

## АНАЛІЗ МЕТОДІВ КОДУВАННЯ МОВИ ДЛЯ ВИКОРИСТАННЯ В РАДІОМЕРЕЖАХ З ПАКЕТНОЮ КОМУТАЦІЄЮ

*В статті розглядаються методи аналого-цифрового перетворення мови в радіомережах з комутацією пакетів, на основі аналізу особливостей функціонування яких обґрунтовуються вимоги до мовних кодеків, визначаються можливі напрямки удосконалення алгоритмів кодування мови.*

*Гурський Т. Г. Анализ методов кодирования речи для использования в радиосетях с пакетной коммутацией. В статье рассматриваются методы аналого-цифрового преобразования речи в радиосетях с коммутацией пакетов, на основе анализа особенностей функционирования которых обосновываются требования к речевым кодекам, определяются возможные направления совершенствования алгоритмов кодирования речи.*

*T. Gursky The analyses of speech coding methods for packet switching radio network use. The paper examines methods of analog-to-digital speech conversion in radio networks with packet switching, based on the analysis of the functioning of substantiation requirements for language codecs; possible directions of algorithm improvement of language coding are recognized.*

**Ключові слова:** методи аналого-цифрового перетворення, кодекси мови, комутація пакетів.

Історично телекомунікаційні послуги надавались споживачам різнорідними мережами, а саме телефонними мережами загального користування, мережами рухомого радіозв'язку та мережами документального електрозв'язку. Ці мережі від початку свого створення до приблизно 2000 року функціонували відносно самостійно. В останні роки намітилася тенденція до конвергенції різнорідних мереж результатом якої повинно стати створення єдиної універсальної мережі наступного покоління [1].

Сучасний ритм життя висуває до засобів персонального зв'язку вимоги до забезпечення надання послуг телефонного та відеозв'язку, високошвидкісного доступу до мережі Internet у будь-якому місці та в будь-який час. Очевидно, що задовольнити ці вимоги можуть лише мережі рухомого радіозв'язку. У той же час, одним із шляхів досягнення високої ефективності використання пропускної спроможності мереж, підвищення швидкості доступу, є застосування єдиного принципу комутації пакетів на всіх ділянках мережі [1–5]. Тому, якщо на етапі становлення мереж передачі даних зв'язківці намагалися за допомогою різноманітних модемів передавати дискретну інформацію через телефонні мережі та пізніше мережі рухомого радіозв'язку, то наприкінці ХХ століття почався зворотний процес: перехід до перетворення мови в дані та їх передача цифровими мережами, орієнтованими на обмін інформацією у формі пакетів.

Аналого-цифрове перетворення мови є важливим завданням при створенні пакетних радіомереж. У телекомунікаційних системах використовуються багато типів мовних кодеків, орієнтованих на мережі з комутацією каналів. Область, в якій ведуться нові розробки – IP-телефонія, для якої створюються більш стійкі до втрат пакетів кодекси. Слід зазначити, що розроблені та випробувані кодекси для мереж з комутацією пакетів орієнтовані на стаціонарні канали. У таких мережах по радіоканалу передача може здійснюватись лише на одній ділянці – для з'єднання з рухомими абонентами. У той же час, в останні роки все більшого поширення набувають так звані мобільні радіомережі з динамічною топологією (Mobile Ad-hoc Networks – MANET) [6, 7]. Робота в умовах швидких змирань та присутності завад у радіоканалах таких мереж висуває додаткові вимоги щодо стійкості до втрат пакетів.

*Тому, метою роботи є проведення аналізу існуючих методів аналого-цифрового перетворення мовних сигналів, особливостей кодування мови в радіомережах із комутацією пакетів та визначення вимог до таких кодеків мови.*

Розглянемо найбільш поширені в телекомунікаційних мережах кодекси.

Перша група аналого-цифрових перетворювачів мови реалізує методи перетворення форми сигналу [8, 9]. Історично першим є кодек стандарту G.711, що використовує принцип імпульсно-кодової модуляції (ІКМ), що включає чотири основних операції: дискретизацію (згідно теореми Котельникова), квантування за рівнем, компандування (компресію сигналу на передачі та експандування на прийомі), та кодування. ІКМ забезпечує найбільш високу якість передачі мови серед усіх вузькосмугових кодеків, але з найбільшою швидкістю – 64 кбіт/с. Диференціальна ІКМ (ДІКМ) дозволяє зменшити необхідну швидкість передачі до 32 кбіт/с за рахунок кодування не абсолютних значень відліків сигналу, а різниці між сусідніми відліками, адаптивна ДІКМ (АДІКМ) передбачає зміну величини кроку квантування в залежності від того, на скільки швидко змінюється огинаюча звукового коливання. Дельта-модуляція представляє собою спрощений випадок ДІКМ з однорозрядним кодуванням.

Іншу принципово відмінну групу методів АЦП мови реалізують вокодері, робота яких заснована на кодуванні параметрів спектрів звуків. Ці методи дозволяють на прийомі відтворити мову „ситнетичним” шляхом [8, 9]. У передавальній частині (аналізаторі) реалізуються такі операції: визначається характер спектра звуку в даний момент часу – тон або шум; для вокалізованих звуків оцінюється частота основного тону; оцінюється форма огинаючої спектра (формантним або смуговим методом). Після квантування значення цих параметрів кодуються і передаються на приймальну сторону. Періодичність, з якою необхідно передавати значення параметрів, визначається, виходячи з мінімальної тривалості звуків української мови, і приймається, як правило, рівною близько 25 мс.

У синтезаторі сигнал відновлюється за допомогою цифрового фільтра, параметри якого встановлюються відповідно до прийнятих оцінок.

У сучасних вокодерах застосовується цифрова обробка сигналів. При цьому в АЦП аналізатора мовний сигнал спочатку перетворюється у сигнал з ІКМ і надалі його обробка здійснюється у цифровому вигляді.

