгоритм виявлення аномалій у часових профілях трафіку комп'ютерних мереж

Розроблений алгоритм призначений для виявлення аномальних процесів у часових профілях трафіку комп'ютерних мереж, тобто для автоматичного визначення моментів часу, коли навантаження в мережі істотно відхиляється від типової поведінки.

На початковому етапі роботи алгоритм виконує ініціалізацію основних параметрів. Зокрема, визначається:

– кількість каналів мережевого трафіку, які підлягають аналізу, загальна тривалість досліджуваного періоду;

– часові масштаби, у яких буде проводитись динамічний модальний аналіз, (довжини ковзних вікон та кроки часових зміщень вікна);

– параметри, що впливають на чутливість алгоритму — наприклад, коефіцієнт порогової детекції для визначення аномалій.

Така підготовка дозволяє адаптувати метод до різних умов функціонування мережі – від коротких тестових фрагментів до тривалого моніторингу добових циклів.

Далі алгоритм переходить до підготовки вхідних даних. У практичних умовах ці дані можуть надходити безпосередньо з мережевих пристроїв моніторингу у вигляді реальних вимірювань (2.19) або у вигляді імітації трафіку (формування багатоканальних часових рядів, у яких відтворюється типова добова мінливість трафіку – у денний час активність знижується, а у вечірній період спостерігається зростання навантаження).

Після підготовки даних алгоритм формує ковзні вікна різної довжини (2.18) з заданим кроком зсуву. Це дозволяє досліджувати трафік одночасно у декількох часових масштабах: коротких, що відображають швидкі коливання або миттєві піки навантаження, середніх — які фіксують періодичні зміни активності користувачів, та довгих, що відповідають добовим або тижневим трендам. Такий підхід забезпечує багатомасштабне охоплення процесів і робить метод стійким до різних типів аномалій.

У межах кожного часового масштабу до даних застосовується динамічний модальний розклад (Dynamic Mode Decomposition, DMD) (2.20). Цей математичний інструмент дозволяє описати часову динаміку мережевого трафіку як сукупність мод — базових структур, кожна з яких характеризує певну частоту, фазу та швидкість зміни процесу. Іншими словами, DMD перетворює часовий ряд на набір «динамічних компонент», які відображають стабільні закономірності в поведінці трафіку. Будь-які порушення цієї структури, наприклад раптові стрибки або появу нетипових коливань, проявляються у зміні спектральних характеристик мод.

Після виконання DMD для всіх часових вікон обчислюються два основні показники, які відображають стабільність динаміки: відхилення між відновленим сигналом і реальними даними (2.23), а також зміна спектра власних значень (2.24). Ці характеристики утворюють своєрідний «портрет поведінки» системи в часі, за яким можна судити про її нормальний або аномальний стан. Оскільки мережевий трафік

містить багато випадкових складових, далі застосовується робастна нормалізація результатів (2.25-2.26) – метод, який базується на медіанних статистиках і нечутливий до окремих викидів. Це дозволяє уникнути помилкових спрацьовувань при наявності шуму або короткочасних нестійких подій.

Наступним етапом є узагальнення результатів з різних часових масштабів. Оскільки кожен масштаб характеризує свій рівень динаміки – від швидких до повільних змін, їх поєднання забезпечує цілісне уявлення про стан системи. Отримані ознаки усереднюються, утворюючи інтегральний показник (2.27), який відображає сумарну стабільність трафіку у поточний момент часу. У моменти, коли цей показник різко зростає, алгоритм інтерпретує це як потенційний прояв аномалії.

Для прийняття остаточного рішення застосовується порогова детекція. Порогове значення визначається автоматично на основі статистичних характеристик показника стабільності (2.33). Якщо у певний момент часу інтегральний показник перевищує це порогове значення, система вважає, що відбулася аномальна подія (2.32). Таке рішення дозволяє гнучко реагувати на зміни трафіку, не потребуючи попереднього навчання або визначення фіксованих порогів вручну.

У фінальній частині алгоритму результати візуалізуються у вигляді графіків, де відображаються вихідні дані, відновлений DMD-сигнал та виявлені аномальні ділянки.

Додатково обчислюються метрики якості, такі як точність і повнота, які дозволяють оцінити ефективність методу у виявленні справжніх аномалій і визначити його стійкість до хибних спрацьовувань (виконується порівняння з виявленими аномаліями за метриками):

$$\text{Precision}=TP/(TP+FP),\ \text{Recall}=TP/(TP+FN),$$

де TP – кількість правильно виявлених аномалій,

FP — помилкові спрацьовування,

FN — пропущені події.

MATLAB надає змогу будувати графіки для кожного каналу трафіку окремо або в інтегрованому вигляді, що полегшує інтерпретацію результатів.

Таким чином, запропонований алгоритм є комплексним підходом до аналізу часових профілів мережевого трафіку. Він поєднує у собі мережевий трафік, мультискейловий підхід до обробки сигналів, динамічний модальний аналіз, робастну нормалізацію та статистичну порогову детекцію. Така архітектура дозволяє забезпечити високу чутливість до різних типів аномалій – як коротких імпульсних сплесків, так і довготривалих відхилень тренду. Реалізація в MATLAB забезпечує модульність, розширюваність і можливість оперативної адаптації алгоритму до реальних умов моніторингу комп'ютерних мереж.

Переваги реалізованого алгоритму:

– Враховує багатомасштабну природу трафіку (короткі, середні, довгі часові цикли);
– Є робастним до шуму та одиничних викидів;
– Може працювати в реальному часі з поточними потоками даних;
– Легко масштабувати для великої кількості каналів;
– MATLAB забезпечує наочну візуалізацію процесів виявлення.

2.5 Висновки до розділу 2

У розділі розроблено та обґрунтовано математичну модель трафіку комп'ютерних мереж, яка описує його як сукупність повільних трендів, локальних коливань, шумових компонентів та аномальних подій. Модель відображає багатомасштабну природу мережевих процесів і враховує як часову, так і просторову структуру багатоканальних потоків, що забезпечує реалістичність подання та придатність до подальшої обробки.

На основі цієї моделі сформульовано метод виявлення аномалій, що ґрунтується на використанні динамічного модального розкладу (DMD) та його мультискейлового узагальнення. Показано, що застосування DMD у кількох часових вікнах різної довжини дає змогу виділяти закономірності у декількох діапазонах —

від короткочасних сплесків до повільних структурних змін. Для оцінювання відхилень використано два ключові показники: енергетичний (похибка реконструкції) та спектральний (зміни власних значень), що разом формують мультискейловий ансамблевий індикатор.

Важливою складовою методу є робастна нормалізація та адаптивне визначення порогу детекції, що дозволяє зменшити вплив шумів, сезонності та випадкових викидів. Розглянуто робастні оцінки енергії, статистичні критерії вибору порогу, а також механізми його динамічного оновлення відповідно до змін трафіку.

Крім того, побудовано імітаційну модель мережевого трафіку з урахуванням добової циклічності, шумових компонентів і контрольованих аномалій, що дала змогу провести первинну верифікацію адекватності методу. На основі розроблених компонентів сформовано узагальнений алгоритм мультискейлового виявлення аномалій, який охоплює етапи підготовки даних, модального аналізу, обчислення показників, нормалізації, порогової детекції та формування інтервалів аномалій.

Таким чином, у розділі сформовано повний теоретичний та методологічний фундамент для реалізації програмного засобу, що здійснює автоматизоване виявлення аномалій у часових профілях трафіку комп’ютерних мереж. Отримані результати створюють основу для практичної реалізації та експериментальних досліджень.

РОЗДІЛ 3
ПРОГРАМНИЙ ЗАСІБ ТА РЕЗУЛЬТАТИ ВИЯВЛЕННЯ АНОМАЛІЙ У ЧАСОВИХ ПРОФІЛЯХ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

3.1 Архітектура програмного засобу

Архітектура розробленого програмного засобу для мультискейлового виявлення аномалій у мережевому трафіку ґрунтується на модульному принципі організації обчислювального процесу. Така структурованість забезпечує чітке розмежування функціональних блоків, кожен із яких виконує власне підзавдання, але при цьому є інтегрованою частиною єдиного обчислювального конвеєра. Загальну схему архітектури програмного засобу подано на рис. 3.1.

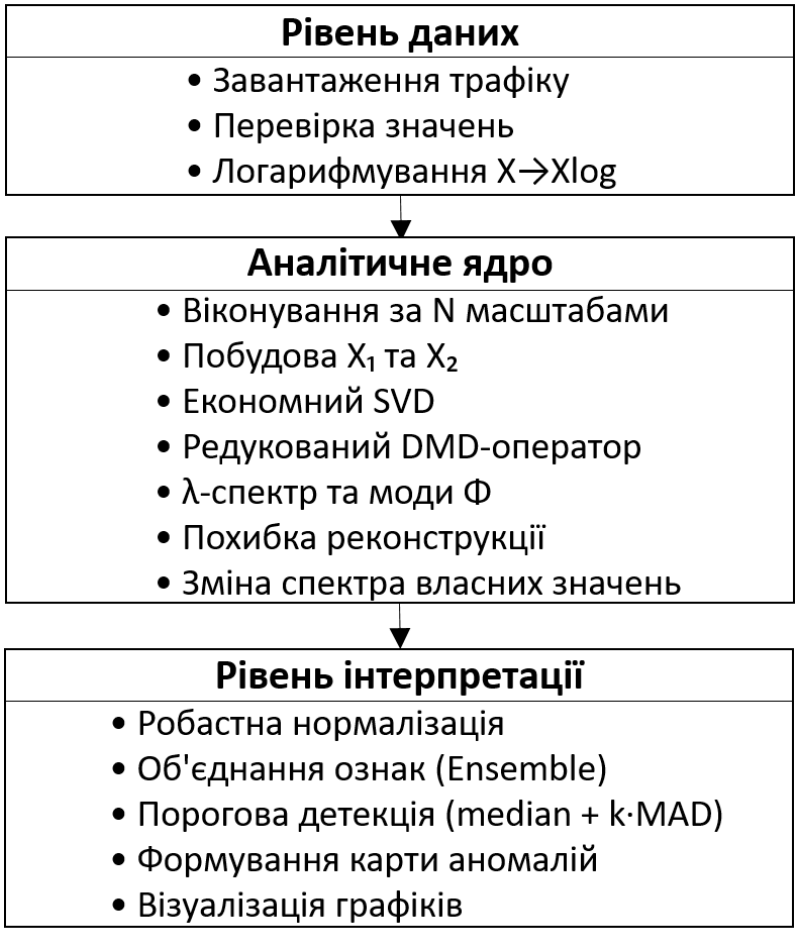


Рис. 3.1. Архітектура програмного засобу EDSD

На верхньому рівні система складається з трьох логічних шарів. Перший шар – рівень даних – відповідає за імпорт багатоканальних часових рядів, їхнє попереднє очищення та приведення до внутрішнього представлення. Саме тут відбувається перевірка коректності вхідних значень та застосування логарифмічного масштабування, яке стабілізує амплітуди та забезпечує чисельну стійкість подальших модальних обчислень.

Другий шар – аналітичне ядро, у межах якого реалізовано мультискейловий аналіз Ensemble DMD-Scale Detection. До цього шару входять модулі ковзного віконування, побудови матриць X1 та X2, виконання економного SVD, побудови редукованого DMD-оператора, обчислення спектра власних значень та формування модальної реконструкції. Для кожного масштабу обчислюються два індикатори: похибка реконструкції та зміна спектра власних значень. Результати з усіх масштабів узгоджуються та передаються на наступний рівень.

Третій шар – рівень інтерпретації результатів – виконує нормалізацію шкальних показників, їх вагове об'єднання в ансамблевий індикатор, порогову детекцію та формування остаточної бінарної оцінки аномальності. Завершальним компонентом є модуль візуалізації, який генерує графіки трафіку, шкальних показників, ансамблевого індикатора та часових інтервалів аномалій.

