

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Навчально-науковий центр заочної форми навчання

(повна назва)

Кафедра Інформаційно-мережної інженерії

(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

Аналіз принципів роботи аудіопротоколів в комп'ютерних мережах

(тема)

Виконав:

здобувач 4 року навчання,
групи ТРІМІЗ-21-1

Євгеній ХІРІН

(власне ім'я, прізвище)

Спеціальність 172 Телекомунікації та
радіотехніка

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма

«Інформаційно-мережна інженерія»

(повна назва освітньої програми)

Керівник доц. Наталія ХАРЧЕНКО

(посада, власне ім'я, прізвище)

Допускається до захисту

Зав. кафедри

(підпис)

Валерій БЕЗРУК

(прізвище, ініціали)

2025 р.

Не містить відомостей заборонених до відкритого публікування.

Здобувач

/ Євген Хірін /

Керівник

/ Наталія Харченко /

Харківський національний університет радіоелектроніки

Навчально-науковий центр заочної форми навчання

Кафедра Інформаційно-мережної інженерії

Рівень вищої освіти перший (бакалаврський)

Спеціальність 172 «Телекомунікації та радіотехніка»

(код і повна назва)

Тип програми освітньо-професійна

Освітня програма «Інформаційно-мережна інженерія»

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 2025 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Хіріну Євгенію Андрійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз принципів роботи аудіопротоколів в комп'ютерних мережах

затверджена наказом університету від 02 травня 2025 р. № 63 Стз

2. Термін подання студентом роботи до екзаменаційної комісії 24 червня 2025р.

3. Вихідні дані до роботи Розглянути принципи роботи протоколів передачі потокового аудіо. Проаналізувати методіку обробки та передачі аудіоданих в комп'ютерних мережах. Провести порівняльний аналіз характеристик різних протоколів AoIP. за критеріями затримки, надійності, масштабованості, якості передачі та складності реалізації для вибору оптимальних аудіопротоколів при їх використанні в інфокомунікаційних мережах передачі даних

4. Перелік питань, що потрібно опрацювати в роботі _____

Вступ

1. Загальний опис телекомунікаційних каналів і методів передачі

2. Методи обробки та передачі аудіоданих

3. Огляд протоколів передачі аудіо

4. Порівняльний аналіз аудіопротоколів

Висновки

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Слайди у форматі Power Point (назва, мета роботи і її актуальність, визначення телекомунікаційних каналів, класифікація каналів зв'язку, методи модуляції та кодування сигналу, процес оцифрування звуку, стиснення аудіоданих, технології потокової передачі звуку, синхронізація аудіо в реальному часі, аудіопротоколи в комп'ютерних мережах, протоколи потокової передачі, протоколи AoIP, критерії порівняння протоколів передачі аудіо, безпосереднє порівняння протоколів передачі аудіо, висновки)

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Ознайомлення із завданням. Уточнення ТЗ.	02.05 – 06.05.25	виконано
2	Підбір літератури за темою роботи.	07.05 – 11.05.25	виконано
3	Виконання розділу 1	12.05 – 19.05.25	виконано
4	Виконання розділу 2	20.05 – 25.05.25	виконано
5	Виконання розділу 3	26.05 – 10.06.25	виконано
6	Виконання розділу 4	11.06 – 16.06.25	виконано
7	Оформлення пояснювальної записки	17.06 – 19.06.25	виконано
8	Оформлення презентаційного матеріалу,	20.06 – 24.06.25	виконано
9	підготовка до захисту у ЕК		

Дата видачі завдання 2 травня 2025 р.

Здобувач _____
(підпис)

Керівник роботи _____
(підпис)

доц. Наталія Харченко
(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 75 с., 16 рис., 5 табл., 17 джерел, 1 додаток.

Об'єкт дослідження – протоколи передачі та обробки аудіоданих в мережі.

Мета роботи – аналіз принципів функціонування та порівняння протоколів передачі та обробки аудіоданих в мережі.

В епоху стрімкого розвитку телекомунікаційних мереж передавання аудіоданих стало невід'ємною частиною побуту суспільства, особливо в таких напрямках як теле-радіомовлення, VoIP-зв'язок, онлайн-відеоконференції, стрімінгові платформи, live-трансляції, дзвінки в месенджерах тощо. Для забезпечення процесів якісної, надійної та ефективної передачі аудіоінформації мережами використовуються спеціальні мережеві аудіопротоколи, що гарантують передачу аудіопотоку високу точність, збереження якості сигналу, мінімальні затримки, синхронність, стійкість до втрат пакетів та адаптацію до змінних умов мережі.

Не можна не зазначити більш профільні напрямки застосування аудіопротоколів, зокрема аудіопротоколів передачі. Які повсюдно використовуються в професійних студіях звукозапису та на величезних сценах навколо всього світу, задля оптимізації сценічної аудіоінфраструктури, задля забезпечення стабільності, надійності та функціональності аудіообладнання. З часом, використання подібних технологій стає все більш економічно-доцільним та менш затратним в контексті використання людського ресурсу. Сучасні технології стрімко замінюють купу аналогового обладнання, більш компактним, лаконічним, енергоефективним та функціональним.

АУДІОДАНИ, АУДІОПРОТОКОЛ, БУФЕРИЗАЦІЯ, ДИСКРЕТИЗАЦІЯ, КВАНТУВАННЯ, КОДУВАННЯ, МЕРЕЖА, ОЦИФРУВАННЯ ЗВУКУ, ПОТІКОВА ПЕРЕДАЧА, ТЕЛЕКОМУНІКАЦІЙНИЙ КАНАЛ.

THE ABSTRACT

Explanatory note 75 pages, 16 fig., 5 tab., 17 sources, 1 app.

Objects of research – protocols for transmitting and processing audio data in network.

The purpose of work – analysis of the principles of functioning and compare the protocols of transmission and processing of audio data in the network.

In the era of rapid development of telecommunication networks, audio data transmission has become an integral part of society, especially in such areas as broadcasting, VoIP communication, online video conferencing, streaming platforms, live broadcasts, calls in instant messengers, etc. To ensure high-quality, reliable, and efficient transmission of audio information over networks, special network audio protocols are used to guarantee high fidelity, signal quality, minimal delays, synchronization, packet loss resistance, and adaptation to changing network conditions.

We should also mention more specialized areas of application of audio protocols, in particular, audio transmission protocols. They are widely used in professional recording studios and on huge stages around the world to optimize the stage audio infrastructure, to ensure the stability, reliability and functionality of audio equipment. Over time, the use of such technologies is becoming more and more economically feasible and less costly in terms of human resources. Modern technologies are rapidly replacing a lot of analog equipment with more compact, concise, energy-efficient and functional equipment, which in turn only makes people's lives easier.

AUDIO DATA, AUDIO PROTOCOL, BUFFERING, SAMPLING, QUANTISATION, CODING, NETWORK, AUDIO DIGITISATION, STREAMING TELECOMMUNICATIONS CHANNEL.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ.....	8
ВСТУП.....	10
1 ЗАГАЛЬНИЙ ОПИС ТЕЛЕКОМУНІКАЦІЙНИХ КАНАЛІВ І МЕТОДІВ	
ПЕРЕДАЧІ	11
1.1 Визначення телекомунікаційних каналів	11
1.2 Класифікація каналів зв'язку	12
1.3. Методи модуляції та кодування сигналу.....	14
1.4. Особливості передачі аудіоінформації через телекомунікаційні канали .	18
2 МЕТОДИ ОБРОБКИ ТА ПЕРЕДАЧІ АУДІОДАНИХ.....	20
2.1. Оцифровування звуку: дискретизація, квантування, кодування	20
2.2. Стиснення аудіоданих	31
2.3. Поточкова передача аудіо: принципи роботи (Streaming).....	37
2.4. Синхронізація аудіо в реальному часі (VoIP, конференц-зв'язок).....	40
3 ОГЛЯД ПРОТОКОЛІВ ПЕРЕДАЧІ АУДІО.....	45
3.1. Визначення та роль аудіопротоколів у комп'ютерних мережах.....	45
3.2. Протоколи потокової передачі аудіо	46
3.3 Протоколи передачі цифрового аудіо IP-мережами (AoIP)	51
4 ПОРІВНЯЛЬНИЙ АНАЛІЗ АУДІОПРОТОКОЛІВ	57
4.1. Критерії для порівняння (затримка, якість передачі, надійність, масштабованість, складність реалізації).....	57
4.2. Порівняння протоколів аудіообробки та передачі	59
ВИСНОВКИ.....	63
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	63
ДОДАТОК А СЛАЙДИ ПРЕЗЕНТАЦІЇ.....	67

ПЕРЕЛІК СКОРОЧЕНЬ

- AES – Advanced Encryption Standard, просунутий стандарт шифрування;
- AIFF – Audio Interchange File Format, формат файлів для обміну аудіоданими;
- ADPCM – Adaptive Differential Pulse-Code Modulation, адаптивна диференціальна імпульсно-кодова модуляція;
- ALAC – Apple Lossless Audio Codec, аудіо кодек без втрат Apple;
- AoIP – Audio over Internet Protocol, передача аудіо через протокол Інтернет;
- ATM – Asynchronous Transfer Mode, асинхронний режим передачі;
- FDM – Frequency-Division Multiplexing, частотне мультиплексування;
- FLAC – Free Lossless Audio Codec, вільний аудіо кодек без втрат;
- FTP – File Transport Protocol, протокол передачі файлів;
- HTTP – Hypertext Transfer Protocol, протокол передачі гіпертексту;
- RTP – Real-Time Transport Protocol, протокол передачі в реальному часі;
- RTSP – Real-Time Streaming Protocol, протокол потокової передачі в реальному часі;
- RTMP – Real-Time Messaging Protocol, протокол обміну повідомленнями в реальному часі;
- DASH – Dynamic Adaptive Streaming over Hypertext Transfer Protocol, динамічна адаптивна потокова передача через протокол передачі гіпертексту;
- HLS – Hypertext Transfer Protocol Live Streaming, пряма трансляція через протокол передачі гіпертексту;
- LPC – Linear Predictive Coding, лінійне прогнозоване кодування;
- NTP – Network Time Protocol, протокол мережевого часу;
- PTP – Precision Time Protocol, протокол точного часу;
- TDM – Time-Division Multiplexing, мультиплексування за часовим поділом;
- VoIP – Voice over Internet Protocol, передача голосу через протокол Інтернет;
- WAV – Waveform Audio File Format, формат аудіо у формі хвилі;

WDM– Wavelength-Division Multiplexing, мультиплексування з поділом за довжиною хвилі;

АІМ - амплітудно-імпульсна модуляція;

АЦП – аналого-цифровий перетворювач;

КІМ – кодово-імпульсна модуляція;

РРЛ – радіорелейний зв'язок;

ФІМ – фазово-імпульсна модуляція;

ЧІМ – частотно-імпульсна модуляція;

ШІМ – широтно-імпульсна модуляція.

ВСТУП

В наші часи не можна уявити життя без зв'язку. Зокрема онлайн-спілкування. Враховуючи умови нашого побуту на даний момент, зв'язок є для нас забезпеченням власного спокою за близьких та рідних людей, які в залежності від конкретної ситуації можуть бути розкидані по абсолютно протилежних закутках нашої планети. І щоб залишатися зі всіма близькими людьми на зв'язку – ми телефонуємо їм, листуємося в месенджерах та за допомогою SMS-повідомлень, надсилаємо голосові повідомлення, тощо. Це є нашими реаліями непростих часів, в які ми зараз живемо. Проте, дуже стрімкий розвиток телекомунікаційних технологій почався ще до цього, наприкінці 2000-х років XXI століття. Першим справжнім випробуванням як людей, так і телекомунікаційних мереж став початок епідемії COVID-19, яка застала нас всіх зненацька. Головним інструментом спілкування та комунікації людей між собою тоді стали онлайн-відеоконференції, дзвінки в месенджерах та live-трансляції в соціальних мережах, цьому передували впроваджені суворі карантинні обмеження. Такі сервіси онлайн-відеоконференцій, як Zoom, Google Meet, Microsoft Teams та інші почали свій стрімкий розвиток та технологічні перегони. Невід'ємною частиною всього вищезазначеного, нарівні з передачею та обробкою відеозображення, є передача та обробка аудіоданих. Це стосується не лише передачі потокового аудіо з різноманітних аудіопристроїв вводу (накшталт мікрофонів, входів аудіоінтерфейсів тощо), а й передачі потокового аудіо яке транслювалося безпосередньо з пристроїв (демонстрація відео-аудіоматеріалу безпосередньо з пристрою).

В цій бакалаврській кваліфікаційній роботі будуть розглянуті та проаналізовані питання принципів роботи протоколів передачі аудіоінформації, їх роль у сучасних комп'ютерних мережах. Також буде проведено порівняльний аналіз аудіопротоколів за критеріями затримки, надійності, масштабованості, якості передачі та складності реалізації.

1 ЗАГАЛЬНИЙ ОПИС ТЕЛЕКОМУНІКАЦІЙНИХ КАНАЛІВ І МЕТОДІВ ПЕРЕДАЧІ

1.1 Визначення телекомунікаційних каналів

Телекомунікаційні канали – це один з основних компонентів телекомунікаційної мережі, та за сумісництвом – шлях, за яким передаються та приймаються дані. Зазвичай телекомунікаційні канали базуються на фізичних провідниках, наприклад – коаксіальний кабель, чи мідний дріт. А оптоволоконні кабелі, здебільшого, використовуються у високошвидкісних мережах, для забезпечення надійного та швидкого зв'язку.

Існує два визначення терміну «канал». Перше з них визначає канал як фізичне середовище, що передає сигнал між передавачами та приймачами. У якості прикладу можна навести атмосферу для аудіо-зв'язку, скляні оптоволокна, коаксіальні кабелі для зв'язку за допомогою електричної напруги та струму, які протікають в ньому, а також через простір за допомогою світла видимого, інфрачервоного або ультрафіолетового спектрів, а також радіохвиль. Це і є визначення «вільного оптичного каналу». Передача радіохвиль незалежна від наявності чи відсутності атмосфери між передавачем і приймачем. До прикладу, радіохвилі можуть без проблем проходити й крізь вакуум, й через будь-яке газоподібне середовище (такі як, повітря, різноманітних газів тощо).

Також телекомунікаційні канали є похідною шляхів передачі в телекомунікаціях, і саме це відкриває для нас усі переваги мультиплексування. Вільний оптичний канал є середовищем передачі інформації. Телекомунікації, яким підпорядковані декілька приймачів та передавачів розроблені для спільного функціонування та використання одного й того ж самого фізичного каналу, що й називають мультиплексними системами. Використання єдиного фізичного середовища передачі шляхом мультиплексування часто дозволяє суттєво зменшити витрати. Мультиплексовані системи, у свою чергу, легко інтегруються в телекомунікаційні мережі, де сигнали можуть бути перенаправлені у вузлових точках до потрібного отримувача.

Термін «канал» також може тлумачитися як «канал зв'язку», який є - різновидом середовища передачі даних [1]. Саме ця властивість дозволяє одночасно передавати кілька потоків інформації в межах одного фізичного

ресурсу. Як приклад наведу радіостанції, де одна транслюватиме сигнал радіосигнал на частоті, приблизно, 90,0 МГц, а інша транслюватиметься на 91,6 МГц. Смуга частотного діапазону для кожної радіостанції, так звана «ширина каналу» складає 180 кГц з опорою на вищезазначених частотах (в МГц). Даний приклад наочно показує що суміжні канали відокремлені одне від одного на приблизно 200 кГц. А різниця між відокремленнями та шириною каналу має назву «захисний інтервал», у вищенаведеному прикладі захисний інтервал складає 20 кГц. Захисний інтервал використовують в якості страхування від вад системи зв'язку.

Технологія, яка забезпечує поділ середовища передачі за частотними діапазонами, відома як «мультиплексування з частотним розділенням» (FDM, Frequency-Division Multiplexing). Водночас варто враховувати також метод «мультиплексування за довжиною хвилі» (WDM, Wavelength-Division Multiplexing,), що широко застосовується в оптичних мережах, де кілька джерел сигналів можуть використовувати спільний фізичний носій.

Ще одним способом поділу середовища є присвоєння кожному передавачу визначеного та повторюваного інтервалу часу, в межах якого він має право передавати дані. Такий підхід лежить в основі тимчасового розподілу каналів.

Даний метод поділу середовища на канали зв'язку називають «Мультиплексуванням з розподілом за часом» (TDM, Time-Division Multiplexing). Та його використовують, як і WDM, в оптичних комунікаціях. Окремі системи радіочастотного зв'язку використовують TDM в FDM.

1.2 Класифікація каналів зв'язку

Існує декілька статей класифікації каналів зв'язку. Перша з яких – за середовищем розповсюдження сигналу. Канали зв'язку поділяються на електричні та оптичні. Електричні в свою чергу представляються у вигляді дротових каналів (кабелі, електричні/коаксіальні дроти) та бездротових (в яких використовуються електромагнітні хвилі, й постають як інфрачервоні або радіоканали) [2].

Друга стаття класифікації – за формою сигналу передачі. В такому випадку канали зв'язку існують аналогові та цифрові. Де в аналогових амплітуда сигналу має безперервну в часі форму будь-якої фізичної природи, а в цифрових

амплітуда постає у цифровій формі сигналу будь-якої фізичної природи (дискретна або імпульсна).

