

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр

(назва освітнього ступеня)

на тему: Метод виявлення змін ймовірнісних характеристик ритмічних
сигналів для комунікаційних систем

Виконав(ла): студент(ка) 6 курсу, групи РАМ-61
спеціальності

172 Електронні комунікації та радіотехніка

(шифр і назва спеціальності)

Максимів Т.Р.
(підпис) (прізвище та ініціали)

Керівник Дедів І.Ю.
(підпис) (прізвище та ініціали)

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

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

Рецензент Дозорська О.Ф.
(підпис) (прізвище та ініціали)

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 Електронні комунікації та радіотехніка

(шифр і назва спеціальності)

студенту Максиміву Тарасу Романовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Метод виявлення змін ймовірнісних характеристик ритмічних сигналів для комунікаційних систем

Керівник роботи Дедів Ірина Юріївна, к.т.н., доцент

(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 11 грудня 2025 року

3. Вихідні дані до роботи Технічне завдання, виявлення змін ймовірнісних характеристик ритмічних сигналів

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

1. Аналітична частина

2. Основна частина

3. Науково-дослідна частина

4. Охорона праці та безпека в надзвичайних ситуаціях

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

6. Консультанти розділів роботи

[illegible]

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

[illegible]

Студент _____
(підпис)

Максимів Тарас Романович
(прізвище та ініціали)

Керівник роботи _____
(підпис)

Дедів Ірина Юріївна

(прізвище та ініціали)

АНОТАЦІЯ

Метод виявлення змін ймовірнісних характеристик ритмічних сигналів для комунікаційних систем // Кваліфікаційна робота магістра // Максимів Тарас Романович // ТНТУ ім. І.Пулюя, ФПТ // Тернопіль, 2025 // с. - 80, рис. - 17, дод. - 2, бібл. - 49.

Ключові слова: ЙМОВІРНІСТЬ, РОЗЛАДКА, СИГНАЛ, ПОРІГ.

В роботі розроблено метод виявлення змін ймовірнісних характеристик ритмічних сигналів для комунікаційних систем. Власне задачу виявлення змін ймовірнісних характеристик сигналів зведено до задачі виявлення розладки, яка характеризує момент або інтервал часу, коли статистичні чи структурні властивості сигналу помітно змінюються порівняно з “нормальним” режимом. В якості моделі таких сигналів використано клас стаціонарних процесів, зокрема кусково стаціонарних. Використано спектрально-кореляційні методи опрацювання, а саме опрацювання проводиться на інтервалах ковзного вікна. Встановлено, що значення варіації оцінок усереднених спектрів значно відрізняються між ділянками сигналу до та після виникнення розладки і шляхом моніторингу їхнього значення при трансляції ковзного вікна по сигналу можна виявити моменти виникнення розладки в сигналі.

ANNOTATION

Method for detecting changes in probabilistic characteristics of rhythmic signals for communication systems // Maksymiv T.R. // TNTU, FPT // Ternopil, 2025 // p. - 80, fig. - 17, appl. - 2, bibl. - 49.

Key words: PROBABILITY, DISRUPTION, SIGNAL, THRESHOLD.

The paper develops a method for detecting changes in the probabilistic characteristics of rhythmic signals for communication systems. The problem of detecting changes in the probabilistic characteristics of signals is reduced to the problem of detecting a disorder, which characterizes the moment or time interval when the statistical or structural properties of the signal change noticeably compared to the “normal” mode. As a model of such signals, a class of stationary processes, in particular piecewise stationary ones, is used. Spectral-correlation processing methods are used, and the processing is carried out on the intervals of a sliding window. It is established that the values of the variation of the estimates of the averaged spectra differ significantly between the signal sections before and after the occurrence of the disorder, and by monitoring their value when translating a sliding window over the signal, it is possible to detect the moments of the occurrence of the disorder in the signal.

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1. АНАЛІТИЧНА ЧАСТИНА.....	10
1.1 Аналіз методів виявлення точок зміни часових рядів.....	10
1.2 Критерії виявлення змін.....	15
1.3 Оцінка продуктивності алгоритмів виявлення змін.....	17
1.4 Висновки до розділу 1.....	22
РОЗДІЛ 2. ОСНОВНА ЧАСТИНА.....	23
2.1 Загальні методи виявлення розладки.....	23
2.2 Огляд методів виявлення змін в часових рядах.....	25
2.3 Порівняння методів.....	42
2.4 Висновки до розділу 2.....	47
РОЗДІЛ 3. НАУКОВО-ДОСЛІДНА ЧАСТИНА.....	49
3.1 Модель тестового сигналу для виявлення змін ймовірнісних характеристик.....	49
3.2 Результати експериментальних досліджень.....	53
3.3 Алгоритм роботи методу.....	58
3.4 Висновки до розділу 3.....	59
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	61
4.1 Охорона праці.....	61
4.2 Безпека в надзвичайних ситуаціях.....	63
4.3 Висновки до розділу.....	66
ВИСНОВКИ.....	67
ПЕРЕЛІК ПОСИЛАНЬ.....	69
ДОДАТКИ	

ВСТУП

Актуальність. Задача виявлення змін ймовірнісних характеристик ритмічних сигналів є актуальною з огляду на те, що ритмічні процеси лежать в основі широкого класу природних і технічних систем, а їхні порушення часто є раннім індикатором зміни стану об'єкта або переходу системи в інший режим функціонування. Ритмічні сигнали (квазіперіодичні коливання) характерні для біомедичних вимірювань, віброакустичної діагностики механізмів, телеметрії енергетичних систем, а також для інформаційно-комунікаційних та сенсорних мереж. У таких застосуваннях зміна ймовірнісних характеристик, що може бути означена як розладка, проявляється як зміна періоду, амплітуди, фазових співвідношень, спектральної структури або рівня шуму, що може бути пов'язано як із нормальними адаптаційними реакціями, так і з відмовами чи патологічними станами.

Науково-практична значущість виявлення розладки зумовлена потребою у своєчасному прийнятті рішень на основі поточкових даних. У біомедичних системах це відповідає завданням раннього попередження про аритмії, ішемічні епізоди, порушення регуляції серцевого ритму або дихання; у технічній діагностиці — виявленню дефектів підшипників, розбалансування, резонансних явищ чи деградації вузлів до настання критичної відмови. Важливо, що затримка у детекції розладки може призводити до підвищення ризиків для безпеки, погіршення якості сервісу або зростання витрат на ремонт, тоді як надмірна кількість хибних спрацювань знижує довіру до систем моніторингу та ускладнює експлуатацію.

Методологічна актуальність задачі пов'язана зі складністю реальних ритмічних сигналів: вони часто є нестационарними, квазіперіодичними та зашумленими, мають міжциклову варіабельність, артефакти вимірювання та можуть демонструвати поступові дрейфи параметрів. Це створює потребу в

розробленні та обґрунтуванні методів, які поєднують чутливість до статистично значущих змін із робастністю до шуму та природної варіабельності. Додатковим чинником є тенденція до впровадження безперервного моніторингу (wearables, промисловий IoT, розподілені сенсорні системи), де алгоритми мають працювати в реальному часі, з обмеженими обчислювальними ресурсами та гарантованими характеристиками за ймовірністю хибної тривоги і середнім часом до виявлення.

Таким чином, виявлення змін ймовірнісних характеристик ритмічних сигналів є актуальною науковою задачею, що знаходиться на перетині статистичної обробки сигналів, теорії прийняття рішень і прикладних систем моніторингу. Її розв'язання забезпечує підвищення надійності діагностики та контролю стану, а також має безпосередню практичну цінність у промислових, інформаційних та телекомунікаційних застосуваннях.

Мета. Розробка методу виявлення змін ймовірнісних характеристик ритмічних сигналів для комунікаційних систем. **Задачі:**

- аналіз способів виявлення змін ймовірнісних характеристик ритмічних сигналів;
- зведення задачі виявлення змін ймовірнісних характеристик ритмічних сигналів до задачі виявлення розладки;
- розроблення методу виявлення розладки ритмічних сигналів;
- експериментальне дослідження розробленого методу виявлення розладки ритмічних сигналів.

Об'єкт дослідження: процес виявлення змін ймовірнісних характеристик ритмічних сигналів.

Предмет дослідження: метод виявлення змін ймовірнісних характеристик ритмічних сигналів.

Наукова новизна. Запропоновано для виявлення розладки використати метод, що ґрунтується на опрацюванні сигналу в межах ковзного вікна із наступним оцінюванням оцінок спектрів потужності.

Практичне значення. Розроблений метод може бути використаний для виявлення часових моментів переходів між різними режимами за змінами ймовірнісних характеристик відповідних сигналів, в безпекових системах для виявлення нетипового трафіку чи кіберзагроз тощо.

РОЗДІЛ 1

АНАЛІТИЧНА ЧАСТИНА

1.1 Аналіз методів виявлення точок зміни часових рядів

1.1.1 Загальні поняття.

Аналіз часових рядів стає дедалі важливішим у різних галузях, включаючи медицину, аерокосмічну галузь, фінанси, бізнес, метеорологію та розваги. Дані часових рядів – це послідовності вимірювань з часом, що описують поведінку систем. Ця поведінка може змінюватися з часом через зовнішні події та/або внутрішні систематичні зміни в динаміці/розподілі [1]. Виявлення точок зміни (CPD) – це проблема знаходження різких змін у даних, коли змінюється властивість часового ряду [2]. Сегментація, виявлення меж, виявлення подій та виявлення аномалій – це подібні концепції, які іноді застосовуються і виявлення точок зміни. На відміну від CPD, оцінка точок зміни намагається моделювати та інтерпретувати відомі зміни в часових рядах, а не ідентифікувати, що зміна відбулася. Основна увага оцінок точок зміни приділяється опису характеру та ступеня відомої зміни.

Моніторинг медичного стану: Безперервний моніторинг здоров'я пацієнтів включає виявлення тенденцій у фізіологічних змінних, таких як частота серцевих скорочень, електроенцефалограма та електрокардіограма, для здійснення автоматизованого моніторингу в режимі реального часу. Дослідження проводять виявлення точок зміни для конкретних медичних проблем, таких як проблеми зі сном, епілепсія, інтерпретація магнітно-резонансної томографії та розуміння діяльності мозку [3].

Виявлення зміни клімату: Методи кліматичного аналізу, моніторингу та прогнозування, що використовують виявлення точок зміни, стали дедалі

важливішими протягом останніх кількох десятиліть через можливе виникнення зміни клімату та збільшення вмісту парникових газів в атмосфері [4].

Розпізнавання мовлення: Розпізнавання мовлення являє собою процес перетворення усних висловлювань у слова або текст. Методи виявлення точок зміни застосовуються тут для сегментації аудіо та розпізнавання меж між тишею, реченнями, словами та шумом [5].

Аналіз зображень: Дослідники та практики збирають дані зображень протягом певного часу або відеодані для відеоспостереження. Виявлення раптових подій, таких як порушення безпеки, можна сформулювати як проблему виявлення точок зміни. Тут спостереження в кожній часовій точці є цифровим кодуванням зображення [6].

Аналіз людської діяльності: Виявлення точок зупинки або переходів активності на основі характеристик спостережуваних даних датчиків з розумних будинків або мобільних пристроїв можна сформулювати як виявлення точок зміни. Ці точки зміни корисні для сегментації діяльності, взаємодії з людьми з мінімізацією перерв, надання послуг, що враховують діяльність, та виявлення змін у поведінці, які дають уявлення про стан здоров'я [7].

Розладка сигналів (часто також кажуть зміна режиму, точка зміни, change-point) — це момент або інтервал часу, коли статистичні або структурні властивості сигналу помітно змінюються порівняно з “нормальним” режимом.

Найчастіші види розладки:

- зміна середнього рівня (mean shift) – сигнал “перейшов” на інший базовий рівень;
- зміна дисперсії / шумності (variance shift);
- зміна тренду – був майже сталий, став зростати/спадати;
- зміна частотного складу / спектра – з’явилась нова гармоніка, змінився основний тон;

- зміна форми розподілу (навіть якщо mean/variance майже ті самі) – “хвости” стали важчі, більше викидів;
- структурна зміна моделі – інший авторегресійний порядок, інша динаміка системи.

Розладки бувають:

- стрибкоподібні (abrupt) або плавні (gradual / drift);
- одноразові або багаторазові;
- офлайн (аналіз запису) або онлайн (виявлення “в реальному часі”).

Загальна постановка задачі. Нехай маємо гіпотези:

- H_0 : сигнал у нормальному режимі (параметри/розподіл сталі);
- H_1 : з деякого моменту τ параметри/розподіл інші.

Мета – знайти τ (точку зміни) або швидко сигналізувати про зміну, контролюючи хибні тривоги.

1.1.2. Передумови виявлення точок зміни.

На рис. 1.1 зображено графік часового ряду, який містить кілька точок зміни. Дані ілюструють довгострокові середньорічні тенденції температури Шпіцбергена за період 1899-2010 років. Дані можна використовувати для виявлення зміни клімату. Цей графік підкреслює спостереження, що клімат Шпіцбергена пройшов через шість різних режимів за цей період. Ми називаємо ці частини часового ряду станами часового ряду або періодами часу, коли параметри, що керують процесом, не змінюються. Два послідовні різні стани розрізняються точкою зміни. Метою виявлення точок зміни є ідентифікація цих меж станів шляхом виявлення точок зміни.

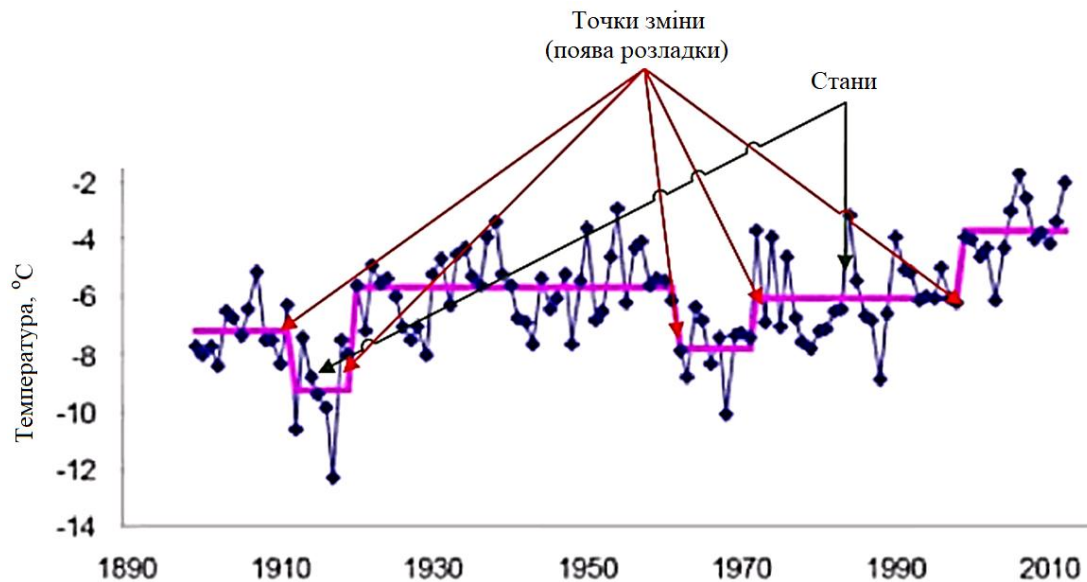


Рис. 1.1. Зразки часових рядів та точок змін (горизонтальні лінії вказують на окремі стани)

1.1.3 Визначення та формулювання проблеми.

Почнемо з представлення визначень ключових термінів, Потік даних часових рядів – це нескінченна послідовність елементів

$$S = \{x_1, \dots, x_i, \dots\}$$

де x_i — d-вимірний вектор даних, що надходить у момент часу i [8].

Стаціонарний часовий ряд — це процес зі скінченною дисперсією, статистичні властивості якого є постійними з часом. Це визначення припускає, що:

- функція середнього значення $\mu_t = E(x_t)$ є постійною величиною та не залежить від часу t .
- функція автоковаріації $\gamma(s, t) = cov(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$ залежить від моментів часу s та t лише через їхню різницю в часі, або $|s - t|$

Незалежні та однаково розподілені (i.i.d.) змінні є взаємно незалежними одна від одної та однаково розподіленими в тому сенсі, що вони походять з

одного й того ж розподілу ймовірностей. I.i.d. часовий ряд є окремим випадком стаціонарного часового ряду.

Враховуючи часовий ряд T фіксованої довжини m (підмножина потоку даних часового ряду) та x_t як вибірку ряду в момент часу t , матрицю WM усіх можливих підпоследовностей довжини k можна побудувати, переміщуючи ковзне вікно розміру k вздовж T та розміщуючи підпоследовність $X_p = \{x_p, x_{p+1}, \dots, x_{p+k}\}$ (рис. 2) в p -тий рядок WM . Розмір результуючої матриці WM is $(m - k + 1) \times n$.

У часовому ряді, використовуючи ковзне вікно X_t як зразок замість x_t , інтервал X_t з матрицею Генкеля $\{X_t, X_{t+1}, \dots, X_{t+n-1}\}$ як показано на рис. 1.2, буде набором з n ретроспективних вибірок підпоследовності, починаючи з моменту часу t [9].

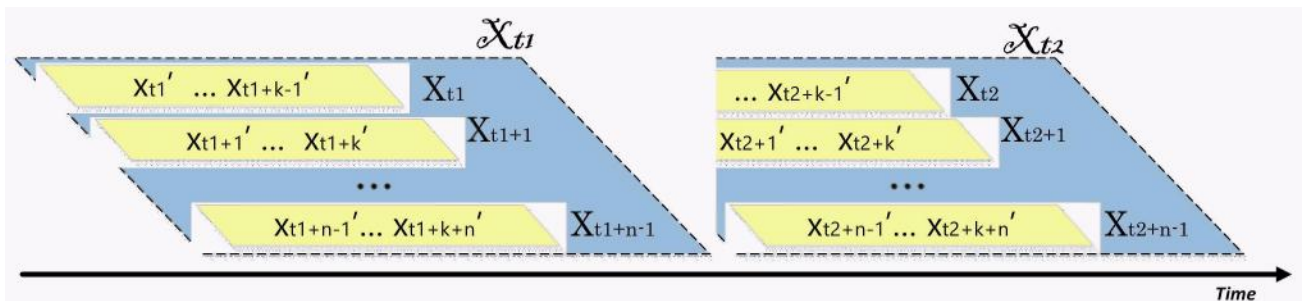


Рис. 1.2. Ілюстративний приклад позначень часових рядів.