Така модульна архітектура дозволяє розвивати кожен компонент незалежно, масштабувати систему для нових сценаріїв аналізу, легко змінювати кількість масштабів, ранги DMD-декомпозиції, параметри порогу та механізми нормалізації. Завдяки цьому програмний засіб придатний як для експериментальних досліджень, так і для практичного розгортання у системах моніторингу трафіку.

3.2 Реалізація прогармного засобу в MATLAB

У цьому підрозділі представлено повну реалізацію програмного засобу для виявлення аномалій на основі мультискейлового методу Ensemble DMD-Scale Detection (EDSD), розробленого в середовищі MATLAB. Опис охоплює всі етапи – від ініціалізації й підготовки даних до формування оцінок аномалій та їхньої

візуалізації. Кожен фрагмент коду пояснюється у контексті логіки алгоритму, обраних параметрів, структури змінних та реалізованих математичних операцій.

Програмний засіб починається з очищення робочого простору, встановлення генератора випадкових чисел та завантаження багатоканального сигналу трафіку з файлу:

```
clearvars; close all; clc;
rng(2025);

%% ---------------- Load signal from file ----------------------
filename = 'trafic_anomal_2chanel.dat';
X = load(filename);  % канали по рядках
[m,T] = size(X);
t = 1:T;
time_hours = t / 500;    % 500 samples/hour → час у годинах
```

На цьому етапі формується змінна X, що містить трафік у вигляді матриці розмірності m × T, де m — кількість каналів, а T — довжина часової серії. Параметр «500 samples/hour» визначає переведення номерів відліків у години, що дозволяє зручно відображати графіки. Основні одиниці вимірювання також встановлюються на початку програми. Масштаби аналізу задаються масивом scales = [6, 18, 2000], де кожне значення відповідає довжині ковзного вікна для певного часового масштабу аналізу. Ранг DMD-розкладу визначається змінною r = 3.

Перед виконанням модального аналізу дані трафіку проходять попереднє перетворення у логарифмічну шкалу. Це є важливою процедурою, оскільки початкові значення мережевого трафіку можуть демонструвати різкі коливання, великі викиди або значну асиметрію розподілу. Такі характеристики ускладнюють подальший лінійний аналіз, зокрема процеси сингулярного розкладу та побудову матриці DMD-оператора. Логарифмування дозволяє знизити контрастність між великими і малими значеннями, забезпечити кращу чисельну стійкість та запобігти домінуванню поодиноких піків у подальшому модальному розкладі.

У реалізації використано функцію log1p, яка обчислює логарифм виразу 1 + X, що є чисельно стабільним способом опрацювання малих позитивних значень.

Попередньо на матрицю застосовується операція max(X,0), що гарантує відсутність від’ємних аргументів у логарифмі. Операція застосовується поелементно до всієї матриці трафіку:

```
%% ---------------- Log transform ----------------
Xlog = log1p(max(X,0));
```

Таке перетворення виконує одразу кілька функцій. По-перше, воно компенсує експоненціальну природу зростання типових мережевих навантажень, переводячи їх у простір, де відносні зміни стають більш вираженими, а абсолютні — менш домінуючими. По-друге, логарифмічна шкала зменшує варіативність даних, вирівнюючи динамічні діапазони між каналами, що сприяє більш коректному виділенню мод у процесі DMD. По-третє, процедура логарифмування відіграє роль «м’якої нормалізації», яка зменшує вплив одиничних пікових сплесків, що могли б призвести до переорієнтації сингулярних компонент на аномальні сегменти замість характерної фонової динаміки.

У контексті мультискейлового аналізу EDSD ця операція є особливо критичною, оскільки стабілізує структуру даних у всіх масштабах одночасно, забезпечуючи узгодженість модальних характеристик, що розраховуються для різних розмірів вікна. Логарифмічне перетворення, таким чином, виступає фундаментальним кроком, що підвищує точність обчислення як модальної реконструкції, так і спектральних характеристик, та зменшує ймовірність хибного збудження ознак аномальності через природну нерівномірність мережевого навантаження.

Подальший фрагмент – ядро методу EDSD. Для кожного масштабу формується ковзне вікно розміром w, всередині якого виконується повний DMD-аналіз. Оскільки кожне вікно має центр center і напівширину halfw, цикл перебирає всі позиції, у яких вікно повністю вміщається в дані. Для кожного вікна формується дві матриці: X1, що містить перші w-1 стовпців та X2 – ті ж стовпці зі зсувом на один крок вперед.

```
for s=1:num_scales
    w = scales(s);
    if w >= T
        warning('Scale %d >= T. Skipping this scale.', w);
        continue;
    end
    halfw = floor(w/2);

    for center = 1+halfw : T-halfw
        idx = (center-halfw):(center+halfw-1);
        Xw = Xlog(:,idx);
        X1 = Xw(:,1:end-1);
        X2 = Xw(:,2:end);
```

Для отримання модальної структури у вікні застосовується економний SVD-розклад. З нього залишаються тільки перші r сингулярних компонент, які містять найбільш енергетично значущу частину динаміки:

```
[U,S,V] = svd(X1,'econ');
rr = min([r size(S,1)]);
U_r = U(:,1:rr); S_r = S(1:rr,1:rr); V_r = V(:,1:rr);
```

Побудова редукованого оператора A_tilde виконується за класичною формулою DMD. На власних значеннях цього оператора ґрунтується спектральний аналіз, а власні вектори визначають моди. Розрахунок мод Phi використовує аналітичну формулу проєкції назад у вихідний простір даних.

```
A_tilde = U_r' * X2 * V_r / S_r;
[W_eigs,D] = eig(A_tilde);
lambda = diag(D);
Phi = X2 * V_r / S_r * W_eigs;
```

Далі оцінюється амплітудний вектор b, що визначає участь кожної моди в реконструкції. Залежно від обумовленості матриці мод використовується стандартне або псевдообернене розв’язання. Це робить алгоритм стійким до погано обумовлених випадків:

```
if cond(Phi) < 1e12
    b = Phi \ Xw(:,1);
else
    b = pinv(Phi) * Xw(:,1);
end
```

Часова динаміка мод відтворюється через піднесення власних значень до степеня, що відповідає індексу часу. На цьому етапі формується реконструйований сигнал X_recon, який згодом використовується для розрахунку енергетичної похибки:

```
time_dynamics = zeros(rr, w);
for k2=1:rr
    time_dynamics(k2,:) = b(k2) * (lambda(k2).^(0:(w-1)));
end
X_recon = real(Phi * time_dynamics);
```

Енергетичний показник, що є ключовим у мультискейловій детекції, визначається як норма різниці між реальним та реконструйованим сигналами у вікні. Це показує, наскільки локальна динаміка відрізняється від звичайної модальної структури, властивої цьому масштабу:

```
recon_err = norm(Xw - X_recon, 'fro') / sqrt(numel(Xw));
```

Другий важливий компонент мультискейлового показника аномальності ґрунтується на відстеженні змін спектра власних значень DMD-оператора у сусідніх вікнах. Оскільки власні значення λ_k мод характеризують локальні динамічні властивості сигналу в межах певного масштабного вікна, їхня стабільність від вікна до вікна свідчить про регулярну, передбачувану поведінку процесу. Навпаки, різкі зміни як фази, так і модуля власних значень є показником порушення сталої динаміки — тобто появи аномалії.

Для кожного вікна обчислюється набір λ_k. Ці значення попередньо сортують за дійсною частиною, що забезпечує стабільність порівняння між вікнами та дозволяє

відстежувати еволюцію саме найенергетичніших мод. Сортування виконується командою:

```
cur_eig = sort(lambda,'ComparisonMethod','real');
```

У цьому контексті відбувається впорядкування власних значень за зростанням дійсної частини, а отже, вищі позиції у векторі відповідають тим модам, які мають найбільший вплив на динаміку процесу всередині вибраного масштабу. Ці моди є найбільш інформативними для детекції аномалій, тому порівняння спектрів фокусується саме на них.

Для поточного масштабу зберігається спектр з попереднього вікна (prev_eigs{s}), і після обчислення нового спектра виконується порівняння найбільш значущих власних значень. Оскільки кількість мод, які варто порівнювати, може бути меншою від загальної кількості доступних власних значень (особливо при малому ранзі), використовується параметр kcomp = min([3 numel(prev) numel(cur_eig)]), що обмежує порівняння трьома найважливішими модами. Саме вони найбільш чутливі до змін у структурі даних.

Зміна спектра обчислюється як сума модулів різниць між абсолютними значеннями відповідних власних значень у сусідніх вікнах:

```
eig_change = sum(abs(abs(cur_eig(end-kcomp+1:end)) - ...
                     abs(prev(end-kcomp+1:end))));
```

Ця операція виконує дві важливі функції. По-перше, використання абсолютних значень відокремлює аналіз амплітудної поведінки від фазових обертів, які можуть бути присутні у комплексних власних значеннях, але не завжди відображають зміну енергетичної структури процесу. По-друге, порівняння саме кінцевих елементів масиву (тобто найбільших за значенням) дозволяє сфокусуватися на модах з найбільшою динамічною значущістю. У разі, коли у даних виникає різкий перехід, збурення або нетипове зростання активності, ці моди зміщуються сильніше за інші.

Таким чином, показник eig_change забезпечує чутливий індикатор структурної нестабільності: великі значення цього параметра означають, що модальна структура локального фрагмента сигналу суттєво відрізнилася від фрагмента, який передував йому, і така зміна є потенційною ознакою аномалії.

У підсумку, показник зміни спектра доповнює показник енергетичної похибки реконструкції. Перший фіксує зміну структури динаміки, другий — зміну форми сигналу. Їхня комбінація у межах методу EDSD дає змогу виявляти як швидкі збурення, так і повільні процеси, що порушують звичний перебіг мережевого трафіку.

Після отримання двох шкалових характеристик (recon_err_scales і eig_changes_scales) слідує етап нормалізації та їхнього вагового об’єднання. Нормалізація виконується за робастною схемою «медіана + MAD», яка забезпечує стійкість до інтенсивних викидів та піків:

```
norm_v1 = (v1 - median(nz1)) ./ (1.4826 * mad(nz1,1) + eps);
norm_v2 = (v2 - median(nz2)) ./ (1.4826 * mad(nz2,1) + eps);
```

Масштабні коефіцієнти (recon_gain і eig_gain) забезпечують баланс між внеском коротких і довгих вікон. Після застосування цих коефіцієнтів формується шкальний показник:

```
score_per_scale(s,:) = ...
    recon_gain * max(0,norm_v1) + 0.9 * (eig_gain * max(0,norm_v2));
```

Ансамблевий показник — середнє усіх шкальних показників, додатково згладжене рухомою медіаною:

```
ensemble_score = mean(score_per_scale,1);
ensemble_score_smooth = movmedian(ensemble_score,5);
```

Порогова детекція в методі EDSD реалізується на основі робастного статистичного критерію «медіана + k·MAD», який забезпечує високу стійкість до викидів і водночас чутливість до стійких або інтенсивних аномалій. На відміну від

класичного правила «середнє ± n·σ», що припускає нормальний розподіл даних і сильно страждає від одиничних піків, робастні статистики медіани та MAD (Median Absolute Deviation) є нечутливими до великих одиничних збурень та забезпечують адекватне оцінювання центральної тенденції навіть за умов значних структурних змін у даних. Це робить їх особливо придатними для задач виявлення аномалій у мережевому трафіку, де присутні різкі, нерівномірні, а іноді й квазі-періодичні сплески.

Медіана ансамблевого показника обчислюється так:

```
med = median(ensemble_score_smooth);
```

Замість середнього використано саме медіану, оскільки вона визначає «центр» розподілу без впливу на нього нетипових великих значень, які у даному контексті часто й є аномальними та не повинні зміщувати базовий фон. Тому медіана формує базову лінію, відносно якої оцінюється відхилення сигналу.