Вокодерні системи з лінійним передбаченням мовлення (Linear Prediction Coding – LPC) працюють з блоками підготовлених відліків. Це призводить до збільшення затримки при передачі, оскільки кодування застосовується не до окремих значень, а до деякого їх набору, який перед початком перетворень необхідно накопичити в буфері.

У LPC-вокодерах передача даних здійснюється на низьких швидкостях – 2,4 і 4,8 кбіт/с. На швидкості 2,4 кбіт/с забезпечується задовільний рівень розбірливості мовлення, але якість і природність недостатні.

Більш складні алгоритми на базі LPC комбінують LPC з елементами кодування звукової хвилі. Ці алгоритми використовують замкнутий LPC-кодер (який називається також „аналіз через синтез” – Analysis-by-Synthesis – Abs). Апаратура передачі виконує процедуру пошуку і знаходить найкращу апроксимацію кожного мовного сегмента. Як тільки така апроксимація визначена, тоді вона у вигляді коду передається на приймальний бік, де використовується для відновлення аналогового сигналу.

Приклади стандартних замкнутих LPC-алгоритмів – це метод лінійного передбачення з кодовим збудженням (code-excited linear prediction – CELP), метод регулярного імпульсного збудження (Regular Pulse Excitation – RPE), який використовується в європейських стільникових системах зі швидкістю 13,2 кбіт/с і метод LD-CELP з низькою затримкою (low delay CELP), прийнятий ІТУ 1992 року як стандарт кодування мовлення G.728 зі швидкістю 16 кбіт/с.

Для різноманітних задач використовують нестандартні методи кодування, зокрема:

- варіанти адаптивного кодування з передбаченням (Adaptive Predictive Coding – APC);
- метод лінійного передбачення з векторним збудженням (Vector-Sum-Excited Linear Prediction – VSELP), запропонований як стандарт для цифрових стільникових систем США, які працюють на швидкості 8 кбіт/с;
- метод лінійного передбачення з предиктивним кодовим збудженням (Predictive Code-Excited Linear Prediction – PCELP).

Якщо не враховувати критерії складності і затримки, то головними досягненнями в кодексах сигналів є: поліпшення якості сигналу при визначеній швидкості передачі й одержання заданої якості сигналу при низьких швидкостях. Для комерційних додатків, де якість передачі мовлення, характерна для ІКМ, служить еталоном, варто особливо виділити задачу одержання прийняттого звучання на все більш низьких швидкостях. Це особливо важливо для кодувальних пристроїв у безпроводових телекомунікаційних системах, які використовують обмежені за пропускну здатністю радіочастотні та супутникові канали зв'язку.

В області мультимедіа та відеотелефонів, які підключаються до комутованих телефонних мереж, використовується стандарт стиску G.723, який дозволяє домогтися дуже суттєвого стиску мовлення при зберіганні достатньо високої якості відтворення сигналу. Він частково базується на новому методі стиску мовлення (Multipulse Maximum Likelihood Quantization – MP-MLQ), а також на методі ACELP (Algebraic code-excited linear prediction).

Метод MP-MLQ відноситься до сімейства алгоритмів AbS. Мовний кодер MP-MLQ використовує LPC-аналізатор 10-го порядку і працює на швидкостях 4,8; 6,4; 7,2 і 8 кбіт/с. Алгоритм MP-MLQ дозволяє розробляти похідні реалізації для швидкостей аж до 4 кбіт/с і більш низьких комунікаційних затримок (до 20 мс), здійснювати кодування на декількох швидкостях і зі змінною швидкістю, виконувати багатоканальну обробку (завдяки низькому обчислювальному навантаженню) та досягати високої якості на швидкості 8 кбіт/с.

На відміну від інших кодерів із низькими інформаційними швидкостями MP-MLQ забезпечує мінімальний рівень спотворень, коли мовний сигнал проходить через два або більш послідовних цикли компресії/декомпресії.

Ефективні алгоритми кодування реалізовані у кодеку MELP (Mixed Excitation Linear Prediction – кодек лінійного передбачення зі змішаним збудженням) зі швидкістю 2,4 кбіт/с. У цьому кодеку використані чотиризмуговий аналіз мовлення і лінійне передбачення.

**Критерії оцінки мовних кодеків.** Кількісна оцінка якості передачі мовлення – це важка та багатогранна задача. Існує багато підходів до проблеми визначення якості. Найбільш широко використовуваний підхід оперує оцінкою *MOS* (*Mean Opinion Score*), яка визначається для конкретного кодека як середня оцінка якості великою групою слухачів за п'ятибальною шкалою [9]. Для прослуховування експертам пред'являються різні звукові фрагменти – мова, музика, мова на фоні різного шуму і т.д. Оцінки інтерпретують таким чином:

4-5 – висока якість; аналогічно якості передачі мови у цифровій мережі з інтеграцією обслуговування (ЦМІО), або ще вище;

3,5-4 – якість телефонної мережі загального користування (ТМЗК) (*toll quality*); аналогічно якості мови, що передається за допомогою кодека АДІКМ при швидкості 32 кбіт/с. Така якість зазвичай забезпечується для передавання більшості телефонних розмов. Мобільні мережі забезпечують якість трохи нижче *toll quality*;

3-3,5 – якість мови, задовільна, проте його погіршення явно помітно на слух;

2,5-3 – мова розбірлива, проте вимагає концентрації уваги для розуміння. Така якість зазвичай забезпечується в системах зв'язку спеціального призначення (наприклад, у збройних силах).

У рамках існуючих технологій якість ТМЗК (*toll quality*) неможливо забезпечити при швидкостях менше 5 кбіт/с.

Усереднена залежність якості від необхідної швидкості передачі для розглянутих трьох основних груп методів АЦП наведено на рис. 1 [8].