Точка зміни являє собою перехід між різними станами в процесі, який генерує дані часового ряду.

Нехай $\{x_m, x_{m+1}, \dots, x_n\}$ є послідовністю змінних часового ряду. Виявлення точки зміни (CPD) можна визначити як задачу перевірки гіпотез між двома альтернативами: нульовою гіпотезою H_0 : «Змін не відбувається» та альтернативною гіпотезою H_A : «Зміна відбувається» [10]

$$1) \quad H_0 : \mathbb{P}_{x_m} = \dots = \mathbb{P}_{x_k} = \dots = \mathbb{P}_{x_n}.$$

2) H_A : Існує $m < k^* < n$ таке, що $\mathbb{P}_{x_m} = \dots = \mathbb{P}_{x_{k^*}} \neq \mathbb{P}_{x_{k^*+1}} = \dots = \mathbb{P}_{x_n}$. де $\mathbb{P}_{x_t}$ — це функція щільності ймовірності початку ковзного вікна в точці x_t , а k^* — точка зміни.

1.2 Критерії виявлення змін

Практичне застосування виявлення точок зміни створює низку нових проблем, які необхідно вирішити.

1.2.1 Онлайн-виявлення.

Алгоритми виявлення точок зміни традиційно класифікуються як «онлайн» або «офлайн». Офлайн-алгоритми розглядають весь набір даних одночасно та оцінюють попередні значення, щоб розпізнати, де відбулася зміна. Метою цього сценарію, як правило, є ідентифікація всіх точок зміни послідовності в пакетному режимі. На відміну від цього, онлайн-алгоритми, або алгоритми реального часу, працюють одночасно з процесом, який вони контролюють, обробляючи кожен точку даних, коли вона стає доступною, з метою виявлення точки зміни якомога швидше після її виникнення, в ідеалі до надходження наступної точки даних.

На практиці жоден алгоритм виявлення точок зміни не працює в ідеальному реальному часі, оскільки він повинен перевірити нові дані, перш ніж визначити, чи відбулася точка зміни між старою та новою точками даних. Однак, різні онлайн-алгоритми потребують різної кількості нових даних, перш ніж зможе відбутися виявлення точки зміни. На основі цього спостереження визначимо новий термін, який буде використовуватися далі. Позначимо як ϵ — алгоритм реального часу - це онлайн-алгоритм, якому потрібно щонайменше ϵ зразків даних у новому пакеті даних, щоб мати змогу знайти точки змін. Офлайн-алгоритм можна розглядати як ∞ — реального часу, а повністю онлайн-алгоритм є 1-реальним часом, оскільки для кожної точки даних він може передбачити, чи

відбудеться точка зміни перед новою точкою даних. Менші значення можуть призвести до сильніших, більш чутливих алгоритмів виявлення точки зміни.

1.2.2 Масштабованість.

Дані реального світу з таких джерел, як діяльність людини чи супутники дистанційного зондування, стають дедалі більшими як за кількістю точок даних, так і за кількістю вимірів. Методи виявлення змін повинні бути розроблені обчислювально ефективним способом, щоб вони могли масштабуватися до величезних розмірів даних [11]. Тому, порівнюється обчислювальна вартість альтернативних алгоритмів CPD, щоб визначити, який з них може досягти оптимального (або достатньо хорошого) рішення якомога швидше. Один із способів порівняння обчислювальної вартості алгоритмів полягає в тому, щоб визначити, чи є алгоритм параметричним, чи непараметричним. Розрізнення параметричних та непараметричних підходів важливе, оскільки непараметричні підходи продемонстрували більший успіх для масивних наборів даних. Крім того, обчислювальна вартість параметричних методів вища, ніж непараметричних, і не так добре масштабується з розміром набору даних.

Параметричний підхід визначає певну функціональну форму, яку має вивчити модель, а потім оцінює невідомі параметри на основі позначених навчальних даних. Після навчання моделі навчальні приклади можна відкинути. На відміну від цього, непараметричні методи не роблять жодних припущень щодо форми базової функції. Однак, всі доступні дані повинні бути наперед збережені [12].

1.2.3 Обмеження алгоритму

Підходи до CPD також можна розрізнити на основі вимог, що накладаються на вхідні дані та алгоритм. Ці обмеження важливі для вибору відповідного методу виявлення точок змін у певній послідовності даних. Обмеження, пов'язані з природою даних часових рядів, можуть впливати зі стаціонарності, незалежної ідентифікації, розмірності або безперервності даних [13].

Деякі алгоритми вимагають інформації про дані, такої як кількість точок змін у даних, кількість станів у системі та характеристики станів системи [14]. Ще одним важливим питанням у параметричних методах є ступінь чутливості алгоритму до вибору початкових значень параметрів.

1.3 Оцінка продуктивності алгоритмів виявлення змін

Для порівняння альтернативних алгоритмів CPD та оцінки очікуваної результуючої продуктивності необхідні показники продуктивності. Було введено багато показників продуктивності для оцінки алгоритмів виявлення точок зміни на основі типу рішень, які вони приймають. Вихідні дані алгоритмів CPD можуть містити наступне:

- рішення «так/ні» щодо точки зміни (алгоритм є бінарним класифікатором)
- ідентифікація точки зміни з різним рівнем точності (тобто точка зміни відбувається протягом x одиниць часу). Цей тип алгоритму використовує багатокласовий класифікатор або методи навчання без учителя.
- час наступної точки зміни (або час усіх точок зміни в послідовності).

У випадку перших двох типів вихідних даних, стандартні методи оцінки алгоритмів навчання з учителем можуть бути використані для оцінки продуктивності детектора точок зміни. Першим кроком в оцінці продуктивності детектора точок зміни з учителем є створення матриці помилок, яка підсумовує фактичні та прогнозовані класи.

Деякі корисні показники продуктивності, які можна використовувати для оцінки алгоритмів CPD, наведено нижче. Хоча вони описані в контексті бінарної класифікації, кожен з них може бути поширений на класифікацію більшої кількості класів, видаючи показники для кожного класу окремо або в комбінації.

1) Точність, розрахована як відношення правильно класифікованих точок даних до загальної кількості точок даних. Цей показник дає загальне уявлення

про продуктивність алгоритму. Супутником точності є коефіцієнт помилок, який обчислюється як $1 - \text{Точність}$. Точність і коефіцієнт помилок не дають уявлення про джерело помилки або розподіл помилок між різними класами. Крім того, вони неефективні для оцінки продуктивності в наборі даних, незбалансованому за класами, що типово для виявлення точок змін, оскільки вони вважають різні типи помилок класифікації однаково важливими. Чутливість і g-середнє є корисними показниками для використання в цьому випадку.

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

2) Чутливість, також відома як повний відлік або істинно позитивний коефіцієнт (коефіцієнт TP). Це стосується частини класу точок зміни, яка була розпізнана правильно.

$$\text{Sensitivity} = \text{Recall} = \text{TP Rate} = \frac{TP}{TP+FN}$$

3) G-середнє. Виявлення точки зміни зазвичай призводить до проблеми навчання з незбалансованим розподілом класів, оскільки відношення змін до загальних даних є малим. Як результат, G-середнє зазвичай використовується як показник ефективності безперервного професійного розвитку (CPD). Тут використовуються як показники чутливості, так і специфічності для оцінки продуктивності алгоритму як з точки зору співвідношення позитивної точності (Sensitivity), так і співвідношення негативної точності (Specificity).

$$G - \text{mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}} = \sqrt{\frac{TP}{TP+FN} \times \frac{TN}{FP+TN}}$$

4) Точність. Розраховується як відношення істинно позитивних точок даних (точок змін) до загальної кількості точок, класифікованих як точки змін.

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

5) F-міра (також відома як f-оцінка або fl-оцінка). Ця міра забезпечує спосіб поєднання точності та повноти як міри загальної ефективності алгоритму CPD. F-міра розраховується як відношення зваженої важливості точності та повноти.

$$F - measure = \frac{(1+\beta)^2 \times Recall \times Precision}{\beta^2 \times Recall + Precision} = \frac{(1+\beta)^2 \times \frac{TP}{TP+FN} \times \frac{TP}{TP+FP}}{\beta^2 \times \frac{TP}{TP+FN} + \frac{TP}{TP+FP}}$$

6) Крива робочих характеристик приймача (ROC). Оцінювання на основі ROC-кривої сприяє явному аналізу компромісу між коефіцієнтами істинно позитивних та хибнопозитивних результатів. Це робиться шляхом побудови двовимірного графіка з коефіцієнтом хибнопозитивних результатів на осі x та коефіцієнтами правдиво позитивних результатів на осі y. Алгоритм CPD створює пару (TP_Rate, FP_Rate), яка відповідає одній точці в ROC-просторі. Один алгоритм, як правило, можна вважати кращим за інший, якщо його точка ближче до координати (0,1) (верхній лівий кут), ніж інша. Щоб оцінити загальну продуктивність алгоритму, можна подивитися на площу під кривою ROC. Загалом, потрібно, щоб коефіцієнт хибнопозитивних результатів був низьким, а коефіцієнт правдиво позитивних результатів - високим. Це означає, що чим ближче до 1 значення AUC, тим сильніший алгоритм. Ще одним корисним показником, який можна отримати з ROC-кривої, є коефіцієнт рівних помилок (EER), який є точкою, де коефіцієнт хибнопозитивних результатів та коефіцієнт хибнонегативних результатів рівні. Ця точка залишається малою завдяки сильному алгоритму.

7) Крива точності та повного відтворення (крива PR). Крива PRC також може бути згенерована та використана для порівняння альтернативних алгоритмів CPD. Крива PR відображає коефіцієнт точності як функцію коефіцієнта повного відтворення. Хоча оптимальна продуктивність алгоритму для кривої ROC позначена точками у верхньому лівому куті простору, оптимальна продуктивність у просторі PR знаходиться поблизу верхнього правого кута. Як і у випадку з ROC, площа під PRC може бути обчислена для порівняння двох алгоритмів та спроби оптимізувати продуктивність CPD. Крива PR, зокрема, забезпечує глибокий аналіз, коли розподіл класів сильно асиметричний.

Якщо різниця в часі між виявленою точкою зміни (CP) та фактичною CP являє собою міру продуктивності (використовуючи методи контрольованого або неконтрольованого CPD), то вищезазначені показники не є відповідним вибором. Оцінка продуктивності цих алгоритмів не така проста, як у попередньому випадку, оскільки немає єдиної мітки, за якою можна виміряти продуктивність алгоритму. Однак для цього випадку існує ряд корисних показників, включаючи:

- середня абсолютна похибка (MAE). Вона безпосередньо вимірює, наскільки близько прогнозована CP до фактичної CP. Абсолютне значення різниці між прогнозованим і фактичним часом підсумовується та нормалізується по кожній з точок КТ.

$$MAE = \frac{\sum_{i=1}^{\#CP} |Predicted(CP) - Actual(CP)|}{\#CP}$$

- середньоквадратична помилка (MSE) – це добре відома альтернатива MAE. У цьому випадку, оскільки помилки зводяться до квадратів, результуючий показник буде дуже великим, якщо в класифікованих даних існує кілька суттєвих викидів.

$$MSE = \frac{\sum_{i=1}^{\#CP} (Predicted(CP) - Actual(CP))^2}{\#CP}$$

- середня знакова різниця (MSD). Окрім обчислення різниці між прогнозованим та фактичним часом CP, цей показник враховує напрямок помилки (прогнозування до або після фактичного часу CP).

$$MSD = \frac{\sum_{i=1}^{\#CP} (Predicted(CP) - Actual(CP))}{\#CP}$$

-середньоквадратична похибка (RMSE). Агрегує різницю між прогнозованою та фактичною похибкою та зводить кожену різницю у квадрат, щоб видалити коефіцієнт знака. Квадратний корінь обчислюється з кінцевої оцінки, щоб компенсувати коефіцієнт масштабування, зведений у квадрат індивідуальних різниць.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{\#CP} (Predicted(CP) - Actual(CP))^2}{\#CP}}$$

Нормалізована середньоквадратична похибка (NRMSE). Цей показник усуває чутливість значень до одиничного розміру прогнозованого значення. NRMSE сприяє більш безпосередньому порівнянню похибки між різними наборами даних та допомагає в інтерпретації показників похибки. Два поширені методи полягають у нормалізації похибки до діапазону спостережуваних КП або нормалізації до середнього значення спостережуваних КП.

$$NRMSE = \frac{RMSE}{MaxLength(Actual\ CP) - MinLength(Actual\ CP)}$$

$$NRMSE = \frac{RMSE}{mean(Actual\ CP)}$$

1.4 Висновки до розділу 1

Проведено аналіз методів виявлення точок зміни часових рядів, основною метою яких є знаходження різких змін у даних, коли змінюється властивість часового ряду. Ця задача є актуальною наприклад для розпізнавання мовлення, аналізу зображень, виявлення сигналів на фоні завад, оцінювання змін стану систем чи режимів їхнього функціонування тощо.

Розглянуто поняття розладки сигналів, що характеризує момент або інтервал часу, коли статистичні чи структурні властивості сигналу помітно змінюються порівняно з “нормальним” режимом.

Розглянуто критерії методів виявлення змін, зокрема можливість виявлення в режимі реального часу чи шляхом оцінювання всього набору даних, масштабованості тощо. Також розглянуто способи оцінки продуктивності алгоритмів виявлення змін.

РОЗДІЛ 2

ОСНОВНА ЧАСТИНА

2.1 Загальні методи виявлення розладки

Проаналізуємо типові та найпоширеніші методи виявлення розладки.

1) Класичні статистичні “контрольні карти” (онлайн).

Добрі, коли зміна проста (mean/variance) і потрібна швидка реакція.

Shewhart chart – ловить великі стрибки (порогові правила).

CUSUM – дуже популярний для малих стійких зсувів середнього/параметра.

EWMA – добре відпрацьовує поступові дрейфи (згладжене накопичення).

Page–Hinkley – варіант для зсуву середнього з шумом.

Переваги цих методів такі: прості, швидкі, легко налаштовувати поріг під рівень false alarm.

Недоліки: вимагають більш-менш стаціонарного “нормального” режиму й часто припущень про шум.

2) Методи на основі відношення правдоподібностей (LR / GLR):

Коли є модель “до” і “після” (або хоча б сімейство моделей).

SPRT (Sequential Probability Ratio Test) – класичний метод для послідовного тестування.

GLR (Generalized Likelihood Ratio) – коли параметри зміни невідомі, їх оцінюють на ходу.

Matched change detector – якщо форма зміни відома (наприклад, стрибок певної величини).

Переваги цих методів такі: оптимальні/наближено оптимальні за певних умов.

Недоліки: модельні припущення можуть бути “крихкими”.

3) Непараметричні / робастні тести (офлайн або онлайн):

Rank-тести (Манна–Уїтні, Вілкоксона) – для зсуву медіани/розподілу.

KS-тест (Kolmogorov–Smirnov) – зміна розподілу загалом.

Тести однорідності дисперсій (Левена тощо) – для зміни “шумності”.

Kernel / MMD – порівняння розподілів у більш загальному вигляді.

Переваги цих методів такі: менше припущень, стійкі до викидів.

Недоліки: інколи потребують більше даних/вікна.

4) Віконні (sliding window) підходи: “ознаки - поріг”

Практичний інженерний шлях:

- сформувати ковзне вікно (наприклад, 0.5–5 с залежно від задачі).

- обчислити ознаки: середнє, RMS, дисперсія; енергія в смугах частот, спектральний центроїд; ентропія, коефіцієнт асиметрії/ексцес; автокореляція, AR-коефіцієнти

- порівняти з базовим режимом: пороги, z-score, відстань Махаланобіса тощо.

Переваги цих методів такі: легко адаптувати під конкретний сенсор/систему.

Недоліки: якість залежить від вибору ознак і правильного “нормального” еталона.

5) Частотно-часові методи (коли є зміни “в спектрі”).

Якщо розладка проявляється в частоті/модуляціях:

- STFT/спектрограма + виявлення стрибків енергії в смугах;

- Вейвлети (DWT/CWT) – добре ловлять локальні зміни та розриви;

- Зміни домінантної частоти / гармонік (для вібрацій, акустики, ЕКГ тощо).

6) Байєсівські підходи (особливо для онлайн).

- BOCPD (Bayesian Online Change Point Detection) – тримає розподіл імовірності точки зміни.

- HMM / перемикальні моделі режимів — “режим 1 / режим 2 / режим 3...”
- Switching Kalman Filter — для динамічних систем.

Переваги цих методів такі: природно працюють з невизначеністю, дають імовірності.

Недоліки: складніші, потрібні припущення/підбір.

7) ML/Deer підходи (коли патерни складні).

Класифікація/детекція аномалій на ознаках або на сирому сигналі

Autoencoder / LSTM-AE — збільшення помилки реконструкції як сигнал розладки

One-class SVM, Isolation Forest — якщо є лише “норма”

Переваги цих методів такі: ловлять складні нелінійні зміни.

Недоліки: потрібні дані, ризик “навчитися на шум”, важче гарантувати рівень хибних тривог.

2.2 Огляд методів виявлення змін в часових рядах

Багато алгоритмів машинного навчання було розроблено, вдосконалено та адаптовано для виявлення точок зміни. Проведемо огляд основних алгоритмів, які зазвичай застосовуються до задачі CPD. Ці методи включають як методи з учителем, так і без учителя, обрані на основі бажаного результату алгоритму.

2.2.1 Методи з учителем.