Далі обчислюється значення MAD — медіана абсолютних відхилень від медіани. У MATLAB функція mad(x,1) повертає середнє абсолютне відхилення (Mean Absolute Deviation), тому множник 1.4826 виконує перетворення до рівняння зі «стандартним відхиленням» у випадку гаусового розподілу:

```
madval = 1.4826 * mad(ensemble_score_smooth,1);
```

Цей коефіцієнт має статистичне походження: при нормальному розподілі MAD × 1.4826 ≈ σ, тому отримане значення можна інтерпретувати як робастний аналог стандартного відхилення. Проте важливо, що на відміну від σ, MAD не реагує на рідкісні сильні піки, які в нашому випадку є саме об’єктом детекції, а не тим, що має впливати на параметри порогу.

Після обчислення медіани та робастного масштабу варіації задається робастний поріг за формулою:

```
thresh = med + k_thresh * madval;
```

Значення k_thresh = 35 підібране емпірично на основі аналізу великої кількості експериментів із різним трафіком. Воно визначає, скільки робастних «стандартних відхилень» потрібно перевищити, щоб сигнал був визнаний аномальним. Велике значення коефіцієнта пояснюється тим, що ансамблевий показник EDSD уже є агрегованим і згладженим, тому його фонові значення дуже низькі, а аномалії — різко виділяються. Тому використання малих множників (наприклад, 3 або 5) спричинило б надмірну чутливість і часті хибні спрацювання, тоді як значення на рівні 30–40 дає змогу відокремити лише ті фрагменти, де масштабні або багаторівневі зміни динаміки є справді суттєвими.

Такий спосіб порогування гарантує, що навіть якщо структура трафіку зміниться (наприклад, загальний рівень навантаження підвищиться), поріг адаптуватиметься до нової медіани та нового розсіювання, зберігаючи здатність виявляти тільки нетипові фрагменти. Це робить EDSD стійким до довготривалих трендів, епізодичних піків та змін у масштабах флуктуацій.

Після первинного порогування застосовуються морфологічні операції згладження та видалення коротких хибних фрагментів:

```
detected = ensemble_score_smooth > thresh;
detected = movmax(detected,15);
detected = movmedian(detected,9) > 0.5;
detected = bwareaopen(detected,20);
```

Завершальний етап роботи програмного засобу полягає у наочному поданні результатів аналізу та виявлених аномалій. Оскільки мультискейловий метод EDSD генерує декілька типів характеристик — вихідний трафік, шкальні оцінки, ансамблевий показник та бінарний індикатор аномалій — кожна з них відображається в окремому графічному вікні для забезпечення інтерпретованості та зручності аналізу.

Перед побудовою графіків встановлюється базовий розмір шрифтів, а також задається структура вікон:

```
fs = 13;  % font size

figure('Position',[100 100 1400 900]);
```

Перший графік демонструє реальні значення трафіку для кожного каналу. Для цього вибирається колірна палітра lines(m), яка гарантує розрізнюваність каналів. Власне побудова відбувається в циклі, де кожен канал відображається окремою кривою. Вивід виконується у фізичних одиницях (терабайти) та в часових позначках, перетворених у години:

```
% 1) Raw channels
figure(1)
subplot(1,1,1);
cols = lines(m); hold on;
for ch=1:m
    plot(time_hours,X(ch,:),'Color',cols(ch,:),'LineWidth',0.8);
end
ylabel(['Трафік (' volume_unit ')'], 'FontSize', fs);
xlabel(['Час (' time_unit ')'], 'FontSize', fs);
title('Завантажений трафік', 'FontSize', fs);
legend(arrayfun(@(i)sprintf('Канал %d',i),1:m,'UniformOutput',false), ...
    'Location','best', 'FontSize', fs);
grid on; axis tight
set(gca,'FontSize',fs)
```

Ця візуалізація дозволяє оцінити загальний характер змін у мережі, визначити корельовані інтервали між каналами та виявити потенційні піки, які можуть бути пов’язані з аномаліями. На цьому етапі дослідник отримує інтуїтивне уявлення про структуру трафіку, що є важливим для інтерпретації результатів подальшої обробки.

Другий графік відображає шкальні показники аномальності, обчислені на етапі мультискейлового модального аналізу. Вони подаються у форматі теплової мапи (heatmap), де кожен рядок відповідає певному часовому масштабу, а інтенсивність

кольору відображає величину показника. Для цього використовуються команди imagesc, axis xy, що забезпечують правильну орієнтацію матриці на графіку:

```
% 2) Scores by scale
figure(2)
subplot(1,1,1);
imagesc(time_hours,1:num_scales,score_per_scale); axis xy;
ylabel('Масштаб', 'FontSize', fs); xlabel(['Час (' time_unit ')'], ...
    'FontSize', fs);
title('Показники аномалій', 'FontSize', fs); colorbar; grid on; axis tight
set(gca,'FontSize',fs)
```

Теплова мапа дозволяє відразу побачити, у якому часовому масштабі відбулася найбільш різка зміна динаміки. Це є ключовим чинником мультискейлового аналізу: короткі аномалії яскравіше відображаються у малих масштабах, а повільні — у великих.

Третя візуалізація подає ансамблевий показник, який є середнім значенням по всіх шкалах після згладжування рухомою медіаною. Цей показник найбільш зручний для користувача, оскільки він представлений як одномірний часовий ряд і слугує інтегральною оцінкою аномальності:

```
% 3) Ensemble + threshold
figure(3)
subplot(1,1,1); hold on;
plot(time_hours, ensemble_score_smooth,'LineWidth',1.2);
yline(thresh,'--r','Поріг', 'FontSize', fs);
ylabel('Аномалійність', 'FontSize', fs); xlabel(['Час (' time_unit ')'], ..
    'FontSize', fs);
title('Ансамблевий показник', 'FontSize', fs); grid on; axis tight
set(gca,'FontSize',fs)
```

На цей графік накладається горизонтальна порогова лінія, обчислена за робастним правилом «медіана + k·MAD». Перетин ансамблевого показника з порогом вказує на потенційно аномальні інтервали, а сам графік дозволяє оцінити, наскільки суттєво аномальний процес виходить за рамки нормальної динаміки.

Четверта фігура відповідає за подання остаточної бінарної індикації аномалій разом із автоматично підсвіченими часовими інтервалами. Спершу відображається бінарний сигнал:

```
% 4) Detections + intervals
figure(4)
subplot(1,1,1); hold on;
plot(time_hours, detected, 'k','LineWidth',1.2);
```

Після цього виконується пошук зв’язних компонент за допомогою bwlabel, що дозволяє визначити межі кожної аномалії. Кожен знайдений інтервал підсвічується напівпрозорим прямокутником, а над ним розміщується текстова позначка з часом початку та завершення аномалії:

```
% HIGHLIGHT INTERVALS
[labeled, num_blobs] = bwlabel(detected);
for i=1:num_blobs
    idx = find(labeled==i);
    t_start = time_hours(idx(1));
    t_end   = time_hours(idx(end));

    fill([t_start t_end t_end t_start], [-0.05 -0.05 1.15 1.15], ...
         [1 0.9 0.4], 'FaceAlpha',0.35, 'EdgeColor','none');

    text((t_start+t_end)/2, 1.05, ...
        sprintf('Аномалія: %.2f–%.2f год', t_start, t_end), ...
        'HorizontalAlignment','center','FontSize', fs, 'Color','k');
end
```

Цей графік є підсумковим результатом роботи всього програмного засобу, оскільки він демонструє не лише момент виникнення аномалії, а й її часову тривалість у реальних одиницях. Крім того, використання прозорої заливки дозволяє оцінити тривалість і структуру аномалії у контексті всього індикаторного сигналу. Завершення візуалізації включає встановлення меж осей, сітки та загального заголовка для всієї групи графіків:

```
xlabel(['Час (' time_unit ')'], 'FontSize', fs);
ylabel('Аномалія', 'FontSize', fs);
title('Виявлені аномалії + часові інтервали', 'FontSize', fs);
grid on; axis tight; ylim([-0.1 1.2]); set(gca,'FontSize',fs)

sgtitle(sprintf('EDSD – Ensemble DMD-Scale Detection (%d каналів)', m), ...
    'FontSize', fs+2);
```

Таким чином, візуалізаційний блок програмного засобу завершує логічний цикл аналізу та дає можливість чітко інтерпретувати отримані показники. З одного боку, перший графік відображає реальний трафік, з другого – мультискейлова теплова карта та ансамблевий показник демонструють, як алгоритм реагує на відхилення, а четвертий графік чітко виділяє ті часові інтервали, які система класифікувала як аномальні. Це забезпечує повну прозорість процесу детекції, дозволяє аналізувати характер і природу порушень, а також оцінювати ефективність методу у практичному застосуванні.

3.3 Результати виявлення аномалій у часових профілях трафіку комп’ютерних мереж

Аналіз вхідних даних розпочинається з дослідження поведінки багатоканального трафіку, який виступає базовим джерелом інформації для подальшої мультискейлової обробки методом Ensemble DMD-Scale Detection (EDSD). У роботі використано два незалежні канали, що відображають типове навантаження комп’ютерної мережі протягом кількох годин.

На рис. 3.2 подано графічне зображення часових рядів цих каналів, яке дозволяє оцінити вихідні властивості сигналів, наявність трендів, флуктуацій та аномальних стрибків, що є критично важливими для коректного налаштування алгоритмів модального аналізу та подальшої детекції.

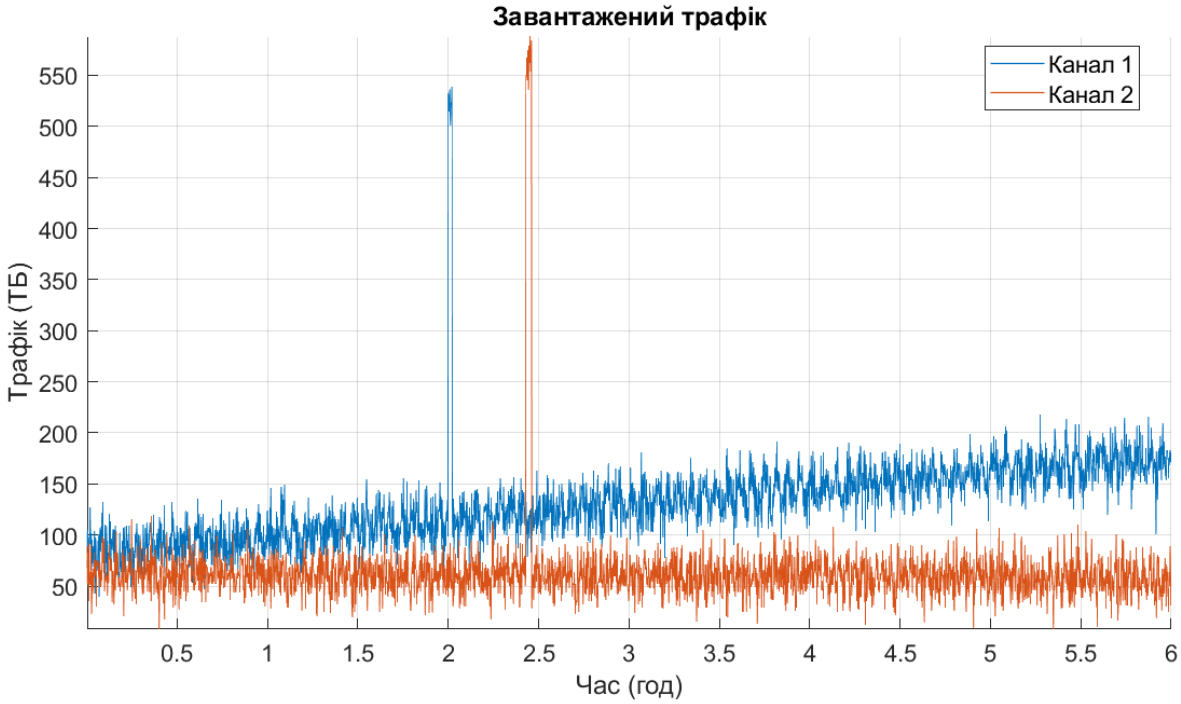


Рис.3.2. Реалізація завантаженого трафіку комп'ютерної мережі (канали – 2)

Як видно з рис.3.2, обидва канали демонструють характерну для мережевих систем добову нерівномірність, накладену на швидкі стохастичні коливання. Перший канал має вищий середній рівень навантаження, що поступово зростає з часом, тоді як другий канал характеризується нижчою інтенсивністю та більш стабільною амплітудою. Для обох каналів спостерігаються виражені аномальні імпульсні піки, які суттєво перевищують фоновий рівень та відображають нетипові події у мережі — раптові сплески трафіку або потенційні збої.

