

Харківський національний університет радіоелектроніки
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

ПОРОШЕНКО АНТОН ІГОРОВИЧ

УДК 004.72:004.93 (043.3)

ДИСЕРТАЦІЯ

**МОДЕЛЬ ТА МЕТОДИ ОБРОБКИ АУДІОСИГНАЛІВ У МЕРЕЖНИХ
СИСТЕМАХ АУДІОАНАЛІТИКИ**

Спеціальність: 126 Інформаційні системи та технології

Галузь знань: 12 Інформаційні технології

Подається на здобуття ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

_____ А. І. Порошенко

Науковий керівник
Коваленко Андрій Анатолійович,
доктор технічних наук, професор

Харків – 2025

АНОТАЦІЯ

Порошенко А. І. Модель та методи обробки аудіосигналів у мережних системах аудіоаналітики. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 126 Інформаційні системи та технології – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2025.

Дисертаційну роботу присвячено підвищенню ефективності обробки аудіосигналів у мережних системах аудіоаналітики. У роботі запропоновано використання математичного моделювання для оптимізації мережних процесів та удосконалено методи передачі і аналізу аудіосигналів, що забезпечують високу точність аналізу аудіосигналів, зниження затримок передачі аудіосигналів та підвищення якості обслуговування в умовах реального часу. Актуальність даної теми визначається тим, що швидкий розвиток технологій та збільшення обсягу аудіоданих призводить до появи нових викликів у передачі та аналізі аудіосигналів у мережних системах аудіоаналітики. Поєднання зростаючої кількості даних, потреби в реальному часі та складності аналізу робить тему підвищення ефективності передачі та аналізу аудіосигналів у мережних системах аудіоаналітики актуальною і важливою для дослідження та розвитку.

Метою дисертаційної роботи є підвищення точності аналізу та зменшення затримок передачі аудіосигналів у мережних системах аудіоаналітики за рахунок оптимізації мережної інфраструктури та удосконалення методів передачі і аналізу аудіосигналів.

Об’єкт дослідження – процеси передачі та аналізу аудіосигналів.

Предмет дослідження – методи передачі та аналізу аудіосигналів у мережних системах аудіоаналітики.

У вступі обґрунтовується актуальність використання мережних систем аудіоаналітики в різноманітних сферах, зокрема безпеки, моніторингових систем, маркетингу та розваг. Звертається увага на важливість передачі аудіосигналів у реальному часі, оскільки затримка може негативно впливати на ефективність та швидкість реакції системи. Розкриваються основні компоненти процесу обробки та аналізу аудіоданих (збір, попередня обробка, передача, вилучення ознак), а також окреслюються завдання дослідження, що включають створення математичної моделі, вибір оптимальної мережної структури, удосконалення методів передачі та аналізу аудіосигналів. Визначаються об'єкт, предмет, мета роботи та описується наукова новизна, яка полягає у новій математичній моделі проходження аудіосигналів та удосконалених методах їх передачі та аналізу. Також наведено сфери практичного застосування отриманих результатів і вказано на їх впровадження в освітній процес та реальні виробничі середовища.

У першому розділі дисертації представлено аналіз сучасного стану розвитку методів обробки аудіосигналів у мережних системах аудіоаналітики. Розглянуто основні напрями та тенденції в цій сфері, проведено огляд існуючих систем, їхніх архітектур і функціональних можливостей, а також визначено ключові вимоги до таких систем, включаючи точність та швидкість обробки. На основі цього аналізу сформульовано завдання дослідження, що спрямоване на розроблення математичної моделі проходження аудіосигналів та удосконалення методів передачі та аналізу аудіосигналів, що забезпечить високу надійність і адаптивність до різних умов.

У другому розділі дисертації досліджено процес проходження аудіосигналів у мережних системах аудіоаналітики, зокрема, їх збір та передачу. Описано математичну модель для оптимізації мережних ресурсів, що враховує ієрархічну структуру системи та різні маршрути передачі даних. Особливу увагу приділено побудові оптимізаційних рівнянь для базових мережних структур, що дозволяють мінімізувати капітальні витрати,

забезпечуючи високу якість обслуговування. Розглянуто вплив неординарного пуасонівського навантаження на ефективність системи та запропоновано методи вибору оптимальної мережної структури, які створюють основу для подальшого вдосконалення систем аудіоаналітики.

У третьому розділі дисертації запропоновано методи для передачі та аналізу аудіосигналів у мережних системах аудіоаналітики. Докладно описано алгоритми управління з'єднаннями, метод комутації віртуальних маршрутів і моделювання джерел навантаження з урахуванням вимог семантичної прозорості та часових обмежень, висвітлено роль управління трафіком і використання сучасних мережних протоколів для дотримання показників якості обслуговування, а також окреслено складнощі, що виникають у зашумленому середовищі та неоднорідній мережній інфраструктурі. Окрему увагу приділено нейронним мережам для аналізу аудіосигналів і застосуванню метрик оцінювання якості. Показано, що запропоновані підходи забезпечують високу швидкодію та точність в режимі реального часу.

У четвертому розділі надано комплексний опис розробленої інформаційної технології обробки аудіосигналів, зокрема висвітлено концептуальні засади побудови, обґрунтовано вибір технологічного стеку та розглянуто підходи до інтеграції компонентів для збору, передачі та аналізу аудіоданих у реальному часі. Детально проаналізовано архітектуру системи, а також принципи взаємодії її складових, спрямовані на забезпечення максимальної швидкодії, масштабованості та надійності. У розділі також наведено результати експериментальної оцінки системи, які підтверджують її здатність до високоточного аналізу аудіосигналів та ефективного функціонування в складних та неоднорідних мережних умовах.

У висновках дисертаційної роботи узагальнено основні результати дослідження, підсумовано значення наукових і практичних досягнень. Показано, що запропоновані модель і методи передачі та аналізу аудіосигналів у мережній системі аудіоаналітики забезпечують підвищення

точності і зменшення затримок передачі аудіосигналів у мережних системах аудіоаналітики.

Методи дослідження. При розв'язанні науково-прикладної задачі використано широкий спектр методів. Так, при розробці моделі процесу проходження аудіосигналів у мережних системах аудіоаналітики використовувались методи математичного моделювання, зокрема аналіз мережних потоків, імовірнісні підходи до обчислення втрат та оптимізаційні алгоритми, що дозволяє врахувати специфіку неординарного навантаження та забезпечити мінімальні капітальні витрати за умови дотримання вимог до якості обслуговування. При розробці методу передачі аудіосигналів у мережних системах аудіоаналітики використовувались методи пакетизації даних, динамічного розподілу ресурсів та оптимізації часової затримки, а також алгоритми формування черг і корекції помилок, які дають змогу забезпечити потрібний рівень якості обслуговування в реальному часі та ефективно керувати обсягами трафіку. При розробці методу аналізу аудіосигналів у мережних системах аудіоаналітики використовувались методи глибинного навчання, насамперед нейронні мережі конволюційної та рекурентної архітектури, що дали змогу розділяти джерела звуку та підвищити точність розпізнавання подій у реальному часі. Оцінка експериментальних даних, отриманих у ході роботи, проводилася на основі методів теорії ймовірності та математичної статистики.

Вибір методів досліджень забезпечив достовірність отриманих результатів і висновків, що підтверджується збіжністю результатів експериментальних досліджень, отриманих при програмній реалізації інформаційної технології обробки аудіосигналів, з теоретичними розрахунками та чисельним моделюванням, а також узгодженням із прогнозними оцінками ефективності запропонованих підходів.

Основні завдання дослідження:

- провести аналіз напрямків розвитку методів обробки аудіосигналів та мережних систем аудіоаналітики;

- розробити математичну модель процесу передачі аудіосигналів у мережних системах аудіоаналітики;
- розробити метод вибору мережної структури системи аудіоаналітики;
- удосконалити метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів.
- удосконалити метод аналізу аудіосигналів на основі багатоподійного класифікатора;
- удосконалити інформаційну технологію обробки аудіосигналів з використанням удосконалених моделі та методів.

Наукова новизна отриманих результатів обумовлена розробленими та удосконаленими моделлю і методами обробки аудіосигналів у мережних системах аудіоаналітики, в межах яких отримані такі нові наукові результати:

1. Вперше запропоновано математичну модель процесу передачі аудіосигналів у мережних системах аудіоаналітики з урахуванням неординарного пуасонівського навантаження, що базується на кількісних результатах аналізу мережних структур та дозволяє підвищити ефективність передачі даних і зменшити витрати на інфраструктуру;

2. Удосконалено метод передачі аудіосигналів, який відрізняється від існуючих використанням комутації віртуальних маршрутів, та дозволяє забезпечити високу якість передачі, знизити затримки у передачі аудіосигналів і підвищити стабільність системи в умовах динамічних мережних навантажень;

3. Отримав подальший розвиток метод аналізу аудіосигналів шляхом застосування багатоподійного класифікатора, заснованого на глибинних нейронних мережах, що дозволяє підвищити точність аналізу аудіосигналів у складних акустичних середовищах;

4. Отримала подальший розвиток інформаційна технологія обробки аудіосигналів у мережних системах аудіоаналітики, яка базується на вдосконалених методах передачі і аналізу аудіосигналів та дозволяє

інтегрувати процеси аналізу аудіосигналів з реальним часом для різних галузей застосування.

Проведено аналітичний огляд предметної області за темою дослідження, проаналізовано методи виявлення як монофонічних імпульсних аудіоподій, так і поліфонічних аудіоподій. Також представлені актуальні проблеми, пов'язані з розробкою таких систем. Здійснено огляд сучасних існуючих рішень для обробки аудіосигналів, сформульовано мету та задачі дослідження дисертаційної роботи.

Запропонована постановка задачі оптимізації мережної системи аудіоаналітики за критерієм мінімуму капітальних витрат. Складені оптимізаційні рівняння для однонаправленої трьохвузлової базової структури, виходячи із завдання мінімізації капітальних витрат. Отримані оптимізаційні рівняння для двонаправленої трьохвузлової базової структури мережі. Проведено узагальнення отриманих оптимізаційних рівнянь на будь-яку структуру мережної системи аудіоаналітики.

Запропоновано метод вибору мережної структури системи аудіоаналітики. Розглянуто випадок оптимізації мережної структури системи аудіоаналітики при неординарному пуасонівському навантаженні. Виконана оптимізація базової мережі системи аудіоаналітики за критерієм мінімізації капітальних витрат та проаналізовані результати оптимізації.

Виконано моделювання джерел навантаження системи. Розглянуто методи управління з'єднаннями та трафіком у мережних системах аудіоаналітики. Досліджено методи забезпечення якості передачі аудіосигналів. Запропоновано метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів.

Обґрунтовано показник якості передачі та аналізу аудіосигналів. Досліджено набори даних для навчання, валідації та тестування методів аналізу аудіосигналів у мережних системах аудіоаналітики. Запропоновано метод аналізу аудіосигналів у мережних системах аудіоаналітики.

Проведено аналіз процесу обробки аудіосигналів у мережній системі аудіоаналітики з метою контролю її ефективності. Запропоновано структуру інформаційної технології обробки аудіосигналів, що враховує процеси передачі та аналізу аудіосигналу.

Практичне значення результатів роботи полягає в тому, що розроблені у роботі модель та методи є науково-практичною основою для подальшого удосконалення мережних систем аудіоаналітики. Представлені на їх основі інженерні методи та алгоритми дають змогу:

- провести комплексну оптимізацію ресурсів мережної системи аудіоаналітики із урахуванням особливостей аудіосигналів та динамічного навантаження;

- зменшити затримки передачі аудіосигналів у мережних системах аудіоаналітики на 10%, що покращить функціонування системи в режимі реального часу;

- підвищити точність результатів аналізу за такими критеріями оцінювання, як F1-score та EER, досягнувши збільшення F1-score на 0,9% та зменшення EER на 0,1 відповідно, що свідчить про покращення якості аналізу аудіосигналів.

Отримані в дисертаційній роботі результати можуть бути використані для автоматичної обробки великих обсягів аудіоданих, реалізації систем безпеки, моніторингу, та медичної діагностики, автоматизації телефонії та клієнтської підтримки. Сфера використання – безпека та аудіоспостереження, медична сфера, моніторингові системи, промисловість та автоматизація.

Результати дисертаційної роботи впроваджено у ТОВ НВЕП «Промінь» під час створення та модернізації мережних систем, що дозволило вдосконалити методи збору й обробки звукових сигналів, та в підсумку призвело до підвищення рівня автоматизації, надійності та безпеки систем. Також отримані результати використовуються в освітньому процесі Харківського національного університету радіоелектроніки на кафедрі електронних обчислювальних машин при підготовці кваліфікаційних робіт, в

курсів проектуванні та в дисципліні «Комп'ютерні мережі» відповідно до освітньо-професійної програми 123 Комп'ютерна інженерія для здобувачів першого (бакалаврського) рівня вищої освіти.

Матеріали дисертації викладено у 16-х публікаціях, з них – 2 статті у наукових фахових виданнях України категорії А що індексуються в Scopus та 3 статті у наукових фахових виданнях України категорії Б; 11 тез доповідей у матеріалах міжнародних науково-технічних та науково-практичних конференцій, серед них одна доповідь, що опублікована в матеріалах міжнародної наукової конференції, яка індексуються в Scopus.

Ключові слова: аудіоаналітика, методи машинного навчання, глибинне навчання, модель, метод, мережі, інформаційна система, комп'ютерна система, системи моніторингу, сигнал, подія, оптимізація, канал, обробка, аналіз.

Список публікацій здобувача

1. Poroshenko A., Kovalenko A. Optimization of a basic network in audio analytics systems // *Advanced Information Systems*. 2023. Vol. 7(1). P. 23–28. doi: <https://doi.org/10.20998/2522-9052.2023.1.04> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).

2. Poroshenko A., Kovalenko A. Audio event analysis method in network-based audio analytics systems // *Advanced Information Systems*. 2024. Vol. 8(4). P. 60–64. doi: <https://doi.org/10.20998/2522-9052.2024.4.08> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).

3. Kovalenko A., Poroshenko A. Analysis of the sound event detection methods and systems // *Advanced Information Systems*. 2022. Vol. 6(1) P. 65–69. doi: <https://doi.org/10.20998/2522-9052.2022.1.11> (Фахове видання категорії Б)

4. Порошенко А. І. Математична модель проходження аудіосигналів у мережних системах аудіоаналітики // *Advanced Information Systems*. 2022.

Т. 6(4). С. 25–29. doi: <https://doi.org/10.20998/2522-9052.2022.4.04> (Фахове видання категорії Б.).

5. Poroshenko A., Kovalenko A. Audio signal transmission method in network-based audio analytics system // Innovative technologies and scientific solutions for industries. 2023. Vol. 4(26). P. 58–67. doi: <https://doi.org/10.30837/ITSSI.2023.26.058> (Фахове видання категорії Б.).

6. Poroshenko A., Kovalenko A., Sedlaček P. Organization of Audio Analytics Systems Topologies // 2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T). Kharkiv, Ukraine, 2022. P. 517-521, doi: <https://doi.org/10.1109/PICST57299.2022.10238687> (входить до міжнародної наукометричної бази Scopus)

7. Порошенко А.І., Коваленко А.А. Методи класифікації ознак аудіосигналів // Проблеми інформатизації: тези доп. 9-ї міжнар. наук.-техн. конф., 18-19 листопада 2021 р. Харків, 2021. Т. 1. С. 90.

8. Порошенко А.І., Коваленко А.А. Методи та підходи до детектування та розпізнавання поліфонічних аудіоподій // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: матеріали 12-ї міжнар. наук.-техн. конф., 27-28 квітня 2022 р. Харків, 2022. Т. 1. С. 37.

9. Порошенко А.І., Коваленко А.А. Методи стиснення аудіоданих у системах аудіоаналітики // Проблеми інформатизації: тези доп. десятої міжнар. наук.-техн. конф., 24-25 листопада 2022 р. Харків, 2022. Т. 1. С. 92.

10. Порошенко А.І., Коваленко А.А. Методи зниження шуму звукового сигналу // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : матеріали 13-ї міжнар. наук.-техн. конф., 26-27 квітня 2023 р. Харків, 2023. Т. 2. С. 54.

11. Порошенко А.І., Коваленко А.А. Використання методів шумопоглинання у системах аудіоаналітики // Проблеми інформатизації: тези

доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. Харків, 2023. Т. 3. С. 49.

12. Порошенко А.І., Коваленко А.А., Мартовицький В.О. Архітектура програмно-апаратної моделі наземного обладнання РСБН типу TACAN // Проблеми інформатизації: тези доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. м. Харків, 2023. Т. 3. С. 106.

13. Порошенко А.І., Коваленко А.А. Variability of delays in network-based audio analytics systems // Сучасні інформаційні системи та технології в цифровому суспільстві: матеріали Міжнародної науково-практичної конференції, 18 - 19 квітня 2024 р. м. Харків, 2024. С. 106.

14. Порошенко А.І., Коваленко А.А. Special features of implementing QoS in network-based audio analytics systems // The 23rd International scientific conference information technologies and management 2024, April 25-26. Riga, Latvia. 2024.

15. Порошенко А.І., Коваленко А.А., Буслов П.В. Можливості використання нейронних мереж для атакна мультісервісні // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 2. С. 48.

16. Порошенко А.І., Коваленко А.А. Метод класифікації звукових подій у мережних системах аудіоаналітики // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 3, С. 8.

ABSTRACT

Poroshenko A. Model and Methods of Audio Signal Processing in Network-Based Audio Analytics Systems. – Qualification scientific work in manuscript form.

Dissertation for the degree of Doctor of Philosophy in the specialty 126 Information Systems and Technologies – Kharkiv National University of Radio Electronics, Ministry of Education and Science of Ukraine, Kharkiv, 2025.

The dissertation is devoted to increasing the efficiency of audio signal processing in network-based audio analytic systems. The work proposes the use of mathematical modeling to optimize network processes and refines methods for the transmission and analysis of audio signals, ensuring high accuracy in audio signal analysis, reduced signal transmission delays, and enhanced quality of service in real-time conditions. The relevance of this topic is determined by the fact that the rapid development of technology and the increasing volume of audio data lead to new challenges in the transmission and analysis of audio signals in network-based audio analytic systems. Combination of a growing amount of data, need for real-time processing, and complexity of analysis makes the issue of enhancing the efficiency of audio signal transmission and analysis in network-based audio analytic systems relevant and important for research and development.

The aim of the dissertation is to improve the accuracy of analysis and reduce the transmission delays of audio signals in network-based audio analytic systems by optimizing the network infrastructure and refining the methods for the transmission and analysis of audio signals.

Object of research – the processes of transmitting and analyzing audio signals.

Subject of research – the methods for transmitting and analyzing audio signals in network-based audio analytic systems.

In the introduction, the relevance of using network-based audio analytic systems in various fields, including security, monitoring systems, marketing, and entertainment, is substantiated. Attention is drawn to the importance of transmitting audio signals in real time, since delays can negatively affect system efficiency and response speed. The main components of the audio data processing and analysis process (collection, preprocessing, transmission, feature extraction) are described, and the research tasks are outlined, which include creating a mathematical model, selecting the optimal network structure, and refining the methods for transmitting and analyzing audio signals. The object, subject, and aim of the work are defined, and the scientific novelty is explained, which lies in the new mathematical model of audio signal transmission and the improved methods for audio signal transmission and analysis. The practical application areas of the obtained results are also presented, along with their implementation in the educational process.

In the first chapter of the dissertation presents an analysis of the current state of development of methods for processing audio signals in network-based audio analytic systems. The main directions and trends in this field are examined, and a review of existing systems, their architectures, and functional capabilities is conducted. Key requirements for such systems are identified, including accuracy and processing speed. Based on this analysis, the research tasks aimed at developing a mathematical model of audio signal transmission and refining the methods for audio signal transmission and analysis, which will ensure high reliability and adaptability under different conditions, are formulated.

In the second chapter of the dissertation, the transmission process of audio signals in network-based audio analytic systems is investigated, particularly their collection and transmission. A mathematical model is described for optimizing network resources, taking into account the hierarchical structure of the system and various data transmission routes. Special attention is paid to constructing optimization equations for basic network structures that allow capital expenditures to be minimized while providing high quality of service. The influence of non-

standard Poisson traffic on system efficiency is considered, and methods for selecting the optimal network structure are proposed, which form the basis for further improvement of audio analytic systems.

In the third chapter of the dissertation, the methods for transmitting and analyzing audio signals in network-based audio analytic systems are proposed. The algorithms for connection management, the method for virtual route switching, and the modeling of load sources are described in detail, taking into account the requirements of semantic transparency and time constraints. The role of traffic management and the use of modern network protocols to meet QoS indicators are highlighted, and the complexities arising in noisy environments or heterogeneous network infrastructures are outlined. Special attention is paid to neural networks for audio signal analysis and the application of quality assessment metrics. It is shown that the proposed approaches ensure high processing speed and system reliability in real time.

In the fourth chapter, a comprehensive description of the developed information technology for audio signal processing in a network-based audio analytic system is provided. In particular, the conceptual principles of system design are presented, the choice of the technology stack is substantiated, and approaches to integrating the components for collecting, transmitting, and analyzing audio data in real time are examined. The system architecture and the principles of interaction among its components, aimed at ensuring maximum processing speed, scalability, and reliability, are analyzed in detail. This chapter also presents the results of an experimental evaluation of the system, which confirm its ability to achieve high-precision audio signal analysis and effective operation under complex and heterogeneous network conditions.

In the conclusions of the dissertation, the main research results are summarized, and the significance of the scientific and practical achievements is highlighted. It is shown that the proposed models and methods for transmitting and analyzing audio signals in a network-based audio analytic system ensure increased accuracy and reduced transmission delays.

Research methods. A wide range of methods was used to solve the applied scientific problem. In developing the model of audio signal transmission in network-based audio analytic systems, mathematical modeling methods were employed, including network flow analysis, probabilistic approaches to loss calculation, and optimization algorithms. These methods make it possible to account for the specifics of non-standard loads and ensure minimal capital costs while meeting quality of service requirements. In developing the method for transmitting audio signals in network-based audio analytic systems, methods of data packetization, dynamic resource allocation, and time delay optimization were used, as well as algorithms for queue management and error correction, which allow maintaining the required quality of service level in real time and effectively managing traffic volumes. In developing the method for analyzing audio signals in network-based audio analytic systems, deep learning methods were used, primarily convolutional and recurrent neural networks, which made it possible to separate sound sources and improve the accuracy of audio signal analysis in real time. The evaluation of experimental data obtained during the work was carried out on the basis of methods from probability theory and mathematical statistics.

The choice of research methods ensured the reliability of the obtained results and conclusions, which is confirmed by the convergence of the experimental research results obtained from the software implementation of the information technology for the network-based audio analytic system with the theoretical calculations and numerical modeling, as well as by their consistency with forecast estimates of the effectiveness of the proposed approaches.

Main research tasks are the following:

- conduct an analysis of development directions of audio signal processing methods and network-based audio analytics systems;
- develop a mathematical model of the process of audio signal transmission in network-based audio analytics systems;
- develop a method for selecting the network structure of an audio analytics system;

- improve the method of audio signals transmission through virtual route switching;
- improve the method of audio signals analysis based on a multievent classifier;
- improve the information technology for audio signals processing using improved methods and model.

Scientific novelty of the obtained results is determined by the developed and refined model and methods for processing audio signals in network-based audio analytic systems, within which the following new scientific findings were achieved:

1. For the first time, a mathematical model of the audio signal transmission process in network-based audio analytic systems has been proposed, taking into account non-standard Poisson traffic, which is based on quantitative results from analyzing network structures and makes it possible to increase data transmission efficiency and reduce infrastructure costs.

2. The method of audio signal transmission has been improved, which differs from existing ones by using virtual route switching and allows ensuring high transmission quality, reducing signal transmission delays, and increasing system stability under dynamic network loads.

3. The method of audio signal analysis has been improved by applying a multievent classifier based on deep neural networks, which makes it possible to improve the accuracy of audio signal analysis in complex acoustic environments.

4. The information technology for audio signal processing in network-based audio analytic systems has been further advanced, which is based on improved methods of audio signal transmission and analysis and makes it possible to integrate real-time audio signal analysis processes for various application domains.

An analytical review of the subject area related to the research topic was carried out, methods for detecting both monophonic impulse audio events and polyphonic audio events were analyzed. Current issues related to the development

of such systems are also presented. An overview of modern existing solutions for audio signal processing was conducted, and the goal and objectives of the dissertation research were formulated.

A formulation of the problem of optimizing the network-based audio analytic system based on the criterion of minimal capital costs is proposed. Optimization equations are derived for a unidirectional three-node basic structure, based on minimizing capital costs. Optimization equations for a bidirectional three-node basic network structure are obtained, and the derived equations are generalized to any structure of network-based audio analytic system.

A method for selecting the network structure of the audio analytic system has been proposed. The case of optimizing the network structure under non-standard Poisson loads has been considered. The basic network of the audio analytic system was optimized based on cost criteria, and the optimization results were analyzed.

Modeling of the system's load sources has been performed. Methods for managing connections and traffic in the network-based audio analytic system have been examined. Methods for ensuring the quality of audio signal transmission have been investigated, and a method for audio signal transmission using virtual route switching has been proposed.

A performance metric for audio signal transmission and analysis has been established. Datasets for training, validation, and testing of audio signal analysis methods in network-based audio analytic systems have been examined, and a method for audio signal analysis in network-based audio analytic system has been proposed.

An analysis of the audio signal processing in the network-based audio analytic system has been conducted to monitor its effectiveness. A structure of the information technology for audio signal processing in network-based audio analytic systems that takes into account the processes of audio signal transmission and analysis has been proposed.

The practical significance of the results lies in the fact that the model and methods developed in this work form a scientific and practical foundation for further improvement of network-based audio analytic systems. The engineering methods and algorithms derived from this work enable to:

- perform comprehensive optimization of the resources of a network-based audio analytic system by taking into account the specific features of audio signals and dynamic load;
- reduce the transmission delays of audio signals in network-based audio analytic systems by 10%, which improves the system's performance in real time;
- improve the accuracy of the analysis results according to such evaluation criteria as F1-score and EER by achieving a 0.9% increase in the F1-score and a 0.1 reduction in the EER, respectively, which indicates an improvement in the quality of audio signal analysis.

The results obtained in the dissertation can be used for the automatic processing of large amounts of audio data, the implementation of security systems, monitoring, and medical diagnostics, as well as for the automation of telephony and customer support. The areas of application include security and audio surveillance, the medical field, monitoring systems, industry, and automation.

The results of the dissertation were implemented at LLC SPEE "Promin" during the creation and modernization of network systems, which made it possible to improve the methods of collecting and processing audio signals, leading to an increase in the level of system automation, reliability, and security. Additionally, the obtained results are used in the educational process of Kharkiv National University of Radio Electronics at the department of electronic computers in the preparation of qualification works, course projects, and in the discipline "Computer networks" in accordance with the educational and professional program 123 Computer Engineering for first-level (bachelor's) education students.

Materials of the dissertation are presented in 16 publications, of which 2 articles in Ukrainian scientific professional journals of Category A that are indexed by Scopus and 3 articles in scientific professional journals of Category B; 11

conference abstracts in the materials of international scientific-technical and scientific-practical conferences, among which one abstract is published in the materials of an international scientific conference that is indexed by Scopus.

Keywords: audio analytics, machine learning methods, deep learning, model, method, networks, information system, computer system, monitoring systems, signal, event, optimization, channel, processing, analysis.

List of publications of the applicant

1. Poroshenko A., Kovalenko A. Optimization of a basic network in audio analytics systems // *Advanced Information Systems*. 2023. Vol. 7(1). P. 23–28. doi: <https://doi.org/10.20998/2522-9052.2023.1.04> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).

2. Poroshenko A., Kovalenko A. Audio event analysis method in network-based audio analytics systems // *Advanced Information Systems*. 2024. Vol. 8(4). P. 60–64. doi: <https://doi.org/10.20998/2522-9052.2024.4.08> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).

3. Kovalenko A., Poroshenko A. Analysis of the sound event detection methods and systems // *Advanced Information Systems*. 2022. Vol. 6(1) P. 65–69. doi: <https://doi.org/10.20998/2522-9052.2022.1.11> (Фахове видання категорії Б)

4. Порошенко А. І. Математична модель проходження аудіосигналів у мережних системах аудіоаналітики // *Advanced Information Systems*. 2022. Т. 6(4). С. 25–29. doi: <https://doi.org/10.20998/2522-9052.2022.4.04> (Фахове видання категорії Б).

5. Poroshenko A., Kovalenko A. Audio signal transmission method in network-based audio analytics system // *Innovative technologies and scientific solutions for industries*. 2023. Vol. 4(26). P. 58–67. doi: <https://doi.org/10.30837/ITSSI.2023.26.058> (Фахове видання категорії Б).

6. Poroshenko A., Kovalenko A., Sedlaček P. Organization of Audio Analytics Systems Topologies // *2022 IEEE 9th International Conference on*

Problems of Infocommunications, Science and Technology (PIC S&T). Kharkiv, Ukraine, 2022. P. 517-521, doi: <https://doi.org/10.1109/PICST57299.2022.10238687> (входить до міжнародної наукометричної бази Scopus)

7. Порошенко А.І., Коваленко А.А. Методи класифікації ознак аудіосигналів // Проблеми інформатизації: тези доп. 9-ї міжнар. наук.-техн. конф., 18-19 листопада 2021 р. Харків, 2021. Т. 1. С. 90.

8. Порошенко А.І., Коваленко А.А. Методи та підходи до детектування та розпізнавання поліфонічних аудіоподій // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: матеріали 12-ї міжнар. наук.-техн. конф., 27-28 квітня 2022 р. Харків, 2022. Т. 1. С. 37.

9. Порошенко А.І., Коваленко А.А. Методи стиснення аудіоданих у системах аудіоаналітики // Проблеми інформатизації: тези доп. десятої міжнар. наук.-техн. конф., 24-25 листопада 2022 р. Харків, 2022. Т. 1. С. 92.

10. Порошенко А.І., Коваленко А.А. Методи зниження шуму звукового сигналу // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : матеріали 13-ї міжнар. наук.-техн. конф., 26-27 квітня 2023 р. Харків, 2023. Т. 2. С. 54.

11. Порошенко А.І., Коваленко А.А. Використання методів шумопоглинання у системах аудіоаналітики // Проблеми інформатизації: тези доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. Харків, 2023. Т. 3. С. 49.

12. Порошенко А.І., Коваленко А.А., Мартовицький В.О. Архітектура програмно-апаратної моделі наземного обладнання РСБН типу TACAN // Проблеми інформатизації: тези доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. м. Харків, 2023. Т. 3. С. 106.

13. Порошенко А.І., Коваленко А.А. Variability of delays in network-based audio analytics systems // Сучасні інформаційні системи та технології в

цифровому суспільстві: матеріали Міжнародної науково-практичної конференції, 18 - 19 квітня 2024 р. м. Харків, 2024. С. 106.

14. Порошенко А.І., Коваленко А.А. Special features of implementing QoS in network-based audio analytics systems // The 23rd International scientific conference information technologies and management 2024, April 25-26. Riga, Latvia. 2024.

15. Порошенко А.І., Коваленко А.А., Буслов П.В. Можливості використання нейронних мереж для атакна мультісервісні // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 2. С. 48.

16. Порошенко А.І., Коваленко А.А. Метод класифікації звукових подій у мережних системах аудіоаналітики // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 3, С. 8.

ЗМІСТ

Перелік умовних скорочень	24
Вступ.....	26
1 Аналіз наукового завдання вибору методів обробки аудіосигналів у мережних системах аудіоаналітики	36
1.1 Аналіз основних напрямків розвитку методів обробки аудіосигналів у мережних системах аудіоаналітики	36
1.2 Аналіз існуючих топологій мережних систем аудіоаналітики	38
1.3 Основні вимоги до мережних систем аудіоаналітики.....	49
1.4 Постановка завдання дослідження	59
Висновки за розділом 1	60
2 Моделювання процесу проходження аудіосигналів у мережних системах аудіоаналітики.....	61
2.1 Визначення особливостей передачі аудіосигналів у мережних системах аудіоаналітики.....	61
2.2 Формалізований опис мережної системи аудіоаналітики	64
2.3 Розроблення математичної моделі процесу проходження аудіосигналів у мережних системах аудіоаналітики	67
2.4 Оптимізація базової мережі системи аудіоаналітики.....	75
2.4.1 Метод вибору мережної структури системи аудіоаналітики	75
2.4.2 Оптимізація мережної структури системи аудіоаналітики при неординарному пуасонівському навантаженні.....	83
2.4.3 Оптимізація базової мережі системи аудіоаналітики за критерієм вартості.....	88
Висновки за розділом 2	98
3 Розроблення методів обробки аудіосигналів у мережних системах аудіоаналітики	100
3.1 Обґрунтування показника якості передачі та аналізу аудіосигналу....	100
3.2 Інформаційна безпека мережної системи аудіоаналітики	102

3.3	Метод передачі аудіосигналів у мережній системі аудіоаналітики ...	109
3.3.1	Моделювання джерел навантаження системи	111
3.3.2	Управління з'єднаннями у мережній системі аудіоаналітики	118
3.3.3	Управління трафіком у мережній системі аудіоаналітики	119
3.3.4	Забезпечення якості передачі аудіосигналу	123
3.3.5	Дотримання часових вимог	128
3.3.6	Метод передачі аудіосигналу за рахунок комутації віртуальних маршрутів.....	132
3.4	Метод аналізу аудіосигналів у мережних системах аудіоаналітики ...	138
3.4.1	Показник якості аналізу аудіосигналу	139
3.4.2	Набір даних	140
3.4.3	Метод аналізу аудіосигналу з використанням багатоподійного класифікатора	142
	Висновки за розділом 3	147
4	Дослідження інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики	149
4.1	Розробка інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики	149
4.2	Програмна реалізація інформаційної технології обробки аудіосигналів	156
4.2.1	Вибір технологічного стеку та архітектури	157
4.2.2	Реалізація основних компонентів.....	160
4.3	Оцінка ефективності запропонованої інформаційної технології.....	164
4.3.1	Оцінка за часом проходження сигналу	164
4.3.2	Оцінка за точністю розпізнавання.....	166
	Висновки за розділом 4	169
	Висновки	170
	Перелік джерел посилання	172
	Додаток А Список публікацій здобувача	186
	Додаток Б Документи, що підтверджують впровадження результатів	190

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

- ВК – Вузли комутації
- ЗЛ – З'єднувальні лінії
- ЛС – Локальний сервер
- МВП – Метод випадкового пошуку
- МЗЛ – Мережі з'єднувальних ліній
- ОПШ – Обхідні проміжні шляхи
- ППАС – Пристрої прийому аудіосигналів
- ПШ – Прямі шляхи
- СП – Семантична прозорість
- ЦОАС – Центр обробки аудіосигналів
- ШКП – Швидка комутація пакетів
- ШОВ – Шлях останнього вибору
- АТМ – Асинхронний режим передачі (англ. Asynchronous Transfer Mode)
- BER – Коефіцієнт помилок по бітам (англ. Bit Error Rate)
- САС – Керування допуском з'єднання (англ. Connection Admission Control)
- CNN – Конволюційна нейронна мережа (англ. Convolutional Neural Network)
- DCASE – Виявлення та класифікація акустичних сцен та подій (англ. Detection and Classification of Acoustic Scenes and Events)
- EER – Помилка виявлення подій (англ. Event Error Rate)
- FC – Управління зворотним зв'язком (англ. Feedback Control)
- FN – Хибно негативні (англ. False Negatives)
- FP – Хибно позитивні (англ. False Positives)
- GCRA – Узагальнений алгоритм регулювання швидкості передачі осередків (англ. Generic Cell Rate Algorithm)

GDPR – Загальний регламент захисту даних (англ. General Data Protection Regulation)

IDS – Системи виявлення вторгнень (англ. Intrusion Detection Systems)

IPS – Системи запобігання вторгненням (англ. Intrusion Prevention Systems)

ITU – Міжнародна спілка електрозв'язку (англ. International Telecommunication Union)

LBA – Алгоритм дірявого відра (англ. Leaky Bucket Algorithm)

NNI – Інтерфейс мережа-вузол (англ. Network-Node Interface)

PER – Коефіцієнт помилок по пакетах (англ. Packet Error Rate)

QoS – Якість обслуговування (англ. Quality of Service)

RNN – Рекурентна нейронна мережа (англ. Recurrent Neural Network)

SRTP – Протокол безпечної передачі в реальному часі (англ. Secure Real-time Transport Protocol)

TLS – Безпека транспортного рівня (англ. Transport Layer Security)

TP – Істинно позитивні (англ. True Positives)

UNI – Інтерфейс користувач-мережа (англ. User-Network Interface)

UPC – Керування параметрами використання (англ. Usage Parameter Control)

UPC/NPC – Управління параметрами навантаження, що створюється користувачем, та управління параметрами мережі (англ. Usage/Network Parameter Control)

VC – Віртуальний канал (англ. Virtual Channel)

VP – Віртуальний шлях (англ. Virtual Path)

VPN – Віртуальна приватна мережа (англ. Virtual Private Network)

VSA – Алгоритм віртуального розкладу (англ. Virtual Scheduling Algorithm)

ВСТУП

Актуальність теми дослідження.

Аудіоаналітика – це технологія, яка використовує алгоритми обробки сигналів та глибоке навчання для аналізу аудіоданих. Ця технологія може бути застосована в різних областях, включаючи безпеку, охорону здоров'я, моніторинг та розваги. Наприклад, вона може використовуватися для виявлення специфічних звуків або мовних патернів у навколишньому середовищі, аналізу емоцій в голосі, розпізнавання мови тощо. До основних компонентів мережної системи аудіоаналітики зазвичай входять збір аудіоданих навколишнього середовища, попередня обробка зібраних даних, передача аудіосигналів до сервера для обробки, вилучення ознак та обробка отриманих аудіосигналів, а також засоби зворотнього зв'язку.

Процес передачі аудіосигналів у мережних системах аудіоаналітики відіграє важливу роль у розвитку та впровадженні сучасних технологій, спрямованих на вдосконалення безпеки, зручності та ефективності різноманітних сфер життя.

Перш ніж проводити аналіз аудіосигналів, спочатку необхідно провести їх попередню обробку та передати дані до сервера. Попередня обробка аудіоданих є критично важливим етапом в процесі аналізу, оскільки вона готує сирі аудіодані для подальшої обробки та аналізу. Попередня обробка може включати фільтрацію шуму, вирівнювання, сегментацію, видалення еха.

У великій мірі важливою є здатність використовуваних методів передачі та аналізу аудіосигналу працювати у режимі, наближеному до реального часу. Затримка при передачі аудіосигналів може негативно впливати на швидкість реакції системи, що є критичним для сценаріїв безпеки або моментального управління. Отже, методи передачі та аналізу аудіосигналів у мережних системах аудіоаналітики стають ключовим

інструментом для покращення різноманітних аспектів повсякденного життя, надаючи можливість використовувати звукову інформацію з максимальною ефективністю та безпекою.

В сучасному світі зростаюча кількість інформації, яка поступає у мережу, вимагає розвитку ефективних методів її обробки. Однією з важливих складових цієї обробки є аналіз аудіосигналів у мережних системах аудіоаналітики, що дозволяє виявляти та класифікувати різноманітні звукові явища в реальному часі або на основі записів, використовуючи різні алгоритми та методи аналізу аудіосигналів.

Розпізнавання аудіоподій у мережних системах аудіоаналітики – це процес ідентифікації та класифікації різних аудіоподій, які відбуваються в аудіозаписах, за допомогою аналізу їхнього змісту та контексту. Цей процес може бути використаний для автоматичного визначення подій, таких як голосові команди, шуми, мовлення, звуки транспорту, вибухи, погодні умови та інші. Метод розпізнавання аудіоподій має широкий спектр застосувань, починаючи від систем безпеки та аудіоспостереження, і закінчуючи автоматизованим аналізом акустичного оточення в медичних дослідженнях.

Отже, актуальною є науково-прикладна задача підвищення ефективності передачі та аналізу аудіосигналів у мережних системах аудіоаналітики в умовах зростаючого обсягу аудіоданих, що потребують обробки та посилення вимог до швидкості та точності аналізу, за рахунок розробки та удосконалення відповідних моделей і методів.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційна робота виконана на кафедрі електронних обчислювальних машин Харківського національного університету радіоелектроніки згідно з планом наукових робіт в рамках держбюджетної НДР «Розроблення системотехнічних основ та робочої моделі побудови наземного комплексу вітчизняної радіотехнічної системи ближньої навігації типу TACAN» (ДР № 0122U200989).

Мета дослідження – підвищення точності аналізу та зменшення затримок передачі аудіосигналів у мережних системах аудіоаналітики за рахунок оптимізації мережної інфраструктури та удосконалення методів передачі і аналізу аудіосигналів.

Для досягнення зазначеної мети поставлені основні завдання дослідження:

- провести аналіз напрямків розвитку методів обробки аудіосигналів та мережних систем аудіоаналітики;
- розробити математичну модель процесу передачі аудіосигналів у мережних системах аудіоаналітики;
- розробити метод вибору мережної структури системи аудіоаналітики;
- удосконалити метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів;
- удосконалити метод аналізу аудіосигналів на основі багатоподійного класифікатора;
- удосконалити інформаційну технологію обробки аудіосигналів з використанням вдосконалених моделі і методів.

Об'єкт дослідження – процеси передачі та аналізу аудіосигналів.

Предмет дослідження – методи передачі та аналізу аудіосигналів у мережних системах аудіоаналітики.

Методи дослідження. При розв'язанні науково-прикладної задачі використано широкий спектр методів. Так, при розробці моделі процесу проходження аудіосигналів у мережних системах аудіоаналітики використовувались методи математичного моделювання, зокрема аналіз мережних потоків, імовірнісні підходи до обчислення втрат та оптимізаційні алгоритми, що дозволяє врахувати специфіку неординарного навантаження та забезпечити мінімальні капітальні витрати за умови дотримання вимог до якості обслуговування. При розробці методу передачі аудіосигналів у мережних системах аудіоаналітики використовувались методи пакетизації

даних, динамічного розподілу ресурсів та оптимізації часової затримки, а також алгоритми формування черг і корекції помилок, які дають змогу забезпечити потрібний рівень якості обслуговування в реальному часі та ефективно керувати обсягами трафіку. При розробці методу аналізу аудіосигналів у мережних системах аудіоаналітики використовувались методи глибинного навчання, насамперед нейронні мережі конволюційної та рекурентної архітектури, що дали змогу розділяти джерела звуку й підвищувати точність аналізу аудіосигналів у реальному часі. Оцінка експериментальних даних, отриманих у ході роботи, проводилася на основі методів теорії ймовірності та математичної статистики.