Алгоритми з учителем – це алгоритми, які вивчають відображення вхідних даних на цільовий атрибут даних, який зазвичай є міткою класу. На рис. 1.3 наведено огляд методів з учителем, що використовуються для виявлення точок зміни. Коли для виявлення точок зміни використовується підхід з учителем, алгоритми навчання можна навчати як бінарні або багатокласові класифікатори. Якщо вказано кількість станів, алгоритм виявлення точок зміни навчається знаходити кожну межу стану. Ковзне вікно переміщується по даних, розглядаючи

кожен можливий поділ між двома точками даних як можливу точку зміни. Хоча цей підхід має простіший етап навчання, необхідно надати достатню кількість та різноманітність навчальних даних для представлення всіх класів. З іншого боку, виявлення кожного класу окремо надає достатньо інформації для визначення як характеру, так і обсягу виявлених змін. Для цієї задачі навчання можна використовувати різноманітні класифікатори. Приклади включають дерево рішень, байєсівський метод, байєсівську мережу, метод опорних векторів, метод найближчого сусіда, приховану марківську модель, умовне випадкове поле та модель гаусової суміші (GMM) [14, 15].

Альтернативою є трактування виявлення точок зміни як задачі бінарних класів, де всі можливі послідовності переходів станів (точок зміни) представляють один клас, а всі послідовності всередині станів представляють другий клас. Хоча в цьому випадку потрібно вивчити лише два класи, це набагато складніша задача навчання, якщо кількість можливих типів переходів велика. Як і в попередньому типі контрольованих підходів, у цьому підході до навчання кожна ознака у вхідному векторі вказує на джерело можливих змін. Отже, будь-який алгоритм навчання з учителем, який генерує інтерпретовану модель (наприклад, дерево рішень або навчальний метод правил), не лише визначить зміну, але й опише природу зміни.

Інший підхід з учителем полягає у використанні віртуального класифікатора. Цей метод виходить за рамки простого виявлення змін та фактичної інтерпретації зміни, що відбувається між двома послідовними вікнами. Віртуальний класифікатор прикріплює гіпотетичну мітку (+1) до кожного зразка з першого вікна та (-1) до кожного зразка з другого вікна, а потім навчає віртуальний класифікатор (ВК) за допомогою будь-якого методу з учителем на основі позначених точок даних. Якщо між двома вікнами є точка зміни, вони повинні бути правильно класифіковані класифікатором, а точність класифікації р повинна бути значно вищою, ніж випадковий шум $\text{rand}-0,5$. Для перевірки

значущості зміни показника використовується обернена функція виживання біноміального розподілу для визначення критичного значення $P < \text{critical}$, за якого очікується, що випробування Бернуллі перевищать rand з рівнем довіри. Нарешті, якщо $p > \text{critical}$, між двома вікнами існує значна зміна. Після виявлення точки зміни класифікатор перенавчається, використовуючи всі зразки у двох сусідніх вікнах. Якщо деякі ознаки відіграють домінуючу роль у класифікаторі, то саме вони характеризують різницю.

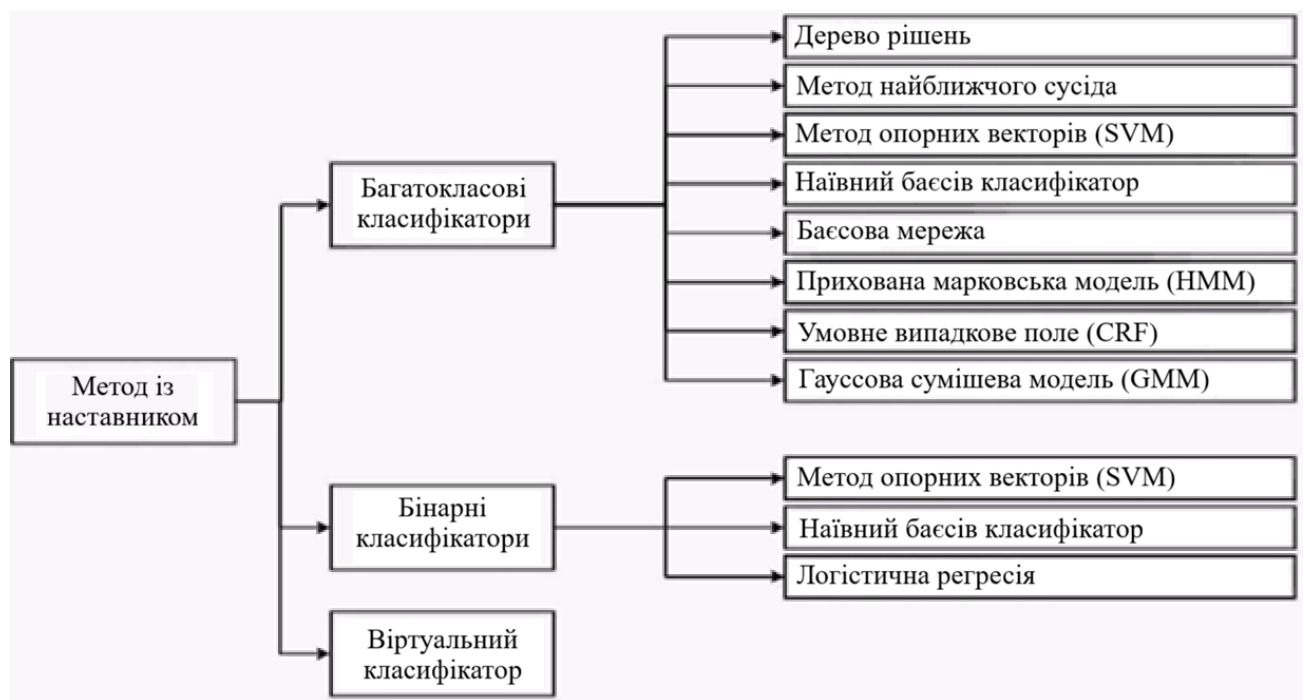


Рис. 2.1. Методи з учителем для виявлення точок зміни.

2.2.2 Методи без учителя.

Алгоритми навчання без учителя зазвичай використовуються для виявлення закономірностей у немаркованих даних. У контексті виявлення точок зміни такі алгоритми можна використовувати для сегментації даних часових рядів, таким чином знаходячи точки зміни на основі статистичних характеристик даних. Сегментація без учителя є привабливою, оскільки вона може обробляти різноманітні ситуації без необхідності попереднього навчання для кожної

ситуації. На рис. 2.1 наведено огляд методів без учителя, які використовувалися для виявлення точок зміни. Раніше описані методи використовують коефіцієнт правдоподібності, заснований на спостереженні, що щільність ймовірності двох послідовних інтервалів однакова, якщо вони належать до одного стану. Іншим традиційним рішенням є моделювання підпростору, яке представляє часовий ряд з використанням просторів станів і таким чином виявляє точки зміни, прогножуючи параметри простору станів. Імовірнісні методи оцінюють розподіли ймовірностей нового інтервалу на основі даних, що спостерігалися з моменту попередньої точки зміни-кандидата. На відміну від цього, методи на основі ядра відображають спостереження на простір ознак вищої вимірності та виявляють точки зміни, порівнюючи однорідність кожної підпослідовності. Метод на основі графіків – це нещодавно запроваджений метод, який представляє спостереження часових рядів у вигляді графіка та застосовує статистичні тести для виявлення точок змін на основі цього представлення. Нарешті, методи кластеризації групують дані часових рядів у відповідні стани та знаходять зміни, виявляючи відмінності між характеристиками станів.

2.2.3 Методи відношення правдоподібності.

Типове статистичне формулювання виявлення точок змін полягає в аналізі розподілів ймовірностей даних до та після точки зміни-кандидата та ідентифікації кандидата як точки зміни, якщо два розподіли суттєво відрізняються. У цих підходах логарифм відношення правдоподібності між двома послідовними інтервалами в даних часових рядів контролюється для виявлення точок змін.

Ця стратегія вимагає двох кроків. По-перше, щільність ймовірності двох послідовних інтервалів обчислюється окремо. По-друге, обчислюється відношення цих щільностей ймовірності. Найбільш відомим алгоритмом точок змін є кумулятивна сума [16], яка накопичує відхилення відносно заданої цілі

вхідних вимірювань та вказує на існування точки зміни, коли кумулятивна сума перевищує заданий поріг.

Пошук змін [17] – це ще один поширений метод, який зводить проблему виявлення точок змін до виявлення викидів на основі часових рядів. Цей метод підганяє модель авторегресії (AR) до даних для представлення статистичної поведінки часового ряду та поступово оновлює оцінки його параметрів, щоб поступово враховувати вплив минулих прикладів. Розглядаючи часовий ряд x_t , можна моделювати часовий ряд, використовуючи режим AR порядку A , таким чином:

$$x_t = \omega x_{t-k}^{t-1} + \varepsilon$$

де $x_{t-k}^{t-1} = (x_{t-1}, x_{t-2}, \dots, x_{t-k})$ попередні спостереження, $\omega = (\omega_1, \dots, \omega_k) \in \mathbb{R}^k$ – константи, а ε – нормальна випадкова величина, що генерується відповідно до гаусового розподілу, подібного до білого шуму. Оновлюючи параметри моделі, обчислюється функція щільності ймовірності в момент часу t , і маємо послідовність щільностей ймовірності $\{p_t: t = 1, 2, \dots\}$. Далі, допоміжний часовий ряд y_t генерується шляхом присвоєння оцінки кожній точці даних.

Ця функція оцінки визначається як середнє значення логарифмічної правдоподібності, $Score(y_t) = -\log p_{t-1}(y_t)$, або статистичного відхилення, $Score(y_t) = d(p_{t-1}, p_t)$, де $d(*,*)$ забезпечується будь-якою з кількох функцій відстані, включаючи варіаційну відстань, відстань Хеллінгера або квадратичну відстань. Нові дані часового ряду представляють різницю між кожною парою послідовних інтервалів часового ряду. Щоб виявити точки змін, потрібно знати, чи є різкі зміни між двома послідовними різницями. Для цього ще одна AR-модель підганяється до часового ряду на основі різниці та нової послідовності функцій щільності ймовірності $\{q_t: t = 1, 2, \dots\}$. Оцінка точки зміни визначається за

допомогою вищезгаданої функції оцінки. Вища оцінка вказує на вищу ймовірність того, що це точка зміни.

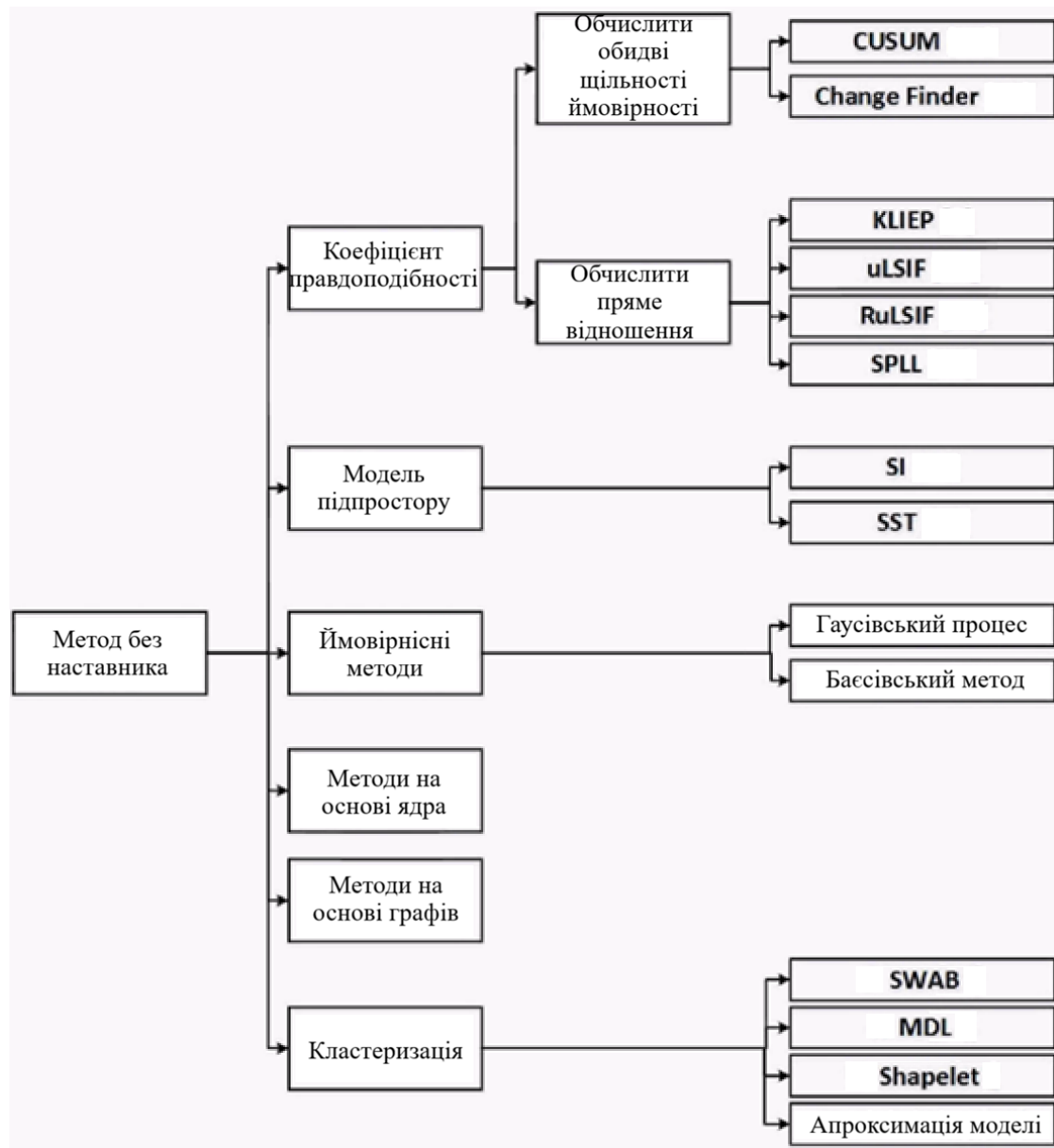


Рис. 1.4. Неконтрольовані методи виявлення точок зміни.

Оскільки ці методи спираються на попередньо розроблені параметричні моделі та є менш гнучкими в реальних сценаріях виявлення точок зміни, деякі нещодавні дослідження вводять більш гнучкі непараметричні варіації, оцінюючи співвідношення густин ймовірності безпосередньо без необхідності виконувати

оцінку густини. Обґрунтування цієї ідеї оцінки співвідношення густин полягає в тому, що знання двох густин передбачає знання співвідношення густин. Однак зворотне твердження невірне: знання співвідношення не обов'язково означає знання двох густин, оскільки таке розкладання не є унікальним. Слідуючи цій ідеї, були розроблені методи прямої оцінки співвідношення густин. Ці методи моделюють співвідношення густин між двома послідовними інтервалами X та X' за допомогою непараметричної моделі гаусового ядра наступним чином:

$$g(X) = \frac{p(X)}{p'(X')} = \sum_{l=1}^n \theta_l K(X, X_l)$$

$$K(X, X') = \exp\left(-\frac{\|X - X'\|^2}{2\sigma^2}\right)$$

де $p(X)$ – розподіл ймовірностей інтервальної області X , $\theta = (\theta_1, \dots, \theta_n)^T$ параметри, які потрібно вивчити зі зразків даних, X – це ковзне вікно, і $\sigma > 0$ – параметр ядра. На етапі навчання параметри θ визначаються таким чином, щоб мінімізувати міру несхожості. Враховуючи оцінку коефіцієнта густини, апроксиматор міри несхожості між двома вибірками X_t і X_{t+n} розраховується на етапі тестування. Чим вищий показник несхожості, тим більша ймовірність того, що точка є точкою зміни.

Популярним вибором для показника несхожості є дивергенція Кульбака-Лейблера (KL):

$$KL[p(x)||p'(x)] = - \int p'(x) \log \frac{p(x)}{p'(x)} dx$$

Процедура оцінки важливості Кульбака-Лейблера (KLIEP) оцінює коефіцієнт густини за допомогою дивергенції KL. Ця задача є задачею опуклої оптимізації, тому єдине глобальне оптимальне рішення можна легко отримати, наприклад, методом градієнтної проекції. Проектований градієнтний спуск

рухається в напрямку негативного градієнта на кожному кроці та проектується на допустимий параметр. Отримане наближення дивергенції KL наведено в наступному рівнянні.

$$\widehat{KL} = \frac{1}{n} \sum_{i=1}^n \log \hat{g}(Y_i)$$

Іншим прямим методом оцінки коефіцієнта густини є uLSIF (Unconstrained Least-Squares Improvance Fixment - метод важливості без обмежень), який використовує дивергенцію Пірсона (PE) як міру несхожості, що показано як:

$$PE[p(x)||p'(x)] = \int p'(x) \left(\frac{p(x)}{p'(x)} - 1 \right)^2 dx$$

Як частина критерію навчання uLSIF, модель коефіцієнта густини підлаштовується під справжній коефіцієнт густини з урахуванням квадратичних втрат. Апроксиматор дивергенції PE має наступний вигляд:

$$\widehat{PE} = -\frac{1}{2n} \sum_{j=1}^n \hat{g}(Y'_j)^2 + \frac{1}{n} \sum_{i=1}^n \hat{g}(Y_i) - \frac{1}{2}$$

Залежно від умови густини другого інтервалу $p'(x)$, значення коефіцієнта густини може бути необмеженим. Щоб подолати цю проблему, α - відносна дивергенція PE для $0 \leq \alpha < 1$ використовується як міра несхожості в підході, відомому як відносний uLSIF (RuLSIF). Міра RuLSIF така:

$$PE_{\alpha}[p(x)||p'(x)] = PE(p(x)||\alpha p(x) + (1 - \alpha)p'(x))$$

α - відносне співвідношення густини зводиться до звичайного співвідношення густини, якщо $\alpha = 0$, і це, як правило, "плавніше", оскільки α стає

більшим. Новизна RuLSIF та, що він завжди обмежений зверху $\frac{1}{\alpha}$, і було показано, що швидкість збіжності для оцінки коефіцієнта відносної густини швидша, ніж у uLSIF [18].