Наявність таких одиничних імпульсів у вихідному сигналі потребує застосування методів, здатних відокремити нормальну динаміку від локальних збурень. Саме тому попереднє візуальне дослідження є важливим етапом, що дозволяє оцінити масштаби варіацій, характер змін та типові особливості кожного каналу перед виконанням логарифмування, мультискейлового віконування, побудови мод та спектрального аналізу в межах алгоритму EDSD.

На рис.3.3 подано мультискейлову теплову карту показників аномалій, отриманих на основі локальних DMD-характеристик для кожного масштабу аналізу.

По осі абсцис відкладено час у годинах, а по осі ординат — індекси масштабів, що відповідають різній тривалості ковзних вікон.

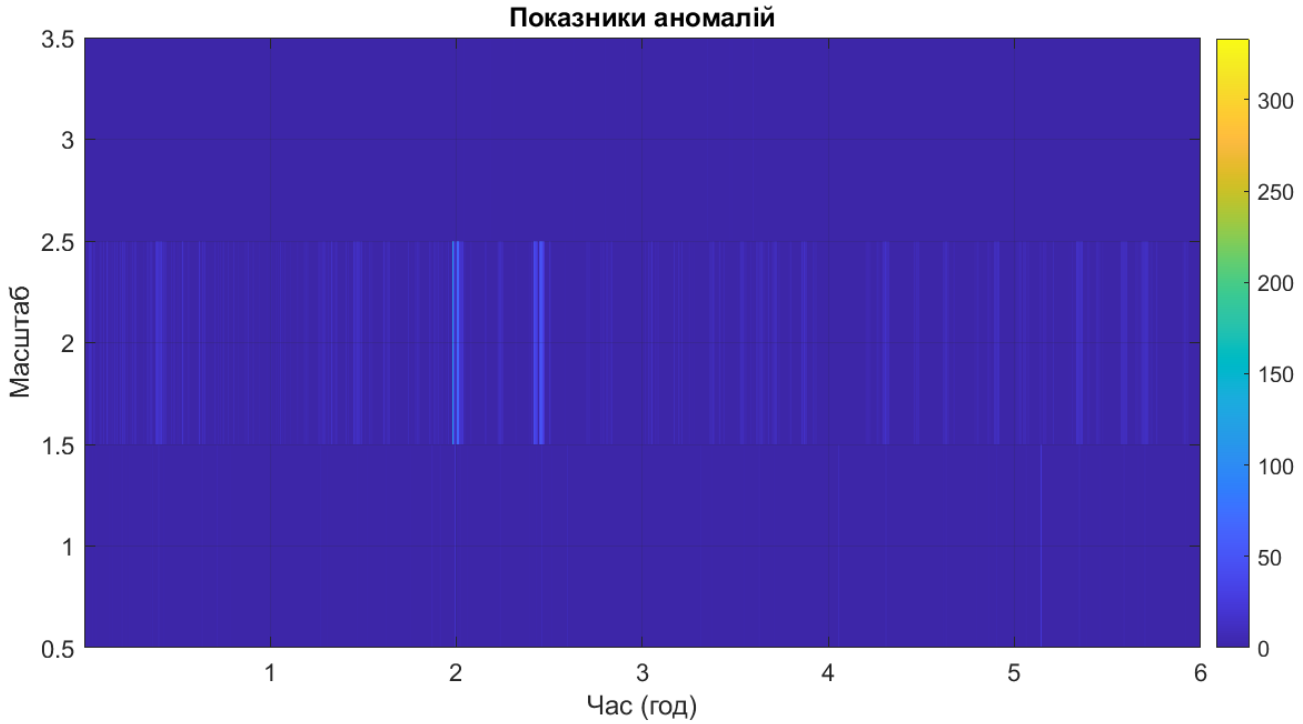


Рис.3.3. Показники аномалій

Яскравіші вертикальні смуги відповідають часовим інтервалам, де алгоритм зафіксував значні відхилення від нормальної поведінки трафіку. Інтенсивніші кольори позначають вищу аномальність, тоді як темно-синій відповідає нормальному режиму.

Такий формат подання дозволяє:

– побачити, на яких часових масштабах проявляється аномалія,

– оцінити її тривалість,

– визначити, чи є подія *локальною*, чи вона впливає на поведінку трафіку у ширших часових масштабах.

Ця теплова карта є ключовою частиною методу EDSD, оскільки дозволяє інтерпретувати поведінку трафіку багатомасштабно та коректно виділяти потенційні аномальні сегменти.

На рис.3.4 подано графік ансамблевого показника аномальності, що є результатом об'єднання всіх шкальних характеристик трафіку у єдиний індикатор.

Рис.3.4. Ансамблевий показник та поріг прийняття рішення

Крива ансамблевого показника (рис.3.4) демонструє, як змінюється інтегральна оцінка відхилення сигналу від нормального режиму у часі.

Пунктирною червоною лінією позначено робастний поріг прийняття рішення, побудований за статистичним правилом «медіана + k·MAD». Якщо значення ансамблевого індикатора перевищує цей поріг, відповідний часовий інтервал класифікується як потенційно аномальний.

На графіку чітко видно інтервали, де ансамблевий показник істотно зростає, що свідчить про наявність стійких відхилень у декількох часових масштабах одночасно. Це дозволяє значно точніше виявляти короткі пікові аномалії, а також довготривалі зміни тренду, порівняно з використанням одношкальних або класичних методів.

Таким чином, ансамблевий показник є ключовим елементом алгоритму EDSD, оскільки забезпечує збалансовану інтерпретацію поведінки трафіку на різних часових рівнях і формує основу для фінального рішення щодо детекції аномалії.

Після формування ансамблевого показника та визначення порогу прийняття рішення (рис.3.4) виконується безпосередня побудова бінарної маски аномальності. На цьому етапі кожному часовому моменту присвоюється значення 1 або 0 залежно

від того, перевищує ансамблевий показник робастний поріг «медіана + k·MAD». Після цього алгоритм групує суміжні одиничні значення у суцільні часові інтервали, що відповідають виявленим аномальним подіям.

Остаточна візуалізація має на меті чітко відобразити результати детекції, виділити виявлені інтервали на часовій осі та показати, як саме метод EDSD позначає ненормальну активність у трафіку. Така форма подання є важливою для інтерпретації результатів, оскільки дозволяє легко оцінити тривалість, початок і кінець кожної аномалії, а також перевірити коректність реакції алгоритму відносно поведінки сигналу.

На рис.3.5 наведено остаточні результати виявлення аномалій методом EDSD у часових профілях двоканального трафіку. По осі абсцис відкладається час у годинах, а по осі ординат — бінарний індикатор аномальності, де значення 1 відповідає виявленій аномалії, а значення 0 — нормальному режиму.

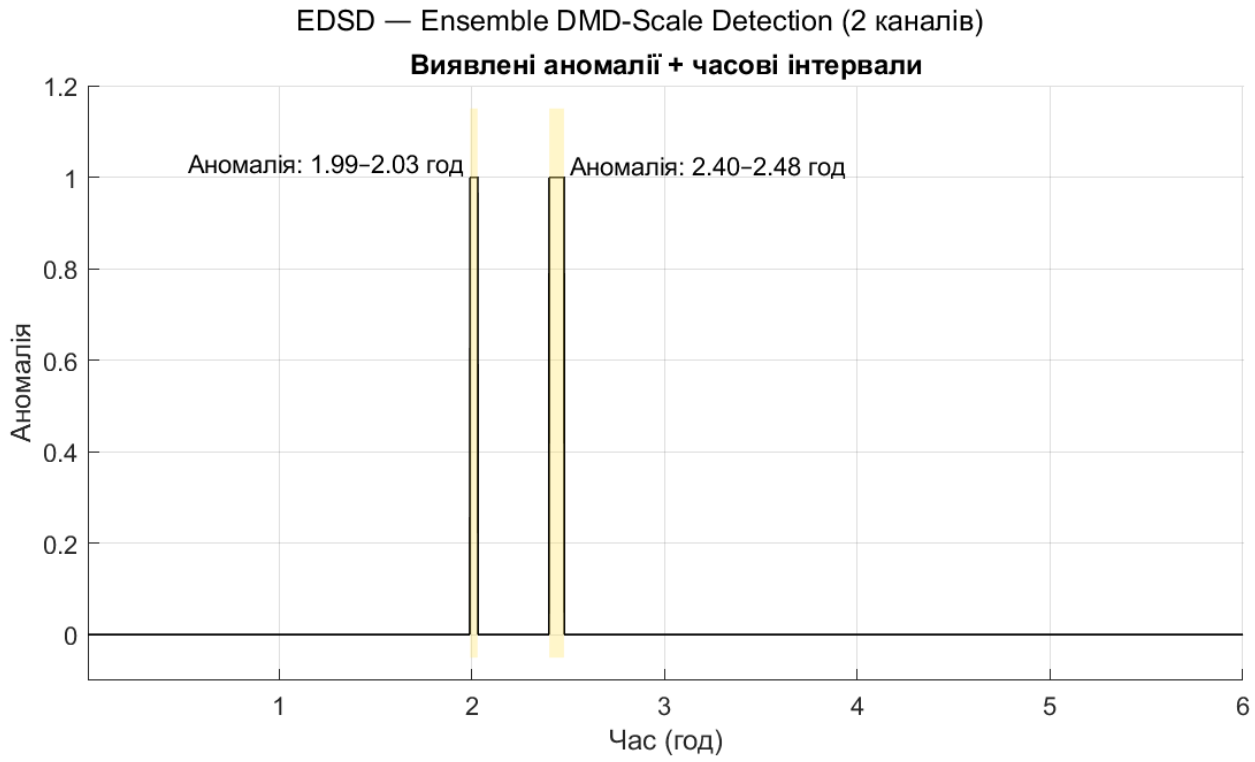


Рис.3.5. Результат виявлення аномалії та часові інтервали (EDSD) (канали - 2)

На графіку (рис.3.5) суцільною чорною лінією зображено бінарну послідовність, що позначає моменти перевищення ансамблевим показником порогу.

Світло-жовті виділені прямокутники позначають об’єднані часові інтервали, у яких спостерігалася ненормальна активність трафіку. Для кожного інтервалу підписано його початок та кінець у годинах, що дає змогу одразу оцінити тривалість аномальних подій.

Як видно з рис.3.5, алгоритм EDSD дозволяє виділити як короткі імпульсні збурення, так і довші періоди відхилення динаміки. Виявлені інтервали узгоджуються з поведінкою, спостереженою у попередніх графіках (рис.3.2–3.4), що підтверджує узгодженість роботи ансамблевого механізму та коректність дії порогової детекції.

Представлена візуалізація на рис.3.5 дозволяє чітко інтерпретувати процес виявлення та оцінити ефективність методу. Виявлені часові інтервали можуть бути використані для подальшого аналізу причин аномальної активності, діагностики мережевого обладнання або виявлення кіберзагроз.

Результати підтверджують, що метод EDSD здатний коректно виокремлювати як короткочасні пікові події, так і довготривалі структурні зміни у трафіку, забезпечуючи високу роздільну здатність для різних часових масштабів. Таким чином, розроблений підхід є придатним для практичного застосування у системах моніторингу комп’ютерних мереж.