Недолік моделі *MOS* полягає в тому, що ця модель не може кількісно та окремо враховувати такі фактори: наскрізну затримку між абонентами; варіацію затримки (джиттер); втрати пакетів; відлуння (якщо в розмовному тракті є перехід з двопроводової схеми передачі до чотирипроводової чи навпаки).

Усунути ці недоліки можна при використанні *E*-моделі, яка дозволяє отримати оцінку якості телефонної передачі на основі результатів вимірювання різних характеристик тракту

передачі пакетів та кінцевого обладнання [10, 11]. Ця модель дозволяє в комплексі врахувати практично всі несприятливі фактори.

У ході розробки  $E$ -моделі була проведена велика кількість випробувань із отриманням суб'єктивних оцінок впливу різних факторів, що спотворюють якість телефонної передачі. Результати цих випробувань були використані в  $E$ -моделі для обчислення об'єктивних оцінок.

Результатом оцінки якості телефонної передачі відповідно до  $E$ -моделі є число, яке називають  $R$ -фактором:

$$R = R_0 - I_s - I_d - I_e + A,$$

де  $R_0 = 93,2$  – вихідне значення  $R$ -фактора;  $I_s$  – спотворення, що вносяться кодеками і шумами в каналі;  $I_d$  – спотворення за рахунок сумарної наскрізної затримки ("з кінця в кінець") у мережі;  $I_e$  – спотворення, що вносяться обладнанням, у тому числі й втрати пакетів;  $A$  – так званий фактор переваги (наприклад, мобільні користувачі можуть погодитись з нижчим рівнем якості, отримуючи додаткові зручності. У більшості випадків розрахунку  $R$ -фактора параметр  $A$  приймається рівним нулю).

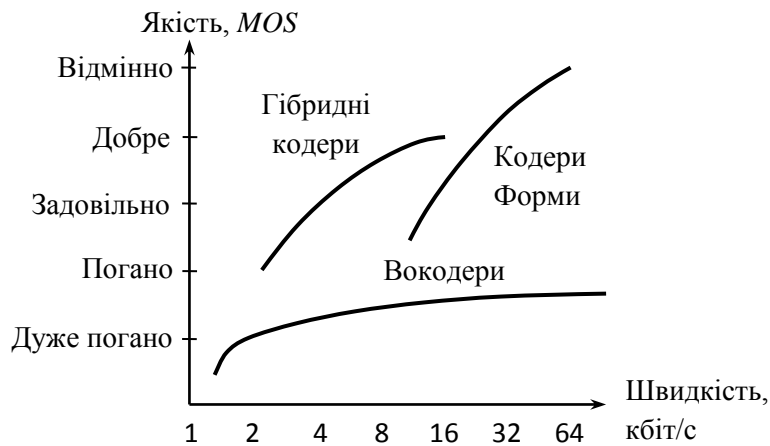


Рис. 1. Узагальнений порівняльний аналіз мовних кодеків

У табл. 1 показана відповідність оцінок  $MOS$  та  $E$ -моделі.

Таблиця 1

Відповідність шкали  $MOS$  та  $E$ -моделі

Якість	$R$	$MOS$
Макс значення шкали	120	5,0
Досяжний максимум ( $G.711$ )	93	4,4
Бізнес-якість	$\geq 80$	$\geq 4,0$
Гранично допустима	70...80	3,6...4,0
Погана	$50...70 \leq$	2,6...3,6
Не рекомендована	$\leq 50$	$\leq 2,5$

Характеристики розглянутих алгоритмів кодування мовлення наведені в табл. 2. Кодек  $iLBC$  буде розглянуто нижче.

Середня суб'єктивна оцінка якості мовлення ілюструє перевагу  $MP-MLQ$  за співвідношенням якості сигналу і швидкості передачі інформації.

У майбутньому очікується поява кодеків із якістю, що не поступаються кодекам стандарту  $G.729$  (8 кбіт/с), але таких, що працюють на швидкостях нижчих ніж 4,8 кбіт/с і з затримкою, яка не перевищує гранично припустиму (0,2 с), встановлену МСЕ для телефонних мереж загального призначення.

Розглянемо кодеки, що спеціально розроблені для використання в  $IP$ -телефонії.

**Кодек *iLBC*.** *iLBC (Internet Low Bitrate Codec)* – це вільний від ліцензійних відрахувань кодек для голосового зв'язку через *Internet*. Кодек призначений для вузькосмугових каналів із швидкістю передачі мови 13,33 кбіт/с при довжині кадру в 30 мс або 15,2 кбіт/с для кадру 20 мс. Кодек *iLBC*, описаний у стандарті *RFC 3951* [12], дозволяє добитися хорошої якості передачі навіть при деяких спотвореннях, які виникають у зв'язку з втратою або затримкою пакетів.

Основні характеристики кодека: частота дискретизації 8 кГц/16 біт (160 відліків для 20-мс кадрів, 240 відліків для 30-мс кадрів); керована реакція на втрату пакетів, затримки і джиттер; фіксована швидкість (15,2 кбіт/с для 20-мс кадрів, 13,33 кбіт/с для 30-мс кадрів); фіксований розмір кадру (304 біта в кадрі для 20-мс кадрів, 400 біт в кадрі для 30-мс кадрів); забезпечується стійкість до втрат пакетів на рівні ІКМ із приховуванням втрат пакетів, як в *G.711*; завантаження процесора на рівні *G.729A* при вищій якості та кращій реакції на втрату пакетів; тестування за ідеальних умов приводить до усередненої суб'єктивної оцінки (*MOS*) у 4,14 для *iLBC* (15,2 кбіт/с), порівняно з оцінкою 4.45 для *G.711* ( $\mu$ -закон).