Третя – за засобами каналотворення. Є радіоканали, дротові, радіорелейні та космічні канали. Радіоканали використовують радіохвилі в якості середовища передачі даних, вони використовують схему передачі сигналу від точки до точки, або так званий «прямий зв'язок». Дротові канали, як стає зрозуміло з назви – виконані у вигляді прокладених дротових мереж, та передають сигнали за допомогою дротів. Радіорелейні канали, на відміну від радіоканалів, можуть утворювати ланцюжок з декількох проміжних станцій та ретрансляторів, й забезпечує доволі надійну стійкість до завад на великих відстанях. Але радіорелейні канали так само використовують як фізичне середовище передачі – радіохвилі. Радіорелейні канали поділяються на канали прямої видимості та на тропосферні канали. Канали прямої видимості використовуються на ділянках РРЛ де встановлена «пряма видимість» між ретрансляторами. Тропосферні канали використовуються у випадках коли умови прямої видимості між «точками» не досягнуті, і тоді використовується простір тропосфери Землі для передачі даних. Космічні канали зв'язку, в свою чергу, використовуються для з'єднання космічних станцій (зокрема і штучних супутників) із наземними станціями, використовуючи для цього радіозв'язок.

Четверта – за відкритістю даних. Існують закодовані канали (тимчасової стійкості, також відомі як «зашумлені», та гарантованої стійкості) та відкриті. Закодовані канали використовують певний різновид шифрування для забезпечення безпеки інформації, що передається, від несанкціонованого доступу до неї або її втрати. Закодовані канали тимчасової стійкості слугують для забезпечення захисту від нетривалих завад. Використовується у VoIP та цифрових засобах передачі аудіоданих. Закодовані канали гарантованої стійкості розраховані на протистояння доволі значним та тривалим завадам, використовується у космічному зв'язку або у стратегічних каналах зв'язку. У відкритих каналах, як зрозуміло з назви, кодування відсутнє. Показовим прикладом постають публічні мережі Wi-Fi, аналогове телебачення, некодовані протоколи (HTTP, Telnet, FTP), перші версії Bluetooth де відсутня автентифікація.

П'ята – за кінцевими джерелами сигналу. Відео, телефонні та телеграфні. Відеоканали зв'язку застосовуються для безпосередньої передачі відеозображення з пристрої фільмування (камери, фотоапарати тощо). В реаліях

нашого часу та ситуації в Україні на даний момент, одним з найпоширеніших прикладів використання відеоканалу є пілотування FPV-дронами, зображення з яких передається на дисплей пілота через радіочастоти (900 МГц; 1,2 ГГц; 2,4 ГГц та 5,8 ГГц) за допомогою передавача. Телефонні та телеграфні канали використовують здебільшого дровий тип зв'язку, проте в них відрізняється діапазон частот (в телефонного каналу це 300-3400 Гц, а в телеграфного – 100-300 Гц). Телеграфні канали розраховані на передачу коротких повідомлень, а телефонні використовуються для мобільного або телефонного зв'язку.

Шоста – за часом перекомутації. Комутовані та некомутовані (оперативні). Комутовані канали є тимчасовими, вони формуються лише для тимчасової передачі даних. Тобто після закінчення передачі вони зачиняються. Некомутовані канали (також їх називаються оперативними) утворюються на тривалий час, з певними виділеними постійними характеристиками.

Сьома – за швидкістю передачі сигналів. Існує два типи каналів за швидкістю передавання сигналу. Це середньошвидкісні (приблизно від 2,4 Кбіт/с до 9,6 Кбіт/с). Вони використовуються в телефонних мережах зв'язку, на телефонних станціях більш нового зразку швидкість може варіюватися від 14 Кбіт/с до 56 Кбіт/с. Та високошвидкісні канали передачі даних, також називаються широкосмуговими. В них швидкість передавання вища за 56 Кбіт/с. Для передачі даних в таких каналах використовуються екрановані, коаксіальні, оптоволоконні кабелі, та радіоканали.

1.3 Методи модуляції та кодування сигналу

Модуляція – це процес, в якому дані з вхідних пристроїв (мікрофони, камери, датчики тощо) перетворюються на електричний або цифровий сигнал, та утворюють певну функцію [1]. Дискретні способи модуляції засновані на дискретизації неперервних процесів і за амплітудою, і за часом. Аналогова модуляція застосовується під час передачі дискретних даних каналами з вузькою смугою частот, де типовим представником є канал тональної частоти, наданий для користування усім користувачам загальним телефонних мереж. Канал тональної частоти охоплює частоти від 300 до 3400 Гц, тобто смуга пропускання такого каналу дорівнює 3100 Гц (рис. 1.1).



Рисунок 1.1 - Канал тональної частоти

Для задовільної якості передачі голосу людини та більшості звуків такого діапазону вистачає, хоч і світ має набагато ширший спектр частот. Таке значне обмеження смуги пропускання викликано використанням засобів ущільнення та комутації каналів в телефонних мережах. Апаратура, що виконує функції модуляції несучої синусоїди на стороні що передає сигнали, та демодуляції на прийомній стороні називається модемом (пристрій модулятор-демодулятор).

Аналогова модуляція є способом фізичного кодування інформації, за використання якого дані кодуються зміною амплітуди, фази або частоти синусоїди несучої частоти.

Основними способами аналогової модуляції є амплітудна, частотна та фазова модуляції. Нижче наведено зображення з показовими прикладами модуляцій та на першій діаграмі (під номером «а») показана послідовність бітів вихідної інформації, там зображені потенціали високого рівня для логічної одиниці та потенціал нульового рівня для логічного нуля. Цей спосіб кодування відомий як «потенційний код», що часто використовують при передачі даних між блоками різноманітних електронних пристроїв. Амплітудна модуляція (номер «б») – де для логічної одиниці обирається один рівень амплітуди синусоїди несучої частоти, для логічного нуля – інший. Такий спосіб досить рідко використовується в чистому вигляді на практиці через низьку завадостійкість, проте часто застосовується в міксі з іншим видом модуляції – фазовою. В частотній модуляції (номер «в») значення 0 та 1 вихідних даних передаються синусоїдами різної частоти, f_0 і f_1 . Цей спосіб модуляції не потребує ніяких складних схем модемів та зазвичай застосовується в низькошвидкісних модемах, які розраховані на роботу в режимах 300 або 1200 біт/с. При фазовій модуляції

(номер «г») значенням даних 0 та 1 відповідають сигнали однакової частоти, проте з різною фазою, наприклад 90 і 270 градусів або 90, 180, 270, 360 градусів.

У високошвидкісних модемах зазвичай методи модуляції комбінуються, фазова з амплітудною наприклад. Найбільш розповсюдженими комбінованими методами модуляції є методи квадратурної амплітудної модуляції. Такі методи засновані на поєднанні фазової модуляції з 8 значеннями величин зміщення фази та амплітудної модуляції з 4 рівнями амплітуди. Проте, не всі з 32 можливих комбінацій сигналу використовуються. В кодах Трелліса, наприклад, допускаються лише від 6 до 8 комбінацій для подання вихідних даних, а інші комбінації – заборонені. Ця надмірність кодування використовується модемом для розрізнення помилкових сигналів, які утворюються внаслідок перекручувань через перешкоди на телефонних каналах, що комутуються значними за амплітудою і тривалими за часом.

В імпульсній модуляції в якості носія модульованих сигналів постають послідовності імпульсів (здебільшого – прямокутної форми). Ці послідовності заповнюються високочастотними коливаннями в бездротових системах передачі даних в радіозв'язку, створюючи тим самим подвійну модуляцію. Такі види модуляції застосовуються для передачі дискретизованих даних. В залежності від вигляду модульованого параметра імпульсну модуляцію поділяють на частотно-імпульсну модуляцію (ЧІМ) $f=1/T$ -var, амплітудно-імпульсну модуляцію (АІМ) A -var, широтно-імпульсну модуляцію (ШІМ) τ 2-var та фазово-імпульсну модуляцію (ФІМ) τ 1-var.

Для імпульсів прямокутної форми найбільш широко використовуються АІМ (рис. 1.2) та ШІМ.

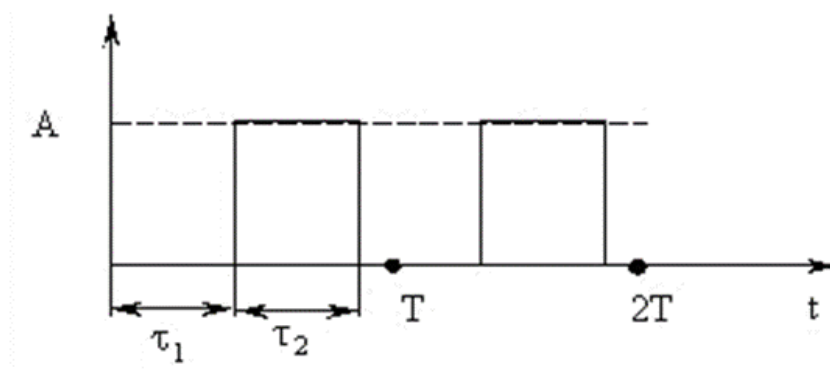


Рисунок 1.2 – Амплітудно-імпульсна модуляція

Модуляцію також розрізняють за характером зв'язку між модулюючим (вхідним) сигналом та модульованим параметром на модуляцію I-го та II-го роду (рис. 1.3).

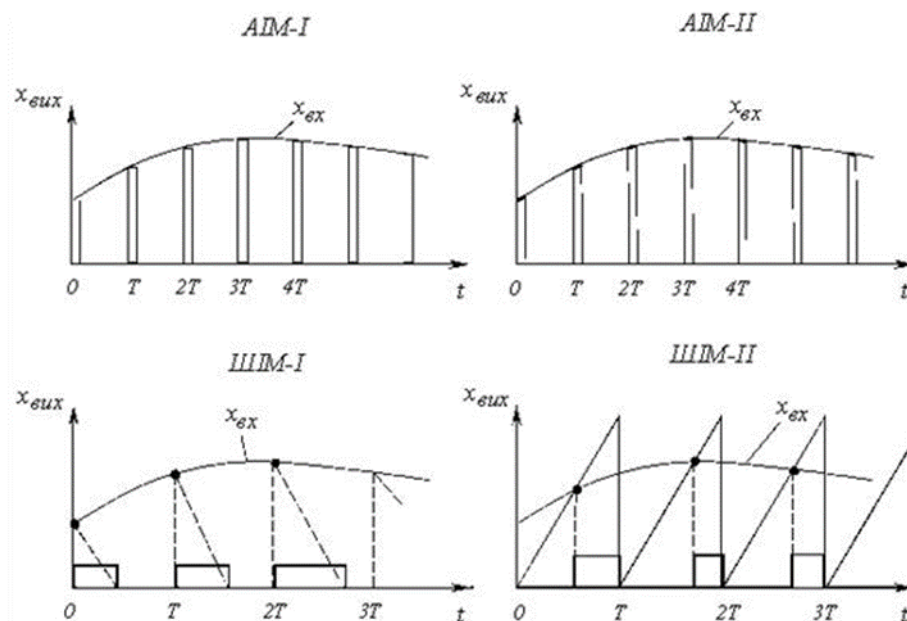


Рисунок 1.3 – Модуляція I та II-го роду

Коди з високою завадостійкістю є одним із найрезультативніших інструментів для гарантування достовірності переданої чи збереженої дискретної інформації. У цій сфері активно розвивається спеціалізована теорія кодування з урахуванням впливу завад. Відомий вчений Клод Шеннон сформулював теорему, що стосується передавання дискретних повідомлень через зашумлені канали. Вона доводить можливість забезпечення надзвичайно низької ймовірності декодувальних помилок за умови використання відповідного кодувального підходу [3].

Під завадостійкими кодами мають на увазі такі, що здатні або виявляти, або виявляти та виправляти помилки, які виникають унаслідок дії шумів. Надмірність, яку вводять у кодові комбінації, і є основою стійкості до завад — тобто не всі елементи комбінації відповідають за передання безпосередньої інформації.

Усі типи завадостійкого кодування умовно класифікують на дві великі групи: блокові (де кожен інформаційний блок має сталу або змінну структуру) та неперервні (або рекурентні, ланцюгові), які працюють із безперервними потоками даних. У блокових схемах кожен сегмент інформації перетворюється на певну кількість сигналів, що кодуються й декодуються окремо. Рівномірні блокові коди

мають фіксовану довжину комбінації, тоді як у нерівномірних вона варіюється. Через складність реалізації останні майже не використовуються на практиці.

Залежно від способу побудови надмірності коди також поділяються на роздільні та нероздільні. У першому випадку інформаційні й перевірочні символи чітко розмежовані: перші несуть зміст, другі забезпечують виявлення й усунення помилок. Такі системи часто називають n -кодами, де n - загальна кількість символів, а k - число інформаційних. Нероздільні коди, навпаки, не передбачають чіткого поділу між типами символів, тому поширення мають значно менше.

Усередині роздільних блокових кодів розрізняють систематичні та несистематичні. Переважна більшість кодів у реальних системах є систематичними: перевірочні символи в них формуються шляхом виконання лінійних обчислень над інформаційною частиною. Зокрема, у двійкових кодах перевірочний символ обирають так, щоб сума з деякими інформаційними символами за модулем два дорівнювала нулю. Процес декодування полягає в перевірці певних груп символів на парність, що дозволяє виявити наявність та можливе місце помилки.

1.4 Особливості передачі аудіоінформації через телекомунікаційні канали

Однією з ключових особливостей передачі аудіосигналів через телекомунікаційні мережі є забезпечення рівноваги між якістю переданого звуку, мінімальною затримкою та безпекою інформації. Підвищення якості аудіопотоку нерідко веде до збільшення затримок, що своєю чергою може спричинити втрату синхронізації між учасниками комунікації. Щоб уникнути цього, варто уникати перевантаження каналів зв'язку. При цьому важливо враховувати і ризики, пов'язані з втратою частини пакетів чи можливим зовнішнім втручанням у процес передавання.

Вище зазначені нюанси особливо актуальні для таких форматів комунікації, як відео- та аудіоконференції в реальному часі, телефонні дзвінки через VoIP, відеовиклики через месенджери, а також онлайн-трансляції. У мережах Ethernet, які складають основу сучасного Інтернету, відсутні вбудовані механізми точного контролю часу доставки пакетів та перевірки цілісності даних відповідно до їхньої актуальності. Як результат — затримки під час передавання потокового контенту можуть перевищувати прийнятні межі для кінцевого користувача.

Ефективне рішення цієї проблеми, що дозволяє уникнути додаткових витрат на оновлення мережевого обладнання, полягає у використанні методів адаптивного потокового стиснення, заснованих на принципах завадостійкої передачі даних. Основна ідея цього підходу полягає у формуванні інформаційних пакетів, які можна частково або з помилками відтворити без необхідності їх повторної відправки сервером. Це суттєво зменшує затримки і втрати.

Для підвищення стійкості до помилок використовуються коди БЧХ (коди Боуза-Чоудхурі-Хоквінгема) [4], а також особливим чином модифікована структура пакету, яка містить додаткові службові дані. Завдяки їм можливо частково реконструювати втрачену інформацію і при цьому зберегти прийнятну якість аудіо та відео. За результатами випробувань, артефакти, що виникають у такому підході, значно менш помітні для слухача або глядача, ніж ситуації, коли звук чи відео перериваються через очікування повторної передачі втраченого пакету.

Особливо ефективною ця стратегія є в умовах українських мереж з асиметричною архітектурою, де запити на повторну передачу потокового контенту взагалі є недопустимими через обмежену смугу пропускання. Значна увага в цьому підході також приділяється алгоритмам аудіокомпресії. Враховуючи важливу роль звукового супроводу в освітніх онлайн-заходах, було впроваджено методи перетворення динамічного діапазону відповідно до особливостей сприйняття мовлення людиною. Це дало змогу ефективно передавати голос навіть у вузькосмугових мережах без істотної втрати розбірливості та якості [5].

2 МЕТОДИ ОБРОБКИ ТА ПЕРЕДАЧІ АУДІОДАНИХ

2.1 Оцифрування звуку: дискретизація, квантування, кодування

Головною метою цифрової обробки звукових сигналів є перетворення аналогових звукових коливань, які реєструються вхідними пристроями (наприклад, мікрофонами або аудіоінтерфейсами), у цифровий формат — тобто у числову послідовність, яку можна передати мережею до отримувача. Після надходження таких аудіоданих вони обробляються за допомогою аналого-цифрового перетворювача (АЦП), який входить до складу приймача.

Процес аналого-цифрового перетворення передбачає трансляцію безперервного сигналу, описаного функцією часу $U(t)$, у серію числових значень $\{U'(t_j)\}$, де $j = 0, 1, 2, 3$ і т. п., що відповідають визначеним моментам часу. Цей процес охоплює два основні етапи: дискретизацію, що забезпечує розбиття сигналу на часові інтервали, та квантування — перетворення амплітуд сигналу у фіксовані рівні згідно з визначеною шкалою [1].

Дискретизація сигналу — це процес, при якому функція, що залежить від неперервної змінної, перетворюється на функцію дискретної змінної. Така дискретна функція дозволяє з певним рівнем точності відтворити оригінальний неперервний сигнал. У більшості випадків дискретні значення є результатами квантування сигналу в заданих точках координатної шкали.

Систему дискретного часу можна визначити як алгоритм, який обробляє вхідну послідовність значень $s(k)$ та формує відповідну вихідну послідовність $y(k)$. Ці системи можуть бути як лінійними, так і нелінійними, а також або інваріантними до зсуву в часі, або такими, що змінюються в часі. Якщо система є лінійною та часово-інваріантною, то вона задовольняє принцип суперпозиції — тобто відповідь системи на суму сигналів дорівнює сумі відповідей на кожен із них, — а також демонструє стабільну реакцію на зсув: якщо вхідний сигнал зміщується в часі, то й вихідна реакція також зміщується на той самий проміжок.