Рис.3.6 ілюструє повну послідовність роботи методу EDSD у єдиному інтегрованому поданні. У верхній частині наведено вихідний багатоканальний трафік, далі — шкальні показники аномальності, ансамблевий індикатор з порогом прийняття рішення та, внизу, остаточні часові інтервали виявлених аномалій. Така паралельна візуалізація дозволяє чітко простежити причинно-наслідкові зв’язки між різними етапами аналізу: від локальних збурень у сигналі — до формування шкальних характеристик, зростання ансамблевого показника та кінцевої детекції.

Рисунок наочно демонструє, що метод EDSD коректно реагує на імпульсні та короткотривалі сплески трафіку, підтверджуючи наявність аномалій у моменти, де спостерігається відхилення у кількох часових масштабах одночасно.

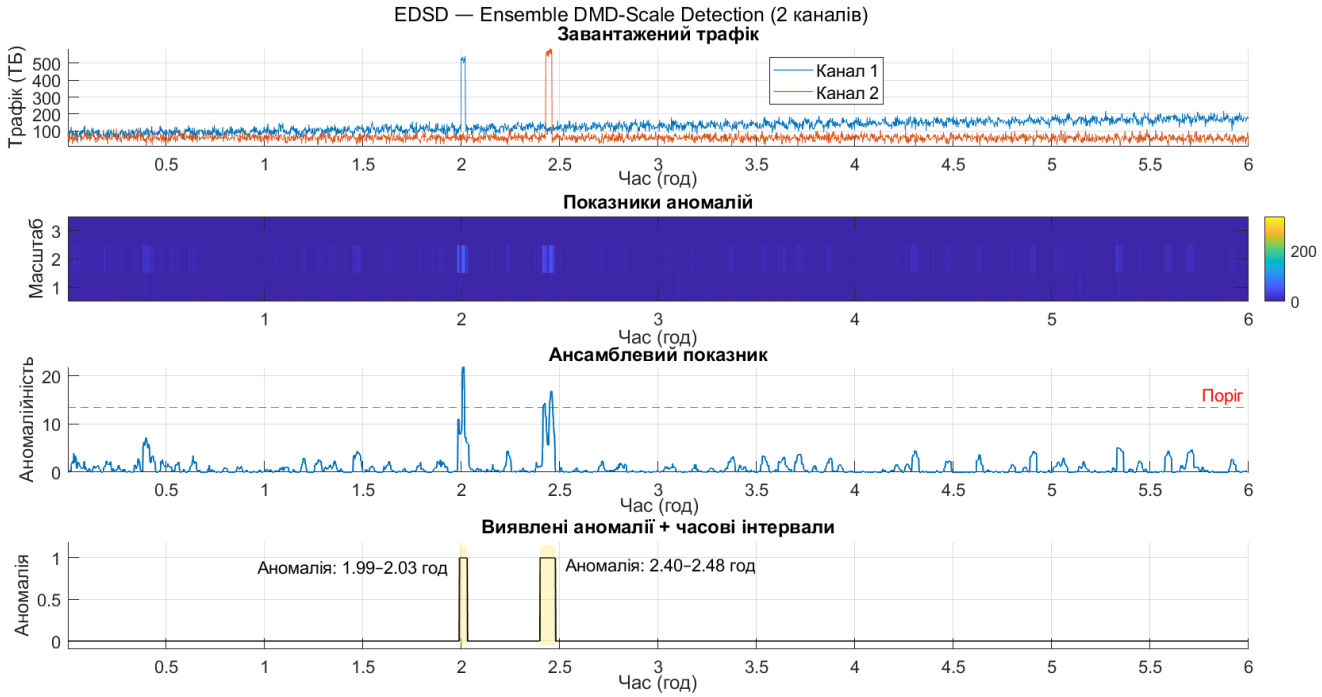


Рис.3.6. Паралельно відображені результати роботи методу EDSD

Завдяки поєднанню мультискейлового аналізу, робастної нормалізації та порогової детекції, результати виявлених інтервалів узгоджуються з динамікою вихідного сигналу, що підкреслює ефективність і надійність методу.

3.4 Висновки до розділу 3

У розділі представлено повну програмну реалізацію мультискейлового методу Ensemble DMD-Scale Detection (EDSD) для автоматизованого виявлення аномалій у часових профілях трафіку комп'ютерних мереж. Запропонована архітектура програмного засобу побудована за модульним принципом і включає шари попередньої обробки, модального аналізу, обчислення шкальних показників, їх робастної нормалізації, порогової детекції та візуалізації результатів. Така структурованість забезпечує гнучкість, можливість масштабування, прозорість логіки роботи та адаптацію під різні сценарії аналізу.

Проведена реалізація в MATLAB показала, що метод EDSD придатний для практичного використання й може працювати з багатоканальним трафіком у реальному часі. Завдяки мультискейловій структурі DMD-декомпозицій програмний

засіб виявляє різні типи аномалій — як короткі імпульсні сплески, так і довготривалі зміни динаміки. Аналіз енергетичних та спектральних показників, а також їхнє ансамблеве об’єднання продемонстрували високу чутливість до структурних змін трафіку та стабільність при наявності шумів і нестаціонарності.

Експериментальні результати підтвердили ефективність алгоритму: виявлені інтервали аномалій узгоджуються з реальними відхиленнями на графіках вихідного сигналу; мультискейловий аналіз підсилює здатність алгоритму відрізняти природні циклічні зміни від справжніх аномалій; робастні механізми нормалізації та порогового відбору забезпечують стійкість до викидів та флуктуацій трафіку. Візуалізація результатів у MATLAB є інформативною та сприяє інтерпретації роботи системи.

Таким чином, розроблений програмний засіб підтвердив працездатність теоретичних рішень, викладених у розділі 2, та довів ефективність запропонованого методу для практичного виявлення аномалій у часових профілях трафіку комп’ютерних мереж. Показані результати створюють підґрунтя для подальшого вдосконалення методу, його оптимізації та інтеграції у реальні системи моніторингу.

РОЗДІЛ 4
ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Охорона праці

Спеціальне навчання та інструктажі працівників із питань охорони праці є одним із основних принципів державної політики в галузі охорони праці і складовою системи управління охороною праці. Вони проводяться з працівниками в процесі їх трудової діяльності.

Згідно з Типовим положенням про навчання з питань охорони праці усі працівники при прийнятті на роботу і періодично в процесі роботи проходять на підприємстві навчання у формі інструктажів із питань охорони праці, вивчають правила надання першої долікарської допомоги потерпілим від нещасних випадків, а також правила поведінки при виникненні аварій.

Працівники, що виконують роботи підвищеної небезпеки (згідно з переліком таких робіт, що затверджується наказом Держпромгірнагляду), а також де є необхідність у професійному відборі, при прийнятті на роботу проходять попереднє спеціальне навчання і перевірку знань з питань охорони праці та періодичне навчання і перевірку знань в термін, встановлений відповідними галузевими нормативними актами, але не рідше одного разу на рік.

На промислових підприємствах України для працівників, що виконують роботи з обслуговування обладнання підвищеної небезпеки, обов'язкове курсове навчання з охорони праці (з обов'язковим іспитом), що проходить безпосередньо на виробництві за затвердженими роботодавцем і погодженими з органами Держпромгірнагляду програмами. Це робот и по обслуговуванню парових та водонагрівальних котлів, виробничих печей та інших теплових установок, устаткування, що працює під тиском, компресорів, холодильних установок, газового обладнання, електричного устаткування, підйомників, підіймальних механізмів, тракторних лопат, буртоукладачів, буртоукривальних машин, автонавантажувачів, електрокарів, тракторів та іншого внутрішнього заводського механізованого транспорту,

газоелектрозварювального обладнання, апаратів дифузії, центрифуг, кислотних та лужних установок, безтарного зберігання сировини, миття харчової сировини, такелажних, монтажних, ремонтних, навантажувально-розвантажувальних та інших робіт. Відповідальність за організацію навчання і перевірку знань на підприємстві покладається на роботодавця, а в структурних підрозділах - на керівників цих підрозділів. Контролює виконання цих завдань відділ охорони праці.

Допуск до роботи осіб, які не пройшли навчання та перевірку знань, забороняється.

Усі посадові особи, відповідно до переліку посад до початку виконання своїх обов'язків і періодично (один раз на три роки) проходять навчання і перевірку знань з питань охорони праці.

Навчання посадових осіб (керівників підприємств і установ та їх заступників тощо), що безпосередньо відповідають за організацію охорони праці на підприємстві чи установі (перелік посадових осіб наведено в додатку до Типового положення про навчання з питань охорони праці), проводиться в навчальних закладах, які мають дозвіл Державного Комітету України по нагляду за охороною праці на проведення такого навчання.

На підприємствах навчання з питань охорони праці організовує відділ охорони праці підприємства, залучаючи до цього працівників відділу охорони праці та спеціалістів, що пройшли навчання і перевірку знань у навчальних закладах або в установах Держпромгірнагляду. Для перевірки знань посадових осіб і спеціалістів наказом по підприємству створюється комісія, очолювана керівником підприємства. До комісії входять керівники (їх заступники) служби охорони праці, виробничо-технічних служб, представники місцевих органів державного нагляду за охороною праці, а також представники профспілкового комітету (комітетів).

Посадові особи та спеціалісти невеликих підприємств, де неможливо провести навчання та утворити комісію по перевірці знань, проходять навчання у відповідних місцевих навчальних закладах або на близьких за їх профілем виробництва підприємствах, а перевірку знань - в комісіях при місцевих органах Держпромгірнагляду.

Працівники, що не пройшли навчання і перевірку знань або при повторній перевірці показали незадовільні знання з питань охорони праці, звільняються з посади, а їх працевлаштування вирішується згідно з чинним законодавством.

Позачергова перевірка знань посадових осіб і спеціалістів проводиться в разі введення в дію або перегляду нормативних актів із питань охорони праці; введення в дію нового устаткування або нових технологічних процесів; при переведенні працівника на іншу роботу, що потребує додаткових знань із питань охорони праці; за вимогою працівника органу державного нагляду за охороною праці, в разі незнання актів про охорону праці.

4.2 Безпека в надзвичайних ситуаціях

У разі загрози чи настання надзвичайних ситуацій необхідним є забезпечення стійкості роботи об’єктів зв’язку, радіомовлення і телебачення до дії уражаючих факторів надзвичайних ситуацій

На виконання вимог Державних будівельних норм ДБН В.1.2-4:2019 «Інженерно-технічні заходи цивільного захисту» (на заміну ДБН В.1.2-4-2006) основними заходами щодо підвищення стійкості забезпечених об’єктів є [24]:

- захист робітників і службовців;
- підвищення стійкості інженерно-технічного комплексу;
- виключення або обмеження можливих негативних наслідків від вторинних факторів ураження;
- організація надійного матеріально-технічного забезпечення і стійких виробничих зв’язків;
- проведення заходів по зниженню можливих втрат і забезпеченню стабільності випуску продукції;
- забезпечення надійності управління виробництвом;
- завчасна підготовка до відновлення порушеного виробництва.

Захист робітників і службовців досягається:

– укриттям в захисних спорудах;

- евакуацією в заміську зону;
- забезпеченістю засобами індивідуального захисту; надійним та своєчасним оповіщенням про небезпеку;
- розробкою режимів захисту;
- герметизацією приміщень;
- проведенням профілактичних медичних заходів.

Підвищення стійкості інженерно-технічного комплексу досягається:

- заглибленням або розміщенням у незавалюваних приміщеннях устаткування і комунікацій;
- підвищенням міцнісних характеристик споруд;
- базуванням підприємства на декількох джерелах електропостачання;
- передбаченням на мережах газопостачання і теплофікації заходів проти витікання газу, пару, води, а також резервних ліній для переключення енергопостачання від інших магістралей (об’єктів).