Таблиця 2

Основні характеристики алгоритмів АЦП мови

Швидкість, кбіт/с	Якість ( <i>MOS</i> )	Погіршення якості для <i>E</i> -моделі ( $I_s$ )	Розмір пакета, мс	Алгоритмічна затримка, мс	Сумарний потік, кбіт/с	Назва стандарту	Рік випуску	Алгоритм кодування	Область застосування
64	4,4	0	20	0	81,2	<i>G.711</i>	1960	ІКМ	Телефонні мережі, <i>VoIP</i>
13,33	4	н/д	30	30	28	<i>iLBC (RFC 3951)</i>	після 2000	–	<i>VoIP, Skype</i>
15,2	4	н/д	20	30	29				
8	3,9	10	20	15	31,2	<i>G.729</i>	1997	<i>CS-ACELP</i>	Телефонні мережі, <i>VoIP</i>
6,3	3,9	15	30	37,5	21,9	<i>G.723.1</i>	1996	<i>MP-MLQ</i>	Телефонні мережі, <i>VoIP</i>
5,3	3,7	19	30	37,5	20,8	<i>G.723.1</i>	1996	<i>ACELP</i>	
13	3,3	20	20	20	35,4	<i>ETSI GSM</i>	1992	<i>RPE-LTP</i>	Стільникова телефонія, <i>VoIP</i>
32	3,8	7	–	–	–	<i>G.726</i>	1984	АДІКМ	Телефонні мережі
16	3,6	7	–	–	–	<i>G.728</i>	1992	<i>LD-CELP</i>	Телефонні мережі
2,4	3,5	н/д	–	–	–	-	1998	<i>MELP</i>	Мін. оборони США
4,8	3,4	н/д	–	–	–	<i>ETSI TETRA</i>	1996	<i>ACELP</i>	Стільникова телефонія
6,4	3,1	н/д	–	–	–	<i>INMARSAT-M</i>	1990	<i>IMBE</i>	Супутникова телефонія

**Кодек *Speex*.** *Speex* – це вільний (неліцензійний) кодек для стиснення мовного сигналу, який може використовуватися в *VoIP* додатках [13].

На відміну від багатьох інших кодеків мови *Speex*, в основному, призначається не для стільникових телефонів, а для використання в *VoIP* і створення файлів із стисненим звуком. *Speex* оптимізований для отримання високоякісного мовного сигналу при низьких швидкостях. Для досягнення цієї мети кодек використовує змінну швидкість і підтримує різні діапазони частот: надширокий (*ultra-wideband*, частота дискретизації 32 кГц), широкий (*wideband*, 16 кГц) і вузький (*narrowband*, якість телефонної лінії – 8 кГц). Спрямованість на *VoIP* замість стільникового зв'язку означає, що *Speex* повинен бути стійкий до втрат пакетів даних, але не до пошкодження їх, оскільки *UDP* (протокол непідтвердженої доставки повідомлень) надає інформацію лише двох видів – дані прибули непошкодженими або ж втрачені. Ця особливість визначає вибір для *Speex* методу кодування *CELP*. До основних

особливостей кодека можна віднести: вільне і відкрите програмне забезпечення, що не має патентних обмежень; динамічне перемикання швидкості та змінна швидкість (*Variable bit-rate, VBR*) – від 2 до 44 кбіт/с; детектор мовної активності (*Voice Activity Detection, VAD*, інтегрований з *VBR*).

**Широкосмугові кодеки.** Переважна більшість кодеків, використовуваних для телефонного зв'язку, займають смугу 300...3400 Гц (так звана вузькосмугова мова). Це обмеження існує майже 100 років, причому самі мережі смуги частот не обмежують (обмеження обумовлене характеристиками застосовуваних перетворювачів). Саме частота 3,4 кГц була прийнята в якості верхньої граничної для комутованої ТМЗК у стандарті цифрової передачі *G.711*. Хоча більша частина енергії найчастіше міститься в голосних звуках, які займають смугу частот нижче 3 кГц, приголосні, що несуть критичну інформацію, часто вимагають смуги частот вище 3 кГц. Тому вузькосмугові системи можуть погіршити розбірливість, наприклад, звуки „с” і „ф” розрізняються тільки за рахунок частот вище 3 кГц. З іншого боку, збільшення смуги частот сигналу до 50...7000 Гц (так званої широкої смуги), покращує розбірливість, і, як наслідок, вимагає від слухача меншої концентрації уваги, а отже, значно знижує втому.

Тому поряд з розробками вузькосмугових мовних кодеків велику увагу фахівців привертають широкосмугові мовні кодеки [8]. Перші широкосмугові кодері *G.722* (48, 56, 64 кбіт/с) були стандартизовані МСЕ 1988 року. Спочатку передбачалося, що вони замінять кодері *G.711*, коли цифрові мережі з інтеграцією обслуговування (ЦМІО) отримають більш широкого розповсюдження. Це кодері зі складною формою сигналу, які працюють з двома піддіапазонами і мають хороші характеристики, але невисокий коефіцієнт стиснення. Потім послідувала рекомендація *G.722.1* (24 і 32 кбіт/с) на кодер, широко використовуваний у даний час у терміналах конференц-зв'язку.

У 2000 році проектом по співпраці в створенні мереж третього покоління *3GPP* (*3G Partnership Project*) був стандартизований кодер *AMR-WB* (*Adaptive Multi-Rate Wide Band*) для застосування в мобільних системах третього покоління. У 2001 році він був стандартизований МСЕ в якості широкосмугового кодера *G.722.2*. Кодер працює з різними швидкостями передачі від 6,6 до 23,85 кбіт/с, але його обчислювальна складність може бути обмежена можливостями реалізації. Дуже важливо, що обидві організації (МСЕ і *3GPP*) прийняли один і той же кодер, так як це усуває необхідність перекодування при зустрічній роботі. У кінцевому рахунку це знизить вартість і поліпшить характеристики передачі „з кінця в кінець”. Також, розширено для широкосмугового застосування кодек *G.729* шляхом створення можливості багатошвидкісної передачі (рекомендація *G.729EV* 2006 р.).