Нещодавно було запропоновано напівпараметричний логарифмічний детектор змін правдоподібності (SPLL) [19] як напівпараметричний детектор змін на основі статистики Кульбака-Лейблера. Припустимо, що дані до точки зміни (вікно W^1) походять з гаусової суміші $P_i(x)$. Критерій виявлення змін виводиться за допомогою верхньої межі логарифмічної правдоподібності даних у другому вікні W^2 , використовуючи індекс компонента з найменшим квадратом відстані Махаланобіса між x та його центром. Якщо W^2 не походить з того ж розподілу W^1 , то середнє значення відстаней відхилятиметься від n (де n - розмірність простору ознак). Значення SPLL, яке більше або менше за заданий діапазон, вказуватиме на зміну. Важливо зазначити, що точність усіх цих методів оцінки знижується шумом даних.

2.2.4 Методи моделювання підпростору.

Інший напрямок досліджень передбачає виявлення точок зміни на аналізі підпросторів, у яких послідовності часових рядів обмежені. Цей підхід має тісний зв'язок з методом ідентифікації системи.

Один з таких методів моделювання підпростору називається ідентифікацією підпростору (ІП) [20]. ІП базується на моделі простору станів системи, яка також явно враховує фактор шуму.

$$\begin{aligned}x(t+1) &= Ax(t) + Ke(t) \\ y(t) &= Cx(t) + e(t)\end{aligned}$$

Тут C та A – системні матриці, $e(t)$ представляє системний шум, а K – стаціонарний коефіцієнт підсилення Калмана. У методах підпростору

використовують різні позначення. Оскільки в цих методах x представляє стани моделі, використовують u як часовий ряд.

В ідентифікації системи розширена матриця спостережуваності – це міра того, наскільки добре внутрішні стани ($x(t)$) системи можна вивести, знаючи її зовнішні виходи ($y(t)$). Тут використовують розширену матрицю спостережуваності як представлення підпростору, в якому дані часових рядів обмежені.

Розширена матриця спостережуваності визначається як:

$$O_k = [C^T \quad (CA)^T \quad \dots \quad (CA^{k-1})^T]$$

Для кожного інтервалу SI оцінює матрицю спостережуваності, використовуючи LQ-факторизацію та розкладання за сингулярними значеннями (SVD) нормалізованої умовної коваріації. LQ-факторизація - це ортогональний розклад матриці на нижні трапецієподібні матриці. SVD матриці A - це розкладання A на добуток трьох матриць $A = UDV^T$, де стовпці U та V ортонормовані, а матриця D є діагональною з додатними дійсними елементами. На наступному кроці обчислюється проміжок між підпросторами та використовується як міра зміни в послідовності часового ряду. Цю міру зміни, D , можна порівняти із заданим порогом, щоб визначити, чи є поточна точка точкою зміни.

$$D = X^T X - X^T U U^T X$$

Тут x представляє матрицю Генкеля нового інтервалу, а U обчислюється за допомогою SVD оціненої розширеної матриці спостережуваності для попереднього інтервалу.

Наступний метод моделі підпростору називається перетворенням сингулярного спектру (SST), яке також базується на моделі простору станів. Однак, на відміну від моделі SI, вона не враховує системний шум. SST визначить матрицю траєкторії на основі матриці Генкеля для кожного вікна, як показано в наступному рівнянні:

$$X = \begin{pmatrix} y_1 & y_2 & \dots & y_K \\ y_2 & y_3 & \dots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \dots & y_N \end{pmatrix}$$

де L – довжина вікна, а K – кількість вікон. Матрицю траєкторій можна розкласти на підматриці за допомогою SVD. Ці підматриці складаються з емпіричних ортогональних функцій сингулярного значення, або функцій EOF, та головних компонент. Оцінки точок зміни на основі відстані визначаються шляхом порівняння сингулярних спектрів двох матриць траєкторій для послідовних інтервалів.

Хоча обидва ці методи моделювання підпростору базуються на заздалегідь визначеній моделі, SST не враховує вплив шуму на систему. Як результат, він чутливіший, ніж SI, до вибору значень параметрів і продемонстрував нижчу точність для деяких наборів даних [20].

2.2.5 Імовірнісні методи.

Ранні байєсівські підходи до виявлення точок зміни були неактуальними. (∞ — real time) та базувалися на ретроспективній сегментації [21]. Один з перших підходів до онлайн-байєсівського виявлення точок зміни (BCPD) був запропонований виходячи з припущення, що послідовність спостережень може бути розділена на неперекриваючі секції станів, а дані в кожному стані в часових рядах є незалежно одинично розрізненими з деяким розподілом ймовірностей. $P(x_t|\eta_p)$.

Порівняно з попередніми методами, які розглядали лише пари послідовних вибірок, VCPD порівнює нові ознаки ковзного вікна з оцінкою, заснованою на всіх попередніх інтервалах з того самого стану. VCPD оцінює апостеріорний розподіл, визначаючи допоміжну змінну довжина пробігу (r_t), яка представляє час, що минув з останньої точки зміни. Враховуючи довжину пробігу в момент часу t , довжина пробігу в наступний момент часу може або скинутися до 0 (якщо точка зміни відбувається в цей момент часу), або збільшитися на 1 (якщо поточний стан триває ще одну одиницю часу). Розподіл довжини пробігу на основі теореми Байєса можна позначити як:

$$P(r_t | x_{1:t}) = \frac{\sum_{r_{t-1}} P(r_t | r_{t-1}) P(x_t | r_{t-1}, x_t^{(r)}) P(r_{t-1}, x_{1:t-1})}{\sum_{r_t} P(r_t, x_{1:t})}$$

де $x_t^{(r)}$ вказує на набір спостережень, пов'язаних з пробігом r_t та $P(r_t | r_{t-1})$, $P(x_t | r_{t-1}, x_t^{(r)})$, і $P(r_{t-1}, x_{1:t-1})$ є апіорною, правдоподібною та рекурсивною компонентами рівняння. Умовна апіорна оцінка не дорівнює нулю лише для двох результатів ($r_t = 0$ or $r_t = r_{t-1} + 1$) і спрощує рівняння.

$$P(r_t | r_{t-1}) = \begin{cases} H(r_{t-1} + 1) & \text{if } r_t = 0 \\ 1 - H(r_{t-1} + 1) & \text{if } r_t = r_{t-1} + 1 \\ 0 & \text{otherwise} \end{cases}$$

У цьому рівнянні, $H(\tau) = \frac{P(\tau)}{\sum_{t=\tau}^{\infty} P(t)}$ — це функція ризику, яка визначається як відношення щільності ймовірності за пробіг до загального значення щільності ймовірності [22]. Член правдоподібності представляє ймовірність того, що найновіша дата належить поточному пробігу. Це найскладніший член для обчислення, і він, як правило, є найбільш обчислювально ефективним, коли використовується спряжена експоненціальна модель.

Після обчислення розподілу довжини пробігу та оновлення відповідної статистики виконується прогнозування точки зміни шляхом порівняння значень ймовірності. Якщо rt має найвищу ймовірність у розподілі, то відбулася точка зміни, і довжина пробігу скидається до $rt = 0$. Якщо ні, довжина пробігу збільшується на одиницю, $rt = rt_{\pm} + 1$.

Пізніше цей метод був поширений на загальний випадок неідентифікованих часових рядів шляхом включення ймовірності різних підпоследовностей даних. Крім того, було введено спрощення, яке зменшує складність алгоритму з n^2 до n за допомогою простого наближення. Ключова ідея полягає в обчисленні спільних вагових коефіцієнтів ймовірності лише для фіксованої кількості вузлів, замість того, щоб обчислювати ці ваги взагалі. $\frac{n(n-1)}{2}$ вузли.

Гаусівський процес (ГП) являє собою ще один ймовірнісний метод для аналізу та прогнозування стаціонарних часових рядів [23]. ГП є узагальненням гаусового розподілу та визначається як сукупність випадкових величин, будь-яка скінченна кількість яких має спільний гаусівський розподіл [24]. У цьому методі спостереження часових рядів $\{x_t\}$ визначаються як зашумлена версія значень гаусової функції розподілу.

$$x_t = f(t) + \epsilon_t$$

У цій функції гаусового розподілу ϵ_t є шумовим членом, який зазвичай вважається гаусовим шумовим членом. $\mathcal{N}(0, \sigma_n^2)$ and $f(t) = \mathcal{GP}(0, K)$ — це функція розподілу GP, що визначається нульовим середнім значенням та коваріаційною функцією K . Зазвичай коваріаційна функція визначається за допомогою набору гіперпараметрів. Широко використовувана коваріаційна функція — це:

$$K(t_1, t_2) = \sigma^2 \exp\left(-\frac{(t_1 - t_2)^2}{2l^2}\right)$$

Враховуючи часовий ряд, функцію GP можна використовувати для прогнозування нормального розподілу в момент часу t . Алгоритм зміни GP використовує гаусівський процес для оцінки прогнозного розподілу в момент часу t , використовуючи спостереження, доступні протягом часу $(t - 1)$. Потім алгоритм обчислює p -значення для фактичного спостереження y_t за еталонного розподілу, $\mathcal{N}(\hat{y}_t, \hat{\sigma}_t^2)$. Поріг $\alpha \in (0,1)$ використовується для визначення, коли фактичне спостереження не відповідає прогнозованому розподілу, що вказує на можливу зміну стану (і, отже, на точку зміни). Використання спостережень, доступних протягом часу $t-1$, для виявлення точок зміни замість використання лише спостережень з останнього стану робить метод GP складнішим і водночас точнішим, ніж BCPD.

2.2.6 Методи на основі ядра

Хоча методи на основі ядра зазвичай використовуються як методи навчання з учителем, деякі дослідження використовують неконтрольовану тестову статистику на основі ядра для перевірки однорідності даних у часових рядах минулих та теперішніх ковзних вікнах. Ці методи відображають спостереження у відтворюючому просторі ядра Гільберта (RKHS). $\mathcal{H}$ пов'язане з відтворюючим ядром $k(\cdot, \cdot)$ та картою ознак $O(X) = k(X, \cdot)$. Потім вони використовують тестову статистику, засновану на дискримінантному відношенні ядра Фішера, як міру однорідності між вікнами.

Розглядаючи два вікна спостережень, емпіричні середні елементи та оператори коваріації для вибірки X довжиною n обчислюються так:

$$\hat{\mu} = \frac{1}{n} \sum_{l=1}^n k(X_l, \cdot)$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{l=1}^n \{k(X_l, \cdot) - \hat{\mu}\} \otimes \{k(X_l, \cdot) - \hat{\mu}\}$$

де оператор тензорного добутку $u \otimes v$ для всіх функцій $f \in \mathcal{H}$ визначено як. $(u \otimes v)f = \langle v, f \rangle_{\mathcal{H}} u$ Тепер дискримінантне відношення Фішера ядра (KFDR) між двома вибірками визначається як [25]:

$$KFDR(X_1^{length\ n_1}, X_2^{length\ n_2}) = \frac{n_1 n_2}{n_1 + n_2} \langle \hat{\mu}_2 - \hat{\mu}_1, (\hat{\Sigma}_w + \gamma I)^{-1} (\hat{\mu}_2 - \hat{\mu}_1) \rangle_{\mathcal{H}}$$

де γ – параметр регуляризації, а

$$\hat{\Sigma}_w = \frac{n_1}{n_1 + n_2} \hat{\Sigma}_1 + \frac{n_2}{n_1 + n_2} \hat{\Sigma}_2$$

Найпростіший спосіб визначити, чи існує точка зміни між двома вікнами – це порівняння коефіцієнта KFDR з пороговим значенням [25]. Інший метод, відомий як запуск стратегії максимального розподілу, обчислює коефіцієнт KFDR між усіма наступними вікнами в кожному інтервалі. Потім максимальне значення порівнюється з порогом для виявлення точки зміни.

Загальним недоліком методів на основі ядра є те, що вони сильно залежать від вибору функції ядра та її параметрів.

2.2.7 Методи на основі графів

Кілька недавніх досліджень показали, що часові ряди можна досліджувати за допомогою інструментів теорії графів. Граф зазвичай виводиться з відстані або узагальненої відмінності у просторі вибірки, при цьому спостереження часових рядів у вигляді вузлів та ребер з'єднують спостереження на основі їхньої відстані. Цей граф можна визначити на основі мінімального охоплюючого дерева, парування мінімальної відстані [26], графа найближчого сусіда або графа видимості.

Графовий фреймворк для виявлення точок зміни – це непараметричний підхід, який застосовує двовибірковий тест на еквівалентному графіку, щоб

визначити, чи є точка зміни в межах спостережень. У цьому методі граф G будується для кожної послідовності даних. Кожне можливе значення t як часу точки зміни поділяє спостереження на два вікна: спостереження, що передують t , та спостереження, після t . Кількість ребер у графі G (R_G), що з'єднує спостереження з цих двох вікон, використовується як індикатор точки зміни, тому менші ребра збільшують ймовірність точки зміни. Оскільки значення R_G залежить від часу t , стандартизована функція (Z_G) визначається як:

$$Z_G(t) = -\frac{R_G(t) - E[R_G(t)]}{\sqrt{\text{VAR}[R_G(t)]}}$$

де $E[.]$ та $\text{VAR}[.]$ – це математичне очікування та дисперсія відповідно. Максимальне значення Z_G серед усіх точок даних на графі визначається як точка-кандидат зміни. Точка зміни приймається, якщо максимуми перевищують заданий поріг. Цей метод є потужним для високорозмірних даних з меншою кількістю припущень щодо параметрів. Однак він не використовує багато інформації з самих спостережень часових рядів, натомість спираючись на визначення відповідної структури графіка.

2.2.8 Методи кластеризації

З іншої точки зору, проблему виявлення точки зміни можна розглядати як проблему кластеризації з відомою або невідомою кількістю кластерів, таким чином, що спостереження всередині кластерів були розподілені однаково, а спостереження між сусідніми кластерами – ні. Якщо точка даних у момент часу t належить до іншого кластера, ніж точка даних у момент часу $t+1$, тоді між двома спостереженнями виникає точка зміни.

Один підхід до кластеризації, що використовується для виявлення точки зміни, поєднує методи ковзного вікна та метод «знизу вгору» в алгоритм під назвою SWAB (Sliding Window and Bottom-up) [27]. Оригінальний підхід «знизу

вгору» спочатку обробляє кожну точку даних як окрему підпоследовність, а потім об'єднує підпоследовності з відповідною вартістю об'єднання, доки не буде виконано критерії зупинки. На противагу цьому, SWAB підтримує буфер розміром w для зберігання достатньої кількості даних для 5-6 підпоследовностей. Метод «знизу вгору» застосовується до даних у буфері, і виводиться результуюча підпоследовність, що залишилася ліворуч. Дані, що відповідають виведеній підпоследовності, видаляються з буфера та замінюються наступними даними в последовності.

Другий підхід кластеризації групує підпоследовності на основі мінімальної довжини опису. Довжина опису DL часового ряду T довжиною m – це загальна кількість бітів для подання последовності, або:

$$DL(T) = m * H(T)$$

де $H(T)$ – ентропія часового ряду.

Виявлення точок зміни на основі MDL – це висхідний пошук у просторі кластерів, який може включати підпоследовності різної довжини та не вимагає вказівки кількості кластерів.

Інший спосіб кластеризації даних часових рядів як спосіб пошуку точок зміни - метод Shapelet. Неконтрольований шейплет, або U-шейплет S , – це невеликий шаблон у часовому ряді T , для якого відстань між S та частиною часового ряду значно менша, ніж відстань між S та рештою часового ряду. Кластеризація на основі шейплетів, яка намагається кластеризувати дані на основі форми всього часового ряду, шукає U-шейплет, який може відокремити та видалити підпоследовність часового ряду від решти набору даних. Алгоритм ітеративно повторює цей пошук серед даних, що залишилися, доки не залишиться даних для розділення. Для вилучення U-шейплетів використовується алгоритм пошуку, який намагається максимізувати проміжок розділення між

двома підмножинами даних. Потім будь-який алгоритм кластеризації, такий як k-середні з евклідовою функцією відстані, може бути використаний для кластеризації часового ряду та пошуку точок змін.

Ще один підхід до кластеризації часових рядів – це підгонка моделі, в якій зміна вважається такою, що відбувається, коли новий елемент даних або блок елементів даних не вписується в жодний з існуючих кластерів. Припускаючи, що потік даних $\{x_1, \dots, x_i, \dots\}$. Точка зміни відбувається після точки даних x_t , якщо наступний логічний вираз є істинним:

$$change = \bigwedge_{j=1}^K \left[d(x_{i+1}, center(C_j)) > radius(C_j) \right]$$

де $d(x_{i+1}, center(C_j))$ – евклідова відстань між нововведеною точкою даних x_{i+1} та центром кластера C_j , $radius(C_j)$ – радіус кластера j , K – кількість кластерів, а $\bigwedge$ – логічний символ. Радіус кластера C з n точками даних та середнім значенням μ дорівнює:

$$radius(C) = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}$$

2.3 Порівняння методів

Порівнюємо методи CPD на основі кількох часто використовуваних критеріїв.

2.3.1 Онлайн проти офлайн.

Одним із важливих критеріїв для виявлення точок зміни є здатність ідентифікувати точку зміни в реальному часі або майже в реальному часі. Повні офлайн-алгоритми застосовні при обробці всього часового ряду одночасно, а s —

алгоритми реального часу повинні переглядати щонайменше ϵ точок даних перед виділенням точки зміни. Значення ϵ залежить від характеру алгоритму та кількості вхідних даних, необхідних для кожного кроку. Онлайн-алгоритми обробляють дані в межах ковзного вікна розміром n . Для цих підходів n має бути достатньо великим, щоб зберігати дані, необхідні для представлення стану часового ряду, але водночас достатньо малим, щоб все ще відповідати вимозі ϵ .

Методи з наглядом. Після обробки достатньої кількості навчальних даних ці методи передбачатимуть, чи є СР у поточному вікні. Тому можна стверджувати, що методи з наглядом працюють у n -реальному часі.

Методи відношення правдоподібності. Ці методи базуються на порівнянні щільностей ймовірності між двома послідовними інтервалами. Коли з'являється нова ретроспективна підпоследовність, нове обчислення поверне результат, тому можна сказати, що ці методи є $n+k$ -реального часу.

Модель підпростору. Нові інтервали в цих методах обчислюються так само, як і для методів правдоподібності. Як результат, ці методи також є $n+k$ -реального часу.