Вибір методів досліджень забезпечив достовірність отриманих результатів і висновків, що підтверджується збіжністю результатів експериментальних досліджень, отриманих при програмній реалізації інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики, з теоретичними розрахунками та чисельним моделюванням, а також узгодженням із прогнозними оцінками ефективності запропонованих підходів.

Наукова новизна отриманих результатів обумовлена розробленими та удосконаленими моделлю і методами обробки аудіосигналів у мережних системах аудіоаналітики, в межах яких отримані такі нові наукові результати:

1. Вперше запропоновано математичну модель процесу передачі аудіосигналів у мережних системах аудіоаналітики з урахуванням неординарного пуасонівського навантаження, що базується на кількісних результатах аналізу мережних структур та дозволяє підвищити ефективність передачі даних і зменшити витрати на інфраструктуру;

2. Удосконалено метод передачі аудіосигналів, який відрізняється від існуючих використанням комутації віртуальних маршрутів, та дозволяє забезпечити високу якість передачі, знизити затримки у передачі сигналів і підвищити стабільність системи в умовах динамічних мережних навантажень;

3. Отримав подальший розвиток метод аналізу аудіосигналів шляхом застосування багатоподійного класифікатора, заснованого на глибинних нейронних мережах, що дозволяє підвищити точність аудіосигналів у складних акустичних середовищах;

4. Отримала подальший розвиток інформаційна технологія обробки аудіосигналів у мережних системах аудіоаналітики, яка базується на вдосконалених методах передачі і аналізу аудіосигналів та дозволяє інтегрувати процеси аналізу аудіосигналів з реальним часом для різних галузей застосування.

Практичне значення результатів роботи полягає в тому, що розроблені у роботі модель та методи є науково-практичною основою для подальшого удосконалення мережних систем аудіоаналітики. Представлені на їх основі методи та алгоритми дають змогу:

- провести комплексну оптимізацію ресурсів мережної системи аудіоаналітики із урахуванням особливостей аудіосигналів та динамічного навантаження;

- зменшити затримки передачі аудіосигналів у мережних системах аудіоаналітики на 10%, що покращить функціонування системи в режимі реального часу;

- підвищити точність результатів аналізу за такими критеріями оцінювання, як F1-score та EER, досягнувши збільшення F1-score на 0,9% та зменшення EER на 0,1 відповідно, що свідчить про покращення якості аналізу аудіосигналів.

Отримані в дисертаційній роботі результати можуть бути використані для автоматичної обробки великих обсягів аудіоданих, реалізації систем безпеки, моніторингу, та медичної діагностики, автоматизації телефонії та клієнтської підтримки. Сфера використання – безпека та аудіопостереження, медична сфера, моніторингові системи, промисловість та автоматизація.

Результати дисертаційної роботи впроваджено у ТОВ НВЕП «Промінь» під час створення та модернізації мережних систем, що дозволило

вдосконалити методи збору й обробки звукових сигналів, та в підсумку призвело до підвищення рівня автоматизації, надійності та безпеки систем. Також отримані результати використовуються в освітньому процесі Харківського національного університету радіоелектроніки на кафедрі електронних обчислювальних машин при підготовці кваліфікаційних робіт, в курсовому проектуванні та в дисципліні «Комп'ютерні мережі» відповідно до освітньо-професійної програми 123 Комп'ютерна інженерія для здобувачів першого (бакалаврського) рівня вищої освіти.

В [1] досліджується оптимізація базових мережних систем аудіоаналітики. Вона включає метод вибору мережної структури системи аудіоаналітики та наводяться приклади застосування оптимізаційних методів у мережних системах аудіоаналітики. В [2] пропонується метод аналізу аудіосигналів у мережних системах аудіоаналітики з використанням багатоподійного класифікатора. Методика базується на використанні алгоритмів глибокого навчання для класифікації аудіосигналів, які містять різні звукові події. В [3] проводиться аналіз різних методів та систем виявлення звукових подій. Вона охоплює огляд сучасних технологій та алгоритмів, що використовуються для автоматичного виявлення та класифікації аудіосигналів та виконується їх порівняння. В [4] описується математична модель передачі аудіосигналів у мережних системах аудіоаналітики. Вона вивчає основні аспекти передачі даних у таких системах, включаючи затримки, пропускну здатність та ефективність. Також розглядається вплив різних факторів, таких як обсяг даних та характеристики мережі, на процес передачі та обробки аудіосигналів. В статті [5] описується метод передачі аудіосигналу за рахунок комутації віртуальних маршрутів. Також досліджується вплив різних факторів, таких як джерело аудіосигналу та характеристики мережі, на якість передачі та аналізу даних.

В [6] розглядається організації топологій систем аудіоаналітики. В цих тезах досліджуються різні способи побудови мережних систем, що використовуються для аналізу аудіосигналів. Доповідь включає огляд

централізованих та розподілених архітектур, їх переваги та недоліки, а також питання масштабованості та надійності. В [7] розглядаються різні методи класифікації ознак аудіосигналів, описуються різні характеристики аудіосигналів та методи їх використання для класифікації, включаючи часові, частотні та спектральні характеристики. В [8] досліджуються методи та підходи до виявлення та розпізнавання поліфонічних аудіоподій, описуються алгоритми, що використовуються для розпізнавання складних аудіосигналів, що містять декілька звукових джерел одночасно. В [9] розглядаються різні методи стиснення аудіоданих у системах аудіоаналітики та їх вплив на якість аналізу аудіоданих. В [10] описуються алгоритми та методи фільтрації шуму для покращення якості аудіосигналів у системах аудіоаналітики. В [11] досліджуються різні методи шумопоглинання та їх використання у системах аудіоаналітики, розглядається ефективність різних алгоритмів шумопоглинання та їхній вплив на якість аналізу аудіосигналів. В [12] розглядається архітектура програмно-апаратної моделі наземного обладнання РСБН типу TACAN, описується склад та функціональні можливості цього обладнання, його основні характеристики та принципи роботи. В [13] досліджуються фактори, що впливають на загальні затримки в мережних системах аудіоаналітики, та методи боротьби з затримками. В [14] досліджуються особливості забезпечення якості обслуговування у мережних системах аудіоаналітики. В роботі розглядаються методи класифікації трафіку та управління чергами для забезпечення стабільності передачі аудіоданих в режимі реального часу. В [15] розглядаються можливості застосування нейронних мереж для здійснення атак на мультисервісні мережі, акцентуючи увагу на зростанні значущості штучного інтелекту у кіберзагрозах. В [16] представлено метод класифікації звукових подій з використанням глибинних нейронних мереж, що підвищує точність класифікації.

Особистий внесок здобувача. Всі основні результати дисертаційної роботи, які виносяться на захист, отримано автором особисто. В [1] здобувач

провів дослідження з оптимізації базових мережних систем аудіоаналітики, запропонувавши метод вибору мережної структури та навівши приклади застосування оптимізаційного методу у мережних системах аудіоаналітики. В [2] удосконалив метод аналізу аудіосигналів у мережних системах аудіоаналітики з використанням багатоподійного класифікатора, застосувавши алгоритми глибинного навчання для класифікації аудіосигналів, що містять різні звукові події. В роботі [3] здобувач провів аналіз різних методів та систем виявлення звукових подій, зосереджуючись на сучасних технологіях та алгоритмах, що використовуються для автоматичного виявлення та класифікації аудіосигналів. В роботі [4] здобувач розробив математичну модель процесу передачі аудіосигналів у мережних системах аудіоаналітики, звернувши увагу на основні аспекти передачі даних, такі як затримки, пропускну здатність та витрати. В [5] здобувач удосконалив метод передачі аудіосигналу за рахунок комутації віртуальних маршрутів та дослідив вплив різних факторів, таких як джерело аудіосигналу та характеристики мережі, на якість передачі та аналізу даних. В тезах доповідей [6], що виконані у співавторстві, здобувач дослідив різні способи організації топологій систем аудіоаналітики, зосереджуючись на побудові мережних систем для аналізу аудіосигналів. В [7] здобувач розглянув різні методи класифікації ознак аудіосигналів, визначивши різні характеристики аудіосигналів та методи їх використання для класифікації, включаючи часові, частотні та спектральні характеристики. В [8] здобувач дослідив методи та підходи до виявлення та розпізнавання поліфонічних аудіоподій, аналізуючи алгоритми, що використовуються для розпізнавання складних аудіосигналів, що містять декілька звукових джерел одночасно. В [9, 10, 11] здобувач розглянув різні методи стиснення аудіоданих, фільтрації шуму та шумопоглинання у системах аудіоаналітики для покращення якості аудіосигналів у системах аудіоаналітики. В тезах доповідей [12], що виконані у співавторстві, здобувач провів аналіз результатів та їх візуалізацію. В [13] здобувач дослідив фактори, що впливають на загальні затримки в мережних

системах аудіоаналітики, та розглянув методи боротьби з затримками. В роботі [14] здобувачем проаналізовано методи забезпечення якості обслуговування у мережних аудіоаналітичних системах, включаючи класифікацію трафіку та управління чергами для оптимізації передачі аудіоданих у режимі реального часу. В роботі [15] здобувач здійснив аналіз можливостей застосування нейронних мереж у кіберзагрозах, дослідив сценарії кібератак із використанням штучного інтелекту та запропонував підходи для підвищення ефективності систем виявлення аномалій. В [16] здобувач розробив та випробував метод класифікації звукових подій у мережних системах аудіоаналітики на основі глибоких нейронних мереж, досягнувши підвищеної точності класифікації.

Апробація результатів дисертації.

Основні положення дисертаційної роботи представлено на таких міжнародних конференціях і форумах:

- на Міжнародній науково-технічній конференції «Проблеми інформатизації», Черкаси – Баку – Бельсько-Бяла – Харків, 2021, 2022, 2023, 2024 р.;

- на Міжнародній науково-технічній конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління», Баку – Харків – Жиліна, 2022, 2023, 2024 р.;

- на Міжнародній конференції «Problems of Infocommunications, Science and Technology (PIC S&T)», 2022, 2024 р.;

- на Міжнародній науково-практичній конференції «Сучасні інформаційні системи та технології в цифровому суспільстві», Харків: ХНЕУ імені Семена Кузнеця, 2024 р.;

- на Міжнародній конференції «Information technology and management», Ріга, Латвія, 2024 р.

Публікації. Матеріали дисертації викладено у 16-х публікаціях, з них – 2 статті у наукових фахових виданнях України категорії А що індексуються в Scopus та 3 статті у наукових фахових виданнях України категорії Б; 11 тез

доповідей у матеріалах міжнародних науково-технічних та науково-практичних конференцій, серед них одна доповідь, що опублікована в матеріалах міжнародної наукової конференції, яка індексуються в Scopus..

Структура та обсяг роботи. Дисертація складається зі вступу, 4 розділів, висновків, списку використаних джерел, додатків. Загальний обсяг роботи складає 193 сторінки тексту, що містять 150 сторінок основного тексту, анотації на 21 сторінках, 50 рисунків, 13 таблиць, список використаних джерел з 105 найменувань на 14 сторінках, 2 додатків на 8 сторінках.

1 АНАЛІЗ НАУКОВОГО ЗАВДАННЯ ВИБОРУ МЕТОДІВ ОБРОБКИ АУДІОСИГНАЛІВ У МЕРЕЖНИХ СИСТЕМАХ АУДІОАНАЛІТИКИ

Виявлення та розпізнавання гучних звуків та характерних шумів може істотно підвищити рівень безпеки та забезпечити своєчасне реагування на різні аварійні ситуації. Виявлення аудіоподій є першим кроком обробки аудіосигналів з безперервним вхідним потоком інформації. Перший розділ дисертації присвячено аналізу основних методів обробки аудіосигналів у мережних системах аудіоаналітики. Спочатку розглянуто основні напрями розвитку в цій сфері, акцентуючи на процесі аналізу аудіосигналів. Подано огляд існуючих топологій мережних систем аудіоаналітики з урахуванням їхніх архітектурних особливостей, що дозволяє зрозуміти, як еволюція методів обробки аудіосигналів впливає на вдосконалення функціональності таких систем. Визначено ключові вимоги до мережних систем аудіоаналітики, зокрема щодо продуктивності, точності й швидкості обробки, що обґрунтовує необхідність використання високоефективних методів обробки аудіоданих. На основі проведеного аналізу сформульовано завдання дослідження, яке полягає в розробці оптимізованих методів передачі та аналізу аудіосигналів у мережних системах аудіоаналітики. Цей розділ закладає основу для подальшого розроблення та вдосконалення підходів до обробки аудіоданих, спрямованих на підвищення ефективності та надійності мережних аудіосистем.

1.1 Аналіз основних напрямків розвитку методів обробки аудіосигналів у мережних системах аудіоаналітики

Мережна система аудіоаналітики є інформаційною системою, що складається із сукупності апаратних і програмних засобів. Вона забезпечує захоплення, передавання, аналіз та зберігання аудіосигналів у реальному часі

за допомогою мережних технологій. Усе це реалізується на основі розподілених комп'ютерних систем, які організують обробку та зберігання даних.

Розвиток методів обробки аудіосигналів у мережних системах аудіоаналітики відіграє ключову роль у сучасних технологіях, адже надає можливість автоматично аналізувати звук для розпізнавання мовлення, виявлення різних акустичних подій, усунення шумів та виконання багатьох інших завдань. Одним із найважливіших напрямів тут є розпізнавання мови та аналіз мовлення [17], де застосовують методи машинного навчання, зокрема глибокі і рекурентні нейронні мережі. Сучасні системи вже вміють працювати в реальному часі й навіть здатні визначати емоційний стан мовця. Ще одним важливим завданням є виявлення акустичних подій [18], наприклад голосових команд, свистків чи вибухів, що суттєво допомагає в системах безпеки та моніторингу довкілля. Крім того, існують методики для покращення загальної якості звуку за допомогою фільтрації шумів та усунення артефактів [19].

Також значне місце посідає виявлення аномальних чи критичних ситуацій, адже аналіз аудіосигналів дає змогу розпізнавати нетипові події або небезпечні стани. У медицині це, зокрема, аналіз серцевих звуків [20] для оперативного виявлення порушень у роботі серця, таких як шуми чи аритмії. Завдяки методам машинного навчання і вдосконаленим алгоритмам обробки сигналів можливе швидке визначення критичних станів та вчасна діагностика. Широкі перспективи відкриває і аналіз інтонації та емоційного стану людини [21], що має попит у медицині, психології й розважальній сфері.

Ще одним цікавим напрямом вважають синтез мовлення [22], коли на основі вхідних даних формують штучно згенерований голос. Такі технології корисні для створення голосових помічників або для людей, які втратили здатність говорити. Також дедалі частіше поєднують аудіо- та відеоаналітику [23], формуючи комплексні системи для відстеження та

аналізу поведінкових особливостей. Усі ці підходи й розробки постійно вдосконалюються завдяки прогресу в машинному навчанні, нейронних мережах і обчислювальних технологіях, а їх застосування охоплює широке коло галузей від промислових і безпекових рішень до медицини та сфери розваг.

1.2 Аналіз існуючих топологій мережних систем аудіоаналітики

Гучні крики з відтінками агресії або паніки, звук розбитого скла, сильний гуркіт чи постріли майже завжди свідчать про загрозу на будь-якому об'єкті. Пристрої, що мають вбудовані алгоритми аудіоаналітики й реагують на тривожні звукові хвилі, допомагають вчасно виявити вторгнення, несанкціоновані злами, дорожньо-транспортні пригоди, агресивну поведінку в громадських місцях, збройні напади, пограбування й навіть терористичні акти. Саме завдяки такій функціональності подібні системи аудіоаналітики затребувані й у приватному секторі, і в невеликих магазинах, і в масштабних торговельних центрах, а також на транспортних об'єктах, промислових підприємствах та в багатьох інших напрямках.

Часто аудіоаналітика поєднується з відеоаналітикою, оскільки вони доповнюють одна одну. У порівнянні з аналізом відеопотоків, обробка аудіоданих є менш затратною і значно простішою з погляду використання обчислювальних ресурсів. Також, аудіоаналітика майже не дає хибних спрацьовувань і зазвичай забезпечує результати, які легко тлумачити однозначно. Завдяки технологіям, що дають змогу приймати звук (нерідко паралельно з оперативною обробкою), можна розробляти автономні рішення для профілактики злочинів, підтримки безпеки та розслідування надзвичайних ситуацій. У комплексі з відеоспостереженням аудіоаналітика стає вкрай важливим доповненням для більшості сучасних моніторингових систем, адже оператор безпеки отримує повніше уявлення про те, що відбувається, коли має одночасно і візуальні, і звукові дані. Так само, як різні

алгоритми відеоаналітики дають змогу фіксувати певні події та генерувати тривожні сигнали, аналіз аудіопотоку дає можливість відстежувати незвичайні звуки й оперативно реагувати на них у разі потреби.

Між прийомом та записом звуку є різниця. Як правило, засоби аудіоаналітики не ведуть постійний запис звуку. Зазвичай вони лише обробляють вхідний сигнал у пошуках характерних структур, гучності чи тембру. Якщо засоби аналітики працюють на периферії системи, цифрові аудіодані можуть взагалі не залишати периферії, до центру обробки надходять лише результати аналізу – метадані або сигнали керування. Звук можна приймати та використовувати, не записуючи. Прийом звуку означає його оцифровку і перетворення у форму, придатну для програмної обробки. Це робиться шляхом перетворення звукових хвиль у повітрі в аналоговий електричний сигнал за допомогою мікрофона, перетворення аналогового сигналу на цифрову форму та передачі отриманого цифрового сигналу на обробку.

Якщо прийнятий звук не зберігається на постійному носії, наприклад, флеш-накопичувачі або жорсткому диску, запису не відбувається. Запис може бути не потрібний у багатьох варіантах застосування, наприклад, коли прийнятий звук прослуховується в реальному часі оператором. У деяких ситуаціях можуть бути конкретні причини відмовитися від запису звуку. На запис та простий прийом звуку можуть накладатися різні законодавчі обмеження [24]. Як і у випадку відеоспостереження, перед встановленням аудіоспостереження необхідно вивчити та зрозуміти юридичні аспекти такої системи. У разі потреби потрібно подати відповідні заявки та отримати дозволи. Там, де необхідно, мають бути встановлені інформаційні знаки та повідомлення.

Використання та запис звуку можуть бути заборонені або можуть вимагати спеціального розгляду з різних причин відповідно до національного законодавства або різного роду місцевими законами та нормативами. У деяких регіонах та ситуаціях прийом звуку може бути дозволено, а запис

заборонено. Для своєї роботи засоби аудіоаналітики тимчасово буферизують аудіодані. Багато систем можна налаштувати таким чином, щоб записувати лише дані, буферизовані незадовго до і після моменту виявлення події, щоб забезпечити надійну перевірку та, можливо, зберегти звук як свідчення для подальшого розслідування.

Майже всі події у світі супроводжуються характерними звуками. Звуковий сигнал є показником, який вказує оточенню, що певна подія трапилася, він вказує на факт випадку події та дозволяє отримати необхідну інформацію щодо неї. Цей звуковий сигнал сприймає певне обладнання для прийому звуку. Людський слух є одним із найдосконаліших інструментів для виявлення та аналізу звуків. Людське вухо та мозок здатні виявляти та інтерпретувати звуки навіть у дуже шумних умовах, коли більшість машинних алгоритмів відмовляють. За допомогою слуху люди здатні отримувати просторову інформацію про обстановку, визначати звідки прийшов звук і чи рухається його джерело. Маючи два вуха, людина може визначити, чи прийшов звук ліворуч, праворуч або з якогось іншого напрямку. Людський мозок аналізує різницю між сигналами від обох вух, миттєво вловлюючи затримки мікросекундного масштабу виявлення подій. Хоча сучасні периферійні пристрої аудіоаналітичних систем не дозволяють досягти результатів, подібних до людського вуха, ці системи будуються з оглядом на його структуру. Після прийому звуку, інформацію готують до наступного кроку обробки. Структурна схема типів підготовки даних наведена на рис. 1.1.

Перетворення означає те, що аудіодані абстрагуються та перетворюються, наприклад, у візуальну інформацію діаграми звукового спектру, спектрограму [25]. Отриманий звук можна візуалізувати на моніторі у вигляді діаграм спектра звуку. Така візуалізація можлива завдяки виконанню перетворення Фур'є для вхідного сигналу. Крім того, швидке перетворення Фур'є може бути виконане за допомогою сучасних графічних процесорів [26]. Така візуалізація звуку може бути корисною у ситуаціях,

коли необхідно одночасно стежити за звуками з множини джерел, наприклад, з кількох палат у лікарні. У той час як слухати відразу велику кількість джерел може бути складно, спостерігати на екрані візуалізації набагато простіше.

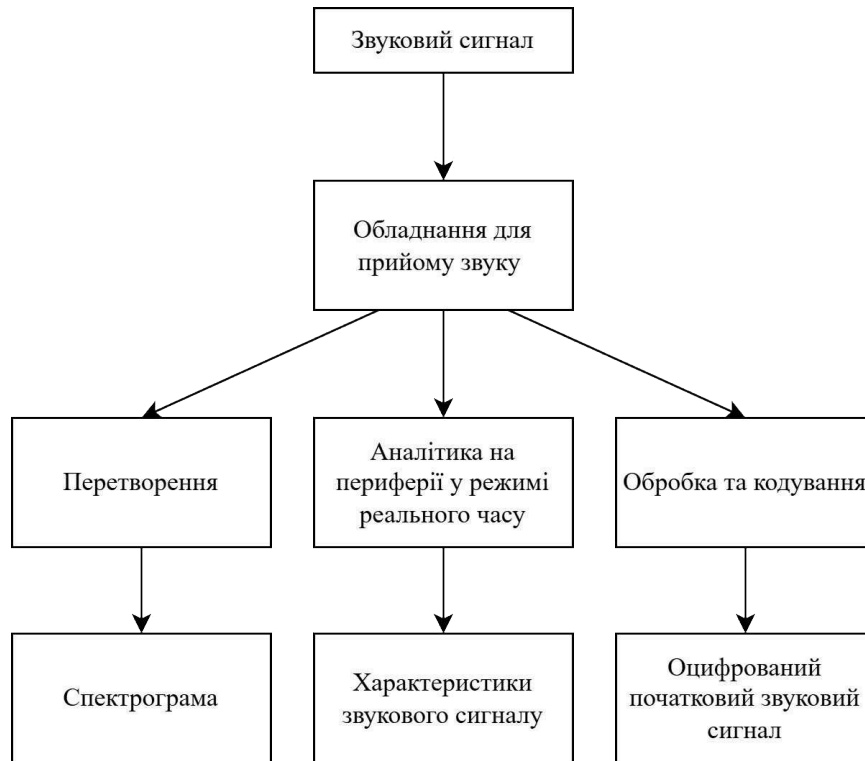


Рисунок 1.1 – Структурна схема типів підготовки даних

Аналітика на периферії у режимі реального часу дозволяє діяти віддалено замість того, щоб пересилати звукові дані до центрів обробки даних або хмар для фільтрації та аналізу. За своїм принципом, аналітика на периферії схожа на IoT [27]. Додатково, така аналітика знімає вимоги до каналу передачі даних, тому що кількість даних, що описують характеристики звукового сигналу набагато менша, ніж звукових сигнал загалом. Для обробки звуку на периферії системи можна використовувати класифікатор звуків. Результатом його є дані, що описують характеристики звуку. Відновити вихідний звук за цими метаданими неможливо. Детектор звуків дозволяє розпізнавати характерні структури, рівні гучності, частоти та видає інформацію про стан. Також існує можливість використання нечіткого

класифікатора на етапі класифікації, але зміна типу класифікатора вимагає модифікації попереднього етапу перетворення, оскільки атрибути для класифікації мають бути нечіткими [28].

У ситуаціях, коли потрібно використовувати вихідні аудіодані (без перетворення та аналізу), вони зазвичай піддаються певній обробці та кодуванню відповідно до моделі застосування. Модель застосування може передбачати збереження аудіоданих на периферії системи або передачу аудіопотоку на зовнішній клієнт (сервер або хмара) для додаткової обробки. В аудіоаналітиці обробка сигналу зазвичай проводиться для поліпшення передачі, підвищення ефективності зберігання або суб'єктивної якості, для виділення або виявлення його складових. Для обробки застосовуються програмні алгоритми, що змінюють та аналізують аудіосигнал різними способами.

Програмні алгоритми можна використовувати, наприклад, для поліпшення чутності сигналу шляхом автоматичного вирівнювання рівня гучності, зміни спектрального складу сигналу з допомогою еквайзера, обмеження рівня сигналу шляхом видалення певних частот або амплітуд. Також можна зменшити обсяг даних за допомогою стиснення або забезпечити конфіденційність шляхом шифрування, аналогічно до методів стиснення відеоінформації [29]. Окрім цього, існує низка алгоритмів, призначених виключно для обробки людської мови [30].

Коли аудіодані потрібно зберігати або передавати, зазвичай використовуються алгоритми стиснення даних. Метод стиснення без втрат передбачає архівування аудіозапису за допомогою алгоритмів, при яких дані не втрачаються, і можливо відновити аудіозапис із бітовою точністю. Прикладами кодеків, що використовують стиснення без втрат, є FLAC та ALAC. Метод стиснення з втратами передбачає, що немає сенсу зберігати повну інформацію про аудіозапис. Наприклад, людське вухо може не чути тихі звуки після гучних або занадто високі чи низькі частоти. Прикладом кодека зі стисненням з втратами є MPEG 1 Layer III або просто MP3.

Обрання оптимальної топології має значний вплив на загальну продуктивність та надійність. Залежно від сценарію використання, інформаційна система може набувати вигляду локального сервера, периферійної обробки або хмарного рішення. Кожен із цих варіантів передбачає застосування різних комп'ютерних систем (від вбудованих пристроїв до високопродуктивних серверів), які по-різному виконують задачі аналізу та зберігання аудіоданих. Розташування засобів аналітики в системі важливе з цілого ряду причин [31]. Зокрема, з точки зору конфіденційності та дотримання нормативних вимог щодо персональних даних, важливо, де програмне забезпечення аналізує аудіодані. Існують ситуації, коли аудіодані не можна пересилати по мережі і критично важливо, щоб аудіодані, що приймаються (але не зберігаються) аналізувалися локально. Якщо обробка вимагає значних обчислювальних ресурсів, які на периферії системи недоступні, може знадобитися передача цифрових аудіоданих у хмарне середовище або сервер.

Існуючі рішення для розташування засобів аудіоаналітики [6,31] наведені на рис. 1.2. Після того, як аудіоподія трапилася та виконався прийом звукового сигналу на периферійних пристроях, є три варіанти, де саме буде виконуватися аналітика, а саме виконання аналітики на периферії, на локальному сервері (ЛС), на хмарному сервері.

Аналітика на периферії передбачає виконання всіх аналітичних дій на периферійних пристроях. Якщо засоби аналітики працюють на периферії системи (мікрофон, камера, тощо), цифрові аудіодані не залишають цю периферію. У разі прийому звуку без запису, через мережу передається тільки результат аналітичної обробки – метадані або сигнали управління. Аналітика на периферії системи передбачає виконання всіх обчислень безпосередньо на пристроях, що приймають звук. Якщо комп'ютерна система пристрою достатньо потужна, можна застосовувати складні алгоритми фільтрації та класифікації аудіосигналів у реальному часі без передавання сирих даних до сервера.

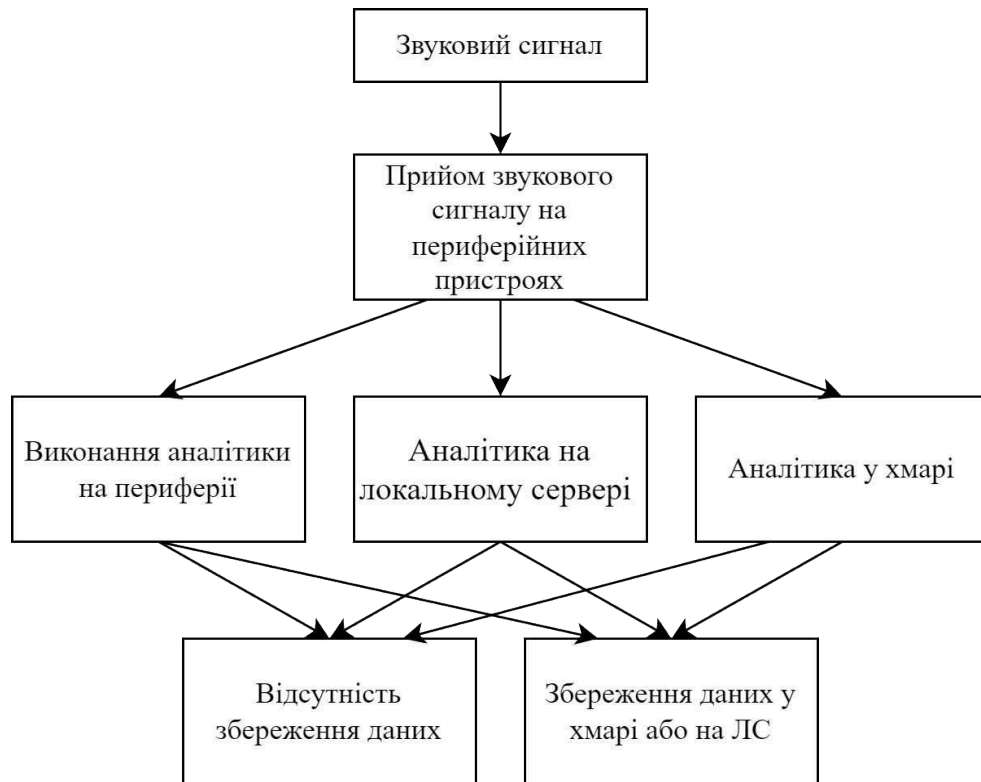


Рисунок 1.2 – Існуючі рішення для розташування засобів аудіоаналітики

Аналітика на локальному сервері передбачає використання локального серверу та локальної мережі, що з'єднує сервер та периферійні пристрої. Під час обробки на локальному сервері цифровий аудіосигнал виходить за межі периферії. На периферії системи може виконуватися попередня обробка сигналу, що абстрагує або деперсоналізує його до метаданих. У цьому випадку комп'ютерна система повністю контролюється власником, що полегшує дотримання норм безпеки та конфіденційності. Додатково необхідно забезпечити дотримання правил і нормативних вимог щодо зберігання та передачі персональних даних. При використанні локального серверу немає потреби покладатися на підключення до інтернету, а також обмежуватися пропускнуною спроможністю інтернет-каналу.

Хмарна аналітика передбачає використання хмарного середовища для проведення аналітики та зберігання даних. Крім того, для забезпечення безпеки всієї інфраструктури системи може використовуватися хмарний міжмережевий екран [32]. Цифрові аудіодані також можуть передаватися на сервер у хмарному середовищі. Як і у випадку аналітики на локальному

сервері, аудіоінформація може піддаватися попередньому перетворенню на метадані. Хмарні системи часто є децентралізованими, тому важливо враховувати навантаження на ресурси в системах хмарних обчислень [33], питання конфіденційності, а також здійснювати моніторинг відповідності регуляторним вимогам.

Після проведення аналітики виникає питання збереження даних. Збереження даних може бути організовано як на локальному сервері, та і у хмарному середовищі. Зазвичай, у системах аудіоаналітики зберігається інформація про аудіоподії у вигляді метаданих. Але існують випадки, коли зберігати дані про аудіоподії заборонено законами країни, або навпаки необхідно зберігати у початковому форматі. У певних ситуаціях запис та зберігання аудіоданих може бути корисним джерелом додаткових свідочств. Наприклад, на записі може бути зафіксовано розмови зловмисників або звуки стрілянини. Запис звуку може показати, хто що говорив, скільки пострілів було зроблено і дати іншу цікаву для розслідування інформацію. При записі звуку в контексті розслідувань слід подбати про те, щоб зберігати вихідні дані, не обробляючи їх (на відміну інших контекстів, коли це може бути необхідно або корисно).

Порівняльна характеристика існуючих рішень для розташування засобів аудіоаналітики з точки зору користувача наведена у табл. 1.1. Виконання аналітики на периферії вимагає встановлення додаткових обчислювальних модулів на кожен периферійний пристрій та зазвичай не використовується, так як не дозволяє отримати високий показник точності. Системи, в яких аналітика виконується не на периферії та вимагає передачі даних (як перетворених, так і неперетворених) через мережу, називаються мережними системами аудіоаналітики.

Головною відмінністю хмарного та локального серверу є необхідні витрати на обладнання, персонал, що обслуговує це обладнання та необхідне програмне забезпечення. У випадку локального серверу необхідно закупити обладнання, що дозволить проводити необхідну аналітику, у той час як при

використанні хмарного серверу за це відповідає постачальник послуги та надає клієнту необхідні обчислювані ресурси, в залежності від орендної плати.

Таблиця 1.1 – Порівняльна характеристика

Характеристика	Аналітика		
	На периферії	На ЛС	Хмарна
Витрати на обладнання	Високі	Високі	Відсутні
Необхідність у персоналі	Необхідні	Необхідні	Відсутня
Витрати на обслуговування ПЗ	Необхідні	Необхідні	Відсутні
Масштабованість	Вимагає витрат	Вимагає витрат	Не вимагає витрат
Резервне копіювання даних	Проблематичне	Присутнє	Присутнє
Доступність даних	На робочому місті	На робочому місті	Виключно за наявністю інтернету
Безпека даних	Вимагає витрат на систему безпеки	Вимагає витрат на систему безпеки	Безпека – це репутація постачальника послуг
Залежність від інтернету	Невелика	Невелика	Повна залежність
Швидкість роботи	У межах потужності обладнання	У межах потужності обладнання	Залежить від потреб
Законодавчі та нормативні вимоги	Легко вирішуються	Вимагає витрат	Вимагає витрат
Орендна плата	Відсутня	Відсутня	Присутня

Зменшення чи нарощування ресурсів можливе завдяки гнучкості хмарного сервера. З фізичним сервером це супроводжується певними складнощами. Локальний сервер налаштовують згідно з точними специфікаціями з точки зору швидкості процесора, дискового простору, пам'яті. Щоб внести зміни в ці параметри, потрібно зупинити роботу сервера та додати необхідне обладнання. Щоб знизити витрати, необхідно якомога точніше визначити кількість ресурсів, які знадобляться для безперебійної роботи.

Резервне копіювання даних у локальному серверу вимагає додаткових витрат на організацію цього процесу. Якщо в один день сервер вийде з ладу, то для відновлення всієї системи та перенесення файлів на інший сервер

потрібно багато часу. Цей час простою може виявитися критичним. Хмарні клієнти можуть не турбуватися про втрату інформації ні при поломці інфраструктури, ні при катастрофах. Багаторівнева система резервування забезпечує ефективне та надійне зберігання даних. Додатково, файли знаходяться не в одному місці, а автоматично розподіляються на декількох фізичних машинах.

До даних, що зберігаються у хмарі, можна отримати доступ незалежно від розташування, але виключно за наявності інтернету. У випадку локального серверу, правами та можливостями доступу до даних займається постачальник аудіоаналітичних послуг. При використанні локального серверу немає потреби покладатися на підключення до інтернету, а також обмежуватися пропускнуою спроможністю інтернет-каналу. Швидкість роботи під час використання такого сервера обмежується межами потужності фізичного обладнання. А у випадку із хмарним сервером вона змінюється в залежності від потреб.

У хмарі використовуються ті ж самі методи захисту даних, що й у локальному середовищі. Щоб хакери не мали доступу до даних, необхідно постійно здійснювати ряд дій: виявлення загроз, шифрування конфіденційних даних, моніторинг у режимі реального часу, багаторівнева ідентифікація. І якщо дані зберігаються на сервері хмар, то відповідальність за це розподіляється між користувачем і постачальником.

Основною проблемою для встановлення систем аудіоаналітики є нормативні та правові вимоги. Запис звуку може бути заборонено або може вимагати спеціального розгляду відповідно до національного законодавства або різноманітних місцевих законів і нормативів. У деяких регіонах та ситуаціях прийом звуку може бути дозволено, а запис заборонено. Наприклад, закони та нормативи в США різні в залежності від штата. У деяких штатах для ведення аудіозапису потрібна згода однієї сторони. В інших штатах потрібна згода всіх (або обох) сторін, тобто перед веденням аудіозапису необхідно отримати згоду всіх учасників. У країнах Європи

аудіоспостереження регулюється національними законами. Тому перш за все необхідно вивчити закони та нормативи, що діють у конкретній країні. Записи звуку можуть містити особисті дані, що підпадають під дію регламенту GDPR. Вирішити ці нормативні проблеми дозволяє використання лише прийому та тимчасової буферизації даних, без запису. Хоча аналітика може проводитись лише на периферійних пристроях, це не дозволяє досягти високої точності та потребує додаткових витрат на обладнання. Можливим рішенням цієї проблеми є отримання характеристик сигналу на периферійному пристрої, та передачі цих метаданих то аналітичного серверу (локального або у хмарі).

Основними аргументами при виборі локального сервера є висока швидкість передачі даних та постійна доступність. При використанні хмарного сервера немає потреби витрачатися на дороге серверне обладнання та його подальше обслуговування. Достатньо обрати один із хмарних серверів, який ідеально підходить вимогам, та платити орендну плату. Є кілька недоліків хмарних серверів, а саме те, що швидкість передачі та обробки даних залежить від швидкості інтернету та високий ризик атак зловмисників, так як відповідальність розподіляється між користувачем та постачальником послуг.

Програми для аудіоаналітики обробляють прийнятий звуковий сигнал, знаходячи та обробляючи певну інформацію. Вони застосовуються, наприклад, для виявлення таких подій, як стрілянина, биття скла або агресивна поведінка. Наприклад, така програма може за результатами обробки звукового сигналу відповісти на запитання: «Чи був постріл?». При виявленні події система зазвичай автоматично повідомляє персонал візуальним сигналом або увімкненням тривоги. Це забезпечує своєчасне повідомлення та оперативне реагування на події.

1.3 Основні вимоги до мережних систем аудіоаналітики

Гучні звуки та характерні шуми часто виникають перед або під час різноманітних протиправних дій, техногенних інцидентів чи природних катаклізмів. Уміння вчасно виявляти такі сигнали дозволяє значно зміцнити безпеку та забезпечити швидку реакцію на надзвичайні події.

Загалом процес розпізнавання аудіосигналів можна розглянути як низку послідовних кроків. Спочатку у вхідному потоці визначають момент появи аудіоподії, потім із початкового сигналу виділяють характерні ознаки, а на завершальному етапі відбувається класифікація цих ознак [34]. На рис. 1.3 представлено схему такого процесу.

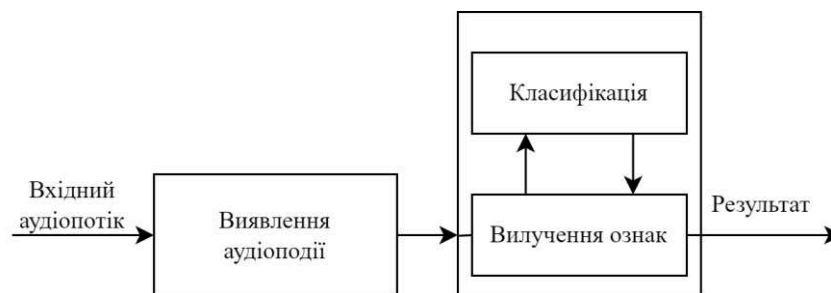


Рисунок 1.3 – Процес виявлення та розпізнавання звуків

Завдання виявлення може бути складнішим, ніж завдання класифікації [35]. Це головним чином пояснюється тим, що для виконання завдання виявлення потрібно не лише розрізнити категорії подій, але й виділяти цільові категорії подій серед багатого фону звуків. Крім того, під час класифікації доступний глобальний контекст подій, тоді як для завдання виявлення цей контекст спершу необхідно визначити, зазвичай використовуючи ненадійні локальні ознаки аудіосигналу.

У свою чергу, у складних системах виявлення аудіоподії може використовуватися верифікація виявлених аудіоподій. Система верифікації використовується для того, щоб зменшити ризик виникнення невідомої аудіоподії на етапі класифікації. Після виявлення аудіоподії система має

доступ до передбачуваних меж виявленої події, тобто має доступ до її глобального контексту, і може розрахувати її глобальні ознаки та виконати верифікацію. У результаті такого підходу суперечливість між даними навчання та тестування зменшується, що призводить до покращення роботи класифікатора. При використанні верифікації також зменшується кількість помилкових спрацьовувань, оскільки перевіряється відповідність між мітками виявлення та мітками класифікації. Процес виявлення аудіоподій наведено на рис. 1.4.

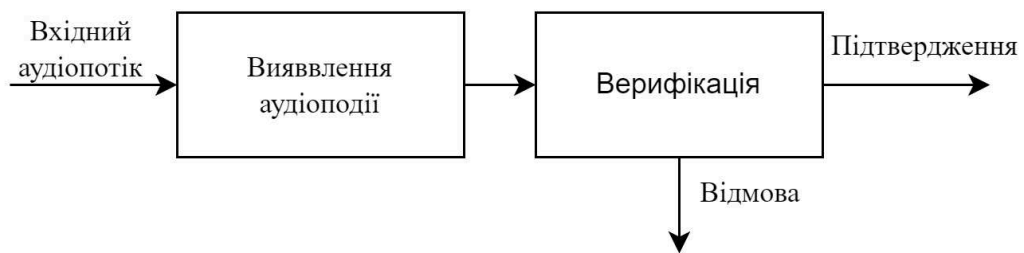


Рисунок 1.4 – Процес виявлення аудіоподій

Найчастіше системи автоматичного виявлення та класифікації аудіоподій розробляються для конкретних задач та оточень [36]. Існує ряд проблем при використанні системи для обробки декількох оточень та великого набору категорій подій. Дисперсія для кожного оточення та категорії ускладнює завдання автоматичної класифікації аудіоподій. Додатково це завдання ускладнює наявність аудіоподій, що перекриваються. Такі аудіоподії називаються поліфонічними. Приклад поліфонічних аудіоподій наведено на рис. 1.5.