Суть дискретизації аналогових сигналів полягає в тому, що неперервність у часі аналогової функції $s(t)$ замінюється послідовністю коротких імпульсів, амплітудні значення яких визначаються за допомогою вагових функцій, або безпосередньо вибітками (відліками) миттєвих значень сигналу $s(t)$ у моменти

часу T_n . Подання сигналу $s(t)$ на інтервалі T є сукупністю дискретних значень c_n та записується у вигляді:

$$(c_1, c_2, \dots, c_n) = A[s(t)],$$

де A – оператор дискретизації.

Запис операції відновлення сигналу $s(t)$:

$$s(t) = B [(c_1, c_2, \dots, c_n)].$$

Вибір операторів A та B визначається необхідною точністю відновлення сигналу. Найбільш простими є лінійні оператори. У загальному випадку:

$$c_n = \int q_n(t) s(t) dt,$$

де $q_n(t)$ – система вагових функцій.

У більшості практичних випадків не вимагається абсолютно точне відтворення сигналу. Натомість зазвичай ставиться задача забезпечення найменшого можливого обсягу інформації, за якого зберігається можливість реконструкції сигналу у безперервному вигляді з допустимим рівнем похибки. Такий підхід є особливо важливим у сфері дистанційного збору та обробки даних, під час передачі інформації каналами зв'язку, а також при її архівації на тривале зберігання. Одним з ефективних підходів до вирішення цього завдання є дискретизація сигналів, заснована на критерії максимально допустимого відхилення.

В процесі дискретизації за критерієм найбільшого відхилення задається допустиме значення похибки відновлення сигналу σ . При відновленні сигналу неперервна функція $s(t)$ апроксимується, як правило, степеневими поліномами n -го порядку. Похибка відновлення функції $s(t)$ поліномом $s_a(t)$ визначається залишковим членом $L(t)$:

$$L(t) = s(t) - s_a(t).$$

Крок дискретизації вибирається з умови забезпечення $L(t) < \sigma$ по всьому інтервалу визначення функції $s(t)$. Як правило, динаміка функції $s(t)$ може суттєво змінюватися в різні моменти часу за інтервалом реєстрації, при цьому крок дискретизації також може змінюватися, за умови не перевищення заданої похибки на кожному кроці. При встановленому значенні σ зменшення числа відліків забезпечується підвищенням степені апроксимувального багаточлена. На практиці, як правило, обмежуються сходящимим, лінійним та параболічним апроксимаційними поліномами відповідно нульового, першого і другого ступенів.

Як було сказано вище, як інтерполювальні багаточлени використовують багаточлени Лагранжа, а як екстраполювальні багаточлени використовують багаточлени Тейлора. Для багаточлена Тейлора нульового ступеня умови відновлення сигналу практично не відрізняються від багаточлена Лагранжа, за винятком напряму (від поточного зареєстрованого відліку і вперед до t). Для багаточленів Тейлора більш високих степенів при відновленні сигналу крім відліку $s(t)$ використовуються також відповідні значення похідних в точці відліку. Відновлення сигналу багаточленами Тейлора відбувається без затримки в часі. Однак при використанні багаточленів вище нульового ступеня для точного відновлення сигналу порівняно з інтерполяційними методами потрібно в два рази вища частота дискретизації.

Частота рівномірної дискретизації інформації розраховується за граничними значеннями частотних характеристик сигналів. Адаптивна дискретизація орієнтована на динамічні характеристики сигналу, що дозволяє забезпечити його відновлення при мінімальній кількості вибірок. В основі принципів адаптивної дискретизації лежить спостереження за поточною похибкою відновлення сигналу. Найбільш широке застосування отримали алгоритми дискретизації з адаптацією по довжині інтервалу апроксимації. Суть дискретизації полягає в послідовному нарощенні інтервалу апроксимації з неперервним порівнянням сигналу $s(t)$ з відтворювальною функцією $s_a(t)$. При досягненні заданого значення нарощення інтервалу зупиняють, і проводиться відлік значень $s(t)$, тобто дискретизація є нерівномірною. Для відтворення сигналів нерегулярної дискретизації зазвичай використовують степеневі алгебраїчні поліноми нульового чи першого ступеня в інтерполяційному чи в екстраполяційному варіантах.

Найбільш простою є техніка адаптивної дискретизації з використанням багаточлена нульового ступеня. На момент t_j початку кожного інтервалу проксимувальний поліном $s_a(t)$ приймається рівним $s(t_i)$, обчислюється поточна різниця

$$L(t) = s(t) - s_a(t),$$

та проводиться порівняння її значення із заданим значенням. При фіксуванні рівності $L(t)$ та заданого значення проводиться черговий відлік і починається наступний інтервал.

При використанні апроксимувального багаточлена першого ступеня обчислюється значення

$$s_a(t) = s(t_j) + s'(t_j),$$

де $s'(t)$ – похідна сигналу.

Момент чергового відліку визначається виконанням рівності

$$s(t) - s(t_j) - s'(t_j) = \sigma.$$

Варто мати на увазі, що даний алгоритм неефективний при наявності високочастотних завад, до яких досить чутлива операція диференціювання.

Дискретизація випадкових сигналів належить до класу непростих задач аналізу випадкових сигналів. Тому обмежимося лише згадуванням основних особливостей такого аналізу на прикладі дискретизації квантованих за рівнем випадкових сигналів.

Таким чином, для обчислення оптимального кроку рівномірної дискретизації достатньо знати лише величину інтервалу квантування та кореляційну функцію вимірюваного процесу.

При оперативних змінах із запам'ятовуванням результату лише останньої відлікової операції оптимальна тривалість наступного кроку практично не залежить від цього результату і дорівнює оптимальному кроку рівномірної дискретизації вимірюваного процесу.

В околі точки оптимуму існує деякий досить широкий інтервал приблизно рівнозначних значень кроку дискретності Δt . Так в межах $\pm 10\%$ відхилень Δt від оптимального значення Δt зміни імовірнісного показника якості $p(\Delta t)$ не перевищують 1%.

Оптимальний крок дискретизації визначається так

$$R(\Delta t_0) = \sqrt{1 - 0,9 \left(\frac{\Delta x}{\sigma}\right)^2},$$

де Δx – величина кванта; σ – середньоквадратичне відхилення процесу.

Дана формула дозволяє визначити оптимальний крок дискретизації безпосередньо на графіку кореляційної функції. На практиці переважно використовують інше співвідношення:

$$\Delta t_0 = 0,95 \frac{\Delta x}{\omega \sigma},$$

$$\text{де } \omega = \sqrt{-\left.\frac{d^2 R(t)}{dt^2}\right|_{t=0}}.$$

Усереднена за усіма номерам k похибка визначається як

$$\delta = 0,01 \left(\frac{\Delta x}{\sigma}\right)^4.$$

Варто зазначити, що при $\frac{\Delta x}{\sigma} < 0,4-0,5$ розрахункову похибку оптимального значення критерію якості вимірювальної системи можна оцінити однією похибкою, яка складає приблизно 0,5%.

Також, однією з найпоширеніших є рівномірна форма дискретизації, в основі якої лежить теорема відліків. Згідно з цією теоремою як коефіцієнти a_j потрібно використовувати миттєві значення сигналу $U(t_j)$ в дискретні моменти часу $t_j = j\omega t$, а період дискретизації вибирати з умови: $t = 1/2F_m$, де F_m – максимальна частота спектра сигналу, що перетворюється. Тоді отримаємо теорему відліків:

$$U(t) = \sum_{j=-\infty}^{\infty} U(j\omega\Delta t) \frac{\sin[2\pi F_m(t-j\Delta t)]}{2\pi F_m(t-j\Delta t)}.$$

Для сигналів із суворо обмеженим спектральним діапазоном наведене співвідношення є точною рівністю. Утім, спектри більшості реальних сигналів лише асимптотично наближаються до нуля, не обриваючись різко. При рівномірній дискретизації таких сигналів у процесах обробки інформації можуть виникати характерні високочастотні спотворення, зумовлені ефектом вибірки. Щоб зменшити ці небажані спотворення, існує два основні підходи: підвищення частоти дискретизації або застосування фільтра нижніх частот до входу аналого-цифрового перетворювача (АЦП), з метою попереднього обмеження спектру сигналу перед його цифровізацією.

У загальному випадку вибір частоти дискретизації буде залежати також від вигляду функції $f_j(t)$, що використовується в першій формулі розділу та допустимого рівня похибок, які виникають при відновленні початкового сигналу за його відліками. Усе це необхідно враховувати при виборі частоти дискретизації, яка визначає необхідну швидкодію АЦП. Часто цей параметр задають розробнику АЦП.

Розглянемо більш детально роль аналого-цифрового перетворювача (АЦП) у процесі дискретизації сигналів. Для сигналів із вузьким спектральним діапазоном операцію дискретизації часто можна реалізувати безпосередньо за допомогою АЦП, поєднуючи її з квантуванням. Важливою особливістю такого підходу є те, що через скінченну тривалість перетворення та непередбачуваність точного моменту його завершення, що, у свою чергу, залежить від характеристик вхідного сигналу, неможливо забезпечити точне зіставлення між часовими мітками та відповідними значеннями відліків. У випадку обробки змінних у часі сигналів це призводить до появи специфічних похибок динамічного характеру. Для їх кількісної оцінки вводиться поняття апертурної похибки, яку зазвичай описують через апертурний час.

Апертурний час t_a – це час, протягом якого зберігається невизначеність між значенням вибірки та часом, до якого вона відноситься [1].

Апертурна невизначеність проявляється як похибка при фіксації значення сигналу у визначений момент часу або як похибка визначення самого моменту вимірювання при відомому значенні сигналу. У випадку рівномірної дискретизації така невизначеність спричиняє появу амплітудних помилок, які

отримали назву апертурних. Їхня величина відповідає зміні сигналу протягом періоду апертурного часу.

З іншого боку, апертурну невизначеність можна розглядати як варіацію істинних моментів відліку сигналу навколо ідеальних, рівномірно розподілених на часовій осі. У результаті цього виникає ефект коливання періоду дискретизації, що порушує умови теореми дискретизації і призводить до формування вже згаданих апертурних похибок у системах цифрової обробки.

Цифрове представлення окремих значень аналогового сигналу передбачає виконання процесів квантування та кодування. Ці операції можуть реалізовуватись за допомогою різних методів: послідовного, паралельного чи комбінованого (послідовно-паралельного) способу наближення числового значення до вимірюваної величини.

Дискретизація аналогових сигналів з перетворенням в цифрову форму пов'язана з квантуванням сигналів. Суть квантування полягає в заміні нескінченної множини можливих значень функції, в загальному випадку випадкових, скінченною множиною цифрових відліків, і виконується округленням миттєвих значень вхідної функції $s(t)$ в моменти часу t до найближчих значень $s_i(t_i) = n_i r$, де r – крок квантування шкали цифрових відліків. Квантування з постійним кроком r називається рівномірним. Математично операція квантування може бути виражена формулою:

$$s_i(t_i) = \left[\frac{s(t_i)}{r} + \frac{1}{2} \right] r,$$

де результат обчислення в дужках округляється до цілого значення.

Квантування за амплітудою – це процес заміни реальних (виміряних) значень амплітуди сигналу значеннями, наближеними з деякою точністю. Кожен з $2N$ можливих рівнів називається рівнем квантування, а відстань між двома найближчими рівнями квантування називається кроком квантування. У випадку лінійного розбиття амплітудної шкали на рівні, квантування називають лінійним (однорідним). На рис. 2.1. наведено приклад такого квантування.

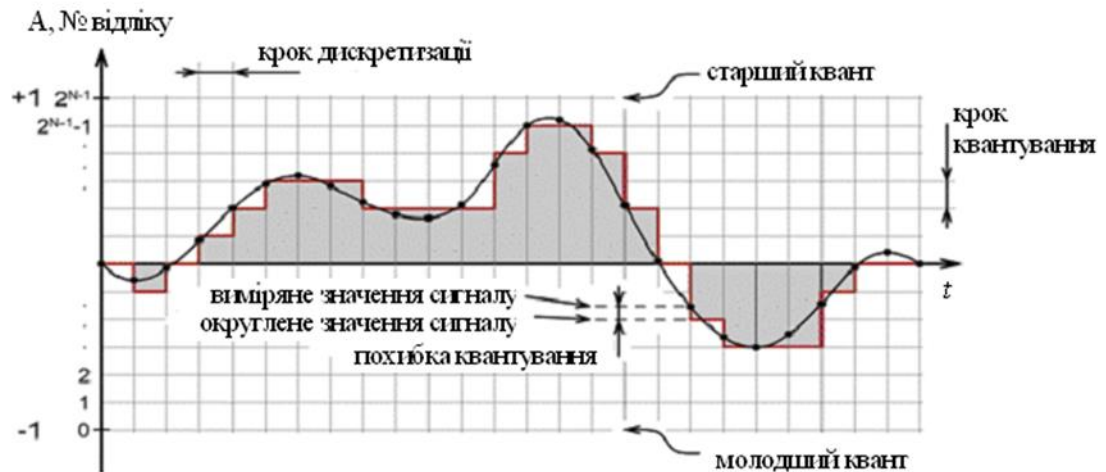


Рисунок 2.1 - Приклад квантування за амплітудою

Результатом такого оцифрування є ступінчатий сигнал, що складається з прямокутників, кожен з яких має ширину, рівну величині кроку дискретизації та висоту, рівну виміряному значенню амплітуди сигналу. Очевидно, що точність округлення залежить від вибраної кількості ($2N$) рівнів квантування, яка, в свою чергу, залежить від кількості бітів (N), відведених для запису значення амплітуди. Число N має назву розрядність квантування (маючи на увазі кількість розрядів, тобто бітів, в кожному слові), а здобуті в результаті округлення значень амплітуди числа – відліками чи семплами (от англ. «sample» – «замір»). Вважається, що квантування з

При квантуванні сигналів у великому динамічному діапазоні значень крок квантування може бути й нерівномірним, наприклад, логарифмічним, тобто пропорційним логарифму значень вхідного сигналу. Установлений діапазон шкали квантування від s_{min} до s_{max} і крок квантування r визначають число поділок шкали $N = \frac{(s_{min} - s_{max})}{r}$ і відповідно цифрову розрядність квантування. В результаті дискретизації й квантування неперервна функція $s(t)$ замінюється числовою послідовністю $\{s(kt)\}$. Похибка округлення $\varepsilon_l = s(t_l) - s_i(kt)$ знаходиться в межах $-kt/2 < \varepsilon < kt/2$ та має назву шум квантування. Необхідна точність квантування оцінюється за впливом на подальшу обробку сигналів виниклого шуму квантування.

При достатньо малому кроці квантування будь-яке значення в його межах може вважатися рівномірним, при цьому значення ε розподілені за рівномірним законом.

Відповідно дисперсія та середнє квадратичне значення шуму квантування:

$$\sigma^2 = kt/12, \quad \varepsilon \approx 0,3kt.$$

При заданому рівні шуму квантування з використанням попередньої формули неважко визначити допустиме значення кроку квантування.

Вхідний сигнал містить, як правило, адитивну суміш власне сигналу $s(t)$ та завади $q(t)$ з дисперсією відповідно σ_q^2 . Якщо завади не корельовані із сигналом, то після квантування сумарна дисперсія шумів:

$$\sigma_{\Sigma}^2 = \sigma_q^2 + \sigma^2.$$

На практиці крок квантування обирається таким, щоб не відбувалось помітної зміни відношення між сигналом та шумом – $\sigma_y^2 \gg \sigma^2$.

Перейдемо до розгляду позиційних систем числення. Позиційна система числення з основою P має P цифр C_0, C_1, \dots, C_{P-1} , що позначають натуральні числа від 0 до $P-1$. Ці записи та позначені ними числа – Значення цих записів (чисел) має назву «однорозрядні».

У програмуванні широко застосовується шістнадцяткова система, в якій перші 10 цифр арабські, а наступні шість, А, В, С, D, Е, F, позначають числа, десятковий запис яких 10, 11, 12, 13, 14, 15, відповідно.

Число P у P -ковій системі позначається дворозрядним записом C_0C_1 , число $P+1$ – записом C_1C_1 тощо. Наприклад, 10, 11, ... , 99 у десятковій системі, 10, 11 у двійковій, 10, 11, ... , 1F, 20, ... , FF у 16-ковій. Наприклад, 100, 101, ... , 999 у десятковій системі, 100, 101, 110, 111 у двійковій, 100, 101, ... , FFF у 16-ковій.

Наприклад, двійковий запис (10011)₂ позначає число, яке в десятковому записі має вигляд $1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 19$. 16-ковий запис (1BC)₁₆ позначає десяткове $1 \times 16^2 + 11 \times 16 + 12 = 444$.

Права цифра в записі числа позначає кількість одиниць і називається молодшою, ліва – кількість чисел P^k і називається старшою.

Ми звикли до десяткового подання чисел, і саме воно, головним чином, використовується в програмах, але в комп'ютері числа, як правило, подаються в

двійковій системі. Таким чином, виникає необхідність створювати двійкове подання числа за його десятковим записом і навпаки.

За P -ковим записом натурального числа N можна побудувати десяткове подання, обчисливши значення полінома за допомогою операцій множення та додавання в десятковій системі.