Ймовірність виникнення на підприємстві пожеж залежить в основному від ступеня вогнестійкості будинків і споруд, пожежо- і вибухо- небезпечності, щільності забудови будинків і споруд і ступеня їх руйнувань.

При розробці заходів по виключенню або обмеженню дії на об’єкт та його елементи вторинних факторів поразки слід враховувати специфічні властивості об’єктів. Так, на об'єктах нафтопереробної і нафтохімічної промисловості внаслідок руйнування і ушкодження місткостей, трубопроводів і арматури можливе виникнення вибухонебезпечних газоповітряних сумішей, які можуть вибухнути і призвести до руйнування устаткування, будівельних конструкцій, будинків і споруд.

Виникнення і характер пожеж залежать не тільки від категорії виробництва по пожежній небезпеці, але й від ступеня вогнестійкості будинків і споруд. За ступенями вогнестійкості будинки і споруди поділяються на п’ять груп (I, II, III, IV і V).

Лікувальні корпуси психіатричних лікарень і диспансерів мають бути не нижче III ступеня вогнестійкості.

Будинки лікувальних закладів на 60 і менше ліжок та амбулаторно-поліклінічні заклади на 90 відвідувань за зміну дозволяється проектувати IV, V ступеня вогнестійкості з рубленими чи брущатими стінами.

Приміщення лікувальних, амбулаторно-поліклінічних закладів і аптек (крім приміщень медичного персоналу громадських будинків і споруд, аптечних кіосків) в разі розміщення їх в будинках іншого призначення мають бути відокремлені від решти приміщень протипожежними стінами 1-го типу і мати самостійні виходи назовні.

Ступінь вогнестійкості будинків і споруд визначається мінімальними межами вогнестійкості будівельних конструкцій і загоряння матеріалів, із яких ці конструкції збудовані. Вогнестійкість будівельних конструкцій визначається часом згоряння (запалювання) і вимірюється в годинах. Всі будівельні матеріали, як відомо, поділяються на три групи: такі, що згоряють, важко згоряють і не згоряють.

Для виключення або обмеження можливості виникнення і поширення пожеж, отруєнь та інших вторинних факторів ураження на підприємствах проводяться заходи по протипожежній та інженерно-технічній профілактиці. Вони включають заходи, що запобігають розповсюдженню пожежі (обладнання протипожежних розривів навколо нагрівальних печей і установок, створення і підготовка сил і засобів пожежогасіння, винесення складів паливо-мастильних матеріалів, вибухонебезпечних речовин за територію об’єкта та зниження їх запасів, обвалювання місткостей і складів СДОР, обладнання устроїв для їх нейтралізації, підготовка резервуарів з водою та прийняття інших необхідних мір захисту).

При оцінці стійкості об’єкта у НС необхідно також враховувати розташування об’єкта відносно меж зон можливого катастрофічного затоплення при зруйнуванні гребель гідровузлів і його наслідків [25].

На стійкість роботи об’єктів значний вплив може створити радіоактивне забруднення місцевості, тому для захисту робітників розроблюються необхідні режими радіаційного захисту.

Це такі заходи, як зниження запасів (на території підприємства) матеріально-технічних засобів, створення в заміській зоні необхідних запасів і резервів сировини,

палива, устаткування, комплектуючих виробів, запасних деталей, будівельних матеріалів, рухомих електричних і компресорних станцій для проведення рятівних і невідкладних робіт: підготовка підприємства до переводу на резервні (автономні) джерела електро-, паро- і водопостачання, завчасне дослідження можливостей використання місцевих джерел сировини, палива, комплектуючих деталей та інших необхідних для виробництва матеріалів.

Важливим заходом є завчасна підготовка до відновлення порушеного виробництва, для чого на об’єкті розробляється план відновлювальних робіт, що дозволить значно скоротити час на проведення робіт у випадку часткового зруйнування об’єкта.

4.3 Висновки до розділу 4

У підрозділі з охорони праці проаналізовано питання спеціального навчання та перевірки знань з питань охорони праці працівників як виконують роботи з підвищеної небезпеки.

У підрозділі з безпеки в надзвичайних ситуаціях проаналізовано питання аналізу заходів по підвищенню стійкості об’єктів зв’язку, радіомовлення і телебачення до дії уражаючих факторів надзвичайних ситуацій.

ЗАГАЛЬНІ ВИСНОВКИ

У кваліфікаційній роботі виконано комплексне дослідження процесів виявлення аномалій у часових профілях мережевого трафіку та розроблено метод і програмний засіб, що забезпечують підвищену точність і стійкість детекції в умовах нестаціонарності та багатомасштабної динаміки мережі.

Результати:

1. Проведено детальний аналіз і класифікацію існуючих методів виявлення аномалій, який здійснено шляхом систематизації статистичних, спектральних, фрактальних, евристичних та машинних підходів, у результаті чого визначено їхні обмеження та сформульовано науково-технічну проблему, що потребує створення нового мультискейлового методу.

2. Розроблено математичну модель мережевого трафіку, побудовану на поєднанні добових і тижневих циклів, локальних коливань, шумових складових і аномальних процесів та реалізовану у багатоканальному форматі, що забезпечило адекватне відтворення часової структури та природної варіативності трафіку.

3. Створено метод і алгоритм виявлення аномалій у часових профілях трафіку, реалізований на основі мультискейлового динамічного модального розкладу (DMD) з робастною нормалізацією та пороговою детекцією, що дало змогу надійно виділяти короткочасні й довготривалі відхилення від типових закономірностей.

4. Реалізовано програмний засіб для автоматизованого аналізу та виявлення аномалій, який розроблено в середовищі MATLAB із використанням модульної архітектури, що включає обчислення шкальних характеристик, формування ансамблевих показників та наочну візуалізацію модальних спектрів і аномальних інтервалів, завдяки чому забезпечено зручність практичного застосування.

5. Проведено експериментальні дослідження ефективності методу, які виконано за допомогою моделювання трафіку з добовою сезонністю та контрольованими аномаліями, і які показали високу точність детекції, стійкість до шумів та здатність коректно розрізняти природні циклічні зміни і справжні аномальні події.

ПЕРЕЛІК ПОСИЛАНЬ

1. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey. // *ACM Computing Surveys*. – 2009. – Vol. 41(3). – P. 1–58.

2. Ahmed M., Mahmood A. N., Hu J. A survey of network anomaly detection techniques. // *Journal of Network and Computer Applications*. – 2016. – Vol. 60. – P. 19–31.

3. Thottan M., Ji C. Anomaly detection in IP networks. // *IEEE Transactions on Signal Processing*. – 2003. – Vol. 51(8). – P. 2191–2204.

4. Patcha A., Park J. An overview of anomaly detection techniques: Existing solutions and latest technological trends. // *Computer Networks*. – 2007. – Vol. 51(12). – P. 3448–3470.

5. Lakhina A., Crovella M., Diot C. Diagnosing network-wide traffic anomalies. // *ACM SIGCOMM Computer Communication Review*. – 2004. – Vol. 34(4). – P. 219–230.

6. Kim S., Lee J., Kim S. Network traffic anomaly detection based on statistical characteristics of traffic features. // *Computer Communications*. – 2020. – Vol. 161. – P. 70–84.

7. Papoulis A., Pillai S. Probability, Random Variables, and Stochastic Processes. – 4th ed. – New York: McGraw-Hill, 2002. – 852 p.

8. Silverman B. W. Density Estimation for Statistics and Data Analysis. – London: Chapman & Hall, 1986. – 175 p.

9. Box G. E. P., Jenkins G. M., Reinsel G. C. Time Series Analysis: Forecasting and Control. – Hoboken: Wiley, 2016. – 712 p.

10. Cortes C., Vapnik V. Support-vector networks. // *Machine Learning*. – 1995. – Vol. 20. – P. 273–297.

11. MacQueen J. Some methods for classification and analysis of multivariate observations. // *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. – 1967. – Vol. 1. – P. 281–297.

12. Schölkopf B., Platt J., Smola A., Shawe-Taylor J., Williamson R. Estimating the support of a high-dimensional distribution. // *Neural Computation*. – 2001. – Vol. 13(7). – P. 1443–1471.

13. Hochreiter S., Schmidhuber J. Long short-term memory. // *Neural Computation*. – 1997. – Vol. 9(8). – P. 1735–1780.

14. Oppenheim A. V., Schafer R. W. Discrete-Time Signal Processing. – 3rd ed. – Pearson, 2009. – 1120 p.

15. Kutz J. N., Brunton S. L., Brunton B. W., Proctor J. L. Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems. – SIAM, 2016. – 234 p.

16. Willinger W., Paxson V., Taqqu M. S. Self-similarity and heavy tails: Structural modeling of network traffic. // *A Practical Guide to Heavy Tails*. – 1998. – P. 27–53.

17. Kantelhardt J. W. Fractal and multifractal time series. // *Physica A: Statistical Mechanics and its Applications*. – 2012. – Vol. 391(4). – P. 1270–1292.

18. Roesch M. Snort – Lightweight Intrusion Detection for Networks. // *Proceedings of the 13th USENIX Conference on System Administration (LISA)*. – 1999. – P. 229–238.

19. Khvostivskyy M., Osukhivska H., Khvostivska L., Lobur T., Velychko D., Lupenko S., Hovorushchenko T. Mathematical modelling of daily computer network traffic. The 1st International Workshop on Information Technologies: Theoretical and Applied Problems, ITTAP 2021. CEUR Workshop Proceedings. Ternopil, Ukraine, November 16-18, 2021. Vol. 3039. P.107-111. ISSN 1613-0073.

20. Khvostivska L., Khvostivskyi M., Dediv I., Yatskiv V., Palaniza Y. Method, Algorithm and Computer Tool for Synphase Detection of Radio Signals in Telecommunication Networks with Noises. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.173-180. ISSN 1613-0073.

21. Khvostivska L., Khvostivskyi M., Dediv I. Mathematical, algorithmic and software support for signals wavelet detection in electronic communications. Proceedings of the 2nd International Workshop on Computer Information Technologies in Industry 4.0

(CITI 2024). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2024. Vol. 3742. P.223-234. ISSN 1613-0073.

22. Khvostivskyy M., Osukhivska H., Khvostivska L., Lobur T., Velychko D. Mathematical modelling of daily computer network traffic. ITTAP-2021: Information Technologies: Theoretical and Applied Problems. The 1st International Workshop (November 16-18, 2021). Ternopil, UKRAINE. P.107-111. DOI: 10.1425/jsdtl.

23. Khvostivskyi, M., Khvostivska, L., Dediv, I., Yatskiv, V., & Kuchvara, O. (2025). Method and computer tool for synphase prediction of computer network traffic load level. Proceedings of the 3rd International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2025), 18–19 September 2025, Ternopil, Ukraine. CEUR Workshop Proceedings (Vol. 4057, pp. 261–269). CEUR-WS. ISSN 1613-0073.

24. Liliya Khvostivska, Mykola Khvostivskyy, Vasyl Dunetc, Iryna Dediv. Mathematical and Algorithmic Support of Detection Useful Radiosignals in Telecommunication Networks. Proceedings of the 2nd International Workshop on Information Technologies: Theoretical and Applied Problems (ITTAP 2022). Ternopil, Ukraine, November 22-24, 2022. P.314-318. ISSN 1613-0073.

25. Voloshchuk A., Osukhivska H., Khvostivskyi M., Sverstiuk A. (2025). Analysis of electricity consumption using the component method of periodically correlated random processes. *Computer Systems and Information Technologies*, (3), 74–82. https://doi.org/10.31891/csit-2025-3-8. ISSN 2710-0766.

26. Хвостівський В., Осухівська Г., Хвостівська Л. Програмне забезпечення системи опрацювання мережевого трафіку. Матеріали IX науково-технічної конференції «Інформаційні моделі, системи та технології» Тернопільського національного технічного університету імені Івана Пулюя, (Тернопіль, 8-9 грудня 2021р.). Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2020. С.102.