Не зважаючи на те, що у реальному житті переважно трапляються поліфонічні аудіоподії, методи виявлення монофонічних аудіоподій мають право на існування, наприклад, у охоронних системах для виявлення пострілів.



Рисунок 1.5 – Поліфонічні аудіоподії

Однією з основних проблем методів виявлення та класифікації звукових подій є проблема навчальних даних [37]. Для забезпечення високої точності виявлення та класифікації необхідно використовувати набір даних, що складається з чітко маркованих звукових сцен, які містять часові мітки початку та завершення аудіоподії. Проте така суворе маркіровка достатньо великого набору даних є складним завданням, а анотації з подібними мітками, ймовірно, міститимуть людські помилки та непослідовності, особливо з огляду на неоднозначність у сприйнятті початку та кінця деяких звукових подій. Одним із рішень цієї проблеми може бути штучне створення звукових сцен та їх позначення з чітким маркуванням [38]. Але це може призвести до невідповідності між штучними даними для навчання та записаними реальними даними, які використовуються під час розв'язання задачі, що, у свою чергу, може спричинити збої в роботі класифікатора. Іншим рішенням цієї проблеми може бути використання слабо маркованих даних, які не містять часових міток, а лише інформацію про наявність події в записі [39]. Проте це рішення має низку недоліків, пов'язаних із різною тривалістю аудіоподій та наявністю фонових шумів.

Більшість методів виявлення імпульсних аудіоподій засновані на визначенні енергії для набору послідовних блоків аудіо сигналу, що не перекриваються. Енергія e для k -го блоку сигналу, що складається з N відліків, визначається як:

$$e(k) = \frac{1}{N} \sum_{i=0}^{N-1} x^2(i + kN), k = 0, 1, \dots$$

У якості прикладу, на рис. 1.6 наведено представлення звукового сигналу з гучними імпульсами, що сталися після 0,01 с та 0,41 с від початку відліку відповідно та графік залежності енергії від часу.

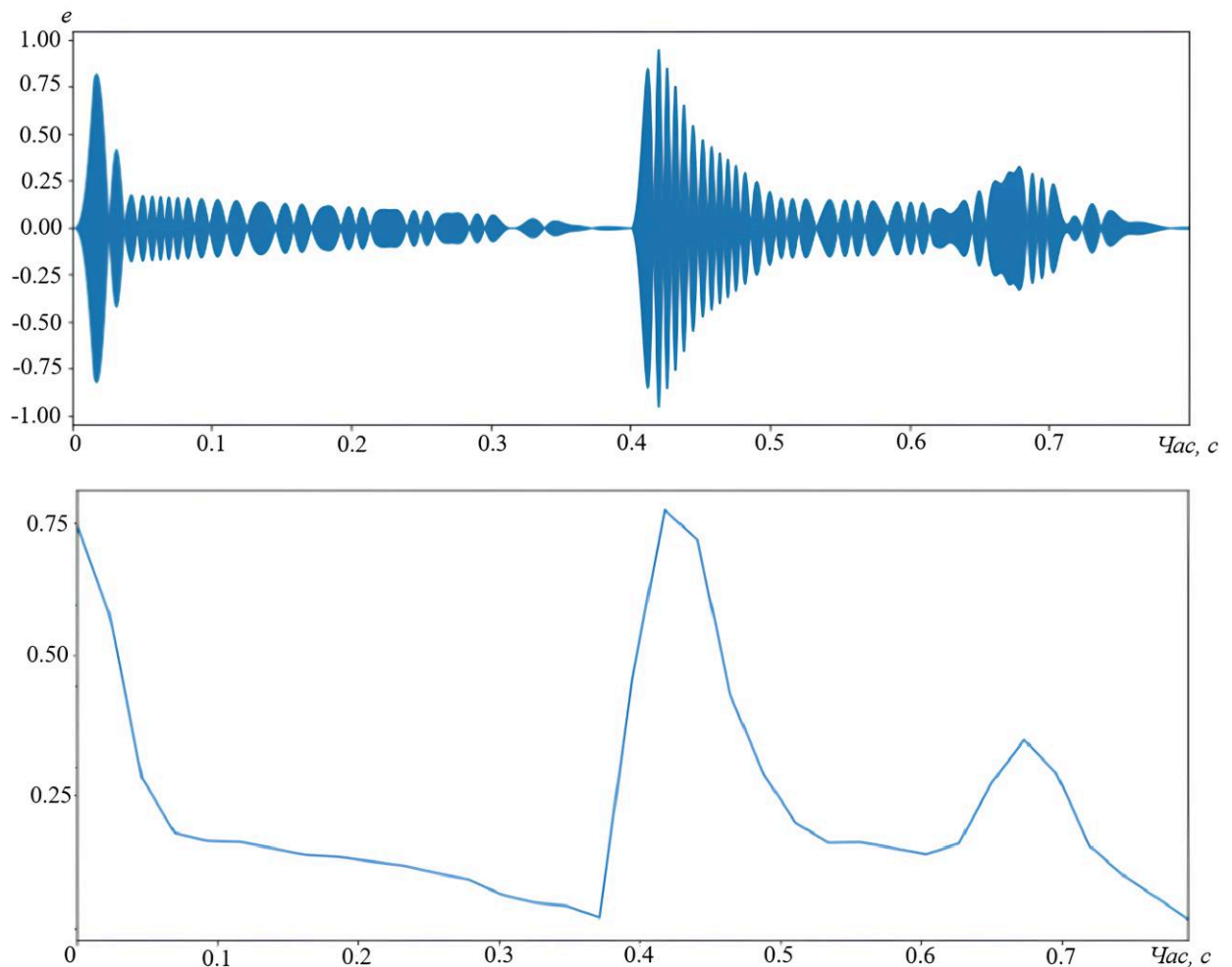


Рисунок 1.6 – Представлення звукового сигналу та графік залежності енергії від часу

Методики для автоматичного визначення блоку, що відповідає різкому імпульсному звуку, різняться залежно від використаних підходів. Серед найпоширеніших метод на основі стандартного відхилення нормованих енергетичних значень блоків, метод із застосуванням медіанного фільтра та

метод із динамічним порогом для аналізу енергії блоків. В основі підходу, що використовує стандартне відхилення нормованих значень, лежить процедура нормування початкових енергій блоків у діапазон від нуля до одиниці:

$$e_{norm}(j) = \frac{e_{frame}(j) - \min_j(e_{frame}(j))}{\max_j(e_{frame}(j) - \min_j(e_{frame}(j)))},$$

де e_{frame} – значення енергії для фрейму.

Далі обчислюється стандартне відхилення отриманого набору значень:

$$\text{var}(k) = \frac{1}{L-1} \sum_{j=0}^{L-2} [e_{norm}(j, k) - \bar{e}_{norm}(k)]^2.$$

Під час надходження нового значення енергії здійснюється його перенормування до цього інтервалу.

Ще одним підходом до виявлення аудіоподій є метод, заснований на використанні медіанного фільтра. Для виявлення блока з імпульсною подією застосовується умовний варіант медіанного фільтра:

$$mf(k) = MED_{i-l-L+1}^k e(i);$$

$$cmf(k) = \begin{cases} mf(k), & \text{якщо } (mf(k) - e(k-d)) > th; \\ e(k-d), & \text{якщо } (mf(k) - e(k-d)) \leq th, \end{cases}$$

де $e(i)$ – послідовність дискретизованих значень вхідного аудіосигналу;

MED – медіанна операція, що обчислює медіанне значення;

$mf(k)$ – вихід медіанного фільтра для моменту часу k ;

d – зсув індексу відносно поточного моменту k ;

th – порогове значення.

Завдяки обчисленню різниці між сигналом після умовної медіанної фільтрації та початковим сигналом (із часовим зміщенням) можна автоматично визначити блок із імпульсною подією.

У підході з динамічним порогом пропонується використовувати середню потужність блоків і середньоквадратичне відхилення як критерій для виявлення імпульсного сигналу. Автоматичне спрацювання відбувається тоді, коли поточна потужність блоку перевищує порог, визначений формулою:

$$th = par * std + m,$$

де par – параметр чутливості алгоритму;

std – середньоквадратичне відхилення потужності блоків;

m – коефіцієнт зміщення.

На рис. 1.7, 1.8 та 1.9 наведено графіки залежності енергії від часу для кожного з описаних методів. Імпульсні події трапилися після 0,01 с та 0,41 с від початку відліку відповідно.

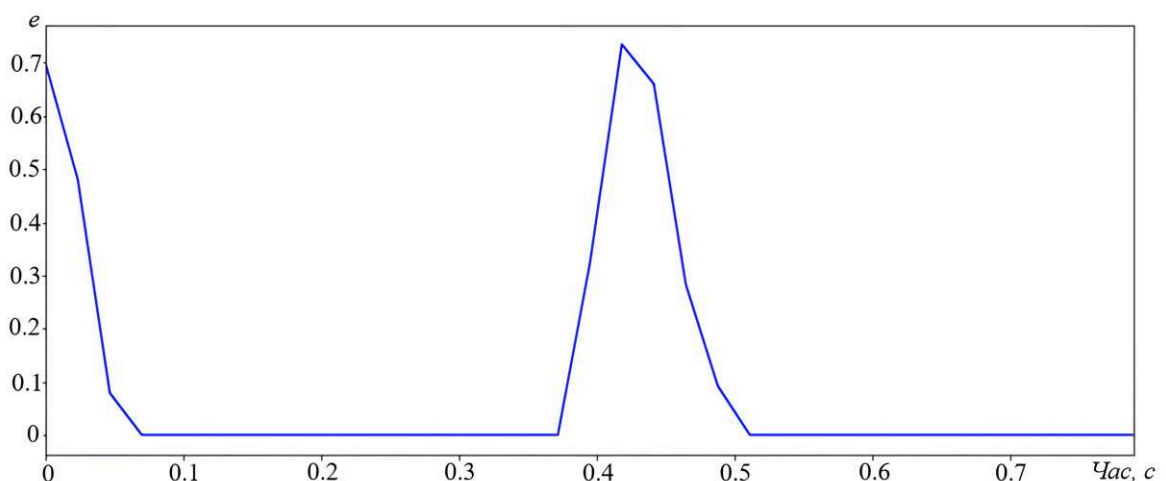


Рисунок 1.7 – Графік залежності енергії від часу при використанні методу на основі стандартного відхилення нормованих значень

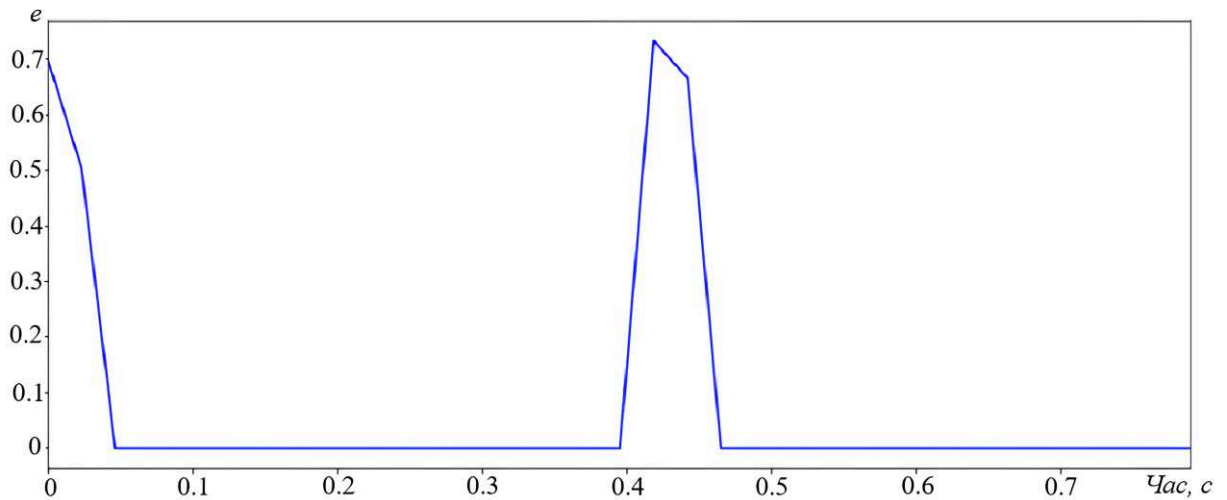


Рисунок 1.8 – Графік залежності енергії від часу з використанням методу на основі медіанного фільтру

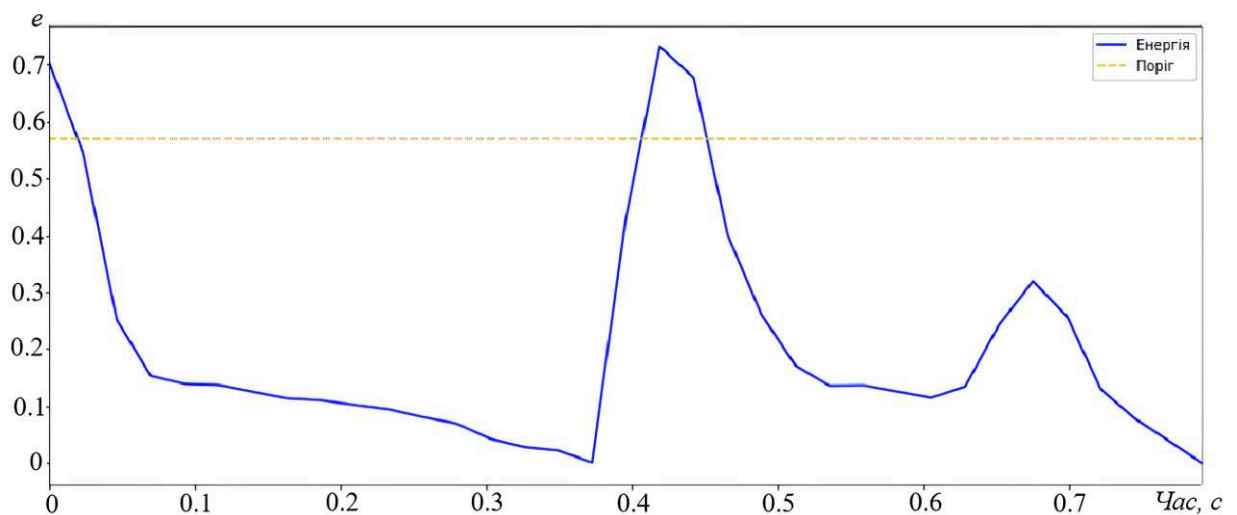


Рисунок 1.9 – Графік залежності енергії від часу з використанням методу на основі динамічного порогу

Точність виявлення аудіоподій при зростанні рівня шуму наведена у табл. 1.2. Як набір даних використовується набір із 476 імпульсних аудіоподій, кожна з яких має відповідне маркування з часовими мітками. Після виявлення аудіоподії розраховуються додаткові ознаки, і виявлена аудіоподія класифікується за одним із відомих методів класифікації [40, 41].

Таблиця 1.2 – Точність виявлення аудіоподій при зростанні рівня шуму

Відношення сигнал/шум, dB	Стандартне відхилення, %	Медіанний фільтр, %	Динамічний поріг, %
5	100	100	100
0	99,79	99,37	99,58
-5	97,69	97,06	97,48
-10	81,52	80,05	81,1
-15	41,81	40,34	40,55
-20	0,42	0,21	0,63

Перевагою методу на основі стандартного відхилення є його стійкість до шумів та можливість виявлення повільно мінливого сигналу, аналізуючи середнє значення нормованих блоків потужності. До переваг використання медіанного фільтра відноситься лише його проста структура, що дозволяє без зусиль зробити апаратну та програмну реалізації з порівняно малою обчислювальною складністю. Він має ряд недоліків, пов'язаних з тим, що фільтр не працює в умовах одиничних імпульсних шумів та при збільшенні розмірів вікна фільтра відбувається розмиття крутих змін сигналу.

Метод на основі динамічного порогу має переваги обох минулих методів, та позбавлений декількох недоліків. Але, незважаючи на це, використовувати його досить складно. Пов'язане це з вимогою завдання параметру чутливості алгоритму, що ускладнює апаратну реалізацію. Необхідність розрахунку параметру чутливості алгоритму з оглядом на місцезнаходження та направлення обладнання робить цей метод складним до реалізації, але ефективним при правильному застосуванні.

У сучасних системах аудіоаналітики виявлення та класифікація тісно пов'язані між собою. Такі системи представлені на змаганнях і семінарах Detection and Classification of Acoustic Scenes and Events (DCASE), які проводяться щороку. Наприклад, у змаганнях DCASE2021 є завдання 4 під назвою «Виявлення звукових подій та їх розділення в домашніх умовах».

Мета цього завдання – оцінити системи виявлення звукових подій за допомогою реальних даних з нечіткими мітками та строго маркованих синтетичних даних. Завдання оцінює системи виявлення звукових подій, навчені на слабо маркованих даних, які не містять часових міток. Системи мають визначати не лише клас події, а й її локалізацію в часі, оскільки в аудіозаписі може бути кілька подій. Учасникам надаються ізольовані звукові події, фони звуків та скрипти для розробки навчального набору з строго маркованими синтетичними даними.

Одне з базових рішень [37] використовує модель «суворий вчитель» на основі нейронних мереж. Ця модель є комбінацією двох моделей з однаковою архітектурою: студентської та вчительської. Вчительська модель допомагає студентській під час навчання, тоді як студентська є фінальною моделлю. Під час навчання обидві моделі отримують однакові вхідні дані, але до вхідних даних вчительської моделі додається гаусівський шум. Це дозволяє студентській моделі навчатися на основі функції втрат, яка забезпечує узгодженість як для сильних, так і для слабких прогнозів для всіх кліпів у пакеті.

Також є спроби покращити виявлення звукових подій за допомогою розділення звуків [42], що використовується як попередня обробка для системи виявлення звукових подій. Завдання розділення звуків полягає у відновленні або реконструкції одного або кількох оригінальних сигналів, змішаних з іншими внаслідок лінійного або згорткового процесу. Це дослідження має багато практичних застосувань, включаючи покращення якості звуку, усунення шумів, реміксинг музики, створення просторового звучання, ремастеринг тощо. Проте останні дослідження показали, що універсальне розділення звуків може використовуватися для розділення звуків довільних класів [43,44]. Поєднання розділення звуків із виявленням звукових подій демонструє потенціал для покращення роботи систем виявлення звукових подій, хоча його переваги досі обмежені через

невідповідність умов навчання для розділення звуків і умов тестування для виявлення звукових подій [42].

Для завдання DCASE 2021 Task 4 існує багато систем виявлення звукових подій, заснованих на базовій системі [45-50].

У 2021 році система, що отримала найвищий рейтинг, запропонувала три основні зміни щодо базового рішення [45]: використання селективної одиниці, використання «м'якого виходу» виявлення шляхом налаштування параметра температури в сигмоїдній функції та використання кількох методів аугментації даних. Ці зміни дозволили мережі адаптивно налаштовуватися як для коротких, так і для довгих подій, що в цілому покращило стабільність і надійність роботи системи.

Система, що посіла друге місце [46], пропонує модель виявлення звукових подій, засновану на самонавчанні з використанням моделі «noisy student». Використовувалася модель RCRNN mean-teacher для прогнозування цільової мітки кожного аудіокліпу, застосовуючи аугментацію даних із шумом на ознаках, шум через випадкові відключення та напівконтрольовану функцію втрат для міток.

Система, що посіла третє місце [47], використовує як рекурентну структуру, так і трансформерну структуру для моделювання складної динаміки реальних звуків у домашньому середовищі. Це було зроблено для підвищення загальної продуктивності порівняно з базовим рішенням, оскільки різні моделі демонструють різну ефективність за різних умов. Крім того, використовуються напівконтрольоване навчання «mean-teacher» і різні методи аугментації даних.

Аналіз розглянутих методів свідчить, що для мережних систем аудіоаналітики визначальним є дотримання низки важливих вимог, аби вони залишалися максимально ефективними та гнучкими під час використання в різних галузях. Насамперед ідеться про високу точність розпізнавання, адже надійність і коректність результатів має першочергове значення. Окрім цього, система повинна адаптуватися до мінливих умов і бути

масштабованою для обробки великих масивів даних, що особливо актуально в мережних сценаріях. Як і будь-яка інформаційна система, мережна система аудіоаналітики повинна відповідати вимогам точності, продуктивності й масштабованості. У той же час різні комп'ютерні системи, що беруть участь у її роботі (локальні сервери, хмарні платформи, периферійні пристрої), мають дотримуватися принципів безпеки та конфіденційності, особливо якщо йдеться про обробку персональних даних.

1.4 Постановка завдання дослідження

На основі аналізу сучасного стану розвитку методів обробки аудіосигналів у мережних системах аудіоаналітики можна зробити висновок щодо актуальності дослідження такої обробки та необхідності подальшого дослідження з метою підвищення її ефективності. У зв'язку з цим метою дисертаційної роботи є підвищення точності розпізнавання аудіосигналу за рахунок удосконалення методів передачі та аналізу аудіосигналів у мережних системах аудіоаналітики. Для досягнення поставленої мети необхідно вирішити такі завдання:

- провести аналіз напрямків розвитку методів обробки аудіосигналів та мережних систем аудіоаналітики;
- розробити модель процесу проходження аудіосигналів у мережних системах аудіоаналітики;
- розробити метод вибору мережної структури системи аудіоаналітики;
- удосконалити метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів.
- удосконалити метод аналізу аудіосигналів на основі багатоподійного класифікатора;
- удосконалити інформаційну технологію обробки аудіосигналів з використанням удосконалених методів і моделей.

Вирішенню поставлених завдань, що виникають в процесі розробки моделі та методів обробки аудіосигналів, присвячено наступні розділи.

Висновки за розділом 1

У даному розділі було проведено аналіз основних аспектів вибору методів обробки аудіосигналів у мережних системах аудіоаналітики:

1. Було розглянуто основні напрямки розвитку методів обробки аудіосигналів у контексті їх використання в мережних системах аудіоаналітики.

2. Був проведений огляд існуючих мережних систем аудіоаналітики з урахуванням їх архітектури та функціональності. Цей аналіз дозволив зрозуміти, що розвиток методів обробки аудіосигналів пов'язаний з постійним удосконаленням архітектури та функціональності мережних систем;

3. Окремо були визначені основні вимоги до мережних систем аудіоаналітики з точки зору обробки аудіосигналів, такі як продуктивність, точність, швидкість обробки тощо. Враховуючи ці вимоги, можна зробити висновок про необхідність розробки та застосування відповідних методів обробки аудіоданих у мережних системах аудіоаналітики;

4. На основі проведеного аналізу було сформульовано постановку завдання дослідження методів обробки аудіосигналів у мережних системах аудіоаналітики. Це завдання буде вирішуватися в подальших розділах даної роботи з метою розробки ефективних методів обробки аудіоданих у мережних системах аудіоаналітики.

Перелік використаних джерел у даному розділі наведено у повному переліку використаних джерел під номерами: 18–51.

2 МОДЕЛЮВАННЯ ПРОЦЕСУ ПРОХОДЖЕННЯ АУДІОСИГНАЛІВ У МЕРЕЖНИХ СИСТЕМАХ АУДІОАНАЛІТИКИ

У другому розділі розглянуто моделювання процесу проходження аудіосигналів у мережних системах аудіоаналітики. Основну увагу приділено аналізу топології мережі, особливостям передачі аудіосигналів, та розробці математичної моделі. Запропоновано ієрархічну структуру мережі, що дозволяє підвищити надійність системи, враховуючи різні маршрути передачі даних між вузлами. У процесі моделювання побудовано оптимізаційні рівняння для базових структур мереж, які дозволяють мінімізувати капітальні витрати на побудову. Окрему увагу приділено впливу неординарного пуасонівського навантаження на продуктивність системи, що враховує особливості аудіосигналів та специфіку обробки в реальному часі. Запропонована модель створює основу для вибору оптимальної мережної структури та є важливим внеском у підвищення ефективності систем аудіоаналітики.

2.1 Визначення особливостей передачі аудіосигналів у мережних системах аудіоаналітики

При проектуванні та впровадженні мережних систем аудіоаналітики одним із ключових етапів є організація ефективної передачі аудіосигналів від джерел (мікрофонів, сенсорів чи пристроїв збору даних) до вузлів аналізу або сховищ. Висока якість та достовірність звукової інформації має вирішальне значення для правильного функціонування системи, адже від цього залежить точність виявлення та розпізнавання подій, мовних сигналів, акустичних образів тощо. Нижче наведено основні особливості та вимоги, які визначають специфіку передачі аудіосигналів у мережних системах аудіоаналітики:

- метод комутації [51-55];

- смуга пропускання та обсяг даних [56];
- затримка і пакетна передача [57, 58];
- якість обслуговування (QoS) [59] та пріоритетність трафіку;
- надійність передачі та корекція помилок [60-62];
- стиснення [63] та кодування [64] аудіосигналів;
- безпека передачі даних [65];
- масштабованість та гнучкість системи [66];
- синхронізація часових міток;
- оптимізація витрат ресурсів [67-69].

У сучасних цифрових мережах майже використовується комутація пакетів, коли аудіодані розбиваються на невеликі пакети й кожен пакет передається через мережу самостійно. Це дає гнучкість у використанні спільних ресурсів, проте може призводити до нерівномірної затримки та втрат пакетів. Традиційні мережі базувалися на комутації каналів, коли між двома вузлами виділявся постійний канал зв'язку. Такий підхід гарантує фіксовану смугу пропускання, але використовує ресурси неефективно. Для аудіозастосунків реального часу (онлайн-конференції, охоронні системи) зазвичай використовується комутація пакетів з додатковими протоколами реального часу та механізмами забезпечення якості обслуговування (QoS), які допомагають мінімізувати затримки та втрати. Комутація каналів є стандартом, але поступається ефективністю у випадку, коли одночасно передаються численні незалежні потоки, або коли пропускна здатність має бути гнучко розподілена між багатьма користувачами. Комутація повідомлень та осередків має більш вузьке застосування і зазвичай використовується у спеціалізованих мережах або в ретроспективних системах. Аудіосигнали, особливо у високій роздільній здатності, потребують широкої смуги пропускання. Зі збільшенням кількості каналів (наприклад, у багатоканальних мікрофонних масивах) різко зростає загальний бітрейт. Вибір формату передачі впливає на обсяг передаваних даних та навантаження на мережу. При цьому слід збалансувати потреби в

якості звуку та пропускну здатність. Застосування методів кодування дає змогу зменшити обсяг даних. Обираючи кодек, необхідно враховувати компроміс між затримкою кодування, споживанням ресурсів та якістю відтворення. У системах аудіоаналітики часто використовують низькі та середні бітрейти для передачі, достатні для розпізнавання подій або мови, але недостатні для студійної якості звуку.

У реальних мережах передача аудіо здійснюється пакетами. Затримка при передачі може бути критичною для систем реального часу, зокрема при виявленні тривожних звуків або дистанційному контролі. Для зменшення затримок застосовують протоколи передачі у режимі реального часу, а також алгоритми адаптації швидкості передачі залежно від поточного стану мережі. В умовах завантажених мереж важливо реалізувати механізми QoS, аби гарантувати мінімальну затримку для аудіопакетів та знизити ймовірність втрат. Маршрутизатори та комутатори, що підтримують пріоритезацію трафіку, дозволяють виділяти чіткий канал або клас обслуговування для аудіоданих. Випадкові втрати пакетів, зокрема в бездротових сегментах, можуть призводити до спотворення аудіосигналів або їх переривання. Використання спеціальних алгоритмів корекції помилок або повторна передача втрачених пакетів допомагають покращити якість, однак підвищують затримку та мережне навантаження.

Важливо забезпечити захист переданої інформації від несанкціонованого доступу та прослуховування, для цього використовують протоколи шифрування. Автентифікація джерела сигналу дає змогу переконатися у достовірності обладнання та запобігти підміні аудіоданих. При збільшенні кількості вузлів чи вимог до системи необхідно забезпечити розширення мережної інфраструктури без втрати продуктивності. Використання гнучких технологій маршрутизації, хмарних серверів обробки та технологій балансування навантаження дає змогу системі динамічно підлаштовуватися під зміни.

2.2 Формалізований опис мережної системи аудіоаналітики

Мережна система аудіоаналітики призначена для збирання, передачі, обробки та аналізу аудіоданих у реальному часі або в режимі відкладеного доступу. Формалізований опис такої системи можна представити за допомогою структурної моделі, яка включає наступні основні компоненти:

1. Джерела аудіоданих S : множина джерел $S = \{S_1, S_2, \dots, S_n\}$, де S_i – окремий аудіосенсор; функція збирання даних $f(S): T \rightarrow D$, де T – часовий інтервал, D – набір зібраних аудіоданих;

2. Мережа передачі даних N : топологія мережі, граф $G = (V, E)$, де V – вузли (пристрої), E – канали зв'язку; протоколи передачі; P_N – набір правил, які визначають спосіб передачі даних по мережі;

3. Сервер аналізу даних P : обчислювальні ресурси: $R_p = \{CPU, RAM, Storage\}$; сховище даних: база даних DB , де зберігаються отримані та оброблені дані, а також результати аналізу R ;

4. Модуль аудіоаналітики A : алгоритми аналізу: набір функцій $F_A = \{f_1, f_2, \dots, f_n\}$, де кожна функція f_i відповідає за певний аспект аналізу; моделі машинного або глибинного навчання: $M = \{m_1, m_2, \dots, m_n\}$, які використовуються для покращення точності аналізу;

5. Інтерфейс користувача UI : компоненти інтерфейсу $UI = \{ui_1, ui_2, \dots, ui_n\}$, що забезпечують взаємодію користувача з системою; API: набір кінцевих точок $API = \{api_1, api_2, \dots, api_n\}$ для інтеграції з іншими системами.

У загальному вигляді, формалізований опис мережної системи аудіоаналітики наведено на рис. 2.1.

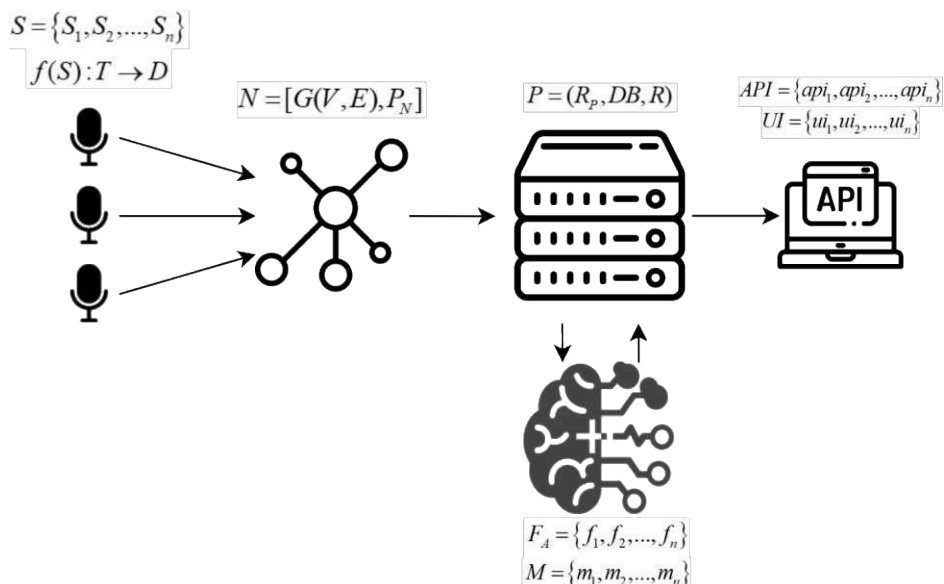


Рисунок 2.1 – Формалізований опис мережної системи аудіоаналітики

Процес функціонування мережної системи аудіоаналітики є складним та багатоступеневим, включає в себе кілька ключових етапів, які взаємодіють між собою для досягнення кінцевої мети, а саме отримання цінної інформації з аудіоданих. Починається все з джерел аудіоданих, представлених множиною сенсорів $S = \{S_1, S_2, \dots, S_n\}$. Кожен сенсор збирає аудіосигнали у певному часовому інтервалі T , реалізуючи функцію збирання даних $f(S): T \rightarrow D$, де D – набір зібраних аудіоданих.

Після збору даних, вони потребують передачі до центрального сервера для подальшої обробки. Цей етап здійснюється через мережу передачі даних N , яка має топологію, описану графом $G = (V, E)$, де V – вузли (пристрої), E – канали зв'язку; Передача даних регулюється протоколами P_N , які визначають правила та методи транспортування аудіоданих, враховуючи можливі затримки, втрати та забезпечуючи безпеку передачі.

Коли дані досягають сервера обробки P , вони зберігаються у сховищі даних DB для подальшої обробки. Сервер має обчислювальні ресурси $R_p = \{CPU, RAM, Storage\}$, які дозволяють виконувати складні алгоритми аналізу. Модуль аудіоаналітики A застосовує набір функцій

$F_A = \{f_1, f_2, \dots, f_n\}$, кожна з яких відповідає за конкретний аспект обробки аудіоданих. Це можуть бути алгоритми фільтрації шуму, розпізнавання мовлення, виявлення аномальних звуків або інших подій, що представляють інтерес.

Крім того, модуль аудіоаналітики може використовувати моделі машинного навчання $M = \{m_1, m_2, \dots, m_n\}$ для покращення точності та ефективності аналізу. Ці моделі можуть бути попередньо навчені на великому обсязі даних і здатні адаптуватися до нових патернів у аудіосигналах.

Результати аналізу R отримані після застосування функцій обробки, зберігаються у базі даних DB для подальшого використання. На фінальному етапі ці результати представлені користувачу через інтерфейс I . Інтерфейс користувача забезпечує зручний та інформативний доступ до даних, дозволяючи переглядати, аналізувати та взаємодіяти з інформацією. Це може бути веб-інтерфейс, мобільний додаток або інші засоби візуалізації даних.

Формально, весь процес можна описати як композицію функцій $I(A(P(N(S))))$, де кожен компонент системи є функцією від виходу попереднього. Це означає, що дані, зібрані сенсорами, проходять через мережу, обробляються на сервері, аналізуються модулем аудіоаналітики і, нарешті, представлені користувачу.

Важливою частиною процесу є забезпечення надійності та безпеки на кожному етапі. Це включає в себе захист даних при передачі по мережі через використання шифрування та інших методів безпеки, а також забезпечення резервування та відмовостійкості системи для безперервної роботи.

Крім того, система повинна бути масштабованою, щоб легко додавати нові сенсори або обробляти більші обсяги даних без втрати продуктивності. Це досягається через оптимізацію алгоритмів, використання ефективних протоколів передачі даних та розподіл обчислювальних ресурсів.

2.3 Розроблення математичної моделі процесу проходження аудіосигналів у мережних системах аудіоаналітики

Розподілена мережна система аудіоаналітики має ієрархічну структуру, що наведена на рис. 2.2. Нижній рівень складають пристрої прийому аудіосигналів (ППАС), котрі з'єднуються з пристроями наступного ієрархічного рівня – вузлами комутації (ВК), за допомогою мережі з'єднувальних ліній (МЗЛ), яка, в свою чергу, передає прийняті потоки аудіоінформації до центру обробки аудіосигналів (ЦОАС).

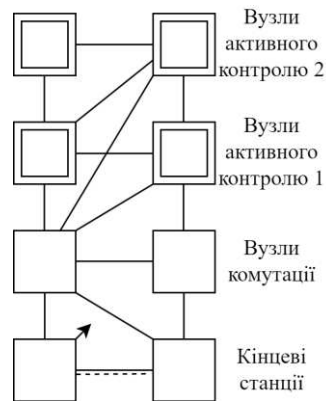


Рисунок 2.2 – Структура ієрархічної мережі

З'єднання в МЗЛ здійснюються в режимі багатоканальної комутації. При цьому є кілька можливих шляхів для обслуговування навантаження між кожною парою станцій. Схема вибору шляхів для обслуговування навантаження передбачається фіксованою і має вигляд, представлений на рис. 2.3, де із вузла 1 до вузла 4 є такі шляхи: 1-4 – прямий шлях; 1-3-4 – обхідний шлях; 1-2-3-4 – резервний додатковий шлях. Таким чином, при досить великому навантаженні між вузлами утворюються прямі шляхи високого використання, на яких обслуговується значна частина поступаючого навантаження. Надлишкове навантаження обслуговується на обхідних проміжних шляхах та на резервних додаткових шляхах. Якість обслуговування абонентів у мережі визначається нормованою величиною втрат на шляху останнього вибору.

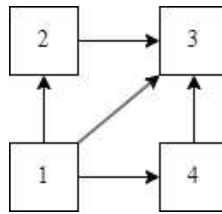


Рисунок 2.3 – Можливі шляхи із вузла 1 до вузла 4

Задано: топологію МЗЛ (розташування ППАС, ВК і з'єднувальних ліній (ЗЛ)); алгоритм вибору шляхів під час обслуговування навантаження між кожною парою вузлів; величини навантажень, які необхідно обслужити, а також тип і характер початкових заявок на обслуговування; структуру побудови ЗЛ і вартість одного каналу в ній; нормовану величину втрат на шляхах останнього вибору. Потрібно: визначити ємності пучків каналів ЗЛ – $N_j, j = \overline{1, L}$, де L – число ЗЛ у мережі, за яких забезпечується мінімум капітальних витрат C на побудову мережної системи аудіоаналітики та за забезпечення необхідної якості обслуговування.

Розглянемо найважливіші базові мережні структури, що використовуються під час розв'язання задачі оптимізації мережної системи аудіоаналітики, основними особливостями якої є:

- використання цифрових систем передачі при створенні ЗЛ;
- модульна зміна ємності пучків каналів ЗЛ;
- наявність багатоканальних систем комутації в ВК;
- наявність вхідних навантажень, які задаються парою (R, Z) , де R – інтенсивність навантаження, Z – коефіцієнт дисперсії.

У мережі з обхідними напрямками типу, що представлена на рис. 2.1, відповідно до схеми вибору обхідних шляхів, що представлена на рис. 2.2, можна виділити дві характерні ситуації, до яких зводяться всі можливі випадки вибору обхідних шляхів:

- навантаження, яке необхідно обслужити між вихідною станцією і станцією призначення, надходить спочатку на пряму колію (ЗЛ, що безпосередньо з'єднує станції). Якщо на цьому шляху навантаження, що

надходить, не може бути обслуговане з необхідною якістю обслуговування, то надлишкове навантаження з прямого шляху надходить на дообслуговування на обхідний шлях/

- для навантаження, що не було обслуговано, обирається наступний обхідний шлях із більшою кількістю транзитних ВК.

Розглянемо дві базові мережні структури: 3-вузлову мережу (рис. 2.4) та 4-вузлову мережу (рис. 2.5).

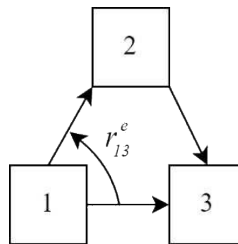


Рисунок 2.4 – Організація обходів на однонаправленій 3-вузловій мережі

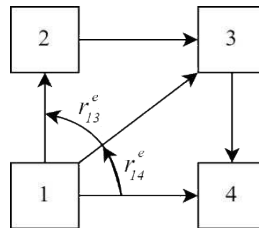


Рисунок 2.5 – Організація обходів на однонаправленій 4-вузловій мережі

Виділемо базові поняття:

- маргінальне використання H , що визначає збільшення обслугованого навантаження A_0 при одиничному збільшенні ємності пучка каналів N і постійному значенні вхідного навантаження A (у цьому випадку: $r = f(A, N)$ – надлишкове навантаження, $A_0 = A - r$ – обслуговане навантаження):

$$H = \left(\frac{\partial A_0}{\partial N} \right)_{A=const} = - \left(\frac{\partial r}{\partial N} \right)_{A=const} ;$$

- маргінальна ємність β , що визначає приріст вхідного навантаження при одиничному збільшенні ємності пучка в умовах постійних втрат E на цьому пучку:

$$\beta = \left(\frac{\partial A}{\partial N} \right)_{E=\text{const}} ;$$

- маргінальне надлишкове навантаження γ , що визначає інтенсивність зміни надлишкового навантаження у разі зміни вхідного навантаження і постійної ємності пучка:

$$\gamma = \left(\frac{\partial r}{\partial A} \right)_{N=\text{const}} ;$$

- маргінальне вхідне навантаження φ , яке визначає приріст вхідного навантаження при одиничному збільшенні ємності пучка та постійному значенні надлишкового навантаження:

$$\varphi = \left(\frac{\partial A}{\partial N} \right)_{r=\text{const}} .$$

Під час визначення перелічених величин припускають, що ємності пучків каналів на шляху можуть бути нецілочисловими.

Для найпростішої базової мережної структури оптимізована функція вартості C :

$$C = C_1 N_1 + C_2 N_2 + C_3 N_3. \quad (2.1)$$

У цьому випадку вартість C є функцією тільки однієї змінної N_1 , тому що під час завдання N_1 можна визначити параметри навантаження, що надходить на шлях останнього вибору, а потім визначити ємності N_2 й N_3 відповідно до заданої величини втрат на цьому шляху – $P_{\text{дод}}$. Таким чином, для визначення оптимального значення функції (2.1) необхідно розв'язати таке рівняння (за заданих навантаження A_1 , що надходить на пучок каналів N_1 шляху першого вибору; втрат на пучку каналів $N_2 - E_2 = P_{\text{дод}}$ і пучку каналів $N_3 - E_3 = P_{\text{дод}}$ шляху останнього вибору):

$$\frac{\partial C}{\partial N_1} = 0,$$

$$C_1 + C_2 \left(\frac{\partial N_2}{\partial N_1} \right) + C_3 \left(\frac{\partial N_3}{\partial N_1} \right) = 0. \quad (2.2)$$

У результаті низки перетворень рівняння (2.2) набуває вигляду:

$$\frac{C_1}{H_1} = \frac{C_2}{\beta_2} + \frac{C_3}{\beta_3}. \quad (2.3)$$

Рівняння (2.3) є оптимізаційним рівнянням для 3-вузлової базової структури, що наведена на рис. 2.6.