P -кове подання чисел, менших 1, має вигляд $0, a_{-1}, 0, a_{-2}, \dots, 0, a_{-i}$, де a_{-i} – P -кові цифри. Цей запис позначає дійсне число – значення виразу $a_{-1}P^{-1} + a_{-2}P^{-2} + \dots + a_{-i}P^{-i}$.

Якщо основа P має прості дільники, відмінні від 2 і 5, то число зі скінченним P -ковим записом зображується нескінченним, але періодичним десятковим дробом. Якщо ж простими дільниками P є тільки 2 і 5, то й десятковий дріб скінченний.

В процесі цифрового кодування дискретної інформації використовуються потенціальні та імпульсні коди.

У потенціальних кодах для подання логічних одиниць і нулів застосовується тільки значення потенціалу сигналу, а його перепади, що формують закінчені імпульси, до уваги не беруться. Імпульсні коди дозволяють подати двійкові дані або імпульсами певної полярності, або частиною імпульсу – перепадом потенціалу певного напрямку.

До методів цифрового кодування існують деякі важливі вимоги. Так, при використанні прямокутних імпульсів для передачі дискретної інформації необхідно обрати такий спосіб кодування, який одночасно вирішував кілька завдань, а саме мав при одній і тій же бітовій швидкості найменшу ширину спектра результуючого сигналу; забезпечував синхронізацію між передавачем та приймачем; мав здатність розпізнавати помилки; мав низьку вартість реалізації.

Звуження спектра сигналу дозволяє підвищити ефективність передавання даних по каналу з фіксованою смугою пропускання, забезпечуючи більшу швидкість передавання на тому самому фізичному середовищі. Водночас часто ставиться вимога відсутності постійної складової сигналу, тобто запобігання проходженню постійного струму між джерелом сигналу та приймачем. Це обумовлено використанням трансформаторних схем, що здійснюють гальванічну розв'язку, та не пропускають постійну компоненту.

Синхронізація між джерелом та приймачем сигналу є критично важливою для правильного зчитування даних у відповідні моменти часу. Ця задача

особливо актуальна в мережевих системах, де пристрої можуть бути значно віддаленими один від одного. У випадку близько розташованих елементів, як-от модулі всередині комп'ютера або комп'ютер та периферійний пристрій, часто використовується окрема лінія синхронізації (рис. 2.2), яка передає тактові імпульси. Завдяки цьому зчитування інформації відбувається чітко в моменти надходження тактового сигналу.

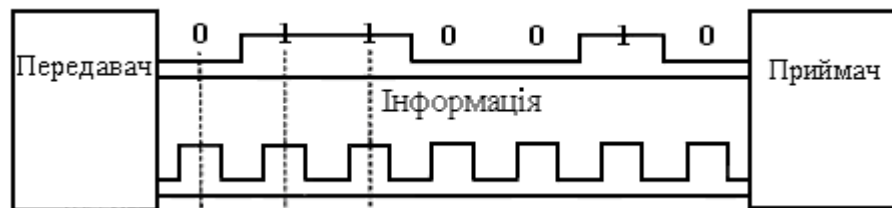


Рисунок 2.2 - Синхронізація приймача та передавача на невеликих відстанях

Використання схем із тактовими імпульсами в комп'ютерних мережах часто викликає труднощі, пов'язані з неоднорідністю властивостей провідників у кабельних лініях. На значних відстанях спостерігається нерівномірність поширення сигналу, що може призвести до зміщення тактового імпульсу відносно сигналу даних настільки, що окремий біт буде або пропущено, або зчитано двічі. Ще одним фактором, який спонукає до відмови від окремих тактуючих ліній у мережах, є потреба в зменшенні кількості провідників у кабелі, що дозволяє знизити вартість інфраструктури.

З цієї причини в мережевих технологіях активно застосовуються сигнали із вбудованою синхронізацією, які самі по собі містять часову інформацію для коректного зчитування даних. Кожна раптова зміна рівня сигналу (так званий фронт) може бути використана для вирівнювання синхронізації між передавачем і приймачем. У разі використання несучих сигналів синусоїдальної форми зміни амплітуди надають приймачу можливість визначити оптимальний момент для обробки інформації, тим самим забезпечуючи самосинхронізацію.

Оскільки фізичний рівень обмежений у можливості корекції помилок, ця функція зазвичай передається протоколам вищих рівнів — таким як каналний, мережний, транспортний або прикладний. Проте, виявлення помилок ще на фізичному рівні дозволяє зекономити час, оскільки дозволяє відразу відхилити кадр із помилками, не чекаючи завершення його передачі до буфера.

2.2. Стиснення аудіоданих

Причинами необхідності стиснення аудіоданих є економія пам'яті при збереженні аудіоінформації та низька пропускна здатність каналів для передачі цифрової інформації на відстань.

Будь-яке стиснення інформації призводить до погіршення її якості. Проте в процесі еволюції людський слух навчився пристосовуватися до певних видів перешкод, не помічаючи їх присутності в отриманій звуковій інформації. Слух людини має логарифмічну чутливість, тобто рівень сприйманого шуму залежить від загального рівня сигналу, також він є нелінійною системою. Тому Міжнародний консультативний комітет з телеграфу та телефонії (ССІТТ) при розробці рекомендації G.711 використовував підхід стиснення звукових зразків (цифрових кодів), заснований на властивостях людського слуху [6].

Рекомендації ССІТТ G.711 впроваджують два алгоритми для перетворення звукових сигналів: μ -law і A-law. Кодування μ -law широко використовується в США, Канаді та Японії, а кодування A-law використовується в Європі. Обидва ці алгоритми перетворюють оригінальні зразки вихідної послідовності сигналу РСМ у зразки байтів. Кожен зразок вихідної послідовності перетворюється в один байт.

У рекомендаціях ССІТТ μ -law і A-law наведені в табличній формі. Для представлення числа виділено 3 поля: поле знаку, поле мантиси та поле порядку (рис. 2.3). Поле порядку містить ступінь, до якого слід підвести число 2, щоб помножити результат цієї операції на мантису, щоб отримати справжнє абсолютне значення цього числа у форматі з фіксованою комою.

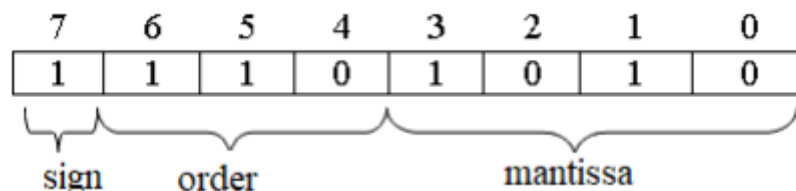


Рисунок 2.3 - Приклад байтової структури для μ -law і A-law

A-law - це алгоритм стиснення з втратами, який використовується для стиснення аудіоданих, який можна виразити формулами:

$$y = \frac{Ax}{1+\ln A}, \text{ for } x \leq \frac{1}{A} \quad (2.1)$$

$$y = \frac{1+\ln Ax}{1+\ln A}, \text{ for } \frac{1}{A} \leq x \leq 1 \quad (2/2),$$

де A – параметр стиснення. В Європі $A=87,6$.

Схема роботи A-law може бути такою:

1) коли число від'ємне, воно змінюється, і припускається, що $s = 0$, інакше $s = 1$;

2) 16-бітове число перетворюється на 8-бітне згідно з табл. 2.1 (s - знаковий біт; зірочки вказують на біти, які втрачаються під час стиснення);

Таблиця 2.1 - Перетворення чисел A-law

Початковий номер	Стиснуте число
s000 0000 wxyz ****	S000 wxyz
s000 0001 wxyz ****	S001 wxyz
s000 001w xyz * ****	S010 wxyz
s000 01wx yz ** ****	S011 wxyz
s000 1wxy z *** ****	s100 wxyz
s001wxyz **** ****	s101 wxyz
s01w xyz * **** ****	s110 wxyz
s1wx yz ** **** ****	s111 wxyz

3) кожен другий біт потрібно інвертувати, починаючи з правого боку (табл. 2.2). μ -law - це аналоговий алгоритм стиснення, який використовується в системах цифрового зв'язку в Північній Америці та Японії для зміни динамічного діапазону аналогового мовного сигналу перед оцифруванням, що виражається формулою:

$$y = \frac{\ln(1+\mu|x|)}{\ln(1+\mu)}, \text{ де } \mu = 255. \quad (2.3)$$

Під час кодування звукових даних за допомогою алгоритмів типу μ -law або A-law, кожен окремий семпл обробляється автономно. Таке стиснення, базуючись на психоакустичних властивостях слухової системи людини, зменшує обсяг інформації приблизно вдвічі. Цей засіб підходить для широкого спектра звукових сигналів, включно з музичними, не спричиняючи помітного зниження якості. Проте, така форма кодування не забезпечує суттєвого зменшення обсягу

даних, оскільки не враховує внутрішню структуру самого сигналу. Для досягнення вищого рівня стискання застосовують методи, що кодують не самі значення сигналу, а різницю між послідовними відліками – як, наприклад, у дельта-модуляції.

Таблиця 2.2 - Перетворення чисел μ -law

Початковий номер	Стиснуте число
від +8158 до + 4063 через 16 інтервалів з 256	0x80+інтервал номер
від +4062 до + 2015 через 16 інтервалів з 128	0x90+інтервал номер
від +2014 до + 991 через 16 інтервалів з 64	0xA0+інтервал номер
від +990 до +479 через 16 інтервалів з 32	0xB0+інтервал номер
від +478 до +223 через 16 інтервалів з 16	0xC0+інтервал номер
від +222 до +95 через 16 інтервалів з 8	0xD0+інтервал номер
від +94 до +31 через 16 інтервалів з 4	0xE0+інтервал номер
від +30 до +1 через 15 інтервалів з 2	0xF0+інтервал номер
0	0xFF
-1	0x7F
від -31 до -2 через 15 інтервалів з 2	0x70+інтервал номер
від -95 до -32 через 16 інтервалів з 4	0x60+інтервал номер
від -223 до -96 через 16 інтервалів з 8	0x50+інтервал номер
від -479 до -224 через 16 інтервалів з 16	0x40+інтервал номер
від -991 до -480 через 16 інтервалів з 32	0x30+інтервал номер
від -2015 до -992 через 16 інтервалів з 64	0x20+інтервал номер
від -4063 до -2016 через 16 інтервалів з 128	0x10+інтервал номер
від -8159 до -4064 у 16 інтервалах по 256	0x00+ номер інтервалу

Подальший розвиток цієї ідеї привів до створення більш ефективних алгоритмів, таких як адаптивна диференціальна імпульсно-кодова модуляція (ADPCM). У цьому випадку система формує передбачуване значення кожного наступного семплу й передає лише величину помилки прогнозу. Такий підхід дозволяє зменшити кількість бітів, які використовуються на одну вибірку, до 4, що відповідає пропускну здатності 32 кбіт/с при збереженні прийнятної якості відтворення звуку [1].

Одним із найбільш розвинених методів є лінійне прогнозує кодування (LPC). У цій моделі мова представлена не у вигляді сирих семплів, а через параметри фільтра, що імітує мовний тракт, та параметри сигналу, який цей фільтр активує. Основним інструментом виступає фільтр лінійного передбачення, позначений передавальною функцією $A(z)$. Під час кодування

розраховуються параметри фільтра та активуючого сигналу, які потім передаються по каналу. На приймальному боці ці параметри використовуються для синтезу сигналу через аналогічну модель фільтра.

Суть роботи такого кодувальника полягає в наступному: мовний сигнал розбивається на сегменти довжиною приблизно 20 мс; для кожного сегмента проводиться оцінка параметрів фільтра та активації; у найпростішому варіанті активуючий сигнал утворюється як залишок після передбачення, отриманий шляхом пропускання сегмента через фільтр $A(z)$; обчислені параметри потім кодуються та передаються (рис. 2.4). Конкретні варіації LPC-кодеків відрізняються способами формування сигналу збудження, вибором параметрів моделі та іншими технічними деталями.

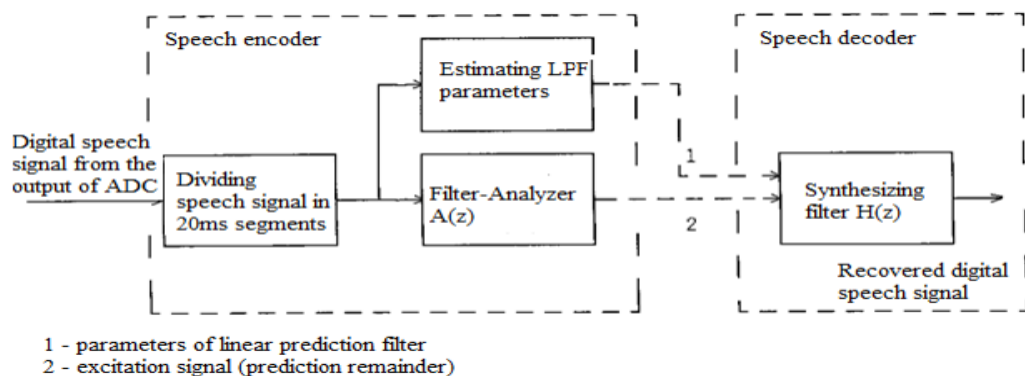


Рисунок 2.4 - Кодування мовлення на основі методу лінійного передбачення

З певного погляду, процес стиснення звукової інформації можна трактувати як одну з форм кодування. На початковому етапі такого кодування вхідні аудіосемпли перетворюються з часової області у частотну за допомогою багатофазного фільтруючого банку, який розбиває сигнал на 32 піддіапазони з однаковою шириною смуги. Такий підхід забезпечує ефективне перетворення з високою частотною та часовою деталізацією. Проте для реалізації цього механізму довелося прийняти низку компромісів.

Перший компроміс полягає в рівномірному поділі спектра на підсмуги. Хоча це значно спрощує реалізацію фільтрів, такий розподіл не відповідає характеристикам людського слуху, який має різну чутливість до звуків на різних частотах. Оптимальніше було б розподіляти частоти відповідно до критичних смуг, які визначаються особливостями сприйняття гучності в присутності

сусідніх звуків. Водночас окремі низькочастотні смуги охоплюють кілька критичних зон, ускладнюючи правильний розподіл бітів для квантованих значень у межах таких підсмуг. Якщо один піддіапазон (позначимо його як X) перекриває декілька критичних смуг, то система призначення бітів орієнтується на смугу з найнижчим порогом маскування шумів для подальшого квантування сигналу в межах X .

Другий компроміс стосується використання однакових фільтрів на боці приймача, тобто у декодері. Це призводить до втрат деякої частини інформації при зворотному перетворенні з частотної області в часову. Незважаючи на це, фахівці, що займалися розробкою фільтрів, доклали зусиль до мінімізації таких втрат, щоб зберегти прийнятну якість звуку.

Третій компроміс стосується конкретних фільтрів, які мають суттєве частотне перекриття, через що звук може потрапити у два фільтри та породити сигнали в двох замість однієї з 32 піддіапазонів. Багатофазний банк фільтрів використовує буфер X для зберігання 512 вхідних семплів та проміжні структури даних. Алгоритм читає чергові 32 семпли з вхідного файлу та заносить їх у буфер, одночасно зрушуючи його. Сигнали $S[i]$ для 32 піддіапазонів обчислюються за формулою. На Рис. 2,1 показано п'ять основних кроків багатофазного алгоритму фільтрування.

t для 32 підсмуги обчислюються за формулою:

$$S_i[i] = \sum_{k=0}^{63} \sum_{j=0}^7 M_{i,k}(C[k+64j] \times X[k+64j]), i=3, \dots, 31.$$

У даному випадку, $S[i] [t]$ описує сигнал у підсмугі i в момент часу t . Вектор C складається з 512 коефіцієнтів вікна аналізу, які задані згідно стандарту. Матриця аналізу, позначена як M , складається з компонентів.

$$M_{ik} = \cos\left(\frac{(2i+1)(k-16)\pi}{64}\right), i=0, \dots, 31; k=0, \dots, 63.$$

Таблиця 2.3 - Багаторазовий банк фільтрів

1	Помістити 32 нових семпла в буфер FIFO X .
2	Вікно семплів $Z_i = C_i$, для $i=0, \dots, 511$.
3	Проміжне обчислювання $Y_i = \sum_{k=0}^7 Z_{i+64k}$, для $i = 0, \dots, 63$
4	Обчислення 32 сигналів: $Y_i = \sum_{k=0}^{63} M_{ik} Y_k$, для $i = 0, \dots, 31$.
5	Дати вихід для 32 підсмугових сигналу S_i .

У рівнянні вираз у круглих дужках не залежить від i , а числа $M_{i,k}$ не залежать від j , що відображається в компромісі, що дозволяє зменшити кількість арифметичних операцій. Це призводить до того, що 32 сигнали $S[i]t$ обчислюються за допомогою $512+32*64=2560$ множень та $64*7+32*63=2464$ додавань на один сигнал. Децимація семплів полягає у проріджуванні 32 вихідних сигналів до одного на фільтр. Графічне зображення роботи кодера та декодера при багатозафазному фільтруванні представлено на рис.2.5.