Ймовірнісні методи. Ці методи спираються лише на одне ковзне вікно для виявлення СР, тому вони є методами n -реального часу.

Методи на основі ядра. Методи ядра без учителя базуються на ковзних вікнах. Однак, як і методи відношення правдоподібності, вони потребують ретроспективної підпоследовності даних, тому вони є методами $n+k$ -реального часу.

Кластеризація. Метод SWAB є комбінацією ковзного вікна та методу «знизу вгору». SWAB підтримує буфер розміром w . Метод «знизу вгору» застосовується до даних у буфері. Як результат, SWAB є методом w -реального часу. Методи на основі MDL та Shapelet потребують одночасного доступу до всього часового ряду, тому вони працюють офлайн або в нескінченному

реальному часі. Метод апроксимації моделі залежить від одного вікна і тому працює в n -реальному часі.

Метод на основі графів. Цей метод створює графік з одного вікна. Точка зміни ідентифікується, якщо вона існує в поточному вікні, таким чином метод працює в n -реальному часі.

2.3.2 Масштабованість.

Другим важливим критерієм є обчислювальна потужність алгоритмів виявлення точок зміни. Загалом, зі збільшенням розмірності часового ряду непараметричні методи збільшують обчислювальну потужність і будуть менш дорогими, ніж параметричні методи. Дуже важко охарактеризувати потужність методів з наглядом, оскільки є дві складності. Це час виконання етапу навчання та час виконання етапу виявлення точок зміни.

2.3.3 Обмеження навчання.

Більшість методів відношення правдоподібності (крім SPLL) та всі методи моделі підпростору спочатку були розроблені для одновимірних часових рядів. Таким чином, у випадку d -вимірного часового ряду ці методи об'єднують усі виміри разом і генерують одновимірний ряд з вектором значень розміру d . Хоча для інших алгоритмів немає обмежень на розмірність часових рядів, збільшення кількості вимірів збільшить обчислювальні витрати алгоритму.

Усі алгоритми приймають як дискретні, так і неперервні вхідні дані часових рядів. Одним винятком є метод на основі MDL, який працює лише з дискретними вхідними значеннями.

Підходи до навчання з учителем для CPD працюють виходячи з припущення, що перехідний період можна виявити незалежно від поточного стану часового ряду. На відміну від цього, алгоритм навчання без учителя працює виходячи з припущення, що розподіл даних часових рядів змінюється до та після кожної точки зміни. Хоча дані з учителем часто перевершують методи без учителя у виявленні точок зміни, вони залежать від достатньої якості та кількості

навчальних даних, які не завжди доступні для реальних даних. Багатокласові алгоритми з учителем є єдиною групою, якій потрібно знати кількість можливих станів часових рядів.

Загалом, непараметричні методи CPD є більш надійними, ніж параметричні, оскільки параметричні методи значною мірою залежать від вибору параметрів. Крім того, проблема CPD стає складнішою для параметричних методів, коли дані мають від середньої до високої розмірності.

Більшість алгоритмів CPD без учителя працюють з обмеженими типами даних часових рядів. Деякі з них працюють лише для стаціонарних або незалежно розмежованих наборів даних, а інші пропонують параметричні версії для нестаціонарних наборів даних часових рядів. Відповідні параметричні версії використовують коефіцієнт забування, щоб усунути вплив старих спостережень.

2.3.4 Оцінка продуктивності.

Для вимірювання продуктивності алгоритмів CPD було використано кілька штучних та реальних наборів даних. Важливо зазначити, що об'єктивне порівняння продуктивності різних методів CPD дуже складне через використання цих різних наборів даних.

Більшість досліджень не надають жодних порівнянь, а в деяких випадках навіть показників продуктивності. Наприклад, немає доступних результатів для методів SPLL та кластеризації. Аналогічно, експериментальні результати для CPD на основі графів доступні лише для різних структур графів, щоб продемонструвати той факт, що точність сильно залежить від структури графа. Дослідження, що включають аналіз продуктивності, як правило, обчислюють відстань між фактичними та виявленими CP та використовують дискретні показники, такі як точність, прецизійність та повнота, для оцінки алгоритмів.

Таблиця 2.1

Порівняння ефективності методів виявлення розладки
(тип сигналу × тип розладки)

Тип сигналу	Тип розладки	Найефективніші методи (рекоменд.)	Чому ефективні (коротко)	Обмеження/коли гірше	Типова вихідна метрика/ статистика	Оцінка складності	Очікувана точність локалізації меж
Стационарний майже i.i.d.	Стрибок середнього (mean shift)	CUSUM; GLR-CUSUM; (офлайн) PELT/DP piecewise-constant	Модельно-оптимальні; мала затримка при заданих хибних тривогах	Автокореляції/викиди → хибні тривоги без корекції	$\ell_t, S_t = \max(0, S_{t-1} + \ell_t);$ (офлайн) $\sum (x_t - \bar{x})^2 + \beta m$	Онлайн $O(T)$; офлайн PELT ~ $O(T)$ (часто), DP $O(T^2)$	Висока (офлайн), середня (онлайн, залежить від порогу)
Стационарний майже i.i.d.	Малий стійкий зсув середнього	CUSUM/GLR-CUSUM; (офлайн) PELT	CUSUM особливо чутливий до малих змін	Shenhart погано для малих зсувів	Як вище	Онлайн $O(T)$; PELT ~ $O(T)$	Висока (офлайн), середня (онлайн)
Стационарний/повільно нестационарний	Плавний дрейф	EWMA; BOSCPD; Page-Hinkley	Згладжування/ймовірнісне накопичення краще для дрейфів	EWMA з малим $\lambda \rightarrow$ більша затримка; BOSCPD складніше налаштувати	$y_t = \lambda x_t + (1 - \lambda)y_{t-1}; \forall t: t=0$	$\chi_{(100)}$	EWMA/PH $O(T)$; BOSCPD типово $O(T \cdot R)$ (з обрізанням R)
Трендові ряди (сенсори/телеметрія)	Зміна тренду/нахилу	GLR (лінійна модель); (офлайн) PELT/DP piecewise-linear; Page-Hinkley	Сегментні моделі добре локалізують злам	При високому шумі потрібні довші вікна	$C = \sum (x_t - (at + b))^2 + \beta m$	PELT часто ~ $O(T)$; DP $O(T^2)$; GLR віконний $O(Tw)$	Середня-висока (офлайн), середня (онлайн)
Автокорельовані (AR/стан-простір)	Зміна динаміки/кореляцій	GLR(AR/state-space); Switching Kalman; NMM/BOSCPD	Ураховування залежностей зменшує хибні спрацювання	Потрібна ідентифікація/підбір моделі	$\log L$ для AR/SSM; постеріор режимів	Залежить від моделі: типово $O(Tp^2)$ або $O(TK^2)$	Середня-висока (офлайн), середня (онлайн)
Ритмічні/квазіперіодичні (монофонія)	Зміна частоти/періоду (pitch)	Гармонічна сегментна модель + PELT/DP; або $f_0(t) + \text{PELT}$	Частота визначена → точна локалізація	Лемається при поліфонії/перкусії	$C(s : e) = \min_f \text{RSS}(f); \text{CPD на } \hat{f}_0(t)$	Гармонічний пошук по ступі: $(O \cap \text{cdot})$	$\backslash \text{mathcal{F}}$
Поліфонічний аудіосигнал	Гармонія/набір висот (акорди/тональність)	Chroma/nRSPR → cosine distance/novelty → PELT; SSM+novelty	Стійке до тембру/гучності, добре в поліфонії	Не бачить "тембрових" змін без гармонічної різниці	$d_t = 1 - \frac{\langle e_t, e_{\text{ref}} \rangle}{\ e_t\ \ e_{\text{ref}}\ }; \text{ novelty з SSM}$	Ознаки: $O(Td)$; PELT ~ $O(T)$; SSM наївно $O(T^2)$ (можна прискорювати)	Висока (для гармонічних меж), середня (для дрібних змін)
Поліфонічний аудіосигнал	Зміна спектральної фактури/регістра	SQT/log-мел → $L_2/KL/JIS$ відстані → PELT	Універсально ловить перерозподіл енергії по частотах	Чутливий до гучності/міксу → потрібна нормалізація	$d_t = \ \hat{s}_L - \hat{s}_R\ _2$ або D_{KL}	Ознаки: STFT $O(T \log N)$; дистанції $O(Td)$; PELT ~ $O(T)$	Середня-висока (залежить від вікна), зазвичай середня
Поліфонічний аудіосигнал	Зміна темпу/beat (період=brp)	Onset → tempo/beat $\widehat{BPM}(t) \rightarrow \text{PELT/WBS}$	Пряма робота з періодом	Погано при нечітких атаках, нестійкий onset	$\widehat{BPM}(t) = \frac{w(t)}{\tau(t)}$	Onset+tempogram: $O(T \log T)$ або $O(Tw)$; PELT ~ $O(T)$; WBS $O(T \log T)$ (типово)	Середня (межа розміта в ритмі), інколи середня-низька
Широкосмугові/подієві аудіо	Частотно-часові режими	Вейвлет/спектрограма → енергія в смугах → PELT; або MMD/energy	Вейвлет ловить локальні TF-зміни; MMD робастний	Вибір масштабу/вікна критичний; MMD важкий	$E_0(t); \text{MMD}^2$	Вейвлет $O(T^2 - O(T \log T))$; MMD наївно $O(Tw^2)$	Середня (залежить від масштабу)
Дані з викидами/важкими хвостами	Будь-яка	Rank/KS; MMD/energy; робастні Bayesіan; ознаки+робастні порого	Менше хибних тривог на "брудних" даних	Можє бути більша затримка/менша точність меж	Rank/KS; MMD^2 ; робастні оцінки	Rank/KS $O(Tw \log w)$ (віконно); MMD $O(Tw^2)$; робастні моделі залежать від реалізації	Середня (часто), інколи середня-низька

Підсумовуючи, варто зазначити, що методи з учителем, як правило, точніші, ніж методи без учителя, якщо існує достатньо навчальних даних, а ряд є стаціонарним. Якщо ці умови не виконуються, методи без учителя є кориснішими. Немає комплексного порівняння продуктивності між методами без учителя, але з експериментальних результатів видно, що RulSIF постійно забезпечує високу точність. Оскільки методи на основі ядра, моделі підпростору, CUSUM, AR та методи кластеризації спираються на параметри для моделювання динаміки часових рядів, вони не демонструють хорошої продуктивності для шумних даних або високодинамічних систем.

Більшість алгоритмів без учителя накладають обмеження на типи методів часових рядів, які можна обробляти. Одним помітним винятком є метод AR. Крім того, деякі з цих методів мають параметричні версії для нестационарних даних, що робить їх чутливими до вибору параметрів. Для високовимірних часових рядів моделі відношення правдоподібності та підпростору не є найкращим вибором, оскільки вони не можуть безпосередньо обробляти багатовимірні дані. У цьому випадку більш перспективними є графові або ймовірнісні методи.

2.4 Висновки до розділу 2

Показано, що одним з важливих питань для алгоритмів виявлення часових моментів змін ймовірнісних характеристик є потреба в розробленні алгоритмів із роботою в режимі реального часу та з мінімальними затримками. У багатьох реальних застосуваннях точки змін використовуються для вибору та виконання своєчасних дій, тому якомога швидше знайти точки змін є критично важливим. Альтернативою розглянутих алгоритмів є використання методів, які потребують менших розмірів вікон для розрахунку часових точок змін, таких як байєсівські методи.

Ще однією відкритою проблемою є стійкість алгоритму. Встановлено, що майже для всіх методів виявлення змін для опрацювання використовується метод ковзного вікна. Однак, практично немає обґрунтування вибору критерію, за яким вибирається поріг розрізнення ділянок сигналів та часових рядів для виявлення часових моментів появи змін.

РОЗДІЛ 3

НАУКОВО-ДОСЛІДНА ЧАСТИНА

3.1 Модель тестового сигналу для виявлення змін ймовірнісних характеристик

Стрибкоподібні зміни ймовірнісних характеристик сигналів, які відбуваються в невідомі моменти часу, можна означити терміном «розладка», і як математичну модель сигналів доцільно використати кусково стаціонарний випадковий процес.

Кусково стаціонарний випадковий процес (piecewise stationary / locally stationary в “changepoint” сенсі) — це випадковий процес, який не є стаціонарним на всій осі часу, але існує розбиття часу на скінченну кількість інтервалів, на кожному з яких процес стаціонарний (зазвичай у строгому або широкому сенсі), а між інтервалами параметри/закон змінюються.

Нехай $\{X_t\}_{t \in \mathbb{Z}}$ — процес. Існують моменти зміни

$$0 = \tau_0 < \tau_1 < \dots < \tau_m < \tau_{m+1} = T,$$

такі що для кожного сегмента $j = 0, \dots, m$ процес на інтервалі

$$t \in (\tau_j, \tau_{j+1}]$$

є стаціонарним: тобто його скінченновимірні розподіли інваріантні до зсуву всередині сегмента. Еквівалентно:

$$(X_{t_1}, \dots, X_{t_k}) \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_k+h})$$

для всіх k , всіх наборів t_i і всіх зсувів h , які не виводять індекси за межі сегмента.

Практично частіше використовують широкую стаціонарність на сегменті.

Кусково стаціонарність у широкому сенсі (2-й порядок)

Процес кусково стаціонарний у широкому сенсі, якщо на кожному сегменті j існують сталі (для цього сегмента) моменти 1-го та 2-го порядку:

$$\begin{aligned}\mathbb{E}[X_t] &= \mu_j, \quad t \in (\tau_j, \tau_{j+1}], \\ \text{Cov}(X_t, X_{t+h}) &= \gamma_j(h), \quad t, t+h \in (\tau_j, \tau_{j+1}].\end{aligned}$$

Тобто середнє й автоковаріаційна функція залежать від номера сегмента, але не залежать від абсолютного часу всередині сегмента.

Якщо $\gamma_j(h)$ абсолютно підсумовна, можна визначити спектральну щільність сегмента:

$$S_j(\omega) = \sum_{h=-\infty}^{\infty} \gamma_j(h) e^{-i\omega h}.$$

3) Типова модель кусково стаціонарного процесу (через “режими”)

Зручно записувати процес як такий, що змінює параметри моделі у моменти τ_j . Наприклад:

- кусково стаціонарний “білий шум” (змінна μ, σ^2)

$$X_t = \mu_j + \sigma_j \varepsilon_t, \quad t \in (\tau_j, \tau_{j+1}],$$

де ε_t — i.i.d. з нульовим середнім і одиничною дисперсією.

- кусково стаціонарний AR(p)

$$X_t = \sum_{k=1}^p \phi_{j,k} X_{t-k} + \varepsilon_t, \quad t \in (\tau_j, \tau_{j+1}],$$

де $\{\phi_{j,k}\}$ сталі в межах сегмента, але можуть змінюватись між сегментами.

Стаціонарність “порушується” лише у точках зміни τ_j .

У кожному сегменті можна застосовувати інструменти стаціонарного аналізу (оцінка спектра, автокореляції, AR-параметрів тощо).

Задача виявлення розладки — це фактично оцінити τ_j та/або параметри $\mu_j, \gamma_j(h) / \theta_j$.

Відмінність від “локально стаціонарного” процесу у часово-частотному сенсі така: іноді “локально стаціонарний” означає, що параметри змінюються плавно з часом (без стрибків). Кусково стаціонарний — це окремий важливий випадок, де зміни стрибкоподібні і їх можна описати кінцевою кількістю сегментів.

Кусково стаціонарний процес можна записати як:

$$\Xi_n(t) = (\xi_1(t), \xi_2(t), \dots, \xi_n(t))$$

де: $\Xi_n(t)$ — означений на $t \in [a, b]$ процес; $B_k, k = \overline{1, n}$ розділяє відрізок $[a, b]$

точками $t_1, t_2, \dots, t_k$; $I_{B_k}(t) = \begin{cases} 0, & t \notin B_k; \\ 1, & t \in B_k, \end{cases}$. Тоді такий процес:

$$\xi_{\Sigma}(t) = \sum_{k=1}^n \xi_k(t) I_{B_k}(t)$$

означається як процес розладки з часовими точками розладки $t_1, t_2, \dots, t_n$.

Таким чином для виявлення змін імовірнісних характеристик сигналу, а відповідно і розладки, достатньо знайти точки $t_1, t_2, \dots, t_n$ та оцінити на інтервалах між цими точками рівні змін імовірнісних характеристик.

В роботі використано спектрально-кореляційні методи опрацювання, а саме опрацювання проводиться на інтервалах ковзного вікна. На рис. 3.1 показано приклад сигналу із трьома ділянками, де крайні дві ділянки характеризують один і той самий стаціонарний процес, а середня ділянка із збільшеним рівнем характеризує початок та закінчення розладки.

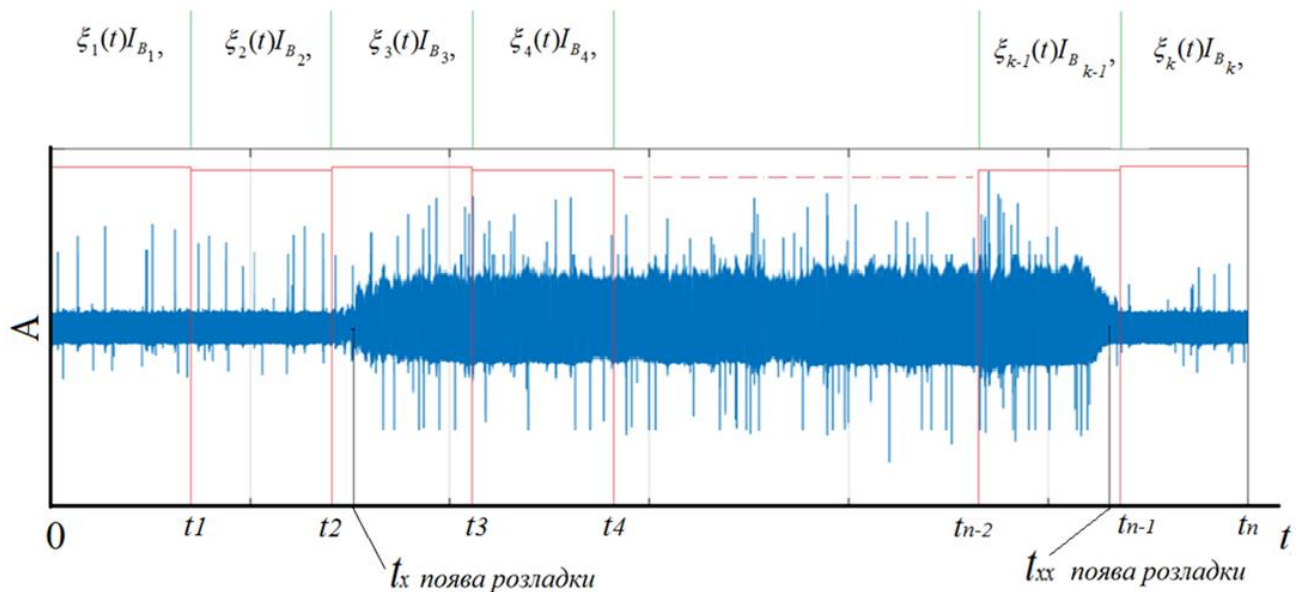


Рис. 3.1. Ілюстрація присутньої розладки сигналу

Таким чином, для виявлення розладки пропонується проводити сигнал в межах ковзного вікна, яке зміщується по сигналу, і всередині якого проводиться спектрально-кореляційна обробка. На основі цієї обробки за сприбкоподібними змінами ймовірнісних характеристик приймається і рішення про появу розладки.