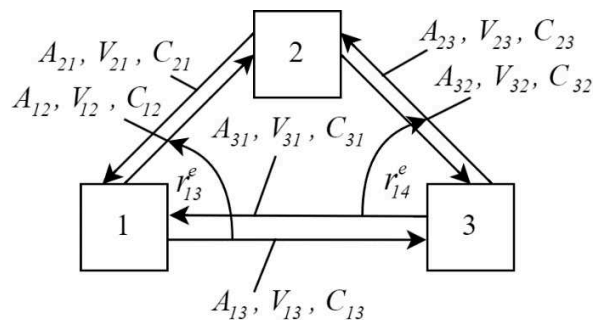


Рисунок 2.6 – Організація обходів на двонаправленій 3-вузловій мережі

Розглянемо цю базову структуру за умови, що навантаження обслуговують не тільки між вузлами 1 і 3, але також між вузлами 3 і 1. Функція вартості для цієї базової структури:

$$C = C_{12}(V_{12} + V_{21}) + C_{23}(V_{23} + V_{32}) + C_{13}(V_{13} + V_{31}),$$

де C_{ij} – вартість організації каналу передачі даних між вузлами i та j ;

V_{ij} – інтенсивність навантаження, що надходить від вузла i до вузла j .

Оптимізаційні рівняння для цієї базової структури стосовно ємностей пучків V_{13} і V_{31} мають вигляд:

$$\begin{cases} \frac{\partial C}{\partial V_{13}} = \frac{\partial [C_{12}(V_{12} + V_{21})]}{\partial V_{13}} + \frac{\partial [C_{23}(V_{23} + V_{32})]}{\partial V_{13}} + \frac{\partial [C_{13}(V_{13} + V_{31})]}{\partial V_{13}} = 0; \\ \frac{\partial C}{\partial V_{31}} = \frac{\partial [C_{12}(V_{12} + V_{21})]}{\partial V_{31}} + \frac{\partial [C_{23}(V_{23} + V_{32})]}{\partial V_{31}} + \frac{\partial [C_{13}(V_{13} + V_{31})]}{\partial V_{31}} = 0. \end{cases}$$

Низка перетворень приводить до оптимізаційного рівняння такого вигляду:

$$H_{13} \left(\frac{C_{12}}{\beta_{12}} + \frac{C_{23}}{\beta_{23}} \right) + H_{31} \left(\frac{C_{12}}{\beta_{12}} + \frac{C_{23}}{\beta_{23}} \right) = (n + m)C_{13}, \quad (2.4)$$

де n , m – приріст числа каналів у ЗЛ між вузлами (3,1) і вузлами (1,3) відповідно.

Якщо приріст ЗЛ між цими вузлами проводиться блоками певної ємності M , то $(m + n) = M$.

З оптимізаційного рівняння (2.4) для базової структури у вигляді орієнтованої 3-вузлової мережі можна отримати як окремі випадки всі відомі критерії оптимізації для подібних базових структур, що свідчить про його достатню універсальність.

Якщо позначити вартість ЗЛ, що складають обхідні шляхи, які розташовані нижче пучка каналів 1, як C' , а ЗЛ, що складають шляхи, які розташовані вище пучка каналів 3, як C'' , то функція вартості наведеного фрагмента мережі матиме вигляд:

$$C = C' + C_1 N_1 + C_2 N_2 + C_3 N_3 + C'' \quad (2.5)$$

Для визначення стаціонарної точки функції (2.5) необхідно прирівняти до нуля її часткові похідні:

$$\frac{\partial C'}{\partial N_1} = \frac{\partial C'}{\partial N_3} = \frac{\partial N_2}{\partial N_3} = 0.$$

Загальний вигляд оптимізаційних рівнянь для розглянутого фрагмента мережі такий:

$$C_1 + C_2 \frac{\partial N_2}{\partial N_1} + \frac{\partial C''}{\partial N_1} = 0;$$

$$C_3 + \frac{\partial C''}{\partial N_3} = 0.$$

Таким чином показано, що критерій мінімуму вартості для 3-вузлової базової мережної структури з достатнім ступенем точності можна застосувати для будь-якого шляху високого використання в мережі більшої розмірності. Під час складання та розв'язання оптимізаційних рівнянь необхідно особливу увагу приділити вибору відповідних апроксимацій, застосування яких дає змогу забезпечити припустиму точність обчислень і водночас скоротити їхню кількість і трудомісткість переборів. Цю базову структуру можна розглядати як декомпозицію двох 3-вузлових базових структур, пов'язаних певною функціональною залежністю:

$$\frac{C_{13}}{H_{13}} = \frac{C_{12}}{\beta_{13}} + \frac{C_{23}}{\beta_{13}}; \quad (2.6)$$

$$\frac{C_{14}}{H_{14}} = \frac{C_{13}}{\beta_{13}} + \frac{C_{34}}{\beta_{34}}; \quad (2.7)$$

$$\beta_{13} = \left(\frac{\partial (R_{13} = f(r_{14}^e))}{\partial V_{13}} \right)_{\substack{E_{13}=\text{const} \\ Z_{13}=\text{const} \\ F_d=\text{const}}}$$

де $R_{13} = (A_{13} + r_{14}^e)$ – інтенсивність навантаження, що надходить на пучок каналів між вузлами (1,3);

Z_{13} – коефіцієнт дисперсії навантаження на даному пучку; r_{14}^e – надлишкове навантаження з пучка каналів між вузлами (1,4);

F_d – коефіцієнт доступності до пучка V_{13} ;

E_{13} – втрати на пучку V_{13} , отримані після визначення оптимального значення ємності пучка V_{14} .

Таким чином, функціональна залежність між рівняннями (2.6) і (2.7) визначається двома виразами:

$$\frac{\partial r_{13}^e}{\partial V_{13}} = (A_{13} + r_{14}^e) \frac{\partial E_{13}(A_{13} + r_{14}^e)}{\partial V_{13}};$$

$$\beta_{13} = \frac{\partial R_{13}}{\partial V_{13}} = \frac{\partial (A_{13} + r_{14}^e)}{\partial V_{13}}; \quad (2.8)$$

де A_{ij} – базовий потік між вузлами i та j ;

r_{ij}^e – ефективна інтенсивність трафіку від вузла i до вузла j .

2.4 Оптимізація базової мережі системи аудіоаналітики

Розв'язання оптимізаційних рівнянь (2.2), (2.4), (2.6), (2.7) для відповідних базових мережних структур дає змогу визначити оптимальні структурні параметри цих фрагментів мережі за заданого обмеження на якість обслуговування. Однак ітераційна процедура отримання такого розв'язку навіть для найпростішої 3-вузлової структури є досить трудомісткою, що призводить до необхідності використання апроксимацій, які дають змогу знизити обсяг обчислень.

2.4.1 Метод вибору мережної структури системи аудіоаналітики

Розглянемо спочатку алгоритм обчислення ємностей пучків каналів у 3-вузловій базовій структурі. Відповідно до оптимізаційного рівняння (2.3) для цієї базової структури процедура його розв'язування має ітераційний характер і включає наступні операції:

1. Задати початкові значення β_2 та β_3 ;
2. Обчислити значення H_1 відповідно до рівняння (2.3):

$$H_1 = C_1 / (C_2 / \beta_2 + C_3 / \beta_3);$$

3. Обчислити ємність прямого пучка каналів (першого вибору):

$V_1 = f(A_1, F_{d1})$ для значення H_1 , отриманого в п. 2 за умови, коли

$$H_1 = - \left(\frac{\partial r_1^e(A_1 V_1)}{\partial V_1} \right)_{\substack{A=\text{const} \\ E_{d1}=\text{const}}} = -A_1 \frac{\partial E_{V_1}(A_1)}{\partial V_1},$$

оскільки $r_1^e(A_1 V_1) = A_1 E_{V_1}(A)$;

4. Обчислити: параметри надлишкового навантаження з пучка каналів першого вибору V_1 , що надходить на перший обхідний напрямок (шлях високого використання (2,3)) – r_1^e, Z_1^e ; параметри навантаження, що надійшли на обслуговування на другу ЗЛ: $r_2 = A_2 + r_1^e$; $Z_2 = (A_2 + r_1^e Z_1^e) / r_2$; ємність пучка каналів на 2-ій ЗЛ – V_2 за умови, що втрати на пучку $V_2 : P_2 = P_{\text{доп}}$; середню величину навантаження, обслуженого на пучку V_2 : $Y_2 = r_2 - r_2^e$ і параметри навантаження, що надходить на третю ЗЛ: $r_3 = A_3 + Y_2$; $Z_3 = (r_3 + Y_2 Z_1^e) / r_1^e$; ємність пучка каналів на 3-ій ЗЛ – V_3 за умови, що втрати на пучку $V_3 : P_3 = P_{\text{доп}}$;

5. Обчислити значення $\beta_i = \left(\frac{\partial r_1^e}{\partial V_1} \right)_{\substack{Z_i = \text{const} \\ F_{d_i} = \text{const} \\ i=2,3}}$;

6. Порівняти значення $\beta_i (i = 2, 3)$ з заданою величиною ε : якщо $\beta_2 = \beta_3 \leq \varepsilon$, то закінчити обчислення; якщо $\beta_2 = \beta_3 > \varepsilon$, то перейти до п. 2.

Z-апроксимація є одним із найефективніших методів для розрахунку ймовірності втрат у ВК із внутрішніми блокуваннями на мережах з обхідними шляхами та визначення ємностей пучків каналів. Відповідно до цієї апроксимації ймовірності втрат викликів на пучку каналів ємністю V та ефективною доступністю d_e на який надходить навантаження з параметрами $(R_i, Z_i), Z_i > 0$, визначається як:

$$P = E\left(\frac{V}{Z_i}, \frac{R_i}{Z_i}, \frac{d_e}{Z_i}\right) = E\left(\frac{V}{Z_i}, \frac{R_i}{Z_i}, \frac{FdV}{Z_i}\right), \quad (2.9)$$

де $E(\bullet, \bullet, \bullet)$ – формула Ерланга для ідеального неповнодоступного включення.

За формулою (2.9) обчислюють надлишкове r_1^e й обслуговане Y навантаження за умови, що на пучок каналів надходить непуасонівське навантаження з параметрами (R_i, Z_i^e) .

Крім того, можна отримати оцінки коефіцієнтів дисперсії надлишкового Z_i^e й Z_y обслугованого навантажень за умови, що на пучок каналів надходить навантаження з параметрами (R_i, Z_i) .

Під час розв'язання задач оптимізації мережі в процесі її проектування коефіцієнт доступності F_d вважається відомим, а також задаються допустима величина втрат викликів на шляху останнього вибору і параметри початкового навантаження (R, Z) . При цьому (2.9) є нелінійним рівнянням з одним невідомим – ємністю пучка каналів V . При цьому ємність V визначається з урахуванням структури ВК.

Таким чином за допомогою формули (2.9) можна визначити ємності пучків каналів у пучках високого використання і пучків каналів у шляхах останнього вибору. З наведеного алгоритму розрахунку 3-вузлової базової структури випливає, що ємності пучків каналів першого вибору (на прямих шляхах) знаходяться з розв'язання рівняння:

$$\frac{\partial R_1}{\partial V_1} = H_1,$$

при фіксованому навантаженні A_1 , що надходить на цей пучок, і коефіцієнті доступності F , де параметр H_1 є константою і визначається в процесі оптимізації.

Процес обчислення є ітераційною процедурою. Оскільки на даному етапі оптимізації оцінюються тільки межі зміни змінних, то припустимою є деяка похибка у визначенні ємностей пучків, тому можна використовувати наближені формули, які дають змогу уникнути великої кількості ітерацій.

Так, для визначення максимальної ємності пучка V_{\max} першого вибору за умови $F_d \approx 1$, коли систему можна вважати неблокуючою, а $P=1\%$, запропоновано апроксимацію вигляду степеневі функції, коефіцієнти якої залежать від маргінального використання цього пучка каналів (для базової структури, представленої на рис. 2.6):

$$V_{\max} = 1,5 + 2,67(A_{ij}^{0,41} + A_{ji}^{0,41}) + 0,9855(A_{ij}^{0,41} + A_{ji}^{0,41}). \quad (2.10)$$

Для $P \leq 3\%$, що відповідає $F_d \approx 0,8$:

$$V_{\max} = 4,6 \left(A_{ij}^{0,499} A_{ij}^{0,075} + A_{ji}^{0,499} A_{ji}^{0,075} \right). \quad (2.11)$$

Формули (2.10) та (2.11) отримані в результаті застосування лінійної регресії при апроксимації формули (2.9).

Якщо пучки каналів мають модульну структуру з ємністю модуля, що дорівнює M , то число модулів K_{\max} у цьому випадку:

$$K_{\max} = \left[\frac{V_{\max}}{M} \right] + 1,$$

де $[\bullet]$ – ціла частина числа.

Аналогічні формули можна використовувати при визначенні ємності пучків каналів на обхідних шляхах наступного вибору, які не є одночасно шляхами останнього вибору. Наприклад, для базової структури, представленої на рис. 2.6, ці формули мають такий вигляд:

$$V_{\max} = 0,75 + 2,65 \left(\frac{R_i}{Z_i} \right)^{0,41} + 0,9855 \left(\frac{R_i}{Z_i} \right), \quad \text{для } F_d \approx 1;$$

$$V_{\max} = 4,6 \left(\frac{R_i}{Z_i} \right)^{0,499} \left(\frac{R_i}{Z_i} \right)^{0,075}, \quad \text{для } F_d \approx 0,8.$$

Ємності пучків каналів останнього вибору обчислюють за фіксованих значень величини втрат $P_{\text{дод}} = P_n$ для шляху останнього вибору та коефіцієнта доступності F_d під час визначення похідної виду:

$$\left(\frac{\partial Y}{\partial V} \right)_{\substack{P_H = \text{const} \\ F_d = \text{const}}} \quad (2.12)$$

Під час обґрунтування спільності критерію оптимізації 3-вузлової базової структури було здійснено перехід до скінченних різниць під час визначення похідних. Однак під час проведення обчислень заміна похідних кінцевими різницями може призвести до похибки, величина якої залежить від співвідношень структурних і навантажувальних параметрів. У випадках, коли потрібне знання похідних типу $\partial f(P) / \partial V$, де $f(P)$ – деяка функція, залежна від імовірності втрат, завдання може бути зведене до використання скінченних різниць лише за умови, що величина ΔV досить мала $\Delta V \approx 0,01$. Однак вимога вибору $\Delta V \ll 1$ призводить до необхідності використання інтегральних подань для таких формул як перша формула Ерланга і формула для ідеального неповнодоступного включення Ерланга.

Під час обчислення ємностей повнодоступних пучків каналів можна використовувати апроксимацію першої формули Ерланга:

$$E_V(A) \approx (1-S) \sqrt{\frac{S + A(1-S)^2}{V(1-S)^2}} \exp \left\{ V \ln \frac{AS}{V} + V - AS \right\},$$

$$\text{де } S = \frac{(V + A + 1) - \sqrt{(V + A + 1)^2 - 4AV}}{2A}.$$

З огляду на те, що шукана похідна (2.12) обчислюється при фіксованому значенні $P_H = \text{const}$, можна записати:

$$\frac{\partial Y}{\partial V} = (1 - P_H) \left[\frac{\partial R(P_H)}{\partial V} \right]_{F_d = \text{const}}, \quad (2.13)$$

де $R(P_H)$ – поступаюче непуасонівське навантаження в точці P_H .

З виразу (2.13) випливає, що похідна $\partial R / \partial V$ визначається в точці P_H , тому можна прийняти:

$$\frac{\partial R}{\partial V} \approx \frac{R(V + \Delta V)_{P_H = \text{const}} - R(V - \Delta V)_{P_H = \text{const}}}{2\Delta V},$$

де $R(V + \Delta V)$ и $R(V - \Delta V)$ – значення навантаження R у точках $(V + \Delta V)$ та $(V - \Delta V)$, відповідно, за умовою $P_H = \text{const}$.

Оскільки для реальних систем $F_d \leq 1$, то для визначення величин $R(V + \Delta V)$ и $R(V - \Delta V)$ використовують співвідношення (2.9), яке пов'язує між собою втрати, навантаження, ємність пучка і величину доступності.

Враховуючи (2.9) та те, що $P_H = \text{const}$, можна записати:

$$\sum_{x=\frac{VF_d}{Z}}^{\frac{V}{Z}} [x]_{\frac{V}{Z}} \sigma(x) = P_H, \quad (2.14)$$

де $[x]_{\frac{V}{Z}}$ и $\sigma(x)$ – функції R та Z .

На підставі отриманих раніше співвідношень, (2.14) можна відобразити як рівняння вигляду $R = f(V)$, за умовою $F_d = const$, $Z = const$. Розв'язати рівняння (2.14) можна з використанням ітераційної процедури, що передбачає знаходження похідної виду:

$$F' \left(\frac{R}{Z} \right) \approx - \frac{\partial \left(\sum_{x=\frac{VF_d}{Z}}^{\frac{V}{Z}} [x]_{\frac{V}{Z}} \sigma(x) \right)}{\partial \left(\frac{R}{Z} \right)},$$

що є досить трудомістким завданням.

Для скорочення обсягу обчислень було використано апроксимацію, що дає хороші результати для значень F_d , які перебувають у межах $0,8 \leq F_d \leq 1$:

$$F' \left(\frac{R}{Z} \right) \approx - \frac{\partial \left(\frac{V}{Z} \right)}{\partial \left(\frac{R}{Z} \right)} = - \frac{\partial E_{\frac{V}{Z}} \left(\frac{R}{Z} \right)}{\partial \left(\frac{R}{Z} \right)},$$

$$\text{де } \frac{\partial E_{\frac{V}{Z}} \left(\frac{R}{Z} \right)}{\partial \left(\frac{R}{Z} \right)} = E_{\frac{V}{Z}} \left(\frac{R}{Z} \right) \left[E_{\frac{V}{Z}} \left(\frac{R}{Z} \right) + \frac{V}{R} - 1 \right].$$

Дослідження характеру цього ітераційного процесу розв'язання рівняння (2.14) показало, що ця апроксимація для $F'(R/Z)$ за умови $P_H = 1\%$ забезпечує добру збіжність при відносній похибці у визначенні R , що дорівнює $\varepsilon = 0,0005$.

Визначення похідної $\left(\frac{\partial Y}{\partial V}\right)_{\substack{A=const \\ F_d=const}}$ зводиться до визначення похідної

$\left(-\frac{\partial P}{\partial V}\right)_{\substack{A=const \\ F_d=const}}$, оскільки:

$$\frac{\partial Y}{\partial V} = \frac{\partial [A(1-P)]}{\partial V} = \frac{\partial A}{\partial V}(1-P) = -A \frac{\partial P}{\partial V}.$$

Таким чином, для шуканої похідної може бути знайдено тільки чисельне значення з використанням інтегрального відображення цієї формули.

Після визначення ємностей пучків каналів ЗЛ, що становлять шлях останнього вибору, можна обчислити маргінальні навантаження β_2 і β_3 :

$$\beta_i = \left(\frac{\partial R_i}{\partial V_i}\right) \text{ за умовою } Z_i = const, E_i = const, F_d = const \text{ та } i = 2, 3.$$

Під час розв'язання оптимізаційного рівняння (2.4) для базової структури, представленої на рис. 2.6, завдання зводиться до розгляду двох 3-вузлових базових структур, на які декомпозиється структура, яку розглядають.. Ємності пучків каналів у прямому і зворотному напрямках ЗЛ визначають виходячи з умови модульності їхньої зміни, тобто число модулів у пучку дорівнює:

$$K_{ij(\max)} = \left[\frac{V_{ji}^* + V_{ij}^*}{M} \right] + 1,$$

де $V_{ij(ji)}^*$ – ємність пучка, отримана під час роздільної оптимізації;

M – ємність модуля первинної групи каналів.

2.4.2 Оптимізація мережної структури системи аудіоаналітики при неординарному пуасонівському навантаженні

Проведений аналіз різних стратегій поділу спільного прямого пучка (шляхи 1-го вибору) показав, що оптимальною є стратегія, яка відповідає розподілу пропорційно до розв'язків, отриманих унаслідок роздільної оптимізації двох 3-вузлових структур, як ефективніша в сенсі досягнення глобального мінімуму та мінімального числа обчислень. У такому випадку величини V_{13} та V_{31} визначаються пропорційно ємностям відповідних пучків, отриманих під час роздільної оптимізації, V_{13}^* та V_{31}^* :

$$V_{13} = \frac{V_{13}^*}{V_{13}^* + V_{31}^*} V; \quad V_{31} = \frac{V_{31}^*}{V_{13}^* + V_{31}^*} V,$$

де V – ємність пучка з урахуванням модульності побудови ЗЛ.

Розв'язання оптимізаційних рівнянь (2.6) та (2.7) для 4-вузлової базової мережної структури (рис. 2.7) з урахуванням функціональної залежності між рівняннями, що описується виразами (2.8) і (2.9), здійснюється відповідно з наступним алгоритмом:

1. Обрати початкове наближення для величини надлишкового навантаження з прямого шляху (1,4) $r_{14}^e = 0$;
2. Визначити величину навантаження, що надходить на ЗЛ (1,3) першого обхідного шляху $r_{13} = A_{13} + r_{14}^e$;
3. Вирішити оптимізаційне рівняння (2.6) для того, щоб визначити ємності пучків каналів (1,3) – V_{13} , (1,2) – V_{12} та (2,3) – V_{23} ;
4. Розв'язати оптимізаційне рівняння (2.7) за умовою, що V_{13} відомо (обчислено в п. 3). У цьому разі під час обчислення значення β_{13} передбачається, що втрати на пучку каналів (1,3) визначаються за формулою:

$$E_{13} = E(r_{13}, Z_{13}, V_{13}, d_{13}).$$

У результаті розв'язання рівняння (2.7) будуть визначені ємності пучків каналів (1,4) – V_{14} – (3,4) – V_{34} та нове значення величини надлишкового навантаження з пучка (1,4), r'_{14} ;

5. Перевірити: якщо $\left| \frac{r_{14}^e - r'_{14}}{r_{14}^e} \right| < \varepsilon$ то закінчити процес обчислення; в

іншому разі прийняти $r_{14}^e = r'_{14}$ та перейти до п. 2.

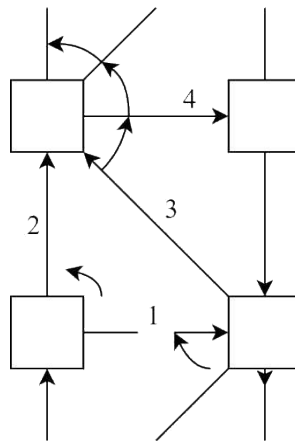


Рисунок 2.7 – 4-вузлова базова мережна структура

При створенні мережних систем аудіоаналітики особливо актуальним є питання про оптимальний розподіл пучків каналів на мережі. Застосування точних аналітичних моделей ускладнене через складну функціональну залежність між імовірністю втрат і навантажувальними та структурними параметрами. Крім того, застосування строгих аналітичних моделей вимагає великого обсягу обчислень, що обмежує розміри аналізованої мережі. Тому найбільш доцільним виглядає застосування наближених моделей, перевага яких полягає в уніформності обчислення характеристик та високій швидкості.

Будемо вважати, що на пучку між вузлами 1 і 3 має місце сукупність K неординарних пуассонівських навантажень, які замінимо на еквівалентні

рекурентні навантаження (тобто навантаження, що створюються рекурентними потоками викликів), які задаються парою $(R_i, Z_i), i = \overline{1, K}$. У цьому випадку рекурентне навантаження між вузлами 1 і 3:

$$R_1 = \sum_{i=1}^K R_i, Z_1 = \frac{\sum_{i=1}^K R_i Z_i}{R_1}.$$

Вважатимемо, що між вузлами 1 та 3 має місце власне навантаження (R_1, Z_1) , між вузлами 1 та 2 – власне пуасонівське навантаження A_2 , і між вузлами 2 та 3 – власне пуасонівське навантаження – A_3 (рис. 2.4).

Таким чином, оптимізаційне рівняння у випадку 3-вузлової мережі, представленої на рис. 2.6, матиме вигляд:

$$\frac{C_1}{H'_1} = \frac{C_2}{\beta_2} + \frac{C_3}{\beta_3}. \quad (2.15)$$

У разі використання оптимізаційного рівняння (2.15) алгоритм визначення ємності пучка каналів шляху першого вибору V_1 за значенням похідної H'_1 зі співвідношення (2.15):

$$H' = \frac{C_1}{\frac{C_2}{\beta_2} + \frac{C_3}{\beta_3}}. \quad (2.16)$$

З іншого боку, значення цієї похідної можна наближено визначити як:

$$H' = \frac{r_{1,2}^e - r_{1,1}^e}{2\Delta V}, \quad (2.17)$$

де $r_{1,1}^e = f((R_1, Z_1), V_1 + \Delta V)$ – надлишкове навантаження, що виникає на пучку каналів шляху першого вибору, ємність якого дорівнює $V_1 + \Delta V$;

$r_{1,2}^e = f((R_1, Z_1), V_1 - \Delta V)$ – надлишкове навантаження, що виникає на пучку ємністю $V_1 - \Delta V$.

Величину надлишкових навантажень $r_{1,i}^e$, $i = 1, 2$ можна визначити за формулою (2.9):

$$r_{1,1}^e = R_1 E \left(\frac{R_1}{Z_1}, \frac{V_1 + \Delta V}{Z_1}, \frac{F_d(V_1 + \Delta V)}{Z_1} \right);$$

$$r_{1,2}^e = R_1 E \left(\frac{R_1}{Z_1}, \frac{V_1 - \Delta V}{Z_1}, \frac{F_d(V_1 - \Delta V)}{Z_1} \right).$$

Оскільки H'_1 визначається зі співвідношення (2.16), то величину ємності V_1 пучка шляху першого вибору можна знайти ітераційно з розв'язання рівняння (2.17), що в цьому випадку набуде вигляду:

$$H'_1 = \frac{R_1 \left[E \left(\frac{R_1}{Z_1}, \frac{V_1 - \Delta V}{Z_1}, \frac{F_d(V_1 - \Delta V)}{Z_1} \right) - E \left(\frac{R_1}{Z_1}, \frac{V_1 + \Delta V}{Z_1}, \frac{F_d(V_1 + \Delta V)}{Z_1} \right) \right]}{2\Delta V}, \quad (2.18)$$

де $\Delta V = const$ та обирається відповідно до допустимої похибки обчислень.

Для визначення значення V_1 з рівняння (2.18) було використано таке ітераційне рівняння:

$$V_1^{[i+1]} = V_1^{[i]} (1 - (H_{11} - H_{12})), \quad (2.19)$$

де $H_{11} = const$ та знаходиться зі співвідношення (2.24), а H_{12} – з (2.26).

Як початкове наближення для величини ємності пучка використовуємо величину $V_1^{[0]} = R_1 \sqrt{Z_1}$. Як показали дослідження, процес є збіжним та

осцилюючим. При цьому число ітерацій у разі безпосереднього використання ітераційного рівняння (2.19) не перевищує 10-ти, а в разі прискорення збіжності не перевищує 5-ти.

Розглянутий алгоритм було перевірено під час розв'язання задачі оптимізації односпрямованої 3-вузлової мережної структури, для якої було отримано попередньо точний розв'язок. Вплив непуасонівського навантаження на характер оптимізації порівнюємо на прикладі 3-вузлової мережі, для якої: $A_2 = A_3 = 5$ Ерл, $C_1 = 1$ и $C_2 = C_3 = 1,5$ у. о., $R_1 = 5$ Ерл, а $Z_1 = 1, 2$ або 4. Результати оптимізації для трьох значень Z_1 зведено в табл. 2.1, у якій наведено результати оптимізації мережі, отримані шляхом повного перебору. Під час оптимізації мережі пучки каналів шляху останнього вибору визначалися при фіксованих середньозважених втратах в 1%.

Таблиця 2.1 – Результати оптимізації мережі шляхом повного перебору

Параметр	Коефіцієнт дисперсії навантаження					
	$Z_1 = 1$		$Z_1 = 2$		$Z_1 = 3$	
	аналітичне	точне	аналітичне	точне	аналітичне	точне
V_1	7	7	9	10	13	16
V_2	13	12	12	12	14	12
V_3	12	12	13	12	13	12
C	425	430	440	460	510	520
$\sigma, \%$	1.18		4.55		1.96	

Як видно з аналізу даних табл. 2.1, метод, узагальнений на випадок непуасонівських навантажень, дає результати, близькі до оптимальних (похибка не більше 6%). Збільшення коефіцієнта дисперсії Z_1 призводить до того, що пучок каналів першого вибору під час оптимізації матиме більшу ємність за тих самих умов порівняно з випадком, коли $Z_1 = 1$.

Таким чином, розглянутий вище алгоритм дає змогу оптимізувати 3-вузлові односпрямовані мережі. Описаний підхід поширюється на інші базові мережні структури. При застосуванні даного підходу до цих структур, зміни стосуватимуться лише визначення пучків каналів шляхів високого використання (прямий шлях та обхідний шлях високого використання). Після того, як буде проведено оптимізацію базових мережних структур та мережі загалом, можна проводити детальний аналіз якості обслуговування індивідуальних потоків викликів, тобто обчислювати ймовірності індивідуальних навантажень (інтенсивності та коефіцієнти дисперсії).

2.4.3 Оптимізація базової мережі системи аудіоаналітики за критерієм вартості

Оскільки розглянутий вище метод оптимізації мережі зв'язку за критерієм мінімальної вартості ґрунтується на низці припущень, то він є наближеним. Оцінимо точність цього методу.

Число каналів на мережі комутації прийнято розраховувати на основі критерію мінімуму капітальних витрат при фіксованій якості обслуговування викликів на пучках каналів шляху останнього вибору (ШОВ) і дотримання вимоги кратності сумарної кількості каналів у вихідному та вхідному пучках прямих шляхів (ПШ) числу каналів первинного модуля системи передачі. Так, при використанні аналогових систем передачі первинний модуль відповідає 12-ти каналам, а при використанні цифрових систем передачі – 13-ти каналам. При дотриманні цих вимог зазвичай приймають такі спрощувальні допущення:

1. Пучки каналів є повнодоступними та неблокованими;
2. Функція капітальних витрат є безперервною і диференційованою;
3. Параметри оптимізаційного рівняння залежать тільки від першого моменту розподілу навантаження.

Дані припущення дають змогу отримати оптимізаційні рівняння в порівняно простому вигляді та ефективно їх використовувати під час автоматизації проектування інформаційної мережі. Однак при цьому виникає деяка похибка у визначенні ємності пучків, оцінити котру аналітично досить важко. Похибку можна оцінити опосередковано. У якості таких методів можна використовувати пошукові методи оптимізації, наприклад, метод випадкового пошуку (МВП), який часто ефективно застосовують під час аналізу окремих фрагментів мережі. Зіставивши результати оптимізації мережі за двома методами, можна приблизно оцінити похибку.

Під цільовою функцією C розглядаються сумарні капітальні витрати на організацію мережі зв'язку:

$$C = \sum_{i \in N} C_i V_i, \quad (2.20)$$

де C_i – капітальні витрати на організацію одного каналу i -го напрямку;

V_i – ємність пучка каналів в i -му напрямку;

N_i – множина усіх напрямків.

Будемо вважати, що всі пучки каналів ПШ визначаються таким чином, що сума каналів у вихідному і вхідному пучках між кожною парою вузлів кратна прийнятому модулю первинної групи системи передачі M :

$$V_{\text{вх.пш}} + V_{\text{вих.пш}} = kM,$$

де k – множник кратності ($k > 0$), причому розподіл каналів по кожному з цих пучків здійснюється на основі забезпечення мінімуму сумарної надмірності навантаження. Ємність пучків каналів обхідних проміжних шляхів (ОПШ) також кратна M .

Пучки каналів ШОВ є пучками високої якості і їх можна визначити з точністю до одного каналу за ймовірності втрат $P = 0.01$. На пучки каналів

ОПШ надходять непуасонівські надлишкові навантаження, які задаватимемо їхніми першими двома моментами розподілу – математичним очікуванням інтенсивності навантаження R_i , та його дисперсією σ_i^2 . Для зручності надалі використовуватимемо коефіцієнт дисперсії $Z_i = \sigma_i^2 / R_i$.

Під час застосування МВП цільову функцію (2.20) необхідно уточнити, пов'язавши ємність пучків $V_i (i \in N)$ з параметрами навантаження, що надходить на цей пучок, імовірністю виникнення надлишкового навантаження та структурою вузла комутації. Вплив структури вузла комутації на V_i , будемо враховувати за допомогою формули:

$$P = E \left(\frac{R_i}{Z_i}, \frac{V_i}{Z_i}, \frac{d_{ei}}{Z_i} \right), \quad (2.21)$$

де $E(\cdot, \cdot, \cdot)$ – третя формула Ерланга;

P – ймовірність виникнення надлишкового навантаження;

V_i – ємність пучка каналів;

R_i – інтенсивність навантаження, що надходить на цей пучок;

d_{ei} – ефективна доступність у напрямку шукання;

Z_i – коефіцієнт дисперсії навантаження, що надходить.

Ефективна доступність d_{ei} пов'язана з ємністю пучка каналів коефіцієнтом доступності F_d :

$$d_{ei} = F_d V_i.$$

За наявних норм на якість обслуговування коефіцієнт доступності можна прийняти за $0,8 \leq F_d \leq 1$.

Щоб скористатися формулою (2.21), необхідно знати параметри навантаження, що надходить на пучок V_i , а саме пару R_i, Z_i . Будемо вважати, що навантаження, які надходять на цей пучок, незалежні. Тоді:

$$R_i = \sum_{j=1}^H R_{ij}; \quad \sigma_i^2 = \sum_{j=1}^H \sigma_{ij}^2;$$

$$Z_i = \frac{\sigma_i^2}{R_i} = \sum_{j=1}^H \sigma_{ij}^2 / \sum_{j=1}^H R_{ij},$$

де R_{ij}, σ_{ij}^2 – математичне очікування та дисперсія навантаження, що надходить у напрямку i від напрямку j ($i, j = \overline{1, H}$) відповідно.

Завдання відшукування параметрів непуасонівського навантаження, що надходить на блоковані пучки каналів, досі не знайшло суворого розв'язання. Наближений розв'язок цієї задачі:

$$R_H = RP; \tag{2.22}$$

$$R_0 = R - R_H; \tag{2.23}$$

$$Z_H = Z + \rho R_H F_d \tag{2.24}$$

$$\rho = \left[P \left(\frac{d_e}{Z} + 1 - Y \right) \right]^{-1} - 1; \tag{2.25}$$

$$P = E \left(\frac{Y}{1-P}, \frac{d_e}{Z}, \frac{d_e}{Z} \right); \tag{2.26}$$

$$Z_0 = Z^2 \sigma_0^2 / R_0; \tag{2.27}$$

$$\sigma_0^2 = \left(\frac{R}{Z} \right)^2 - \left(\frac{R}{Z} \right) \left[\frac{V}{Z} \right] \left(\frac{V}{F_d} + \frac{R}{Z} \right) + \left(\frac{R_0}{Z} \right) - \left(\frac{R_0}{Z} \right)^2, \tag{2.28}$$

де R, Z – параметри навантаження, що надходить на деякий напрямок шукання;

R_0, Z_0 – параметри навантаження, обслугованого цим напрямком;

R_H, Z_H – параметри навантаження, не обслугованого розглянутим напрямком і переданого на дообслуговування на обхідний напрямок зв'язку.

Перейдемо до опису методу, використовуваного для оптимізації цільової функції. З огляду на математичну природу розв'язуваної задачі будемо використовувати покроковий випадковий пошук, а саме метод із класу локальних методів крокового пошуку. Локальність обраного методу зумовлена незалежністю процесу пошуку від його передісторії.

Аналіз методів локального випадкового крокового пошуку показує, що для розв'язуваної задачі досить ефективно можна застосувати так званий алгоритм із поверненням при невдалому кроці, особливістю якого є знижена тенденція блукання в області мети. Вдалим кроком називатимемо такий крок, при якому чисельне значення цільової функції буде меншим за її значення на попередньому кроці. Нехай у просторі параметрів, що оптимізуються, зроблено деякий випадковий крок із точки, що характеризується вектором каналів \bar{V}_i і значенням функції мети $C(\bar{V}_i)$. У цьому разі нова випадкова точка характеризуватиметься вектором \bar{V}_{i+1} і новим значенням цільової функції $C(\bar{V}_{i+1})$. Якщо $C(\bar{V}_{i+1}) < C(\bar{V}_i)$, то зроблений крок вважається вдалим, якщо $C(\bar{V}_{i+1}) \geq C(\bar{V}_i)$ то крок вважається невдалим і здійснюється повернення на попередній етап оптимізації. Формула цього алгоритму має такий вигляд:

$$C(\bar{V}_{i+1}) = C(\bar{V} + \Delta \bar{V}_{i+1}),$$

$$\text{де } \Delta \bar{V}_{i+1} = \begin{cases} \alpha \bar{\xi} & \text{якщо } C(V_i) \leq C(V_{i-1}); \\ -\Delta \bar{V}_i & \text{якщо } C(V_i) > C(V_{i-1}); \end{cases}$$

$\bar{\alpha}$ – вектор масштабу, який визначається можливими межами зміни вектора змінних;

$\bar{\xi}$ – одиничний L -вимірний випадковий вектор.

Розглянемо особливості застосування оптимізації цільової функції (мінімізації функції капітальних витрат) за допомогою алгоритму випадкового пошуку фрагмента мережі, що містить сім станцій і два вузли (рис. 2.8).

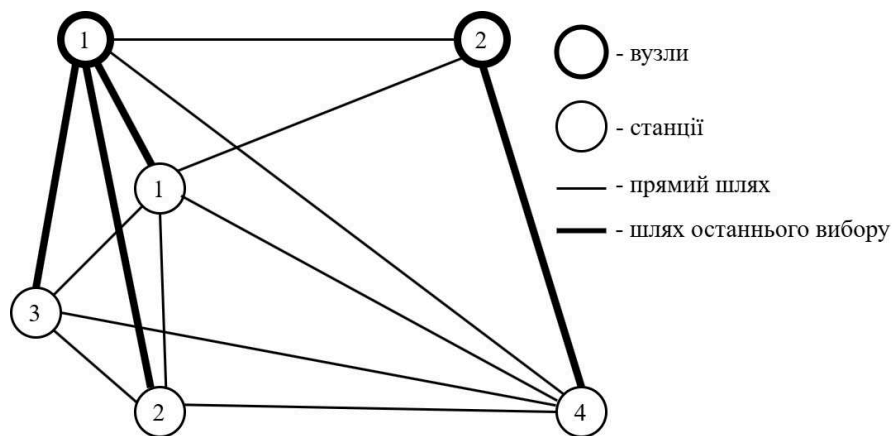


Рисунок 2.8 – Схема 2-вузлової мережі

Задамо такий алгоритм маршрутизації: за необхідності встановлення з'єднання між двома станціями використовувати принцип найкоротшого шляху. Це означає, що на першому етапі з'єднання робиться спроба його встановлення прямим пучком каналів між розглянутими станціями. Якщо ця спроба виявилася невдалою, то робиться спроба встановити з'єднання обхідним шляхом. При цьому можливі два випадки:

1. Станції включені в один і той самий вузол, тоді між станціями є тільки один обхідний шлях – ШОВ з імовірністю втрат на кожній ділянці цього шляху $P = 1\%$;
2. Станції включені в різні вузли, тоді між станціями є два обхідні шляхи: ОПШ і ШОВ.

Під час оптимізації цільової функції необхідно визначити кількість незалежних змінних. Відповідно до вимог, що висуваються до проектування

мережі зв'язку, для i -ї станції сумарне число вхідних і вихідних каналів, що утворюють пучки ПШ, які використовуються для зв'язку з j -ю станцією, повинно обиратися кратним модулю ущільнення (30 каналів); з кратністю одиниця обираються пучки ШОВ.

Всі пучки ПШ так само, як і пучки ОПШ, які обирають із кратністю 12, у цьому випадку є незалежними змінними. Це означає, що розмірність простору, в якому буде проводитися оптимізація, $L = 49$. Ця величина зумовлена: числом усіх ПШ між станціями, що дорівнює 42; числом пучків ОПШ, що дорівнює 6, між i -ю станцією ($i = 1, \dots, 6$) та другим вузлом і пучком ОПШ між сьомою станцією та першим вузлом.

Розглянемо детальніше етапи роботи алгоритму мінімізації капітальних витрат на створення мережі за допомогою методу випадкового пошуку. Для того щоб застосувати алгоритм методу випадкового пошуку з поверненням у початкову точку у разі невдалого кроку, потрібно вибрати початкове значення цільової функції. Необхідне число кроків для реалізації пасивного випадкового пошуку може бути оцінено за допомогою формули:

$$N = \frac{\ln(1-W)}{\ln(1-q)},$$

де W – імовірність знаходження значення цільової функції в обраній області її найкращих значень;

q – розмір області найкращих значень функції мети.

Нехай область найкращих значень цільової функції $q = 0,010$, тоді за ймовірності знаходження одного зі значень цільової функції $W = 0,95$ отримуємо, що для пошуку мінімуму цільової функції необхідно $N = 30$ кроків. Оскільки для цієї задачі максимальна зміна сумарної ємності пучків каналів шляхів першого і другого вибору відповідає одному модулю, це і визначає величину зміщення вектору за кожною координатою в L -вимірному просторі:

$$\bar{\xi} = \begin{cases} -1, & \text{якщо } 0 \leq \alpha \leq 1/3; \\ 0, & \text{якщо } 1/3 \leq \alpha \leq 2/3; \\ +1, & \text{якщо } 2/3 \leq \alpha \leq 1, \end{cases}$$

де α – випадкове число, рівномірно розподілене між 0 і 1.