**Кодек *Opus*.** Очікується, що *Skype* найближчим часом перейде на використання нового аудіокодека *Opus* [14], орієнтованого для роботи в безпроводових мережах. Головна технічна перевага *Opus* полягає в знайденому балансі між компресією аудіосигналу і його якістю, що актуально в умовах передачі в мережах мобільних операторів. Кодек використовує гнучкий алгоритм адаптації у разі зміни пропускну здатності каналу – наприклад, при переході з *3G*-сигналу на *Wi-Fi* з'єднання – і, в подальшому, може забезпечити передачу розмови в *CD*-якості. Також застосовані спеціальні алгоритми для боротьби з втратою пакетів при обмежених можливостях безпроводової мережі. Кодек є безкоштовним і може бути вільно ліцензований сторонніми розробниками для використання в *VoIP*-додатках. На рис. 2 представлено порівняння якості *Opus* з іншими кодеками [15].

Основні технічні характеристики: підтримує змінну швидкість – від 6 до 510 кбіт/с; частота дискретизації від 8 до 48 кГц; розмір кадру від 2,5 мс до 60 мс.

Розглянемо основні особливості реалізації аналого-цифрового перетворення в радіомережах з комутацією пакетів та визначимо вимоги до кодеків мови в цих мережах.

Послідовність перетворень мовного сигналу при його передачі в *IP*-мережі полягає в наступному (рис. 3):

– на передачі: дискретизація та цифровізація аналогового мовного сигналу, його компресія у відповідності зі стандартами МСЕ, інкапсуляція в пакети, які передаються по мережі у віддалений маршрутизатор;

– на прийомі – зворотні перетворення: сортування пакетів та їх перетворення у безперервну послідовність, декомпресія цифрового сигналу та його перетворення в аналоговий.

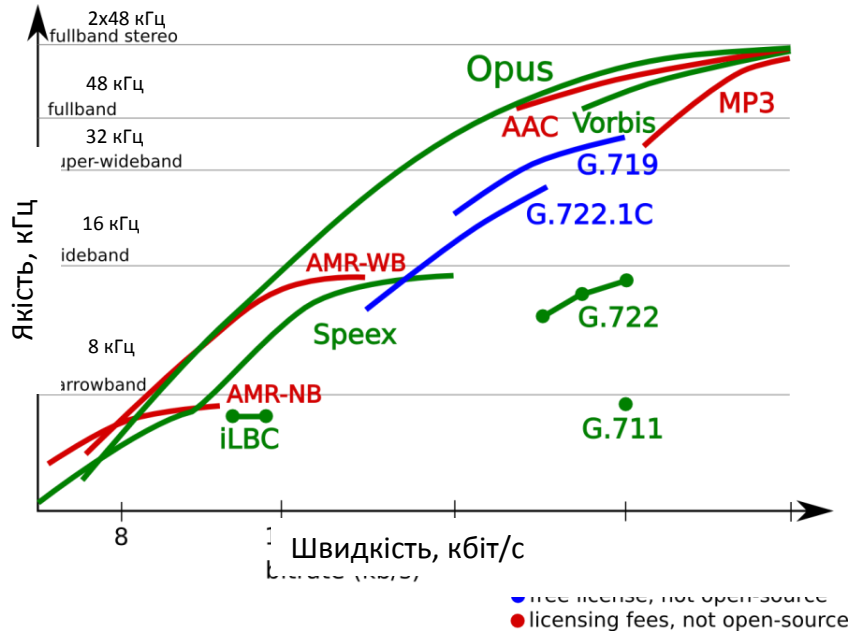


Рис. 2. Порівняльний аналіз якості кодеків для VoIP

З рис. 3 видно, що при формуванні пакета до корисної інформації додаються заголовки протоколів RTP (Real-time Transport Protocol), UDP (User Datagram Protocol), IP (Internet Protocol), технології Ethernet. Детально призначення цих протоколів та заголовків розглядається в [1, 4, 5].

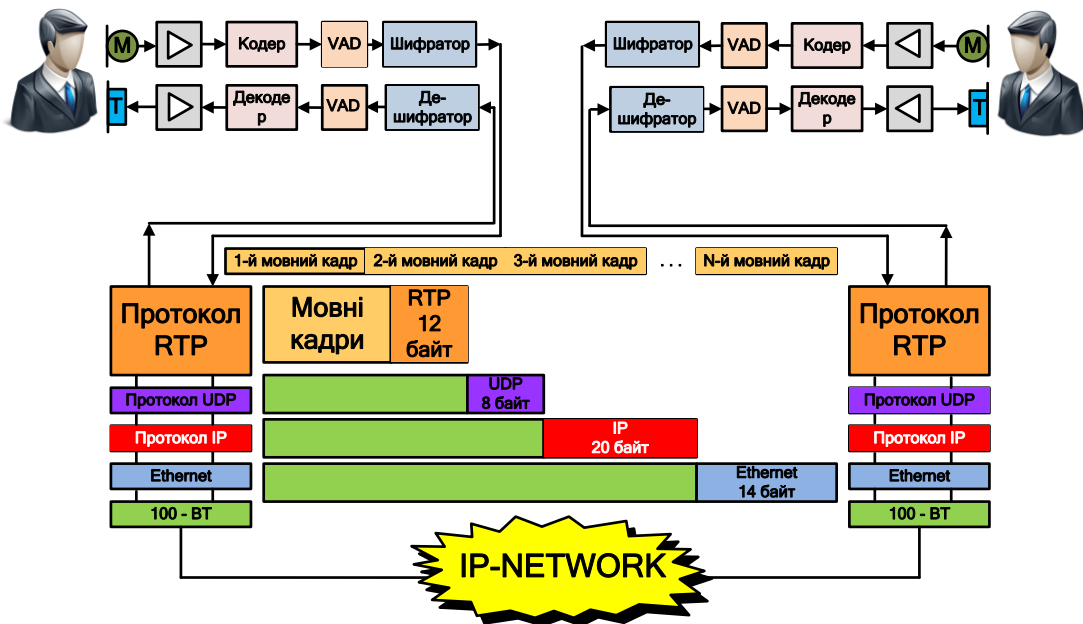


Рис. 3. Схема підготовки мовної інформації в кінцевих терміналах та шлюзах

**Проблема затримки сигналу.** Серед багатьох чинників, що впливають на якість передачі мови, можна відзначити затримку сигналу в терміналах і вузлах мережі.