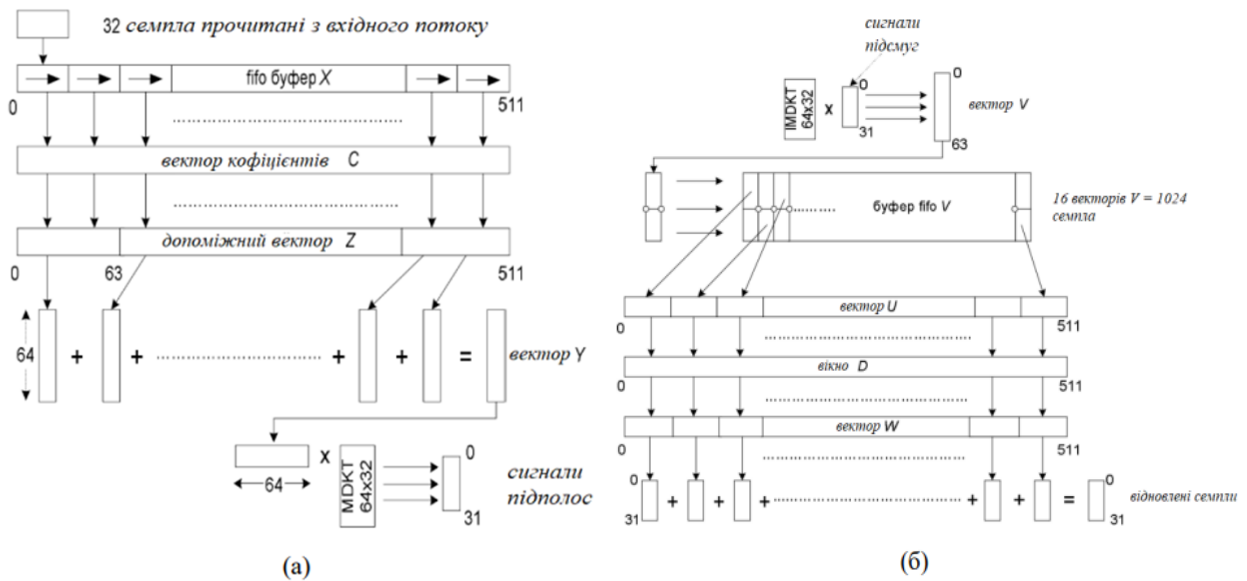


Рисунок 2.5 - Кодер звуку MPEG 9а) та його декодер (б)

Буфер X складається з 64 сегментів по 32 звукових семпла в кожному, який зсувається на один сегмент праворуч перед читанням наступних 32 нових семплів із вхідного файлу. Після множення буфера X на коефіцієнти вікна, результат поміщається в вектор Z, який потім множиться на матрицю MDCT. Результат цього множення складається в остаточний вектор із 32 компонентів, що представляють сигнали підсмути. Операції декодера зображені на частині (b) малюнку. Спочатку група з 32 сигналів підсмути помножена на матрицю IMDCT, що складається з компонент $N_{i,k}$, і результат записується у вектор V. Вектор V містить два сегменти по 32 числа у кожному і заноситься у буфер FIFO V зліва. Буфер V має 16 комірок для останніх векторів V (тобто для $16 * 64$ або 1024 чисел). Новий вектор U формується з 32 альтернативних сегментів буфера V, як показано на малюнку, і множиться на 512 коефіцієнтів D_i вікна синтезу (аналогічно коефіцієнтам C_i вікна аналізу, використовуваного кодером), щоб

отримати вектор W . Цей вектор ділиться на 16 сегментів по 32 компоненти в кожному, і всі сегменти складаються. Результатом є 32 реконструйованих звукових семпла. Компоненти матриці синтезу IMDCT задаються формулою:

$$N_{ik} = \cos\left(\frac{(2i+1)(k+16)\pi}{64}\right), i=0, \dots, 63; k=0, \dots, 31.$$

На стадії фільтрування сигнали підполос об'єднуються в кадри, кожен з яких містить 1152 сигналів (табл. 2.4). Далі ці сигнали піддаються масштабуванню та квантуванню, використовуючи психоакустичну модель і алгоритм призначення бітів. Після квантування, отримані квантовані величини, коефіцієнти масштабування та інформація про квантування записуються в стислий файл з використанням кодування Хаффмана [7].

Таблиця 2.4 - Реконструкція аудіосемплів

1.	Прочитати 32 нових сигналів $S_i, i = 0, \dots, 31$
2.	Зрушити буфер $FIFO V_i = V_{i-64}, i = 1023, \dots, 64$.
3.	Помножити: $V_i = \sum_{k=0}^{32} N_{ik} S_k, i = 0, \dots, 63$.
4.	Побудувати 512 компонентів вектор U : для $i = 0$ до 7 виконати; для $j = 0$ до 31 виконати; $U_{j+64i} = U_{j+128i}$; $U_{32+j+64i} = U_{96+j+128i}$.
5.	Побудувати W . Для $i = 0$ до 511 виконати, $W_i = U_i \times D_i$
6.	Обчислити 32 семпла. Для $j = 0$ до 31 виконати $S_i = \sum_{i=0}^{15} W_{j+32i}$
7.	Видати на вхід 32 семпла S_j .

2.3 Потокова передача аудіо: принципи роботи (Streaming)

Передача потокового аудіо, що також називається стримінгом, це потік мультимедійних даних, що отримується кінцевим користувачем від провайдера безперервно через Інтернет. Мультимедіа потоки існують двох типів – так звані «живі», live потоки та потоки за запитом. Live потоки доступні певний невеликий період часу, наприклад – живі ефіри зі спортивних матчів, прямих репортажів новин, тобто подій які відбуваються конкретно в цей час.

Розмір потокової мультимедійної інформації, що зберігається обчислюється в залежності від швидкості переданої інформації та тривалості інформації за наступною формулою (для одного користувача та файлу):

Розмір сховища = тривалість мультимедійних даних (у секундах) * бітрейт (у кбіт/с) / (8 * 1024); (8 * 1024 = 1 Мб). Як приклад наведено аудіо файл формату .mp3, тривалістю 5 хвилин (300 секунд) та максимальним значенням бітрейту для цього формату – 320 кбіт/с. Таким чином, розрахунок займаного місця на носії буде наступним: $(300 \text{ с} * 320 \text{ кбіт/с}) / (8 * 1024) = 11,72 \text{ Мб}$ місця на носії.

Якщо файл, що зберігається на сервері з режимом передачі за запитом, буде проглядатися 1000 людей одночасно за протоколом Unicast (1 клієнт - 1 з'єднання), то сервер повинен мати таку пропускну здатність: $320 \text{ кбіт/с} * 1000 = 320.000 \text{ кбіт/с} = 320 \text{ Мбіт/с}$ мережевого інтерфейсу. Це еквівалент близько 11 Гб інформації на годину. Зрозуміло, що при використанні протоколу Multicast навантаження на сервер набагато нижче, так як для передачі інформації всім клієнтам використовується єдиний потік. Тому, подібний потік буде займати лише 320 кбіт/с мережевого інтерфейсу сервера [8].

Також, з передачею потокової мультимедійної інформації пов'язаний важливий процес під назвою буферизація. Буферизація – це процес завантаження частини потоку достроково, щоб користувач міг безперервно прослуховувати, продивлятися, одним слово – використовувати, контент, навіть якщо з'єднання короткочасне [9]. Буферизація особливо корисна у випадках, де швидкість мережі доволі низька або нестабільна (існує ймовірність від'єднання мережі). Поняття затримки імплементовано в явище потокової передачі даних, і є невід'ємною її частиною. Затримка – це час, протягом якого інформація надходить до кінцевого користувача, багатьом затримка також відома як «ring».

Не варто й оминати поділ потокової передачі на передачу з втратами та без втрат (так звана lossless). Передача з втратами має місце на існування за умов обмеженої пропускну здатності мережі або коли місце для збереження аудіо обмежене. Це можуть бути відео- та аудіоконференції, телефонні розмови, онлайн-відеотрансляції. Однак, очевидним недоліком передачі з втратами є якість передаваного звуку. Це є наслідком роботи алгоритмів стиснення даних в процесі їх передачі. Передача аудіо без втрат підходить до питання передаваного аудіопотоку трохи з іншого боку. Як зрозуміло з матеріалу наведеного вище – передача без втрат застосовується в тих випадках, коли для користувача є

важливим якість передаваних даних. Здебільшого, враховуючи наш сучасний побут, це актуально під час прослуховування пісень, аудіотреків. Його очевидним недоліком є те, що трафік такого формату передачі значно більший, за стандартний формат стисненого аудіо. Стрімкий розвиток сучасних технологій, на щастя для всіх нас, дозволяє використовувати певні методи стиснення даних, що зберігають фактично 100% оригінальних даних, при цьому зменшуючи вихідний розмір передаваних даних.

На даний момент зустрічається все більше і більше стрімінгових майданчиків, що надають всі можливості Lossless-аудіо користувачам. Найпопулярнішими з них є Spotify, Apple Music, YouTube Music, Tidal та інші. Щоб зробити потокове передавання без втрат можливим, ці майданчики мають використовувати спеціальні аудіокодеки без втрат. Такі кодеки стискають файли без втрати важливої частини аудіоданих, при цьому зберігаючи оригінальну якість звуку, зменшуючи їх розмір для більш зручного зберігання та потокового відтворення (передавання).

Найпоширенішими з цих кодеків в потоковому мовленні без втрат є:

- WAV (Waveform Audio File Format) – одна з найпопулярніших форм стиснення звуку без втрат у професійному аудіо-продакшені. Має діапазон бітрейту від 353 Кбіт/с (для 8-бітного моно-файлу на частоті семплування 8 КГц) і аж до 9,2 Мбіт/с (для 32-бітного стерео-файлу з частотою семплування 192 КГц). Тобто цей формат є абсолютно не стиснутим;

- FLAC (Free Lossless Audio Codec) – є одним з найбільш гнучких кодеків, що пропонує високоякісне стиснення аудіо. Підтримується широким спектром аудіопристроїв та ПЗ. Бітрейт даного кодеку варіюється від 800 Кбіт/с до 1400 Кбіт/с;

- ALAC (Apple Lossless Audio Codec) – це власний кодек Apple inc. Подібно до FLAC, він забезпечує стиснення аудіо без втрат, хоча в основному використовується для потокового передавання в Apple Music. Значення бітрейту варіюється від ~900 Кбіт/с (для 16 біт в 44.1 КГц) і до 6444 Кбіт/с (для 24 біт в 192 КГц);

- AIFF (Audio Interchange File Format) – AIFF схожий на WAV за ситуаціями використання, популярністю та гнучкістю, бо так само використовується у аудіовиробництві, зведенні та мастерингу. Значення бітрейту наближені до WAV.

Хоча може здатися, що це не так вже й багато, але потокове передавання без втрат є революційним для поточкових сервісів, а ще краще для аудіофілів, які серйозно ставляться до того, як вони слухають музику. З появою поточкових сервісів без втрат, таких як Apple Music, Tidal та Amazon Music HD, насолоджуватися музикою з високою якістю стало простіше, ніж будь-коли (прощайте, солодкі MP3). Звичайно, хоча для поточкового передавання без втрат може знадобитися сумісне обладнання та вища пропускна здатність, проте, за цими технологіями майбутнє.

2.4 Синхронізація аудіо в реальному часі (VoIP, конференц-зв'язок)

Синхронізація є ключовим компонентом при передаванні мультимедійного контенту через мережу. Вона відіграє вирішальну роль у забезпеченні своєчасного відтворення аудіо на обох кінцях з'єднання між абонентами, а також у точній координації відтворення кількох медіапотоків із визначеним часовим зміщенням, що важливо для комфортного сприйняття користувачем. Це особливо актуально, наприклад, під час перегляду фільмів, де одночасно передаються відео- і аудіодоріжки.

Залежно від контексту синхронізації мультимедійних потоків у мережевому середовищі, виділяють три основні типи:

- 1) синхронізація всередині одного потоку (Intra-stream synchronization);
- 2) синхронізація між різними потоками (Inter-stream synchronization);
- 3) синхронізація між декількома кінцевими точками (також відома як групова) - Inter-destination synchronization.

Синхронізація всередині потоку відповідає за узгодження часових зв'язків між медіаодинацями в межах одного потоку, тобто гарантує відсутність часових збоїв, таких як пропуски чи перекриття кадрів під час відтворення.

Синхронізація між потоками забезпечує узгодження різних, але взаємопов'язаних потоків, наприклад, коли необхідно синхронно відтворити відео й аудіо. Типовим прикладом є синхронізація звуку з мімікою під час розмови у відео - так звана синхронізація губ [10].

Останній тип - синхронізація між приймачами (або між кінцевими пунктами) - стосується одночасного відтворення медіа на кількох пристроях. Це критично, наприклад, при трансляції звуку на декілька динаміків: навіть

незначна затримка на одному з них може викликати відчутне спотворення звучання, якщо пристрої розташовані на близькій відстані.

Після досягнення синхронізації необхідно забезпечити її подальшу підтримку. Через те, що апаратні й програмні годинники не є абсолютно точними, їх хід може дещо відрізнятись, що призводить до поступового порушення синхронності, навіть якщо початкове узгодження було ідеальним. Така деградація синхронізації з часом викликає додаткові труднощі, зокрема: як повторно забезпечити синхронне відтворення без створення небажаних звукових артефактів.

Для подолання цієї проблеми використовуються різні стратегії. Прості підходи можуть просто встановлювати годинник на нове значення, не враховуючи можливі спотворення. Натомість більш складні методи поступово коригують хід годинника за допомогою спеціальних алгоритмів, що дозволяє підтримувати синхронізацію з меншими втратами якості.

Щоб досягти фактичної синхронізації, необхідно організувати відтворення кожного фрагменту аудіо на всіх пристроях у точно визначений момент, орієнтуючись на єдиний глобальний часовий стандарт. Якщо кожен аудіопаке́т буде позначено часовою міткою, що відповідає певному моменту в майбутньому, а всі пристрої матимуть синхронізовані годинники, то вдасться забезпечити узгоджене відтворення. Синхронізацію годинників між пристроями зазвичай реалізують за допомогою спеціалізованих протоколів передачі часу.

Втім, важливо не лише досягти початкової точності синхронізації, а й підтримувати її стабільність протягом тривалого періоду. На якість синхронного звучання впливають численні фактори, включно з мережею, навантаженням на обладнання та навіть температурними змінами, які можуть викликати зсуви у ході годинника.

Слід також враховувати, що кожен пристрій має власне джерело часу - це може бути як апаратний годинник реального часу, так і програмна реалізація, що впливає на точність і стабільність часової синхронізації в багатопрістроєвих системах.

Під час роботи програмного забезпечення використання годинника реального часу не є обов'язковим. Натомість застосовується системний годинник - програмна конструкція, що оновлює свої значення на основі апаратних таймерів. Такі таймери формуються з урахуванням тактових імпульсів, які генеруються кварцовими генераторами. Проте точність таких

генераторів може варіюватися через їхню фізичну природу та якість виготовлення. Саме тому виникає потреба у постійному коригуванні системного часу відповідно до загального (еталонного) часу.

Для забезпечення такої синхронізації використовують спеціальні протоколи, серед яких найрозповсюдженішими є NTP (Network Time Protocol) і PTP (Precision Time Protocol).

Коли мова йде про синхронізацію годинників, варто підкреслити, що ключовим є не абсолютна точність відносно об'єктивного часу, а ступінь узгодженості між годинниками на різних пристроях. Тобто важливішим є не збіг з "реальним" часом, а мінімальна різниця між локальними годинниками, що й визначає ефективність синхронізованої роботи системи.

Протокол NTP (Network Time Protocol) є одним із ключових механізмів для узгодження системного часу на комп'ютерах через мережі з нестабільною затримкою, які функціонують на основі передачі даних у вигляді пакетів. NTP передбачає як формати обміну часовими даними, так і алгоритмічні методи визначення найбільш точного джерела часу та корекції внутрішнього годинника.

Цей протокол підтримується більшістю сучасних операційних систем. Наприклад, сам проєкт NTP надає повноцінне програмне забезпечення для платформ типу Unix, а Microsoft включила підтримку NTP у свої пізніші версії Windows.

Протокол побудований за принципом «майстер-підлеглий» (або «ведучий-ведений»), де годинники організовані у вигляді ієрархічних рівнів, або «шарів». Чим менше значення рівня, тим вища точність джерела часу. До першого рівня належать сервери, що мають пряме з'єднання з еталонними джерелами (рівень 0) — наприклад, з атомними годинниками або GPS-годинниками. Кожен пристрій, що синхронізується з сервером певного рівня n , отримує рівень $n+1$. Чим далі пристрій від джерела референтного часу, тим меншою вважається точність, і максимально допустимий рівень у системі - п'ятнадцятий.

Несинхронізований комп'ютер зазвичай позначається як рівень 16. Ієрархічна структура рівнів дозволяє уникнути замкнених ланцюгів у процесі синхронізації. Для налаштування годинника комп'ютера В відповідно до часу сервера А протокол NTP використовує чотири часові позначки, які передаються у двох обмінюваних пакетах.