Після цього визначається кількість модулів на пучках першого та другого виборів, а потім сумарна кількість вихідних і вхідних каналів. Ця сумарна кількість каналів розподіляється по вихідних та вхідних зв'язках так, щоб забезпечити мінімум сумарного надлишкового навантаження. Його параметри на відповідних пучках каналів визначають за формулами (2.22)-(2.28). Укрупнену блок-схему алгоритму наведено на рис. 2.9.

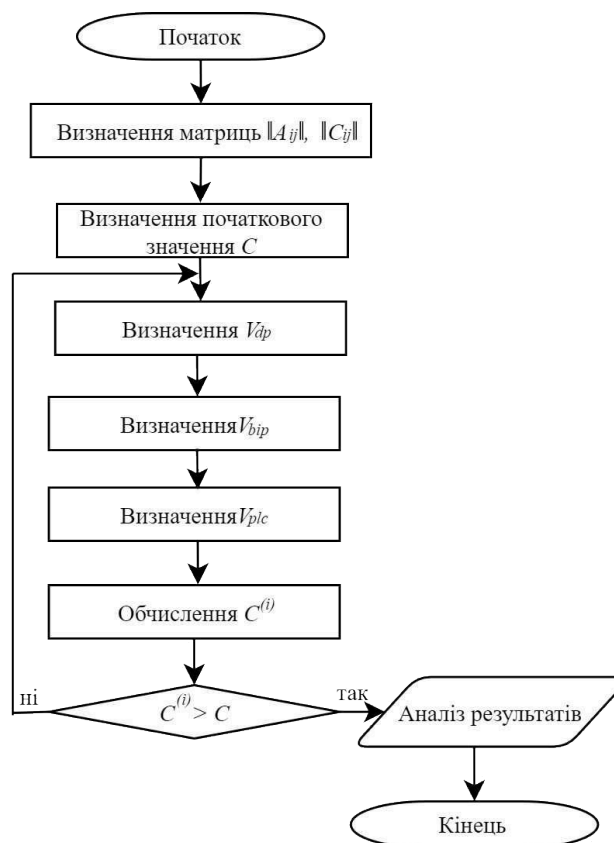


Рисунок 2.9 – Укрупнена блок-схема алгоритму оптимізації мережі за критерієм вартості

Початковими даними для оптимізації були: матриця навантажень $\|A_{ij}\|$ між станціями (табл. 2.2), матриця вартостей одного каналу між вузлами та станціями мережі $\|C_{ij}\|$ (табл. 2.3).

Таблиця 2.2 – Матриця навантажень між станціями

i	j			
	C_1	C_2	C_3	C_4
C_1	0	49,2	1,5	0,6
C_2	6,1	0	95,2	13,2
C_3	1,9	14,1	0	0,9
C_4	0,4	79,8	0,7	0

Таблиця 2.3 – Матриця вартостей між вузлами та станціями мережі

i	j					
	C_1	C_2	C_3	C_4	Y_1	Y_2
C_1	0	3426,5	2453,1	5321,8	3823,7	6502,5
C_2	3922,7	0	1437,4	4562,8	2752,3	5530,3
C_3	2680,1	1263,0	0	4208,5	1833,0	5987,5
C_4	4679,1	4232,0	3899,4	0	4005,4	3211,0
Y_1	3855,8	2600,3	1866,5	4682,0	0	6016,8
Y_2	7005,5	5226,9	6054,1	3022,4	6384,3	0

Як початкове наближення було використано оптимізовану матрицю пучків каналів, наведену в табл. 2.4. Вартість оптимізації цього варіанта мережі становить 1462,36 у. о.

Таблиця 2.4 – Матриця пучків каналів

i	j					
	C_1	C_2	C_3	C_4	Y_1	Y_2
C_1	0	54	4	0	11	0
C_2	8	0	112	14	12	0
C_3	3	18	0	0	8	0
C_4	0	78	0	0	7	6
Y_1	8	17	11	0	0	9
Y_2	0	0	0	10	7	0

Результат оптимізації мережі за допомогою МВП ілюструє матриця каналів (табл. 2.5), що відповідає 182-м крокам пошуку. Капітальні витрати цього варіанта розподілу каналів становлять $C = 1451,96$ у. о.

Для зазначених вище умов було проведено також оптимізацію мережі за МВП у припущенні, що на всіх станціях і вузлах допускаються внутрішні блокування, величина яких визначається вибором $F_d = 0,8$. Результат оптимізації ємностей пучків каналів мережі за допомогою МВП наведено в табл. 2.6. Аналіз табл. 2.6 показує, що перехід від повнодоступних пучків ($F_d = 1$) каналів на мережі до пучків із блокуваннями ($F_d = 0,8$) збільшує витрати приблизно на 7% ($C = 1551,92$ у. о.).

Таблиця 2.5 – Результат оптимізації мережі за допомогою МВП

i	j					
	C_1	C_2	C_3	C_4	Y_1	Y_2
C_1	0	59	0	0	10	0
C_2	13	0	106	14	16	0
C_3	0	30	0	0	10	0
C_4	0	81	0	0	8	3
Y_1	8	16	9	0	0	12
Y_2	0	0	0	12	5	0

Таблиця 2.6 – Оптимізація мережі за МВП з блокуваннями

i	j					
	C_1	C_2	C_3	C_4	Y_1	Y_2
C_1	0	59	4	0	9	0
C_2	13	0	115	11	24	0
C_3	4	31	0	0	21	0
C_4	0	72	0	0	17	4
Y_1	7	15	8	0	0	7
Y_2	0	0	0	7	5	0

Загальна структурна схема розробленої моделі наведена на рис. 2.10.

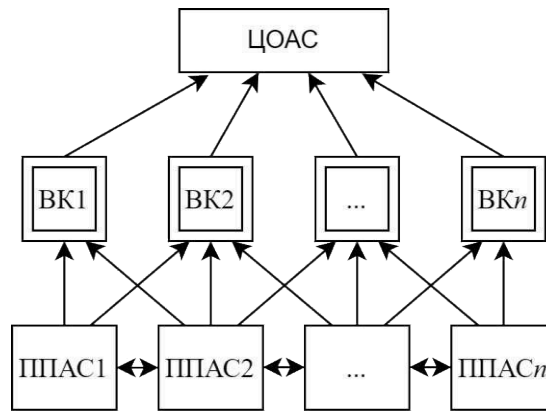


Рисунок 2.10 – Загальна структурна схема розробленої моделі

Основні елементи цієї моделі: що містить:

- ЦОАС – основний блок (зазвичай верхній рівень схеми), який отримує та обробляє аудіодані. Здійснює обчислення, зберігає дані, виконує алгоритми обробки (розпізнавання звуків, виявлення подій тощо).

- вузли комутації – проміжні блоки, через які здійснюється маршрутизація та передача аудіосигналів у мережі. Прямі шляхи або обхідні проміжні шляхи пов’язують кожен ВК між собою та з ЦОАС. Визначаються ємністю пучків каналів.

- пристрої приймання аудіосигналів – джерела аудіосигналів (мікрофони, сенсори тощо), де відбувається первинний збір даних. Кожен ППАС може бути підключений до одного чи кількох ВК. Передає аудіодані в мережу для подальшої обробки в ЦОАС.

Висновки за розділом 2

У результаті проведених досліджень отримано низку важливих результатів, а саме:

1. Запропонована постановка задачі оптимізації мережі з’єднувальних ліній за критерієм мінімуму капітальних витрат;
2. Складені оптимізаційні рівняння для однонаправленої трьохвузлової базової структури, виходячи із завдання мінімізації

капітальних витрат. Ці рівняння є базою для аналізу найпростішої (у структурному розумінні) конфігурації мережі. Отримані формули дозволяють знаходити оптимальний розподіл каналів із мінімальними затратами на обмеженому прикладі, який надалі можна розширити на складніші мережі;

3. Отримані оптимізаційні рівняння для двонаправленої трьохвузлової базової структури. На відміну від однонаправленої, двонаправлена структура відображає реальніші сценарії передачі даних у різних напрямках. Отримані рівняння враховують це двостороннє навантаження, що дозволяє точніше оцінити необхідну ємність кожного пучка каналів і зменшити втрати;

4. Проведено узагальнення раніше отриманих оптимізаційних рівнянь на будь-яку структуру мережі з'єднувальних ліній, що дає змогу поширювати результати досліджень на складніші топології;

5. Запропоновано метод вибору мережної структури системи аудіоаналітики, який враховує баланс між вартістю, продуктивністю та якістю обслуговування;

6. Розглянуто оптимізацію мережної структури системи аудіоаналітики в умовах неординарного пуасонівського навантаження, що дає змогу точніше прогнозувати й оцінювати роботу системи;

7. Виконана оптимізація базової мережі системи аудіоаналітики за критерієм вартості та проаналізовані результати оптимізації. Завдяки цій оптимізації показано реальну придатність описаних моделей та методів. Аналіз результатів дозволяє зрозуміти, як змінюються витрати й пропускна здатність під час реалізації запропонованих алгоритмів, і оцінити їх ефективність у реальному застосуванні.

Перелік використаних джерел у даному розділі наведено у повному переліку використаних джерел під номерами: 52–70.

3 РОЗРОБЛЕННЯ МЕТОДІВ ОБРОБКИ АУДІОСИГНАЛІВ У МЕРЕЖНИХ СИСТЕМАХ АУДІОАНАЛІТИКИ

У цьому розділі розглянуто методи, що забезпечують ефективну обробку аудіосигналів у мережних системах аудіоаналітики. Запропоновані рішення спрямовані на удосконалення передачі та аналізу аудіосигналів з урахуванням специфічних вимог до точності, швидкодії та стійкості системи. Особлива увага приділяється розробці методів управління трафіком і з'єднаннями, а також методам передачі та аналізу аудіосигналів, які є ключовими елементами для забезпечення високої якості аналізу аудіосигналів.

3.1 Обґрунтування показника якості передачі та аналізу аудіосигналу

Обґрунтування показника якості передачі та аналізу аудіосигналу відіграє визначальну роль у проектуванні систем аудіоаналітики, адже саме від цього залежить, наскільки ефективно вони виконуватимуть низку важливих завдань. До таких показників належить висока точність розпізнавання, адекватна фільтрація шумів та швидка обробка даних у режимі реального часу. Критично важливо враховувати надійність роботи системи навіть за складних умов навколишнього середовища, коли присутні значні акустичні перешкоди.

Для деяких сфер застосування, особливо де потрібна миттєва реакція, наприклад у голосових чатах чи відеоконференціях, одним із ключових показників якості стає мінімальна затримка між надходженням аудіоданих та їх обробкою. Якщо затримка є завеликою, користувачі можуть відчути дискомфорт, а у випадках критично важливих систем (наприклад, системи моніторингові системи) це може призвести до невчасного реагування на небезпечні події.

Висока точність розпізнавання насамперед пов'язана із здатністю системи вчасно та правильно ідентифікувати потрібні звукові події, а також проводити коректний аналіз мовлення. Це особливо актуально для застосувань, де автоматизація покликана зменшити людський фактор, покращити якість обслуговування або підвищити безпеку. Навіть дрібні помилки в таких системах можуть мати небажані наслідки, тому поширюється використання складних нейронних мереж чи гібридних моделей обробки, щоб поліпшити здатність алгоритмів розпізнавати сигнали.

Водночас надмірне ускладнення алгоритмів може спричинити відчутне збільшення часу, необхідного для обробки аудіоданих, що є неприйнятним для задач, де реальний час є критичним фактором. Задля досягнення балансу між швидкістю та точністю вдаються до різноманітних інженерних підходів, а саме оптимізують архітектури нейронних мереж, застосовують розподілені обчислення або спрощують математичні моделі там, де це можливо без значного зниження коректності результатів. Окрім того, важливо, щоб система аудіоаналітики була стійкою до змін у середовищі. Наприклад, вона має коректно працювати за наявності різноманітних шумів або перешкод, зумовлених погодними умовами, особливостями приміщення чи наявністю інших джерел звуку. Розв'язання цих викликів передбачає застосування спеціалізованих методів попередньої обробки, зокрема шумозаглушення, частотного аналізу та адаптивних фільтрів.

Найпоширенішим показником якості є саме точність розпізнавання аудіосигналу. Саме вона визначає, наскільки обґрунтованим буде впровадження технології у виробничі або сервісні процеси, і якою мірою система зможе автоматично виконувати завдання без втручання фахівців. Паралельно з цим велике значення мають часові обмеження, оскільки навіть найточніший алгоритм може виявитися марним, якщо результати надходять із надто великою затримкою.

3.2 Інформаційна безпека мережної системи аудіоаналітики

Зі зростанням обсягів аудіоданих, зумовленим розвитком таких технологій, як інтернет речей та хмарні обчислення, ризики, пов'язані з безпекою, стають дедалі складнішими. Мережі аудіоаналітики тепер повинні обробляти великі обсяги даних у реальному часі, що ставить нові вимоги до захищеності каналів передачі та процесів зберігання. Однією з найбільших загроз у цій сфері є вразливість самих аудіоданих, які можуть містити чутливу інформацію про користувачів, а саме їхні голоси, розмови, місцезнаходження та інші персональні деталі. Це робить системи аудіоаналітики цікавим об'єктом для атак, що можуть призвести до серйозних порушень конфіденційності та витоку даних. Останні досягнення підкреслюють необхідність у більш потужних протоколах шифрування та адаптивніших системах безпеки, здатних реагувати на загрози в реальному часі, особливо враховуючи експоненційне зростання обсягів інформації. Розподілені обчислювальні середовища, такі як мережні системи аудіоаналітики, розглядаються в [70].

Шифрування забезпечує конфіденційність даних, роблячи їх недоступними для зловмисників навіть у разі перехоплення. Методи шифрування даних для забезпечення безпеки обговорюються в [71,72]. Крім того, досліджуються нові методи шифрування, такі як квантове шифрування, для підвищення захисту без шкоди для ефективності робочих процесів обробки даних [73,74].

Ідентифікація та авторизація є важливими компонентами безпеки в мережних системах аудіоаналітики, оскільки вони гарантують, що доступ до конфіденційних даних і ресурсів отримують лише авторизовані користувачі. Досліджуються розширені методи ідентифікації, включаючи біометричну аутентифікацію та багатофакторну аутентифікацію, для подальшого зміцнення цих систем. Питання, пов'язані з ідентифікацією та авторизацією, розглядаються в [75,76].

Методи машинного навчання широко застосовуються в системах аудіоаналітики для вирішення завдань, таких як розпізнавання мовлення, класифікація звуків і виявлення аномалій у аудіоданих. Одним із найпоширеніших підходів є глибоке навчання, яке використовує нейронні мережі для аналізу аудіосигналів. Алгоритми глибокого навчання дозволяють ефективно обробляти великі обсяги даних та автоматично навчатися на їх основі, що робить їх ідеальними для аудіоаналітики, де важливими є якість та швидкість обробки. Сучасні методи машинного навчання та пов'язані з ними виклики обговорюються в [77, 78].

Крім технічних аспектів, аудіоаналітика також пов'язана з важливими етичними питаннями. Наприклад, зростає занепокоєння щодо збирання аудіоданих без відома користувачів або їх використання в неетичних цілях. У відповідь на це сучасні стандарти конфіденційності вимагають прозорості в обробці даних та надання контролю користувачам щодо того, як їхні дані збираються і використовуються.

Паралельно із загрозами розвиваються і технології захисту даних. Ці методи дозволяють мінімізувати ризики несанкціонованого доступу до даних, але швидке зростання складності кіберзагроз вимагає постійної адаптації захисних заходів.

Спрощена блок-схема процесу обробки даних у мережній системі аудіоаналітики зображена на рис. 3.1.

Перший етап представляє пристрої збору даних, такі як мікрофони, які відповідають за захоплення аудіосигналів. Ці пристрої забезпечують початкове отримання звукової інформації, що потім передається до системи для подальшої обробки.

Далі відбувається передача захоплених аудіоданих від пристроїв збору до центральної системи або сервера для обробки. У цьому процесі важливим аспектом є забезпечення захищеної передачі даних, що досягається за допомогою технологій шифрування.

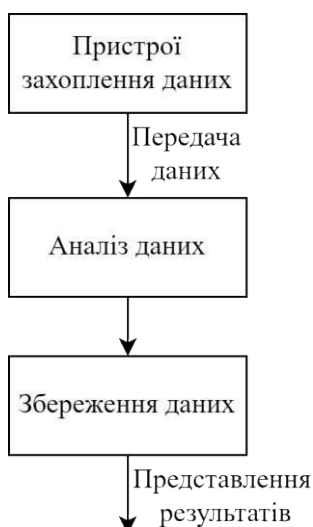


Рисунок 3.1 – Спрощена блок-схема процесу обробки даних у мережній системі аудіоаналітики

Наступним етапом є аналіз даних, де здійснюється обробка та аналіз зібраних аудіоданих. Це може включати розпізнавання мовлення, виявлення аномальних звуків або інший аналіз для перетворення сирих даних у корисну інформацію. На цьому етапі використовуються алгоритми машинного навчання або штучного інтелекту для досягнення точних результатів.

Після аналізу дані переходять до блоку збереження даних, де вони зберігаються для подальшого використання або доступу. Зберігання може відбуватися на локальних серверах або в хмарних сховищах. На цьому етапі важливо забезпечити захист даних від несанкціонованого доступу та використання відповідних механізмів контролю доступу та шифрування.

Останнім етапом є представлення результатів, коли результати аналізу аудіоданих доставляються кінцевим користувачам або інтегруються в інші системи для подальшого використання. Це може включати виведення результатів на графічний інтерфейс користувача або створення звітів для подальшого аналізу.

На кожному з етапів, представлених на схемі, існують потенційні можливості для атак. Можуть бути проведені фізичні атаки на самі пристрої, такі як мікрофони. Зловмисники можуть отримати доступ до цих пристроїв,

змінити їх або замінити на підроблені для захоплення або передачі неправильних даних. Крім того, можливі атаки типу перехоплення на рівні прошивки, що дозволяє контролювати пристрої дистанційно.

Етап передачі даних є одним з найбільш вразливих етапів. Можливі атаки типу «людина посередині», де зловмисник перехоплює передані дані між пристроями захоплення та центральним сервером. Без належного шифрування передані дані можуть бути зламані, змінені або записані. Навіть при шифруванні можуть виникнути вразливості, якщо використовуються слабкі протоколи або ключі.

Під час обробки аудіоданих на сервері можливі атаки на рівні алгоритмів машинного навчання або інших інструментів аналітики. Зловмисники можуть спробувати виконати атаки, що спотворюють результати обробки даних шляхом подачі спеціально змінених вхідних даних. Також є ризики інсайдерських атак, коли співробітники або підрядники можуть вплинути на результати обробки.

Основна загроза на етапі збереження даних пов'язана з несанкціонованим доступом до зашифрованих даних. Можливі атаки на рівні сховищ або баз даних, де зловмисники можуть отримати доступ до даних, використовуючи вразливості в системах контролю доступу. Також існує ризик фізичних атак на сервери, де зберігаються дані, або злому хмарних сховищ. На етапі отримання результатів можливі атаки на кінцевих користувачів, коли зловмисники можуть підробити результати або перехопити звіти, щоб викрасти або змінити критичну інформацію. Атаки можуть також бути на рівні інтерфейсу користувача (наприклад, через атаки на веб-додатки), що дозволяє маніпулювати відображеними результатами.

Законодавче регулювання обробки аудіоданих є важливим елементом забезпечення безпеки та конфіденційності інформації у мережних системах аудіоаналітики. Основним документом, який регулює обробку персональних даних у Європейському Союзі, є General Data Protection Regulation (GDPR) [79]. Цей регламент встановлює суворі правила щодо збору,

зберігання, обробки та передачі персональних даних, зокрема аудіоданих, що можуть бути використані для ідентифікації особи. Усі принципи, визначені в статтях 5.1-2 GDPR, є важливими для обробки аудіоданих.

Обробка аудіоданих повинна бути законною, зокрема, на підставі отриманої згоди або іншої правової підстави (наприклад, законного інтересу або контракту). Організації повинні інформувати користувачів про те, що їхні аудіодані збираються і використовуються для конкретних цілей. Прозорість є особливо важливою, коли йдеться про запис аудіо або моніторинг.

Аудіодані можуть збиратися лише для чітко визначених цілей, наприклад, для розпізнавання мовлення, покращення якості обслуговування клієнтів або безпеки. Ці цілі мають бути повідомлені користувачам заздалегідь, і подальша обробка цих даних для інших цілей без згоди користувачів є незаконною.

В аудіоаналітиці слід збирати лише необхідний обсяг аудіоданих для досягнення поставленої мети. Це означає, що записи слід скорочувати або фільтрувати, щоб не зберігати зайву інформацію.

У контексті аудіоданих точність має значення, коли йдеться про правильне збереження результатів аналізу. Якщо аудіодані використовуються для розпізнавання мовлення або підтвердження ідентичності, ці дані повинні бути точними і вчасно оновленими, щоб уникнути помилок або неправомірних рішень. Аудіодані повинні зберігатися лише протягом часу, необхідного для досягнення мети, з якою вони були зібрані. Наприклад, якщо аудіо використовується для покращення якості обслуговування клієнтів, воно може бути видалено після завершення аналізу. Якщо аудіодані більше не потрібні, їх необхідно видалити або анонімізувати. Важливо забезпечити захист аудіоданих від несанкціонованого доступу або витоків. Аудіодані часто містять конфіденційну інформацію, тому їх потрібно шифрувати під час передачі та зберігання. Важливо також обмежувати доступ до цих даних лише авторизованим користувачам.

Крім GDPR, існує множина інших законів, які регулюють захист персональних даних по всьому світу. Наприклад, у Бразилії діє General Personal Data Protection Law, що має багато схожостей із GDPR. У Сполучених Штатах у штаті Каліфорнія запроваджено California Consumer Privacy Act, а в Індії розроблено Personal Data Protection Bill 2019. Кожне з цих законодавств має на меті захист конфіденційності даних своїх громадян, хоча регулювання можуть відрізнятися залежно від конкретного регіону.

Забезпечення безпеки аудіоданих потребує комплексного підходу, що охоплює всі етапи роботи мережних систем аудіоаналітики. Основними методами забезпечення безпеки є шифрування, контроль доступу та моніторинг мережі.

На етапі збору аудіоданих важливо забезпечити захист пристроїв, які виконують запис звуку. Це включає в себе фізичний захист пристроїв від несанкціонованого доступу. Таким чином, аудіодані захищаються з моменту їхнього захоплення, що зменшує ризик перехоплення даних на джерелі.

Під час передачі аудіоданих по мережі критично важливо забезпечити їх захист від перехоплення, несанкціонованого доступу або маніпуляцій. Аудіодані, особливо якщо вони містять конфіденційну інформацію, можуть бути вразливими до атак під час передачі між джерелами даних та центральними серверами обробки. Для захисту таких даних зазвичай використовуються протоколи шифрування та віртуальні приватні мережі (VPN). Основними протоколами для передачі зашифрованих аудіоданих є Secure Real-time Transport Protocol (SRTP) та Transport Layer Security (TLS). SRTP забезпечує шифрування, автентифікацію та цілісність даних, що передаються по мережі, зокрема в таких застосунках, як інтернет-телефонія, відеоконференції, потокове передавання аудіо- та відеопотоків, а також системи аудіоаналітики. TLS захищає дані під час їх передачі від перехоплення, модифікації та несанкціонованого доступу, забезпечуючи три ключові аспекти: конфіденційність, цілісність і автентифікацію.

Використання VPN є одним із ключових методів захисту під час передачі аудіоданих по мережі, особливо коли йдеться про конфіденційні або персональні дані. VPN забезпечує створення зашифрованого каналу передачі даних між двома точками через загальнодоступні або приватні мережі, що робить цю технологію важливою складовою захисту у системах аудіоаналітики. За допомогою VPN шифруються не лише аудіодані, але й метадані про їх передачу, такі як IP-адреси та маршрути, що ускладнює перехоплення трафіку та виявлення інформації про учасників комунікації. У системах аудіоаналітики VPN може застосовуватися для захисту як при передачі аудіоданих між джерелами даних та серверами обробки, так і при доступі користувачів до системи з віддалених місць. Контроль доступу є певною системою механізмів, що забезпечують обмеження доступу до ресурсів лише для авторизованих користувачів. Важливою частиною політики контролю доступу є встановлення правил та процедур для автентифікації користувачів, управління їхніми правами доступу, а також відстеження їхньої активності.

Для забезпечення безпеки мережі використовуються засоби моніторингу мережі. Метою моніторингу є виявлення аномалій, спроб несанкціонованого доступу, атак або інших загроз, а також контроль за використанням ресурсів мережі, щоб забезпечити її безперебійне функціонування. Важливим аспектом мережного моніторингу є можливість виявлення інцидентів у реальному часі. Для цього використовуються спеціалізовані інструменти, такі як Intrusion Detection Systems (IDS) та Intrusion Prevention Systems (IPS). IDS системи аналізують трафік і повідомляють про підозрілу активність, наприклад, спроби вторгнення або відхилення від звичайного мережного трафіку. IPS системи можуть автоматично вживати заходів для блокування підозрілої активності, запобігаючи загрозам до того, як вони вплинуть на мережу.

Ключові виклики для забезпечення безпеки у мережних системах аудіоаналітики пов'язані з розвитком штучного інтелекту. З розвитком

технологій зловмисники починають використовувати машинне навчання для покращення своїх атак. Системи аудіоаналітики зазвичай використовуються моделі машинного навчання для обробки аудіосигналів, що відкриває нові можливості для атак зловмисників.

Атаки на мережу з використанням машинного навчання є небезпечною тенденцією. Зловмисники використовують алгоритми для автоматизованого виявлення вразливостей у мережних протоколах і системах. Алгоритми машинного та глибинного навчання можуть навчатися на великій кількості даних про мережний трафік, щоб знайти схеми, які можуть бути експлуатовані. Це дозволяє зловмисникам здійснювати більш точні та цілеспрямовані атаки.

З іншого боку, атаки на моделі машинного навчання стають серйозною проблемою для аудіоаналітичних систем, які залежать від аналізу даних і розпізнавання мовлення. Атаки на моделі машинного навчання також можуть включати отруєння даних, коли зловмисники вводять шкідливі дані під час навчання моделі, що призводить до того, що система вивчає хибні шаблони. Ще однією формою атаки є модельна інверсія, за якої зловмисник може відновити інформацію про вхідні дані або навіть витягти особисту інформацію з моделі, базуючись на її виведених результатах. Загрози, пов'язані з машинним навчанням, у майбутньому залишатимуться одним із головних викликів для безпеки мережних систем аудіоаналітики.

3.3 Метод передачі аудіосигналів у мережній системі аудіоаналітики

Передача аудіосигналів у мережних системах аудіоаналітики вимагає ефективних методів, які забезпечують мінімальні затримки, високу якість сигналу та стійкість до різноманітних перешкод. У сучасних умовах зростання обсягу аудіоданих і підвищення вимог до якості обслуговування стають ключовими аспектами. Зокрема, у системах, що працюють у режимі реального часу, критично важливою є здатність підтримувати стабільний

зв'язок і обробляти дані з високою швидкістю, що робить необхідним використання інноваційних підходів до управління трафіком та оптимізації передачі.

Ефективність системи аудіоаналітики залежить від точності розпізнавання аудіосигналів. Задача оцінювання точності методу обробки аудіосигналу наведена у роботі [80]. Метод передачі повинен забезпечувати стабільну передачу якісних звукових даних, щоб система могла правильно ідентифікувати та аналізувати аудіоінформацію.

Затримка та можливість працювати у режимі реального часу є критичними параметрами для ефективності методу передачі аудіосигналу у мережній системі аудіоаналітики, особливо у вимогливих застосуваннях, таких як моніторингові системи. Додатково, метод передачі повинен бути оптимізованим для мінімізації споживаної ширини смуги, особливо в умовах великої кількості даних, що передаються. Це робить систему більш ефективною та економічно вигідною. Ці питання розглянуті у роботі [81].

Метод передачі повинен демонструвати стійкість до помилок та шумів, які можуть виникнути під час передачі аудіосигналів через мережу. Забезпечення надійної передачі навіть при наявності різних перешкод є ключовим аспектом ефективності. Питання компресії аудіосигналу та аудіо кодекси розглянуті у роботах [82, 83].

Здатність методу передачі адаптуватися до збільшення обсягу даних та масштабу системи є важливим фактором в контексті розвитку та розширення застосувань мережної системи аудіоаналітики. Питання масштабованості складних мереж розглянуті у статтях [84, 85] та результати цих робіт можна застосувати при побудові мережних систем аудіоаналітики.

У зв'язку зі збільшенням кількості підключених пристроїв та обсягу обміну аудіоданими важливим є розробка методів захисту від потенційних атак та забезпечення конфіденційності переданих даних. Питання безпеки подібних мереж розглядаються у роботі [86].

3.3.1 Моделювання джерел навантаження системи

В епоху постійного розвитку технологій звуку та мережних систем, важливість передачі аудіосигналів у мережній системі аудіоаналітики стає незаперечною. Метод передачі аудіосигналу у мережній системі аудіоаналітики відіграє важливу роль у розвитку та впровадженні сучасних технологій, спрямованих на вдосконалення безпеки, зручності та ефективності різноманітних сфер нашого життя.

Мережні системи аудіоаналітики використовуються для нагляду та виявлення подій в громадських місцях, офісах, транспорті тощо. Ефективність цих систем залежить від точності та швидкості передачі аудіосигналів, що дозволяє реагувати на небезпеку або порушення в режимі реального часу. У сфері управління будівлями та інфраструктурними об'єктами використовуються мережні системи, які аналізують аудіосигнали для виявлення змін у середовищі. За допомогою передачі аудіосигналів у мережній системі можливо реалізувати інтерактивні системи управління та взаємодії, такі як голосові асистенти або системи розпізнавання голосу, що роблять використання технологій більш зручним та доступним. У медичній сфері передача аудіосигналів у мережній системі використовується для реєстрації та аналізу звуків, що може бути корисним у діагностиці різних станів або моніторингу пацієнтів. У великій мірі важливою є здатність методу передачі працювати у реальному часі. Затримка при передачі аудіосигналів може негативно впливати на швидкість реакції системи, що критично у сценаріях безпеки або моментального управління. Отже, метод передачі аудіосигналу у мережній системі аудіоаналітики стає ключовим інструментом для покращення різноманітних аспектів повсякденного життя, надаючи можливість використовувати звукову інформацію з максимальною ефективністю та безпекою.

Створення мережної системи аудіоаналітики призводить до появи нових класів джерел навантаження, які передають пакетизовані звукові дані.

Тому без побудови достатньо адекватних математичних моделей не можна побудувати мережну систему аудіоаналітики, що добре працює. Головною особливістю джерел навантаження мережних систем аудіоаналітики є те, що вони характеризуються різними швидкостями передачі інформації. Необхідна швидкість передачі джерела навантаження може мати випадковий характер як у момент виникнення заявки обслуговування, так і у процесі передачі. Цю обставину необхідно враховувати під час побудови мережної системи аудіоаналітики. Навантаження, створюване джерелами, вже не є звичайним пуассонівським навантаженням, і не є також маркованим пуассонівським навантаженням.

Введення нових моделей роботи джерел навантаження потребує нових описів для самого навантаження. Розробка аналітичних моделей навантаження повинна мати досить загальний характер, враховуючи багато параметрів джерел навантаження. В основу таких математичних моделей покладено марківський та спеціальний напівмарківський процеси.

При описі моделей навантажень, що виникають у мережних системах аудіоаналітики, використовується трирівнева модель обробки навантажень, що представлена на рис. 3.2.

Перший рівень, рівень викликів та відбоїв, відповідає за тривалість сеансу зв'язку. На цьому рівні необхідно враховувати активність окремих джерел навантаження, знати інтенсивність надходження викликів від них та інтенсивність їх обслуговування. Оскільки мережна система аудіоаналітики є мультисервисною мережею, вона характеризується великою різноманітністю джерел навантаження та їх кількістю. Перший рівень оперує зі шкалою часу секунди-години і безпосередньо пов'язан з визначенням необхідного числа віртуальних каналів для забезпечення всіх видів комунікаційного сервісу.

Другий рівень розшифровує сеанс зв'язку. Він визначає характер бітового потоку в моменти обслуговування користувача та відповідає моментам його активності та пасивності. Розподіл тривалостей пауз між бітовими потоками визначає берстність повідомлення. Бітовий потік, що

генерується джерелом, розділяється на осередки постійної довжини, що утворюються пакетизатором. Другий рівень оперує шкалою часу, що вимірюється секундами. На цьому рівні вирішується питання допуску виклику до мережі.

Третій рівень є рівнем передачі осередків, що надходять в транспортну мережу. Трансльовані осередки від усіх джерел навантаження утворюють трафік мережної системи аудіоаналітики. Статистичний закон розподілу цього трафіку сьогодні є предметом інтенсивних досліджень. Даний рівень характеризується типовими інтервалами часу, що становлять мілісекунди, і він безпосередньо визначає ємність буферних накопичувачів, ймовірність втрат осередків та час їх очікування у черзі.

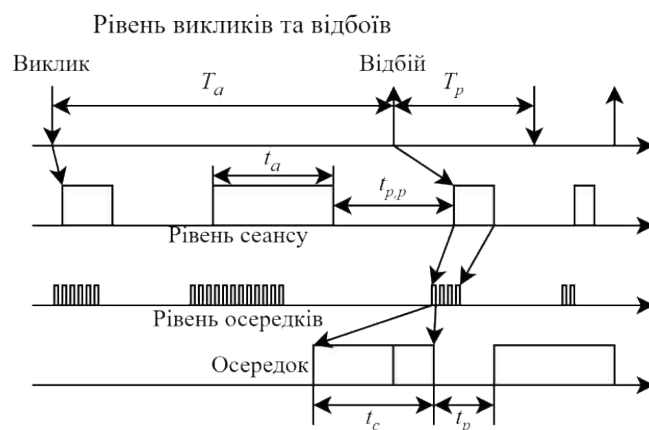


Рисунок 3.2 – Трирівнева модель навантаження:

T_a – час активного стану користувача; T_p – час пасивного стану користувача;

t_a – час активної передачі інформації користувача; t_p – тривалість паузи у передачі інформації; $t_{p,p}$ – тривалість паузи між передачею осередків; t_c –

час передачі осередка.

Кожен із рівнів характеризується певними моделями навантаження та відповідно характеристиками якості обслуговування. Мережна система аудіоаналітики є складною інформаційною системою, для якої аналітичні моделі тільки почали розроблятися порівняно недавно. Для перевірки

аналітичних моделей навантаження та якості обслуговування користувачів широко використовують імітаційне моделювання. У мережних систем аудіоаналітики QoS оцінюють не тільки як ймовірністю втрат інформаційних осередків, а й ймовірністю їх затримки, та навіть величиною джиттера затримки.

Як правила допуску виклику до обслуговування, так і сама стратегія обслуговування суттєво впливають на зазначені характеристики якості обслуговування. Тому проблема підтримки QoS (втрат повідомлення, затримки та джиттера) у мережній системі аудіоаналітики вирішується на всіх трьох рівнях обслуговування [87, 88].

Вивчення комплексу проблем першого рівня показує, що на першому етапі при нормальному функціонуванні обладнання мережної системи аудіоаналітики, її можна починати проектувати умовно як мережу комутації каналів. На цьому етапі можна оцінити пропускі здібності віртуальних шляхів для різних напрямів зв'язку та ефективність їх використання, вирішити питання про резервування ширини смуги бітової швидкості передачі для окремих класів користувачів, та про прийнятну якість обслуговування. Вивчення комплексу проблем другого рівня дозволяє усувати перевантаження бітових потоків та забезпечити необхідну якість надання послуги користувачеві.

Загальна ідеологія дослідження рівня інформаційного осередку, як правило, базується на припущенні, що потік осередків є безперервним із певною заданою середньою швидкістю. Один із підходів дослідження цієї проблеми заснований на двох припущеннях:

- мала ємність буферного накопичувача та відносно мала швидкість надходження інформаційних осередків, у яких ймовірність надходження двох осередків на один шлях була б мала;
- ймовірність перевантаження, з якою виникає велика кількість необслугованих осередків в один і той же час, дуже мала.

Ці припущення дають можливість моделювати рівень за допомогою бернуллієвського та пуассонівського випадкових процесів. Зазначені припущення призводять до добре вивчених моделей масового обслуговування з обмеженою ємністю буфера і детермінованим часом обслуговування. У таких моделях ймовірність втрат осередків головним чином залежить від інтенсивності навантаження та меншою мірою від числа джерел, що створюють навантаження. Крім того, такі моделі не є чутливими до пропорції осередків, що генеруються користувачами різних класів.

При вивченні рівня інтервалу слід звернути увагу на вивчення поведінки черг при перевантаженнях. Один з підходів у вивченні цього рівня базується на апроксимації, що отримала назву апроксимації «рідкий потік», при якій передбачається, що інформація надходить безперервно в період активності, і пристрій з буфером працюють таким чином, що характер вихідного потоку аналогічний вхідному. У цьому випадку використовують марківські моделі, що дозволяють отримати рішення у замкнутому вигляді для нескінченних черг та чисельне рішення для кінцевих черг.

Зі спрощенням аналітичної моделі потоку з'являється можливість врахувати такі важливі параметри, як загальна кількість джерел навантаження, максимальна швидкість передачі, кількість активних джерел та коефіцієнт активності або берстність джерел навантаження певного класу, яка, як було показано вище, характеризується значною варіацією для різних швидкостей передачі і видів інформації.

Від системи керування трафіком потрібна певна гнучкість, тому що необхідно підтримувати заданий рівень QoS. Складність вирішення цієї проблеми зумовлена великою різноманітністю класів користувачів. При цьому від системи керування вимагається, з одного боку, відносна простота реалізації і з іншого – робастність, тобто забезпечення високої ефективності керування ресурсами при зміні трафічної ситуації на мережі.

При організації доступу до мережної системи аудіоаналітики виділяють два аспекти:

- правило доступу виклику в мережі;
- стратегія обслуговування виклику, коли виклик вже прийнято до обслуговування.

Перший аспект пов'язаний із вивченням впливу пріоритетів на якість обслуговування, другий – з рівномірністю видачі інформаційних осередків у мережу. Правила допуску виклику в мережу ґрунтуються на знаннях про стан ланки, тобто числа прийнятих до обслуговування викликів різного типу. Рішення про прийняття наступного виклику приймається, якщо ймовірність втрат інформаційних осередків лежить у допустимих межах ($P = 10^{-5} - 10^{-9}$).

Щоб задовольнити ці аспекти, в мережній системі аудіоаналітики необхідно реалізувати дві функції управління та контролю трафіком в мережі: управління доступом користувача до мережі та управління параметрами використання.

Застосування асинхронного режиму передавання (АТМ) в мережних системах аудіоаналітики замислювалося не тільки як використання ефективного методу передачі, а й як гнучка технологія організації мультисервісного обслуговування користувачів. Використовуючи можливості, що надаються АТМ, створюються програмно-апаратні елементи, які можна використовувати як окремі будівельні сервісні блоки. В результаті цього з'являється гнучкий доступ до мережних ресурсів при забезпеченні компромісу між реалізацією та вартістю. Введення категорій АТМ-сервісу дозволило встановлювати віртуальні з'єднання каналів (VC) та шляхів (VP) з різною якістю надання послуги. Ця особливість мережних систем аудіоаналітики призвела до організації переговорів між мережею та кінцевим пристроєм для кожного нового з'єднання.

Введення категорій сервісу на рівні АТМ дозволяє організувати класи з'єднань, які мають гомогенні характеристики з точки зору зразка трафіку, що приймається до обслуговування, вимог QoS і можливого використання механізмів управління, роблячи цей клас з'єднань раціональним для

прийнятого способу розподілу ресурсів. Для організації класів з'єднань необхідні спеціальні базові підтримуючі механізми:

- керування допуском з'єднання (CAC);
- керування параметрами використання (UPC);
- управління зворотним зв'язком;
- управління розподілом ресурсів.

Механізм CAC визначений як набір дій, що здійснюються мережею протягом фази обробки виклику (фаза установки віртуального з'єднання) і протягом фази узгодження, щоб визначити, чи прийнята заявка на з'єднання або відкинута. Мережні ресурси (ширина смуги бітової швидкості передачі та розмір буфера) резервуються під час обробки заявки у кожному вузлі, якщо це потрібно прийнятої категорії послуги.

Механізм UPC визначений як набір дій виконуваних мережею для моніторингу за трафіком і управління ним, а також для створення ATM-з'єднання від користувачів до мережного інтерфейсу. Основна мета UPC полягає у захисті мережних ресурсів від можливості неправильного обслуговування користувача, при якому може змінитися QoS інших уже прийнятих до обслуговування користувачів, з'єднання яких вже встановлено. Механізм UPC базується на використанні алгоритму Generic Cell Rate, який контролює кожен осередок, що надходить, на відповідність його контракту, тобто прийнятому відповідності параметрів трафіку для дозволеного з'єднання. Якщо виявляються порушення параметрів трафіку, то приймаються відповідні дії – затримка, позначка осередку або навіть його режекція.

Механізм управління зворотним зв'язком (FC) визначається як множина дій, що здійснюються мережними та кінцевими системами (можливо у взаємодії), для регулювання трафіку з'єднання згідно з умовами, прийнятими мережними елементами. Специфічні процедури керування CAC пов'язані з категоріями ATM-сервісу. Управління розподілом ресурсу найефективніше здійснюється з допомогою запровадження резервування

ширини бітової швидкості передачі. Управління розподілом ресурсів найефективніше здійснюється з допомогою запровадження резервування ширини бітової швидкості передачі.

3.3.2 Управління з'єднаннями у мережній системі аудіоаналітики

Алгоритм управління допуску користувача в мережу САС визначає дії, що виконуються мережею, що дозволяють, в кінцевому рахунку, встановити з'єднання. За допомогою алгоритму САС з'єднання встановлюється тільки в тому випадку, якщо мережа може підтримувати це з'єднання з необхідним QoS і не погіршувати якості обслуговування вже встановлених з'єднань. Алгоритм дозволяє блокувати допуск у мережу нових з'єднань, якщо в ній зростають скупчення осередків у чергах. Він здійснює розподіл пікових значень бітової швидкості передачі осередків. За допомогою цього алгоритму можна зменшити можливість спроби повторного виклику. Алгоритм САС передбачає також можливість функціонування, коли абонент у зв'язку вимагатиме обслуговування з вищими параметрами, попередньо здійснивши перереєстрацію.