27. Методичні рекомендації до виконання, оформлення та захисту кваліфікаційних робіт для здобувачів другого (магістерського) рівня вищої освіти за спеціальністю 172 «Електронні комунікації та радіотехніка» галузі знань

17 «Електроніка, автоматизація та електронні комунікації»/уклад.: Дунець В.Л., Хвостівська Л.В., Дедів І.Ю. Тернопіль: ТНТУ, 2024. 56 с.

28. Стручок В.С. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної бо та заочної (дистанційної) форм навчання / В.С.Стручок. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.

29. Khvostivska L., Herasymchuk Yu., Zemledukh V,. Vorobets I. Method for Automated Detection of Load Levels in Computer Networks. Collection of Scientific Papers with the Proceedings of the 2nd International Scientific and Practical Conference «Global Trends in Science: Research, Innovation and Development» (September 29 – October 1, 2025, Varna, Bulgaria). European Open Science Space, 2025. P.145-149. DOI 10.70286/EOSS-29.09.2025. ISBN 979-8-89704-966-0.

ДОДАТКИ

ДОДАТОК А

Теза конференція та сертифікат учасника

UDC 01.1

Collection of Scientific Papers with the Proceedings of the 2nd International Scientific and Practical Conference «Global Trends in Science: Research, Innovation and Development» (September 29 – October 1, 2025, Varna, Bulgaria). European Open Science Space, 2025. 166 p.

ISBN 979-8-89704-966-0 (series)
DOI 10.70286/EOSS-29.09.2025

The conference is included in the Academic Research Index ReserchBib International catalog of scientific conferences.

The conference is registered in the database of scientific and technical events of UkrISTEI to be held on the territory of Ukraine (Certificate №552 dated 16.06.2025).

The materials of the conference are publicly available under the terms of the CC BY-NC 4.0 International license.

The materials of the collection are presented in the author's edition and printed in the original language. The authors of the published materials bear full responsibility for the authenticity of the given facts, proper names, geographical names, quotations, economic and statistical data, industry terminology, and other information.

ISBN 979-8-89704-966-0 (series)

EUROPEAN OPEN SCIENCE SPACE

© Participants of the conference, 2025
© Collection of scientific papers, 2025
© European Open Science Space, 2025

Proceedings of the 2nd International Scientific and Practical Conference
"Global Trends in Science: Research, Innovation and Development"
September 29 – October 1, 2025
Varna, Bulgaria

EUROPEAN OPEN SCIENCE SPACE

Section: Technical Sciences

METHOD FOR AUTOMATED DETECTION OF LOAD LEVELS IN COMPUTER NETWORKS

Khvostivska Liliia
Ph.D., Associate Professor
Herasymchuk Yurii
Student
Zemledukh Vladyslav
Student
Vorobets Ihor
Student
Ternopil Ivan Puluj National Technical University, Ukraine

The rapid development of information technologies [1] causes a significant increase in the load on computer networks. Modern telecommunication systems deal with traffic that is formed under the influence of cyclical patterns and random fluctuations. To ensure stable operation of the network, it is necessary not only to predict its states, but also to automatically detect load levels, which allows for timely identification of critical modes and make proactive decisions on resource management.

Traditional methods are based [2] on threshold criteria that do not take into account the stochastic nature of traffic and do not provide control over the probability of classification errors. In this paper, an approach to automated detection of load levels is proposed, which is based on the periodically correlated stochastic process (PCSP) model, in-phase/component signal analysis, and the Neumann–Pearson criterion.

Network load classification methods can be conditionally divided into several groups:

– Stochastic models (Poisson, Markov, ARIMA) describe randomness, but do not reproduce cyclicality.

– Fractal models (fBm) take into account self-similarity, but do not take into account diurnal fluctuations [5].

– Machine learning methods [3] provide high accuracy, but require large computational resources.

– Hybrid approaches combine different models, but are difficult to implement in practice [6].

The disadvantage of most of them is the use of simple thresholding schemes, which reduces the reliability of classification. The proposed automated detection system overcomes these limitations by integrating in-phase/component analysis and the Neumann-Pearson criterion.

Proceedings of the 2nd International Scientific and Practical Conference
"Global Trends in Science: Research, Innovation and Development"
September 29 – October 1, 2025
Varna, Bulgaria
EUROPEAN OPEN SCIENCE SPACE

Traffic load is considered as a periodically correlated stochastic process:

$$\xi(t) = S(t) + X(t), \quad t \in \mathrm{R} \tag{1}$$

where $S(t)$ – a deterministic periodic component with a daily period T, reflecting a repeating cyclicity;

$X(t)$ – stochastic component.

Thus, the model combines the regular structure of traffic and its random nature.

To move from the model to the processing tools, the random component $\xi(t)$ is given in the form of a harmonic expansion:

$$\xi(t) = \sum_{k \in Z} \xi_k(t) e^{-\frac{j2\pi k}{T}}, \quad t \in \mathrm{R} \tag{2}$$

where $\xi_k(t)$ – stochastic coefficients that vary with time, and exponential multipliers describe the periodic structure of the process.

This approach allows us to represent the random part as a superposition of in-phase harmonic components.

In-phase and component methods use these harmonic components to construct correlation features.

For centered traffic signals, a spectral-correlation representation is calculated for harmonic number k:

$$\hat{B}_k(u) = \frac{1}{N} \sum_{t \in Z} \overset{0}{\xi}(t) \overset{0}{\xi}(t-u) e^{-\frac{j2\pi k}{T}}, \tag{3}$$

where $\overset{0}{\xi}(t)$ – centered traffic signal relative to the mathematical expectation, $\overset{0}{\xi}(t) = \xi(t) - E[\xi(t)]$;

N – sample length; u – time shift;

k – harmonic number (correlation component number); Z – indexing area.

The correlation components $\hat{B}_k(u)$, which are calculated for centered traffic signals, reflect the time-frequency characteristics of the network load. However, in real data there are

– noise components (random fluctuations of short duration);

– anomalous peaks (sudden jumps due to local events) [4, 7];

– limited sample size, leading to statistical instability of estimates.

If raw correlation components $\hat{B}_k(u)$ are used, the classification results may be overly sensitive to random fluctuations.

To avoid this, a procedure of averaging over components was introduced:

$$\overline{C}(u) = \frac{1}{N} \sum_{k=1}^{K} B_k(u), \tag{4}$$

where $\hat{B}_k(u)$ – estimate of the correlation component on the u-th interval;

K – number of averaging components.

Proceedings of the 2nd International Scientific and Practical Conference
"Global Trends in Science: Research, Innovation and Development"
September 29 – October 1, 2025
Varna, Bulgaria
EUROPEAN OPEN SCIENCE SPACE

That is, averaging correlation components is a kind of «filter» of statistical features: it cuts off short-term random fluctuations and leaves long-term patterns that are truly informative for further classification of load levels.

As a result, a feature vector is formed:

$$C(u) = \left(\overline{C}(u_1), \overline{C}(u_2), ..., \overline{C}(u_m)\right), \tag{5}$$

which preserves the temporal dynamics of the load and is the basis for classification.

Automated detection of load levels in computer networks is formulated as a statistical hypothesis testing problem. In the simplest case, binary classification is considered:

– H_0: the load is average (normal network operation mode);

– H_1: the load is abnormal (minimum or critical).

In order to decide between hypotheses, a likelihood ratio is introduced:

$$\Lambda(C(u)) = \frac{f(C(u) \mid H_1)}{f(C(u) \mid H_0)}, \tag{6}$$

where $C(u)$ – vector of averaged in-phase/component correlation features;

$f(C(u) \mid H_1)$, $f(C(u) \mid H_0)$ – probability density of the appearance of these features when the hypotheses H_0 and H_1 are true.

The decision (classification) rule has the form:

$$\Lambda(C(u)) \underset{H0}{\overset{H1}{\lessgtr}} \eta, \tag{7}$$

where η – threshold value determined from the condition of controlling the probability of incorrect decisions (determined by the Neumann-Pearson criterion) [8].

This approach allows you to minimize the probability of missing a dangerous condition while controlling the probability of false alarms.

According to the Neumann–Pearson criterion, the optimal rule is considered to be one that minimizes the probability of missing an abnormal state (type II error):

$$\beta = P(\textit{рішення } H0 \mid H1) = \int_{\Lambda(C(u))<\eta} p(x \mid H1)dx, \tag{8}$$

provided that the probability of a false alarm (type I error):

$$\alpha = P(\textit{рішення } H1 \mid H0) = \int_{\Lambda(v)>\eta} p(x \mid H0)dx, \tag{9}$$

does not exceed a pre-set level (threshold) α0:

$$\alpha \le \alpha_0, \tag{10}$$

where α_0 – false alarm threshold.

Then the detection reliability (the probability of correctly detecting an abnormal condition) is defined as:

$$p_d = 1 - \beta. \tag{11}$$

Thus, the detection method controls the risk of false alarms (α) while minimizing the probability of missing a dangerous condition (β).

Proceedings of the 2nd International Scientific and Practical Conference
"Global Trends in Science: Research, Innovation and Development"
September 29 – October 1, 2025
Varna, Bulgaria
EUROPEAN OPEN SCIENCE SPACE

Since in practical conditions it is necessary to distinguish at least three network states, the problem is generalized to the multi-class case. The following classes are distinguished:

A_1: minimum load ($C(u) < \theta 1$) – suitable for preventive work

A_2: average load ($\theta 1 \le C(u) \le \theta 2$) – normal mode

A_3: critical load ($C(u) > \theta 2$) – a condition that carries the risk of overload

Interval classification rule:

$$A_1 : C(u) < \theta 1,\ A_2 : \theta 1 \le C(u) \le \theta 2,\ A_3 : C(u) > \theta 2. \quad (12)$$

Maximum likelihood classification for a multidimensional feature vector $C(u)$:

$$A(C(u)) = \arg\max_i p(C(u) | A_i). \quad (13)$$

Threshold values are calculated at the intersection points of the densities:

$$p(C(u) | A_1) = p(C(u) | C_2) \Rightarrow \theta_1, \quad (14)$$

$$p(C(u) | A_2) = p(C(u) | C_3) \Rightarrow \theta_2, \quad (15)$$

The combination of the likelihood ratio and the Neumann-Pearson criterion allows:

– ensure statistical optimality of decision-making;

– control the probability of misclassifications;

– generate automated decisions regarding load levels with high confidence.

As a result, the method not only detects critical modes, but also does so based on a rigorous mathematical model, taking into account the balance between first and second type errors.

The proposed method for automated detection of load levels is based on the PCVP model, centered signal values, estimation of correlation components according to the formula $\hat{B}_k(u)$, averaging over harmonic components, and application of the Neumann–Pearson criterion.

The method allows controlling the probability of type I errors, minimizing the omission of dangerous states, and applying both interval and probabilistic rules of multi-class classification.

References

1. Leland, W. E., Taqqu, M. S., Willinger, W., & Wilson, D. V. (1994). On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM Transactions on Networking, 2(1), 1–15. DOI: 10.1109/90.282603.

2. Papagiannaki, K., Moon, S., Fraleigh, C., Thiran, P., Tobagi, F., & Diot, C. (2003). Analysis of measured single-hop delay from an operational backbone network. IEEE Journal on Selected Areas in Communications, 21(6), 908–921. DOI: 10.1109/JSAC.2003.814650.

3. Hsieh, H. Y., & Sivakumar, R. (2002). Performance comparison of cellular and multi-hop wireless networks: A quantitative study. ACM SIGMETRICS Performance Evaluation Review, 30(1), 113–122. DOI: 10.1145/511399.511388.

4. Li, J., Mohapatra, P., & Chuah, C.-N. (2005). Online anomaly detection in network traffic using wavelet analysis. IEEE INFOCOM 2005, 6, 2815–2825. DOI: 10.1109/INFCOM.2005.1498537.