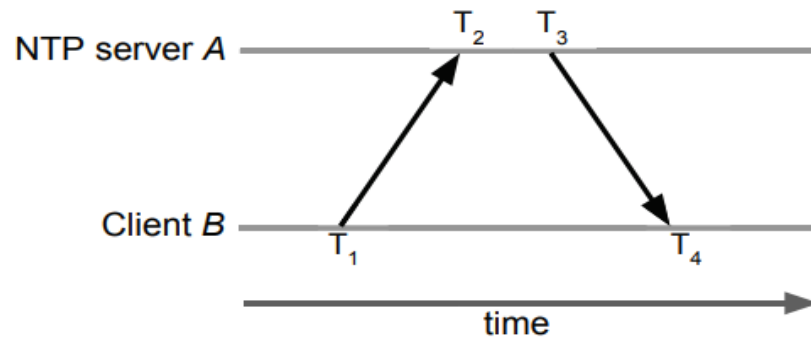


Рисунок 2.6 - Процес обміну пакетів між «майстром» та «підлеглим» в NTP

У випадку одноадресного режиму (див. рис. 2.6) ініціатор В надсилає повідомлення до А, фіксуючи час відправлення як T_1 . Сервер приймає його о T_2 і надсилає відповідь у момент T_3 . Коли клієнт отримує відповідь, він реєструє час T_4 . Таким чином, клієнт В має доступ до всіх чотирьох часових міток (T_1 – T_4), необхідних для обчислення затримки δ та зсуву годинника θ відносно сервера [10].

Для того щоб обчислений час вважався надійним, клієнт має виконати кілька успішних обмінів, які відповідають вимогам протоколу. В середньому довіреним вважається сервер, якщо таких коректних обмінів відбулося близько п'яти. Після цього годинник клієнта починає синхронізуватися з сервером або за допомогою різкого коригування, або поступової адаптації — залежно від виявленого зсуву. Частота синхронізації налаштовується: спочатку пакети надсилаються кожні 60 секунд, а згодом інтервал може збільшитися до 20 хвилин, щоб знизити навантаження на мережу.

Однією з головних проблем при визначенні зсуву часу є асиметричність затримки: маршрути в прямому та зворотному напрямках можуть мати різну тривалість. Така різниця у затримках маскується під звичайне зміщення часу, тому її важко виявити. У випадках з неоднорідними каналами, наприклад, коли зв'язок в одному напрямку здійснюється через супутник, а в іншому – по наземній лінії, похибка може сягати 100 мс і більше. Єдиний спосіб зменшити такі помилки – мати заздалегідь відому інформацію про маршрути в мережі. Проте, у сприятливих умовах, як-от у локальних Ethernet-мережах з невеликим навантаженням, можливо досягти точності до 100 мкс, тоді як у глобальних інтернет-з'єднаннях – до кількох десятків мілісекунд [10].

PTP (Precision Time Protocol) призначений для синхронізації годинників у мережевому середовищі з високим рівнем точності. У локальних мережах (LAN) він може забезпечувати синхронізацію із точністю до субмікросекунд, що робить його придатним для застосування в системах управління, вимірювання, телекомунікацій, фінансових транзакцій та інших критичних застосунках, де GPS-сигнали недоступні.

Хоча обидва протоколи – PTP і NTP – мають спільну мету (синхронізацію часу), вони суттєво відрізняються за реалізацією. Зокрема, PTP оперує часовими значеннями з роздільністю до наносекунд, що свідчить про його орієнтацію на точне середовище. Протокол також висуває вимоги до типу мережевого середовища – ідеальним вважається Ethernet з мінімальними затримками. Для досягнення максимальної точності необхідна апаратна підтримка: мережева карта, комутатор або маршрутизатор повинні бути здатні до апаратного маркування часу. Завдяки цьому помилки синхронізації можуть бути зведені до рівня наносекунд, тоді як у суто програмному режимі вони зазвичай становлять від сотень мікросекунд до мілісекунд [10].

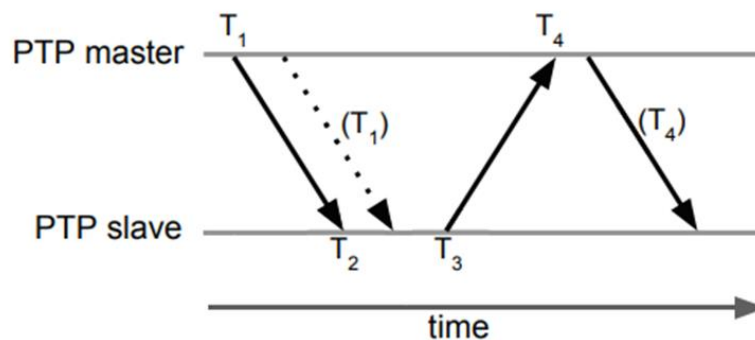


Рисунок 2.7 - Процес обміну пакетів між «майстром» та «підлеглим» в PTP

На відміну від NTP, який використовує систему стратифікації (страти) для визначення авторитетності джерела часу, PTP застосовує алгоритм «найкращого майстра». Це означає, що кожен вузол у мережі здатний автоматично визначити, чи повинен він діяти як ведучий (гросмейстер) або підлеглий, залежно від своєї точності та доступності інших вузлів. Якщо поточний гросмейстер втрачає якість сигналу, наприклад, через втрату GPS-з'єднання, система автоматично переналаштовується, і інший вузол бере на себе роль ведучого. Така гнучкість забезпечує стійку синхронізацію навіть у випадках збоїв.

3 ОГЛЯД ПРОТОКОЛІВ ПЕРЕДАЧІ АУДІО

3.1 Визначення та роль аудіопротоколів у комп'ютерних мережах

У сучасному цифровому середовищі, де передавання інформації дедалі частіше відбувається через комп'ютерні мережі, передача звуку набуває особливого значення. Це стосується не лише побутового використання, але й професійних сфер: від дистанційної освіти до мультимедійного мовлення. Здійснення якісного звукового зв'язку можливе завдяки аудіопротоколам, які забезпечують надійне, якісне та швидке передавання аудіоданих через мережу.

Аудіопротоколи - це набір технічних правил, що визначають спосіб пакування, маршрутизації, обробки й передачі звукових сигналів між пристроями через мережу [9, с.698]. Вони працюють на основі моделі OSI (Open Systems Interconnection), і можуть реалізовуватись на кількох її рівнях залежно від конкретного стандарту чи технології. Основними задачами покладеними на аудіопротоколи є синхронізація передаваних даних з усіх кінців мережі, мережева маршрутизація аудіопотоків, масштабованість (більше стосується AoIP) та сумісність з апаратним обладнанням, ОС та ПЗ.

Залежно від сфери використання, аудіопротоколи поділяються на:

- Streaming-протоколи (RTP, RTSP, HLS, WebRTC), які використовуються у глобальних мережах (Інтернет), забезпечують передавання аудіопотоків до кінцевого користувача незалежно від його пристрою або умов мережі;
- протоколи AoIP (Audio over IP) — Dante, RAVENNA, AVB, які працюють у локальних або професійних мережах із високими вимогами до якості звуку.

Одним із ключових параметрів для аудіопротоколів є затримка передавання (latency), тобто час між передачею та відтворенням сигналу. Іншим критичним аспектом є якість аудіо, що визначається форматом кодування (кодеком) та режимом стискання. Наприклад, у протоколах VoIP на кшталт RTP часто застосовуються кодеки з втратами, тоді як AoIP орієнтований на безстиснене передавання.

Streaming-протоколи, як-от RTP (Real-time Transport Protocol) та RTSP (Real-Time Streaming Protocol), широко використовуються для голосового зв'язку, відеоконференцій, потокового мовлення. Вони компенсують втрати пакетів, адаптуються до зміни пропускної здатності та забезпечують безперервність потоку [10]. HLS (HTTP Live Streaming) і DASH дозволяють динамічно змінювати якість передавання в залежності від мережевих умов, що особливо важливо для мобільних пристроїв.

WebRTC (Web Real-Time Communication) - один із найновіших і найперспективніших протоколів для аудіо-/відеозв'язку в реальному часі через браузері без плагінів. Його застосовують в онлайн-конференціях, чатах, телемедицині [12].

Натомість протоколи AoIP, як-от Dante, RAVENNA, AVB, Livewire, CobraNet, AES67, призначені для високоточної передачі звуку з мінімальною затримкою (менше 5 мс) та високою частотою дискретизації (до 192 кГц, 24 біти). Такі системи застосовуються в студіях звукозапису, у концертних залах та на радіостанціях. Наприклад, AES67 — це відкритий стандарт, що забезпечує сумісність між різними AoIP-системами, дозволяючи передавати аудіо без стиснення в реальному часі.

Завдяки цим протоколам забезпечується:

- висока якість звуку у професійних сферах;
- адаптивна передача в умовах змінного навантаження;
- масштабованість для різних платформ, пристроїв та фізичних умов конкретної ситуації.

Отже, аудіопротоколи є критичним інструментом у сучасних комп'ютерних мережах, визначаючи не лише якість передачі, а й загальну ефективність комунікаційних систем у найрізноманітніших галузях — від домашнього користування до професійного мовлення.

3.2 Протоколи потокової передачі аудіо

Одним з основних протоколів, що забезпечують передачу мультимедійного контенту в реальному часі через IP-мережі, є RTP (Real-Time Protocol). Він призначений для потокової трансляції аудіо- та відеоінформації та функціонує на прикладному рівні моделі OSI. У заголовку RTP-пакета міститься службова інформація, яка дозволяє на приймальній стороні правильно

реконструювати звук або зображення, зокрема відомості про використований кодек (наприклад, JPEG або MPEG), мітки часу та номери пакетів. Ці параметри необхідні для коректного впорядкування, декодування та, при потребі, інтерполяції відсутніх фрагментів даних.

Протокол RTP не має фіксованого порту, однак дотримується правило: для RTP-потоків використовується парне значення порту, а відповідний непарний порт зарезервовано для допоміжного протоколу RTCP. Динамічне призначення портів ускладнює проходження RTP-трафіку через міжмережеві екрани, тому для вирішення цього питання часто застосовується STUN-сервер, що допомагає узгодити параметри з'єднання.

RTCP (Real-Time Control Protocol) — це допоміжний протокол, що працює разом із RTP і забезпечує застосунок механізмами адаптації до змін у мережевому середовищі. Наприклад, коли надходить сигнал про збільшення мережевого навантаження або зменшення доступної пропускної здатності, застосунок може відповідно знизити свої вимоги, наприклад, шляхом погіршення якості переданого медіа. Якщо мережеве навантаження зменшиться, застосунок знову зможе повернутися до первинних налаштувань якості. У процесі передачі RTP-потоків критично важливо отримувати зворотний зв'язок, оскільки приймачі можуть не підтримувати вказані параметри якості або виникає необхідність коригувати параметри передачі в режимі реального часу. Протокол RTCP, який функціонує через UDP і використовує власний порт, якраз і забезпечує такий тип керування та моніторингу.

RTSP (Real-Time Streaming Protocol) — це спеціалізований протокол керування передачею мультимедійних потоків, розроблений для інтерактивної взаємодії з аудіо- та відеоконтентом у реальному часі через мережу. Він дає змогу встановлювати, керувати та контролювати медіасесанси між клієнтами (наприклад, програвачами або мобільними застосунками) і серверами, які надають контент. За допомогою RTSP користувачі можуть запускати, призупиняти, продовжувати або зупиняти відтворення поточкових даних, а також змінювати характеристики потоку — такі як бітрейт, роздільна здатність чи частота кадрів. Цей протокол фактично реалізує функціонал дистанційного керування, подібний до дій з локальним медіаплеєром, але в мережевому середовищі.

RTMP (Real-Time Messaging Protocol) — це власницький протокол, призначений для потокової передачі даних, зокрема відео- та аудіоконтенту,

який надходить, наприклад, з вебкамер через інтернет. Попри появу новітніх технологій, RTMP досі активно використовується для передачі закодованого відео на популярні соціальні платформи, сервіси онлайн-трансляцій та медіасервери.

Протокол RTMP зазвичай застосовується для організації початкового етапу трансляції, відомого як «перша миля» — тобто доставки відеопотоку від джерела (транслятора) до медіасерверу. Для забезпечення доставки контенту кінцевому споживачу, так званої «останньої милі», використовуються інші рішення, як-от HLS (HTTP Live Streaming) або WebSocket, що краще адаптовані до сучасних пристроїв і мережевих умов.

RTMP функціонує поверх TCP-з'єднання, забезпечуючи стійке та стабільне з'єднання з мінімальними затримками. Для підвищення ефективності передавання даних цей протокол ділить потоки на менші сегменти. Зазвичай аудіо фрагментується на пакети по 64 байти, а відео — на пакети по 128 байтів, хоча розмір може узгоджуватись динамічно.

RTMP також підтримує кілька логічних (віртуальних) каналів, кожен із яких відповідає за окремий тип передаваних даних. Наприклад, один канал призначено для відео, інший — для аудіо, ще один — для процедур віддаленого виклику (RPC), а також є спеціальний канал для службових повідомлень. Це розділення дозволяє забезпечити ефективне керування передачею даних та зменшує конфлікти між потоками.

Завдяки малій затримці протокол RTMP залишається популярним вибором для трансляцій у реальному часі — таких як стрими подій, онлайн-семінари та геймерські трансляції [14].

HLS (HTTP Live Streaming) — це протокол потокової трансляції з підтримкою адаптивного бітрейту, який базується на HTTP та був створений компанією Apple Inc. Хоча спочатку він був орієнтований на пристрої Apple, згодом отримав широку підтримку з боку різних операційних систем і пристроїв — від Android до Smart TV та медіаплеєрів інших виробників.

Протокол HLS має просту структуру, яка не потребує спеціалізованого обладнання для реалізації. Процес трансляції починається з кодування відео за допомогою сучасних відеокодеків, таких як H.264 або H.265, які дозволяють зменшити розмір файлів без значного погіршення якості. Після цього відео розбивається на невеликі відрізки тривалістю, як правило, від 5 до 10 секунд,

хоча тривалість може адаптуватися під специфіку контенту або швидкість мережі.

Кожен такий фрагмент функціонує як незалежна одиниця, яку можна окремо завантажити та програти. Це дозволяє системі більш гнучко реагувати на проблеми зі зв'язком — у разі помилки можна повторно запитати лише потрібний фрагмент, не перезавантажуючи весь потік.

TCP (Transmission Control Protocol) — це один із ключових транспортних протоколів, який використовується в комп'ютерних мережах для надійної передачі даних. На відміну від альтернативного протоколу UDP, TCP гарантує цілісність і послідовність переданих даних шляхом встановлення логічного з'єднання між двома хостами. TCP належить до категорії протоколів із попереднім встановленням з'єднання.

Протокол TCP отримує дані від вищих рівнів мережевої моделі OSI (наприклад, HTTP або FTP) і розділяє їх на частини — сегменти, до яких додається службова інформація, зокрема номер послідовності. Сформовані сегменти потім інкапсулюються в IP-пакети й передаються через мережу до хоста-приймача.

На приймальній стороні кожен IP-пакет перевіряється на відповідність контрольній сумі, а отримані TCP-сегменти розміщуються в правильному порядку. Якщо деякі частини втрачені або пошкоджені, хост-приймач може ініціювати повторну передачу конкретного сегмента. Завдяки цій процедурі забезпечується повнота отриманої інформації. Коли всі дані зібрані, TCP передає їх тому ж протоколу верхнього рівня, від якого вони походили на стороні відправника.

UDP (User Datagram Protocol) — це протокол транспортного рівня, який забезпечує обмін даними без встановлення з'єднання. На відміну від TCP, який гарантує доставку, контроль потоку та правильний порядок передачі, UDP позбавлений подібних механізмів. Він не перевіряє цілісність, не здійснює повторну передачу втрачених пакетів і не контролює перевантаження, що дозволяє досягти високої швидкості та мінімальних затримок.

Завдяки своїй простоті UDP активно застосовується у сфері, де швидкість передачі важливіша за її надійність. Наприклад, у потоковій трансляції аудіо та відео, мережних іграх, DNS-запитах або протоколах керування мережею (наприклад, SNMP). Його легковажна природа дозволяє уникнути складностей,

пов'язаних з узгодженням стану з'єднання між клієнтом і сервером, що критично для застосунків, орієнтованих на реальний час.

З іншого боку, відсутність перевірки доставки означає, що пакети можуть загубитися, прибути двічі або не в тому порядку. Тому відповідальність за обробку помилок та відновлення даних покладається на рівень застосунків. Це робить UDP непридатним для використання в ситуаціях, де потрібна гарантована доставка, як-от передача файлів або надсилання електронної пошти.

Отже, UDP — це високопродуктивне рішення для комунікацій у реальному часі, де ключовим пріоритетом є швидкодія. Водночас важливо розуміти межі його застосування, щоб уникнути ризиків, пов'язаних із втратою чи пошкодженням даних. Розробникам варто ретельно зважувати переваги UDP у своїх проєктах і при необхідності передбачати додаткові механізми компенсації його недоліків.