3.2 Результати експериментальних досліджень

На наступному етапі було завантажено тестовий сигнал в Матлаб і проведено його опрацювання. Для цього використано ділянку зашумленого аудіосигналу із підвищенням його рівня після дванадцятої секунди (рис. 3.2).

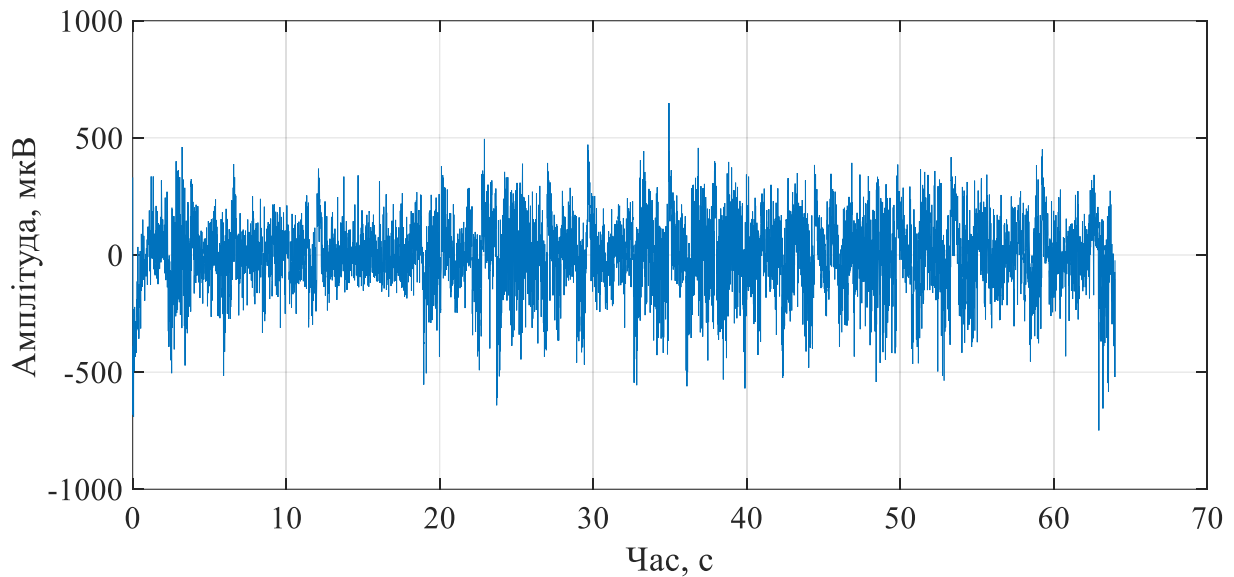


Рис. 3.2. Вигляд ділянки зашумленого аудіосигналу із підвищенням його рівня після дванадцятої секунди

На рис. 3.3 умовно показано сигнал та прямокутниками показано трансляції ковзного вікна.

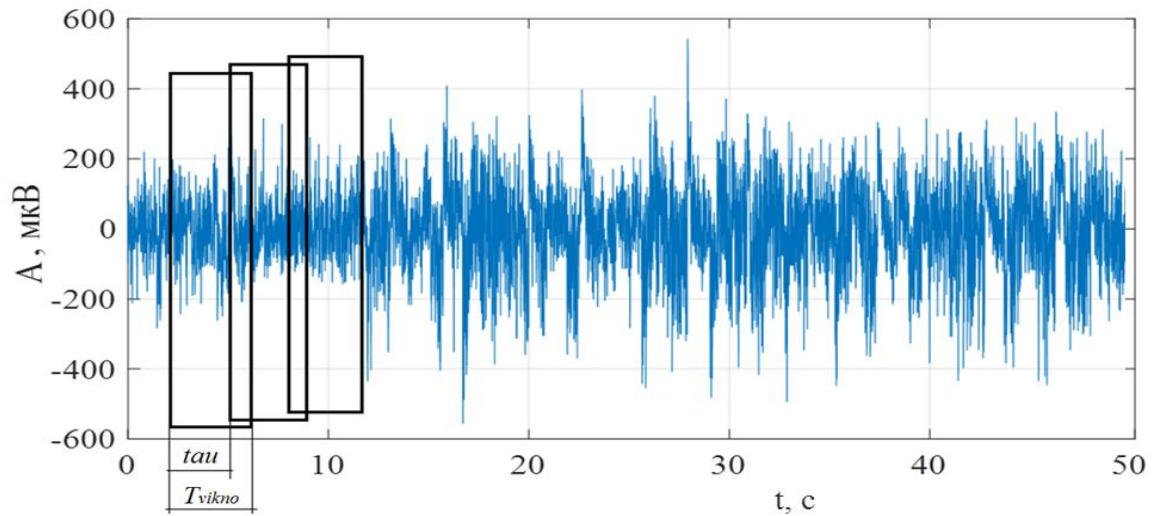


Рис. 3.3. Умовне переміщення ковзного вікна по сигналу

Для виявлення розладки в структурі тестового сигналу в межах кожної трансляції ковзного вікна було обчислено спектр потужності, а далі було проведено його усереднення як за потужністю, так і за часом. В результаті отримувалась точка, яка далі відкладалась на часовій осі відповідно до розміщення відповідного вікна. Це проілюстровано на рис. 3.4.

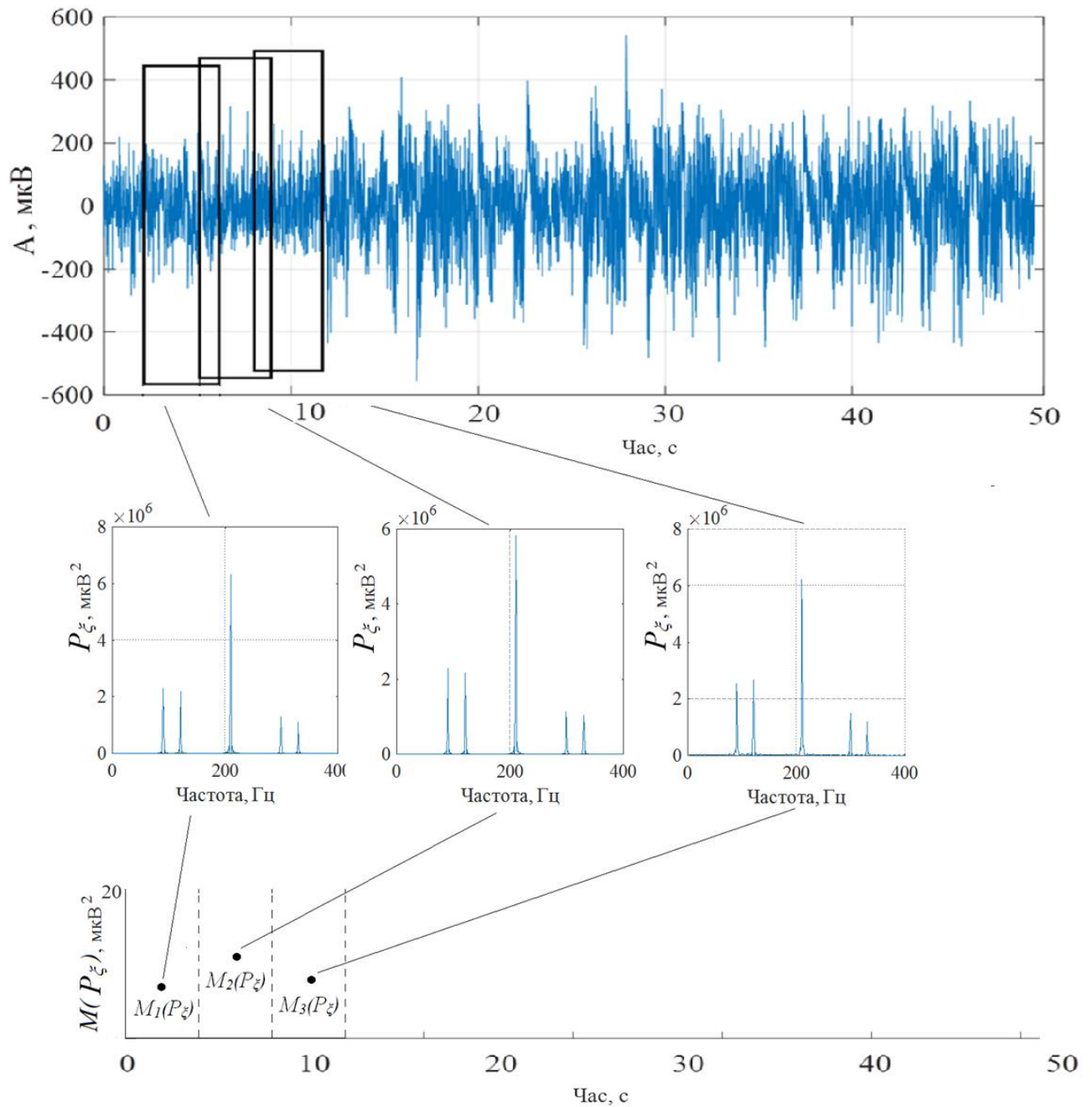


Рис. 3.4. Суть пропонованого методу

На рис. 3.5 показано власне тестовий сигнал з ознаками появи розладки після 12 секунди та графік зміни усередненої в межах трансляцій ковзного вікна потужності.

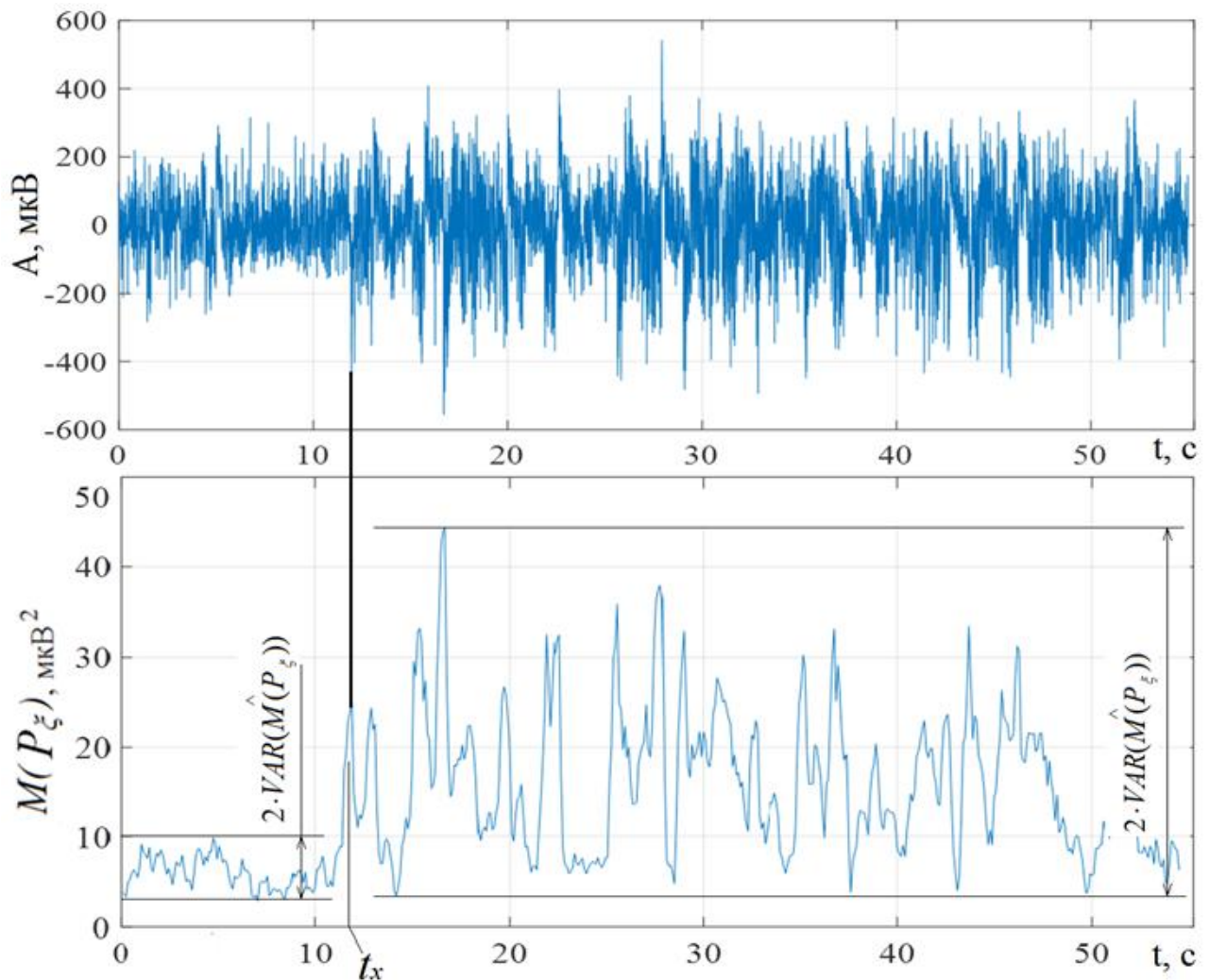


Рис. 3.5. Тестовий сигнал з ознаками появи розладки після 12 секунди та графік зміни усередненої в межах трансляцій ковзного вікна потужності

Також на рис. 3.6 показані оцінки усередненої потужності для двох випадків: вікна не перекриваються (а) та вікна перекриваються на половину ширини вікна (б)

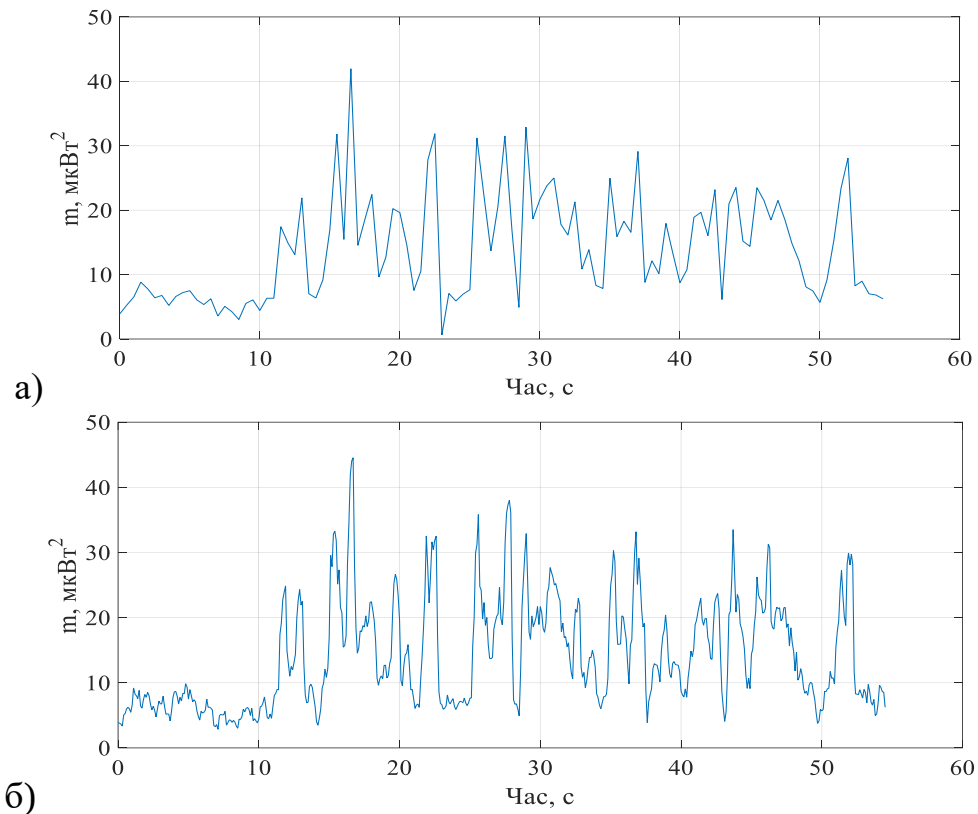


Рис. 3.6. Оцінки усередненої потужності для двох випадків: вікна не перекриваються (а) та вікна перекриваються на половину ширини вікна (б)

З рис. 3.6 можна зробити висновок, що при перекритті вікон точність визначення моменту розладки стає вищою. Максимальною ж точність буде при зсуві вікна на один відлік, але і сам результуючий графік буде дещо розмитим.

Для кількісного визначення часового моменту появи розладки використано поріг, значення якого припускається рівними варіації оцінок усереднених спектрів $VAR(\hat{M}_P)$.

Його значення та відхилення для двох ділянок сигналу до та після розладки є наступні:

$$VAR(\hat{M}_P)_{\text{norm}} = (2,9522 \pm 10\%) \text{ мкВ}^2;$$

$$VAR(\hat{M}_P)_{\text{disruption}} = (83,3164 \pm 10\%) \text{ мкВ}^2$$

В ході експериментальних досліджень встановлено, що оцінки $\hat{VAR}(M_P)$ значно відрізняються між ділянками сигналу до та після виникнення розладки і шляхом моніторингу їхнього значення при трансляції ковзного вікна по сигналу справді можна виявити моменти виникнення розладки в сигналі.