Дії алгоритму САС виконуються як у початковому етапі обслуговування виклику, так і на етапі переузгодження умов обслуговування виклику. Алгоритм САС вирішує питання, чи надати віртуальне каналне з'єднання VC або віртуальне з'єднання шляху VP споживачеві. Робота даного алгоритму забезпечує баланс між ефективністю та QoS.

Алгоритм САС для вирішення питання про прийняття чи відхилення з'єднання має визначити:

- прийнятні межі обсягу трафіку;
- необхідний клас QoS.

Ця інформація може бути переглянута протягом існування з'єднання. Мережа може обмежити частоту цих домовленостей із користувачем.

На етапі реалізації алгоритму САС визначається так званий трафічний контракт. Цей контракт полягає у необхідності дотримання набору параметрів обслуговування, включаючи параметри QoS. Відповідно до рекомендацій міжнародної спілки електрозв'язку (ITU), виконання контракту контролюється шляхом його перевірки на виконання відповідності прийнятому правилу обслуговування та мережному QoS. Контракт фіксує користувача, з яким можна змінити його подальше обслуговування.

Набір параметрів, що описують трафік, називається дескриптором трафіку з'єднання, що встановлюється. Дескриптор трафіку регламентований ITU і пов'язаний із встановленням єдиного з'єднання VP або з'єднання VS. Згідно з рекомендаціями ITU дескриптор трафіку визначає:

- параметри трафіку;
- відповідність параметрів трафіку перевірочному правилу;
- мережна QoS.

На етапі встановлення з'єднання узгоджується інформація, що міститься у специфікації трафічного контракту. Таким чином, дескриптор трафіку визначає можливість компромісних умов встановлення з'єднання між користувачем та мережею.

Отже, алгоритм САС представляє множину дій, які мають бути здійснені на етапі встановлення з'єднання для того, щоб прийняти або відкинути обслуговування користувача. Заявка на встановлення з'єднання приймається до обслуговування лише в тому випадку, коли є достатні ресурси для реалізації з'єднання по всій мережі з необхідним QoS. При цьому обов'язковою вимогою є збереження прийнятої якості обслуговування вже встановлених з'єднань.

3.3.3 Управління трафіком у мережній системі аудіоаналітики

Управління параметрами навантаження, що створюється користувачем, та управління параметрами мережі (UPC/NPC) здійснюється на рівні

інтерфейсів користувач-мережа (UNI) та мережа-вузол (NNI). Це управління являє собою набір дій, які повинні бути здійснені мережею для спостереження та управління трафіком, створюваним АТМ-з'єднаннями. Результат управління формулюється в термінах трафіку осередків та ефективності їхньої маршрутизації. Основна мета управління параметрами трафіку, що обслуговується, полягає в контролі за кожним встановленим з'єднанням щодо його більш суворої відповідності трафічному контракту. Ідеальний алгоритм UPC/NPC має відповідати наступним вимогам:

- здатність виявляти будь-які неприпустимі трафічні ситуації;
- швидка реакція на погіршення параметрів трафіку;
- проста реалізація.

Серед методів, що пропонуються для забезпечення цих вимог, значну увагу приділяється методу, який отримав умовну назву «діряве відро». Іншим методом, що широко застосовується в мережах передачі даних, є так званий метод ковзного вікна. Для побудови процедури управління застосовують два підходи: статистичний та заснований на використанні деяких апріорних правил. Недоліком статистичного підходу є необхідність тривалого спостереження та повільна реакція управління. Тому на практиці застосовуються операційні підходи.

У рекомендації 1.371 ІТУ представлено дві еквівалентні версії операційного алгоритму – так званого узагальненого алгоритму регулювання швидкості передачі осередків (GCRA). Перша версія отримала назву алгоритму віртуального розкладу (VSA), друга – алгоритму дірявого відра (LBA). Розглянемо ці версії докладніше.

Обидві версії GCRA для будь-якої послідовності моментів надходження осередків ($t_a \geq 1$) визначають осередки, які можуть бути передані за ланкою або їх передача буде затримана. В алгоритмі віртуального розкладу VSA використовують два параметри: I – збільшення; L – деяка межа його збільшення та множина проміжних змінних $GCRA(I, L)$.

В алгоритмі VSA обчислюється теоретично передбачуваний момент появи осередків TAT при припущенні, що осередки рівномірно розподілені в часі, причому відстань між двома послідовними осередками дорівнює I в момент активності джерела. Якщо фактичний час надходження перевищує $t = TAT - L$, де L – деяке прийнятне значення, тоді осередок допускається, інакше осередок затримується.

Схема алгоритму віртуального розкладу наведена на рис. 3.3.

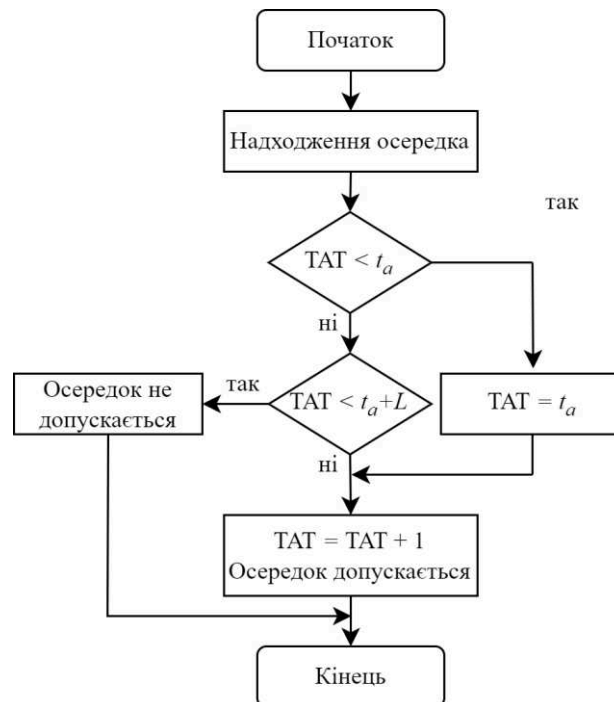


Рисунок 3.3 – Алгоритм віртуального розкладу

Розглянемо алгоритм дірявого відра. На рис. 3.4 показано, як берстний потік осередків заповнює відро В1. Якщо відро В1 не переповнено, то осередки в середовище передачі надходять із гарантованою швидкістю. Якщо В1 виявляється переповненим, то надлишковий потік надходить у відро В2. Якщо відро В2 не переповнено, то з В2 потік осередків надходить у середовище передачі з довільною швидкістю, але меншою, ніж гарантована. Якщо В2 виявляється переповненим, то надлишковий потік режектується.

Описаній схемі керування потоком відповідає алгоритм, наведений на рис. 3.5.

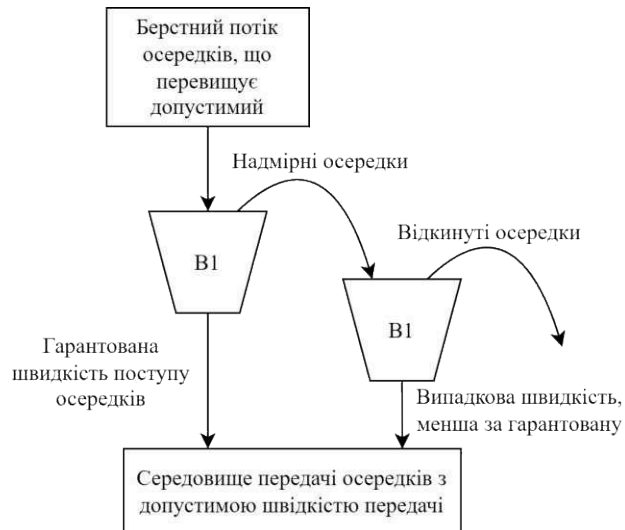


Рисунок 3.4 – Фізична інтерпретація методу дірявого відра

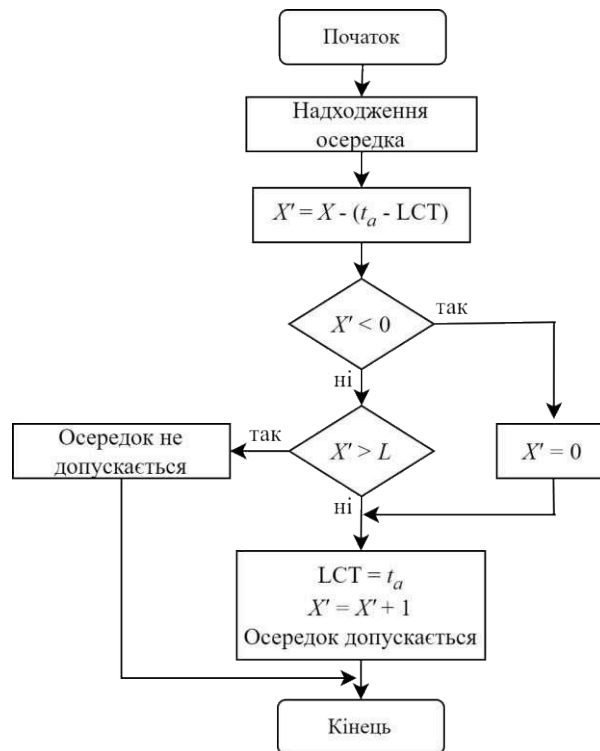


Рисунок 3.5 – Алгоритм управління бітовим потоком «діряве відро»

Часова схема процедури управління дірявого відра ілюструється схемами на рис. 3.6 та рис. 3.7. Рис. 3.6 відповідає випадку заповнення відра, а рис. 3.7 – випадку вставки осередків у середовище передачі. На відміну від механізму «ковзне вікно», механізм «діряве відро» оперує з індивідуальними осередками. Потік осередків копіюється в так звану псевдочергу. Черга

обслуговується з деякою швидкістю, яка вважається швидкістю протікання відра. Коли буфер наповниться, осередки вихідного потоку затримуються. Кожному встановленому з'єднанню ставиться у відповідність лічильник, вміст якого збільшується на одиницю у разі надходження інформаційного осередку і відповідно зменшується, якщо у ланці передачі має місце прийнятна поточна швидкість біта. Лічильник має деякий поріг L . Якщо вміст лічильника дорівнює L , то доступ осередків до мережі припиняється, і вони надходять у буферний накопичувач.

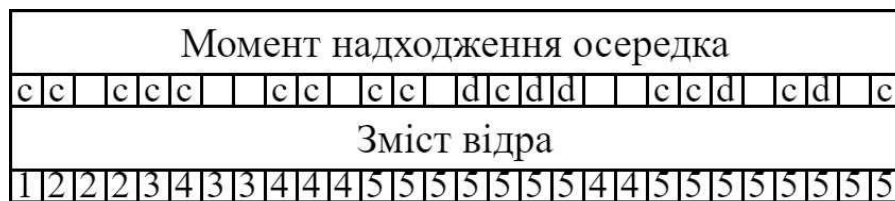


Рисунок 3.6 – Часова інтерпретація процедури «діряве відро» на етапі заповнення відра: c – допущений осередок; d – затриманий осередок; M – ємність відра ($M = 5$); $a = 0,33$ осередків/слот;

$$\#c = \min(\#находжень, M - Q); \#d = \min(0, \#находжень - \#c)$$

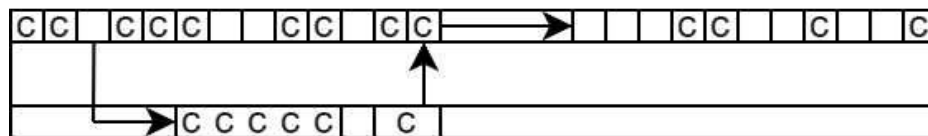


Рисунок 3.7 – Часова інтерпретація процедури «діряве відро» на етапі вставки осередків

3.3.4 Забезпечення якості передачі аудіосигналу

Семантична прозорість (СП) мережі визначається як здатність мережі транспортувати повідомлення від джерела до одержувача з прийнятною кількістю помилок. При заданій системі передачі прозорість мережі забезпечується:

- кодуванням повідомлення;

- повтором повідомлення на вимогу приймача;
- комбінацією прийомів п.1 та 2.

Найважливішою характеристикою, яка визначає СП, є можливість спотворення бітів інформації, так званий коефіцієнт помилок по бітам (BER):

$$BER = \frac{\text{Загальна кількість помилкових бітів}}{\text{Загальна кількість переданих бітів}}.$$

Для різних систем передачі вимір BER практично зводиться до набору статистики. У сучасних волоконно-оптичних лініях зв'язку (ВОЛЗ) величина $BER = 10^{-9}$. При збільшенні ймовірності BER відбувається збільшення числа повторних передач пошкоджених пакетів. Підвищення навантаження при цьому можна оцінити за формулою, що враховує використання протоколу керування «ковзне вікно»:

$$R(n) = \frac{W}{2} \frac{1 - (1 - BER)^{nL}}{(1 - BER)^{nL}},$$

де W – розмір вікна;

n – число ланок передачі;

L – довжина пакету, біт.

На рис. 3.8 представлено залежність зміни навантаження від ймовірності помилки BER на ланці передачі. Графік побудований для випадку п'яти ланок передачі. На цьому графіку трафік оцінюється стосовно вихідного трафіку. При малій величині BER управління від ланки до ланки мало ефективно, тому що ймовірність помилок виявляється вкрай малою. Найбільш ефективним у цьому разі є управління від кінця до кінця.

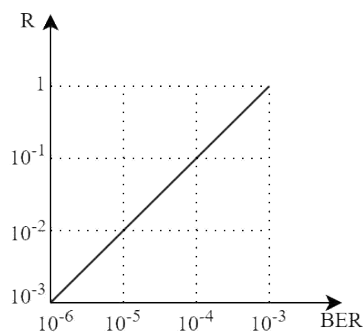


Рисунок 3.8 – Залежність зміни навантаження від ймовірності помилки BER на ланці передачі

Рис. 3.9 та 3.10 ілюструють випадок управління на ланці та управління від кінця до кінця. Рис. 3.9 відповідає випадку, коли довжина пакета становить 53 октети, а рис. 3.10 – коли довжина пакета становить 1000 октетів. Чим менша величина BER у каналі, тим ефективнішим виявляється управління від кінця до кінця порівняно з управлінням на ланці.

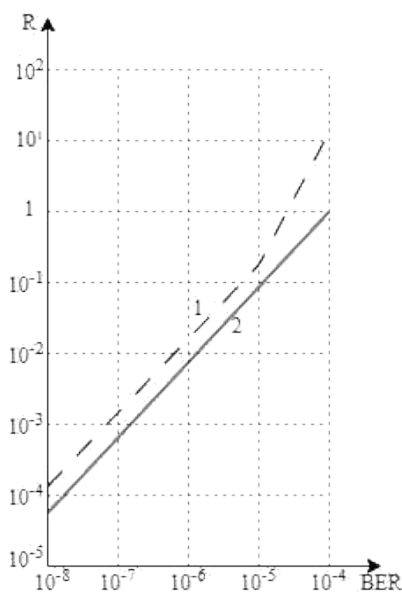


Рисунок 3.9 – Залежність зміни навантаження ланки від зміни ймовірності помилки на ланці при довжині пакета 53 октет: 1 – управління від кінця до кінця, п'ять ланок; 2 – управління на одній ланці передачі

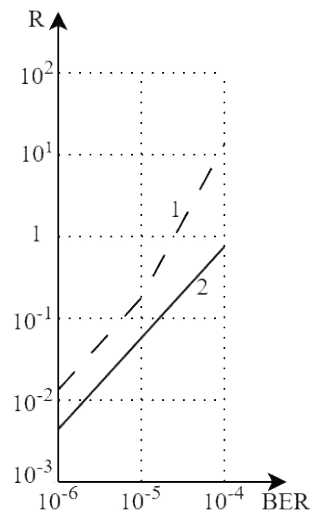


Рисунок 3.10 – Залежність зміни навантаження ланки від зміни ймовірності помилки на ланці при довжині пакета 1000 октетів: 1 – управління від кінця до кінця, п'ять ланок передачі; 2 – управління на одній ланці передачі

На рис. 3.11 представлена залежність ефективності управління від кінця до кінця в порівнянні з управлінням на ланці для стандартного осередку довжиною $53 \times 8 = 424$.

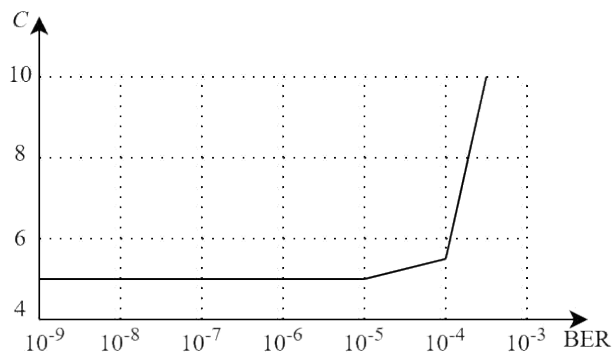


Рисунок 3.11 – Залежність зміни ефективності від ймовірності помилки

Ефективність можна оцінити відношенням навантажень тракту з керуванням на n ланках і керуванням на одній ланці:

$$C = \frac{R(n)}{R(1)}.$$

При цьому передбачається, що між користувачами є п'ять ланок передачі ($n = 5$), тобто:

$$C = \frac{R(5)}{R(1)}.$$

Із графіка, представленого на рис. 3.10 слід, що до величини $BER \leq 10^{-5}$ управління від кінця до кінця практично не відрізняється від управління на ланці. Якщо $BER \leq 10^{-9}$ то $C = 5,0$, а якщо $BER \leq 10^{-5}$ то $C = 5,04$.

У пакетно-орієнтованих мережах використовують іншу оцінку, а саме коефіцієнт помилок по пакетах (PER). Її обчислюють як число помилково прийнятих пакетів до загального числа переданих пакетів:

$$PER = \frac{\text{Кількість пакетів, переданих з помилками}}{\text{Загальна кількість переданих пакетів}}.$$

Величина PER визначається за деякий нормований час і істотно залежить від властивостей передаючого середовища. Переповнення буферних накопичувачів може також призвести до збільшення навантаження на ланці. Однак на практиці ймовірність втрат осередків виявляється екстремально малою ($P_{cell} \leq 10^{-9}$). Про величину ємності буферних накопичувачів, що забезпечують малу ймовірність втрат осередків, можна судити з рис. 3.12, для розрахунку якого було використано модель $M/D/1/L$ ($L \neq \infty$).

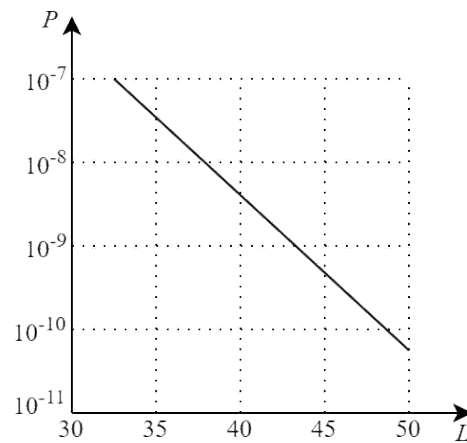


Рис. 3.12. Залежність ймовірності втрат осередків від зміни ємності буферного накопичувача при $a = 0,8$ Ерл

3.3.5 Дотримання часових вимог

Часова прозорість мережі оцінюється величиною затримки повідомлення та джиттера затримки повідомлення від кінця до кінця. Затримка істотно впливає на якість зв'язку для користувачів, що здійснюють передачу в реальному часі мови або відео. Великі затримки для мовних повідомлень проявляються у вигляді ефекту відлуння. Розмір затримки сигналу під час передачі від кінця до кінця нормується.

Час затримки складається із двох компонентів:

- затримки, зумовленої середовищем передачі;
- затримки, зумовленої обробкою повідомлень (пакетів). У цифрових системах комутації величина затримки обмежена значенням $T_{цск} \leq 450$ мкс (Q.607, ITU-T).

На загальні затримки у мережних системах аудіоаналітики впливають декілька факторів, а саме втрати пакетів, затримки передачі та джиттер. Метрика втрат вказує на те, скільки з відправлених джерелом пакетів дійшло адресата. Причиною втрат може бути проблема в інтерфейсі або кабелі, навантаження мережі, бітові помилки та ін. У випадки мережних систем аудіоаналітики, при виникненні втрат пакетів необхідна їх повторна

відправка до отримувача. Затримка передачі є часом, який потрібний даним, щоб дістатися від джерела до одержувача. Вона складається з затримки пакетизації, затримки середовища, затримки якості обслуговування та затримки обробки пакетів. Ці затримки є критичними для мережних систем аудіоаналітики, що працюють у режимі реального часу, і їх управління та оптимізація є важливими аспектами для забезпечення ефективності та надійності мережної системи аудіоаналітики.

Джитер є різницею у затримках між доставкою послідовних пакетів. У мережних системах аудіоаналітики аналоговий сигнал цифрується та розбивається на пакети даних. Приймаючий пристрій має невеликий буфер фіксованого розміру, куди укладаються послідовно отримані пакети. Для відновлення аналогового сигналу необхідно мати достатню кількість пакетів. У випадку змінних затримок наступний блок даних може не прийти вчасно, що еквівалентно його втраті, і сигнал не може бути відновлений.

Боротьба з затримками в мережі може включати в себе різні стратегії та техніки, спрямовані на оптимізацію та поліпшення швидкодії передачі даних, наприклад, використання механізмів QoS. Однак впровадження QoS в мережних системах аудіоаналітики має свої особливості через їх складнішу архітектуру.

Зазвичай, основні етапи забезпечення QoS включають в себе класифікацію та маркування трафіку, полісінг, застосування механізмів черг та вирівнювання трафіку. Через класифікацію та маркування виконується розподіл трафіку за класами для визначення пріоритету обробки. Полісінг використовується для обмеження або регулювання обсягу передачі даних через мережний інтерфейс до певного рівня, відповідно до заздалегідь встановлених правил і параметрів. У випадку, якщо трафік потребує обмеження або пріоритезації, пакети можуть бути поміщені у черги згідно з їхнім QoS-класом. Вирівнювання трафіку включає в себе регулювання рівня вихідного трафіку з метою зменшення переповнення обсягу трафіку.

Ці механізми впровадження QoS не завжди підходять сучасним мережним системам аудіоаналітики через їх складнішу архітектуру. В контексті мережних систем аудіоаналітики важливо врахувати специфіку пакетизації та депакетизації даних, а також можливого шифрування, що може впливати на механізми керування чергами і забезпечення QoS. Оскільки весь захищений трафік проходить через пакетизатор, важливо встановити механізми керування чергами як на виході, так і на вході цього пристрою. Це дозволить керувати потоками даних перед пакетизацією та після депакетизації, забезпечуючи потрібний рівень QoS.

Розмір черги і параметри керування чергами повинні бути встановлені з урахуванням вимог пропускної здатності пакетизатора та криптографічних алгоритмів, за наявності. Це може включати в себе визначення максимальної допустимої затримки та обсягу черги для кожного типу трафіку. Використання кількох черг з різними рівнями пріоритету дозволяє обробляти високопріоритетний трафік з мінімальною затримкою, при цьому забезпечуючи доступ до ресурсів і для менш критичних типів трафіку. Але, необхідно уникнути того, щоб менш важливий трафік був блокований в черзі і не отримував достатньо ресурсів через інтенсивний потік більш пріоритетних даних.

Час затримки, зумовлений обробкою повідомлень, у першому наближенні можна оцінити за методом Кроммеліна, згідно з яким, ймовірність того, що фактичний час очікування перевищить допустимий час t_a , визначається у наступному вигляді:

$$P(> t_a) = 1 - (1 - \lambda) \left[\sum_{k=0}^{[t]} \frac{(\lambda(k-t))^k}{k!} e^{\lambda(t-k)} \right].$$

Вид функції $P(> t_a)$ для $\lambda = 0,2, 0,4$ та $0,8$ як функція часу очікування представлена на рис. 3.14.

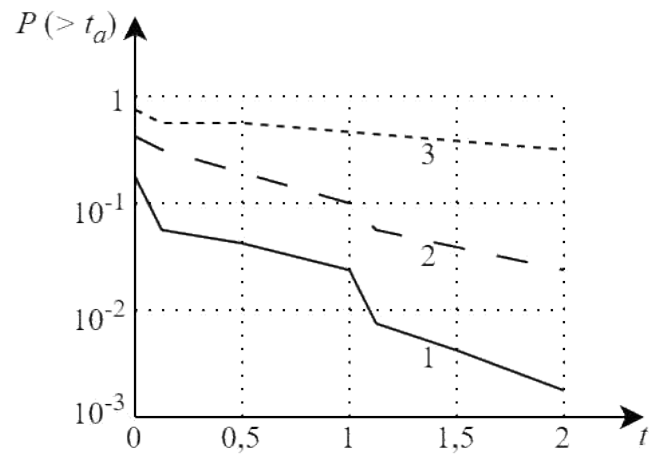


Рисунок 3.14 – Зміна ймовірності $P(>t_a)$ для трьох значень навантаження:
0,2 Ерл; 0,4 Ерл; 0,8 Ерл

Залежність середнього часу перебування у черзі T_c затриманих осередків від зміни навантаження a показано на рис. 3.15.

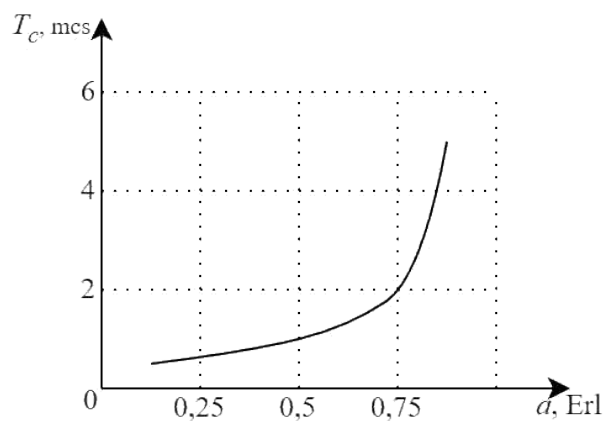


Рисунок 3.15 – Залежність середнього часу затримки від зміни інтенсивності навантаження

Про характер затримок, мкс, для гбїт/с з'єднання різних швидкостей передачі на ланці передачі (1 і 10 Гбїт/с) можна судити по табл. 3.1. Дані табл. 3.1 отримані при ймовірності втрати пакету 10^{-10} і інтенсивності поступаючого навантаження $a = 0,8$ Ерл.

Таблиця 3.1 – Залежність затримок від швидкості передачі на ланці

Затримки, мкс	Швидкість передачі на ланці, Гбіт /с					
	1			10		
	Розмір пакету, байт					
	256	512	256	512	256	512
Передачі	4000	4000	4000	4000	4000	4000
Фіксована на вузлі комутації	16	32	16	32	16	32
У черзі та при депакектизації	100	200	100	200	100	200
Пакетизації	500	1000	500	1000	500	1000
У синхронній мережі	900	900	900	900	900	900

3.3.6 Метод передачі аудіосигналу за рахунок комутації віртуальних маршрутів

Метод передачі аудіосигналу в мережній системі аудіоаналітики за рахунок комутації віртуальних маршрутів базується на принципі швидкої комутації пакетів. Цей метод розроблено з метою оптимізації передачі аудіосигналів у мережній системі аудіоаналітики. Основною метою є забезпечення ефективності та точності аудіоаналітики у режимі реального часу. При створенні цього методу було розроблено математичні моделі для адекватного відображення нових класів джерел навантаження, які передають пакетизовані звукові дані та було виконано аналіз і вдосконалення методів управління з'єднаннями та трафіком в мережній системі для використання ресурсів. У результаті було впроваджено механізми управління доступом користувача до мережі та управління параметрами використання ресурсів, а також створено ефективний метод комутації, який базується на використанні віртуальних маршрутів для направлення пакетів аудіосигналу.

Для комутації двох терміналів між собою потрібна не тільки комутація віртуальних каналів, але й віртуальних шляхів. Схематично показано на рис. 3.16 та 3.17, де VP-комутатор – комутатор віртуальних шляхів, VP/VC-комутатор – комутатор віртуальних шляхів та віртуальних каналів. VP/VC-комутатор (рис. 3.16) складається з комутатора віртуальних шляхів (VP-комутатора) та комутатора віртуальних каналів (VC-комутатора).

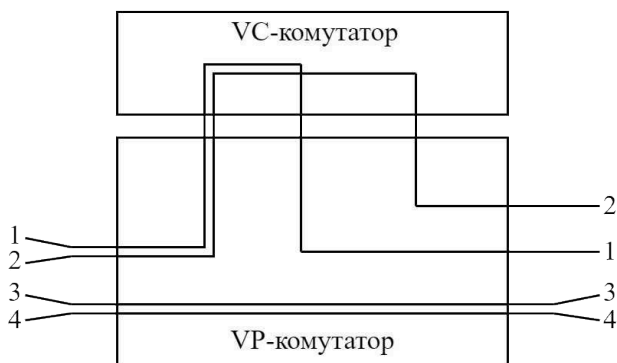


Рисунок 3.16 – Концепція комутації віртуальних шляхів та віртуальних каналів

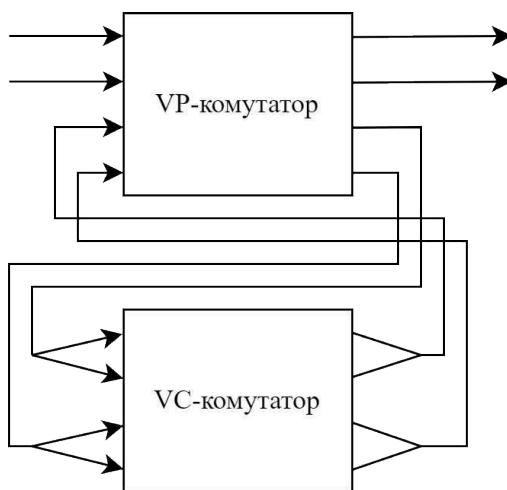


Рисунок 3.17 – Схема комутації віртуальних шляхів та каналів

Принцип швидкої комутації пакетів (ШКП) наведено на рис. 3.18. Осередок, що надходить на ВК, містить інформаційне поле та заголовок. На ВК, за допомогою спеціального процесу маршрутизації, осередку приписуються біти додаткової адреси, що представляють адресу маршруту

просування осередка за системою комутації. У міру проходження осередка каскадами системи комутації біти додаткової адреси використовуються для вибору маршруту. Осередок на виході містить лише інформаційне поле та заголовок

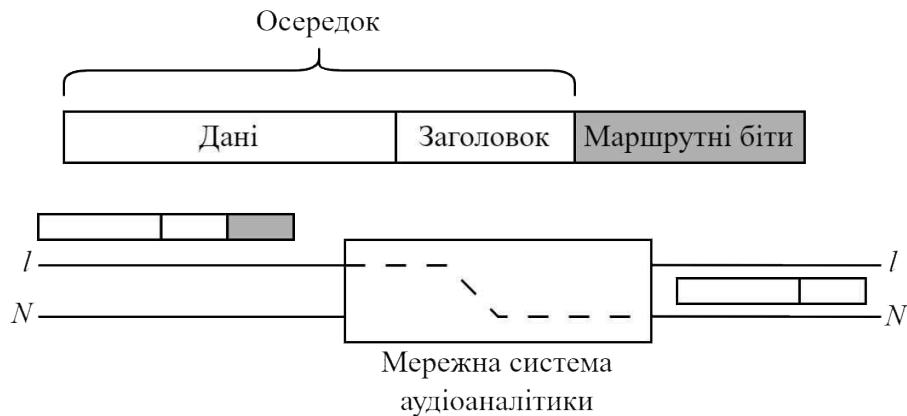


Рисунок 3.18 – Принцип швидкої комутації пакетів

З'єднання комутаторів VP та VP/VC формує з'єднання на мережній системі аудіоаналітики, що наведене на рис. 3.19.

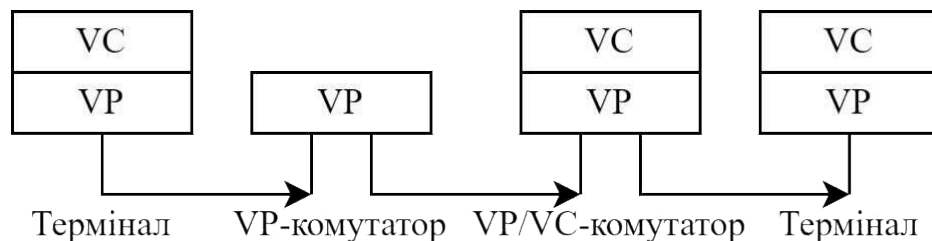


Рисунок 3.19 – Схема комутації віртуальних шляхів та каналів на мережі

На рис. 3.20 представлена умовна залежність витрат керування від швидкості передачі для різних видів комутації. Як випливає з рис. 3.20, витрати на управління істотно більші, ніж у разі комутації каналів та багатоканальної комутації, але менші, ніж у разі комутації пакетів у режимі віртуального каналу та дейтаграмному режимі.

Дана комутаційна система є системою великої структурної складності. Тому всі реальні методи розрахунку її пропускної спроможності наближені.

При цьому виникають суттєві труднощі перевірки цих методів за допомогою імітаційного моделювання.

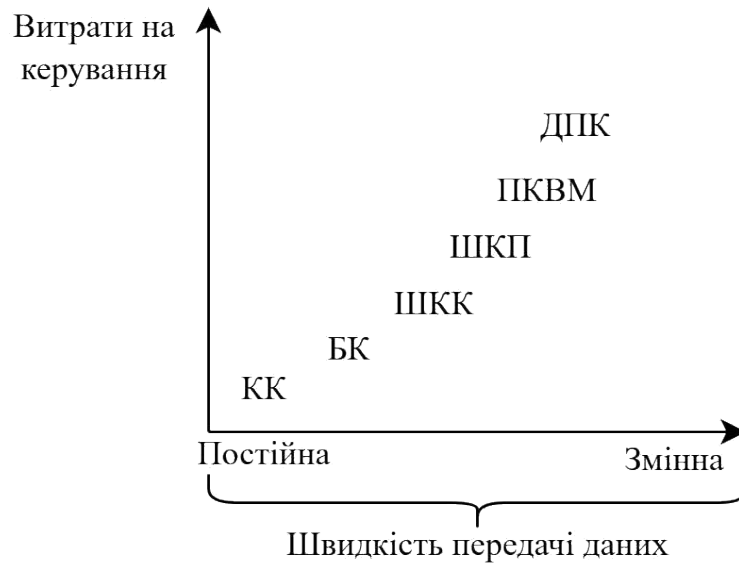


Рисунок 3.20 – Залежність витрат на керування від швидкості передачі для різних режимів комутації: КК – комутація каналів; БК – багатоканальна комутація; ШКК – швидка комутація каналів; ШКП – швидка комутація пакетів; ПКВМ – пакетна комутація віртуальними маршрутами; ДПК – дейтаграмна пакетна комутація

Можливості використання імітаційного моделювання вузлів мережної системи аудіоаналітики дуже обмежені. Це пояснюється тією обставиною, що допустимий рівень втрат інформаційних осередків дуже низький. Допустима ймовірність втрат осередків лежить у діапазоні $P \leq 10^{-9} - 10^{-10}$. У цьому діапазоні втрат при сучасних комп'ютерах досягти достовірних статистичних результатів поки неможливо. У зв'язку з цим проводяться інтенсивні дослідження у галузі моделювання рідкісних подій.

При побудові імовірнісних моделей мережних систем аудіоаналітики використовують різноманітних апроксимації: марківські процеси, комбінаторні підходи, методи, засновані на декомпозиції системи на підсистеми з більш простою структурою. У деяких випадках при оцінці

затримок інформаційних осередків застосовують апроксимацію Вінера-Хопфа, що дозволяє іноді отримувати досить прості формули за умови, коли вхідний потік заявок є марківським модульованим пуассонівським потоком.

На якість обслуговування інформаційних осередків в мережній системі аудіоаналітики істотно впливає маршрутизація. На ВК мережної системи аудіоаналітики здійснюється так званий режим ШКП. Цьому режиму комутації присутні як властивості режиму комутації каналів, так і режиму комутації пакетів. Реалізація таких систем здійснюється на основі мереж типу Баньян, що самомаршрутизуються. Кожен каскад такої системи складається з комутаторів із двома входами та виходами. Спочатку такі мережі призначалися для побудови багатопроцесорних обчислювальних систем, де вони працювали в режимі разової комутації, при якому з'єднання встановлюються за попередньо складеним списком. Застосування цих мереж безпосередньо для ШКП не є можливим через виникаючі у системі конфлікти інформаційних осередків, тобто ситуації, коли виникає необхідність передачі двох інформаційних осередків по тому самому виходу комутатора. Для усунення конфліктів на кожному вході комутатора необхідно встановлювати буферний накопичувач, у якому утворюється черга з інформаційних осередків. Це призводить до того, що час затримки інформаційного осередку під час передачі за алгоритмом від кінця до кінця може перевищити допустиме. Щоб зменшити час затримки осередків, слід зменшити кількість послідовно включених у системі буферів.

Схема проходження аудіосигналу наведена на рис. 3.21.

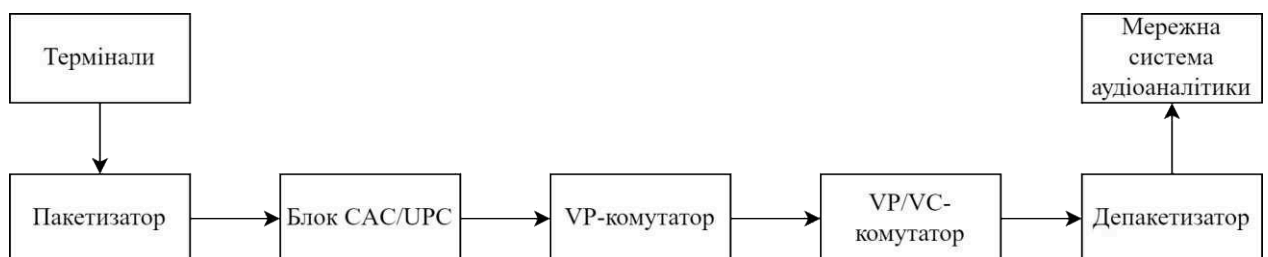


Рисунок 3.21 – Схема проходження аудіосигналу

Метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів:

1. Поділ аудіосигналу на осередки фіксованого розміру;
2. Оцінка можливості передачі аудіосигналу;
3. Визначення можливих віртуальних маршрутів для кожного осередка;
4. Передача осередків згідно з визначеними та скомутованими віртуальними маршрутами;
5. Моніторинг та управління трафіком;
6. Збір осередків та депакетизація;
7. Відновлення аудіосигналу для подальшого аналізу.

Термінали (джерело аудіо) слугують відправною точкою, де формується вхідний аудіосигнал. Потім дані надходять до пакетизатора, який виконує розбиває аудіопотік на логічні фрагменти з відповідними заголовками або службовою інформацією. Після цього блок SAC/UPC перевіряє, чи здатна мережа впоратися з новим потоком: якщо ресурси достатні, трафік приймається; якщо ні – відхиляється. На рівні VP-комутатора відбувається маршрутизація за віртуальними шляхами: так мережа може одночасно обслуговувати різні групи каналів у спільному тунелі. Потім дані проходять через інфраструктуру мережі аудіоаналітики, яка може містити декілька проміжних вузлів з аналогічною логікою. На приймальному боці віддалений VP/VC-комутатор остаточно спрямовує аудіопотік до блоку депакетизації, де зібрані пакети повертаються у безперервний потік і аналізуються терміналом-приймачем. Така послідовність дій дає змогу підтримувати низькі затримки, мінімізувати втрати та забезпечувати потрібну якість обслуговування аудіосигналу в режимі реального часу.

3.4 Метод аналізу аудіосигналів у мережних системах аудіоаналітики

Аналіз аудіосигналів у мережних системах аудіоаналітики передбачає два етапи. По-перше, здійснюється виявлення аудіоподій та їх попередня обробка. По-друге, виконується класифікація цих подій на основі сучасних алгоритмів глибокого навчання.

Перед проведенням класифікації аудіоподій необхідно виявити та провести попередню обробку. Попередня обробка аудіоданих є критичним етапом у процесі аналізу аудіо, оскільки вона готує сирі аудіодані до подальшої обробки та аналізу. Попередня обробка може включати фільтрацію шуму, вирівнювання, сегментацію та видалення відлуння. Особливості попередньої обробки описані в [89].

Аналіз аудіосигналів у мережних системах аудіоаналітики є важливою складовою, яка дозволяє виділяти та ідентифікувати звукові події в реальному часі. Це завдання потребує високої точності в умовах різноманітних фонів та можливих перешкод, адже аудіоподії можуть мати різні рівні інтенсивності, тривалості та характерні ознаки. Використання спеціалізованих алгоритмів і методів класифікації забезпечує здатність системи коректно обробляти складні звукові потоки, відкриваючи можливості для покращення безпеки, автоматизації та інших застосувань у різних сферах.

Якість розробленої моделі машинного або глибокого навчання значною мірою залежить від якості та релевантності набору даних, на якому вона була навчена. Методи машинного навчання розглядаються в [90], а їх застосування до аудіо в [91, 92]. Основні набори даних для навчання моделей поділу звуків і аналізу аудіосигналів наведені в [93-95].

У реальному світі аудіосигнали часто містять суміш різних звуків – мовлення, музики, фонового шуму тощо. Розділення звуків дозволяє ізолювати ці компоненти для більш точного аналізу. Особливості сучасних методів поділу звуків детально розглянуто в [96].

Завдяки щорічному конкурсу DCASE Challenge [97] методи аналізу аудіосигналів постійно вдосконалюються. DCASE Challenge пропонує уніфіковані критерії для оцінювання та порівняння різних методів розпізнавання аудіоподій. Кожного року перед дослідниками встають все складніші завдання стимулюючи розвиток більш передових технологій.