Розробка мережевих протоколів потокового мовлення викликає ряд проблем:

- протоколи типу датаграм, зокрема User Datagram Protocol (UDP), передають мультимедійні дані у вигляді послідовності окремих невеликих пакетів, які відзначаються простотою реалізації та високою ефективністю, однак стандарт UDP не передбачає механізмів гарантії доставки інформації до одержувача. Це створює суттєві труднощі для програм, які отримують ці дані, оскільки виявлення та корекція помилок лягає на їхній функціонал. У разі втрати деяких пакетів можливе припинення або значне порушення цілісності потоку [13];

- протоколи RTSP, RTP та RTCP були спеціально створені для передавання мультимедійного контенту через мережу. При цьому RTP і RTCP функціонують поверх UDP-протоколу;

- протоколи з підвищеною надійністю, зокрема TCP, забезпечують точність доставки даних користувачам потокового відео чи аудіо. Проте, за умов значної кількості помилок під час з'єднання або підтвердження отримання, дані можуть втратити свою актуальність. До того ж, повторне надсилання пошкоджених фрагментів спричиняє додаткову затримку. Щоб компенсувати це, часто застосовується буферизація контенту на боці користувача;

- протоколи типу Unicast передають індивідуальну копію медіаданих кожному окремому користувачу. Такий підхід є ефективним для більшості

користувачів Інтернету, проте створює суттєві труднощі з масштабуванням серверів при великій кількості підключень;

- Multicast-протоколи були створені з метою зменшення навантаження на сервер та економії пропускної здатності при передачі мультимедіа багатьом одержувачам одночасно. У цьому випадку один потік даних розповсюджується на групу клієнтів. Проте можливість реалізації групової трансляції залежить від особливостей мережевої інфраструктури. До недоліків цього підходу можна віднести неможливість забезпечити індивідуальний запит користувача, як у сервісах відео на вимогу, а також відсутність контролю за відтворенням з боку глядача. Цю проблему частково вирішують через використання кешувальних серверів і буферизацію потоку на стороні клієнта за допомогою спеціального програмного забезпечення;

- технологія Multicast забезпечує можливість передачі єдиного потоку даних одночасно кільком користувачам у межах мережі. Одним із викликів при впровадженні такої схеми трансляції є правильна конфігурація мережевого обладнання, зокрема маршрутизаторів, для пересилання ширококомовних пакетів між різними мережевими сегментами. Якщо організація, яка надає доступ до потокового контенту, має повний контроль над маршрутом між сервером і споживачами (наприклад, у навчальних закладах, державних структурах або корпоративних мережах), то можливо застосувати протоколи маршрутизації на кшталт IGMP або PIM для ефективною доставки медіаданих у різні локальні мережі;

- протоколи однорангової передачі (P2P) також можуть бути залучені для розповсюдження заздалегідь записаного контенту між комп'ютерами користувачів. Такий підхід зменшує навантаження на центральний сервер, однак лінія зв'язку між сервером і одним із клієнтів може стати «вузьким місцем» у загальній системі доставки потокової інформації.

3.3 Протоколи передачі цифрового аудіо IP-мережами (AoIP)

Передача цифрового звуку через IP-мережі стала суттєвим проривом у сфері аудіокомунікацій, особливо у професійному середовищі. Залежно від типу протоколу та обладнання, ця технологія дозволяє передавати великі обсяги аудіоінформації в режимі реального часу, без стиснення, із мінімальними затримками, а також із можливістю гнучкого налаштування параметрів мережі

відповідно до конкретного завдання. Однією з ключових переваг AoIP-рішень є значне зменшення кількості аналогових кабелів — замість десятків або сотень метрів проводів можна обійтися лише кількома мережевими з'єднаннями. Це значно спрощує інфраструктуру студій, сценічних майданчиків або мобільних систем, покращуючи зручність та оптимізуючи простір.

AoIP-протоколи функціонують на основі низки технологій, що забезпечують стабільну та якісну передачу звукового сигналу з низькою латентністю. Одним із таких стандартів є AES67, який служить спільною платформою для сумісності між різними AoIP-системами та дозволяє передавання несжатого аудіо у мережах третього рівня (Layer 3). Ще один важливий компонент — RTPv2 (Precision Time Protocol), що забезпечує високоточну синхронізацію годинників між пристроями, з точністю до субмікросекунди, що є критично важливим для аудіопередач у реальному часі [15].

Також у мережах із підтримкою AoIP використовуються IGMP (Internet Group Management Protocol) та Multicast-технології, які дають змогу ефективно розподіляти аудіопотоки на декілька приймачів, мінімізуючи навантаження на мережу. А механізм DiffServ QoS (Quality of Service) дозволяє виставляти пріоритети для аудіоданих у загальному мережевому трафіку, що гарантує стабільну якість відтворення навіть за умов високої завантаженості.

Серед численних переваг, які забезпечує використання технології Audio over IP, варто відзначити передусім високу якість передаваного звуку. Цифрове аудіо передається без стиснення, з високою роздільною здатністю. Наприклад, специфікація AES67 передбачає підтримку аудіо з частотою дискретизації до 192 кГц при 24-бітній розрядності, що відповідає найвищим стандартам професійного звукового середовища.

Ще однією визначальною перевагою є масштабованість. У системах AoIP легко додавати нові пристрої або розширювати кількість аудіоканалів без необхідності прокладки додаткових фізичних з'єднань. Якщо в аналогових системах кожен канал потребує окремого кабелю та інфраструктури, то в AoIP достатньо лише підключити пристрій до наявної мережі. Більше того, цифровий сигнал можна розподіляти на декілька приймачів без втрат у якості або потреби в додатковому обладнанні.

Важливою особливістю є також можливість передачі аудіосигналу на великі відстані — до сотень кілометрів — за допомогою стандартних мережеских

технологій (наприклад, Ethernet або оптоволокна), що відкриває нові горизонти для побудови складних аудіомереж.

Системи AoIP мають вбудовані механізми підвищення надійності з'єднання, серед яких — резервні маршрути, відмовостійкі топології та автоматичне перемикання на альтернативні канали у разі збоїв. Зменшення кількості фізичних кабелів також позитивно впливає на стабільність та зменшує ризик механічних пошкоджень або перешкод.

Ще однією перевагою є централізоване управління, яке дозволяє адмініструвати усю аудіомережу з одного або декількох вузлів, що значно спрощує налаштування, моніторинг та обслуговування системи.

Окрім того, цифрові аудіомережі менш чутливі до електромагнітних завад та шумів, які часто виникають при передачі аналогового сигналу, особливо на великих відстанях.

Водночас, як і будь-яка технологія, AoIP має свої обмеження. Її ефективне використання потребує глибших знань з комп'ютерних мереж та доступу до інфраструктури, яка працює на рівні 3 моделі OSI (мережевий рівень). Це передбачає залучення мережевого обладнання з підтримкою специфічних протоколів.

Керування AoIP-системою часто потребує наявності виділених пристроїв, які відповідають за маршрутизацію та синхронізацію аудіопотоків у мережі.

Ще одним викликом є неповна стандартизація у цій галузі. Через велику кількість несумісних між собою протоколів (Dante, Ravenna, Livewire+, AVB та інші), інтеграція обладнання від різних виробників може бути складною та вимагати проміжних пристроїв або програмного забезпечення. Незважаючи на існування стандарту AES67, який частково вирішує проблему сумісності, на практиці створення універсальної системи залишається технічно складним завданням.

Навіть зі стрімким розвитком мережевих та аудіотехнологій, й досі, багато існуючих аудіопристроїв класу Hi-End не підтримують AoIP технології, що обмежує їх застосування на певних ринках. Та не можна оминати й фактичну вартість спеціалізованого обладнання для побудови AoIP інфраструктури. Хоч воно і стає все доступнішим для кінцевого користувача, проте це явище ще не досягло того рівня, коли його можна назвати «доступним».

Більшість аудіо мереж на основі Ethernet використовують лише перші два рівні: фізичний і канальний. Фізичний рівень описує електричні та фізичні

характеристики пристроїв, що утворюють мережу. Канальний рівень описує спосіб передачі даних у мережі, якщо зв'язок здійснюється від вузла до вузла в межах локальної мережі.

Третій (мережевий) рівень вступає в дію, коли потрібні більш складні мережі. Він описує, як відбуватиметься комунікація в різних мережах, і визначає адреси для кожного вузла, маршрутизацію через проміжні вузли (маршрутизатори), розмір пакетів даних і підтримання відповідної якості обслуговування (описано нижче). Цей рівень використовується для доставки звуку в таких протоколах, як Dante, де використовується IP. Цей додатковий рівень додає гнучкості, дозволяючи аудіоданим подорожувати мережею разом з іншим мережесим трафіком.

Говорячи про AoIP обов'язково треба згадати про мережеві топології, бо вони є одним з основних принципів функціонування мереж такого типу. Найбазовішими топологіями є:

1) топологія типу «Шина» (рис. 3.1). Кожен пристрій в мережі з'єднаний одним кабелем. Ця топологія рідко використовується для сучасних Ethernet-з'єднань. USB та IEEE 1394 (Firewire) є системами віртуальної шини [16];

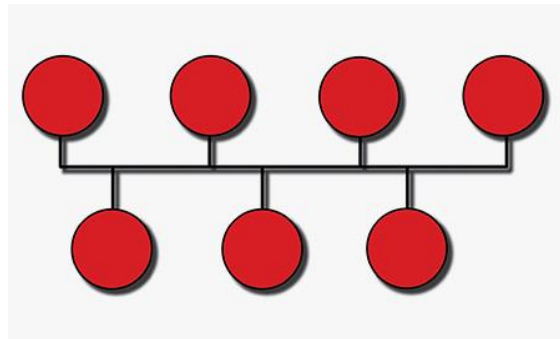


Рисунок 3.1 - Топологія типу «Шина»

2) Топологія типу «Зірка» (рис. 3.2). Кожен пристрій підключено до центрального концентратора. Хаби також можуть бути з'єднані між собою, створюючи топологію «зірка із зірок» [16];

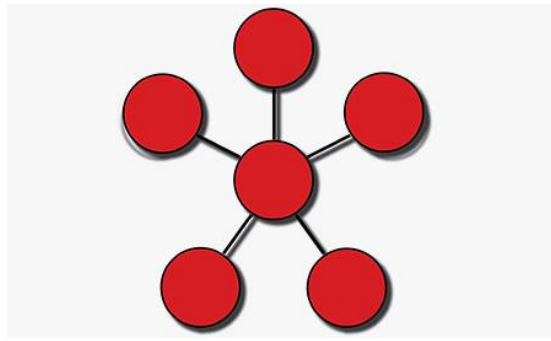


Рисунок 3.2 - Топологія типу «Зірка»

3) топологія типу «Шлейф» (рис. 3.3). Пристрої з'єднуються послідовно, кінець в кінець. Це найпростіший спосіб додати більше пристроїв до мережі, за винятком топології «Зірка» [16];

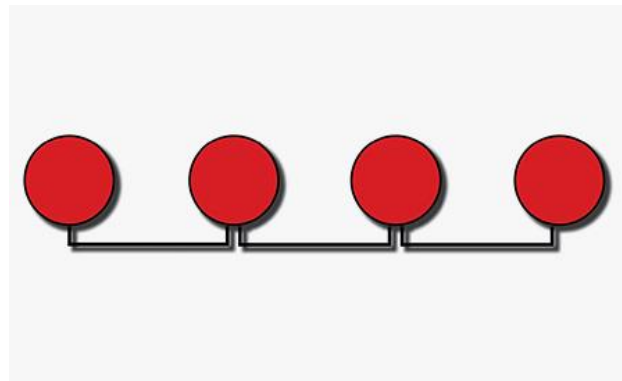


Рисунок 3.3 - Топологія типу «Шлейф»

4) топологія типу «Кільце» (рис. 3.4). Принцип схожий до Шлейфу, проте тут останній пристрій з'єднаний з першим, утворюючи замкнуту контур або кільце[16];

5) топологія типу «Дерево» (рис. 3.5). В цій топології є центральний вузол, з'єднаний з одним або декількома вторинними вузлами, які далі поєднуються з третинними вузлами і так далі. Ця топологія може розширюватися до нескінченності, поки не утворяться кільця [16];

6) Топологія типу «Сітка» (рис. 3.6) [16]. Вона дозволяє підключати всі пристрої до будь-якого пристрою в мережі. Щоб уникнути втрати даних у комірчастих з'єднаннях, потрібні складніші алгоритми маршрутизації.

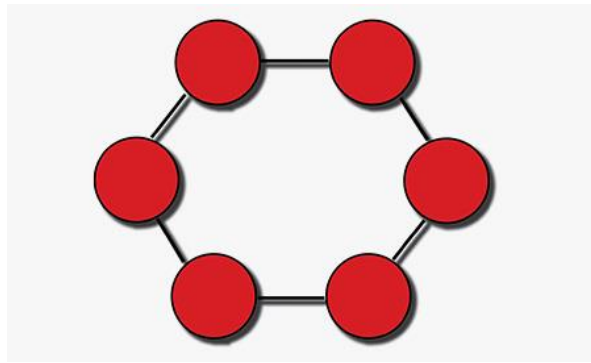


Рисунок 3.4 - Топологія типу «Кільце»

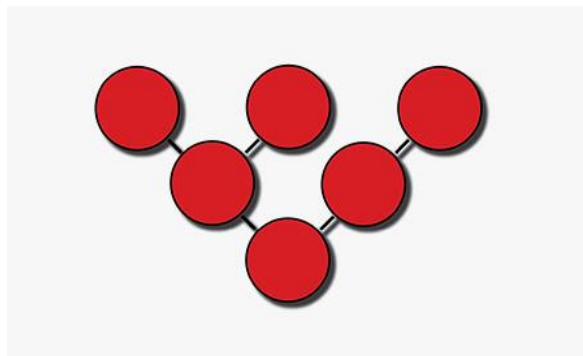


Рисунок 3.5 - Топологія типу «Дерево»

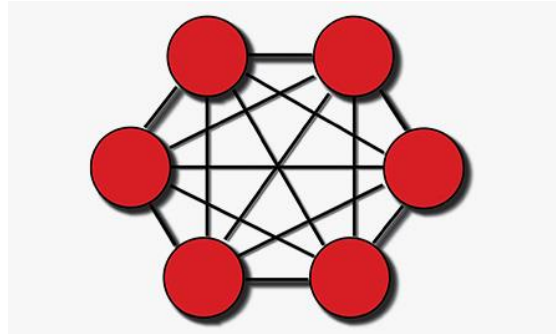


Рисунок 3.6 - Топологія типу «Сітка»

На даний момент вже існує доволі багато протоколів технології AoIP. Найвідомішими з них є: AES3 (AES/EBU), AES47, AES50, A-Net, AVB, CobraNet, Dante, EtherSound ES-100, EtherSound ES-Gigam HyperMAC, LightViper, Livewire, MADI (AES10), Nexus, Optocore, SANE, REAC, RockNet. Всі вони відрізняються використанням методом передавання, топологією, кількістю передаючих каналів, максимальною частотою дискретизації, значенням затримки, кабелями що використовують, конекторами, максимальною відстанню дії та вимогами до наявності пропрієтарного обладнання для їх функціонування.

4 ПОРІВНЯЛЬНИЙ АНАЛІЗ АУДІОПРОТОКОЛІВ

4.1 Критерії для порівняння (затримка, якість передачі, надійність, масштабованість, складність реалізації)

У сучасних комп'ютерних мережах мультимедійний контент (аудіо, відео, анімація, інтерактивні елементи) передається в режимі реального часу або за запитом. Для забезпечення якісного та ефективного передавання мультимедіа розроблено різноманітні мережеві протоколи: RTP, RTSP, HTTP Live Streaming (HLS), MPEG-DASH, WebRTC, SIP, а також протоколи рівня доступу та синхронізації — наприклад, RTP, TCP, UDP. Для того, щоб об'єктивно оцінити доцільність використання того чи іншого протоколу в конкретній мережевій ситуації, треба проаналізувати порівняльні критерії, які охоплюють технічні та експлуатаційні аспекти.

Важливим критерієм є затримка передачі (Latency) – час від моменту генерації звукового сигналу на передавальному пристрої до його відтворення на приймальному пристрої, який вимірюється в мілісекундах (мс) і є критично важливим у реальному часі. Низькою вважається затримка 1–50 мс, яка необхідна для голосового зв'язку, музичного виступу, онлайн-конференцій. Висока затримка 100–5000 мс підходить для потокового мовлення та аудіоархівів. Протоколи RTP і WebRTC забезпечують дуже низьку затримку, тоді як HLS і DASH мають більшу затримку через буферизацію [17]. AoIP-протоколи, як-от Dante чи AVB, досягають затримки менше 2 мс завдяки апаратній оптимізації.

Наприклад, у відеозв'язку через WebRTC між браузерами Chrome і Firefox спостерігається затримка в межах 20–40 мс, що дозволяє підтримувати природну розмову. Для HLS затримка зазвичай становить 6–30 секунд, тому він більше підходить для трансляцій, а не для взаємодії в реальному часі.

Також критерієм є якість передачі аудіосигналу (Quality of Service), яка залежить від типу використовуваних кодеків, таких як Opus, G.722, AAC, PCM, рівня стиснення (зі втратами чи без, частоти дискретизації та бітової глибини). AoIP-протоколи передають безстиснене аудіо з частотою до 192 кГц та 24-бітною точністю, що забезпечує студійну якість. У свою чергу, VoIP-протоколи, такі як RTP, можуть використовувати кодеки зі стисненням, жертвуючи

частиною якості заради меншого навантаження на мережу. Наприклад, RTP у парі з кодеком Opus забезпечує якісне аудіо при 32 кбіт/с, а DASH дозволяє змінювати якість відео під час трансляції, адаптуючись до змін пропускної здатності інтернет-каналу.

Надійність (Reliability) – критерій, який визначає здатність протоколу зберігати цілісність аудіопотоку, відновлювати пакети при втраті або пошкодженні, ефективно працювати при змінних умовах мережі. HLS і DASH мають високу надійність завдяки буферизації та адаптивній зміні бітрейту, RTP забезпечує певну надійність через використання контрольних протоколів (наприклад, RTCP), але не має вбудованого механізму перезапиту. AoIP залежить від якості мережі (часто працює в ізольованій VLAN), але за правильного налаштування досягає дуже високої стабільності. Наприклад, HLS автоматично буферизує контент, щоб уникнути зупинок при тимчасових мережевих збоях. WebRTC застосовує механізми FEC для відновлення втрачених пакетів у голосовому трафіку.