3.3 Алгоритм роботи методу

На рис. 3.5 наведено алгоритм роботи пропонованого методу.

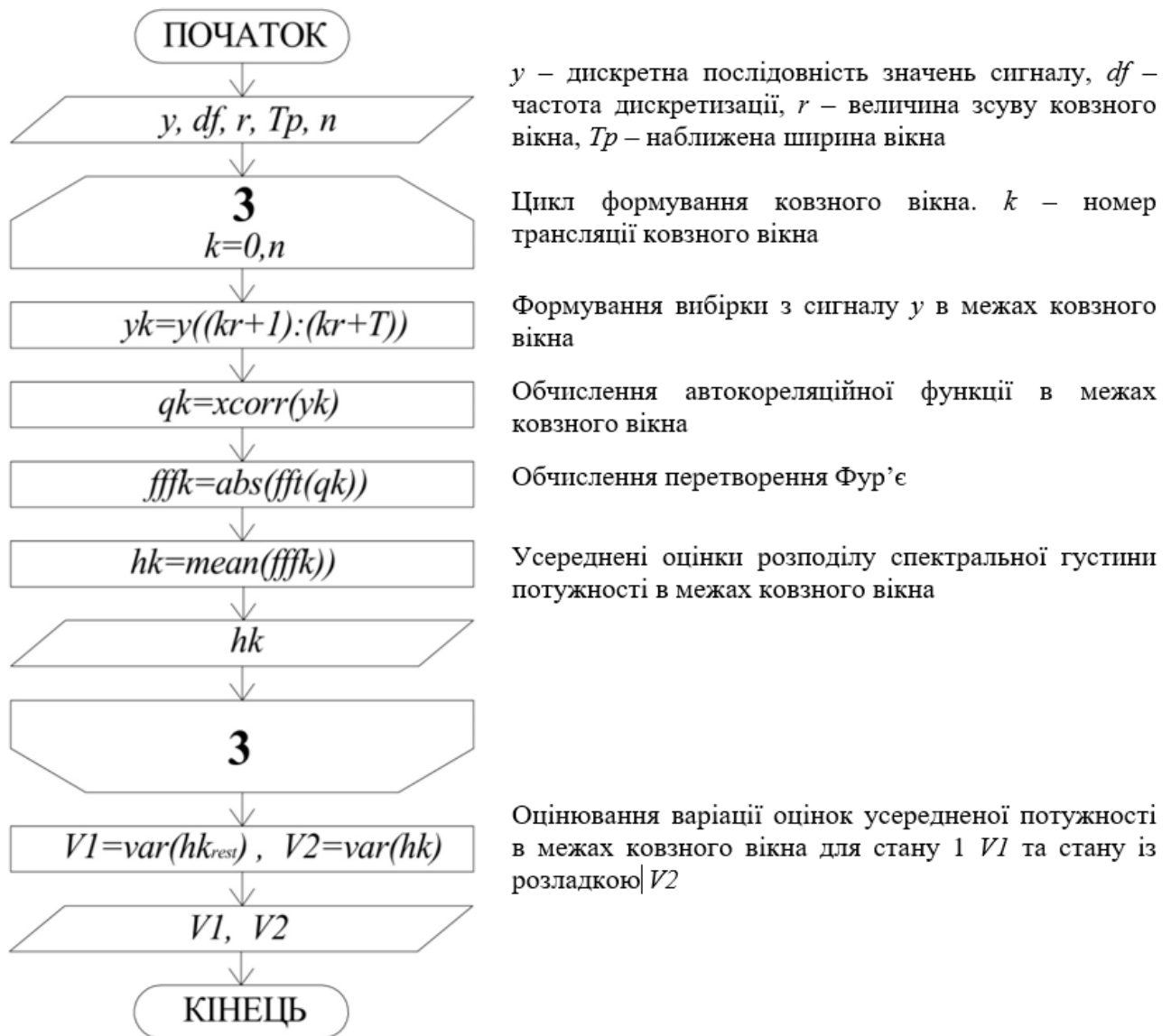


Рис. 3.7. Алгоритм роботи пропонованого методу

Використавши пропонований алгоритм, маючи обчислені значення варіації усередненої потужності для ділянок сигналу без розладки, приймаючи їх за значення порогу, можна виявляти часові моменти розладки.

3.4 Висновки до розділу 3

Оскільки, розладка трактується як стрибкоподібні зміни ймовірнісних характеристик сигналів, в якості ймовірнісної моделі таких сигналів доцільно використати клас стаціонарних процесів, зокрема кусково стаціонарних. Це накладає значні обмеження, оскільки потребує підтвердження стаціонарності сигналів, однак для більшості прикладних задач припущення про стаціонарність на коротких проміжках часу є допустимим.

В роботі використано спектрально-кореляційні методи опрацювання, а саме опрацювання проводиться на інтервалах ковзного вікна.

Таким чином, для виявлення розладки пропонується проводити сигнал в межах ковзного вікна, яке зміщується по сигналу, і всередині якого проводиться спектрально-кореляційна обробка. На основі цієї обробки за стрибкоподібними змінами ймовірнісних характеристик приймається і рішення про появу розладки.

Було завантажено тестовий сигнал в Матлаб і проведено його опрацювання. Для виявлення розладки в структурі тестового сигналу в межах кожної трансляції ковзного вікна було обчислено спектр потужності, а далі було проведено його усереднення як за потужністю, так і за часом. В результаті отримувалась точка, яка далі відкладалась на часовій осі відповідно до розміщення відповідного вікна.

Для кількісного визначення часового моменту появи розладки використано поріг, значення якого припускається рівними варіації оцінок усереднених спектрів.

В ході експериментальних досліджень встановлено, що значення варіації оцінок усереднених спектрів значно відрізняються між ділянками сигналу до та після виникнення розладки і шляхом моніторингу їхнього значення при трансляції ковзного вікна по сигналу можна виявити моменти виникнення розладки в сигналі.

Розроблено алгоритм пропонованого методу, на основі використання якого, маючи обчислені значення варіації усередненої потужності для ділянок сигналу без розладки та приймаючи їх за значення порогу, можна виявляти часові моменти розладки.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Охорона праці

Планування заходів з охорони праці. Види планування та контролю стану охорони праці. Виявлення, оцінка та зменшення ризиків небезпечних подій.

Метою планування заходів з охорони праці є визначення необхідних вкладень у заходи з охорони праці для ефективного впливу на стан охорони праці.

Система планів з охорони праці окремого підприємства може включати:

- перспективне планування (на період, більший одного року) ;
- поточне планування (на рік) ;
- оперативне планування (детальні плани, спрямовані на вирішення конкретних питань працезахоронної діяльності на підприємстві в короткостроковому, до одного року, періоді).

Планування в охороні праці може включати:

- визначення цілей діяльності з охорони праці на підприємстві та засобів їх досягнення;
- вибір методів і базових показників, за допомогою яких може здійснюватися оцінка необхідних вкладень в охорону праці;
- розрахунок суми вкладень у заходи з охорони праці та раціональний розподіл цієї суми за напрямками діяльності;
- забезпечення організації контролю виконання плану (при необхідності здійснення коригування запланованих показників) ;
- здійснення постійного контролю умов і безпеки праці на підприємстві та оперативне реагування на відхилення від нормативних вимог.

Перспективне планування вміщує найбільш важливі, трудомісткі і довгострокові за терміном виконання заходи з охорони праці, виконання яких, як

правило, вимагає сумісної роботи кількох підрозділів підприємства. Можливість виконання заходів перспективного плану повинна бути підтверджена обґрунтованим розрахунком необхідного матеріально-технічного забезпечення і фінансових витрат з зазначенням джерел фінансування.

До перспективних планів належить комплексний план покращення умов праці і санітарно-оздоровчих заходів, що передбачає створення, відповідно до нормативних актів з охорони праці, умов праці, пов'язаних з перспективними змінами підприємства. Таке планування, як правило, розраховане на термін від 2 до 5 років. Реалізація цих планів забезпечується через річні плани номенклатурних заходів з охорони праці, які вносяться до угоди, що є невід'ємною частиною колективного договору.

Поточне планування здійснюється у межах календарного року через розроблення відповідних заходів у розділі «Охорона праці» колективного договору.

Поточні плани передбачають реалізацію заходів із покращення умов праці, створення кращих побутових і соціальних умов на виробництві. Ці плани обов'язково забезпечуються фінансуванням згідно з розробленими кошторисами.

Питання охорони праці можуть віддзеркалюватися в інших поточних планах, які підприємства та організації можуть складати на вимогу трудових колективів:

- план соціального розвитку колективу;
- наукової організації праці;
- механізації важких і ручних робіт;
- охорони праці жінок;
- підготовки підприємства до робіт в осінньо-зимовий період;
- підвищення культури виробництва та ін.

Оперативне планування роботи з охорони праці здійснюється за підсумками контролю стану охорони праці в структурних підрозділах і на підприємстві в цілому.

Оперативні плани складаються для швидкого виправлення виявлених в процесі державного, відомчого і громадського контролю недоліків в стані охорони праці, а також для ліквідації наслідків аварій або стихійного лиха.

Оперативні заходи щодо усунення виявлених недоліків зазначаються безпосередньо у наказі власника підприємства, який видається за підсумками контролю, або у плані заходів, як додатку до наказу.

Організаційно-методичну роботу щодо складання перспективних, поточних та оперативних планів здійснює служба (спеціаліст) охорони праці.

4.2 Безпека в надзвичайних ситуаціях

Пожежна безпека на підприємстві

Пожежа - неконтрольоване горіння поза спеціальним вогнищем, яке призводить до матеріальної шкоди.

Пожежна безпека – стан об'єкта, при якому з регламентованою ймовірністю виключається можливість виникнення та розвиток пожежі і впливу на людей її небезпечних факторів, а також забезпечується захист матеріальних цінностей.

Причинами пожеж та вибухів на підприємстві є порушення правил і норм пожежної безпеки, невиконання Закону “Про пожежну безпеку”.

Небезпечними факторами пожежі і вибуху, які можуть призвести до травми, отруєння, загибелі або матеріальних збитків є відкритий вогонь, іскри, підвищена температура, токсичні продукти горіння, дим, низький вміст кисню, обвалення будинків і споруд.

За стан пожежної безпеки на підприємстві відповідають її керівники, начальники цехів, майстри та інші керівники.

На підприємствах існує два види пожежної охорони: професійна і воєнізована. Воєнізована охорона створюється на об'єктах з підвищеною небезпекою. Крім того на підприємствах для посилення пожежної охорони організовуються добровільні пожежні дружини і команди, добровільні пожежні товариства і пожежно-технічні комісії з числа робітників та службовців. При Міністерстві внутрішніх справ існує управління пожежної охорони (УПО) і його органи на місцях. До складу УПО входить Державний пожежний нагляд який здійснює:

Контроль за станом пожежної безпеки

Розробляє і погоджує протипожежні норми і правила та контролює їх виконання в проектах і безпосередньо на об'єктах народного господарства

Проводить розслідування і облік пожеж

Організовує протипожежну профілактику.

Протипожежна профілактика – це комплекс організаційних і технічних заходів, які спрямовані на здійснення безпеки людей, на попередження пожеж, локалізацію їх поширення, а також створення умов для успішного гасіння пожежі.

Відповідальним керівником робіт по ліквідації пожеж і аварій на підприємстві є головний інженер. Начальник структурного підрозділу, в якому виникла пожежа, є відповідальним виконавцем робіт по її ліквідації.

Протипожежні вимоги до будинків і споруд [1]

Виходячи з властивостей речовин і матеріалів, умов їх застосування і обробки і у відповідності із ОНТП 24-86 “Визначення категорій приміщень і будівель по вибухопожежній і пожежній небезпеці” приміщення по вибухопожежній і пожежній небезпеці діляться на п'ять категорій – А, Б, В, Г, Д.

До категорії А належать приміщення, де перебувають спалимі та легкозаймисті рідини з температурою спалаху, що не перевищує 28°C, а також речовини і матеріали здатні вибухати і горіти при взаємодії з водою, киснем або одне з одним; при утворенні вибухонебезпечних сумішей розвивається розрахунковий надлишковий тиск вибуху 5 кПа.

До категорії Б належать приміщення, в яких є пил та волокна, легкозаймисті рідини з температурою спалаху понад 28°C та спалимі рідини в такій кількості, що можуть утворюватися вибухонебезпечні пилоповітряні та пароповітряні суміші, при займанні яких розвивається розрахунковий надлишковий тиск вибуху 5 кПа.

До категорії В належать приміщення, де перебувають спалимі та важкоспалимі рідини, тверді спалимі та важкоспалимі речовини та матеріали (в тому числі пил та волокна), а також речовини і матеріали які здатні при взаємодії з водою, киснем повітря та одне з одним тільки горіти (за умови, що ці приміщення не відносяться до категорії А чи Б).

До категорії Г належать приміщення, в яких є неспалимі речовини та матеріали в гарячому, розпеченому або розплавленому стані, а також спалимі гази, рідини та тверді речовини, які спалюються або утилізуються як паливо; процес їх обробки супроводжується виділенням променевої теплоти, іскор та полум'я.

До категорії Д належать приміщення, в яких є неспалимі речовини та матеріали у холодному стані.

На розвиток пожежі у приміщеннях та спорудах значно впливає здатність окремих будівельних елементів чинити опір впливу тепла, тобто їх вогнестійкість.

Вогнестійкість – здатність будівельних конструкцій чинити опір дії високої температури, утворенню наскрізних тріщин та поширенню вогню в

умовах пожежі і виконувати при цьому свої звичайні експлуатаційні функції. Вогнестійкість конструкцій будівель характеризується межею вогнестійкості.

Межа вогнестійкості – це час, на протязі якого конструкція може витримати дію вогню, а потім вже починається деформація.

Всі будівлі і споруди за ступенем вогнестійкості за СНиП 2.01.02-85 поділяють на 5 ступеней.

Будинок може належати до того або іншого ступеня вогнестійкості, якщо значення меж вогнестійкості і меж поширення вогню усіх конструкцій не перевищує значень вимог СНиП 2.01.02-85. [1]

4.3 Висновки до розділу

В розділі «Охорона праці та безпека в надзвичайних ситуаціях» розглянуто планування заходів з охорони праці. Види планування та контролю стану охорони праці. Виявлення, оцінка та зменшення ризиків небезпечних подій, а також питання пожежної безпеки на підприємстві.

ВИСНОВКИ

Проведено аналіз методів виявлення точок зміни часових рядів, основною метою яких є знаходження різких змін у даних, коли змінюється властивість часового ряду. Ця задача є актуальною наприклад для розпізнавання мовлення, аналізу зображень, виявлення сигналів на фоні завад, оцінювання змін стану систем чи режимів їхнього функціонування тощо.

Власне задачу виявлення змін ймовірнісних характеристик сигналів зведено до задачі виявлення розладки, яка характеризує момент або інтервал часу, коли статистичні чи структурні властивості сигналу помітно змінюються порівняно з “нормальним” режимом.

Показано, що одним з важливих питань для алгоритмів виявлення часових моментів змін ймовірнісних характеристик є потреба в розробленні алгоритмів із роботою в режимі реального часу та з мінімальними затримками. У багатьох реальних застосуваннях точки змін використовуються для вибору та виконання своєчасних дій, тому якомога швидше знайти точки змін є критично важливим. Альтернативою розглянутих алгоритмів є використання методів, які потребують менших розмірів вікон для розрахунку часових точок змін, таких як байєсівські методи.

Оскільки, розладка трактується як стрибкоподібні зміни ймовірнісних характеристик сигналів, в якості моделі таких сигналів доцільно використати клас стаціонарних процесів, зокрема кусково стаціонарних. Це накладає значні обмеження, оскільки потребує підтвердження стаціонарності сигналів, однак для більшості прикладних задач припущення про стаціонарність на коротких проміжках часу є допустимим. В роботі використано спектрально-кореляційні методи опрацювання, а саме опрацювання проводиться на інтервалах ковзного вікна.

Для виявлення розладки запропоновано проводити опрацювання сигналу в межах ковзного вікна, яке зміщується по сигналу, і всередині якого проводиться спектрально-кореляційна обробка. На основі цієї обробки за сприйнятливими змінами ймовірнісних характеристик приймається і рішення про появу розладки.

Для виявлення розладки в структурі тестового сигналу в межах кожної трансляції ковзного вікна було обчислено спектр потужності, а далі було проведено його усереднення як за потужністю, так і за часом. В результаті отримувалась точка, яка далі відкладалась на часовій осі відповідно до розміщення відповідного вікна. Для кількісного визначення часового моменту появи розладки використано поріг, значення якого припускається рівними варіації оцінок усереднених спектрів.

Встановлено, що значення варіації оцінок усереднених спектрів значно відрізняються між ділянками сигналу до та після виникнення розладки і шляхом моніторингу їхнього значення при трансляції ковзного вікна по сигналу можна виявити моменти виникнення розладки в сигналі.

Розроблено алгоритм пропонованого методу, на основі використання якого, маючи обчислені значення варіації усередненої потужності для ділянок сигналу без розладки та приймаючи їх за значення порогу, можна виявляти часові моменти розладки.