Вплив затримки на якість передачі мови можна оцінити за допомогою кривої, отриманої з використанням моделі характеристик передачі згідно з рекомендаціями МСЕ-Т G.107. Ця крива, що дозволяє оцінити вплив затримки на якість мови, представлена в рекомендації МСЕ-Т G.114 [14] і показана на рис. 4.

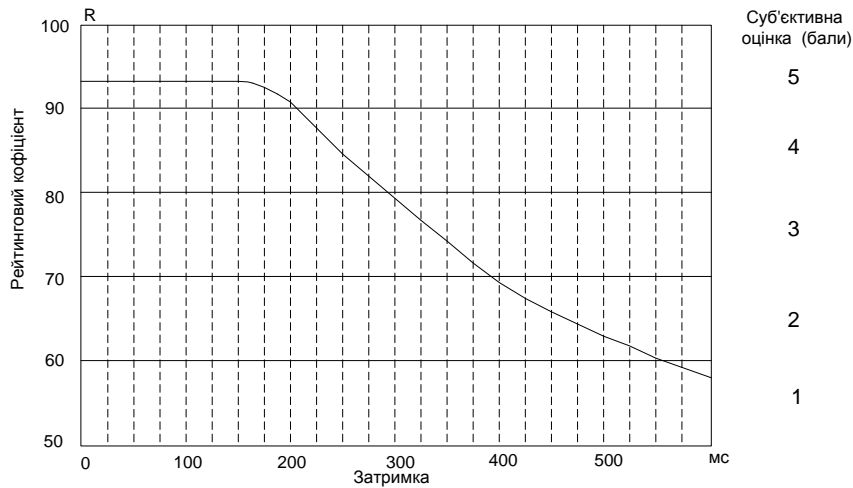


Рис. 4. Залежність якості передачі мови від затримки сигналу

По осі ординат ліворуч відкладені значення рейтингового коефіцієнта  $R$  згідно  $E$ -моделі, по осі ординат праворуч – рівні „прийнятності” якості передачі для користувача, пропорційні значенням  $R$  і виражені в балах. При рейтинговому коефіцієнті  $R < 50$  якість настільки погана, що не підлягає будь-якій оцінці. Затримка понад 400-500 мс призводить до практичної неможливості здійснення діалогового режиму в мовному з’єднанні. Затримана відповідь співрозмовника викликає психологічну реакцію у вигляді перепитування, в той же час надходить затримана відповідь, і нормальний діалог порушується.

Бажаною є затримка, яка не перевищує 150 мс, оскільки крім затримки слід враховувати й інші погіршуючі фактори. Затримка, що вноситься стандартними кодеками ІКМ, незначна і становить менше 0,4 мс.

Істотне збільшення затримки, у порівнянні зі стандартними кодеками ІКМ, дають низькошвидкісні кодеки. Зокрема, тільки самі кодеки в терміналах GSM вносять затримку в 60 мс. Процедура формування кадрів на радіоінтерфейсі додає ще 35 мс затримки.

Очевидно, що затримки в низькошвидкісних кодексах можуть зробити додаткові перекодування практично неприпустимими. При цьому не можна забувати про значні затримки у таких мережевих елементах як шлюзи, маршрутизатори тощо.

Таким чином, розробникам доводиться робити вибір між якістю зв’язку і шириною каналу, а завдання проектування найбільш якісних кодеків для низькошвидкісних каналів з високим рівнем завад є, безперечно, актуальним.

**Придушення періодів мовчання.** При діалозі один його учасник говорить, в середньому, тільки 35 відсотків часу. Таким чином, якщо застосувати алгоритми, які дозволяють зменшити об’єм інформації, що передається в періоди мовчання, то можна значно звузити необхідну смугу пропускання. У двосторонній розмові такі заходи дозволяють досягти скорочення об’єму інформації до 50 %, а в децентралізованих багатоадресних конференціях (за рахунок більшої кількості тих, що говорять) – і більше.

Потрібно відзначити, що визначення меж пауз у мові є дуже істотним для ефективної синхронізації передавальної та приймальної сторін: приймач може, трохи змінюючи тривалість пауз, проводити підстроювання швидкості відтворення для кожного окремого



сеансу зв'язку, що виключає необхідність синхронізації тактових генераторів усіх елементів мережі, як це має місце в ТМЗК.

Для визначення періодів часу, коли користувач говорить необхідний детектор мовної активності – *VAD*. Він повинен володіти малим часом реакції, щоб не допускати втрат початкових слів і не пропускати даремні фрагменти мовчання в кінці речень (фраз); у той же час *VAD* не повинен спрацьовувати від дії фонового шуму. *VAD* оцінює енергію вхідного сигналу і, якщо вона перевищує деякий поріг, активізує передачу. Якби детектор відкидав усю інформацію до моменту поки енергія сигналу не стала вища за поріг, то відбувалося б відрізування початкової частини періоду активності. Тому реалізації *VAD* вимагають збереження в пам'яті декількох мілісекунд інформації, щоб мати можливість запустити передачу до початку періоду активності. Це збільшує, в деякій мірі, затримку проходження сигналу, проте її можна мінімізувати або звести до нуля в кодерах, що працюють з блоками відліків.