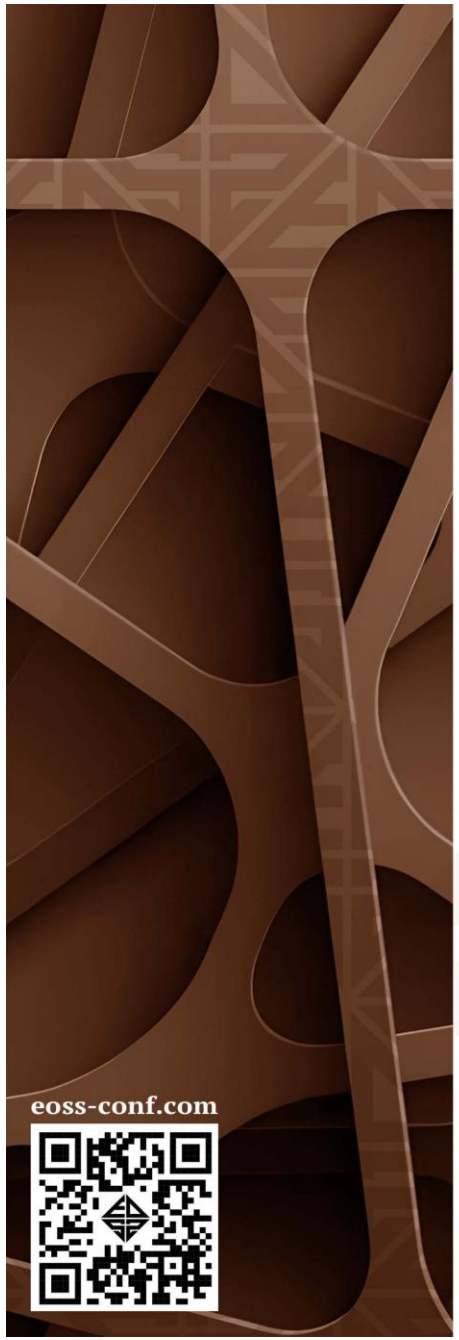
Proceedings of the 2nd International Scientific and Practical Conference
"Global Trends in Science: Research, Innovation and Development"
September 29 – October 1, 2025
Varna, Bulgaria

5. Abry, P., Veitch, D., & Flandrin, P. (1998). Long-range dependence: Revisiting aggregation with wavelets. Journal of Time Series Analysis, 19(3), 253-266. DOI: 10.1111/1467-9892.00090.
6. Khvostivska L., Khvostivskyi M., Dediv I., Yatskiv V., Palaniza Y. Method, Algorithm and Computer Tool for Synphase Detection of Radio Signals in Telecommunication Networks with Noises. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.173-180. ISSN 1613-0073.
7. Khvostivska L., Khvostivskyi M., Dediv I. Mathematical, algorithmic and software support for signals wavelet detection in electronic communications. Proceedings of the 2nd International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2024). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2024. Vol. 3742. P.223-234. ISSN 1613-0073.
8. Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society of London. Series A, 231(694-706), 289–337. DOI: 10.1098/rsta.1933.0009.

МОДЕРНІЗАЦІЯ ТА ДОСЛІДЖЕННЯ ПОКАЗНИКІВ ЯКОСТІ ЕЛЕКТРОПРИВОДА МЕХАНІЗМУ РОТОРНОГО ВАГОНОПЕРЕКИДАЧА З ВИКОРИСТАННЯМ СУЧАСНИХ СИСТЕМ ЕЛЕКТРОПРИВОДУ

Нежурін В.І.
доцент, к.т.н.
Майстренко Д.Н.
магістрант
Салюк О.Д.
магістрант
Куваєв В.Ю.
старший викладач
Кафедра електричної інженерії
Український державний університет науки і технологій, Дніпро, Україна

На сучасному етапі складної економічної ситуації в металургійній промисловості української держави склалися сприятливі та необхідні обставини для модернізації діючих виробничих фондів основного технологічного обладнання металургійних підприємств. Вдосконалення напівпровідникової техніки, розвиток, теорії керування створює умови для розробки комплектних



EUROPEAN OPEN SCIENCE SPACE

CERTIFICATE
of participation

Ihor Vorobets

took part in the 2nd International Scientific and Practical Conference

«GLOBAL TRENDS IN SCIENCE: RESEARCH, INNOVATION AND DEVELOPMENT»

24 Hours of Participation
(0,8 ECTS credits)

Head of the
organizing committee
Helen Volokitina

EUROPEAN OPEN SCIENCE SPACE 3382508349

EOSS-25/0929-080 September 29 – October 1, 2025, Varna, Bulgaria

ДОДАТОК Б

Скрипт програмного забезпечення

```
% ================================================================
% Виявлення аномалій в багатоканальному трафіку комп’ютерної мережі
% ================================================================

clearvars; close all; clc;
rng(2025);

%% ---------------- Load signal from file ----------------------
filename = 'trafic_anomal_2chanel.dat';
X = load(filename);  % канали по рядках
[m,T] = size(X);
t = 1:T;
time_hours = t / 500;    % 500 samples/hour → час у годинах
fprintf('Loaded %d channels, %d samples from %s\n', m, T, filename);

time_unit = 'год';
volume_unit = 'МБ/с';
scales = [6, 18, 2000];
num_scales = numel(scales);
r = 3;  % DMD rank

%% ---------------- Log transform ----------------
Xlog = log1p(max(X,0));

%% ---------------- EDSD detection ----------------
recon_err_scales = zeros(num_scales,T);
eig_changes_scales = zeros(num_scales,T);
prev_eigs = cell(num_scales,1);

for s=1:num_scales
    w = scales(s);
    if w >= T
        warning('Scale %d >= T. Skipping this scale.', w);
        continue;
    end
    halfw = floor(w/2);

    for center = 1+halfw : T-halfw
        idx = (center-halfw):(center+halfw-1);
        Xw = Xlog(:,idx);
        X1 = Xw(:,1:end-1);
        X2 = Xw(:,2:end);

        [U,S,V] = svd(X1,'econ');
        rr = min([r size(S,1)]);
        U_r = U(:,1:rr); S_r = S(1:rr,1:rr); V_r = V(:,1:rr);
```

```
        A_tilde = U_r' * X2 * V_r / S_r;
        [W_eigs,D] = eig(A_tilde);
        lambda = diag(D);
        Phi = X2 * V_r / S_r * W_eigs;

        if cond(Phi) < 1e12
            b = Phi \ Xw(:,1);
        else
            b = pinv(Phi) * Xw(:,1);
        end

        time_dynamics = zeros(rr, w);
        for k2=1:rr
            time_dynamics(k2,:) = b(k2) * (lambda(k2).^(0:(w-1)));
        end
        X_recon = real(Phi * time_dynamics);

        recon_err = norm(Xw - X_recon, 'fro') / sqrt(numel(Xw));

        cur_eig = sort(lambda,'ComparisonMethod','real');
        if isempty(prev_eigs{s})
            eig_change = 0;
        else
            prev = sort(prev_eigs{s},'ComparisonMethod','real');
            kcomp = min([3 numel(prev) numel(cur_eig)]);
            if kcomp<=0
                eig_change = 0;
            else
                eig_change = sum(abs(abs(cur_eig(end-kcomp+1:end)) - ...
                                     abs(prev(end-kcomp+1:end))));
            end
        end

        recon_err_scales(s,center) = recon_err;
        eig_changes_scales(s,center) = eig_change;
        prev_eigs{s} = cur_eig;
    end
end

%% ---------------- Normalize & aggregate ----------------
score_per_scale = zeros(num_scales,T);
for s=1:num_scales
    v1 = recon_err_scales(s,:);
    v2 = eig_changes_scales(s,:);

    nz1 = v1(v1~=0); nz2 = v2(v2~=0);
    if isempty(nz1), norm_v1 = zeros(size(v1)); else
        norm_v1 = (v1 - median(nz1)) ./ (1.4826 * mad(nz1,1) + eps);
    end
    if isempty(nz2), norm_v2 = zeros(size(v2)); else
```

```
        norm_v2 = (v2 - median(nz2)) ./ (1.4826 * mad(nz2,1) + eps);
    end

    recon_gain = max(scales) / scales(s);
    eig_gain = (scales(s) / max(scales)) * 6;

    if max(abs(norm_v1))>eps, norm_v1 = norm_v1 / (max(abs(norm_v1))+eps);
end
    if max(abs(norm_v2))>eps, norm_v2 = norm_v2 / (max(abs(norm_v2))+eps);
end

    score_per_scale(s,:) = ...
        recon_gain * max(0,norm_v1) + 0.9 * (eig_gain * max(0,norm_v2));
end

ensemble_score = mean(score_per_scale,1);
ensemble_score_smooth = movmedian(ensemble_score,5);

%% ---------------- Thresholding ----------------
med = median(ensemble_score_smooth);
madval = 1.4826 * mad(ensemble_score_smooth,1);
k_thresh = 35;
thresh = med + k_thresh * madval;

detected = ensemble_score_smooth > thresh;
detected = movmax(detected,15);
detected = movmedian(detected,9) > 0.5;
detected = bwareaopen(detected,20);

%% ---------------- Visualization (4 plots, FontSize=14) ----------------
fs = 13;  % font size

figure('Position',[100 100 1400 900]);

% 1) Raw channels
subplot(4,1,1);
cols = lines(m); hold on;
for ch=1:m
    plot(time_hours,X(ch,:),'Color',cols(ch,:),'LineWidth',0.8);
end
ylabel(['Трафік (' volume_unit ')'], 'FontSize', fs);
xlabel(['Час (' time_unit ')'], 'FontSize', fs);
title('Завантажений трафік', 'FontSize', fs);
legend(arrayfun(@(i)sprintf('Канал %d',i),1:m,'UniformOutput',false),
'Location','best', 'FontSize', fs);
grid on; axis tight
set(gca,'FontSize',fs)

% 2) Scores by scale
subplot(4,1,2);
imagesc(time_hours,1:num_scales,score_per_scale); axis xy;
```

```
ylabel('Масштаб', 'FontSize', fs); xlabel(['Час (' time_unit ')'],
'FontSize', fs);
title('Показники аномалій', 'FontSize', fs); colorbar; grid on; axis tight
set(gca,'FontSize',fs)

% 3) Ensemble + threshold
subplot(4,1,3); hold on;
plot(time_hours, ensemble_score_smooth,'LineWidth',1.2);
yline(thresh,'--r','Поріг', 'FontSize', fs);
ylabel('Аномалійність', 'FontSize', fs); xlabel(['Час (' time_unit ')'],
'FontSize', fs);
title('Ансамблевий показник', 'FontSize', fs); grid on; axis tight
set(gca,'FontSize',fs)

% 4) Detections + intervals
subplot(4,1,4); hold on;
plot(time_hours, detected, 'k','LineWidth',1.2);

% HIGHLIGHT INTERVALS
[labeled, num_blobs] = bwlabel(detected);
for i=1:num_blobs
    idx = find(labeled==i);
    t_start = time_hours(idx(1));
    t_end   = time_hours(idx(end));

    fill([t_start t_end t_end t_start], [-0.05 -0.05 1.15 1.15], ...
         [1 0.9 0.4], 'FaceAlpha',0.35, 'EdgeColor','none');

    text((t_start+t_end)/2, 1.05, ...
        sprintf('Аномалія: %.2f–%.2f год', t_start, t_end), ...
        'HorizontalAlignment','center','FontSize', fs, 'Color','k');
end

xlabel(['Час (' time_unit ')'], 'FontSize', fs);
ylabel('Аномалія', 'FontSize', fs);
title('Виявлені аномалії + часові інтервали', 'FontSize', fs);
grid on; axis tight; ylim([-0.1 1.2]); set(gca,'FontSize',fs)

sgtitle(sprintf('EDSD – Ensemble DMD-Scale Detection (%d каналів)', m),
'FontSize', fs+2);
```