3.4.1 Показник якості аналізу аудіосигналу

Для оцінювання точності моделі використовується метрики F1-score та Event Error Rate. Ключовою перевагою цих метрик є їх здатність забезпечити детальний аналіз якості виявлення подій, дозволяючи розробникам краще розуміти та оптимізувати свої системи для конкретних вимог і сценаріїв застосування. F1-score розглядається як гармонічне середнє між точністю та повнотою. Для розрахунку цих значень використовуються істинно позитивні (TP), хибно позитивні (FP) та хибно негативні (FN) класифікації. Точність є мірою того, скільки з вибраних елементів є правильними. Значення точності розраховується як:

$$P = \frac{TP}{TP + FP}.$$

Повнота є мірою того, скільки з актуальних елементів було вибрано:

$$R = \frac{TP}{TP + FN}.$$

Значення F1-score розраховується як:

$$F1 = \frac{2PR}{P + R}.$$

Error Rate як метрика в задачах аналізу аудіосигналів використовує значення замін S , вставок I , видалень D та кількості подій у довідкових даних N для визначення ефективності моделі. Ця метрика особливо важлива у сценаріях, де точне розпізнавання та класифікація аудіоподій є критичними. Значення замін S – це кількість випадків, коли модель неправильно ідентифікувала одну подію як іншу. Значення вставок I являє собою помилкові спрацьовування, коли модель виявила подію, якої насправді не було. Значення видалень D – це кількість випадків, коли модель пропустила реальну подію. Довідкові події N дорівнює загальній кількості подій, які мали б бути виявлені, відповідно до довідкових даних. На основі цих значень обчислюється загальна помилка виявлення подій (EER), яка зазвичай виражається за допомогою наступної формули:

$$EER = \frac{S + D + I}{N}.$$

Ця метрика дозволяє оцінити загальну ефективність системи у виявленні та розпізнаванні аудіоподій, враховуючи всі типи помилок. Низьке значення EER вказує на високу точність та ефективність системи, тоді як високе значення EER свідчить про проблеми у виявленні подій.

3.4.2 Набір даних

FSD50K є великим набором даних, який використовується для завдань, пов'язаних із розпізнаванням звуків. Цей відкритий набір даних містить понад 51 000 аудіокліпів загальною тривалістю понад 100 годин, помічених вручну з використанням 200 класів з онтології AudioSet. Ліцензування аудіокліпів FSD50K під Creative Commons, включаючи хвили звуку, робить їх вільно розповсюджуваними та доступними. Ця відкрита доступність є значною перевагою порівняно з деякими іншими датасетами, роблячи його

незамінним ресурсом для дослідників та розробників по всьому світу. Звуки у FSD50K покривають широкий спектр категорій, включаючи природні звуки, музику, людські голоси, міські звуки та багато інших. Кожен звук у наборі даних супроводжується метаданими, які включають інформацію про джерело, ліцензію, теги, описи та інше.

Спочатку необхідно підготувати дані з FSD50K для використання у моделі. Це включає в себе вибір відповідних аудіозаписів із набору даних, які відповідають класам подій, які потрібно розпізнати. Для навчання методу розпізнавання аудіоподій було використано 15 класів звуків з гучними аудіоподіями типу Alarm, Gunshot, Gong та подібних. Загальна кількість датасету складає 11763 записів. Всі записи мають відповідне маркування та часові характеристики.

Поділ даних на тренувальну, валідаційну та тестову вибірки є важливою частиною процесу підготовки даних для задач машинного навчання. Тренувальна вибірка є основним набором даних, який використовується для тренування моделі. Тренувальна вибірка повинна бути репрезентативною і включати широкий спектр можливих випадків, які модель може зустріти у реальному світі.

Валідаційна вибірка використовується для тонкого налаштування параметрів моделі та для перевірки її продуктивності під час тренування. Валідаційна вибірка допомагає визначити, коли модель починає перенавчатися на тренувальних даних.

Тестова вибірка є незалежним набором даних, який використовується для оцінки кінцевої продуктивності моделі після тренування та налаштування. Ця вибірка дозволяє оцінити, як модель буде працювати на нових, невідомих даних, симулюючи реальне використання. Тестова вибірка повинна бути абсолютно незалежною від тренувальної та валідаційної вибірок.

Загальний датасет було поділено на тренувальну, валідаційну та тестову вибірки у співвідношенні 70-20-10. Додатково, кожна з вибірок було

об'єднано у один аудіофайл зі зберіганням маркувань, для моделювання роботи у режимі реального часу.

3.4.3 Метод аналізу аудіосигналу з використанням багатоподійного класифікатора

Узагальнена схема методу, що пропонується наведена на рис. 3.22.

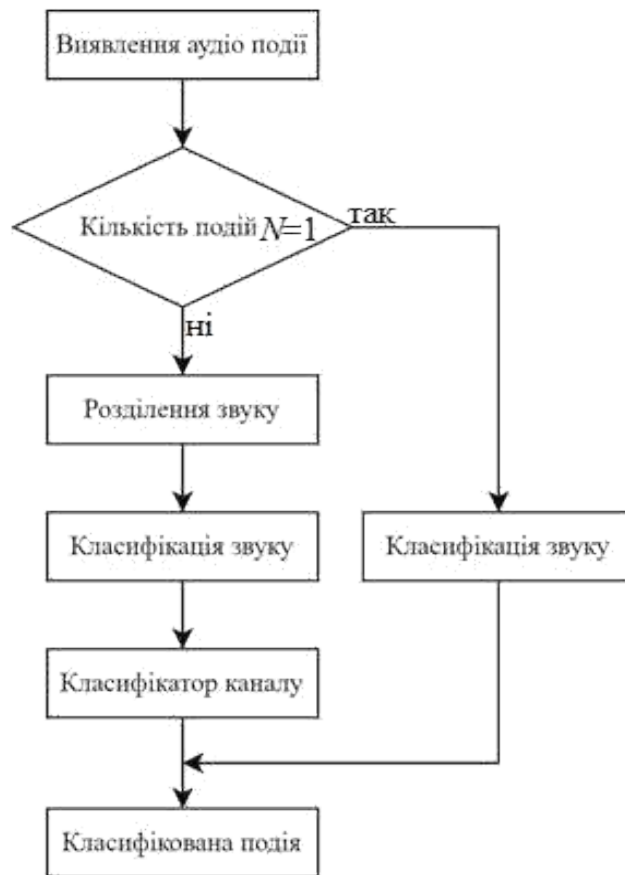


Рисунок 3.22 – Узагальнена схема методу

Завдання поділу звуків полягає в отриманні аудіодоріжок окремих джерел звуку по одному аудіозапису з кількома джерелами. У загальному випадку, її можна представити наступним чином:

$$a_{mixed} = \sum_{i=0}^N a_i .$$

Задача поділу звуків поляє в визначенні компонент a_i основного аудіофайлу a_{mixed} . Для рішення цієї задачі, аудіодані представляються у вигляді спектрограми. Після визначення аномалії у вхідному потоку аудіоданих за допомогою порогового значення шуму та фільтрації, аудіодані буферизуються та подаються до спеціального багатоподійного класифікатора, який визначає кількість одночасних аудіоподій. Якщо кількість аудіоподій N дорівнює 1, то вона одразу подається до класифікатора аудіоподій, інакше, через операцію розділення звуку визначаються декілька одночасних аудіоподій та отримуються їх відповідні характеристики. Після цього, отримані аудіодані проходять через класифікатор звуків, і результатом є N аудіоподій та ще одна від нерозподіленого запису. Окрема модель під назвою класифікатор каналу отримує всі аудіодані та вибирає основну аудіоподію.

На етапі багатоподійного класифікатора важливо якнайшвидше визначити, якою з гілок піде алгоритм. У випадку поліфонічної аудіоподії цей етап може суттєво збільшити загальний час обробки аудіосигналу. Виходячи з цього, модель має бути досить маленькою. Як багатоподійний класифікатор використовувалася модифікація ResNet18 на спектрограмах з кількістю входів 1 та трьома класами на виході, відповідно до кількості можливих одночасних аудіоподій. Випадки, коли $N > 3$ рідкісні та вимагаються складних рішень для своєї обробки.

На етапі розділення звуку використовується метод розділення звуку за допомогою конволюційної нейронної мережі (CNN) та використання бінарних масок. Навчання CNN відбувається на основі спектрограм. Замість безпосередньої реконструкції звуків, CNN використовується для створення бінарних масок. Бінарна маска визначає присутність або відсутність конкретного джерела звуку в кожному часовому кадрі та частотному діапазоні спектрограми. Отримана бінарна маска застосовується до вхідної спектрограми для виділення інтенсивності тих частот, які відповідають цільовому джерелу звуку. Це дозволяє відділити цільовий звук від решти

міксу. CNN складається з двох блоків. Кожен блок містить два конволюційні шари. Конволюційні шари використовуються для виявлення локальних характеристик у вхідних даних, таких як краї, текстури або специфічні шаблони в спектрограмі. Використання декількох конволюційних шарів дозволяє мережі вчитися складнішим та абстрактнішим характеристикам даних. Кожен блок має один шар макс-пулінгу. Цей шар зменшує розмірність даних, вибираючи найбільш значущі елементи у вікні пулінгу. Це допомагає зменшити обчислювальну складність та контролювати перенавчання. Після кожного блоку макс-пулінгу застосовується дропаут. Це робить мережу менш чутливою до специфічних характеристик тренувальних даних, тим самим запобігаючи перенавчанню. Після пропускання через конволюційні та макс-пулінг шари, дані перетворюються в одновимірний формат та подаються до повнозв'язного шару. Цей шар інтегрує інформацію з усіх попередньо вилучених ознак, дозволяючи моделі вчинити комплексні висновки на основі загальної інформації. Використання ще одного шару дропауту зі значенням 0.2 перед останнім повнозв'язним шаром додатково знижує ризик перенавчання. Останній повнозв'язний шар генерує вектор значень, який відповідає кількості частотних кошиків у типовій спектрограмі. Вектор представляє бінарні маски для кожного частотного компонента в розрізі часу.

Для задачі класифікації звуків використовуються базова комбінаційна модель на основі ResNet та RNN. ResNet адаптована від ResNet моделі для вирішення задачі розпізнавання зображення. Комбінування ResNet та RNN дозволяє створити модель, яка спроможна працювати з аудіо даними, здійснюючи при цьому як просторовий аналіз звуку (за допомогою ResNet), так і аналіз залежностей в часі (за допомогою RNN). Це дозволяє досягнути високої точності в задачах класифікації звуків. Спочатку використовується ResNet для аналізу звукових спектрів, а потім отримані ознаки передаються до RNN для аналізу послідовностей в часі та вирішення конкретної задачі класифікації.

Класифікатор каналу виконує задачу класифікації над всіма отриманими міксами. На цьому етапі доступні всі можливі ознаки аудіосигналу, бінарні маски та результат класифікації.

Для порівняння розробленого методу з існуючими, навчання виконувалось на наборі даних з DCASE 2023 завдання 4. У цьому завданні використовувався власний набір даних з слабо маркованими даними.

Після побудови, навчання та тестування моделі були отримані результати F1-score та EER, що дорівнюють 49,6% та 0.44 відповідно (табл. 3.2). Ці результати є корашенням у порівнянні, наприклад, з базовим у DCASE challenge task 3, де були отримані 48.7% та 0.54 відповідно.

Таблиця 3.2 – Результати F1-score та EER

Метод	F1-score	EER
The nerc-slip system for sound event localization and detection	62.7%	0.33
Attention mechanism network and data augmentation	58.5%	0.35
The distillation system based on ResNet-Conformer model guided by a ResNet-GRU	51.4%	0.40
Запропонований	49,6%	0.44
Baseline	48.7%	0.54
One audio augmentation chain proposed for sound event localization and detection	45.0%	0.48
A framework for SELD using conformer and multi-ACCDOA strategies	33.1%	0.56
Based on omi-dimensional dynamic convolution and feature pyramid attention module	22.1%	0.64

Основна перевага розробленого методу – швидкий аналіз монофонічних аудіоподій через використання простого багатоподійного класифікатора. При тестуванні методу, максимальний час, що був витрачений на аналіз дорівнює 55 мс, а середній 50 мс. Також, якість аналізу повинна бути стійкою до шуму та інших перешкод, через наявність блоку sound separator. Забезпечення стабільної роботи системи в умовах шуму допомагає підтримувати високу якість обробки аудіоданих в реальних

умовах. Особливу складність для системи складають велика кількість одночасних аудіоподій та класифікація гучних звуків з високою тональністю. І навпаки, система показує хороші результати при класифікацію гучних басів.

Результати дослідження вказують на потенціал такого підходу у вирішенні задачі обробки аудіосигналів у мережній системі аудіоаналітики. Подальші дослідження та розвиток цього напрямку можуть призвести до покращення та оптимізації результатів у майбутньому. У майбутніх дослідженнях можливо розглянути розширення та оптимізацію використаних нейронних мереж та класифікаторів. У поточній роботі використовувалася відносно проста модель з обмеженою кількістю шарів та параметрів для компактності і ефективності тренування. Збільшення глибини та додавання додаткових шарів може допомогти нейронній мережі виявляти ще більше важливих особливостей у вихідних даних.

Структурна схема запропонованого методу наведена на рис. 3.23.

На схемі процес розпочинається з виявлення аудіоподії у вхідному сигналі. Після цього відбувається оцінка кількості одночасних подій: якщо визначено одну подію ($N=1$), сигнал одразу спрямовується на етап класифікації. Коли ж подій виявляється декілька, проводиться розділення сигналу на складові частини, кожна з яких класифікується окремо. Потім усі проаналізовані події надходять до класифікатора каналу, який визначає основну або найважливішу з них. Результатом у будь-якому випадку є кінцева класифікована подія, що відображає успішне розпізнавання та ідентифікацію звуку.

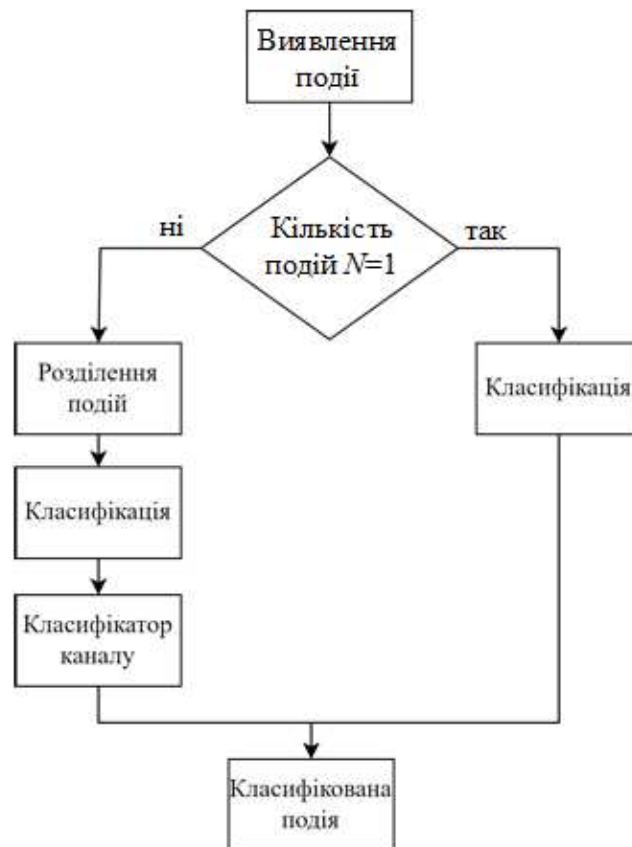


Рисунок 3.23 – Структурна схема запропонованого методу

Висновки за розділом 3

У цьому розділі було детально розглянуто основні методи та підходи, які забезпечують ефективну передачу та аналіз аудіосигналів у мережних системах аудіоаналітики. Результати досліджень підкреслюють важливість оптимізації мережних параметрів для забезпечення мінімальних затримок, стійкості до шуму, а також надійного аналізу аудіосигналів, що є критичними для роботи в режимі реального часу. Ефективність методу обробки аудіосигналів у мережній системі аудіоаналітики також є важливим аспектом. Метод повинен забезпечувати високу точність в розпізнаванні подій та бути стійким до шумів і здатним працювати в різних середовищах та з різними видами аудіосигналів.

Отже, у цьому розділі отримані такі результати:

1. Виконано моделювання джерел навантаження системи;

2. Розглянуто управління з'єднаннями та трафіком у мережній системі аудіоаналітики;
3. Досліджено методи забезпечення якості передачі аудіосигналу;
4. Запропоновано метод передачі аудіосигналів за рахунок комутації віртуальних маршрутів;
5. Обґрунтовано показник якості передачі та аналізу аудіосигналу;
6. Досліджено набори даних для навчання, валідації та тестування методів розпізнавання аудіоподій у мережних системах аудіоаналітики;
7. Запропоновано метод аналізу аудіосигналів у мережних системах аудіоаналітики.

Запропоновані методи передачі та аналізу аудіосигналів у мережних системах аудіоаналітики здатні значно покращити ефективність і надійність таких систем. Завдяки розробленим підходам забезпечується висока швидкість обробки аудіосигналів, що робить можливим їхнє застосування в задачах, де критично важливими є точність, швидкодія та стійкість до перешкод.

Перелік використаних джерел у даному розділі наведено у повному переліку використаних джерел під номерами: 71–98.

4 ДОСЛІДЖЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ОБРОБКИ АУДІОСИГНАЛІВ У МЕРЕЖНИХ СИСТЕМАХ АУДІОАНАЛІТИКИ

У цьому розділі представлено результати експериментальних досліджень розробки інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики. Проведено детальний аналіз отриманих метрик продуктивності, включаючи оцінювання швидкості проходження сигналу та оцінювання точності розпізнавання.

4.1 Розробка інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики

Інформаційні технології охоплюють різні методи, засоби та процеси, які використовуються для створення, збирання, обробки, зберігання, захисту та передачі інформації. Це включає як апаратне, так і програмне забезпечення, мережну інфраструктуру, бази даних та інформаційні системи. Інформаційні технології є основою сучасного бізнесу, оскільки дозволяють автоматизувати багато процесів, зменшувати витрати, підвищувати ефективність та покращувати комунікацію.

Інформаційна технологія обробки аудіосигналів об'єднує апаратне та програмне забезпечення для збору, аналізу та інтерпретації звукових даних у мережному середовищі. Це дозволяє вирішувати різні завдання у сфері безпеки, моніторингу, медицини та інших галузях. Мережна система аудіоаналітики включає різноманітні компоненти. Перш за все, це аудіосенсори, такі як мікрофони, що розташовуються в різних місцях для збору звукових даних. Дані з цих сенсорів передаються через мережне обладнання, яке може включати комутатори, маршрутизатори та бездротові технології. Зібрані аудіодані обробляються на серверах за допомогою спеціалізованого програмного забезпечення, яке може виконувати різні види

аналізу: розпізнавання мови, виявлення аномалій, класифікацію звуків. Результати цього аналізу зберігаються у базах даних та можуть бути візуалізовані через веб-інтерфейси або мобільні додатки.

Перед тим як приступати до проектування інформаційної технології обробки аудіосигналів, важливо розібратись з розташуванням засобів аналітики та топологією мережі. Розташування аналітичних засобів може мати значний вплив на ефективність їх використання та можливість отримання корисної інформації. Топологія мережі визначає, як компоненти системи підключені один до одного та як дані передаються в мережі.

Основними компонентами інформаційної технології обробки аудіосигналів є засоби передачі інформації та засоби обробки інформації. Засоби передачі інформації забезпечують передачу аудіоданих через мережі з одного пристрою на інший. Вони можуть включати мережне обладнання, таке як маршрутизатори, комутатори, а також протоколи передачі даних, а також спеціалізовані протоколи для передачі аудіоданих. Засоби обробки інформації відповідають за обробку отриманих аудіоданих. Вони можуть включати програмне забезпечення для обробки сигналів, алгоритми аналізу аудіо, системи штучного інтелекту для розпізнавання мови або аналізу аудіоданих, бази даних для зберігання оброблених даних та засоби візуалізації для представлення результатів аналізу. Розташування систем аудіоаналітики можна реалізувати за трьома основними підходами: на периферії, на локальному сервері та в хмарі [98-100].

Аудіоаналітика на периферії передбачає виконання всіх аналітичних дій на периферійних пристроях, таких як мікрофони або камери. Цей підхід дозволяє зменшити навантаження на мережу, оскільки дані не потребують передачі до центрального сервера. Перевагою є те, що цифрові аудіодані не залишають периферію, що підвищує конфіденційність. Однак, цей підхід потребує значних витрат на встановлення обчислювальних модулів на кожному периферійному пристрої, що може бути складним у реалізації.

Локальний сервер використовується для обробки цифрового аудіосигналу, який надходить з периферійних пристроїв. Цей підхід дозволяє централізовано керувати обробкою даних та забезпечує високу швидкість передачі інформації в межах локальної мережі. Локальні сервери зазвичай є частиною закритої системи, що сприяє дотриманню конфіденційності. Однак, витрати на обладнання та обслуговування локальних серверів можуть бути значними. Крім того, необхідно забезпечити дотримання всіх нормативних вимог щодо зберігання та передачі персональних даних.

Хмарна аналітика передбачає використання хмарних серверів для обробки та зберігання даних. Цей підхід забезпечує гнучке масштабування та зниження витрат на обладнання, оскільки обчислювальні ресурси надаються як послуга. Хмарні рішення знімають необхідність у постійному технічному обслуговуванні обладнання, що спрощує його експлуатацію. Проте, для використання хмарних серверів необхідне надійне інтернет-з'єднання, і питання безпеки даних стає критично важливим, оскільки відповідальність за збереження даних розподіляється між користувачем і постачальником послуг. Також потрібно враховувати законодавчі вимоги щодо передачі та зберігання даних у хмарі. Контекстна діаграма інформаційної технології обробки аудіосигналів представлена на рис. 4.1.



Рисунок 4.1 – Контекстна діаграма інформаційної технології обробки аудіосигналів

У контекстній діаграмі інформаційної технології обробки аудіосигналів можна виділити кілька ключових параметрів: вхідні дані, вихідні дані, параметри управління та механізми.

Вхідні дані включають аудіосигнали, які є первинними даними, параметри налаштувань, що визначають конфігурацію системи, такі як частотні діапазони, рівні чутливості та алгоритми обробки, а також метадані, які забезпечують додаткову контекстуальну інформацію для підтримки аналізу та зберігання даних.

Вихідні дані представлені обробленими даними, що включають результати аналізу аудіосигналів, а також звіти та сповіщення, які інформують користувачів про виявлені події та результати аудіоаналізу.

Управління здійснюється через політики безпеки, що регулюють доступ до даних і забезпечують їх захист, методи передачі аудіоданих, які встановлюють стандарти для надійної та ефективної передачі аудіоінформації через мережу, а також методи обробки аудіоданих, такі як шумозаглушення, нормалізація звуку, виділення ознак, класифікація звуків та розпізнавання мови.

Механізми включають апаратні засоби, такі як мікрофони, сервери та мережне обладнання, що забезпечують фізичну інфраструктуру системи, та програмне забезпечення, яке реалізує методи передачі, аналізу та візуалізації.

Можна провести декомпозицію контекстної діаграми на більш дрібні, керовані частини або підфункції. Це дозволить проаналізувати кожну частину системи окремо, що сприяє більш ефективному управлінню та розробці системи. За допомогою декомпозиції функцій можна ідентифікувати окремі завдання, їхні входи, виходи та взаємодії між ними, що полегшує виявлення проблем та оптимізацію роботи системи. На верхньому рівні можна виділити загальні функції, такі як збір, передача, обробка, зберігання та візуалізація аудіоданих. Декомпозиція функцій інформаційної технології обробки аудіосигналів наведена на рис. 4.2.

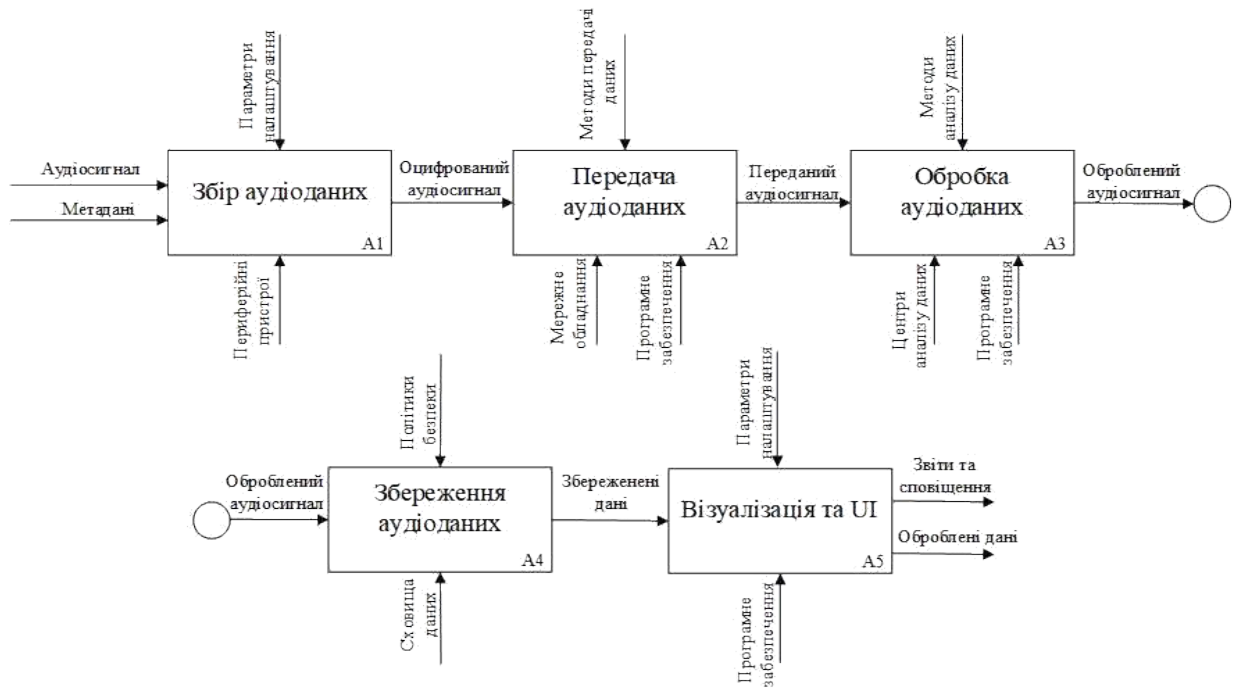


Рисунок 4.2 – Декомпозиція функцій інформаційної технології обробки аудіосигналів

Збір аудіоданих виконується за допомогою мікрофонів, що розташовані в зоні моніторингу. Ці пристрої фіксують акустичні сигнали з навколишнього середовища. Система налаштовується на певний рівень чутливості та частотний діапазон, щоб забезпечити адекватний запис звуків. Параметри налаштувань, такі як чутливість мікрофонів та частотні діапазони, контролюють, як саме здійснюється цей збір. На виході процесу збору аудіоданих утворюються оцифровані аудіосигнали, готові для подальшої передачі.

Передача аудіоданих відбувається через мережу до центральних серверів для подальшої обробки. Аудіодані оцифровуються та упорядковуються в пакети для ефективної передачі. Цей процес керується методами передачі даних, які забезпечують надійну та ефективну передачу даних через мережу. На виході процесу передачі отримуються передані аудіодані, готові для обробки на центральних серверах. Обробка аудіоданих включає застосування різних алгоритмів для аналізу отриманих звукових сигналів. Алгоритми обробки можуть включати шумозаглушення,

нормалізацію звуку, виділення ознак, класифікацію звуків та розпізнавання мови. Параметри обробки визначають, як ці алгоритми застосовуються до аудіоданих. На виході цього процесу отримуються оброблені дані, які містять результати аналізу, такі як виявлені події та розпізнані звуки.

Зберігання даних здійснюється в базах даних або хмарних сховищах. Оброблені дані зберігаються разом із метаданими, що забезпечує можливість подальшого доступу та аналізу. Процес зберігання даних керується політиками безпеки, що включають шифрування даних, контроль доступу та забезпечення цілісності даних. На виході процесу зберігання отримуються збережені дані, готові для використання в звітах або для подальшого аналізу.

Візуалізація та інтерфейси користувача забезпечують доступ до результатів аналізу для кінцевих користувачів. Оброблені дані та звіти відображаються через графічні інтерфейси, панелі управління та інші засоби візуалізації. Параметри візуалізації визначають, як саме результати аналізу будуть представлені користувачам. На виході цього процесу користувачі отримують звіти та сповіщення, що інформують їх про результати аналізу та виявлені події, дозволяючи їм приймати відповідні заходи.

Кожна з цих функцій може бути деталізована на нижчих рівнях. Декомпозицію низького рівня функцій інформаційної технології обробки аудіосигналів представлено на рис. 4.3.

Збір даних можна поділити на функції налаштування мікрофонів, запису звуків, фільтрації шуму та виявлення аудіоподій. Налаштування мікрофонів включає калібрування чутливості мікрофонів і встановлення частотного діапазону для запису звуків. Це необхідно для забезпечення точного і чіткого захоплення звукових сигналів. Запис звуків здійснюється через оцифрування звукових сигналів, що надходять від мікрофонів, і їх перетворення у цифровий формат. Це дозволяє зберегти звукові дані у вигляді, придатному для подальшої обробки. Фільтрація шуму займається видаленням небажаних шумів і перешкод з записаних аудіоданих, що покращує їх якість і полегшує подальший аналіз. Виявлення аудіоподій

аналізує аудіодані для виявлення значущих подій, таких як голоси, сигнали тривоги, розбиття скла та інші звукові аномалії.



Рисунок 4.3 – Декомпозиція низького рівня функцій інформаційної технології обробки аудіосигналів

Передачу даних можна поділити на пакетизацію даних, передачу пакетів та забезпечення необхідного рівня QoS. Пакетизація даних передбачає розбивку аудіосигналів на окремі пакети для ефективної передачі через мережу. Це дозволяє оптимізувати використання мережних ресурсів і забезпечити стійкість до втрат даних. Передача пакетів включає використання мережних протоколів для забезпечення надійності та ефективності передачі аудіоданих. Вибір протоколу залежить від вимог до швидкості передачі та стійкості до помилок. Забезпечення QoS включає управління пріоритетами та ресурсами мережі для забезпечення надійної і якісної передачі аудіоданих.

Зберігання даних можна поділити на збереження аудіоданих, збереження метаданих та забезпечення безпеки даних. Збереження даних включає збереження оброблених аудіоданих у базах даних або хмарних сховищах. Це забезпечує доступність даних для подальшого аналізу та використання. Збереження метаданих передбачає запис додаткової інформації про аудіодані, такої як час запису, місце розташування та інші релевантні характеристики, для забезпечення можливості подальшого доступу та аналізу даних. Безпека даних включає заходи для захисту збережених даних, такі як шифрування, контроль доступу та забезпечення цілісності даних. Це важливо для запобігання несанкціонованому доступу та втраті даних.

Візуалізацію та інтерфейси користувача можна поділити на налаштування параметрів візуалізації, відображення даних та сповіщення користувачів. Налаштування параметрів візуалізації включає визначення, як саме результати аналізу будуть представлені користувачам. Це може включати вибір форматів графіків, таблиць, кольорових схем та інших параметрів візуалізації. Відображення даних передбачає представлення оброблених аудіоданих та звітів через графічні інтерфейси, панелі управління та інші засоби візуалізації. Це дозволяє користувачам легко зрозуміти та інтерпретувати результати аналізу.

4.2 Програмна реалізація інформаційної технології обробки аудіосигналів

Розвиток сучасних інформаційних технологій та зростання обсягів аудіоданих обумовлюють необхідність створення мережних систем аудіоаналітики, здатних ефективно обробляти та аналізувати звукові сигнали в режимі реального часу. Такі системи мають забезпечувати високу точність розпізнавання аудіоподій, мінімізуючи затримки при передачі та обробці даних. Реалізація інформаційної технології обробки аудіосигналів включає

інтеграцію компонентів збору, передачі, попередньої обробки сигналів та застосування моделей машинного навчання для класифікації звукових подій. Особлива увага приділяється оптимізації архітектури системи та алгоритмів обробки для підвищення продуктивності та надійності в умовах реальних мережних середовищ.

4.2.1 Вибір технологічного стеку та архітектури

Під час реалізації інформаційної технології обробки аудіосигналів надзвичайно важливо правильно визначити набір інструментів і платформ, які гарантуватимуть високу ефективність, можливість масштабування, належний рівень безпеки та дотримання сучасних стандартів. Від того, які саме технології обирає розробник, залежать всі етапи обробки, від збору та передавання аудіоданих до їхньої подальшого аналізу, зберігання та представлення результатів.

Вибір технічних рішень зазвичай зумовлений кількома ключовими чинниками. Насамперед ідеться про продуктивність та масштабованість, адже система має обробляти великі обсяги аудіоданих у режимі реального часу, реагуючи на збільшення навантаження завдяки технологіям контейнеризації та оркестрації. Безпека не менш важлива, оскільки в аудіозаписах може міститися конфіденційна інформація. Використання сучасних криптографічних протоколів і рішень для моніторингу підвищує захищеність даних на всіх рівнях.

Крім того, розробники часто обирають мікросервісну архітектуру, а також популярні фреймворки, що робить систему гнучкішою та значно спрощує інтеграцію з різними зовнішніми модулями або сервісами. Застосування технологій з активною спільнотою фахівців сприяє вчасному впровадженню оновлень і швидкому розв'язанню потенційних проблем.

Основною мовою програмування для серверної частини та обробки аудіоданих було обрано Python [101]. Цей вибір зумовлений широкими

можливостями Python у сфері наукових обчислень, обробки сигналів та машинного навчання. Мова має велику кількість бібліотек та фреймворків, таких як TensorFlow, PyTorch, Keras для машинного навчання, та LibROSA, PyAudio для обробки аудіосигналів. Python також відзначається простотою синтаксису та активною спільнотою розробників, що полегшує процес розробки та підтримки коду.

Для реалізації веб-сервісів та RESTful API було використано фреймворк Django [102]. Django надає потужні інструменти для швидкої розробки веб-додатків, забезпечує високий рівень безпеки та підтримує масштабованість. Фреймворк має вбудовані засоби для роботи з базами даних, автентифікації користувачів, управління сесіями та інші важливі компоненти веб-додатків.

У ролі бази даних для зберігання структурованих даних та метаданих було обрано PostgreSQL [103]. Ця система керування базами даних відома своєю надійністю, продуктивністю та відповідністю стандартам SQL. PostgreSQL підтримує складні запити, транзакції та забезпечує високий рівень безпеки даних, що є критично важливим для системи, яка обробляє конфіденційну інформацію.

Для забезпечення безпечної передачі аудіоданих було обрано протоколи SRTP та TLS. SRTP забезпечує шифрування, автентифікацію та цілісність даних, що передаються в реальному часі, що важливо для передачі аудіосигналів. TLS використовується для захисту даних під час передачі між клієнтом і сервером, запобігаючи атакам типу «людина посередині».

Для створення захищених каналів передачі даних між периферійними пристроями та центральним сервером було використано OpenVPN [104]. Це рішення дозволяє ізолювати мережний трафік та запобігти несанкціонованому доступу ззовні, що підвищує загальний рівень безпеки системи.

У сфері обробки аудіосигналів та реалізації моделей машинного навчання було використано бібліотеки та фреймворки, такі як TensorFlow та

PyTorch. Вони надають потужні засоби для розробки, навчання та розгортання нейронних мереж, що дозволяє ефективно вирішувати задачі розпізнавання аудіоподій. Бібліотека LibROSA використовується для аналізу та обробки аудіо, надаючи широкий спектр інструментів для роботи зі спектрограмами, мел-частотними коефіцієнтами та іншими ознаками, необхідними для аудіоаналітики.

Для розробки інтерфейсу користувача було обрано JavaScript з використанням фреймворку React [105]. React дозволяє створювати динамічні та інтерактивні інтерфейси користувача, що підвищує зручність та ефективність роботи з системою. Використання компонентного підходу в React сприяє повторному використанню коду та спрощує підтримку додатку. Обраний технологічний стек забезпечує необхідний рівень продуктивності, безпеки та масштабованості, що дозволяє системі ефективно функціонувати у різних умовах.

Архітектурна діаграма системи наведена на рис. 4.4. Архітектура системи побудована на мікросервісному підході, що дозволяє розділити систему на незалежні компоненти, які можуть розвиватися та масштабуватися окремо. Основними компонентами системи є периферійні пристрої збору даних, мережний шар, сервер обробки даних, база даних, інтерфейс користувача та система безпеки.

Периферійні пристрої збору даних – це мікрофони з вбудованими модулями обробки, які виконують попереднє фільтрування шуму та пакетизацію даних. Комунікаційний шар включає мережне обладнання та протоколи для передачі даних до центрального сервера з використанням шифрування SRTP та TLS. Сервер обробки даних відповідає за аналіз аудіоданих з використанням моделей машинного навчання, розроблених у попередніх розділах. База даних слугує сховищем для оброблених даних, результатів аналізу та метаданих, забезпечуючи надійне зберігання та швидкий доступ до інформації. Інтерфейс користувача реалізований як веб-додаток, який надає доступ до результатів аналізу, звітів та налаштувань

системи. Система безпеки забезпечує автентифікацію, авторизацію, шифрування даних та моніторинг безпеки для захисту системи від потенційних загроз.



Рисунок 4.4 – Архітектурна діаграма системи

4.2.2 Реалізація основних компонентів

Реалізація інформаційної технології обробки аудіосигналів включає кілька ключових компонентів, які тісно взаємодіють між собою для забезпечення ефективного збору, передачі, обробки та зберігання аудіоданих. Кожен з цих компонентів відіграє критичну роль у загальній функціональності системи, і їхня реалізація вимагає детального підходу до технічних і безпекових аспектів.

На етапі збору даних використовуються периферійні пристрої, обладнані високочутливими мікрофонами та мікроконтролерами. Ці пристрої виконують початкову обробку аудіосигналу, що включає оцифрування сигналу з частотою дискретизації 44,1 кГц, фільтрацію шуму та нормалізацію рівня сигналу. Для підвищення якості запису та зменшення впливу навколишнього шуму використовуються алгоритми цифрової обробки сигналів, такі як адаптивне фільтрування та спектральне відсіювання.

Після початкової обробки аудіодані пакетизуються згідно з протоколом SRTP, що забезпечує шифрування та цілісність даних. Кожен пакет містить

метадані, такі як часові мітки та ідентифікатори пристроїв, що полегшує синхронізацію та подальший аналіз даних на сервері.

Передача даних від периферійних пристроїв до центрального сервера здійснюється через захищене з'єднання з використанням VPN, створеної на базі OpenVPN. Це забезпечує додатковий рівень безпеки, захищаючи дані від перехоплення та несанкціонованого доступу під час передачі через загальнодоступні мережі.

Центральний сервер відповідає за глибоку обробку аудіоданих, отриманих від периферійних пристроїв. Серверна інфраструктура побудована на базі високопродуктивних обчислювальних вузлів з підтримкою графічних процесорів, що дозволяє прискорити виконання алгоритмів машинного навчання.

На сервері розгорнуто програмне забезпечення, розроблене на мові Python з використанням фреймворків TensorFlow та PyTorch. Основою системи розпізнавання є метод, описаний у розділі 3.4. Перед подачею на вхід нейронної мережі аудіодані перетворюються у спектрограми з використанням швидкого перетворення Фур'є. Спектрограми нормалізуються та масштабуються для покращення стабільності моделі. Для підвищення стійкості до шумів та різноманітних акустичних умов застосовуються методи аугментації даних. Результати обробки, а також сирі аудіодані та метадані зберігаються у базі даних PostgreSQL. Структура бази даних розроблена з урахуванням вимог нормалізації та оптимізації запитів. Дані розділені на логічні таблиці, що містять інформацію про аудіоподії, пристрої збору даних, користувачів системи та інші сутності. Загальний алгоритм обробки аудіоданих наведено на рис. 4.5.



Рисунок 4.5 – Алгоритм обробки аудіоданих

Загальна схема передачі даних наведена на рис. 4.6.

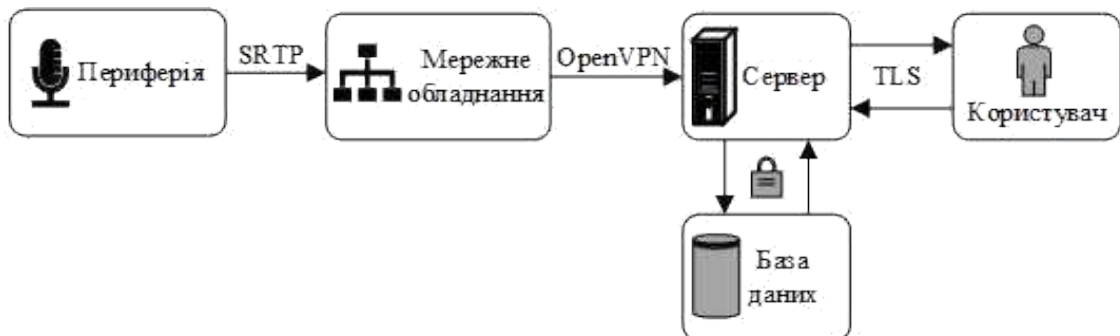


Рисунок 4.6 – Схема передачі даних

Для відповідності вимогам GDPR та інших законодавчих актів щодо захисту персональних даних впроваджено механізми видалення та анонімізації даних за запитом користувачів. Ведеться журнал подій, що фіксує всі дії з даними, що дозволяє проводити аудит та виявляти можливі порушення безпеки.

Результати обробки, а також сирі аудіодані та метадані зберігаються у базі даних PostgreSQL. Структура бази даних розроблена з урахуванням вимог нормалізації та оптимізації запитів. Дані розділені на логічні таблиці, що містять інформацію про аудіоподії, пристрої збору даних, користувачів системи та інші сутності.

Для забезпечення безпеки даних застосовуються механізми шифрування як на рівні передачі (з використанням TLS), так і на рівні зберігання. Шифрування даних у стані спокою реалізовано за допомогою функціональності Transparent Data Encryption, яка шифрує файли бази даних на рівні файлової системи. Це запобігає доступу до даних у випадку компрометації фізичних носіїв або резервних копій.

Контроль доступу до даних здійснюється за допомогою ролей та прав доступу, налаштованих у PostgreSQL. Кожному користувачу системи призначається роль з певними привілеями, що обмежує його дії в системі відповідно до принципу найменших привілеїв. Аутентифікація користувачів реалізована з використанням протоколу OAuth 2.0, що забезпечує безпечний та стандартизований механізм управління сесіями та токенами доступу.