Підтримка переривистої передачі (*Discontinuous Transmission – DTX*) дозволяє кодеку припинити передачу пакетів у той момент, коли *VAD* виявив період мовчання. Найбільш досконалі кодери не припиняють передачу повністю, а переходять у режим передачі набагато меншого об'єму інформації (інтенсивність, спектральні характеристики), потрібної для того, щоб декодер на віддаленому кінці міг відновити фоновий шум.

Генератор комфортного шуму (*Comfort Noise Generator – CNG*) служить для генерації фонового шуму. У момент, коли в мові активного учасника бесіди починається період мовчання, термінали тих абонентів, що слухають, можуть просто відключити відтворення звуку. Проте це було б нерозсудливо. Якщо в трубці виникає абсолютна тиша, тобто фоновий шум (наприклад, шум вулиці), який був чутний під час розмови, раптово зникає, то тому, хто слухає, здається, що з'єднання порушилося, і він зазвичай починає питати, чи чує його співрозмовник. Генератор *CNG* дозволяє уникнути таких неприємних ефектів.

Більш прості кодеки просто припиняють передачу в період мовчання, і декодер генерує який-небудь шум з рівнем, що дорівнює мінімальному рівню, відміченому в період мовної активності. Досконаліші кодеки (*G.723.1 Annex A, G.729 Annex B*) мають можливість надавати віддаленому декодеру інформацію для відновлення шуму з параметрами, близькими до тих, що фактично спостерігалися.

**Розмір кадру.** Більшість вузькосмугових кодеків обробляють мовну інформацію блоками, що називаються кадрами (*frames*), і їм необхідно проводити попередній аналіз відліків, що поступають безпосередньо за відліками в блоці, який вони в даний момент кодують. Розмір кадру важливий, оскільки мінімальна теоретично досяжна затримка передачі інформації (алгоритмічна затримка) визначається сумою цього параметра і довжини буфера попереднього аналізу. Насправді, процесори цифрової обробки сигналів, які виконують алгоритм кодування, мають кінцеву продуктивність, так що реальна затримка сигналу більше теоретичної.

Можна, здавалося б, зробити висновок, що кодеки з меншим розміром кадру є кращими за таким важливим критерієм, як мінімізація затримки. Проте якщо врахувати, що відбувається при передачі інформації мережею, то до кадру, сформованого кодеком, додається багато додаткової інформації (див. рис. 3) – заголовки *IP* (20 байтів), *UDP* (8 байтів), *RTP* (12 байтів). Для кодека з тривалістю кадру 30 мс посилка таких кадрів по мережі призвела б до передачі надмірної інформації зі швидкістю 10,6 кбіт/с, що перевищує швидкість передачі мовної інформації у більшості вузькосмугових кодеків. Тому, зазвичай, використовується пересилка декількох кадрів у пакеті, при цьому їх кількість обмежена максимально допустимою затримкою. У більшості випадків в одному пакеті передається до 60 мс мовної інформації. Чим менше тривалість кадру, тим більше кадрів доводиться упаковувати в один пакет, тобто затримка визначається зовсім не довжиною кадру, а практично прийнятним об'ємом корисного навантаження в пакеті. Крім того, кодеки з більшою довжиною кадру ефективніші, оскільки тут діє загальний принцип: чим довше спостерігається явище (мовний сигнал), тим краще воно може бути змодельоване.

**Чутливість до втрат кадрів.** Втрати пакетів є невід’ємним атрибутом *IP*-мереж. Оскільки пакети містять кадри, сформовані кодеком, то це викликає втрати кадрів. Але втрати пакетів і втрати кадрів не обов’язково безпосередньо пов’язані між собою, оскільки існують підходи (такі як застосування кодів з виправленням помилок), що дозволяють зменшити число втрачених кадрів при даному числі втрачених пакетів. Потрібна для цього додаткова службова інформація розподіляється між декількома пакетами, так, що при втраті деякого числа пакетів кадри можуть бути відновлені.

Але позитивний ефект від введення надмірності для боротьби з втратами пакетів не є так легко досяжним, оскільки втрати в *IP*-мережах відбуваються серіями пакетів, тобто значно вірогідніше те, що буде втрачено відразу декілька пакетів підряд, ніж те, що втрачені пакети розподіляться в послідовності переданих пакетів поодиночці. Отже, якщо застосовувати прості схеми введення надмірності (наприклад, повторюючи кожен кадр у двох пакетах, що послідовно передаються), то в реальних умовах вони, хоч і збільшать об’єм надмірної інформації, але, швидше за все, виявляться даремними. Крім того, введення надмірності негативно позначається на затримці відтворення сигналу. Наприклад, якщо ми повторюємо один і той же кадр у чотирьох пакетах підряд, щоб забезпечити можливість відновлення інформації при втраті трьох підряд переданих пакетів, то декодер вимушений підтримувати буфер із чотирьох пакетів, що вносить значну додаткову затримку відтворення.

Вплив втрат кадрів на якість відтвореної мови залежить від використовуваного кодека. Якщо втрачений кадр, що складається з  $N$  мовних відліків кодека *G.711*, то на приймальному кінці буде відмічений пропуск звукового фрагмента тривалістю  $N \times 125$  мкс. Якщо використовується досконаліший вузькосмуговий кодек, то втрата одного кадру може позначитися на відтворенні декількох наступних, оскільки декодеру буде потрібен час для того, щоб досягти синхронізації з кодером – втрата кадру тривалістю 20 мс може призводити до чутного ефекту протягом 150 мс і більше.

Кодеки типу *G.723.1* розроблені так, що вони функціонують без істотного погіршення якості в умовах некорельованих втрат до 3 % кадрів, проте при перевищенні цього порогу якість погіршується катастрофічно.