Масштабованість (Scalability) – критерій, який визначає здатність протоколу обробляти зростаючу кількість користувачів або пристроїв без втрати продуктивності. Протоколи типу HLS і HTTP-DASH мають можливість добре масштабуватися завдяки архітектурі на основі CDN (Content Delivery Network). Протокол WebRTC, навпаки, не масштабований без складних серверних рішень (SFU, MCU). У AoIP-системах масштабованість залежить від пропускної здатності мережі та кількості каналів; у складних проєктах використовуються спеціальні комутатори з підтримкою QoS. Наприклад, HLS транслюється через CDN, що дозволяє обслуговувати тисячі глядачів одночасно. WebRTC без серверів SFU підходить лише для малих груп (2–5 користувачів).

Складність реалізації (Implementation Complexity) є важливим критерієм для розробників та інженерів та включає необхідність спеціального обладнання, сумісність з існуючими мережами, складність налаштування та підтримки. Протоколи, що працюють поверх HTTP (HLS, DASH), мають відносно просту реалізацію та високу сумісність. AoIP-протоколи часто вимагають спеціалізованих комутаторів, синхронізації годинників (через RTP), окремих VLAN, тому мають високу вартість впровадження. Наприклад, інтеграція HLS у вебсайт потребує лише HTTP-сервера і плеєра, а впровадження WebRTC вимагає налаштування NAT-traversal, STUN/TURN серверів і сертифікатів шифрування.

Також критеріями є сумісність з платформами та безпека. HLS ефективно підтримується на iOS, WebRTC — у сучасних браузерях (Chrome, Safari), а RTSP потребує окремих програм-плеєрів (VLC, QuickTime). WebRTC шифрує весь трафік за допомогою DTLS і SRTP, тоді як DASH через HTTPS також забезпечує захист від стороннього втручання.

Таким чином, жоден з протоколів не є універсальним та підходить для виконання певних конкретних завдань. В процесі вибору протоколу треба враховувати умови використання: за потреби низької затримки оптимальним вибором можуть бути WebRTC, RTP або AoIP-протоколи. При необхідності масштабованості та стабільності доцільно обирати HLS, DASH - протоколи. Для забезпечення високої якості без стиснення раціональним вибором можуть бути Dante, AVB, AES67 - протоколи.

4.2 Порівняння протоколів аудіообробки та передачі

Більшість протоколів потокової передачі аудіо не можна порівняти між собою, бо велика частина цих протоколів працюють кожен в своїй площині, та виконує власні конкретні задачі. Проте, як мінімум, такі параметри як затримка, адаптивність, сумісність з програмним забезпеченням або пристроями, наявність буферизації можна піддати порівнянню в різних протоколах.

Для початку згадаємо, для чого ж конкретно призначені кожен протокол потокової аудіопередачі. RTP – протокол реального часу, використовується у VoIP та стримінгу; RTSP – здійснює контроль потоків (команди відтворення, паузи, та закінчення аудіоматеріалу); RTMP – використовується для трансляцій в реальному часі; HLS – протокол потокової передачі через HTTP від Apple inc.; DASH – відкритий стандарт адаптивного стримінгу; TCP та UDP – транспортні протоколи на яких базується стримінг.

RTP – низька затримка (від 10 до 100 мс), має високу сумісність (і відео тракт, і VoIP, і SIP). RTSP – значення затримки коливаються від низьких до середніх. Для використання протоколу потрібен плеєр. Застосовується в IP-камерах та локальному стримінгу. RTMP – значення затримки від 2 до 5 секунд, використовується в Live-стримінгу, YouTube, Twitch та подібних платформах (через сервер). HLS – протокол адаптивного стримінгу, значення затримки коливаються від 5 до 25 секунд, що є доволі великою затримкою, проте стається з-за того що HLS підтримує буферизацію, яка і має в своїй основі певну

затримку, протокол може працювати з адаптивною передачею даних (адаптуватися під конкретні умови та під конкретного користувача) і сумісний зі всіма сучасними браузерями, використовується на онлайн-радіо, подкастах та в Apple Music. DASH – як і HLS має на меті організацію адаптивного стрімінгу, має таку ж високу затримку, так само має можливість працювати з адаптивною передачею даних та має буферизацію. Працює з HTML5 плеєрами, використовується в YouTube, веб-плеєрах та VoD. В TCP затримка середня, сумісність висока, є базовим протоколом для RTMP, HLS та DASH. UDP – низька затримка, використовується в RTP та VoIP.

Порівняння протоколів AoIP потрібне для побудови оптимальної інфраструктури передачі цифрового аудіо (радіостанції, студії звукозапису). В процесі порівняння AoIP протоколів найважливішими критеріями є – метод транспортування, використаний тип топології, ліміт кількості каналів передачі, максимальна частота дискретизації, значення затримки, кабель та роз'єми, що використовуються в мережах цих протоколів, максимальна відстань передачі сигналу за допомогою цих технологій та протоколів [16].

Детальне порівняння найрозповсюдженіших протоколів передачі цифрового аудіо (AoIP) наводимо в окремій таблиці (див. табл. 4.1).

У AoIP усі протоколи розташовані на трьох рівнях OSI (Layer 1, Layer 2 та Layer 3). Протоколи першого рівня використовують кабельні комунікації та сигнальні компоненти Ethernet, проте не використовують структуру Ethernet фреймів. На цьому рівні протоколи доволі часто використовують власний контроль доступу до середовища (MAC), що відрізняється від того, що використовується в Ethernet, в підсумку, це створює певні проблеми сумісності, та вимагає виділеної мережі для протоколу. Відкритими стандартами на цьому рівні є AES50 та MaGIC, пропрієтарними – ULTRANET, A-Net, HyperMAC та AudioRail. Протоколи другого рівня інкапсулюють аудіодані в стандарті фрейми Ethernet. Це явище має назву «Audio over Ethernet». Більшість протоколів використовують стандартні комутатори та концентратори, хоча деякі все ж вимагають, щоб мережа (або VLAN) була виділена для програми розповсюдження аудіо. Відкритими стандартами на цьому рівні є AES51 та AVB який використовує IEEE 1722 AV профіль транспортного протоколу, пропрієтарними – CobraNet, EtherSound, REAC, SoundGrid та dSNAKE. Третій рівень протоколів інкапсулює аудіодані в пакети мережевого рівня (третій рівень OSI). Це не обмежує вибір протоколу лише IP-протоколом третього рівня. В

деяких випадках пакети аудіоданих третього рівня додатково пакуються всередину пакетів транспортного рівня (четвертий рівень OSI), зазвичай це UDP або RTP.

Таблиця 4.1 - Порівняння найпопулярніших протоколів аудіо передачі (AoIP)

Назва	Метод передачі	Топологія	Кількість каналів (макс.)	Значення дискретизації (макс.)	Затримка	Кабель	Роз'єм	Відстань передачі (макс.)
AES47	ATM	Сітка	Обмежено пропуск.здатністю	192 КГц	125 μ s за хоп	Cat5 або оптика	RJ45 або оптичні типу SC/LC	100 м з Cat5/2 км з MM/70 км з SM
AES50	Фіз. рівень Ethernet	Точка-Точка	48	384 КГц	63 μ s	Cat5	RJ45	100 м
A-Net	Фіз. рівень Ethernet	Шлейф, зірка або комбінована	64	192 КГц	322 μ s + 1,32 μ s за хоп включно A/D та D/A	Cat5e або оптика	RJ45 або оптичні типу SC/LC	120 м з Cat5e/2 км з MM/70 км з SM
AVB	Будь-який носій IP	Будь-яка IP	Обмежено пропуск.здатністю	Обмежено пропуск.здатністю	2 мс для 7-хопової мережі зі 100 Мбіт/с Ethernet	Cat5 або оптика	RJ45 або оптичні типу SC/LC	100 м з Cat5/2 км з MM/70 км з SM
CobraNet	Канальний рівень	Будь-яка IP	Обмежено пропуск.здатністю	96 КГц	1,33 та 5,33 мс	Cat5 або оптика	RJ45 або оптичні типу SC/LC	100 м з Cat5/2 км з MM/70 км з SM
Dante	Будь-який носій IP	Будь-яка IP	Обмежено пропуск.здатністю	192 КГц	Від 0,3 мс до 2,33 мс	Cat5 або оптика	RJ45 або оптичні типу SC/LC	100 м з Cat5/2 км з MM/70 км з SM
Livewire	Канальний рівень	Дерево	Обмежено пропуск.здатністю	48 КГц	2,75 мс включаючи A/D та D/A	Cat6 або оптика	RJ45 або оптичні типу SC/LC	100 м з Cat6/2 км з MM/70 км з SM
Optocore	Виділене оптоволокно	Кільце	512	96 КГц	41,6 μ s	Оптика	SC/LC оптичні	700 м з MM/110 км з SM

Використання RTP або UDP для передачі аудіоданих дозволяє розповсюджувати їх через стандартні комп'ютерні маршрутизатори, в свою чергу це надає можливість економічно-оптимізовано побудувати велику аудіомережу з використанням комерційного готового обладнання. Хоч такі IP-пакети можуть проходити через Інтернет, проте більшість протоколів третього рівня не можуть забезпечити надійну передачу через Інтернет з-за обмеженої пропускної здатності, значної наскрізної затримки та втрат пакетів, з якими потік даних може зіткнутися протягом проходження крізь терени Інтернету. Аналогічні причини стають на заваді передачі аудіо третього рівня через бездротову локальну мережу та не підтримуються більшістю рішень в цій

професійній галузі. Відкритими стандартами тут є AES67, AVB що використовує IEEE 1733, NetJack, RAVENNA. Пропрієтарними на цьому рівні є Dante, Q-LAN, Livewire та WheatNet-IP. Не можна не згадати такі розробки, як RockNet (використовує Cat5 кабель), Hydra2 (використовує кабель Cat5e або оптоволокно за допомогою SFP-трансиверів. Технологія MADI, що використовує 75-омний коаксіальний кабель з роз'ємами BNC або оптоволокно для передачі до 64 каналів цифрового аудіо у з'єднанні «точка-точка». AES47 забезпечує аудіомережу, передаючи аудіосигнал AES3 через мережу АТМ (Асинхронного Режиму Передавання) за допомогою структурованих мережевих кабелів (як мідних, так і оптоволоконних). Ця технологія широко використовувалася підрядниками, які забезпечували аудіозв'язок ВВС в режимі реального часу по всій Великобританії.

ВИСНОВКИ

В межах виконання бакалаврської кваліфікаційної роботи був проведений аналіз з різних точок зору будови та функціонування існуючих мережевих аудіопротоколів, що дозволило отримати глибоке розуміння їх принципів роботи та використання у сучасних локальних та глобальних мережах. У першому розділі було сформульоване загальне уявлення про телекомунікаційні канали як основу передачі звукової інформації, наведено класифікацію каналів зв'язку, розглянуто основи модуляції та кодування сигналу, а також визначено ключові особливості, які впливають на якість переданої аудіоінформації.

У другому розділі проаналізовані методи обробки та передачі аудіоданих, таких, як оцифрування (дискретизація, квантування та кодування), стиснення та безпосередньо передача аудіотракту. Також розглянуто принципи потокової передачі аудіо та виклики, пов'язані із синхронізацією в реальному часі, що є особливо важливим у випадку VoIP та онлайн-конференцій.

Третій розділ присвячений огляду та аналізу мережевих аудіопротоколів потокової передачі AoIP, які використовуються в сучасних комп'ютерних мережах. Визначена роль кожного з аудіопротоколів потокової передачі у забезпеченні безперебійної та якісної передачі звуку, а також запропоновані рішення для оптимальної та стабільної передачі аудіосигналу.

У четвертому розділі проведено порівняльний аналіз обраних аудіопротоколів за критеріями затримки, надійності, масштабованості, якості передачі та складності реалізації. Встановлено, що кожен протокол має свої переваги та обмеження, а ефективність їх застосування залежить від конкретного контексту — технічних умов, вимог до якості, бюджету та програмного/апаратного середовища.

Узагальнюючи результати дослідження, можна зробити висновок, що аудіопротоколи є невід'ємною частиною сучасних мереж та інтернету, грамотний вибір та застосування аудіопротоколів у комп'ютерних мережах є вирішальним чинником для забезпечення стабільного, якісного та синхронізованого передавання звукової інформації. З огляду на стрімкий розвиток мережевих технологій, важливим залишається подальше вдосконалення стандартів та впровадження гнучких протоколів, які зможуть адаптуватися до вимог нових форматів, платформ та користувацьких сценаріїв.

Розуміння функціонування аудіопротоколів та впливу їх на мережеві технології має важливе значення для подальшого розвитку сучасного інформаційного простору. Дослідження в цій галузі дає можливість вдосконалити мережеві технології та забезпечити їхню стабільність та ефективність у майбутньому. Варто зазначити, що в контексті швидкого розвитку мережевих та аудіо технологій зі зростаючої складністю мережевих інфраструктур, важливо постійно вдосконалювати аудіопротоколи та адаптувати їх до нових вимог та викликів. Також варто прагнути уніфікації стандартів для забезпечення сумісності між різними системами та пристроями, що сприяє їхній більш ефективній роботі та взаємодії. Зокрема, інноваційні технології, такі як AoIP, VoIP та Streaming-протоколи, відіграють важливу роль у вдосконаленні телекомунікаційних мереж. У майбутньому важливо продовжувати дослідження у галузі мережевих аудіопротоколів та активно впроваджувати нові розробки з метою забезпечення стабільності, ще більшої доступності, безпеки та ефективності передачі звукової інформації. Такий підхід дозволить забезпечити подальший розвиток та інновації у галузі інтернету та мережевих технологій, що є ключовим для подальшого успішного функціонування сучасного світу.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Білінський Й. Й., Огороднік К. В., Юкиш М. Й. Електронні системи: навч. посіб. Вінниця : ВНТУ, 2011. 208 с.
2. Учасники проєктів Вікімедіа. Канал зв'язку – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Канал_зв'язку (дата звернення: 18.06.2025).
3. Shannon C. E. A Mathematical Theory of Communication. Reprinted with corrections from The Bell System Technical Journal. 1948. No. 27. P. 379–443. <https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>
4. Introduction to BCH Codes. <https://www.ece.unb.ca/>.
5. Зубань Ю.О., Крючко Є.В. Система передачі аудіо-відео інформації засобами мережі інтернет на основі адаптивного алгоритму стиску. Електронні засоби та дистанційні технології для навчання протягом життя: тези доповідей VIII Міжнародної науково-методичної конференції, 15-16 листопада 2012 р., м. Суми. Суми: СумДУ, 2012. С. 121-122.
6. RFC 7655: RTP Payload Format for G.711.0. *IETF Datatracker*. URL: <https://datatracker.ietf.org/doc/html/rfc7655> (date of access: 19.06.2025).
7. A Study on Data Compression Algorithms for Its Efficiency Analysis / C. Rodrigues et al. *Smart Innovation, Systems and Technologies*. Singapore, 2021. P. 475–488. URL: https://doi.org/10.1007/978-981-16-3675-2_36 (date of access: 19.06.2025).
8. Учасники проєктів Вікімедіа. Потокowe мультимедіа – Вікіпедія. *Вікіпедія*. URL: <https://uk.wikipedia.org/wiki/Телекомунікації> (дата звернення: 18.06.2025).
9. Tanenbaum A. S., Wetherall D. J. Computer networks. Boston, Massachusetts 02116 : Pearson Education, Inc., 2011. 962 p.
10. Lauri C., Malmgren J. Synchronization of streamed audio between multiple playback devices over an unmanaged IP network. Lund : Department of Electrical and Information Technology Lund University, 2015. 89 p.
11. RFC 3550: RTP: A Transport Protocol for Real-Time Applications. IETF Datatracker. URL: <https://datatracker.ietf.org/doc/html/rfc3550> (date of access: 22.06.2025).
12. WebRTC: Real-Time Communication in Browsers. W3C. URL: <https://www.w3.org/TR/2025/REC-webrtc-20250313/> (date of access: 22.06.2025).

22.06.2025 Голь В.Д., Ірха М.С. Системи передачі даних: конспект лекцій. Київ: ІСЗЗІ КПІ ім. Ігоря Сікорського, 2021. 126 с.

13. Breaking Down UDP: A Comprehensive Guide to the User Datagram Protocol | CQR. *CQR*. URL: <https://cqr.company/ua/wiki/protocols/breaking-down-udp-a-comprehensive-guide-to-the-user-datagram-protocol/> (date of access: 21.06.2025).

14. Understanding RTMP for Live Streaming: A Comprehensive Guide. *Embedding IP Camera Live Video Stream into web page - IPCamLive.com*. URL: <https://www.ipcamlive.com/understanding-live-streaming-rtmp> (date of access: 21.06.2025).

15. What is AoIP? - Nuprime-X. Nuprime Audio MultiChannel Products. URL: <https://nuprime-x.com/what-is-aoip/> (date of access: 23.06.2025).

16. Digital Audio Protocols – AV.technology. AV.technology. URL: <https://av.technology/tutorials/digital-audio-protocols> (date of access: 23.06.2025).

17. Schulzrinne H., Casner S., Frederick R., Jacobson V. RTP: A Transport Protocol for Real-Time Applications [Електронний ресурс] RFC 3550. — IETF, 2003.

[https://www.researchgate.net/publication/287227808 RTP A Transport Protocol for Real-Time Applications](https://www.researchgate.net/publication/287227808_RTP_A_Transport_Protocol_for_Real-Time_Applications)