Інтерфейс користувача розроблено як веб-додаток, що забезпечує доступність з різних пристроїв, включаючи комп'ютери, планшети та смартфони. Для фронтенду використано фреймворк React, який дозволяє створювати динамічні та інтерактивні користувацькі інтерфейси. Використання компонентного підходу в React сприяє підтримованості коду та його повторному використанню. Дизайн інтерфейсу орієнтований на інтуїтивність та зручність використання. Користувачі можуть переглядати результати аналізу аудіоданих у вигляді графіків, таблиць та інтерактивних

візуалізацій. Реалізовано систему сповіщень, яка інформує користувачів про виявлені події, такі як аномальні звуки або сигнали тривоги.

Для бекенду веб-додатку використовується Django REST Framework, що надає API для взаємодії з фронтендом. API реалізує всі необхідні операції, включаючи отримання даних, створення нових записів, оновлення та видалення. Комунікація між фронтендом та бекендом відбувається через захищений протокол HTTPS з використанням TLS.

Крім того, впроваджено журналювання всіх важливих подій у системі, включаючи дії користувачів, системні помилки та інші критичні події. Журнали зберігаються у централізованому сховищі та можуть бути використані для аналізу інцидентів безпеки та оптимізації роботи системи.

4.3 Оцінка ефективності запропонованої інформаційної технології

Оцінка ефективності запропонованої інформаційної технології є ключовим етапом у підтвердженні її доцільності та конкурентоспроможності. У цьому розділі проведено детальний аналіз системи за кількома критеріями: час проходження сигналу та точність розпізнавання. Цей аналіз дозволяє виявити переваги та недоліки технології, а також визначити напрями для подальшого вдосконалення.

4.3.1 Оцінка за часом проходження сигналу

Часом проходження сигналу вважається проміжок від моменту фіксації звуку мікрофоном до формування остаточного рішення системи щодо розпізнавання аудіоподій. Метою дослідження є визначення середніх, максимальних та мінімальних затримок на всіх етапах, а також оцінка придатності системи до роботи в режимі, наближеному до реального часу.

Завдяки попередній обробці на периферійних пристроях та ефективній пакетизації даних, затримки на етапі збору мінімізуються. Використання

протоколів SRTP та оптимізованих мережних шляхів дозволяє забезпечити низьку затримку при передачі даних. Шифрування даних незначно впливає на загальний час передачі завдяки використанню ефективних алгоритмів шифрування. Для проведення дослідження було виділено шість основних етапів обробки аудіосигналу:

1. Збір даних – фіксація звуку;
2. Оцифрування – перетворення аналогового сигналу в цифрову форму;
3. Передача даних – відправка оцифрованого сигналу через мережу до сервера з використанням удосконаленого методу передачі аудіосигналів;
4. Підготовка даних – нормалізація рівнів, видалення шуму, формування набору ознак;
5. Обробка моделлю – застосування методу аналізу аудіосигналу;
6. Пост-обробка та ухвалення рішення – формування кінцевого висновку про аудіоподію та запису до бази даних.

На кожному з етапів система фіксувала проміжок часу за допомогою вбудованих таймерів. Експерименти проводилися в умовах локальної мережі зі швидкістю передачі даних 1 Гбіт/с, при середньому навантаженні на канал. Загальна значення затримки кожного етапу наведені у табл. 4.1.

Таблиця 4.1 – Значення затримок кожного етапу

Етап обробки	Середня затримка, мс	Максимальна затримка, мс	Мінімальна затримка, мс
Збір даних	4	5	3
Оцифрування	2	3	2
Передача даних	28	45	15
Підготовка даних	14	17	10
Обробка моделлю	45	58	36
Пост-обробка та ухвалення рішення	11	13	8
Загальний час затримки	104	141	74

Оптимізація моделей машинного навчання, включаючи використання графічних процесорів, дозволяє знизити час обробки аудіофрагментів. Середній час обробки одного аудіофрагмента становить близько 104 мс, що відповідає вимогам режиму реального часу. Це досягається завдяки використанню ефективних алгоритмів. Загальний час проходження сигналу від моменту його фіксації до отримання результату не перевищує 141 мс. Це забезпечує можливість швидкого реагування на виявлені події та підходить для застосувань, де оперативність є критичною.

Діаграма розподілу затримок наведена на рис. 4.7.

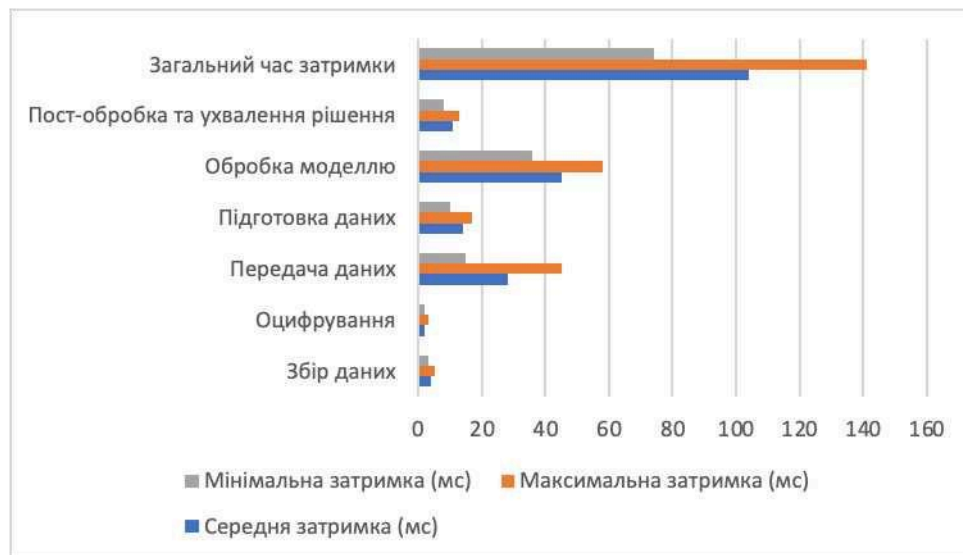


Рисунок 4.7 – Діаграма розподілу затримок

Запропонована інформаційна технологія обробки аудіосигналів демонструє високу ефективність за часом проходження сигналу, що робить її придатною для застосування в режимі реального часу, де затримки повинні бути мінімальними.

4.3.2 Оцінка за точністю розпізнавання

Точність розпізнавання аудіоподій є ключовим показником ефективності системи аудіоаналітики. Для оцінки точності

використовувалися метрики F1-score та EER. Ці метрики дозволяють об'єктивно оцінити продуктивність системи. Матриця плутанини для оцінки точності розпізнавання наведена у табл. 4.2.

Таблиця 4.2 – Матриця плутанини для оцінки точності розпізнавання

Фактичний \ Передбачений	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Разом
1. Accelerating	60	3	3	3	0	0	0	6	0	0	0	0	5	0	0	80
2. Aircraft	2	59	3	2	0	0	3	3	0	0	0	1	0	5	0	78
3. Alarm	3	2	60	2	0	4	0	3	0	0	0	0	0	5	0	79
4. Animal	2	0	0	58	3	0	2	0	5	3	0	2	0	1	0	76
5. Applause	0	1	0	2	62	0	1	4	5	3	1	0	0	0	2	81
6. Bell	1	2	4	0	0	57	2	3	0	1	0	0	0	3	2	75
7. Vehicle	5	4	0	0	1	0	59	1	2	1	0	0	3	2	0	78
8. Musical Instr.	1	3	4	4	8	10	0	66	3	5	0	0	0	1	2	107
9. Domestic Sounds	0	2	5	5	4	3	1	0	56	0	0	0	0	0	1	77
10. Human Voice	2	0	0	3	2	0	0	0	3	61	1	5	3	0	0	80
11. Water	1	0	0	3	0	0	0	0	2	4	60	0	5	2	2	79
12. Human Group	0	1	0	5	9	0	0	0	3	10	0	59	0	0	0	87
13. Engine	7	5	0	0	0	0	6	1	0	0	3	0	58	1	0	81
14. Explosion	5	1	5	1	0	3	4	0	1	0	0	0	2	59	2	83
15. Glass	0	0	3	1	1	3	1	5	3	0	0	0	0	2	56	75
Разом	89	83	87	89	90	80	79	92	83	88	65	67	76	81	67	1216

Результати розрахунків F1-score та EER наведені у табл. 4.3.

Таблиця 4.3 – Результати розрахунків F1-score та EER

Клас	EER	F1-score
1. Accelerating	0.25	0.71
2. Aircraft	0.24	0.73
3. Alarm	0.24	0.72
4. Animal	0.24	0.7
5. Applause	0.23	0.73
6. Bell	0.24	0.74
7. Vehicle	0.24	0.75
8. Musical Instr.	0.38	0.66
9. Domestic Sounds	0.27	0.7
10. Human Voice	0.24	0.73
11. Water	0.24	0.83
12. Human Group	0.32	0.77
13. Engine	0.28	0.74
14. Explosion	0.29	0.72
15. Glass	0.25	0.79
Середнє значення	0.27	0.73

Масштаб вибірки для класу «8 Musical Instrument» 107 зразків є найбільшим показником, а 66 правильно класифікованих записів становлять лише близько 61.7%. Це може вказувати на велику різноманітність аудіофрагментів у межах цього класу, що ускладнює розпізнавання. «12 Human Group» також помітно відстає, точність всього 67.8%. Причиною може бути схожість із «10 Human Voice» або брак чітких відмінних ознак у навчанні. У ході тестування система досягла F1-score на рівні 73.46% та EER 0,27. Ці показники є покращенням порівняно з попередніми експериментами та відповідають або перевершують результати аналогічних систем у галузі. Підвищення точності було досягнуто завдяки оптимізації моделей, використанню додаткових даних для навчання та впровадженню методів аугментації даних.

Діаграма розподілу F1-score та EER наведена на рис. 4.8

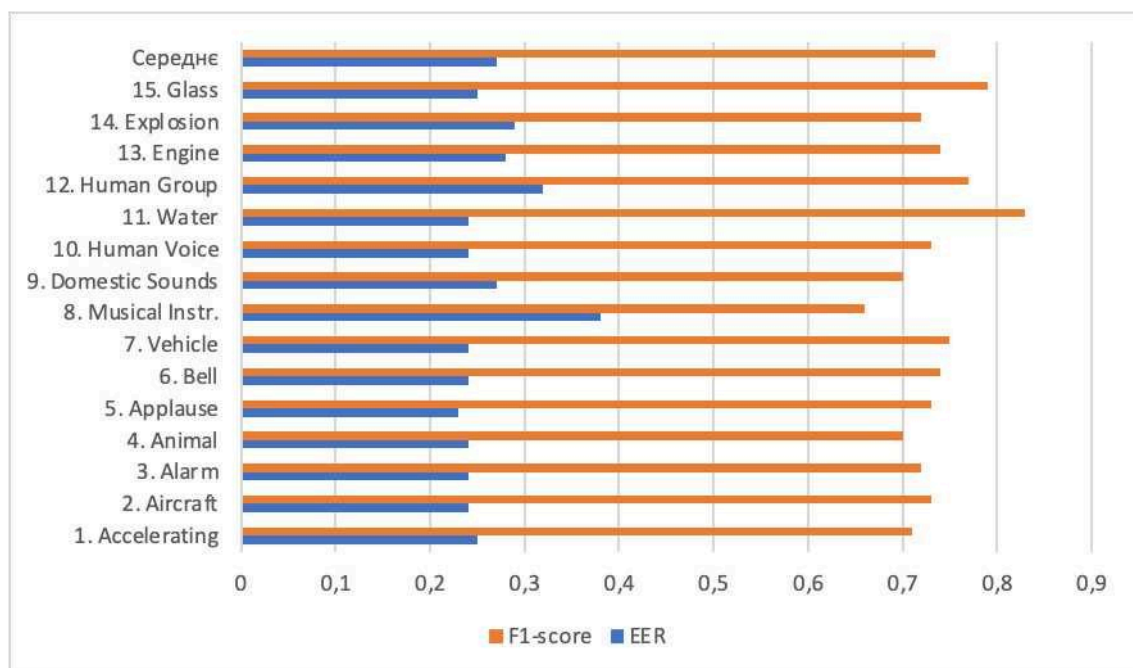


Рисунок 4.8 – Діаграма розподілу F1-score та EER

Запропонована інформаційна технологія обробки аудіосигналів забезпечує конкурентоспроможну точність розпізнавання аудіоподій, що

підтверджує її ефективність для практичного використання у різних сферах, де необхідно здійснювати точний та швидкий аналіз аудіосигналів.

Висновки за розділом 4

У цьому розділі було проведено дослідження та розробку інформаційної технології обробки аудіосигналів у мережних системах аудіоаналітики. У цьому розділі були досягнуті наступні результати:

1. Описано інформаційну технологію обробки аудіосигналів;
2. Виконано вибір технологічного стеку та архітектури системи та розроблено архітектуру системи;
3. Реалізовано основні компоненти інформаційної технології обробки аудіосигналів;
4. Виконано оцінку ефективності запропонованої інформаційної технології за такими критеріями як час проходження сигналу та точність розпізнавання.

Запропонована інформаційна технологія обробки аудіосигналів демонструє високу продуктивність, надійність та гнучкість, що робить її перспективною для застосування у сферах безпеки, моніторингу та інших галузях, де потрібне точне розпізнавання звукових подій в режимі реального часу. Хоча інформаційна технологія демонструє високу ефективність, її продуктивність може знижуватись у середовищах із високим рівнем шуму або при одночасному виникненні декількох аудіоподій.

Перелік використаних джерел у даному розділі наведено у повному переліку використаних джерел під номерами: 99–106.

ВИСНОВКИ

У дисертаційній роботі розв'язана актуальна науково-прикладна задача підвищення точності аналізу та зменшення затримок передачі аудіосигналів у мережних системах аудіоаналітики шляхом оптимізації мережної інфраструктури та удосконалення методів передачі й аналізу аудіосигналів. При цьому були отримані такі наукові та практичні результати:

1. Проведено аналіз сучасних напрямків розвитку методів обробки аудіосигналів і мережних систем аудіоаналітики. Визначено переваги й недоліки існуючих систем, обґрунтовано актуальність і необхідність удосконалення методів передачі та аналізу аудіосигналів у реальному часі;

2. Розроблено математичну модель процесу проходження аудіосигналів у мережних системах аудіоаналітики, що дозволило оптимізувати мережну структуру за критерієм мінімуму капітальних витрат. Це підвищує ефективність передачі даних і зменшує витрати на мережну інфраструктуру;

3. Розроблено метод вибору мережної структури системи аудіоаналітики, що враховує ієрархічну структуру та різні типи навантажень;

4. Удосконалено метод передачі аудіосигналів за рахунок застосування комутації віртуальних маршрутів. Запропонований підхід дозволяє забезпечити високу якість передачі, знизити затримки передачі аудіосигналів на 10% та підвищити стабільність роботи системи в умовах динамічних навантажень;

5. Удосконалено метод аналізу аудіосигналів на основі багатоподійного класифікатора. Це дозволило підвищити точність аналізу аудіосигналів, а саме покращивши показник F1-score на 0,9% і зменшивши рівень помилок EER на 0,1;

6. Удосконалено інформаційну технологію обробки аудіосигналів, яка інтегрує запропоновані методи передачі і аналізу та дозволяє покращити

процеси автоматичної обробки у різних галузях застосування;

7. Проведено апробацію результатів дослідження, що підтвердило ефективність запропонованих моделі та методів в умовах реальних виробничих та освітніх середовищ.

Результати дисертаційної роботи підтверджено експериментальними дослідженнями, що засвідчили їхню достовірність і відповідність теоретичним розрахункам. Отримані методи та модель становлять науково-практичну основу для розвитку та впровадження мережних систем аудіоаналітики у сферах безпеки, медичної діагностики, аудіоспостереження, клієнтської підтримки та моніторингових систем.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Poroshenko A., Kovalenko A. Optimization of a basic network in audio analytics systems // *Advanced Information Systems*. 2023. Vol. 7(1). P. 23–28. doi: <https://doi.org/10.20998/2522-9052.2023.1.04> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).
2. Poroshenko A., Kovalenko A. Audio event analysis method in network-based audio analytics systems // *Advanced Information Systems*. 2024. Vol. 8(4). P. 60–64. doi: <https://doi.org/10.20998/2522-9052.2024.4.08> (Фахове видання категорії А, входить до міжнародної наукометричної бази Scopus, 4 квартиль).
3. Kovalenko A., Poroshenko A. Analysis of the sound event detection methods and systems // *Advanced Information Systems*. 2022. Vol. 6(1) P. 65–69. doi: <https://doi.org/10.20998/2522-9052.2022.1.11> (Фахове видання категорії Б)
4. Порошенко А. І. Математична модель проходження аудіосигналів у мережних системах аудіоаналітики // *Advanced Information Systems*. 2022. Т. 6(4). С. 25–29. doi: <https://doi.org/10.20998/2522-9052.2022.4.04> (Фахове видання категорії Б).
5. Poroshenko A., Kovalenko A. Audio signal transmission method in network-based audio analytics system // *Innovative technologies and scientific solutions for industries*. 2023. Vol. 4(26). P. 58–67. doi: <https://doi.org/10.30837/ITSSI.2023.26.058> (Фахове видання категорії Б).
6. Poroshenko A., Kovalenko A., Sedlaček P. Organization of Audio Analytics Systems Topologies // *2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*. Kharkiv, Ukraine, 2022. P. 517-521, doi: <https://doi.org/10.1109/PICST57299.2022.10238687>

7. Порошенко А.І., Коваленко А.А. Методи класифікації ознак аудіосигналів // Проблеми інформатизації: тези доп. 9-ї міжнар. наук.-техн. конф., 18-19 листопада 2021 р. Харків, 2021. Т. 1. С. 90.

8. Порошенко А.І., Коваленко А.А. Методи та підходи до детектування та розпізнавання поліфонічних аудіоподій // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: матеріали 12-ї міжнар. наук.-техн. конф., 27-28 квітня 2022 р. Харків, 2022. Т. 1. С. 37.

9. Порошенко А.І., Коваленко А.А. Методи стиснення аудіоданих у системах аудіоаналітики // Проблеми інформатизації: тези доп. десятої міжнар. наук.-техн. конф., 24-25 листопада 2022 р. Харків, 2022. Т. 1. С. 92.

10. Порошенко А.І., Коваленко А.А. Методи зниження шуму звукового сигналу // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : матеріали 13-ї міжнар. наук.-техн. конф., 26-27 квітня 2023 р. Харків, 2023. Т. 2. С. 54.

11. Порошенко А.І., Коваленко А.А. Використання методів шумопоглинання у системах аудіоаналітики // Проблеми інформатизації: тези доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. Харків, 2023. Т. 3. С. 49.

12. Порошенко А.І., Коваленко А.А., Мартовицький В.О. Архітектура програмно-апаратної моделі наземного обладнання РСБН типу TACAN // Проблеми інформатизації: тези доп. 11-ї міжнар. наук.-техн. конф., 16-17 листопада 2023 р. м. Харків, 2023. Т. 3. С. 106.

13. Порошенко А.І., Коваленко А.А. Variability of delays in network-based audio analytics systems // Сучасні інформаційні системи та технології в цифровому суспільстві: матеріали Міжнародної науково-практичної конференції, 18 - 19 квітня 2024 р. м. Харків, 2024. С. 106.

14. Порошенко А.І., Коваленко А.А. Special features of implementing QoS in network-based audio analytics systems // The 23rd International scientific

conference information technologies and management 2024, April 25-26. Riga, Latvia. 2024.

15. Порошенко А.І., Коваленко А.А., Буслов П.В. Можливості використання нейронних мереж для атак на мультисервісні // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 2. С. 48.

16. Порошенко А.І., Коваленко А.А. Метод класифікації звукових подій у мережних системах аудіоаналітики // Проблеми інформатизації: тези доп. 12-ї міжнар. наук.-техн. конф., 21-22 листопада 2024 р. м. Харків, 2024. Т. 3, С. 8.

17. Alharbi S. Automatic Speech Recognition: Systematic Literature Review // IEEE Access. 2021. Vol. 9, P. 131858-131876. doi: <https://doi.org/10.1109/ACCESS.2021.3112535>

18. Mesaros A., Heittola T., Virtanen T., Plumbley M.D. Sound Event Detection: A tutorial // IEEE Signal Processing Magazine. 2021. Vol. 38. №, 5. P. 67-83. doi: <https://doi.org/10.1109/MSP.2021.3090678>

19. Gul S., Khan M. S. A Survey of Audio Enhancement Algorithms for Music, Speech, Bioacoustics, Biomedical, Industrial, and Environmental Sounds by Image U-Net // IEEE Access. 2023. Vol. 11. P. 144456-144483 doi: <https://doi.org/10.1109/ACCESS.2023.3344813>

20. Ren Z., Chang Y., Nguyen T. T., Tan Y., Qian K., Schuller B. W. A Comprehensive Survey on Heart Sound Analysis in the Deep Learning Era // IEEE Computational Intelligence Magazine. 2024. Vol. 19, №. 3. P. 42-57 doi: <https://doi.org/10.1109/MCI.2024.3401309>

21. Prajapati Y. J., Gandhi P. P., Degadwala S. A Review - ML and DL Classifiers for Emotion Detection in Audio and Speech Data // 2022 International Conference on Inventive Computation Technologies (ICICT). Nepal, 2022. P. 63-69. doi: <https://doi.org/10.1109/ICICT54344.2022.9850614>

22. Triantafyllopoulos A. An Overview of Affective Speech Synthesis and Conversion in the Deep Learning Era // Proceedings of the IEEE. 2023. Vol. 111, № 10. P. 1355-1381. doi: <https://doi.org/10.1109/JPROC.2023.3250266>
23. Zhu H., Luo MD., Wang R. Deep Audio-visual Learning: A Survey // Int. J. Autom. Comput. 2021. Vol. 18. P. 351–376. doi: <https://doi.org/10.1007/s11633-021-1293-0>
24. Kaminski M. E. Privacy and the Right to Record. BUL Rev. 2017. 167 pp.
25. Arnold S. The Audio Spectrum. In: Radio and Radar Astronomy Projects for Beginners // The Patrick Moore Practical Astronomy Series. Springer, Cham. 2021. P. 61-65. doi: https://doi.org/10.1007/978-3-030-54906-0_3
26. Ulyanov O. M., Plakhov M. S., Shevtsova A. I., Ulyanova O. O., Skoryk A. A., Tkachev V. N. Testing the speed of the FFT using the NVIDIA graphic cards // 2015 International Young Scientists Forum on Applied Physics (YSF). 2015. P. 1-1. doi: <https://doi.org/10.1109/YSF.2015.7333253>
27. Klimushin P., Solianyuk T., Kolisnyk T., Mozhaiv O. Potential application of hardware protected symmetric authentication microcircuits to ensure the security of internet of things // Advanced Information Systems. 2021. Vol. 5, № 3. P. 103–111, doi: <https://doi.org/10.20998/2522-9052.2021.3.14>
28. Rabcan J., Levashenko V., Zaitseva E., Kvassay M., Subbotin S. Non-destructive diagnostic of aircraft engine blades by Fuzzy Decision Tree // Engineering Structures. 2019. Vol. 197. doi: <https://doi.org/10.1016/j.engstruct.2019.109396>
29. Kuchuk H., Kovalenko A., Ibrahim B.F., Ruban I. Adaptive compression method for video information // International Journal of Advanced Trends in Computer Science and Engineering. 2019. Vol. 8, № 1.2. P. 66-69. DOI: <https://doi.org/10.30534/ijatcse/2019/1181.22019>
30. Barkovska O., Kholiev V., Lytvynenko V. Study of noise reduction methods in the sound sequence when solving the speech-to-text problem //

Advanced Information Systems. 2022. Vol. 6, № 1. P. 48–54. doi: <https://doi.org/10.20998/2522-9052.2022.1.08>

31. Ashouri M., Davidsson P., Spalazzese R. Cloud, Edge, or Both? Towards Decision Support for Designing IoT Applications // 2018 Fifth International Conference on Internet of Things: Systems, Management and Security. Valencia, Spain, 2018. P. 155-162. doi: <https://doi.org/10.1109/IoTSMS.2018.8554827>

32. Tkachov V., Hunko M., Volotka V. Scenarios for Implementation of Nested Virtualization Technology in Task of Improving Cloud Firewall Fault Tolerance // 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T). 2019. P. 759-763, doi: <https://doi.org/10.1109/PICST47496.2019.9061473>.

33. Davydov V., Hrebeniuk D. Development the resources load variation forecasting method within cloud computing systems // Advanced Information Systems. 2020. Vol. 4, №, 4. P. 128–135. <https://doi.org/10.20998/2522-9052.2020.4.18>

34. Dufaux A., Besacier L., Ansorge M., Pellandini F. Automatic sound detection and recognition for noisy environment // 2000 10th European Signal Processing Conference. Tampere, Finland, 2000. P. 1-4.

35. Phan H. What makes audio event detection harder than classification? // 2017 25th European Signal Processing Conference (EUSIPCO). 2017. P. 2739-2743. doi: <https://doi.org/10.23919/EUSIPCO.2017.8081709>.

36. Rahman S.U., Khan A., Abbas S. Hybrid system for automatic detection of gunshots in indoor environment // Multimed Tools Appl. 2021. Vol. 80. P. 4143–4153. doi: <https://doi.org/10.1007/s11042-020-09936-w>

37. Turpault N., Serizel R. Training Sound Event Detection On A Heterogeneous Dataset // DCASE Workshop. Tokyo, Japan, 2020.

38. Salamon J., MacConnell D., Cartwright M., Li P., Bello J. P. Scaper: A library for soundscape synthesis and augmentation // 2017 IEEE Workshop on

Applications of Signal Processing to Audio and Acoustics (WASPAA). NY, USA, 2017. P. 344-348. doi: <https://doi.org/10.1109/WASPAA.2017.8170052>.

39. Shah A., Kumar A., Hauptmann A. G., Raj B. A closer look at weak label learning for audio events // arXiv:1804.09288. [Online]. Available: <http://arxiv.org/abs/1804.09288>

40. Kumar K., Chaturvedi K. An Audio Classification Approach using Feature extraction neural network classification Approach // 2nd International Conference on Data, Engineering and Applications (IDEA). 2020. P. 1-6. doi: <https://doi.org/10.1109/IDEA49133.2020.9170702>.

41. Hirata K., Kato T., Oshima R. Classification of Environmental Sounds Using Convolutional Neural Network with Bispectral Analysis // 2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS). 2019. P. 1-2, doi: <https://doi.org/10.1109/ISPACS48206.2019.8986304>.

42. Turpault N., Wisdom S., Erdogan H., Hershey J.R., Serizel R. Improving Sound Event Detection In Domestic Environments Using Sound Separation // DCASE Workshop 2020 - Detection and Classification of Acoustic Scenes and Events. Tokyo, Japan, 2020.

43. Tzinis E., Wisdom S., Hershey J. R., Jansen A., Ellis D. P. W. Improving Universal Sound Separation Using Sound Classification // 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2020. P. 96-100. doi: <https://doi.org/10.1109/ICASSP40776.2020.9053921>.

44. Sose S., Mali S., Mahajan S. P. Sound Source Separation Using Neural Network // 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). 2019. P. 1-5, doi: <https://doi.org/10.1109/ICCCNT45670.2019.8944614>.

45. Zheng X., Chen H., Song Y. Zheng USTC teams submission for dcase2021 task4 semi-supervised sound event detection // DCASE2021 Challenge, Tech. Rep, 2021.

46. Kim N. K., Kim H. K. Self-training with noisy student model and semi-supervised loss function for dcase 2021 challenge task 4 // arXiv preprint. 2021. arXiv:2107.02569.
47. Lu R. Integrating advantages of recurrent and transformer structures for sound event detection in multiple scenarios // Tech. Rep., DCASE2021 Challenge. 2021.
48. Ebbers J., Haeb-Umbach R. Self-Trained Audio Tagging and Sound Event Detection in Domestic Environments // Proceedings of the Detection and Classification of Acoustic Scenes and Events 2021 Workshop (DCASE2021). 2021. P. 15-19.
49. Nam H. Heavily augmented sound event detection utilizing weak predictions // arXiv preprint. 2021. arXiv:2107.03649.
50. Tian G. Sound event detection using metric learning and focal loss for dcase 2021 task 4. // Tech. Rep., DCASE2021 Challenge, 2021.
51. Hunter D. K., Andonovic I. Approaches to optical Internet packet switching // IEEE Communications Magazine. 2000. Vol. 38, № 9. P. 116-122. doi: <https://doi.org/10.1109/35.868150>
52. Chiesa M., Kamisiński A., Rak J., Rétvári G., Schmid S. A Survey of Fast-Recovery Mechanisms in Packet-Switched Networks // IEEE Communications Surveys & Tutorials. 2021. Vol. 23, № 2. P. 1253-1301. doi: <https://doi.org/10.1109/COMST.2021.3063980>
53. Adamo Di Packet switching in quantum networks: A path to the quantum Internet // Physical Review Research. 2022. Vol. 4. P. 43–64. doi: <https://doi.org/10.1103/PhysRevResearch.4.043064>
54. Yoo S. J. B. Optical Packet and Burst Switching Technologies for the Future Photonic Internet // Journal of Lightwave Technology. 2006. Vol. 24, № 12. P. 4468-4492. doi: <https://doi.org/10.1109/JLT.2006.886060>
55. Manzato A. G., da Fonseca N. L. S. A survey of channel switching schemes for IPTV // IEEE Communications Magazine. 2013. Vol. 51, № 8, P. 120-127. doi: <https://doi.org/10.1109/MCOM.2013.6576349>

56. Yoo W., Sim A. Network bandwidth utilization forecast model on high bandwidth networks // 2015 International Conference on Computing, Networking and Communications (ICNC). Garden Grove, CA, USA, 2015. P. 494-498, doi: <https://doi.org/10.1109/ICCNC.2015.7069393>.

57. Shukla S., Hassan M.F., Tran D.C. Improving latency in Internet-of-Things and cloud computing for real-time data transmission: a systematic literature review (SLR) // Cluster Comput. 2023. Vol. 26. P. 2657–2680. doi: <https://doi.org/10.1007/s10586-021-03279-3>

58. Park J., Samarakoon S., Shiri H. Extreme ultra-reliable and low-latency communication // Nat Electron 5. 2022. P. 133–141. doi: <https://doi.org/10.1038/s41928-022-00728-8>

59. Singh M., Baranwal G. Quality of Service (QoS) in Internet of Things // 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU). Bhimtal, India, 2018. P. 1-6. doi: <https://doi.org/10.1109/IoT-SIU.2018.8519862>.

60. Cavalcanti J. C. Microphone and audio compression effects on acoustic voice analysis: A pilot study // Journal of Voice. 2023. Vol. 37, № 2. P. 162-172, doi: <https://doi.org/10.1016/j.jvoice.2020.12.005>

61. Shi Y. Calculation of Network System Reliability Based on Improved Disjointed Minimal Path Set // 2010 International Conference on E-Product E-Service and E-Entertainment. Henan, China, 2010. P. 1-4. doi: <https://doi.org/10.1109/ICEEE.2010.5660486>.

62. Yevdokymenko M., Manasse M., Zalushniy D., Sleiman B. Analysis of methods for assessing the reliability and security of infocommunication network // 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). Kharkiv, Ukraine, 2017. P. 199-202. doi: <https://doi.org/10.1109/INFOCOMMST.2017.8246379>.

63. Mohanty R. K., Sahoo S. P., Kabat M. R. A Network Reliability based Secure Routing Protocol (NRSRP) for Secure Transmission in Wireless Body Area Network // 2023 8th International Conference on Communication and Electronics

Systems (ICCES). Coimbatore, India, 2023. P. 663-668, doi: <https://doi.org/10.1109/ICCES57224.2023.10192691>

64. Zeghidour N., Luebs A., Omran A., Skoglund J., Tagliasacchi M. SoundStream: An End-to-End Neural Audio Codec // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2022. Vol. 30. P. 495-507. doi: <https://doi.org/10.1109/TASLP.2021.3129994>

65. Jian MS., Wu J.M.T. Hybrid Internet of Things (IoT) data transmission security corresponding to device verification // Ambient Intell Human Comput. 2021. doi: <https://doi.org/10.1007/s12652-021-03122-y>

66. Zhang C. Scalable Traffic Engineering for Higher Throughput in Heavily-loaded Software Defined Networks // 2020 IEEE/IFIP Network Operations and Management Symposium. Budapest, Hungary, 2020. P. 1-7, doi: <https://doi.org/10.1109/NOMS47738.2020.9110259>.

67. Pratt C.W. The concept of marginal overflow in alternate routing // The 5th ITC. New-York, June, 1967. P. 52–58.

68. Wallstrom A. Methods for Optimizing Alternative Routing Networks // Ericsson Technics. № 1. P. 3–29.

69. Bailey A. Network Technology for Digital Audio // Routledge. 2001. doi: <https://doi.org/10.4324/9780080512556>

70. Mamchych O., Volk M. Estimation of power consumption of mobile devices in cloud computing // Innovative technologies and scientific solutions for industries. 2023. Vol. 1. № 23. P. 72–82. DOI: <https://doi.org/10.30837/ITSSI.2023.23.07>

71. Mota A. V., Azam S., Shanmugam B., Yeo K. C., Kannoorpatti K. Comparative analysis of different techniques of encryption for secured data transmission // 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI). Chennai, India, 2017. P. 231-237. doi: <https://doi.org/10.1109/ICPCSI.2017.8392158>.

72. Chaudhry S. An Encryption-based Secure Framework for Data Transmission in IoT // 2018 7th International Conference on Reliability, Infocom

Technologies and Optimization (Trends and Future Directions) (ICRITO). Noida, India, 2018. P. 743-747. doi: <https://doi.org/10.1109/ICRITO.2018.8748523>.

73. Alagic G., Broadbent A., Fefferman B., Gagliardini T., Schaffner C., St. Jules M. Computational Security of Quantum Encryption // Information Theoretic Security. Lecture Notes in Computer Science. Springer, Cham. 2016. Vol. 10015. doi: https://doi.org/10.1007/978-3-319-49175-2_3

74. Xu D., Yu K., Ritcey J. A. Cross-Layer Device Authentication With Quantum Encryption for 5G Enabled IIoT in Industry 4.0 // IEEE Transactions on Industrial Informatics. 2022. Vol. 18, № 9. P. 6368-6378. doi: <https://doi.org/10.1109/TII.2021.3130163>.

75. Martovytskyi V., Sievierinov O., Liashenko O., Koltun Y., Liashenko S., Kis V., Sukhoteplyi V., Nosyk A., Konov D., Yevstrat D. Devising an approach to the identification of system users by their behavior using machine learning methods // Eastern-European Journal of Enterprise Technologies. 2022. Vol. 3, № 117. P. 23–34. doi: <https://doi.org/10.15587/1729-4061.2022.259099>

76. Barkovska O., Movsesian I., Yeromina N., Liashenko O., Tkachenko D. System of individual multidimensional biometric authentication // International Journal of Emerging Trends in Engineering Research. 2019 Vol. 7, № 12. P. 812–817. doi: <https://doi.org/10.30534/ijeter/2019/147122019>

77. Kholiev V. and Barkovska O. Comparative analysis of neural network models for the problem of speaker recognition // Innovative technologies and scientific solutions for industries. 2023. Vol. 2, № 24. P. 172–178. doi: <https://doi.org/10.30837/ITSSI.2023.24.172>.

78. Mykhailichenko I., Ivashchenko H., Barkovska O., Liashenko O. Application of Deep Neural Network for Real-Time Voice Command Recognition // 2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek). Kharkiv, Ukraine, 2022. P. 1-4, doi: <https://doi.org/10.1109/KhPIWeek57572.2022.9916473>.

79. European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27

April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation - GDPR). Official Journal of the European Union. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

80. Barkovska O. Performance study of the text analysis module in the proposed model of automatic speaker's speech annotation // *Computer systems and information technologies*. 2022. Vol. 4. P. 13–19. DOI: <https://doi.org/10.31891/csit-2022-4-2>.

81. Kuchuk N. Mathematical Methods of Reliability Analysis of the Network Structures: Securing QoS on Hyperconverged Networks for Traffic Anomalies // *Future Intent-Based Networking. Lecture Notes in Electrical Engineering*. 2022. Vol. 831. P. 223–241. DOI: https://doi.org/10.1007/978-3-030-92435-5_13.

82. Ghido F., Tabus I. Sparse Modeling for Lossless Audio Compression // *IEEE Transactions on Audio, Speech, and Language Processing*. 2013. Vol. 21. № 1. P. 14–28. DOI: <https://doi.org/10.1109/TASL.2012.2211014>.

83. Huang H., Shu H., Yu R. Lossless audio compression in the new IEEE Standard for Advanced Audio Coding // *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2014. P. 6934–6938, DOI: <https://doi.org/10.1109/ICASSP.2014.6854944>.

84. Petrosky E. E., Michaels A. J., Ridge D. B. Network Scalability Comparison of IEEE 802.15.4 and Receiver-Assigned CDMA // *IEEE Internet of Things Journal*. 2019. Vol. 6. № 4. P. 6060–6069. DOI: <https://doi.org/10.1109/JIOT.2018.2884455>.

85. Meng X., Pappas V., Zhang L. Improving the Scalability of Data Center Networks with Traffic-aware Virtual Machine Placement // *2010 Proceedings IEEE INFOCOM*. 2010. P. 1–9, DOI: <https://doi.org/10.1109/INFCOM.2010.5461930>.

86. Amarudin, Ferdiana R., Widyawan. A. Systematic Literature Review of Intrusion Detection System for Network Security: Research Trends, Datasets and

Methods // 2020 4th International Conference on Informatics and Computational Sciences (ICICoS). 2020. P. 1–6, <https://doi.org/DOI:10.1109/ICICoS51170.2020.9299068>.

87. Szymanski A. Grade-of-service-based routing in optical networks [Quality-of-Service-Based Routing Algorithms for Heterogeneous Networks] // IEEE Communications Magazine. 2007. Vol. 45. № 2. P. 82–87. DOI: <https://doi.org/10.1109/MCOM.2007.313400>.

88. Liu C. H., Gkelias A., Leung K. K. Connection admission control and grade of service for QoS routing in mesh networks // 2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications. 2008. P. 1–5, DOI: <https://doi.org/10.1109/PIMRC.2008.4699895>.

89. Barkovska O., Havrashenko A. Analysis of the influence of selected audio pre-processing stages on accuracy of speaker language recognition // Innovative technologies and scientific solutions for industries. 2023. Vol. 4, № 26. P. 16–23. doi: <https://doi.org/10.30837/ITSSI.2023.26.016>.

90. Koosha S. and Mahyar A. Machine Learning and Deep Learning: A Review of Methods and Applications // World Information Technology and Engineering Journal. 2023. Vol. 10, № 7. P. 3897-3904. Available at SSRN: <https://ssrn.com/abstract=4458723>

91. Grumiaux P.A., Kitić S., Girin L., Guérin A. A survey of sound source localization with deep learning methods // Acoust. Soc. Am. 2022. Vol. 152 P. 107-151. doi: <https://doi.org/10.1121/10.0011809>

92. Zaman K., Sah M., Direkoglu C., Unoki M. A Survey of Audio Classification Using Deep Learning // IEEE Access. 2023. Vol. 11. P. 106620-106649. doi: <https://doi.org/10.1109/ACCESS.2023.3318015>.

93. Fonseca E., Favory X., Pons J., Font F., Serra X. FSD50K: An Open Dataset of Human-Labeled Sound Events // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2022. Vol. 30. P. 829-852. doi: <https://doi.org/10.1109/TASLP.2021.3133208>.

94. Karol J. Piczak. ESC: Dataset for Environmental Sound Classification // Proceedings of the 23rd ACM international conference on Multimedia (MM '15). NY, USA, 2015. P. 1015–1018. doi: <https://doi.org/10.1145/2733373.2806390>

95. Foster P., Sigtia S., Krstulovic S., Barker J., Plumbley M. D. Chime-home: A dataset for sound source recognition in a domestic environment // 2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). NY, USA, 2015. P. 1-5. doi: <https://doi.org/10.1109/WASPAA.2015.7336899>.

96. Kavalero I. Universal Sound Separation // 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). New Paltz, NY, USA, 2019. P. 175-179. doi: <https://doi.org/10.1109/WASPAA.2019.8937253>.

97. DCASE Challenge 2023 - Sound Event Localization and Detection Evaluated in Real Spatial Sound Scenes. URL: <https://dcase.community/challenge2023/task-sound-event-localization-and-detection-evaluated-in-real-spatial-sound-scenes> (дата звернення: 28.03.2025).

98. Hunko M., Tkachov V., Kuchuk H., Kovalenko A. Advantages of Fog Computing: A Comparative Analysis with Cloud Computing for Enhanced Edge Computing Capabilities // 2023 IEEE 4th KhPI Week on Advanced Technology. 2023. P. 1-5. doi: <https://doi.org/10.1109/KhPIWeek61412.2023.10312948>

99. Kuchuk H., Malokhvii E. Integration of IOT with Cloud, Fog, and Edge Computing: A Review // Advanced Information Systems. 2024. Vol. 8, № 2. P. 65–78. doi: <https://doi.org/10.20998/2522-9052.2024.2.08>

100. Petrovska I., Kuchuk H. Adaptive resource allocation method for data processing and security in cloud environment // Advanced Information Systems. 2023. Vol. 7, № 3. P. 67–73. doi: <https://doi.org/10.20998/2522-9052.2023.3.10>

101. Welcome to Python.org. URL: <https://www.python.org> (дата звернення: 16.01.2025).

102. The web framework for perfectionists with deadlines. URL: <https://www.djangoproject.com> (дата звернення: 16.01.2025).

103. PostgreSQL: The world's most advanced open source database. URL: <https://www.postgresql.org> (дата звернення: 16.01.2025).

104. Business VPN For Secure Networking. URL: <https://openvpn.net> (дата звернення: 16.01.2025).

105. The library for web and native user interfaces. URL: <https://react.dev> (дата звернення: 16.01.2025).