Як показує аналіз літератури [3–6, 8], якість відтворення мови на прийомі значно залежить від вибору схеми завадостійкого кодування, при цьому ефективність завадостійкого кодека в цьому випадку доцільно оцінювати не енергетичним вирашем (або кількістю виправлених помилок). Це пояснюється тим, що мовний кодек повинен бути стійкий до втрат пакетів даних, але не до пошкодження їх, оскільки *UDP* (протокол негарантованої доставки повідомлень) надає інформацію лише двох видів – дані прибули неушкодженими або ж втрачені.

**Мінімізація швидкості передачі.** Обмежений спектр каналів радіозв’язку вимагає зменшення швидкості передачі на виході кодера при забезпеченні заданої якості відтворення мови. Тому перспективні кодери для пакетних радіомереж повинні підтримувати змінну швидкість (*VBR*) та пристосовуватись до наявного частотного ресурсу з можливістю забезпечення в достатньо широкому каналі якості „*fullband*” за необхідності, а в умовах частотних обмежень – мінімально необхідної вузькосмугової („*narrowband*”) передачі.

Таким чином, кодеки мови, що розроблені та впроваджені в засоби телекомунікацій не в повній мірі відповідають вимогам, що висуваються мережами наступного покоління. Використання їх у складі засобів рухомого радіозв’язку висуває додаткові вимоги щодо стійкості до втрат пакетів та завадостійкості в цілому.

У результаті проведеного аналізу методів аналого-цифрового перетворення мовних сигналів та особливостей передачі мовних повідомлень у радіомережах із комутацією пакетів можна сформулювати вимоги, які повинні забезпечуватись при практичній реалізації кодека: мінімізація затримок, що вносяться при аналого-цифровому перетворенні, підтримка змінної швидкості передачі, наявність детектора мовної активності, підтримка переривистої передачі, наявність генератора комфортного шуму, оптимізація розміру кадрів, зменшення

чутливості до втрат кадрів та пакетів, у першу чергу, за рахунок вибору ефективного методу завадостійкого кодування.

Перспективними напрямками подальших досліджень є:

– дослідження ефективності функціонування існуючих мовних кодеків у пакетних радіомережах, у тому числі, при впливі несприятливих факторів, що призводять до втрат пакетів, наприклад, перевантаження у мережі, функціонуванні в умовах багатопроменевості та доплерівського зміщення частоти, природних та штучних завад;

– вибір найбільш ефективних кодеків для різних сигнальних та завадових ситуацій, розробка кодека з використанням відомих та вдосконалених методів, здатного адаптивно змінювати швидкість та алгоритм кодування з метою підвищення стійкості щодо впливу вказаних несприятливих факторів.

#### ЛІТЕРАТУРА

1. Гольдштейн Б. С. Сети связи: Учебник для ВУЗов / Б. С. Гольдштейн, Н. А. Соколов, Г. Г. Яновский. – СПб.: БХВ-Петербург, 2010. – 400 с., илл.
2. Єрохін В. Ф. Аналіз і прогноз розвитку систем мобільного зв'язку загального користування / В. Ф. Єрохін, Б. А. Гиндич, О. В. Кувшинов // Збірник наукових праць ВІПІ НТУУ „КПІ”. – № 1. – 2011. – С. 54 – 64.
3. Современные беспроводные сети: состояние и перспективы развития / [Гепко И. А., Олейник В. Ф., Чайка Ю. Д., Бондаренко А. В.]. – К.: „ЕКМО”, 2009. – 672 с.
4. Олифер В.Г. Новые технологии и оборудование IP-сетей / В. Г. Олифер, Н. А. Олифер. – СПб.: БХВ-Петербург, 2000. – 512 с.
5. Нейман В. И. Интернет-телефония и перспективы ее развития / В. И. Нейман, Д. А. Селезнев // Электросвязь. – 2008. – №1. – с. 6–9.
6. Enable VoIP Survivability for Future Tactical Airborne Networks. [T. Le, S. Cook, G. Nadynski and oth.]. – In Proc. of MILCOM 2010. – P. 760 – 765.
7. Романюк В. А. Интеллектуальная маршрутизация в мобильных радиосетях (MANET) / Романюк В. А., Шацко П. В., Сова О. Я. [и др.] // Зв'язок. – 2011. – № 2(94). – С. 24 – 31.
8. Шелухин О. И. Цифровая обработка и передача речи / О. И. Шелухин, Н. Ф. Лукьянцев. – М.: Радио и связь, 2000. – 454 с.
9. Теорія електричного зв'язку. Ч. 1: Основи теорії сигналів та розподілу інформації: Підручник / [Кувшинов О. В., Лівенцев С. П., Лежнюк О. П., та ін.]. – К.: ВІПІ НТУУ „КПІ”, 2008. – 331 с.
10. ITU-T Rec. G. 107. The E-Model, a computational model for use in transmission planning (12/1998).
11. Жученко О.С. Дослідження впливу характеристик продуктивності мережі технологічного зв'язку залізничного транспорту на якість телефонної передачі / О. С. Жученко, М. А. Абакумов, А. О. Соловйов // Інформаційно-керуючі системи на залізничному транспорті. – 2010. – №5. – С. 28 – 32.
12. Internet Low Bit Rate Codec (iLBC) [Електронний ресурс]. – Режим доступу: <http://www.ietf.org/rfc/rfc3951.txt>.
13. Сжатие аудио [Електронний ресурс]. – Режим доступу: <http://audiocoding.ru/форматы/speech.html>.
14. Definition of the Opus Audio Codec [Електронний ресурс]. – Режим доступу: <http://tools.ietf.org/html/rfc6716>.
15. Opus Interactive Audio Codec [Електронний ресурс]. – Режим доступу: <http://www.opus-codec.org>.
16. ITU-T Rec. G.114. International telephone connections and circuits – General Recommendations on the transmission quality for an entire international telephone connection (01/2003).