ПЕРЕЛІК ПОСИЛАНЬ

1. G. D. Montanez, S. Amizadeh, and N. Laptev, “Inertial Hidden Markov Models: Modeling Change in Multivariate Time Series,” in AAAI Conference on Artificial Intelligence, 2015, pp. 1819–1825.
2. Y. Kawahara and M. Sugiyama, “Sequential Change-Point Detection Based on Direct Density-Ratio Estimation,” in SIAM International Conference on Data Mining, 2009, pp. 389–400.
3. P. Yang, G. Dumont, and J. M. Ansermino, “Adaptive change detection in heart rate trend monitoring in anesthetized children,” IEEE Trans. Biomed. Eng., vol. 53, no. 11, pp. 2211–9, Nov. 2006
4. J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, “A Review and Comparison of Changepoint Detection Techniques for Climate Data,” J. Appl. Meteorol. Climatol., vol. 46, no. 6, pp. 900–915, Jun. 2007.
5. M. F. R. Chowdhury, S.-A. Selouani, and D. O’Shaughnessy, “Bayesian on-line spectral change point detection: a soft computing approach for on-line ASR,” Int. J. Speech Technol., vol. 15, no. 1, pp. 5–23, Oct. 2011
6. R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, “Image change detection algorithms: a systematic survey,” IEEE Trans. Image Process., vol. 14, no. 3, pp. 294–307, Mar. 2005.
7. E. Keogh and J. Lin, “Clustering of time-series subsequences is meaningless: implications for previous and future research,” Knowl. Inf. Syst., vol. 8, no. 2, pp. 154–177, Aug. 2004.
8. D.-H. Tran, “Automated Change Detection and Reactive Clustering in Multivariate Streaming Data,” Nov. 2013.

9. S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation.,” *Neural Netw.*, vol. 43, pp. 72–83, Jul. 2013.
10. H. Chen and N. Zhang, “Graph-Based Change-Point Detection,” *Ann. Stat.*, vol. 43, no. 1, pp. 139–176, Sep. 2014.
11. Z. Harchaoui, E. Moulines, and F. R. Bach, “Kernel Change-point Analysis,” in *Advances in Neural Information Processing Systems*, 2009, pp. 609–616
12. V. C. Raykar, “Scalable machine learning for massive datasets: Fast summation algorithms,” University of Maryland, College Park, 2007.
13. T. Rakthanmanon, E. J. Keogh, S. Lonardi, and S. Evans, “Time Series Epenthesis: Clustering Time Series Streams Requires Ignoring Some Data,” in *IEEE 11th International Conference on Data Mining*, 2011, pp. 547–556.
14. S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, “Using mobile phones to determine transportation modes,” *ACM Trans. Sens. Networks*, vol. 6, no. 2, pp. 1–27, Feb. 2010
15. I. Cleland, M. Han, C. Nugent, H. Lee, S. McClean, S. Zhang, and S. Lee, “Evaluation of prompted annotation of activity data recorded from a smart phone.,” *Sensors (Basel)*, vol. 14, no. 9, pp. 15861–79, Jan. 2014.
16. H. Cho and P. Fryzlewicz, “Multiple-change-point detection for high dimensional time series via sparsified binary segmentation,” *J. R. Stat. Soc. Ser. B (Statistical Methodol.)*, vol. 77, no. 2, pp. 475–507, Mar. 2015.
17. K. Yamanishi and J. Takeuchi, “A unifying framework for detecting outliers and change points from non-stationary time series data,” in *8th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02*, 2002, p. 676
18. M. Yamada, A. Kimura, F. Naya, and H. Sawada, “Change-point detection with feature selection in high-dimensional time-series data,” in *International Joint Conference on Artificial Intelligence*, 2013.

19. L. I. Kuncheva, "Change Detection in Streaming Multivariate Data Using Likelihood Detectors," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 5, pp. 1175–1180, May 2013.
20. Y. Kawahara, T. Yairi, and K. Machida, "Change-Point Detection in Time-Series Data Based on Subspace Identification," in *7th IEEE International Conference on Data Mining (ICDM 2007)*, 2007, pp. 559–564.
21. S. Chib, "Estimation and Comparison of multiple change point models," *J. Econom.*, vol. 86, no. 2, pp. 221–241, 1998.
22. B. A. Tan, P. Gerstoft, C. Yardim, and W. S. Hodgkiss, "Change-point detection for recursive Bayesian geoacoustic inversions.," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1962–70, Apr. 2015.
23. Y. Saatzi, R. D. Turner, and C. E. Rasmussen, "Gaussian Process Change Point Models," in *International Conference on Machine Learning*, 2010, pp. 927–934.
24. S. Brahim-Belhouari and A. Bermak, "Gaussian process for nonstationary time series prediction," *Comput. Stat. Data Anal.*, vol. 47, no. 4, pp. 705–712, Nov. 2004.
25. Z. Harchaoui, F. Vallet, A. Lung-Yut-Fong, and O. Cappe, "A regularized kernel-based approach to unsupervised audio segmentation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 1665–1668.
26. P. R. Rosenbaum, "An exact distribution-free test comparing two multivariate distributions based on adjacency," *J. R. Stat. Soc.*, vol. 67, pp. 515–530, 2005.
27. E. Keogh, S. Chu, D. Hart, and M. Pazzani, "An online algorithm for segmenting time series," in *IEEE International Conference on Data Mining*, 2001, pp. 289–296.

28. Dozorskyi V., Dediv I., Sverstiuk S., Nykytyuk V., Karnaukhov A. The Method of Commands Identification to Voice Control of the Electric Wheelchair. 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.233-240. ISSN 1613-0073.

29. 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.

30. Khvostivska L., Khvostivskyi M., Dunets V., Dediv I. Mathematical, algorithmic and software support of synphase detection of radio signals in electronic communication networks with noises. Scientific Journal of TNTU (Tern.), vol 111, no 3, 2023. pp. 48–57.

31. Основи технології радіоелектронних апаратів : навчальний посібник / Р. А. Ткачук, В. Г. Дозорський, Л. Є. Дедів, І. Ю. Дедів. - Тернопіль : Тернопільський національний технічний університет імені Івана Пулюя, 2017. - 336 с.

32. Dozorskyi V., Dediv L., Kovalyk S., Dozorska O., Dediv I. (2024) Design of the endoskeleton of a biocontrolled hand prosthesis. Scientific Journal of TNTU (Tern.), vol. 115, no 3, pp. 100-111.

33. Дедів І. Ю. Метод автентифікації користувачів за параметрами голосових сигналів / І. Ю. Дедів, М. М. Кузик // Збірник тез доповідей VI Міжнародної науково-технічної конференції молодих учених та студентів „Актуальні задачі сучасних технологій“, 16-17 листопада 2017 року. Т.: ТНТУ, 2017. Том 2. С. 47.

34. 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.

35. Математичне та комп'ютерне моделювання електрокардіосигналів у системах голтерівського моніторингу / Л.Є. Дедів, А.С. Сверстюк, І.Ю. Дедів, М.О. Хвостівський, В.Г. Дозорський, Є.Б. Яворська. – Львів: Видавництво «Магнолія - 2006», 2021. – 120 с. ISBN 978-617-574-218-1.

36. Математичне моделювання, методи та програмне забезпечення опрацювання дихальних шумів у комп'ютерних аускультативних діагностичних системах / І.Ю. Дедів, А.С. Сверстюк, Л.Є. Дедів, В.Г. Дозорський, М.О. Хвостівський. – Львів: Видавництво «Магнолія - 2006», 2021. – 126 с. ISBN 978-617-574-219-8.

37. Методичні рекомендації з оформлення кваліфікаційних робіт бакалавра за спеціальністю 172 «Телекомунікації та радіотехніка» уклад.: Дунець В.Л., Хвостівський М.О. Дедів І.Ю. Тернопіль: ТНТУ імені Івана Пулюя, 2021 р. – 72с.

38. Дозорський В.Г., Дозорська О.Ф., Дедів Л.Є., Дедів І.Ю., Паньків І.М., Яворська Є.Б. Структура системи відбору біосигналів для задачі відновлення комунікативної функції людини. Вісник Хмельницького національного університету: технічні науки. – Хмельницький: редакція журналу "Вісник Хмельницького національного університету". – 2019. - №2(271) – с. 183-186.

39. Dozorska O., Yavorska E., Dozorskyi V., Pankiv I., Dediv L. Dediv I. The Method of Indirect Restoration of Human Communicative Function. Proc. of the 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), CADSM'2019, (pp. 19–22). Polyana-Svalyava (Zakarpatty), UKRAINE 978-1-7281-0053-1/19.

40. Дозорська О.Ф., Яворська Є.Б., Дозорський В.Г., Дедів Л.Є., Дедів І.Ю. Метод виявлення ознак основного тону в структурі електроміографічних сигналів для задачі компенсації порушеної комунікативної функції людини», Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування, (81), с. 56-64. doi: 10.20535/RADAP.2020.81.56-64.

41. Khvostivska L., Khvostivsky M., Dunetc V., Dediv I. Mathematical and Algorithmic Support of Detection Useful Radiosignals in Telecommunication Networks. 2nd International Workshop on Information Technologies: Theoretical and Applied Problems, ITTAP 2022. CEUR Workshop Proceedings. Ternopil 22- 24 November 2022. Vol 3309, P. 314-318. ISSN 1613-0073.

42. Гевко О.В., Дозорський В.Г., Дедів Л.Є., Дедів І.Ю., Дозорська О.Ф. Структурний синтез вібромасажної апаратури. Перспективні технології та прилади, № 20, Луцьк, 2022. – с. 23-31.

43. Dozorskyi V., Dediv I., Sverstiuk S., Nykytyuk V., Karnaukhov A. The Method of Commands Identification to Voice Control of the Electric Wheelchair. 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.233-240. ISSN 1613-0073.

44. 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.

45. Khvostivska L., Khvostivskyi M., Dunets V., Dediv I. Mathematical, algorithmic and software support of synphase detection of radio signals in electronic communication networks with noises. Scientific Journal of TNTU (Tern.), vol 111, no 3, 2023. pp. 48–57.

46. Основи технології радіоелектронних апаратів : навчальний посібник / Р. А. Ткачук, В. Г. Дозорський, Л. Є. Дедів, І. Ю. Дедів. - Тернопіль : Тернопільський національний технічний університет імені Івана Пулюя, 2017. - 336 с.

47. Dozorskyi V., Dediv L., Kovalyk S., Dozorska O., Dediv I. (2024) Design of the endoskeleton of a biocontrolled hand prosthesis. Scientific Journal of TNTU (Tern.), vol. 115, no 3, pp. 100-111.

48. Техноекологія та цивільна безпека. Частина «Цивільна безпека». Навчальний посібник / В.С. Стручок, – Тернопіль: ТНТУ ім. І.Пулюя, 2022. – 150 с.

49. 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.

ДОДАТКИ

Додаток А

Текст програми обчислення усереднених оцінок спектральної густини
потужності в межах ковзного вікна

```
x=data;  
T=50;  
for k=0:545  
    x1=x((k*T+1):(k*T+250));  
    z1=xcorr(x1);  
    hold on  
    figure(2)  
    plot(z1)  
    grid on  
    ff1=abs(fft(z1));  
    hold on  
    figure(3)  
    plot(ff1(1:100))  
    grid on  
    d=mean(ff1(1:60));  
end
```

Додаток Б

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ**

МАТЕРІАЛИ

ХІІ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

**«ІНФОРМАЦІЙНІ МОДЕЛІ,
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



17-18 грудня 2025 року

**ТЕРНОПІЛЬ
2025**

СЕКЦІЯ 1. МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

УДК 621.397

О. Керницький; В. Сікула; Т. Максимів; І. Дедів, к. т. н., доц.

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

МЕТОД ПІДВИЩЕННЯ РІВНЯ ВІЗУАЛЬНОГО СПРИЙНЯТТЯ ПОШКОДЖЕНИХ ЗОБРАЖЕНЬ НА ОСНОВІ СТОХАСТИЧНОГО РЕЗОНАНСУ

UDC 621.397

O. Kernitsky; V. Sikula; T. Maksymiv; I. Dediv, PhD., Assoc. Prof.

THE METHOD FOR ENHANCING THE LEVEL OF DAMAGED IMAGES VISUAL PERCEPTION BASED ON STOCHASTIC RESONANCE

Візуальна якість зображень є ключовим поняттям у цифровій обробці зображень, комп'ютерному зорі, системах стиснення та передавання мультимедійних даних. Від того, наскільки якісно відтворюється зображення, залежить комфорт користувача, точність діагностики в медичних застосуваннях, надійність систем відеоспостереження та ефективність автоматизованих систем розпізнавання. Під візуальною якістю зображення зазвичай розуміють ступінь відповідності між сприйнятим спостерігачем зображенням і тим, що вважається еталонним з точки зору людського зору.

На сьогодні розроблено велику кількість методів покращення якості зображень, що в свою чергу, призводить до покращення їхнього візуального сприйняття. Перспективним до обробки зображень, зокрема покращення візуального сприйняття, є застосування принципу стохастичного резонансу, що є явищем, коли додавання шуму до слабого сигналу допомагає краще його виявити. У випадку зображень, до них додається певна кількість випадкового шуму та застосовується нелінійна обробка, і в результаті контури, деталі та структури стають видимішими, особливо у темних або низькоконтрастних ділянках. Однак, основною вимогою цього методу є та, що параметри шуму мають бути оптимальними. Якщо порівнювати цей метод із іншими поширеними, то, наприклад, контрастування й адаптивне гістограмне вирівнювання з обмеженням контрасту є різними версіями детермінованих перетворень яскравості, а метод стохастичного резонансу є нелінійним шумо-спричиненим підсиленням слабких сигналів, де шум входить в структуру методу як необхідний елемент.

Для середовища Matlab було проведено розробку програми, яка підтримує роботу із сіро-рівневими та кольоровими (RGB) зображеннями, передбачає автоматичне оновлення параметрів стохастичного резонансу, забезпечує наочне порівняння отриманого після обробки зображення із оригіналом. Для власне обробки в методі реалізовано алгоритм стохастичного резонансу в плані багаторазового додавання шуму, нелінійної порогової обробки, усереднення результатів та фінального підсилення контрасту.

Як тестове зображення було використано пошкоджене зображення САРТСНА. Встановлено, що при додаванні шуму з допомогою розробленого методу в зображенні починають виокремлюватись краї тексту і він стає більш сприйнятним візуально. Попри те, що на зображенні присутній значний рівень фонового шуму, який був штучно доданий в процесі реалізації стохастичного резонансу, окремі букви тексту стало набагато простіше диференціювати візуально, тобто візуальне сприйняття покращилось.

Таким чином встановлено, що застосування стохастичного резонансу до обробки пошкоджених зображень справді дає можливість підвищення рівня візуального сприйняття пошкоджених зображень, а розроблена програма дає можливість пошуку оптимальних параметрів обробки, зокрема рівня доданого шуму.

ЗМІСТ

СЕКЦІЯ 1. МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

О. Керницький, В. Сікула, Т. Максимів, І. ДедівМЕТОД ПІДВИЩЕННЯ РІВНЯ ВІЗУАЛЬНОГО СПРИЙНЯТТЯ ПОШКОДЖЕНИХ
ЗОБРАЖЕНЬ НА ОСНОВІ СТОХАСТИЧНОГО РЕЗОНАНСУ**O. Kernitsky, V. Sikula, T. Maksymiv, I. Dediv**THE METHOD FOR ENHANCING THE LEVEL OF DAMAGED IMAGES VISUAL
PERCEPTION BASED ON STOCHASTIC RESONANCE 3**А. Кондратюк, Я. Литвиненко**

МЕТОДИ МОДЕЛЮВАННЯ ЕЛЕКТРОКАРДІОСИГНАЛУ З АРТЕФАКТАМИ

A. Kondratyuk, I. Lytvynenko

METHODS FOR MODELING ELECTROCARDIO SIGNAL WITH ARTIFACTS 4

В. Кривень, В. Шаповалов

МАТЕМАТИЧНА МОДЕЛЬ ЕЛЕКТРИЧНОГО ОПОРУ КРУГОВОЇ ПЛАСТИНИ

V. Kryven, V. SharovalovMATHEMATICAL MODEL OF ELECTRICAL RESISTANCE OF A CIRCULAR
PLATE 5**С. Литвиненко, М. Фриз**ПІДХОДИ ДО МОДЕЛЮВАННЯ СПОСОБІВ ПРОСУВАННЯ ЦІЛЬОВОЇ
РЕКЛАМИ**S. Lytvynenko, M. Friz**

APPROACHES TO MODELING METHODS OF TARGETED ADVERTISING 6

І. Плєскун, Д. Кривко, Т. Дубиняк, М. ЯворськаМОДЕЛЮВАННЯ СИСТЕМИ КЕРУВАННЯ ТЕМПЕРАТУРНИМ РЕЖИМОМ
ГОСПОДАРСЬКОГО ПРИМІЩЕННЯ**I. Pleskun, D. Kryvko, T. Dubyniak, M. Yavorska**

MODELING OF A TEMPERATURE CONTROL SYSTEM FOR UTILITY ROOMS 7

М. СтрілецькийМОДЕЛЬ ЛОТКА-ВОЛЬТЕРРА ОЦІНКИ ЗАБРУДНЕНЬ В ЗОНІ
ФУНКЦІОНУВАННЯ МАЛОГО МАШИНОБУДІВНОГО ПІДПРИЄМСТВА**M. Striletskyi**LOTKA-VOLTERRA MODEL FOR ASSESSING POLLUTION IN THE
OPERATIONAL ZONE OF A SMALL MACHINE-BUILDING ENTERPRISE 9**В. Ступка, В. Анненков, І. Дедів**МЕТОД СТИСНЕННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ НА ОСНОВІ ВЕЙВЛЕТ-
КОДУВАННЯ**V. Stupka, V. Annenkov, I. Dediv**

DIGITAL IMAGE COMPRESSION METHOD BASED ON WAVELET CODING 11

В. Сумко, Я. ЛитвиненкоОГЛЯД МЕТОДІВ ІНТЕРПОЛЯЦІЇ СИГНАЛІВ В ЗАДАЧАХ ОЦІНЮВАННЯ ЇХ
РИТМІЧНОЇ СТРУКТУРИ**V. Sumko, I. Lytvynenko**A REVIEW OF SIGNAL INTERPOLATION METHODS FOR THE PURPOSE OF
ESTIMATING THEIR RHYTHMIC STRUCTURE 12**М. Фриз, Б. Млинко**СИСТЕМНИЙ АНАЛІЗ МОДЕЛЕЙ ТА МЕТОДІВ МОНІТОРИНГУ
ЗАБРУДНЕНЬ АТМОСФЕРНОГО ПОВІТРЯ**M. Fryz, B. Mlynko**SYSTEM ANALYSIS OF ATMOSPHERIC AIR POLLUTION MONITORING
MODELS AND METHODS 14